On the gracesize of trees

Shoham Letzter*† Alexey Pokrovskiy* Ella Williams*‡

Abstract

An n-vertex tree T is said to be *graceful* if there exists a bijective labelling $\phi : V(T) \to \{1, \dots, n\}$ such that the edge-differences $\{|\phi(x) - \phi(y)| : xy \in E(T)\}$ are pairwise distinct. The longstanding graceful tree conjecture, posed by Rósa in the 1960s, asserts that every tree is graceful. The *gracesize* of an n-vertex tree T, denoted $\mathrm{gs}(T)$, is the maximum possible number of distinct edge-differences over all bijective labellings $\phi : V(T) \to \{1, \dots, n\}$. The graceful tree conjecture is therefore equivalent to the statement that $\mathrm{gs}(T) = n - 1$ for all n-vertex trees.

We prove an asymptotic version of this conjecture by showing that for every $\varepsilon > 0$, there exists n_0 such that every tree on $n > n_0$ vertices satisfies $\mathrm{gs}(T) \geq (1 - \varepsilon)n$. In other words, every sufficiently large tree admits an almost graceful labelling.

1 Introduction

Within extremal combinatorics, a common question to pose is "which properties must a host graph possess in order to guarantee that it contains a given subgraph?", with classical results such as theorems of Mantel [12] and Dirac [4] providing sufficient conditions for the existence of triangles and spanning cycles respectively. Colouring analogues of these problems have been extensively studied, most notably in Ramsey theory, which investigates the appearance of monochromatic subgraphs in edge-coloured graphs. A complementary perspective is given by anti-Ramsey theory, initiated by Erdős, Simonovits, and Sós [6], which concerns the existence of rainbow subgraphs, in which all edges must receive distinct colours. In this paper, we use this colouring framework to study a classical labelling problem in combinatorics: the graceful tree conjecture. Specifically, we tackle the problem by reformulating the conjecture in terms of finding rainbow trees in appropriately edge-coloured complete graphs, connecting graceful labellings to tools from extremal and probabilistic graph theory.

A *labelling* of a graph G on m edges is an injective mapping $\phi : V(G) \to \mathbb{N}$, and we say ϕ is *graceful* if its image is $\{1, \dots, m+1\}$ and the values of $|\phi(x) - \phi(y)|$ are pairwise distinct over all edges $xy \in E(G)$. Graceful labellings were first introduced by Rósa [17] in 1966 (as β-valuations) and later renamed by Golomb [8]. Given a labelling ϕ, we call the values of $|\phi(x) - \phi(y)|$ for $xy \in E(G)$ the *edge-differences* of ϕ. The main theme of research in this area is to determine which graphs admit graceful labellings. An unpublished result of Erdős (see [9]) reports that almost all graphs do not have this property. On the other hand, trees are an interesting class of graphs to consider, since the set of edge-differences of any graceful labelling ϕ of an n-vertex tree must exactly coincide with the set $\{1, \dots, n-1\}$. This resulted in the graceful tree conjecture, posed in the 1960s, and usually accredited to Rósa.

Conjecture 1.1 (Graceful tree conjecture)**.** *Every tree admits a graceful labelling.*

*Department of Mathematics, University College London, Gower Street, WC1E 6BT, UK. Emails: {s.letzter,a.pokrovskiy,ella.williams.23}@ucl.ac.uk.

†Research supported by the Royal Society.

‡Research supported by the Martingale Foundation.

Graceful labellings were initially motivated as a method for proving Ringel's conjecture [16], a notable conjecture from 1963 which proposed that for every $(n+1)$-vertex tree T, the complete graph K_{2n+1} can be decomposed into edge-disjoint copies of T. Conjecture 1.1 implies a stronger variant of Ringel's conjecture, attributed to Kotzig, suggesting that this decomposition can be done by cyclically shifting the copies of the tree. Indeed, if there exists a graceful labelling of an $(n+1)$-vertex tree T, say $\phi : V(T) \to \{1, \ldots, n+1\}$, then one can construct an embedding of T in K_{2n+1} using only the first $n+1$ vertices, by embedding each vertex of T to its image under ϕ. Cyclically shifting this copy of T $2n+1$ times by adding 1 modulo $2n+1$ to each label at each shift, we get $2n+1$ copies of T, which can be seen to be pairwise edge-disjoint, and which cover all edges of K_{2n+1} using the properties of the graceful labelling ϕ. Ringel's conjecture (and its stronger variant) was recently proven for all sufficiently large n by Montgomery, Pokrovskiy and Sudakov [15] and independently Keevash and Staden [11]. The strength of the graceful tree conjecture illustrates that it stands to be a challenging and interesting problem in itself.

Over the years, research towards Conjecture 1.1 has been fruitful, yet in full generality it remains open. Despite a resolution for certain classes of trees (see e.g. Table 1 of [7]), there seems to be much more difficulty in proving broader statements. The best known general result about Conjecture 1.1 due to Adamaszek, Allen, Grosu and Hladký considers a variant of graceful labellings where a small number of additional labels are allowed. Using terminology of Van Bussel [3], given $m > n$ and an n-vertex tree T, an injective mapping $\phi : V(T) \to \{1, \ldots, m\}$ for which the edge-differences of ϕ are pairwise distinct is called a *range-relaxed graceful labelling* of T. In 2020, they proved the following notable result.

Theorem 1.2 (Adamaszek, Allen, Grosu and Hladký [1])**.** *For all $\varepsilon > 0$ there exist $\eta, n_0 > 0$ such that for all $n > n_0$, the following holds. If T is an n-vertex tree and $\Delta(T) \leqslant \frac{\eta n}{\log n}$, then there exists a range-relaxed graceful labelling $\phi : V(T) \to \{1, \ldots, (1+\varepsilon)n\}$.*

Alternatively, one could relax the notion of graceful labellings by allowing a small number of edge-differences to coincide, and we follow this approach. Introduced by Erdős, Hell and Winkler [5], and denoted by $\mathrm{gs}(T)$, the *gracesize* of an n-vertex tree T is the maximum size of the set of edge-differences over all bijective labellings $\phi : V(T) \to \{1, \ldots, n\}$. This was mentioned as a dual notion to the bandsize of a tree, which instead considers the minimum size of this set. Gracesize naturally relates to the graceful tree conjecture since such a labelling ϕ of T is graceful if and only if the set of its edge-differences has size $|E(T)| = n-1$. Therefore, one can restate the graceful tree conjecture as: every n-vertex tree has gracesize $n-1$. It was suggested in [5] that it may be an interesting problem to find non-trivial lower bounds on the gracesize.

In 1995, Rósa and Širáň [18] proved that $\mathrm{gs}(T) \geqslant 5(n-1)/7$ for every n-vertex tree T. Without too much difficulty one can show that $\mathrm{gs}(T) = (1-o(1))n$ for all trees with maximum degree $o(n/\log n)$ as a corollary of Theorem 1.2[1]. The main result of the paper confirms this bound on the gracesize of all sufficiently large trees, irrespective of maximum degree.

Theorem 1.3. *For all $\varepsilon > 0$ there exists n_0 such that for all $n > n_0$, the following holds. If T is an n-vertex tree then $\mathrm{gs}(T) \geqslant (1-\varepsilon)n$.*

1.1 Reduction argument

Before reformulating the problem to be about edge-coloured graphs, we first show that in order to prove Theorem 1.3, it is enough to verify the following lemma, which is another approximately graceful result. Similar to the style of Theorem 1.2, we consider an extra relaxation on the number of labels used.

[1] One can argue similarly to our proof that Lemma 1.4 implies Theorem 1.3.

Lemma 1.4. *For all $\varepsilon > 0$ there exists N such that for all $n > N$ the following holds. If T is an n-vertex tree then there exists an injective mapping $\phi : V(T) \to \{1, \dots, (1+\varepsilon)n\}$ such that the set of edge-differences of ϕ has size at least $(1-\varepsilon)n$.*

Proof of Theorem 1.3 from Lemma 1.4. We can assume $\varepsilon < 1$ as else the result is trivial, and let n_0 be sufficiently large so that for all $n > n_0$, the value of $(1-\varepsilon/2)n$ is at least as large as the output N of Lemma 1.4 when applied with $\varepsilon/2$. Let T be an n-vertex tree. One-by-one, for $\varepsilon n/2$ steps, remove a leaf from the current subtree of T that remains, to obtain a subtree T^* on $n^* \coloneqq (1-\varepsilon/2)n$ vertices. Apply Lemma 1.4 to T^* with $\varepsilon/2$ playing the role of ε to obtain an injective mapping $\phi^* : V(T^*) \to \{1, \dots, (1+\varepsilon/2)n^*\}$ such that the set of edge-differences of ϕ^* has size at least $(1-\varepsilon/2)n^*$. Since $(1+\varepsilon/2)n^* \leqslant n$ then from this we can define a bijective labelling $\phi : V(T) \to \{1, \dots, n\}$ such that $\phi(x) = \phi^*(x)$ for all $x \in V(T^*)$ and the vertices in $V(T) \setminus V(T^*)$ are bijectively assigned a label from the set of unused labels in $[n]$. Since $E(T^*) \subseteq E(T)$ and $|\phi^*(x) - \phi^*(y)| = |\phi(x) - \phi(y)|$ for all $xy \in E(T^*)$, the set of edge-differences of ϕ is at least as large as the set of edge differences of ϕ^*. Therefore $\mathrm{gs}(T) \geqslant (1-\varepsilon/2)n^* \geqslant (1-\varepsilon)n$, as desired. □

We remark that the relaxation given in Theorem 1.3 is a slightly weaker notion than that of a range-relaxed graceful labelling in Theorem 1.2, in the sense that if a tree has a range-relaxed graceful labelling using $(1+o(1))|T|$ labels, then it has gracesize $(1-o(1))|T|$ (this can be seen via the reduction used in the proof that Lemma 1.4 implies Theorem 1.3), while there is no obvious argument that gracesize $(1-o(1))|T|$ implies the existence of a range-relaxed graceful labelling using $(1+o(1))|T|$ labels. On the other hand, unlike Theorem 1.2, Theorem 1.3 holds for all large trees, including those with high degree vertices.

Combining these two directions, a natural next step towards proving Conjecture 1.1 would be to show that every tree T admits a range-relaxed graceful labelling into $[(1+o(1))|T|]$, either via a reduction from gracesize, or by an independent argument.

1.2 Rainbow reformulation

Recall that we are interested in using edge-coloured graphs to study the graceful tree conjecture. To do this, we can define the difference coloured complete graph, as follows.

Definition 1.5. *Let $X \subseteq \mathbb{N}$. The difference coloured complete graph for X, denoted by K_X, is the edge-coloured complete graph on vertex set X, where edge ij is assigned colour $|i-j|$ for all distinct $i, j \in X$.*

Consider the case when $X = [n] = \{1, \dots, n\}$. We can ask whether $K_{[n]}$ contains a rainbow copy of an n-vertex tree, meaning that every edge of the tree has a distinct colour. Given a tree T and an injective function $\phi : V(T) \to [n]$, an embedding of T in $K_{[n]}$ is uniquely determined from ϕ by embedding a given vertex x to $\phi(x) \in V(K_{[n]})$. Note that an edge $xy \in E(T)$ has colour c under this embedding if and only if $c = |\phi(x) - \phi(y)|$. Similarly an embedding of T uniquely determines such an injective function. In particular if T has n vertices, then observe that ϕ is graceful if and only if the corresponding embedding of T is rainbow. Therefore we can use $K_{[n]}$ to consider the following equivalent version of the graceful tree conjecture.

Conjecture 1.6 (Rainbow version of the graceful tree conjecture)**.** *$K_{[n]}$ contains a rainbow copy of every n-vertex tree T.*

In this context, our main result (Theorem 1.3) corresponds to the statement that every n-vertex tree T can be embedded into $K_{[n]}$ such that at least $(1-o(1))n$ distinct colours appear on its edges. Since we have already shown that Theorem 1.3 holds assuming the correctness of Lemma 1.4, proving this lemma will be the main focus for the remainder of this paper. We wish to attack it from a coloured perspective, so, rather than proving Lemma 1.4 directly, we will actually show that the following equivalent version holds.

Lemma 1.7. *For all $\varepsilon > 0$ there exists n_0 such that for all $n > n_0$ the following holds. If T is an n-vertex tree, then there exists an embedding of T in $K_{[(1+\varepsilon)n]}$ where at least $(1-\varepsilon)n$ distinct colours appear on edges of T.*

2 Preliminaries

2.1 Organisation of the paper

In the remainder of this section, we introduce notation and some standard probabilistic tools. Section 3 will give an overview of the proof strategy for Lemma 1.7, discussing the main methods needed to (i) find a nice structure in a tree T, and (ii) use the properties of this structure to embed parts of T in a rainbow way into the difference coloured complete graph. These two points will be addressed in Sections 4 and 5 respectively. In Section 6, we combine everything to prove Lemma 1.7.

2.2 Notation

For all $n \in \mathbb{N}$, let $[n] \coloneqq \{1, \ldots, n\}$. For any $a, b, c \in \mathbb{R}$ we write $a = b \pm c$ to mean $b - c \leqslant a \leqslant b + c$. If we say that a statement holds whenever $0 < a \ll b \leqslant 1$, then there exists a non-decreasing function $f : (0, 1] \to (0, 1]$ such that the statement holds for all $0 < a, b \leqslant 1$ with $a \leqslant f(b)$. We similarly consider a hierarchy of constants $0 < b_1 \ll b_2 \ll \ldots \ll b_k < 1$, and these constants must be chosen from right to left. For every constant b in a hierarchy it will be implicitly assumed that both $b \in (0, 1)$ and $1/b \in \mathbb{N}$ are satisfied.

If G is a graph we denote its vertex set and edge set by $V(G)$ and $E(G)$ respectively, and use the notation $|G| \coloneqq |V(G)|$ and $e(G) \coloneqq |E(G)|$. The degree and neighbourhood of a vertex $v \in V(G)$ are denoted by $d_G(v)$ and $N_G(v)$ respectively, and the neighbourhood of a set $U \subseteq V(G)$ is $N_G(U) = \bigcup_{u \in U} N_G(u)$. We may omit the subscripts where context is clear. The minimum degree of G is $\delta(G)$ and the maximum degree is $\Delta(G)$. Given a set $U \subseteq V(G)$, we write $G[U]$ for the induced subgraph of G on U. We use $G \setminus U$ to denote the subgraph $G[V(G) \setminus U]$, and for a single vertex v, we write $G - v$ instead of $G \setminus \{v\}$. For an edge subset $M \subseteq E(G)$, $V(M)$ denotes the set of vertices which are contained in some edge belonging to M, and we say these vertices are covered by M. For any edge $xy \in E(G)$, $G - xy$ is the spanning subgraph obtained by deleting the edge xy from G. If a graph G is edge-coloured, then we denote by $C(G)$ the set of colours used on some edge in G. For a set S of colours, we say that G is S-rainbow if all edges in G have a distinct colour from the set S.

For graphs F and G, we write $F + G$ to denote the disjoint union of F and G. The sum is defined analogously for

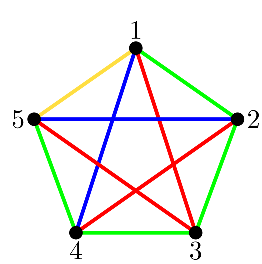

Figure 1: $K_{[5]}$

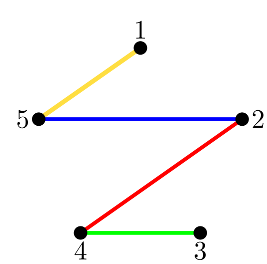

Figure 2: Rainbow copy of P_5

more than two graphs. For $d \in \mathbb{N}$, we write $F \times d$ to mean the graph obtained by taking the disjoint union of d copies of F, that is, $F \times d \cong F + (F + (F + \ldots))$ where the sum is taken d times.

For a hypergraph H we use the analogous notation $V(H), E(H), |H|, e(H), \delta(H), \Delta(H)$ as for graphs, whereby degree will always refer to vertex degree. H is r-uniform if each of its edges has size r, and it is k-partite if $V(H)$ can be partitioned into k parts $U_1, \ldots, U_k$, such that $|e \cap U_i| \leqslant 1$ for every $e \in E(H),\ i \in [k]$. A hypergraph H is said to be linear if any pair of distinct edges in H share at most one common vertex.

2.3 Probabilistic tools

We will need some common probabilistic tools to show the existence of specified properties in edge-coloured graphs and hypergraphs. We write $X \sim \mathrm{Bin}(n, p)$ to mean X follows the Binomial distribution with parameters n and p, and write $\mathbb{P}[\cdot]$ and $\mathbb{E}[\cdot]$ to denote the probability and expectation of an event respectively. Given a set N and $p \in [0, 1]$, a *p-random subset* $S \subseteq N$ is one that is formed by independently keeping every element in N with probability p. Similarly if $p \leqslant 1/k$, then $S_1 \ldots, S_k$ are *pairwise disjoint p-random subsets* of N if they are formed by independently placing each element of N in at most one S_i so that the element is placed in S_1 with probability p, in S_2 with probability p, and so on, and placed in none of the sets with probability $1 - kp$. These sets form a *p-random partition* of N if $p = 1/k$.

Theorem 2.1 (Chernoff bound, see e.g. [10])**.** *Let $X \sim \mathrm{Bin}(n, p)$. For all $\lambda \in (0, 1)$,*

$$\mathbb{P}[|X - \mathbb{E}[X]| \geqslant \lambda \mathbb{E}[X]] \leqslant 2e^{-\frac{\lambda^2 \mathbb{E}X}{3}}.$$

Let $Z_1, \ldots, Z_n$ be independent random variables, each taking values in a set Ω. For a constant c, we say that a function $f : \Omega^n \to \mathbb{R}$ is *c-Lipschitz* if for any $\omega = (\omega_1, \ldots, \omega_n) \in \Omega^n$, changing the outcome of at most one ω_i affects the value of $f(\omega)$ by at most c.

Theorem 2.2 (McDiarmid's inequality [13])**.** *Let $Z_1, \ldots, Z_n$ be independent random variables, each taking values in a set Ω. Let $f : \Omega^n \to \mathbb{R}$ be a c-Lipschitz function and consider the random variable $X = f(Z_1, \ldots, Z_n)$. For all $t > 0$,*

$$\mathbb{P}[|X - \mathbb{E}[X]| \geqslant t] \leqslant 2e^{-\frac{2t^2}{c^2 n}}.$$

3 Proof strategy for Lemma 1.7

We have already observed that a graceful labelling of a tree T with labels in $[n]$ can be thought of interchangeably as a rainbow embedding of T in $K_{[n]}$. Therefore throughout, we will regard labellings and (edge-coloured) embeddings of graphs as equivalent, where the colour of an edge is uniquely determined by taking the absolute difference of the labels assigned to the endvertices of the edge. More explicitly, given two adjacent vertices $x, y \in V(T)$ for some tree T together with a mapping $\phi : V(T) \to [n]$, we will refer to the colour of the edge xy to mean the value of $|\phi(x) - \phi(y)|$.

Let us now give an overview of the main ideas used in our strategy to embed a given n-vertex tree T into $K_{[(1+\varepsilon)n]}$ in an almost rainbow way. First, in Section 4, we find a helpful structure within T which is easier to work with. We split up T by first isolating a set $S_{\mathrm{high}} \subseteq V(T)$, consisting of all vertices of high degree. There cannot be too many high degree vertices, since $e(T) = n - 1$, so we have an upper bound for $|S_{\mathrm{high}}|$. We find a structural argument Lemma 4.1 which roughly states the following: for any subset of vertices $S \subseteq V(T)$ such that $|S|$ is not too large, there exists a set of vertices $W \subseteq V(T) \setminus S$, also of small order, such that deleting $S \cup W$ from T yields many vertex-disjoint copies of some small forest F. Applying this with S_{high} playing the role of S, we can think of the set W as a set of 'waste' vertices in T, of which the lemma tells us there are not too many. So, we have that $T \setminus (S_{\text{high}} \cup W)$ is exactly made up of many copies of some small forest F.

Our new aim is to find a fully rainbow copy of $T \setminus W$ in $K_{[(1+\varepsilon)n]}$, since at the end, the set W will be embedded arbitrarily to the available leftover labels, and this is where we may obtain colour repetitions. There will not be too many repeats since all vertices in W will have low degree, so that the number of edges with an endvertex in W is small. The embedding of $T \setminus W$ is the main focus in Section 5, with the key lemma given as Lemma 5.9. So, let us now summarise this process, noting that $T \setminus W$ is comprised of the set S_{high} and many vertex-disjoint copies of F. We define an auxiliary tree T_{aux} from $T \setminus W$ obtained by contracting the set S_{high} into a single vertex v, and contracting sets of vertices in the copies of F (chosen so that these sets of vertices are copies of the same vertex of F) so that the resulting tree T_{aux} consists of one high degree vertex v adjacent to constantly many copies of F.

We observe that for any integer d, a copy of $F \times d$, that is, the union of d vertex-disjoint copies of F, can be constructed by taking the union of $e(F)$ pairwise edge-disjoint matchings, each of size d, and each of which corresponds to a distinct edge of F. Thus we find tools for embedding rainbow matchings in Lemma 5.4, and use these to find a rainbow embedding of T_{aux} inside $K_{[(1+\varepsilon/2)|T_{\text{aux}}|]}$. We consider a rainbow 'blow up' of this auxiliary tree in $K_{[(1+\varepsilon)n]}$ by thinking of each vertex as a set (these will correspond to the contracted vertex sets) and by replacing each edge with some rainbow matching. With the flexibility from the extra labels, this allows us to manipulate the properties of the blow up and be more careful with which sections we are allowed to embed into, in order to obtain a rainbow embedding of $T \setminus W$ in $K_{[(1+\varepsilon)n]}$, as desired.

For this argument to work we require that both $|S_{\text{high}}|$ is small, and all vertices in W have low degree. The former condition holds using the fact that all vertices in S_{high} have degree bigger than some constant Δ, so since $e(T) = n - 1$ there must be fewer than $2n/\Delta$ of them. But if we choose Δ to be too high, then we cannot get a good upper bound for the maximum degree of the waste vertices. Therefore, for these conditions to be compatible, we need to be particularly careful with how we select the parameters which define these sets. More detail of this is given prior to the full proof, in Section 6.

4 Tree splitting

As mentioned in the previous section, in order to prove Lemma 1.7, we start by finding a way to delete certain vertices in a tree to yield many copies of some forest F, satisfying various additional properties. In this section, our main aim is to prove the following lemma.

Lemma 4.1. *For all $\delta > 0$ there exist $\zeta_0, n_0 > 0$ such that the following holds for all $n > n_0$ and $\zeta < \zeta_0$ satisfying $\zeta n \in \mathbb{N}$. Let T be an n-vertex tree and let $S \subseteq V(T)$ have order at most $\delta n/10$. There exists a set $W \subseteq V(T) \setminus S$ of order at most δn such that $T \setminus W$ satisfies the following properties:*

- (i) *$T \setminus (W \cup S) \cong F \times \zeta n$ for some forest F which has rooted trees as components; and*
- (ii) *for every component of $F \times \zeta n$, the root has at most one neighbour in S, and all other vertices in the component have no neighbour in S.*

We need to introduce some further results that demonstrate various properties about trees and forests, which we will then combine to prove Lemma 4.1 at the end of this section.

Lemma 4.2. *Let T be an n-vertex tree. For every $S \subseteq V(T)$, there is a set $W \subseteq V(T) \setminus S$ satisfying $|W| \leqslant 6|S|$, such that for every component C of $T \setminus (S \cup W)$, there is at most one edge between C and S.*

Proof. Let T_S be the smallest subtree of T which contains all vertices in S. Let $W \coloneqq N_{T_S}(S) \setminus S$. We need to find an upper bound for $|W|$, and will require the following claim.

Claim 4.3. *For any tree H, the number of leaves in H is equal to $\sum_{v: d_H(v) \geqslant 3}(d_H(v) - 2) + 2$.*

Proof of claim. Let ℓ denote the number of leaves in H. By the handshaking lemma we know

$$\begin{aligned} 2e(H) = \sum_v d_H(v) &= \sum_{v: d_H(v) \geqslant 3} d_H(v) + \sum_{v: d_H(v) = 2} d_H(v) + \sum_{v: d_H(v) = 1} d_H(v) \\ &= \sum_{v: d_H(v) \geqslant 3} d_H(v) + \sum_{v: d_H(v) = 2} 2 + \sum_{v: d_H(v) = 1} 1. \end{aligned} \tag{4.1}$$

Expanding out the left hand-side gives

$$2e(H) = 2|H| - 2 = \left(\sum_{v: d_H(v) \geqslant 3} 2 + \sum_{v: d_H(v) = 2} 2 + \sum_{v: d_H(v) = 1} 2 \right) - 2. \tag{4.2}$$

Combining (4.1) and (4.2) and rearranging, we have that

$$\ell = \sum_{v: d_H(v) = 1} 1 = \sum_{v: d_H(v) \geqslant 3} (d_H(v) - 2) + 2,$$

as required. This proves the claim. ■

Since we chose T_S minimally, we must have that all leaves in T_S belong in S. Otherwise we can delete any leaf outside of this set and find a smaller subtree still containing S. Thus the number of leaves in T_S is at most $|S|$. By applying Claim 4.3 to the tree T_S, it follows that

$$\sum_{\substack{v \in S \\ d_{T_S}(v) \geqslant 3}} (d_{T_S}(v) - 2) \leqslant \sum_{\substack{v \in V(T_S) \\ d_{T_S}(v) \geqslant 3}} (d_{T_S}(v) - 2) = (\#\text{ leaves in } T_S) - 2 \leqslant |S| - 2 \leqslant |S|,$$

and rearranging gives

$$\sum_{\substack{v \in S \\ d_{T_S}(v) \geqslant 3}} d_{T_S}(v) \leqslant |S| + \sum_{\substack{v \in S \\ d_{T_S}(v) \geqslant 3}} 2 \ \leqslant |S| + 2|S| = 3|S|.$$

We use this upper bound to deduce that

$$\begin{aligned} |W| \leqslant |N_{T_S}(S)| \leqslant \sum_{v \in S} d_{T_S}(v) &= \sum_{\substack{v \in S \\ d_{T_S}(v) = 1}} 1 + \sum_{\substack{v \in S \\ d_{T_S}(v) = 2}} 2 + \sum_{\substack{v \in S \\ d_{T_S}(v) \geqslant 3}} d_{T_S}(v) \\ &\leqslant |S| + 2|S| + 3|S| \\ &\leqslant 6|S|. \end{aligned}$$

Suppose there exists a component C in $T \setminus (S \cup W)$ which sends two distinct edges into S, say xs and yt for some $x, y \in V(C)$, $s, t \in S$. Since T_S and C are both connected, then there is a path P from s to t contained in T_S, and there is a path Q from x to y contained in C. We know that at least one of these paths is non-empty as otherwise $x = y$, $s = t$ implies $xs = yt$. If $V(P) \cap V(Q) = \emptyset$, then $xQytPsx$ forms a cycle of length at least 3 in T, a contradiction. Otherwise, choose $z \in V(P) \cap V(Q)$ to be of minimal distance from x along the path Q. Then the paths xQz and zPs are edge-disjoint. Since $z \in V(Q) \subseteq C$, then $z \notin S \cup N_{T_S}(S)$, implying that zPs is a path of length at least 2. Therefore $xQzPsx$ forms a cycle of length at least 3 in T, again a contradiction. We can therefore conclude that every component in $T \setminus (S \cup W)$ touches at most one edge going into S. □

We use the following standard observation about trees, providing a proof for completeness.

Fact 4.4. *Every n-vertex tree T contains a vertex $v \in V(T)$ such that all components of $T - v$ have order at most $n/2$.*

Proof. Let us consider an orientation of the tree T given as follows: for every $xy \in E(T)$, direct the edge from x to y if the component of $T - xy$ containing x is smaller than the one containing y (and orient the edge xy arbitrarily if these components have the same size). Every oriented tree contains a sink, that is, a vertex $v \in V(T)$ where v has only in-neighbours. Any component C in $T - v$ contains some neighbour u of v since T is connected, and thus C is smaller than the component of $T - uv$ which contains v. So, $|C| \leqslant n/2$. □

This fact provides a means to prove the inductive step of the next straightforward lemma.

Lemma 4.5. *Let $k \in \mathbb{N} \cup \{0\}$. For every n-vertex tree T there exists a set $W \subseteq V(T)$ of order at most $1+2+\cdots+2^k$, such that all components of $T \setminus W$ have order at most $n/2^k$.*

Proof. We proceed by induction on k. The case $k = 0$ holds trivially. Suppose the statement holds true for some k, and we want to prove it for $k+1$. Consider the set W' obtained by applying the result for k, so that $|W'| \leqslant 1 + 2 + \ldots + 2^k$ and all components of $T \setminus W'$ have order at most $n/2^k$. Let $C_1, \ldots, C_t$ denote the set of components of $T \setminus W'$ which have order larger than $n/2^{k+1}$. Clearly $t \leqslant 2^{k+1}$ since these components are pairwise vertex-disjoint and $|T| = n$. For each $i \in [t]$, we apply Fact 4.4 to C_i to find a vertex $v_i \in V(C_i)$ such that all components of $C_i - v_i$ have order at most $|C_i|/2 \leqslant n/2^{k+1}$. Then $W \coloneqq W' \cup \{v_1, \ldots, v_t\}$ has order $|W'| + t \leqslant 1 + 2 + \cdots + 2^k + 2^{k+1}$ and all components of $T \setminus W$ have order at most $n/2^{k+1}$, as required. □

Lemma 4.6. *Let $m, n \in \mathbb{N}$. For every n-vertex tree T there exists a set $W \subseteq V(T)$ satisfying $|W| \leqslant \frac{4n}{m}$ and such that every component of $T \setminus W$ has order at most m.*

Proof. Let $k \coloneqq \lceil \log_2(n/m) \rceil$. Applying Lemma 4.5 to T we have that there exists a set W of order at most $1 + 2 + \cdots + 2^k = 2^{k+1} - 1 \leqslant 2^{\log_2(n/m)+2} = 4n/m$, such that all components of $T \setminus W$ have order at most $n/2^k \leqslant n/2^{\log_2(n/m)} = m$. □

Finally we find a way to delete vertices from a forest H so that the subgraph obtained consists precisely of disjoint copies of a new forest F with a specified rooted structure.

Lemma 4.7. *Let $m, n \in \mathbb{N}$, $\zeta > 0$ be such that $\zeta n \in \mathbb{N}$ and suppose H is a forest with components of order at most m and in each component we have specified a root. We can delete at most $m^{m+1}\zeta n$ vertices from H to get a copy of $F \times \zeta n$ for some forest F which has rooted trees as components, and such that all roots in $F \times \zeta n$ were also roots in H.*

Proof. By Cayley's formula there exist at most m^{m-1} trees of order at most m, and for each of them there are at most m ways to select a root. So there are at most m^m types of rooted trees amongst the components of H. For each such rooted tree T, we delete at most ζn copies of T so that the remaining number of copies is divisible by ζn. The remaining graph is a copy of $F \times \zeta n$ for some forest F where all components are rooted trees. In total we have deleted at most $m^m \times m \times \zeta n = m^{m+1}\zeta n$ vertices to obtain this. □

We are now ready to prove Lemma 4.1, combining the properties we have seen already.

Proof of Lemma 4.1. Let $\delta > 0$, $m \coloneqq \frac{20}{\delta}$, $\zeta_0 \coloneqq \frac{\delta}{5m^{m+1}}$ and let n_0 be sufficiently large. Let $n > n_0$ and $\zeta < \zeta_0$ be such that $\zeta n \in \mathbb{N}$., and let T and S be as in the statement of the lemma. We construct W by deleting a bounded number of vertices. By Lemma 4.2, there exists a set $W_1 \subseteq V(T) \setminus S$ of order at most $6|S|$ such that every component of $T \setminus (S \cup W_1)$ touches at most one edge that is incident to S. By Lemma 4.6 there exists a set $W_2 \subseteq V(T)$ of order at most $4n/m$ such that all components in $T \setminus W_2$ have order at most m. So, each component in $H \coloneqq T \setminus (S \cup W_1 \cup W_2)$ has at most one vertex belonging in $N_T(S)$, and if such a vertex exists then call this the root, otherwise arbitrarily choose a root within the component. Then we can apply Lemma 4.7 to H to obtain a set $W_3 \subseteq V(H)$ of order at most $m^{m+1}\zeta n$ such that $H \setminus W_3 = T \setminus (S \cup W_1 \cup W_2 \cup W_3) \cong F \times \zeta n$ for some forest F with rooted trees as components, where only the roots belong in $N_T(S)$. Clearly every component of $H \setminus W_3$ is a subgraph of a component of H, so using the property of W_1, we also know that every root has at most one neighbour in S. Take $W \coloneqq (W_1 \cup W_2 \cup W_3) \setminus S$ to get $|W| \leqslant 6|S| + \frac{4n}{m} + m^{m+1}\zeta n \leqslant \left(\frac{6\delta}{10} + \frac{2\delta}{10} + \frac{2\delta}{10}\right) n = \delta n$. Therefore $T \setminus W$ satisfies the desired conditions. □

5 Finding rainbow subgraphs

5.1 Using hypergraphs for matchings

Observe that given a forest F and some integer d, a rainbow copy of $F \times d$ is exactly the union of $e(F)$ colour-disjoint rainbow matchings, where each matching has size d and corresponds to a unique edge in F. By using Lemma 4.1, if we can find rainbow matchings in $K_{[n]}$, this will give us tools for embedding parts of trees in a rainbow way. It will be helpful to consider 3-uniform 3-partite hypergraphs by making the following observation. Consider two disjoint vertex sets $A, B \subseteq [n]$ and a colour set $C \subseteq C(K_{[n]})$, let $K_{[n]}[A, B, C]$ denote the edge-coloured subgraph of $K_{[n]}$ containing only edges between A and B which have colour in C. We can consider the corresponding 3-uniform 3-partite hypergraph, denoted by $H_{[n]}[A, B, C]$, with vertex classes A, B and C such that for all $a \in A,\ b \in B,\ c \in C$, the size-three edge abc is an edge in $H_{[n]}[A, B, C]$ if and only if the edge $ab \in E(K_{[n]})$ has colour c, or equivalently $c = |b - a|$. In particular, there is a mapping $\theta : E(K_{[n]}[A, B, C]) \to E(H_{[n]}[A, B, C])$ which maps an edge $ab \in E(K_{[n]}[A, B, C])$ to the size-three edge $\{a, b, |b-a|\} \in E(H_{[n]}[A, B, C])$ where $|b - a| \in C$. Thus we observe the following.

Observation 5.1. *A subset $M \subseteq E(K_{[n]}[A, B, C])$ is a rainbow matching if and only if $\theta(M)$ is a hypergraph matching in $H_{[n]}[A, B, C]$.*

In order to use this observation, we will first need some preliminary lemmas about $K_{[n]}[A, B, C]$.

Lemma 5.2. *Let $s, n \in \mathbb{N}$ be such that $s \leqslant n/2$ and n is even. Let $A \coloneqq [1, \frac{n}{2}]$, $B \coloneqq [\frac{n}{2}+1, n]$ and $C \coloneqq [n-1]$. There exists a spanning subgraph $G \subseteq K_{[n]}[A, B, C]$ such that $\Delta(G) = 2s$, every vertex in $[2s, \frac{n}{2} - s] \cup [\frac{n}{2} + 2s, n - s]$ has degree $2s$, every colour in $[2s, n-2s]$ appears on exactly s edges, no colour in C appears on more than s edges, and no edge has colour 1.*

Proof. Note that $C = C(K_{[n]})$ so C contains all possible colours in $K_{[n]}$. We will find our desired spanning subgraph G to be contained in $K_{[n]}[A, B, C]$, meaning that all edges will lie between A and B.

For every $c \in C$, let $E_c \coloneqq \left\{\left(\frac{n-c+i}{2}, \frac{n+c+i}{2}\right) : i = 1, 2, \ldots, 2s\right\} \cap \mathbb{Z}^2$. Then $|E_c| = s$ for all $c \in C$. Indeed, both of the values $n - c$ and $n + c$ are odd or they are both even, implying that the terms $\frac{n-c+i}{2}$ and $\frac{n+c+i}{2}$ are integer valued exactly when i is odd or even respectively. Furthermore any edge in $E_c \cap E(K_{[n]})$ has colour c, since for a fixed i, we have $\frac{n+c+i}{2} - \frac{n-c+i}{2} = \frac{2c}{2} = c$.

Now, let $C' \coloneqq [2s, n-2s]$ and fix $c \in C'$. For any $i \in [2s]$, note that

$$\frac{n-c+i}{2} \geqslant \frac{n-(n-2s)+i}{2} \geqslant \frac{2s+1}{2} \geqslant 1,$$
$$\frac{n-c+i}{2} \leqslant \frac{n-2s+i}{2} \leqslant \frac{n-2s+2s}{2} = \frac{n}{2},$$
$$\frac{n+c+i}{2} \geqslant \frac{n+2s+i}{2} \geqslant \frac{n+2s+1}{2} \geqslant \frac{n}{2}+1,$$
$$\frac{n+c+i}{2} \leqslant \frac{n+(n-2s)+i}{2} \leqslant \frac{2n-2s+2s}{2} = n.$$

This shows that $E_c \subseteq A \times B$ for all $c \in C'$, i.e. every pair in E_c forms an edge in $K_{[n]}$ of colour c, going from A to B. Let G be the spanning subgraph of $K_{[n]}$ defined by its edge set

$$E(G) = \bigcup_{c \in C \setminus \{1\}} E_c \cap (A \times B),$$

So G can be constructed by first taking edges from all $E_c \cap E(K_{[n]})$ sets, except for when $c = 1$, and deletes edges not lying between A and B. Thus no edge in G has colour 1. Note that if n is odd then $E_1 \cap (A \times B) = \emptyset$, and if n is even, $E_1 \cap (A \times B) = \{(n/2, n/2+1)\}$, so by not adding this edge set to G we only affect the degrees of at most two vertices, namely $n/2$ and $n/2+1$, neither of which belong in $[2s, \frac{n}{2}-s] \cup [\frac{n}{2}+2s, n-s]$. For any $c \in C'$, since $E_c \subseteq A \times B$, then this implies $E_c \subseteq E(G)$ and in particular every colour $c \in C'$ occurs on exactly $|E_c| = s$ edges of G, as desired. Any colour in $C \setminus C'$ clearly appears on at most s edges. It remains to check the degrees of vertices in A and B.

First consider $A' \coloneqq [2s, \frac{n}{2}-s]$, and fix $x \in A'$. Let $C(x) \coloneqq \{c \in C : \frac{n-c+i}{2} = x \text{ and } \frac{n+c+i}{2} \in B \text{ for some } i \in [2s]\}$. Note that

$$\frac{n-c+i}{2} = x \iff c = n+i-2x,$$

so we can deduce that any $c \in C(x)$ satisfies $c \leqslant n+2s-2x \leqslant n+2s-4s = n-2s$ and $c \geqslant n+1-2x \geqslant n+1-(n-2s) = 1+2s$. This implies $C(x) \subseteq C'$ and $1 \notin C(x)$. In particular $C(x)$ is the set of colours $c \in C$ for which there exists $y \in B$ such that $xy \in E(G)$ has colour c. If there exist two edges in G of colour c that contain x, then there exist distinct $y, y' \in B$ such that xy and xy' are edges of colour c. By choice of A and B, we know $y \geqslant x$ and $y' \geqslant x$ and so $c = y - x = y' - x$, implying $y = y'$, a contradiction. Thus for each colour $c \in C(x)$, there is exactly one edge in G of colour c with x as an endvertex. So in total the number of edges in G containing x is $|C(x)| = 2s$, and in particular we have $d_G(x) = 2s$ for every $x \in A'$.

Similarly, let $B' = [\frac{n}{2}+2s, n-s]$, and fix $y \in B'$. Let $D(y) \coloneqq \{c \in C : \frac{n+c+i}{2} = y \text{ and } \frac{n-c+i}{2} \in A \text{ for some } i \in [2s]\}$. Note that

$$\frac{n+c+i}{2} = y \iff c = 2y-(n+i),$$

so we deduce that any $c \in D(y)$ satisfies $c \leqslant 2y-(n+1) \leqslant 2n-2s-n-1 = n-2s-1$ and $c \geqslant 2y-(n+2s) \geqslant n+4s-n-2s = 2s$. Thus $D(y) \subseteq C'$ and $1 \notin D(y)$. We may argue similarly that there exists exactly one edge in G of each colour in $D(y)$ which contains y. So there are $|D(y)| = 2s$ edges in G containing y and we deduce that $d_G(y) = 2s$ for all $y \in B'$. Finally all vertices in $(A \setminus A') \cup (B \setminus B')$ are contained in at most $2s$ edges by choice of the E_c sets, so $\Delta(G) \leqslant 2s$. Thus the graph G satisfies our desired properties. □

Consider again the subgraph $K_{[n]}[A, B, C]$ of $K_{[n]}$ for disjoint vertex sets $A, B \subseteq [n]$ and a colour set $C \subseteq C(K_{[n]})$. If every element of A is smaller than every element of B, then observe that the corresponding hypergraph $H_{[n]}[A, B, C]$ is linear. We would like to find matchings in this hypergraph, in accordance with Observation 5.1. We will use the following variant of an unpublished result from Pippenger (see e.g. [2]) in order to find almost perfect matchings in linear uniform hypergraphs.

Theorem 5.3. *Let $n^{-1} \ll \gamma \ll \alpha, \mu, r^{-1}$. Every n-vertex linear r-uniform hypergraph H with maximum degree at most $(1+\gamma)\alpha n$ and at most μn vertices of degree less than $(1-\gamma)\alpha n$ has a matching covering all but at most $2\mu n$ vertices.*

Proof. Let H be as in the statement of the lemma. A result of Molloy and Reed (see [14, Theorem 1]), tells us that the chromatic index $\chi'(H)$ is at most $\Delta + c\Delta^{1-1/r}(\log \Delta)^4$ for some constant c, where Δ is the maximum degree of H. Thus there exists a proper edge-colouring of H using at most $\chi'(H) \leqslant (1+\gamma)\alpha n + ((1+\gamma)\alpha n)^{1-1/(r+1)}$ colours, by choosing n sufficiently large. Each colour class forms a matching. Since at least $(1-\mu)n$ vertices in H have degree at least $(1-\gamma)\alpha n$, then $e(H) \geqslant (1-\mu)(1-\gamma)\alpha n^2/r$. Thus there is a colour class of size at least

$$\begin{aligned}\frac{e(H)}{\chi'(H)} \geqslant \frac{(1-\mu)(1-\gamma)\alpha n^2/r}{(1+\gamma)\alpha n + ((1+\gamma)\alpha n)^{1-\frac{1}{r+1}}} &= (1-\mu)\frac{n}{r}\left(\frac{1-\gamma}{1+\gamma+(1+\gamma)^{1-\frac{1}{r+1}}(\alpha n)^{-\frac{1}{r+1}}}\right)\\ &\geqslant (1-\mu)^2\frac{n}{r}\\ &\geqslant (1-2\mu)\frac{n}{r}.\end{aligned}$$

The second inequality holds by choosing n sufficiently large to satisfy $n \geqslant \frac{(1-\mu)(1+\gamma)^r}{\alpha(\mu-2\gamma-\mu\gamma)^{r+1}}$. Since this colour class contains at least $(1-2\mu)n/r$ pairwise disjoint edges, each of size r, then the number of vertices covered by this matching is at least $(1-2\mu)n$. Thus the matching covers all but at most $2\mu n$ vertices in H. □

Lemma 5.4. *Let $n^{-1} \ll p, \mu$ be such that n is even. Suppose $S_1, S_2 \subseteq [n]$ are p-random and disjoint. Then with probability $1-o(1)$ there exists a rainbow matching in $K_{[n]}[S_1, S_2, S_2]$ such that no edge in the matching has colour 1, and it covers all but at most $2\mu \max\{|S_1|, |S_2|\}$ vertices in S_1 and all but at most $2\mu \max\{|S_1|, |S_2|\}$ vertices in S_2.*

Proof. Choose $\gamma \in (0,1)$ such that $n^{-1} \ll \gamma \ll p, \mu$ and let n, S_1, S_2 be as in the statement of the lemma. Let $A \coloneqq [1, \frac{n}{2}]$, $B \coloneqq [\frac{n}{2}+1, n]$ and $C \coloneqq [n-1]$. Let $m \coloneqq 3pn/2$ and $s \coloneqq \mu m/10$. Let $G \subseteq K_{[n]}[A, B, C]$ denote the spanning subgraph obtained by applying Lemma 5.2 and take $R \subseteq H_{[n]}[A, B, C]$ to be the 3-partite 3-uniform hypergraph corresponding to G, that is, there is an edge $abc \in E(R)$ if and only if $c = |b-a|$ and $ab \in E(G)$ for some $a \in A$, $b \in B$ and $c \in C$. We will consider two vertex-disjoint colour-disjoint random subhypergraphs of R and find large matchings within both of these. We will then combine them to form an almost perfect matching M in $H_{[n]}[S_1, S_2, S_2]$. By Observation 5.1, this then corresponds to an almost perfect rainbow matching $\theta^{-1}(M)$ in $K_{[n]}[S_1, S_2, S_2]$. For that purpose, let us define various random subsets of vertices.

For $i \in \{1,2\}$, let $A_i \coloneqq A \cap S_i$ and $B_i \coloneqq B \cap S_i$. Partition C into two $\frac{1}{2}$-random subsets P and Q, and define $P_2 \coloneqq S_2 \cap P$ and $Q_2 \coloneqq S_2 \cap Q$. Note that, for any $a \in A$, $b \in B$ and for any $i, j \in [2]$, we have $\mathbb{P}[a \in A_i] = \mathbb{P}[b \in B_j] = p$ and the corresponding events are independent. Also, for any $c \in C$, we have $\mathbb{P}[c \in P_2] = \mathbb{P}[c \in Q_2] = \frac{p}{2}$. We will consider two 3-partite 3-uniform subhypergraphs of R induced on these subsets, defined by $H_P \coloneqq R[A_1, B_2, P_2]$ and $H_Q \coloneqq R[A_2, B_1, Q_2]$. As noted earlier, both H_P and H_Q are linear hypergraphs, since all elements in A are smaller than all elements in B. We claim that these hypergraphs both contain a large matching with high probability.

Claim 5.5. *With probability $1-o(1)$ there is both a hypergraph matching $M_P \subseteq E(H_P)$ covering all but at most $2\mu|H_P|$ vertices of H_P, and a hypergraph matching $M_Q \subseteq E(H_Q)$ covering all but at most $2\mu|H_Q|$ vertices of H_Q.*

Proof of claim. We will only prove that with high probability the matching M_P with the desired property inside H_P exists, since one can then apply the identical method to H_Q to find the matching M_Q, and take a union bound to get that with high probability they both exist, as desired to prove the claim. In order to find M_P, we will use Theorem 5.3. To apply this, we first need to show that the hypergraph H_P satisfies various properties. We will show that the following properties hold with high probability.

(P1) $|H_P| = (1 \pm \gamma)m$

(P2) H_P has maximum degree at most $(1+\sqrt{\gamma})\frac{\mu p^2}{10}|H_P|$.

(P3) All vertices in the set $\left(A_1 \cap \left[2s, \frac{n}{2}-s\right]\right) \cup \left(B_2 \cap \left[\frac{n}{2}+2s, n-s\right]\right) \cup (P_2 \cap [2s, n-2s])$ have degree in the interval $(1 \pm \sqrt{\gamma})\frac{\mu p^2}{10}|H_P|$ in H_P,

Note that $\mathbb{E}|H_P| = \mathbb{E}[|A_1|] + \mathbb{E}[|B_2|] + \mathbb{E}[|P_2|] = 3pn/2 = m$ by linearity of expectation. So a simple application of the Chernoff bound tells us that (P1) holds with probability at least $1 - 2e^{-\frac{\gamma^2 m}{3}} = 1 - 2e^{-\Omega(n)}$. In order to prove (P2) and (P3) hold with high probability, we will start by considering the degrees of vertices in H_P. We know that $\Delta(G) \leqslant 2s$, and every colour in G appears on at most s edges. In particular this means that $d_R(x) \leqslant 2s$ for every $x \in A \cup B$, and $d_R(c) \leqslant s$ for every $c \in C$. By choice of R we also have the additional properties that $d_R(x) = 2s$ for every $x \in \left(A \cap \left[2s, \frac{n}{2}-s\right]\right) \cup \left(B \cap \left[\frac{n}{2}+2s, n-s\right]\right)$, and $d_R(c) = s$ for every $c \in C \cap [2s, n-2s]$.

- Let $c \in C$. Suppose $abc \in E(R)$ for some $a \in A$ and $b \in B$. The events that $a \in A_1$ and $b \in B_2$ happen independently and with probability p each, so the probability that they both occur is p^2. Thus the expected number of edges $abc \in E(R)$ containing c for which $a \in A_1$ and $b \in B_2$ is $p^2 d_R(c) \leqslant p^2 s$. In particular, since this holds for all $c \in P_2$, and for a fixed $c \in P_2$, the expected degree of c in H_P is the expected number of edges $abc \in E(R)$ containing c for which $a \in A_1$ and $b \in B_2$, we have $\mathbb{E}[d_{H_P}(c)] \leqslant p^2 s$. Furthermore, note that the pairs $a \in A_1$ $b \in B_2$ for which $c = |b-a| = b-a$ are pairwise disjoint, so that the random variables $\mathbf{1}\{abc \in E(R)\}$ over all such pairs are independent. Applying a Chernoff bound (Theorem 2.1), for each $c \in P_2$ we have

$$\mathbb{P}(d_{H_P}(c) > (1+\gamma)p^2 s) \leqslant \mathbb{P}(|d_{H_P}(c) - \mathbb{E}[d_{H_P}(c)]| \geqslant \gamma p^2 s) \leqslant 2e^{-\frac{\gamma^2 p^2 s}{3}} = 2e^{-\Omega(n)}.$$

  Taking a union bound we obtain

$$\mathbb{P}\left(\bigcup_{c \in P_2}\left\{d_{H_P}(c) > (1+\gamma)p^2 s\right\}\right) \leqslant \sum_{c \in P_2} \mathbb{P}\left(d_{H_P}(c) > (1+\gamma)p^2 s\right) \leqslant 2ne^{-\Omega(n)}. \tag{5.1}$$

  Furthermore, for $c \in P_2 \cap [2s, n-2s]$, we have $\mathbb{E}[d_{H_P}(c)] = p^2 d_R(c) = p^2 s$ and by a Chernoff bound (Theorem 2.1) we have

$$\mathbb{P}(|d_{H_P}(c) - p^2 s| \geqslant \gamma p^2 s) \leqslant 2e^{-\frac{\gamma^2 p^2 s}{3}} \leqslant 2e^{-\Omega(n)}.$$

  Again taking a union bound gives

$$\mathbb{P}\left(\bigcup_{c \in P_2 \cap [2s, n-2s]}\left\{d_{H_P}(c) \neq (1 \pm \gamma)p^2 s\right\}\right) \leqslant 2ne^{-\Omega(n)}. \tag{5.2}$$

- Let $b \in B$. Suppose $abc \in E(R)$ for some $a \in A$ and $c \in C$. The events that $a \in A_1$ and that $c = b-a \in P_2$ happen independently with probabilities p and $p/2$ respectively, unless $b - a = a$, in which case there is a dependence. This latter case can only happen at most once, if $b = 2a$. Thus the expected number of edges $abc \in E(R)$ containing b for which $a \in A_1$ and $c \in P_2$ is $\frac{p^2}{2} d_R(b) \pm 1 \leqslant p^2 s + 1$. Since this number of edges is precisely the degree of a vertex $b \in B_2$ in H_P, we have $\mathbb{E}[d_{H_P}(b)] \leqslant p^2 s + 1$ for all $b \in B_2$.

  Consider the product space $\Omega = X^n$ with $X = \{v \in S_2, v \notin S_2\}$ and first suppose $b \in B_2$. Then $d_{H_P}(b)$ is a 2-Lipschitz function from Ω to $\mathbb{R}$. Indeed, for any $v \in [n]$, v is contained in at most two edges alongside b in the hypergraph R (at most one for each of the possibilities that $v \in A$ and $v \in C$), so changing the outcome of whether some vertex v belongs in S_2 or not changes the degree of b by at most 2. We apply McDiarmid's inequality (Theorem 2.2) to obtain

$$\mathbb{P}(d_{H_P}(b) > (1+\gamma)p^2 s) \leqslant \mathbb{P}(|d_{H_P}(b) - \mathbb{E}[d_{H_P}(b)]| \geqslant \gamma p^2 s - 1) \leqslant 2e^{-\frac{(\gamma p^2 s/2)^2}{8s}} \leqslant 2e^{-\Omega(n)}. \tag{5.3}$$

Again we can apply a union bound by ranging over all possible (at most $n/2$) choices for $b \in B_2$ to obtain that with probability at most $ne^{-\Omega(n)}$, some vertex in B_2 has degree greater than $(1+\gamma)p^2 s$ in H_P.

Secondly suppose $b \in B_2 \cap \left[\frac{n}{2} + 2s, n-s\right]$. Then $\mathbb{E}[d_{H_P}(b)] = \frac{p^2}{2} d_R(b) \pm 1 = p^2 s \pm 1$. Again we have that $d_{H_P}(b)$ is a 2-Lipschitz function from the same product space Ω to $\mathbb{R}$, and by the same inequality from the right hand side of 5.3, the probability that $d_{H_P}(b) \neq (1 \pm \gamma)p^2 s$ is at most $2e^{-\Omega(n)}$. Taking a further union bound, with probability at most $ne^{-\Omega(n)}$, some vertex in $B_2 \cap [\frac{n}{2} + 2s, n-s]$ has degree $\neq (1\pm\gamma)p^2 s$ in H_P.

- Let $a \in A$. An identical argument to the case above gives that the expected number of edges $abc \in E(R)$ containing a for which $b \in B_2$ and $c \in P_2$ is $\frac{p^2}{2} d_R(a) \pm 1 \leqslant p^2 s + 1$. Thus $\mathbb{E}[d_{H_P}(a)] \leqslant p^2 s + 1$ for all $a \in A_1$. In the same way, $d_{H_P}(a)$ is a 2-Lipschitz function from Ω to $\mathbb{R}$, on the same product space. Applying Theorem 2.2 and a union bound as before, with probability at most $ne^{-\Omega(n)}$, some vertex in A_1 has degree greater than $(1+\gamma)p^2 s$ in H_P.

  In the same way as before, the probability of some vertex in $A_1 \cap [2s, \frac{n}{2} - s]$ having degree $\neq (1\pm\gamma)p^2 s$ in H_P is at most $ne^{-\Omega(n)}$.

One further application of the union bound tells us that with probability $1 - o(1)$ all vertices in H_P have degree at most $(1+\gamma)p^2 s$, and combining this with (P1) we have that

$$\Delta(H_P) \leqslant (1+\gamma)p^2 s = (1+\gamma)\frac{p^2 \mu m}{10} \leqslant \left(\frac{1+\gamma}{1-\gamma}\right)\frac{p^2\mu}{10}|H_P| \leqslant (1+\sqrt{\gamma})\frac{p^2\mu}{10}|H_P|,$$

where the final inequality holds since $\frac{1+y}{1-y} \leqslant 1 + \sqrt{y}$ for all $0 < y \ll 1$. This proves (P2).

Combining inequalities, as a consequence of a union bound, we deduce that with probability at least $1 - o(1)$, all vertices in the following set have degree $(1 \pm \gamma)p^2 s$ in H_P.

$$\left(A_1 \cap \left[2s, \frac{n}{2} - s\right]\right) \cup \left(B_2 \cap \left[\frac{n}{2} + 2s, n - s\right]\right) \cup (P_2 \cap [2s, n-2s]).$$

Considering this interval of possible degrees, and using (P1) and our choice of γ, gives

$$(1\pm\gamma)p^2 s = (1\pm\gamma)\frac{p^2\mu m}{10} = \left(\frac{1\pm\gamma}{1\pm\gamma}\right)\frac{p^2\mu}{10}|H_P| = (1\pm\sqrt{\gamma})\frac{p^2\mu}{10}|H_P|.$$

This proves (P3). So, all of our objectives are satisfied with high probability, and we are finally ready to apply Theorem 5.3. Taking $\sqrt{\gamma}, \mu$, $\frac{p^2\mu}{10}$ and $|H_P|$ to play the roles of γ, μ α and n respectively, with high probability we can apply Theorem 5.3 to find a matching in H_P covering all but $2\mu|H_P|$ vertices. Denote it by M_P. As mentioned earlier, we can apply the exact same method to H_Q, and apply a union bound, to get that with probability $1 - o(1)$ both M_P and M_Q exist as in the statement of the claim. ■

Note that H_P and H_Q are vertex-disjoint and $V(H_P) \cup V(H_Q) = V(H_{[n]}[S_1, S_2, S_2])$. Then with probability $1 - o(1)$, there is a matching $M \coloneqq M_P \cup M_Q$ in $H_{[n]}[S_1, S_2, S_2]$, and with probability $1 - o(1)$, covering all but at most $2\mu(|H_P| + |H_Q|) = 2\mu|H_{[n]}[S_1, S_2, S_2]|$ vertices in $H_{[n]}[S_1, S_2, S_2]$. Since $H_{[n]}[S_1, S_2, S_2]$ is 3-uniform and 3-partite, then such a matching M covers the same number of vertices in each of its tripartition classes, and therefore covers all but at most $2\mu \max\{|S_1|, |S_2|\}$ vertices in each class. Each vertex class has size at least $\min\{|S_1|, |S_2|\}$. Finally we take $\theta^{-1}(M)$ to be the corresponding rainbow matching in the edge-coloured graph $K_{[n]}$ with all edges between S_1 and S_2 having colour in S_2. Thus we know that with probability $1 - o(1)$, this matching exists, and covers all but at most $2\mu \max\{|S_1|, |S_2|\}$ vertices in both S_1 and in S_2. By construction $\theta^{-1}(M) \subseteq G$ and by Lemma 5.2 no edge in G has colour 1. So it also follows that no edge in $\theta^{-1}(M)$ has colour 1, as desired. □

5.2 Embedding trees with a splitting vertex

We focus on trees with a specific structure that will be particularly useful for embedding the auxiliary tree defined in the proof of Lemma 5.9, and as mentioned in the strategy overview in Section 3.

Lemma 5.6. *Let $n^{-1} \ll \zeta \ll \varepsilon < 1$ and suppose that T is an $(n+1)$-vertex tree containing a vertex v such that all components of $T - v$ have order at most ζ^{-1}. Then there exists a rainbow copy of T in $K_{[(1+\varepsilon)n] \cup \{0\}}$ such that v receives label 0, and no edge has colour 1.*

Proof. Choose $\lambda \in (0,1)$ and $n \in \mathbb{N}$ to satisfy the hierarchy given by

$$n^{-1} \ll \lambda \ll \zeta \ll \varepsilon. \tag{5.4}$$

Let $d := \lfloor \lambda n \rfloor$. Let k denote the number of all possible rooted trees on at most ζ^{-1} vertices, and denote these rooted trees as $T_1, T_2, \ldots, T_k$. Note that by Cayley's formula, there are at most $(\zeta^{-1})^{\zeta^{-1}-1}$ trees on at most ζ^{-1} vertices. For each of these there are at most ζ^{-1} vertices to select as a root. So in total, we have $k \leqslant (\zeta^{-1})^{\zeta^{-1}}$. We can assume by (5.4) that $\lambda \leqslant \varepsilon \zeta^{1/\zeta}/18$ and so in particular $\lambda k \leqslant \varepsilon/18$. Let T be an $(n+1)$-vertex tree and suppose there exists a vertex $v \in V(T)$ satisfying the assumptions of the statement. Then in particular every component of $T - v$ is isomorphic to T_i for some $i \in [k]$ where the root of T_i is the sole neighbour of v in T (there cannot be more than one neighbour of v in this component as else this would create a cycle). For each $i \in [k]$, let $m_i \in \mathbb{N} \cup \{0\}$ count the number of components in $T - v$ which are isomorphic to T_i, with the root neighbouring v. We view the forest $T - v$ as a collection of rooted trees by

$$T - v \cong (T_1 \times m_1) + (T_2 \times m_2) + \ldots + (T_k \times m_k),$$

where v is adjacent to the root of each T_i. For each $i \in [k]$, let m_i' be the smallest integer at least as big as m_i which is divisible by d. So $m_i \leqslant m_i' < m_i + d$. Define F' to be the forest given by

$$F' = (T_1 \times m_1') + (T_2 \times m_2') + \ldots + (T_k \times m_k').$$

It follows that $T - v$ is a subgraph of F'. Let $n' := |F'|$. Since $m_i' < m_i + d$ for every $i \in [k]$, then $n' < n + dk \leqslant (1 + \lambda k)n$. By our choice of each m_i' being divisible by d, we can write $F' = \hat{F} \times d$ for the forest $\hat{F}$ given by

$$\hat{F} = \left(T_1 \times \frac{m_1'}{d}\right) + \left(T_2 \times \frac{m_2'}{d}\right) + \ldots + \left(T_k \times \frac{m_k'}{d}\right).$$

So now we have that $T - v$ is a subgraph of $\hat{F} \times d$ and let us consider the tree T' obtained by adding vertices and edges to T so that that $T' - v$ is isomorphic to $\hat{F} \times d$, where v is adjacent to the root in each component of $\hat{F}$. We have that $|\hat{F} \times d| = n' \leqslant (1 + \lambda k)n$. In particular note that

$$e(\hat{F}) < |\hat{F}| \leqslant (1 + \lambda k)\frac{n}{d} \leqslant (1 + \lambda k)\frac{2n}{\lambda n} \leqslant 2(\lambda^{-1} + \zeta^{-1/\zeta}). \tag{5.5}$$

We will require some extra labels in our complete graph in order to find a rainbow embedding of T'. Let $\delta := \lambda^2$ and choose $\tilde{n}$ to be the smallest even integer that is at least $\frac{(1+\delta)n'}{1-\delta e(\hat{F})}$. We will construct a rainbow copy of T' in $K_{[\tilde{n}] \cup \{0\}}$ by mapping v to 0, and the remaining vertices in $T' \setminus \{v\}$ will be assigned a unique label in $[\tilde{n}]$ so that all edges of T' have a distinct colour in $K_{[\tilde{n}] \cup \{0\}}$. We will also enforce the additional restriction that colour 1 cannot be used. Since $T \subseteq T'$, this defines a rainbow embedding of T in $K_{[\tilde{n}] \cup \{0\}}$ where v receives label 0, by considering the same embedding restricted to $V(T)$. Using (5.5) and recalling that $n' \leqslant (1 + \lambda k)n$, it follows from (5.4) that

$$\tilde{n} \leqslant \left\lceil \frac{(1+\delta)n'}{1 - \delta e(\hat{F})} \right\rceil + 1 \leqslant \frac{(1+2\delta)(1+\lambda k)n}{1 - 2\delta(\lambda^{-1} + \zeta^{-1/\zeta})} \leqslant (1+\varepsilon)n,$$

So, finding a rainbow copy of T in $K_{[\tilde{n}]\cup\{0\}}$ satisfying the required properties will give us our desired embedding of T in $K_{[(1+\varepsilon)n]\cup\{0\}}$.

Let $p := \frac{1}{|\hat{F}|}$ and consider a p-random partition $[\tilde{n}] = S_1 \cup S_2 \cup \ldots \cup S_{|\hat{F}|}$, meaning that each element of $[\tilde{n}]$ is selected independently and uniformly at random with probability p to belong to a single part. Then for each $i \in |\hat{F}|$, $|S_i| \sim \mathrm{Bin}(\tilde{n}, p)$ and $\mathbb{E}|S_i| = p\tilde{n} = \frac{(1+\delta)d}{1-\delta e(\hat{F})} \pm \frac{2d}{n'}$. So applying Theorem 2.1 gives

$$\mathbb{P}\left[\Big||S_i| - \mathbb{E}|S_i|\Big| > \frac{\delta d}{1-\delta e(\hat{F})}\right] \leqslant \mathbb{P}\left[\Big||S_i| - \mathbb{E}|S_i|\Big| > \frac{4\delta}{(1+\delta)}\mathbb{E}|S_i|\right] \leqslant 2e^{-\frac{\delta^2 d}{(1+\delta)(1-\delta e(\hat{F}))}} = 2e^{-\Omega(n)} \tag{5.6}$$

Let $\mu := \frac{\delta\zeta(1-2\delta)}{4(1+2\delta)}$ and note that $n^{-1} \ll \mu$. We have $p = d/n' \geqslant \lambda/2$, and so $n^{-1} \ll p$. Consider any pair $i, j \in [|\hat{F}|]$ such that $i < j$. Apply Lemma 5.4 with $S_i, S_j, \tilde{n}, p$ and μ playing the roles of S_1, S_2, n, p and μ respectively, to deduce that with probability $1 - o(1)$, there is a rainbow matching with edges between S_i and S_j having colours in $S_j \setminus \{1\}$, covering all but at most $2\mu \max\{|S_i|, |S_j|\}$ vertices in both S_i and S_j. Taking a union bound and using (5.6), with positive probability such a matching exists for each pair i, j, and $|S_i| = \frac{(1+\delta)d}{1-\delta e(\hat{F})} \pm \frac{\delta d}{1-\delta e(\hat{F})}$ is satisfied for all $i \in [|\hat{F}|]$. Therefore taking n to be sufficiently large, we can assume that there is a partition $[\tilde{n}] = S_1 \cup \cdots \cup S_{|\hat{F}|}$ satisfying the properties that

$$\frac{d}{1-\delta e(\hat{F})} \leqslant |S_i| \leqslant \frac{(1+2\delta)d}{1-\delta e(\hat{F})} \tag{5.7}$$

for all $i \in [|\hat{F}|]$, and that there is a collection of matchings $\mathcal{M}_0 := \{M_{ij} : i, j \in [|\hat{F}|], i < j\}$ such that each M_{ij} is S_j-rainbow, covers all but at most $2\mu \max\{|S_i|, |S_j|\}$ vertices in both S_i and S_j, with all edges between S_i and S_j, and that no edge in any matching in this collection has colour 1. Fix such a partition for the remainder of the proof.

We will combine these matchings in order to embed the graph $\hat{F} \times d$. Enumerate the vertices of $\hat{F}$ as $u_1, \ldots, u_{|\hat{F}|}$ using the rooted ordering of each component, that is, all roots are given the lowest index, and vertices are arbitrarily labelled in ascending order of distance from a root. Note that any vertex in this ordering has at at most one neighbour amongst the lower indexed vertices, using the property that rooted trees are 1-degenerate. The element 1 belongs to exactly one vertex set S_i in the partition. If this i is such that u_i is a root in a component of $\hat{F}$, then we can assume that it is possible to choose a j such that u_j is not a root, and switch the parts S_i and S_j. Indeed, if there is no such j, then T is an $n+1$ vertex star with centre v and we can give each of the n leaves a distinct label in $\{2, 3, \ldots, n+1\} \subseteq \{2, 3, \ldots, (1+\varepsilon)n\}$ to satisfy the lemma. So now let us assume 1 is in some part S_i for which u_i is not a root. Let $\mathcal{M} := \{M_{ij} : u_i u_j \in e(\hat{F})\}$. We use the following claim.

Claim 5.7. *$K_{[\tilde{n}]}$ contains a rainbow copy of $\hat{F} \times d$ such that if $\psi : V(\hat{F} \times d) \to [\tilde{n}]$ corresponds to this embedding, then the following properties are satisfied:*

- (i) *for every $j \in [|\hat{F}|]$, $\psi(u_j) \in S_j$,*
- (ii) *for every $\ell \in [|\hat{F}|]$ such that u_ℓ is a root in a component of $\hat{F}$, $\psi(u_\ell) \neq 1$ and no edge in the embedding has colour in S_ℓ,*
- (iii) *no edge has colour 1.*

Proof of claim. Recall that a rainbow copy of $\hat{F} \times d$ is exactly the union of $e(\hat{F})$ colour-disjoint rainbow matchings, where each matching has size d and corresponds to a unique edge in $\hat{F}$. So, we use the partition $S_1 \cup \ldots \cup S_{\hat{F}}$ to find a collection of rainbow matchings, each of which contains only edges of colour in S_j for some $j \in [|\hat{F}|]$, and such that no two matchings use the same colour set.

Note that by choice of our ordering, for a fixed index j, there is at most one $i < j$ such that $u_i u_j \in E(\hat{F})$. In particular at most one matching in $\mathcal{M}$ is S_j-rainbow. Since this holds for all j, then the matchings in $\mathcal{M}$ are pairwise colour-disjoint in $K_{[\tilde{n}]}$, and clearly also pairwise edge-disjoint.

For each $t \in [|\hat{F}|]$, let us define the subcollection of matchings $\mathcal{M}_t \coloneqq \{M_{ij} \in \mathcal{M} : i = t \text{ or } j = t\}$. Consider the set of 'bad' vertices in S_t to be those in the set

$$B_t \coloneqq S_t \setminus \bigcap_{M \in \mathcal{M}_t} V(M).$$

By (5.7), it follows that $|S_i| \geqslant |S_j| - \frac{2\delta d}{1-\delta e(\hat{F})}$ and vice versa for any pair i, j. Without loss of generality assume that $|S_j| \geqslant |S_i|$ (else we can easily apply the same argument the other way around), and we consider the lower bound in (5.7) to see that

$$2\mu|S_j| = \frac{\delta\zeta(1-2\delta)}{2(1+2\delta)}|S_j| = \frac{\delta\zeta}{2(1+2\delta)}\left(|S_j| - 2\delta|S_j|\right) \leqslant \frac{\delta\zeta}{2(1+2\delta)}\left(|S_j| - \frac{2\delta d}{1-\delta e(\hat{F})}\right) \leqslant \frac{\delta\zeta}{2(1+2\delta)}|S_i|.$$

So we can conclude that $2\mu \max\{|S_i|, |S_j|\} \leqslant \frac{\delta\zeta}{2(1+2\delta)} \min\{|S_i|, |S_j|\}$, and therefore every matching $M_{ij} \in \mathcal{M}$ covers all but at most $\frac{\delta\zeta}{2(1+2\delta)} \min\{|S_i|, |S_j|\}$ vertices in S_i and all but at most $\frac{\delta\zeta}{2(1+2\delta)} \min\{|S_i|, |S_j|\}$ vertices in S_j.

In particular, for every $i \in [|\hat{F}|]$ we have $|B_i| \leqslant \sum_{M \in \mathcal{M}_i} |S_i \setminus V(M)| \leqslant |\{e \in E(\hat{F}) : u_i \in e\}| \cdot \frac{\delta\zeta}{2(1+2\delta)}|S_i|$. Again using (5.7), we deduce that

$$\sum_{i \in [|\hat{F}|]} |B_i| \leqslant \sum_{i \in [|\hat{F}|]} \sum_{\substack{e \in E(\hat{F}): \\ u_i \in e}} \frac{\delta\zeta}{2(1+2\delta)}|S_i| \leqslant \frac{\delta\zeta}{1+2\delta} e(\hat{F}) \max_j |S_j| \leqslant \delta\zeta e(\hat{F}) \min_j |S_j|. \tag{5.8}$$

Let $H \subseteq K_{[\tilde{n}]}$ be the graph obtained by taking the union of the matchings $\bigcup_{M \in \mathcal{M}} M$ and deleting all connected components which contain a bad vertex. The components of H combine to form copies of $\hat{F}$, of which each copy will contain exactly one vertex in each part. Since every component of $\hat{F}$ has at most ζ^{-1} vertices and using (5.8), then in total we delete at most $\zeta^{-1} \sum_i |B_i| \leqslant \delta e(\hat{F})|S_t|$ vertices from $\bigcup_{M \in \mathcal{M}} M$ within a given part S_t, in order to obtain H. So, using the lower bound from (5.7), we can find at least $(1 - \delta e(\hat{F}))|S_t| \geqslant d$ copies of $\hat{F}$ in total. Choosing exactly d of them we find an embedding of $\hat{F} \times d$ in $K_{[\tilde{n}]}$ which we have shown is rainbow.

Clearly we have avoided colour 1 since no edge in the matchings have colour 1, so property (iii) also holds. The way in which we have constructed this forest means that each matching M_{ij} corresponds to an edge $u_i u_j \in E(\hat{F})$. So, all of the endpoints belong in the corresponding vertex classes of the partition, i.e. for every $j \in [|\hat{F}|]$, each copy of u_j is embedded into the part S_j. Since for any root u_ℓ in some component of $\hat{F}$, u_ℓ has no neighbour in $\hat{F}$ with a lower index in the ordering $u_1, \ldots, u_{|\hat{F}|}$, and so no matching of the form $M_{k\ell}$ with $k < \ell$ has been used to construct the copy of $\hat{F} \times d$. This means there is no S_ℓ-coloured matching. Since the element 1 was specially selected not to belong to S_ℓ if u_ℓ is a root, then no copy of u_ℓ receives label 1 in this process. Thus all conditions of Claim 5.7 are satisfied. ■

So, now let us use the construction from the claim to find our desired copy of T in $K_{[\tilde{n}] \cup \{0\}}$. First we find a copy of T'. Let us extend the mapping ψ of $V(\hat{F} \times d)$ given by Claim 5.7, by adding the assignment $\psi(v) \coloneqq 0$. Then $\psi : V(T') \to [\tilde{n}] \cup \{0\}$ is an injection providing us with an embedding of T' into $K_{[\tilde{n}] \cup \{0\}}$. We aim to prove this is rainbow. Since we already know ψ restricted to $V(\hat{F} \times d)$ is rainbow, it remains to check that by mapping v to the label 0, we create no colour repetitions.

By choice of $\hat{F}$, we know that only the roots within each component of $\hat{F}$ are neighbours of v, and thus by embedding v to 0 in $K_{[\tilde{n}] \cup \{0\}}$, we only add edges between v and copies of roots in $\hat{F}$. These edges will have colour $|\psi(u_\ell) - \psi(v)| = |\psi(u_\ell) - 0| = \psi(u_\ell)$ for some copy of a root $u_\ell \in V(\hat{F})$, and since we know $\psi(u_\ell) \in S_\ell$ from property (i) and $\psi(u_\ell) \neq 1$ from property (ii), then all colours added in this embedding belong in S_ℓ for some ℓ such that u_ℓ is a root, and none of the edges have colour 1. Also by property (ii) of the embedding, these edges will be colour-disjoint from those already embedded within $\hat{F} \times d$. Our resulting copy of T' is therefore rainbow. Finally we restrict this embedding by only considering the vertices of T, which is possible since T is a subtree of T'. This gives us our desired rainbow copy of T in $K_{[\tilde{n}]\cup\{0\}} \subseteq K_{[(1+\varepsilon/2)n]\cup\{0\}}$, such that v is assigned to label 0, and we have avoided using colour 1 on any edge. □

5.3 Extending the rainbow embedding

We hope to use the statement of Lemma 5.6 in order to consider more families of trees. We do this by proving that for certain trees, it is possible to contract vertex sets so that the resulting tree contains a vertex of high degree which separates it into components of small order. Applying Lemma 5.6, the labelling provided will then be extended by converting each edge of the contracted tree into a matching in the original. First, we need the following proposition which will be useful in finding rainbow perfect matchings between intervals.

Proposition 5.8. *Let $i, j, \ell, n \in \mathbb{N}$ be such that ℓ is odd, $i < j$ and $(j+1)\ell \leqslant n$. Consider two disjoint intervals of integers $I_i = \{i\ell+1, \dots, (i+1)\ell\}$ and $I_j = \{j\ell+1, \dots, (j+1)\ell\}$, and an interval of colours, $C_{ij} = \{(j-i)\ell - \lfloor \ell/2 \rfloor, \dots, (j-i)\ell + \lfloor \ell/2 \rfloor\}$. Then there exists a rainbow perfect matching in $K_{[n]}$ from I_i to I_j using colours in C_{ij}.*

Proof. Note that $\ell = \lceil \ell/2 \rceil + \lfloor \ell/2 \rfloor$. We match up the first $\lceil \ell/2 \rceil$ elements of I_i with the first $\lceil \ell/2 \rceil$ elements of I_j, by smallest to largest, second smallest to second largest, and so on, to get a matching M_1 explictily defined as

$$M_1 = \{(i\ell + y, j\ell + \lceil \ell/2 \rceil + 1 - y) : y \in \{1, 2, \dots, \lceil \ell/2 \rceil\}\}.$$

Similarly we match up the last $\lfloor \ell/2 \rfloor$ elements of I_i with the last $\lfloor \ell/2 \rfloor$ elements of I_j in the same way to get a matching M_2 given by

$$M_2 = \{(i\ell + \lceil \ell/2 \rceil + z, (j+1)\ell + 1 - z) : z \in \{1, 2, \dots, \lfloor \ell/2 \rfloor\}\}.$$

See Fig. 3 for an example. Note that $V(M_1) \cup V(M_2) = I_i \cup I_j$ and $V(M_1) \cap V(M_2) = \emptyset$, thus $M := M_1 \cup M_2$ is a perfect matching. It is easy to see that both M_1 and M_2 are individually rainbow, since the colour of the edges strictly decrease as we match up the pairs from smallest to largest and so on. It remains to verify that their union maintains the rainbow property, and all edges have a colour in C_{ij}.

Consider any edge in M_1. It will be in the form $(i\ell + y, j\ell + \lceil \ell/2 \rceil + 1 - y)$ for some $y \in \{1, 2, \dots, \lceil \ell/2 \rceil\}$. Calculating the absolute difference, the colour of this edge is

$$j\ell + \lceil \ell/2 \rceil + 1 - y - i\ell - y = (j-i)\ell + \lceil \ell/2 \rceil - 2y + 1.$$

Therefore $C(M_1) = \{(j-i)\ell + \lceil \ell/2 \rceil - 2y + 1 : y \in \{1, 2, \dots, \lceil \ell/2 \rceil\}\}$. By the same reasoning, we similarly deduce that $C(M_2) = \{(j-i)\ell + \lceil \ell/2 \rceil - 2z : z \in \{1, 2, \dots, \lfloor \ell/2 \rfloor\}\}$, using the identity $\lceil \ell/2 \rceil = \ell - \lceil \ell/2 \rceil + 1$ since ℓ is odd.

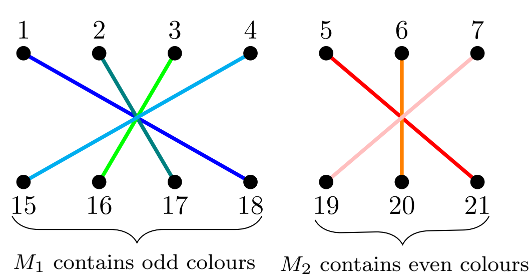

Figure 3: Rainbow matching between intervals I_0 and I_2 for $\ell = 7$, with colour set $\{11, 12, \dots, 17\}$.

Colours in each of these sets have different parities, and thus they are disjoint. So, M is a rainbow perfect matching such that each of its edges has a unique colour in the set

$$C(M_1) \cup C(M_2) = \{(j-i)\ell + \lceil \ell/2 \rceil - w : w \in \{1, 2, \ldots, \ell\}\},$$

and this is equal to C_{ij} as desired. □

Lemma 5.9. *Let $m^{-1} \ll \lambda \ll \zeta \ll \varepsilon \ll 1$ be such that $\zeta m \in \mathbb{N}$, and let T be an m-vertex forest. Suppose there exists a set of vertices $S \subseteq V(T)$ satisfying $|S| \leqslant \lambda m$ and $T \setminus S \cong F \times \zeta m$ for some forest F with rooted components, such that within each component the root has at most one neighbour in S, and all other vertices in the component have no neighbour in S. Then $K_{[(1+\varepsilon)m]}$ contains a rainbow copy of T.*

Proof. Choose $\gamma \in (0,1)$ such that $1/\gamma \in \mathbb{N}$ and $\zeta/\gamma \in \mathbb{N}$, and so that the following hierarchy is satisfied

$$m^{-1} \ll \lambda \ll \gamma \ll \zeta \ll \varepsilon \ll 1. \tag{5.9}$$

Let T be as in the statement of the lemma. Without loss of generality we may assume T is a tree and that for every component of $T \setminus S$, the root has exactly one neighbour in S, by adding edges between roots of components and vertices in S, and edges within S if necessary. Let $d \in \{\lceil \gamma m \rceil, \lceil \gamma m \rceil + 1\}$ be such that $d + 3|S|$ is odd. We will first construct an auxiliary tree T_{aux} obtained from T, such that T_{aux} contains a vertex v satisfying the assumptions of Lemma 5.6, allowing us to find a rainbow embedding of T_{aux} in the difference coloured complete graph on a little more than $|T_{\text{aux}}|$ vertices.

Let us denote $V(F) \coloneqq \{u_1, \ldots, u_{|F|}\}$ using the rooted ordering of F, that is, all roots in the components of F have lowest index, and the remainder are enumerated in ascending order in terms of increasing distance from a root. Since in $T \setminus S$ we have exactly ζm copies of each vertex in F, then we can denote the set of vertices in $T \setminus S$ by

$$V(F \times \zeta m) = \{u_\ell^t : \ell \in [|F|], t \in [\zeta m]\}.$$

We partition $V(F \times \zeta m)$ into $\zeta\gamma^{-1}|F|$ disjoint vertex classes, each of which will correspond to a vertex in T_{aux}. For this purpose, for each $\ell \in [|F|]$ and $r \in [\zeta\gamma^{-1}]$, define

$$U_\ell^r \coloneqq \left\{u_\ell^t : t \in ((r-1)d, rd]\right\}.$$

In particular, for all $\ell \in [|F|]$ we have $|U_\ell^r| = d$ when $r \in [\zeta\gamma^{-1} - 1]$ and $|U_\ell^{\zeta\gamma^{-1}}| \leqslant d$. Let us construct and auxiliary graph F_{aux} on these U_ℓ^r sets that is isomorphic to $F \times \zeta\gamma^{-1}$. We take F_{aux} to be the union of $\zeta\gamma^{-1}$ vertex-disjoint copies of F where for each $r \in [\zeta\gamma^{-1}]$, the vertices in the rth copy of F are given by $\{U_\ell^r : \ell \in [|F|]\}$. More explicitly, we have

$$V(F_{\text{aux}}) = \{U_\ell^r : \ell \in [|F|], r \in [\zeta\gamma^{-1}]\} \quad \text{and} \quad E(F_{\text{aux}}) = \{U_i^r U_j^r : r \in [\zeta\gamma^{-1}], u_i u_j \in E(F), i, j \in [|F|]\}.$$

Let T_{aux} be the tree obtained by adding a new vertex v to F_{aux}, and adding an edge between v and the root vertex in every component of F_{aux}. We can write this explicitly as

$$V(T_{\text{aux}}) = V(F_{\text{aux}}) \cup \{v\} \quad \text{and} \quad E(T_{\text{aux}}) = E(F_{\text{aux}}) \cup \{vU_i^r : r \in [\zeta\gamma^{-1}],\ u_i \text{ is a root in a component of F}\}.$$

Note that since in T each component of $F \times \zeta m$ had at most one edge going into S by assumption, then we do not form any cycles nor multi-edges in T_{aux}, so T_{aux} is indeed a tree. Also, $|F| \leqslant \frac{m}{\zeta m} = \zeta^{-1}$ so in particular every component of F has size at most ζ^{-1}. We know that $|F| \leqslant \zeta^{-1}$ and so $|T_{\text{aux}}| \leqslant \gamma^{-1} + 1$. Since all components of $T_{\text{aux}} \setminus \{v\}$ have size at most ζ^{-1}, then we can apply Lemma 5.6 to T_{aux} with $\varepsilon/2$ playing the role of ε to obtain a rainbow embedding $\phi : V(T_{\text{aux}}) \to \{0, 1, 2, \ldots, (1+\varepsilon/2)\gamma^{-1}\}$ in $K_{[(1+\varepsilon/2)\gamma^{-1}] \cup \{0\}}$ where $\phi(v) = 0$ and we avoid colour 1. We will use this rainbow embedding ϕ of T_{aux} to construct a rainbow embedding ψ of T in $K_{[(1+\varepsilon)m]}$. For convenience, let $\eta \coloneqq (1+\varepsilon/2)\gamma^{-1}$.

We now look only at the labels of $V(F \times \zeta\gamma^{-1})$ given by ϕ, and note that these all belong to $[\eta]$. Let $\tilde{m} \coloneqq (\eta+1)(d+3|S|)$ and consider a partition of $[\tilde{m}]$ into $\eta+1$ disjoint intervals, each of length $d+3|S|$. We denote this by $[\tilde{m}] = I_0 \cup I_1 \cup \ldots \cup I_\eta$ such that for each $i \in \{0,1,\ldots,\eta\}$,

$$I_i \coloneqq \{i(d+3|S|)+1,\ldots,(i+1)(d+3|S|)\}\,.$$

Consider the function $g : V(T) \to \{I_0, I_1, \ldots, I_\eta\}$ defined for each $w \in V(T)$ as follows: If $w \in S$ then let $g(w) = I_0$, otherwise $w \in V(F \times \zeta n)$ and so there exists a unique pair $(\ell, t) \in [|F|] \times [\zeta m]$ such that $w = u_\ell^t$. In this case let $g(w) = I_{\phi(U_\ell^r)}$ for the unique r with $t \in ((r-1)\gamma m, r\gamma m]$ (or equivalently, for the unique r with $w \in U_\ell^r$). We will use g to show that there exists a rainbow embedding ψ of T where every vertex $w \in V(T)$ is assigned a label from the interval $g(w)$. We embed $V(T)$ into $K_{[\tilde{m}]}$ by embedding all vertices in these vertex set in the following order:

$$S, U_1^1, U_1^2, \ldots, U_1^{\zeta\gamma^{-1}}, U_2^1, \ldots, U_2^{\zeta\gamma^{-1}}, \ldots, U_{|F|}^1, \ldots, U_{|F|}^{\zeta\gamma^{-1}}. \tag{5.10}$$

We will ensure the colours used on edges contained in $T[S]$ belong in the interval $C_S \coloneqq [3|S|]$, and that all remaining edges outside of $T[S]$ receive colours chosen from some other disjoint intervals. For this purpose, for every pair $i, j \in [\eta] \cup \{0\}$ with $j > i$, define the colour set

$$C_{ij} \coloneqq \left\{(j-i)(d+3|S|) - \left\lfloor \frac{d+3|S|}{2} \right\rfloor, \ldots, (j-i)(d+3|S|) + \left\lfloor \frac{d+3|S|}{2} \right\rfloor \right\}.$$

Embedding vertices in S. We enumerate the vertices of S as $w_1, \ldots, w_{|S|}$ such that each of these has at most one neighbour amongst the lower indexed vertices, possible since $T[S] \subseteq T$ is 1-degenerate. Each vertex $w_i \in S$ satisfies $g(w_i) = I_0$, and we want to embed S into a subinterval $I_0' \subset I_0$, defined by

$$I_0' = \left\{ \left\lceil \frac{d-3|S|}{2} \right\rceil, \ldots, \left\lceil \frac{d+3|S|}{2} \right\rceil \right\},$$

and such that the resulting embedding ψ is rainbow. We do this greedily. Suppose we have labelled the first j vertices in this ordering for some $0 \leqslant j < |S|$ and we want to assign a label to w_{j+1}. There are $j < |S|$ labels from I_0' unavailable due to previously embedded vertices in S. Since at each step, there is at most one new edge having both of its endvertices already labelled, then we have also used at most j colours on these edges. Each colour used causes at most two labels from I_0' to become unavailable by considering the absolute difference, in total restricting $2j < 2|S|$ colours. Since $|I_0'| \geqslant 3|S|$ then there are suitable labels remaining, and we select $\psi(w_j)$ to be the minimal label from this set of choices. Note that I_0' is an interval of length $3|S|$, and so all colours used in this process for edges in $T[S]$ belong in $C_S = [3|S|]$. Furthermore under ψ, the forest $T[S]$ is rainbow.

Embedding root vertices. We proceed by embedding the vertices which are roots in some component of $F \times \zeta m$, noting that these sets came first (after S) in the ordering (5.10). Consider a vertex set U_ℓ^r where all vertices in this set are copies of u_ℓ for some $\ell \in [|F|]$ such that u_ℓ is a root. Recall that we are assuming every copy of a root vertex has exactly one neighbour in S in T, and that u_ℓ is a neighbour of v in T_{aux}. Choose $j \coloneqq \phi(U_\ell^r)$. We will embed U_ℓ^r into a subinterval $I_j' \subset I_j$, adding only edges with colour in C_{0j}. We define

$$I_j' \coloneqq \{j(d+3|S|)+1, \ldots, (j+1)(d+3|S|) - 3|S|\}\,.$$

Let $d' = |U_\ell^r|$ and recall that $d' \leqslant d$. Just for this step, consider a new enumeration on the vertices in U_ℓ^r by $v_1, v_2, \ldots, v_{d'}$, chosen such that, for the unique neighbour $s_i \in N_T(v_i) \cap S$, we have

$$\psi(s_1) \geqslant \psi(s_2) \geqslant \ldots \geqslant \psi(s_{d'}).$$

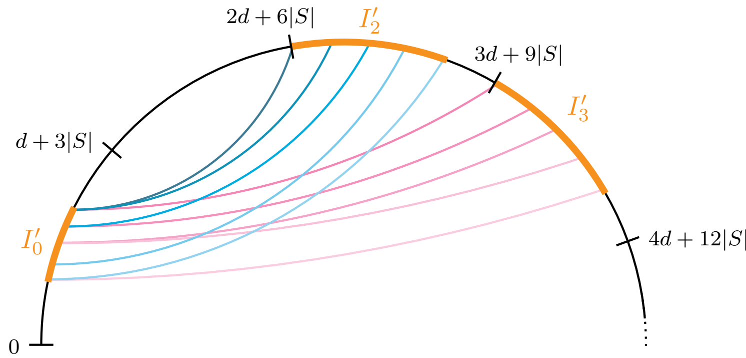

Figure 4: Intervals $I_0, \ldots, I_3$ depicted around the semi-circle, with subintervals I'_0, I'_2 and I'_3 highlighted in orange. Vertices in S are embedded into I'_0. In this example, v is adjacent to the vertices labelled by 2 and by 3 in T_{aux}, and we embed the root vertices corresponding to these vertex sets into I'_2 and I'_3 respectively. Blue edges have a colour in C_{02}, and lighter blue edges have larger colours than darker blue edges. The same holds for pink edges having colours in C_{03}.

Extend the embedding ψ so that the vertices of U^r_ℓ are embedded into I'_j using the ordering $v_1, \ldots, v_{d'}$, so that v_1 receives the smallest label in I'_j, v_2 receives the second smallest label in I'_j, and so on. This is possible because $d' \leqslant d = |I'_j|$. Let $i, k \in [\gamma m]$ be such that $k > i$. Then $\psi(v_i) - \psi(s_i) \leqslant \psi(v_i) - \psi(s_k) < \psi(v_k) - \psi(s_k)$. So, the colour of the edge $v_i s_i$ is strictly smaller than the colour of the edge $v_k s_k$, under the embedding ψ. Since this holds for all such i, k, then any pair of edges added in the process of embedding U^r_ℓ have different colours. See Fig. 4 for an example.

We also wish to verify that these new edges have colours in C_{0j}. It suffices to check the colour of an edge uv where $u \in I'_0$ and $v \in I'_j$. We have $v - u \leqslant (j+1)(d+3|S|) - 3|S| - \left\lceil \frac{d-3|S|}{2} \right\rceil = j(d+3|S|) + \left\lfloor \frac{d+3|S|}{2} \right\rfloor$, and $v - u \geqslant j(d+3|S|) + 1 - \left\lceil \frac{d+3|S|}{2} \right\rceil = j(d+3|S|) - \left\lfloor \frac{d+3|S|}{2} \right\rfloor$. So $v - u \in C_{0j}$, as desired. This concludes our embedding process for vertices which have neighbours in S.

Embedding non-root vertices. In order to embed all remaining vertices, we will require some rainbow matchings. For all $i, j \in [\eta]$ with $j > i$, note that $\tilde{m} \geqslant (j+1)(d+3|S|)$ and we chose d such that $d+3|S|$ is odd, so applying Proposition 5.8 with $\tilde{m}$ and $d+3|S|$ playing the roles of n and ℓ respectively, there exists a rainbow perfect matching in $K_{[\tilde{m}]}[I_i, I_j, C_{ij}]$. Denote this matching by M_{ij}. Since these exist for each pair $i = \phi(U^r_k), j = \phi(U^r_\ell)$ with $U^r_k U^r_\ell \in E(T_{\text{aux}})$, we can consider a collection of such matchings to uniquely define the labels in our embedding of T.

Now consider a vertex set U^r_ℓ which does not contain vertices which have neighbours in S, and assume we have defined the embedding ψ for all previous vertex sets from (5.10). By our chosen ordering, U^r_ℓ has exactly one neighbour in T_{aux} amongst earlier vertex sets from the ordering (5.10), say U^r_k. By construction we know that $\psi(U^r_k) \subseteq I_{\phi(U^r_k)}$, and we will embed U^r_ℓ into $I_{\phi(U^r_\ell)}$ using the corresponding rainbow perfect matching as follows. Let $i = \phi(U^r_k)$ and $j = \phi(U^r_\ell)$. Since every vertex $z \in U^r_\ell$ has exactly one neighbour $y \in U^r_k$, and there is exactly one edge e in M_{ij} that contains $\psi(y)$, then we can use this edge to uniquely determine $\psi(z)$ by choosing $\psi(z)$ to be the vertex in $e \setminus \{\psi(y)\}$. Note by definition of M_{ij} that this ensures that $\psi(z) \in I_j$, that distinct vertices in U^r_ℓ are assigned distinct labels, and that colours of all edges added between U^r_ℓ and U^r_k are distinct and belong to C_{ij}. This completes our embedding ψ of T.

It remains to check that what we have embedded altogether is rainbow and that we have not used too many labels.

Claim 5.10. *Let $i, j, i', j' \in [\eta] \cup \{0\}$ be such that $|j-i| \neq |j'-i'|$. Then $C_{ij} \cap C_{i'j'} = \emptyset$.*

Proof of claim. Without loss of generality assume $j > i$ and $j' > i'$ and $j - i < j' - i'$. In particular since these are integer valued, we know $j - i \leqslant j' - i' - 1$. For any $c \in C_{ij}$, it follows that

$$c \leqslant (j-i)(d+3|S|) + \left\lfloor \frac{d+3|S|}{2} \right\rfloor \leqslant (j'-i'-1)(d+3|S|) + \left\lfloor \frac{d+3|S|}{2} \right\rfloor = (j'-i')(d+3|S|) - \left\lceil \frac{d+3|S|}{2} \right\rceil$$

so $c \notin C_{i'j'}$. Thus there is no c such that $c \in C_{ij} \cap C_{i'j'}$, as desired. ■

Claim 5.11. *If $i, j \in [\eta] \cup \{0\}$ are distinct and $|j - i| \neq 1$, then $C_{ij} \cap C_S = \emptyset$.*

Proof of claim. Without loss of generality suppose $j > i$, and we can assume $j - i \geqslant 2$ since i and j take integer values, and $j - i \neq 1$. For any $c \in C_{ij}$, we have $c \geqslant 2(d + 3|S|) - \left\lfloor \frac{d+3|S|}{2} \right\rfloor > 3|S|$. Since $C_S = [3|S| - 1]$, then $c \notin C_S$ and the statement holds. ■

Our labelling of T_{aux} is rainbow so there are no $i, j, i', j' \in [\eta] \cup \{0\}$ and edges $U_k^r U_\ell^r, U_{k'}^s U_{\ell'}^s \in E(T_{\text{aux}})$ for which $i = \phi(U_k^r)$, $j = \phi(U_\ell^r)$, $i' = \phi(U_{k'}^r)$, $j' = \phi(U_{\ell'}^r)$ and $|j - i| = |\ell - k|$. Recall also that this embedding of T_{aux} avoided colour 1, and so similarly there are no such i and j with $|j - i| = 1$. Whilst embedding S, we used colour set C_S. Whilst embedding root vertices, we used colour sets of the form C_{0j} where $j = \phi(U_\ell^r)$ and $vU_\ell^r \in E(T_{\text{aux}})$ for some $\ell \in [|F|]$ and $r \in [\zeta\gamma^{-1}]$. Whilst embedding all remaining vertices, we used colour sets of the form C_{ij} where $i = \phi(U_k^r)$, $j = \phi(U_\ell^r)$ and $U_k^r U_\ell^r \in E(T_{\text{aux}})$ for some $k, \ell \in [|F|]$ and $r \in [\zeta\gamma^{-1}]$. Thus together Claims 5.10 and 5.11 tell us that the family $\mathcal{C}$ of all colour sets considered are pairwise disjoint. By construction of our labelling of T, we use each colour set $C \in \mathcal{C}$ at most once, and under ψ, the edges with a colour in C are rainbow. Altogether, we deduce that the resulting embedding of T is in fact rainbow, as desired.

Finally let us count the total the number of labels used. We know that $d \leqslant \gamma m + 2$ and $|S| \leqslant \lambda m$. So, using the hierarchy (5.9) we have

$$\begin{aligned}
\tilde{m} = (\eta + 1)(d + 3|S|) &\leqslant ((1 + \varepsilon/2)\gamma^{-1} + 1)(3\lambda m + \gamma m + 2) \\
&\leqslant \left(1 + \varepsilon/2 + \gamma + ((1 + \varepsilon/2)\gamma^{-1} + 1)(3\lambda + 2/m)\right) m \\
&\leqslant (1 + \varepsilon)m,
\end{aligned}$$

as desired. □

6 Proof of Lemma 1.7.

We can now finally collate what we know to prove Lemma 1.7, following the proof strategy given in Section 3. As discussed there, we need to be careful with how we choose our parameters in order to ensure that $|S_{\text{high}}|$ is small and all waste vertices in W have low degree. So let us add in a brief sketch of how we may select these.

At the beginning, we consider many possible choices for a parameter Δ, chosen so that $\Delta_0 \ll \Delta_1 \ll \ldots \ll \Delta_{4\varepsilon^{-1}} \ll n$ and for a given n-vertex tree T, there exist two consecutive Δ_i, Δ_{i+1} satisfying the following property: there are at most $\varepsilon n/2$ edges in T which have an endvertex amongst the set of vertices with degree in the interval $[\Delta_i, \Delta_{i+1})$. Call this 'special' set of vertices $\tilde{V}$, and choose S_{high} to be the set of vertices with degree at least Δ_{i+1}. So, $|S_{\text{high}}| \leqslant 2n/\Delta_{i+1}$. Furthermore, the waste vertices in W either have degree strictly less than Δ_i, or, they belong to $\tilde{V}$. In the former case we can show that the number of edges touching this set is at most $\Delta_i|W| \leqslant \varepsilon n/2$, and in the latter case $\tilde{V}$ was chosen specially so that there are also at most $\varepsilon n/2$ touching this set. Therefore, when we embed W arbitrarily at the end, the number of edges that do not receive a distinct colour is at most εn.

Proof. Let $\varepsilon > 0$ and let $\Delta_0, \Delta_1, \ldots, \Delta_{4\varepsilon^{-1}}$ be a sequence of natural numbers chosen to satisfy

$$\Delta_{4\varepsilon^{-1}}^{-1} \ll \ldots \ll \Delta_1^{-1} \ll \Delta_0^{-1} \ll \varepsilon.$$

Choose N to be sufficiently large with respect to the Δ_i and let $n > N$. Let T be an n-vertex tree. For each $i \in \{0, 1 \ldots, 4\varepsilon^{-1} - 1\}$, let $U_i \coloneqq \{v \in V(T) : d_T(v) \in [\Delta_i, \Delta_{i+1})\}$ and let $E_i \subseteq E(T)$ be the set of edges containing a vertex in U_i. Each edge in $E(T)$ belongs to at most two distinct E_is. Suppose $|E_i| > \frac{\varepsilon n}{2}$ for every $i \in \{0, 1 \ldots, 4\varepsilon^{-1} - 1\}$, then

$$2|E(T)| \geqslant \sum_i |E_i| > 4\varepsilon^{-1}\left(\frac{\varepsilon n}{2}\right) = 2n,$$

a contradiction. So there exists an $i \in \{0, 1 \ldots, 4\varepsilon^{-1} - 1\}$ such that $|E_i| \leqslant \frac{\varepsilon n}{2}$ and for such an i, for convenience let us define $\Delta \coloneqq \Delta_i$, $\tilde{\Delta} \coloneqq \Delta_{i+1}$ and $U \coloneqq U_i$, noting for later that at most $\varepsilon n/2$ edges of T contain a vertex in U, and that every vertex in U has degree less than $\tilde{\Delta}$ in T.

We now additionally choose $\delta, \zeta > 0$ such that ζ is at most the output ζ_0 of Lemma 4.1 when applied with δ, such that $\zeta n \in \mathbb{N}$, and satisfying

$$n^{-1} \ll \tilde{\Delta}^{-1} \ll \zeta \ll \delta \ll \Delta^{-1} \ll \varepsilon. \tag{6.1}$$

It is easily observed that $\delta\Delta < \varepsilon/2$ and $\delta\tilde{\Delta} > 20$.

Let $S_{\text{high}} \coloneqq \{v \in V(T) : d_T(v) \geqslant \tilde{\Delta}\}$ so that $|S_{\text{high}}| \leqslant 2n/\tilde{\Delta} < \delta n/10$. Since we chose ζ accordingly, we can apply Lemma 4.1 to T with S_{high} playing the role of S to obtain a vertex set $W \subseteq V(T) \setminus S_{\text{high}}$ and a forest F with rooted trees as components, such that $|W| \leqslant \delta n$ and $T \setminus (W \cup S_{\text{high}}) \cong F \times \zeta n$, and for every component in $F \times \zeta n$, only the root has a neighbour in S_{high}. Let $\tilde{T} = T \setminus W$, so that $\tilde{T}$ is a forest of the form $S_{\text{high}} \cup (F \times \zeta n)$, and we know that only the roots in $F \times \zeta n$ have a neighbour in S_{high}. Note that $\tilde{n} \coloneqq |\tilde{T}| = |T| - |W| \geqslant (1-\delta)n$. Choose $\tilde{\zeta} \coloneqq \zeta n/\tilde{n}$, so we have $\tilde{\zeta} \in [\zeta, \frac{\zeta}{1-\delta}]$ and $\tilde{\zeta} \ll \varepsilon$, and $\tilde{T} \setminus S_{\text{high}} \cong F \times \zeta n = F \times \tilde{\zeta}\tilde{n}$. Observe that $|S_{\text{high}}| \leqslant \frac{2n}{\tilde{\Delta}} \leqslant \frac{2\tilde{n}}{\tilde{\Delta}(1-\delta)}$. From (6.1) we can assume $\frac{2}{\tilde{\Delta}(1-\delta)} \ll \zeta \leqslant \tilde{\zeta}$. Applying Lemma 5.9 with $\tilde{T}$, $\tilde{n}$, S_{high}, $\frac{2}{\tilde{\Delta}(1-\delta)}$ and $\tilde{\zeta}$ playing the roles of T, m, S, λ and ζ respectively, we find a rainbow copy of $\tilde{T}$ in $K_{[(1+\varepsilon)\tilde{n}]} \subseteq K_{[(1+\varepsilon)n]}$.

All that is left to embed of T is the set W. First, let us embed $W \setminus U$. Since $W \setminus U \subseteq V(T) \setminus (S_{\text{high}} \cup U)$, then every vertex in this set has degree at most Δ. So, the number of edges containing a vertex in $W \setminus U$ is at most $|W \setminus U|\Delta \leqslant |W|\Delta \leqslant \delta\Delta n < \varepsilon n/2$. Therefore, we can arbitrarily embed vertices in W into $K_{[(1+\varepsilon)n]}$, ensuring only unused labels of $[(1+\varepsilon)n]$ are used, causing at most $\varepsilon n/2$ colours to repeat. Finally we must embed $W \cap U$. We can similarly give these vertices an arbitrary label from what remains of $[(1+\varepsilon)n]$. This adds at most $\varepsilon n/2$ edges, each of which may cause a colour to be repeated. Altogether we find a copy of T in $K_{[(1+\varepsilon)n]}$ such that all but at most $\varepsilon n/2 + \varepsilon n/2 = \varepsilon n$ edges have distinct colours. This proves the lemma. □

References

[1] A. Adamaszek, P. Allen, C. Grosu, and J. Hladkỳ. Almost all trees are almost graceful. *Random Structures & Algorithms*, 56(4):948–987, 2020. 2

[2] N. Alon and R. Yuster. On a hypergraph matching problem. *Graphs and Combinatroics*, 21(4):377–384, 2005. 10

[3] F. Van Bussel. Relaxed graceful labellings of trees. *Electronic Journal of Combinatorics*, 9(1):12, 2002. 2

[4] G. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952. 1

[5] P. Erdős, P. Hell, and P. Winkler. Bandwidth versus bandsize. *Annals of Discrete Mathematics*, 41:117–129, 1988. 2

[6] P. Erdős, M. Simonovits, and V.T. Sós. Anti-Ramsey theorems. *Infinite and finite sets*, 2:633–643, 1975. 1

[7] J. Gallian. A dynamic survey of graph labeling. *Electronic Journal of Combinatorics*, 25:9–46, 2022. 2

[8] S. Golomb. How to number a graph. In *Graph Theory and Computing*, pages 23–37. Academic Press, 1972. 1

[9] R. Graham and N. Sloane. On additive bases and harmonious graphs. *Siam Journal on Algebraic and Discrete Methods*, 1, 12 1980. 1

[10] S. Janson, T. Łuczak, and A. Rucinski. *Random Graphs*. Wiley New York, 2000. 5

[11] P. Keevash and K. Staden. Ringel's tree packing conjecture in quasirandom graphs. *Journal of the European Mathematical Society (EMS Publishing)*, 27(5), 2025. 2

[12] W. Mantel. Problem 28. *Wiskundige Opgaven*, 10:60–61, 1907. 1

[13] C. McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141:148–188, 1989. 5

[14] M. Molloy and B. Reed. Near-optimal list colorings. *Random Structures & Algorithms*, 17(3-4):376–402, 2000. 11

[15] R. Montgomery, A. Pokrovskiy, and B. Sudakov. A proof of Ringel's conjecture. *Geometric and Functional Analysis*, 31:1–58, 06 2021. 2

[16] G. Ringel. Theory of graphs and its applications. In *Proceedings of the Symposium Smolenice*, 1963. 2

[17] A. Rosa. On certain valuations of the vertices of a graph. *Theory of Graphs (Internat. Symposium, Rome)*, page 349–355, 1966. 1

[18] A. Rosa and J. Širáň. Bipartite labelings of trees and the gracesize. *Journal of Graph Theory*, 19(2):201–215, 1995. 2