THE TRIPARTITE RAMSEY NUMBER FOR TREES

JULIA BÖTTCHER, JAN HLADKÝ, AND DIANA PIGUET

Abstract. We prove that for all $\varepsilon > 0$ there are $\alpha > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. For any two-colouring of the edges of $K_{n,n,n}$ one colour contains copies of all trees T of order $t \leq (3-\varepsilon)n/2$ and with maximum degree $\Delta(T) \leq n^\alpha$. This confirms a conjecture of Schelp.

1. Introduction and results

The celebrated theorem of Ramsey [11] states that for any finite family of graphs $\mathcal{F}$ the number $R(\mathcal{F})$, defined as the smallest integer m such that in any edge-colouring of K_m with green and red there are either copies of all members of $\mathcal{F}$ in green or in red, exists. In this case we also write $K_m \to \mathcal{F}$ and say that K_m is *Ramsey* for $\mathcal{F}$. Let $\mathcal{T}_t$ denote the class of trees of order t, $\mathcal{T}_t^\Delta$ is its restriction to trees of maximum degree at most Δ. Ajtai, Komlós, Simonovits, and Szemerédi [1] announced a result which implies that $K_{2t-2} \to \mathcal{T}_t$ for large even t and $K_{2t-3} \to \mathcal{T}_t$ for large odd t. This bound is best possible. For the case of odd t this is also a consequence of a theorem by Zhao [14] concerning a conjecture of Loebl (see also [7]).

The graph $K_{R(\mathcal{F})}$ is obviously a Ramsey graph for $\mathcal{F}$ with as few vertices as possible. However, one may still ask, whether there exist graphs with fewer edges which are Ramsey for $\mathcal{F}$. This minimal number of edges is also called *size Ramsey number* and denoted by $R_s(\mathcal{F})$. Trivially $R_s(\mathcal{F}) \leq \binom{R(\mathcal{F})}{2}$, but it turns out that this inequality is often far from tight. The investigation of size Ramsey numbers recently experienced much attention. Trees are considered in [3, 6]. Progress on determining the size Ramsey number for classes of bounded degree graphs was made in [9].

The first and the third author were partially supported by DFG grant TA 309/2-1. The second and the third author were partially supported by DAAD. The second author was partially supported by the grant GAUK 202-10/258009 of the Grant Agency of Charles University.

A question of similar flavour is what happens when we do not confine ourselves to finding Ramsey graphs for $\mathcal{F}$ with few edges but require in addition that they are proper subgraphs of K_m with m very close to $R(\mathcal{F})$. This question has two aspects: a quantitative one (i.e., how many edges can be deleted from K_m so that the remaining graph is still Ramsey) and a structural one (i.e., what is the structure of the edges that may be deleted). Questions of similar nature were explored in [5] when $\mathcal{F}$ consists of an odd cycle and in [4] when $\mathcal{F}$ is a path. Our focus in this paper is on the case when $\mathcal{F}$ is a class of trees.

Schelp [12] posed the following Ramsey-type conjecture about trees in tripartite graphs: For n sufficiently large the tripartite graph $K_{n,n,n}$ is Ramsey for the class $\mathcal{T}_t^{\Delta}$ of trees on $t \leq (3-\varepsilon)n/2$ vertices with maximum degree at most Δ for constant Δ. The conjecture thus asserts that we can delete three cliques of size $m/3$ from a graph K_m with m only slightly larger than $R(\mathcal{T}_t^{\Delta})$ while maintaining the Ramsey property. In addition Schelp asked whether the same remains true when the constant maximum degree bound in the conjecture above is replaced by $\Delta \leq \frac{2}{3}t$ (which is easily seen to be best possible). Our main result is situated in-between these two cases, solving the problem for trees of maximum degree n^{α} for some small α and hence, in particular, answering the first conjecture above.

Theorem 1. *For all $\mu > 0$ there are $\alpha > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$*

$$K_{n,n,n} \rightarrow \mathcal{T}_t^{\Delta},$$

with $\Delta \leq n^{\alpha}$ and $t \leq (3-\mu)n/2$.

We use Szemerédi's regularity lemma [13] to establish this result. Due to the nature of the methods related to this lemma it follows that Theorem 1 remains true when $K_{n,n,n}$ is replaced by a much sparser graph: For any fixed $\mu \in (0,1]$ a random subgraph of $K_{n,n,n}$ with edge probability μ allows for the same conclusion, as long as n is sufficiently large (cf. Section 8).

The proof of Theorem 1 splits into a combinatorial part and a regularity based embedding part. The lemmas we need for the combinatorial part are stated in Section 3 and proved in Section 7. Łuczak [10] first noted that a large connected matching in a cluster graph is a suitable structure for embedding paths. In the present paper, we extend Łuczak's idea and use what we call "odd connected matchings" and "connected fork systems" in the cluster graph.

For the embedding part we formulate an embedding lemma (Lemma 13, see Section 4) that provides rather general conditions for the embedding of trees with growing maximum degree. The proof of this lemma is prepared in Section 5 and presented in Section 6. First, however, we shall introduce all necessary definitions as well as the regularity lemma in the following section.

2. Definitions and Tools

Let $G = (V, E)$ be a graph and $X, X', X'' \subseteq V$ be pairwise disjoint vertex sets. Then we define $E(X) := E \cap \binom{X}{2}$ and $E(X, X') := E \cap (X \times X')$ and write $G[X]$ for the graph with vertex set X and edge set $E(X)$. Similarly, $G[X, X']$ is the bipartite graph with vertex set $X \dot{\cup} X'$ and edge set $E(X, X')$ and $G[X, X', X'']$ is the tripartite graph with vertex set $X \dot{\cup} X' \dot{\cup} X''$ and edge set $E(X, X') \dot{\cup} E(X', X'') \dot{\cup} E(X'', X)$. For convenience we frequently identify graphs G with their edge set $E(G)$ and vice versa. We say that a subgraph G' of G covers a vertex v of G if v is contained in some edge of G'. For a vertex set D and an edge set M we denote by $D \cap M$ the set of vertices from D that appear in some edge of M. We write $\mathrm{N}(v)$ for the neighborhood of a vertex v.

A *matching* M in a graph $G = (V, E)$ is a set of vertex disjoint edges in E and its size is the number of edges in M. For vertices v and vertex sets U covered by M we also write, abusing notation, $v \in M$ and $U \subseteq M$. Sometimes we also consider a matching as a bijection $M \colon V_M \to V_M$ where $V_M \subseteq V$ is the set of vertices covered by M. For $U \subseteq V_M$ we then denote by $M(U)$ the set of vertices $v \in V_M$ such that $uv \in M$ for some $u \in U$.

To make our notation compact we sometimes use subscripts in a non-standard way as illustrated by the following example. Let $A_1, A_2 \subseteq A$ and $B_1, B_2 \subseteq B$ be sets and suppose that $D \in \{A, B\}$ and $i \in [2]$. The symbol D_i then denotes the set A_i if $D = A$ and the set B_i if $D = B$.

2.1. **Regularity.** Let $G = (V, E)$ be a graph and $\varepsilon, d \in [0, 1]$. For disjoint nonempty vertex sets $U, W \subseteq V$ the *density* $d(U, W)$ of the pair (U, W) is the number of edges that run between U and W divided by $|U||W|$. A pair (U, W) with density at least d is (ε, d)*-regular* if $|d(U', W') - d(U, W)| \leq \varepsilon$ for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$. The following lemma states that in dense regular pairs most vertices have many neighbours. This is an immediate consequence of the definition of regular pairs.

Lemma 2. *Let (U, U') be an (ε, d)-regular pair and $X \subseteq U$ with $|X| \geq \varepsilon|U|$. Then less than $\varepsilon|U'|$ vertices in U' have less than $(d - \varepsilon)|X|$ neighbours in X.* □

In the rest of the paper we will say that all other vertices in U are (ε, d)-*typical* with respect to X (or simply typical, when ε and d are clear from the context).

An (ε, d)-*regular partition* of $G = (V, E)$ with *reduced graph* $\mathbb{G} = ([k], E_{\mathbb{G}})$ is a partition $V_0 \dot\cup V_1 \dot\cup \dots \dot\cup V_k$ of V with $|V_0| \leq \varepsilon |V|$, such that (V_i, V_j) is an (ε, d)-regular pair in G whenever $ij \in E_{\mathbb{G}}$. In this case we also say that G has (ε, d)-*reduced graph* $\mathbb{G}$. (Throughout this paper blackboard symbols such as $\mathbb{G}$ or $\mathbb{M}$ denote reduced graphs and their subgraphs.) The partition classes V_i with $i \in [k]$ are also called *clusters* of G and V_0 is the *bin set*. We also call a vertex i of the reduced graph a cluster and identify it with its corresponding set V_i.

Suppose that P is a partition of V. We then say that a partition $V_0 \dot\cup V_1 \dot\cup \dots \dot\cup V_s$ of V *refines* P if for every $i \in [s]$ there exists a member $A \in P$ such that $V_i \subseteq A$. Finally, a partition $V_0 \dot\cup V_1 \dot\cup \dots \dot\cup V_k$ of V is an *equipartition* if $|V_i| = |V_j|$ for all $i, j \in [k]$.

Now we state a version of Szemerédi's celebrated regularity lemma [13]. This lemma takes an n-vertex graph G that is given with some preliminary partition and produces a regular partition of G with $k \leq k_1$ clusters which refines this partition where k_1 does not depend on n.

Lemma 3 (Regularity lemma). *For all $\varepsilon > 0$ and integers k_0 and k_* there is an integer k_1 such that for all graphs $G = (V, E)$ on $n \geq k_1$ vertices the following holds. Let G be given together with a partition $V = V_1^* \dot\cup \dots \dot\cup V_{k_*}^*$ of its vertices. Then there is $k_0 \leq k \leq k_1$ such that G has an ε-regular equipartition $V = V_0 \dot\cup V_1 \dot\cup \dots \dot\cup V_k$ refining $V_1^* \dot\cup \dots \dot\cup V_{k_*}^*$*

We also say that $V = V_1^* \dot\cup \dots \dot\cup V_{k_*}^*$ is a *prepartition* of G.

2.2. **Coloured graphs.** A *coloured graph* G is a graph (V, E) together with a 2-colouring of its edges by red and green. We denote by $G(c)$ the subgraph of G formed by exactly those edges with colour c. Two vertices are *connected* in G if they lie in the same connected component of G and are *c-connected* in G if they are connected in $G(c)$. Let G be a coloured graph and v be a vertex of G and $c \in \{\text{red,green}\}$. Then a vertex u is a *c-neighbour* of v if uv is an edge of colour c in G. The *c-neighbourhood* of v is the set of all c-neighbours of v.

Definition 4 (connected, odd, even). *Let G' be either a subgraph of an uncoloured graph G, or a c-monochromatic subgraph of a coloured graph G. Then we say that G' is* connected *if any two vertices covered by G' are connected, respectively c-connected, in G. Further, the component of*

G, respectively of $G(c)$, containing G' is called the component of G' *and is denoted by $G[G']$. Further, G' is* odd *if there is an odd cycle in $G[G']$, otherwise G' is* even.

Notice that this notion of connected subgraphs differs from the standard one. A red-connected matching is a good example to illustrate this concept: it is a matching with all edges coloured in red and with a path (in the original graph) of red colour between any two vertices covered by the matching. For subgraphs containing edges of different colours the notion of connectedness is not defined.

Definition 5 (fork, fork system)**.** *An r-*fork *(or simply* fork*) is the complete bipartite graph $K_{1,r}$. We also say that an r-fork has r* prongs *and one* center *by which we refer to the vertices in the two partition classes of $K_{1,r}$. A* fork system *F in a graph G is a set of pairwise vertex disjoint forks in G (not necessarily having the same number of prongs). We say that F has* ratio *r if all its forks have at most r prongs. Then we also call F an r-fork system.*

Suppose F is a connected fork system in G. If F is even then the *size* f of F is the order of the bigger bipartition class of $G[F]$. If F is odd then F has size at least f if there is a connected bipartite subgraph G' of G such that F has size f in G'. For a vertex set D in G we say that F is *centered* in D if the centers of the forks in F all lie in D.

Next, we define two properties of coloured graphs that characterise structures (in a reduced graph) suitable for the embedding of trees as we shall see later (cf. Section 4). Roughly speaking, these properties guarantee the existence of large monochromatic connected matchings and fork-systems.

Definition 6 (m-odd, (m, f, r)-good)**.** *Let G be a coloured graph on n vertices. Then G is called m-*odd *if G contains a monochromatic odd connected matching of size at least m. We say that G is (m, f, r)-*good *(in colour c) if G contains a c-coloured connected matching M of size at least m as well as a c-coloured connected fork system F of size at least f, and ratio at most r.*

We further need to define a set of special, so-called extremal, configurations of coloured graphs that will need special treatment in our proofs. To prepare their definition, let K be a graph on n vertices and D, D' be disjoint vertex sets in K. We say that the bipartite graph $K[D, D']$ is *η-complete* if each vertex of $K[D, D']$ is incident to all but at most ηn vertices of the other bipartition class. If K is a coloured graph then $K[D, D']$ is *(η, c)-complete* for some colour c if it is η-complete and all edges in $K[D, D']$ are of colour c. We call a set A *negligible* if $|A| < 2\eta n$. Otherwise, A is *non-negligible.*

Definition 7 (extremal)**.** *Let $K = (V, E)$ be a coloured graph of order $3n$. Suppose that $\eta > 0$ is given. We say that K is a pyramid configuration with parameter η if it satisfies 1 below and a spider configuration if it satisfies 2. In both cases we call K* extremal with parameter η *or* η-extremal. *Otherwise we say that K is* not η-extremal.

(1) pyramid configurations: *There are (not necessarily distinct) colours c, c' and pairwise disjoint subsets $D_1, D_2, D'_1, D'_2 \subseteq V$ of size at most n, with $|D_1|, |D_2| \geq (1-\eta)n$ and $|D'_1| + |D'_2| \geq (1-\eta)n$ where D'_1 and D'_2 are either empty or non-negligible. Further, $K[D_1, D'_1]$ and $K[D_2, D'_2]$ are (η, c)-complete and $K[D_1, D'_2]$, $K[D_2, D'_1]$, and $K[D_1, D_2]$ are η-complete.*

In addition, either $K[D_1, D_2]$ is (η, c')-complete or both $K[D_1, D'_2]$ and $K[D'_1, D_2]$ are (η, c')-complete. In the first case we say the pyramid configuration has a c'-tunnel, *and in the second case that it has a* crossing. *The pairs (D_1, D'_1) and (D_2, D'_2) are also called the* pyramids *of this configuration.*

(2) spider configuration: *There is a colour c and pairwise disjoint subsets $A_1, A_2, B_1, B_2, C_1, C_2 \subseteq V$ such that $|D_1 \cup D_2| \geq (1-\eta)n$ and $K[D_1, D'_2]$ is (η, c)-complete for all $D, D' \in \{A, B, C\}$ with $D \neq D'$, the edges in all these bipartite graphs together form a connected bipartite subgraph K_c of K with (bi)partition classes $A_1 \dot\cup B_1 \dot\cup C_1$ and $A_2 \dot\cup B_2 \dot\cup C_2$. Further there are sets $A_B \dot\cup A_C = A_2$, $B_A \dot\cup B_C = B_2$, and $C_A \dot\cup C_B \dot\cup C_C = C_2$, each of which is either empty or non-negligible, such that the following conditions are satisfied for all $\{D, D', D''\} = \{A, B, C\}$:*

   (a) $|A_1| \geq |B_1| \geq |C_1 \cup C_C|$ *and* $|D_{D'}| = |D'_D| \leq n - |D''_2|$,

   (b) *either* $C_C = \emptyset$ *or* $A_B = \emptyset$,

   (c) *either* $A_2 = \emptyset$ *or* $|A_2 \cup B_2 \cup C_A \cup C_B| \leq (1-\eta)\frac{3}{2}n$,

   (d) *either* $C_1 = \emptyset$ *or* $|A_1 \cup B_1 \cup C_1| < (1-\eta)\frac{3}{2}n$ *or* $|B_1 \cup C_1| \leq (1-\eta)\frac{3}{4}n$.

By $\mathcal{K}^\eta_n$, finally, we denote the class of all spanning subgraphs K of $K_{n,n,n}$ with minimum degree $\delta(K) > (2-\eta)n$. We also call the graphs in this class *η-complete tripartite graphs.*

3. Connected matchings and fork systems

In order to prove Theorem 1 we will use the following structural result about coloured graphs from $\mathcal{K}^\eta_n$. It asserts that such graphs either contain large monochromatic odd connected matchings or appropriate connected fork systems. With the help of the regularity method we will then, in Section 4, use this result (on the reduced graph of a regular partition) to find monochromatic trees. The reason why odd connected matchings and connected fork systems are useful for this task is explained in Section 4.1.

Lemma 8. *For all $\eta' > 0$ there are $\eta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Every coloured graph $K \in \mathcal{K}_n^\eta$ is either $(1-\eta')\frac{3}{4}n$-odd or $\big((1-\eta')n, (1-\eta')\frac{3}{2}n, 3\big)$-good.*

We remark that the dependence of the constant n_0 and η' is only linear, and in fact we can choose $n_0 = \eta'/200$. As we will see below, Lemma 8 is a consequence of the following two lemmas. The first lemma analyses non-extremal members of $\mathcal{K}_n^\eta$.

Lemma 9 (non-extremal configurations)**.** *For all $\eta' > 0$ there are $\eta \in (0, \eta')$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let K be a coloured graph from $\mathcal{K}_n^\eta$ that is not η'-extremal. Then K is $(1-\eta')\frac{3}{4}n$-odd.*

The second lemma handles the extremal configurations.

Lemma 10 (extremal configurations)**.** *For all $\eta' > 0$ there is $\eta \in (0, \eta')$ such that the following holds. Let K be a coloured graph from $\mathcal{K}_n^\eta$ that is η-extremal. Then K is $((1-\eta')n, (1-\eta')\frac{3}{2}n, 3)$-good.*

Proofs of Lemma 9 and 10 are provided in Sections 7.2 and 7.3, respectively. We get Lemma 8 as an easy corollary.

Proof of Lemma 8. Given η' let $\eta_{\mathrm{L10}} < \eta'$ be the constant provided by Lemma 10 for input η' and let η_{L9} be the constant produced by Lemma 9 for input $\eta'_{\mathrm{L9}} := \eta_{\mathrm{L10}}$. Set $\eta := \min\{\eta_{\mathrm{L10}}, \eta_{\mathrm{L9}}\}$ and let $K \in \mathcal{K}_n^\eta$ be a given coloured graph. Then $K \in \mathcal{K}_n^{\eta_{\mathrm{L9}}}$ and by Lemma 9 the graph K is either $(1-\eta'_{\mathrm{L9}})3n/4$-odd (and thus $(1-\eta')3n/4$-odd as oddness is monotone) or η_{L10}-extremal. In the first case we are done and in the second case Lemma 10 implies that K is $\big((1-\eta')n, (1-\eta')\frac{3}{2}n, 3\big)$-good (goodness is also monotone) and we are also done. $\square$

4. Proof of theorem 1

In this section we will first briefly outline the main ideas for the proof of Theorem 1. Then we will state the remaining necessary lemmas, most notably our main embedding result (Lemma 13). These lemmas will be proved in the subsequent sections. At the end of this section we finally provide a proof of Theorem 1.

4.1. **The idea of the proof.** We apply the Regularity Lemma on the coloured graph $K_{n,n,n}$ with prepartition as given by the partition classes of $K_{n,n,n}$. As a result we obtain a coloured reduced graph $\mathbb{K} \in \mathcal{K}_k^\eta$ where the colour of an edge in $\mathbb{K}$ corresponds to the majority colour in the underlying regular pair. Such a regular pair is well-known to possess almost as good embedding properties as a complete bipartite graph. We apply our structural result Lemma 8 and infer that $\mathbb{K}$ is either $(1-\eta')\frac{3}{4}k$-odd or $((1-\eta')k, (1-\eta')\frac{3}{2}k, 3)$-good, i.e., there is a colour, say green, such that $\mathbb{K}$ contains either an odd connected green matching $\mathbb{M}_o$ of size at least $(1-\eta')\frac{3}{4}k$, or it contains a connected greed matching $\mathbb{M}$ of size at least $(1-\eta')k$ and a 3-fork system $\mathbb{F}$ of size at least $(1-\eta')\frac{3}{2}k$. We shall show that using either of these structures we can embed any tree $T \in \mathcal{T}_t^\Delta$ into the green subgraph of $K_{n,n,n}$. As a preparatory step, we cut T into small subtrees (see Lemma 15), called shrubs.

Now let us first consider the case when we have an odd matching $\mathbb{M}_o$. Our aim is to embed each shrub S into a regular pair (A, B) corresponding to an edge $e \in \mathbb{M}_o$. Shrubs are bipartite graphs. Therefore there are two ways of assigning the colour classes of S to the clusters of e. This corresponds to two different *orientations* of S for the embedding in (A, B). Our strategy is to choose orientations for all shrubs (and hence assignments of their colour classes to clusters of edges in $\mathbb{M}_o$) in such a way that every cluster of $V(\mathbb{M}_o)$ receives roughly the same number of vertices of T. We will show that this is possible without "over-filling" any cluster. It follows that we can embed all shrubs into regular pairs corresponding to edges of $\mathbb{M}_o$. The fact that $\mathbb{M}_o$ is connected and *odd* then implies that between any pair of edges in $\mathbb{M}_o$ there are walks of both even and odd length in the reduced graph. We will show that this allows us to connect the shrubs and to obtain a copy of T in the green subgraph of $K_{n,n,n}$.

For the second case, i.e., the case when we have a matching $\mathbb{M}$ as well as a 3-fork system $\mathbb{F}$ the basic strategy remains the same. We assign shrubs to edges of $\mathbb{M}$ or $\mathbb{F}$. In difference to the previous case, however, these substructures of the reduced graph are not odd. This means that we cannot choose the orientations of the shrubs as before. Rather, these orientations are determined by the connections between the shrubs. Therefore, we distinguish the following two situations when embedding the tree T. If the partition classes of T are reasonably balanced, then we use the matching $\mathbb{M}$ for the embedding. If T is unbalanced, on the other hand, we employ the fork system $\mathbb{F}$ and use the prongs of the forks in $\mathbb{F}$ to accommodate the bigger partition class of T and the centers for the smaller.

4.2. **The main embedding lemma.** As indicated, in the proof of the main theorem we will use the regularity lemma in conjunction with an embedding lemma (Lemma 13). This lemma states that a tree T can be embedded into a graph given together with a regular partition if there is a homomorphism from T to the reduced graph of the partition with suitable properties. In the following definition of a valid assignment we specify these properties. Roughly speaking, a valid assignment is a homomorphism h from a tree T to a (reduced) graph $\mathbb{G}$ such that no vertex of $\mathbb{G}$ receives too many vertices of T and that does not "spread" in the tree too quickly in the following sense: for each vertex $x \in V(T)$ we require that the neighbours of x occupy at most two vertices of $\mathbb{G}$.

Definition 11 (valid assignment)**.** *Let $\mathbb{G}$ be a graph on vertex set $[k]$, let T be a tree, $\varrho \in [0,1]$ and $L \in \mathbb{N}$. A mapping $h\colon V(T) \to [k]$ is a (ϱ, L)*-valid assignment *of T to $\mathbb{G}$ if*

(1) *h is a homomorphism from T to $\mathbb{G}$,*
(2) *$|h(\mathrm{N}_T(x))| \leq 2$, for every $x \in V(T)$,*
(3) *$|h^{-1}(i)| < (1-\varrho)L$, for every $i \in [k]$.*

In addition we need the concept of a cut of a tree, which is a set of vertices that cuts the tree into small components which we call shrubs.

Definition 12 (cut, shrubs)**.** *Let $S \in \mathbb{N}$ and T be a tree with vertex set $V(T)$. A set $C \subseteq V(T)$ is an S*-cut *(or simply* cut*) of T if all components of $T - C$ are of size at most S. The components of $T - C$ are called the* shrubs *of T corresponding to C.*

Now we can state the main embedding lemma.

Lemma 13 (main embedding lemma)**.** *Let G be an n-vertex graph with an (ε, d)-reduced graph $\mathbb{G} = ([k], E(\mathbb{G}))$ and let T be a tree with $\Delta(T) \leq \Delta$ and an S-cut C. If T has a $\big(\varrho, (1-\varepsilon)\frac{n}{k}\big)$-valid assignment to $\mathbb{G}$ and $(\frac{1}{10}d\varrho - 10\varepsilon)\frac{n}{k} \geq |C| + S + \Delta$ then $T \subseteq G$.*

The proof of this lemma is deferred to Section 6. Before we can apply it for embedding a tree T in the proof of Theorem 1 we need to construct a valid assignment for T. This is taken care of by the following lemma which states that this is possible if the reduced graph of some regular partition contains an odd connected matching or a suitable fork system. The proof of this lemma is given in Section 5.

Lemma 14 (assignment lemma)**.** *For all $\varepsilon, \mu > 0$ with $\varepsilon < \mu/10$ and for all $k \in \mathbb{N}$ there is $\alpha = \alpha(k) > 0$ and $n_0 = n_0(\mu, \varepsilon, k) \in \mathbb{N}$ such that for all*

$n \geq n_0$, all $r \in \mathbb{N}$, all graphs $\mathbb{G}$ of order k, and all trees T with $\Delta(T) \leq n^\alpha$ the following holds. Assume that either

(1) *$\mathbb{G}$ contains an odd connected matching of size at least m and that $t := |V(T)| \leq (1-\mu)2m\frac{n}{k}$, or*

(6) *$\mathbb{G}$ contains a connected fork system with ratio r and size at least f, and T has colour class sizes t_1 and t_2 with $t_2 \leq t_1 \leq t'$ and $t_2 \leq t'/r$, where $t' = (1-\mu)f\frac{n}{k}$.*

Then there is an $(\varepsilon\frac{n}{k})$-cut C of T with $|C| \leq \varepsilon\frac{n}{k}$ and a $\left(\frac{1}{2}\mu, (1-\varepsilon)\frac{n}{k}\right)$-valid assignment of T to $\mathbb{G}$.

4.3. **The proof.** Now we have all tools we need to prove the main theorem.

Proof of Theorem 1. We start by defining the necessary constants. Given $\mu > 0$, set $\mu' := \eta'$ in such a way that

$$1 - \tfrac{\mu}{3} \leq (1-\eta')^2(1-\mu'). \tag{1}$$

Lemma 8 with input $\eta' > 0$ provides us with $\eta > 0$ and $k_0 \in \mathbb{N}$. The regularity lemma, Lemma 3, with input

$$\varepsilon := \min\{\tfrac{1}{100}\eta^2, \tfrac{1}{10}\eta'^2, 10^{-3}\mu'\} \tag{2}$$

and k_0 and $k_* := 3$ returns a constant k_1. Next we apply Lemma 14 with input $\frac{\varepsilon}{10}$ and μ' separately for each value $3k$ with $k_0 \leq 3k \leq k_1$ and get constants $\alpha(3k)$ and $n_0'(3k)$ for each of these applications. Set $\alpha := \min\{\alpha(3k) \colon k_0 \leq 3k \leq k_1\}$ and $n_0' := \max\{n_0'(3k) \colon k_0 \leq 3k \leq k_1\}$. Finally, choose

$$n_0 := \max\{n_0', k_1, (\tfrac{k_1}{\varepsilon})^{1/(1-\alpha)}\}\,. \tag{3}$$

We are given a complete tripartite graph $K_{n,n,n}$ with $n \geq n_0$ as input whose edges are coloured with green and red. Our goal is to select a colour and show that in this colour we can embed every member of $\mathcal{T}_t^\Delta$ with $\Delta \leq n^\alpha$ and $t \leq (3-\mu)n/2$.

We first select the colour. To this end let G and R be the subgraphs of $K_{n,n,n}$ formed by the green and red edges, respectively. We apply the regularity lemma, Lemma 3, with input $\frac{\varepsilon}{10}$ on the graph G with prepartition $V_1^* \dot\cup V_2^* \dot\cup V_3^*$ as given by the three partition classes of $K_{n,n,n}$. We obtain an $\frac{\varepsilon}{10}$-regular equipartition $V = V_0 \dot\cup V_1 \dot\cup \ldots \dot\cup V_{3k}$ refining this prepartition such that $k_0 \leq 3k \leq k_1$. Each cluster of this partition lies entirely in one of the partition classes of $K_{n,n,n}$. Let $\mathbb{K} = ([3k], E_{\mathbb{K}})$ be the graph that contains edges for all ε-regular cluster pairs that do not lie in the same partition class of $K_{n,n,n}$. Clearly, $\mathbb{K}$ is a tripartite graph. Furthermore, there are less than εk^2 pairs (V_i, V_j) in our regular partition that are not $\frac{\varepsilon}{10}$-regular in G. It follows that at most $2\sqrt{\varepsilon}k$ clusters V_i are contained in more than $\sqrt{\varepsilon}k$ irregular pairs. We move all these clusters and possibly up to $6\sqrt{\varepsilon}k$ additional clusters to the bin set V_0. The additional clusters are chosen in such a way that we obtain in each partition class of $\mathbb{K}$ the same number of clusters. We call the resulting bin set V_0' and denote the remaining clusters by $V_1'\dot\cup\ldots\dot\cup V_{3k'}'$ and the corresponding subgraph of $\mathbb{K}$ by $\mathbb{K}'$. Observe that $k' \geq (1-3\sqrt{\varepsilon})k$. Because each remaining cluster forms an irregular pair with at most $\sqrt{\varepsilon}k \leq 2\sqrt{\varepsilon}k' \leq \eta' k'$ of the remaining clusters we conclude that $\mathbb{K}'$ is a graph from $\mathcal{K}_{k'}^{\eta}$. In addition, it easily follows from the definition of ε-regularity that each pair (V_i, V_j) with $i, j \in [k']$ which is ε-regular in G is also ε-regular in R. This motivates the following "majority" colouring of $\mathbb{K}'$: We colour the edges ij of $\mathbb{K}'$ by green if the ε-regular pair (V_i, V_j) has density at least $\frac{1}{2}$ and by red otherwise. In this way we obtain a coloured graph $\mathbb{K}_c' \in \mathcal{K}_{k'}^{\eta}$.

Now we are in a position to apply Lemma 8 to $\mathbb{K}_c'$. This lemma asserts that $\mathbb{K}_c'$ is either $(1-\eta')\frac{3}{4}k'$-odd or $\big((1-\eta')k', (1-\eta')\frac{3}{2}k', 3\big)$-good. By definition this means that in one of the colours of $\mathbb{K}_c'$, say in green, we

(1) either have an odd connected matching $\mathbb{M}_o$ of size $m_1 \geq (1-\eta')\frac{3}{4}k' \geq (1-\eta')(1-3\sqrt{\varepsilon})\frac{3}{4}k$,

(7) or a connected matching $\mathbb{M}$ of size $m_2 \geq (1-\eta')k' \geq (1-\eta')(1-3\sqrt{\varepsilon})k$ together with a connected fork system $\mathbb{F}$ of size $f \geq (1-\eta')\frac{3}{2}k' \geq (1-\eta')(1-3\sqrt{\varepsilon})\frac{3}{2}k$ and ratio 3.

In the following we use the matchings and fork systems we just obtained to show that we can embed all trees of $\mathcal{T}_t^{\Delta}$ in the corresponding system of regular pairs. For this purpose let $\mathbb{G}$ be the graph on vertex set $[3k]$ that contains precisely all green edges of $\mathbb{K}_c'$. Observe that $\mathbb{G}$ is an $(\varepsilon, 1/2)$-reduced graph for G.

Let $T \in \mathcal{T}_t^{\Delta}$ be a tree of order $t \leq (3-\mu)n/2$ and with maximal degree $\Delta(T) \leq n^{\alpha}$. Now we distinguish two cases, depending on whether we obtained configuration 1 or configuration 7 from Lemma 8. In both cases we plan to appeal to Lemma 14 to show that T has

(4)
an $(\varepsilon\frac{n}{k})$-cut C with $|C| \leq \varepsilon\frac{n}{k}$ and a $(\frac{1}{2}\mu', (1-\varepsilon)\frac{3n}{3k})$-valid assignment to $\mathbb{G}$.

Recall that we fed constants ε, $\mu' > 0$ and $3k$ into this lemma. Assume first that we are in configuration 1. Because $m_1 \geq (1-\eta')(1-2\sqrt{\varepsilon})\frac{3}{4}k$ we have

$$t \leq (3-\mu)\frac{n}{2} \leq 3(1-\tfrac{\mu}{3})\frac{n}{2}\frac{m_1}{(1-\eta')(1-3\sqrt{\varepsilon})\frac{3}{4}k} \overset{(1),(2)}{\leq} (1-\mu')2m_1 \cdot \frac{3n}{3k}\,.$$

Hence by 1 of Lemma 14 applied with the matching $\mathbb{M}_o$ (with n replaced by $\tilde{n} := 3n$ and k replaced by $\tilde{k} := 3k$) we get (4) for T in this case, as $\Delta(T) \le n^\alpha \le \tilde{n}^\alpha$.

If we are in configuration 7, on the other hand, then let $t_1 \ge t_2$ be the sizes of the two colour classes of T. We distinguish two cases, using the two different structures provided in 7. Assume first that $t_2 \le \frac{t}{3}$. Then, we calculate similarly as above that

$$t_2 \le \tfrac{1}{3}t \le (1 - \tfrac{1}{3}\mu)\tfrac{n}{2} \le \tfrac{1}{3}(1-\mu')f\tfrac{3n}{3k}\,, \qquad \text{and} \qquad t_1 \le t \le (1-\mu')f\tfrac{3n}{3k}\,.$$

Otherwise, if $t_2 \ge \frac{t}{3}$ then, similarly,

$$t_2 \le t_1 \le \tfrac{2}{3}t \le (1 - \tfrac{\mu}{3})n \le (1-\mu')m_2 \cdot \tfrac{3n}{3k}\,.$$

Consequently, in both cases we can appeal to 6 of Lemma 14, in the first case applied to $\mathbb{F}$ and in the second to $\mathbb{M}$. We obtain (4) for T as desired.

We finish our proof with an application of the main embedding lemma, Lemma 13. As remarked earlier $\mathbb{G}$ is an $(\varepsilon, 1/2)$-reduced graph for G. We further have (4). For applying Lemma 13 it thus remains to check that $(\frac{1}{2} \cdot \frac{1}{10}\varrho - 10\varepsilon)\frac{n}{k} \ge S + |C| + \Delta$ with $\varrho = \frac{1}{2}\mu'$, $S = \varepsilon\frac{n}{k}$, $|C| \le \varepsilon\frac{n}{k}$, and $\Delta \le n^\alpha$. Indeed,

$$\left(\tfrac{1}{20}\varrho - 10\varepsilon\right)\tfrac{n}{k} = \left(\tfrac{1}{20}\cdot\tfrac{1}{2}\mu' - 10\varepsilon\right)\tfrac{n}{k} \overset{(2)}{\ge} 3\varepsilon\tfrac{n}{k} \overset{(3)}{\ge} \varepsilon\tfrac{n}{k} + \varepsilon\tfrac{n}{k} + n^\alpha\,.$$

So Lemma 13 ensures that $T \subseteq G$, i.e., there is an embedding of T in the subgraph induced by the green edges in $K_{n,n,n}$. □

5. Valid Assignments

In this section we will provide a proof for Lemma 14. The idea is as follows. Given a tree T and a graph G with reduced graph $\mathbb{G}$ we first construct a cut of T that provides us with a collection of small shrubs (see Lemma 15). Then we distribute these shrubs to edges of the given matching or fork-system in $\mathbb{G}$ (see Lemmas 16 and 17). Finally, we slightly modify this assignment in order to obtain a homomorphism from T to $\mathbb{G}$ that satisfies the conditions required for a valid assignment (see Lemma 20).

Lemma 15. *For every $S \in \mathbb{N}$ and for any tree T there is an S-cut of T that has size at most $\frac{|V(T)|}{S}$.*

Proof. To prove Lemma 15 we need the following fact.

Fact 1. *For any $S \in \mathbb{N}$ and any tree T with $|V(T)| > S$, there is a vertex $x \in V(T)$ such that the following holds. If F_x is the forest consisting of all components of $T - x$ with size at most S, then $|V(F_x)| + 1 > S$.*

To see this, root the tree T at an arbitrary vertex x_0. If x_0 does not have the required property, it follows from $|V(T)| > S$ that there exists a component T_1 in $T - x_0$ with $|V(T_1)| > S$. Set $x_1 := \mathrm{N}(x_0) \cap V(T_1)$. Let $F(T_1 - x_1)$ be the forest consisting of the components of $T_1 - x_1$ that have size at most S. Observe that $F(T_1 - x_1)$ is a subgraph of F_{x_1}. So if $|F(T_1 - x_1)| + 1 > S$, then x_1 has the property required by Fact 1. Otherwise there exists a component T_2 in $T_1 - x_1$ of size larger than S. Observe that T_2 is also a component of $T - x_1$. Now repeat the procedure just described by setting $x_2 = \mathrm{N}(x_1) \cap V(T_2)$ and so on, i.e., more generally we obtain trees T_i and vertices $x_i = \mathrm{N}(x_{i-1}) \cap V(T_i)$. As the size of T_i decreases as i increases, there must be an x_i with the property required by Fact 1.

Now we prove Lemma 15. Set $C = \emptyset$. Repeat the following process until it stops. Choose a component T' of $T - C$ with size larger than S. Apply Fact 1 to T' and obtain a cut vertex x of T' together with a forest F_x consisting of components of $T' - x$ that have size at most S and is such that $|V(F_x) \cup \{x\}| > S$. Add x to C and repeat unless there is no component of size larger than S in $T - C$. As $|V(T - C)|$ decreases this process stops. Observe that then C is an S-cut. By the choice of C we obtain

$$|V(T)| = \sum_{x \in C} |V(F_x) \cup \{x\}| > |C| \cdot S,$$

which implies the required bound on the size of C. $\square$

After Lemma 15 provided us with a cut and some corresponding shrubs we will distribute each of these shrubs T_i to an edge e of the odd matching or the fork system in the reduced graph by assigning one colour class of T_i to one end of e and the other colour class to the other end. Here our goal is to distribute the shrubs and their vertices in such a way that no cluster receives too many vertices. The next two lemmas guarantee that this can be done. Lemma 16 takes care of the distribution of the shrubs to the clusters of a matching M and Lemma 17 to those of a fork system F. We provide Lemmas 16 and 17 with numbers $a_{i,1}$ and $a_{i,2}$ as input. These numbers represent the sizes of the colour classes $A_{i,1}$ and $A_{i,2}$ of the shrub T_i. Since we do not need any other information about the shrubs in these lemmas the shrubs T_i do not explicitly appear in their statement. Both lemmas then produces a mapping ϕ representing the assignment of the colour classes $A_{i,1}$ and $A_{i,2}$ to the clusters of M or F.

Lemma 16. *Let $\{a_{i,j}\}_{i \in [s],\, j \in [2]}$ be natural numbers with sum at most t and $a_{i,1} + a_{i,2} \le S$ for all $i \in [s]$, and let M be a matching on vertices $V(M)$.*

Then there is a mapping $\phi\colon [s]\times[2]\to V(M)$ such that $\phi(i,1)\phi(i,2)\in M$ for all $i\in[s]$ and

$$\sum_{(i,j)\in\phi^{-1}(v)} a_{i,j} \le \frac{t}{2|M|} + 2S \qquad \text{for all } v\in V(M). \tag{5}$$

Proof. A simple greedy construction gives the mapping ϕ: We consider the numbers $a_{i,j}$ as weights that are distributed, first among the edges, and then among the vertices of M. For this purpose greedily assign pairs $(a_{i,1}, a_{i,2})$ to the edges of M, in each step choosing an edge with minimum total weight. Then, clearly, no edge receives weight more than $S + t/|M|$. In a second round, do the following for each edge vw of M. For the pairs $(a_{i,1}, a_{i,2})$ that were assigned to e, greedily assign one of the weights of this pair to v and the other one to w, such that the total weight on v and on w are as equal as possible. Hence each of these vertices receives weight at most $\frac{1}{2}(S + t/|M|) + S$ and so the mapping ϕ corresponding to this weight distribution satisfies the desired properties. □

Lemma 17. *Let $\{a_{i,1}\}_{i\in[s]}$ and $\{a_{i,2}\}_{i\in[s]}$ be natural numbers with sum at most t_1 and t_2, respectively. Let $S \le t_1 + t_2 =: t$ and assume that $a_{i,1} + a_{i,2} \le S$ for all $i\in[s]$. Let F be a fork system with ratio at most r and partition classes $V_1(F)$ and $V_2(F)$ where $|V_1(F)| \ge |V_2(F)|$. Then there is a mapping $\phi\colon [s]\times[2]\to V_1(F)\cup V_2(F)$ such that $\phi(i,1)\phi(i,2)\in F$ and $\phi(i,j)\in V_j(F)$ for all $i\in[s], j\in[2]$ satisfying that for all $v_1\in V_1(F)$, $v_2\in V_2(F)$ we have*

$$\sum_{(i,1)\in\phi^{-1}(v_1)} a_{i,1} \le \frac{t_1}{|F|} + \sqrt{12tS|F|} \qquad \textit{and} \qquad \sum_{(i,2)\in\phi^{-1}(v_2)} a_{i,2} \le \frac{rt_2}{|F|} + \sqrt{12tS|F|}\,. \tag{6}$$

In the proof of this lemma we will make use the so-called Hoeffding bound for sums of independent random variables (see, e.g., [2, Theorem A.1.16]).

Theorem 18. *Let $X_1,\dots,X_s$ be independent random variables with $\mathbb{E}X_i = 0$ and $|X_i|\le 1$ for all $i\in[s]$ and let X be their sum. Then $\mathbb{P}[X > a] \le \exp(-a^2/(2s))$.* □

Proof of Lemma 17. For showing this lemma we use a probabilistic argument and again consider the $a_{i,j}$ as weights which are distributed among the vertices of F.

Observe first that we can assume without loss of generality that for all but at most one $i\in[s]$ we have $\frac{1}{2}S \le a_{i,1} + a_{i,2}$ since otherwise we can group weights $a_{i,1}$ together, and also group the corresponding $a_{i,2}$ together, such that this condition is satisfied and continue with these grouped weights. This in turn implies, that $s \leq (2t/S) + 1 \leq 3t/S$.

We start by assigning weights $a_{i,1}$ to vertices of $V_1(F)$ by (randomly) constructing a mapping $\phi_1 \colon [s] \times [1] \to V_1(F)$. To this end, independently and uniformly at random choose for each $i \in [s]$ an image $\phi_1(i,1)$ in $V_1(F)$. Clearly, there is a unique way of extending such a mapping ϕ_1 to a mapping $\phi \colon [s] \times [2] \to V_1(F) \cup V_2(F)$ satisfying $\phi(i,1)\phi(i,2) \in F$. We claim that the probability that ϕ_1 gives rise to a mapping ϕ which satisfies the assertions of the lemma is positive.

Indeed, for any fixed vertex $v = v_1 \in V_1(F)$ or $v = v_2 \in V_2(F)$ let $\sigma(v)$ be the event that the mapping ϕ does not satisfy (6) for v. We will show that $\sigma(v)$ occurs with probability strictly less than $1/(2|F|)$, which clearly implies the claim above. We first consider the case $v = v_1 \in V_1(F)$. For each $i \in [s]$ let $\mathbb{1}_i$ be the indicator variable for the event $\phi(i,1) = v_1$ and define a random variable X_i by setting

$$X_i = \left(\mathbb{1}_i - \tfrac{1}{|F|}\right) \tfrac{a_{i,1}}{S}.$$

Observe that these variables are independent, and satisfy $\mathbb{E}X_i = 0$ and $|X_i| \leq 1$ and so Theorem 18 applied with $a = \sqrt{12t|F|/S}$ asserts that

$$\mathbb{P}\Big[\sum_{i\in[s]} X_i > \sqrt{12t|F|/S}\Big] \leq \exp\left(-\frac{12t|F|}{S \cdot 2s}\right) \leq \exp\left(-2|F|\right) < \frac{1}{2|F|} \tag{7}$$

where we used $s \leq 3t/S$. Now, by definition we have

$$X := \sum_{i\in[s]} X_i = \frac{1}{S} \sum_{(i,1)\in\phi^{-1}(v_1)} a_{i,1} - \frac{t_1}{|F|S},$$

and so, if (6) did not hold for v_1, then we had $X > \sqrt{12t|F|/S}$, which by (7) occurs with probability less than $1/(2|F|)$.

For the case $v = v_2 \in V_2(F)$ we proceed similarly. Let $r' \leq r$ be the number of prongs of the fork that contains v_2. We define indicator variables $\mathbb{1}'_i$ for the events $\phi(i,2) = v_2$ for $i \in [s]$ and random variables

$$Y_i = \left(\mathbb{1}'_i - \tfrac{r'}{|F|}\right) \tfrac{a_{i,2}}{S}.$$

with $\mathbb{E}Y_i = 0$ and $|Y_i| \leq 1$. The rest of the argument showing that $\sigma(v_2)$ occurs with probability strictly less than $1/(2|F|)$ is completely analogous to the case $v = v_1$ above. With this we are done. □

As explained earlier these two previous lemmas will allow us to assign the shrubs of a tree T to edges of a reduced graph $\mathbb{G}$. By applying them we will obtain a mapping ψ from the vertices of T to those of $\mathbb{G}$ that is a homomorphism when restricted to the shrubs of T. The following lemma transforms such a ψ to a homomorphism h from the whole tree T to $\mathbb{G}$ that "almost" coincides with ψ provided the structures of T and $\mathbb{G}$ are "compatible" with respect to ψ in the sense of the following definition.

Definition 19 (walk condition)**.** *Let T be a tree and $C \subseteq V(T)$. A mapping $\psi\colon V(T) \setminus C \to \mathbb{G}$ satisfies the* walk condition *if for any $x, y \in V(T) \setminus C$ such that there is a path $P_{x,y}$ from x to y whose internal vertices are all in C there is a walk $\mathbb{P}_{x,y}$ between $\psi(x)$ and $\psi(y)$ in $\mathbb{G}$ such that the length of $P_{x,y}$ and the length of $\mathbb{P}_{x,y}$ have the same parity.*

Lemma 20. *Let T be a tree with maximal degree Δ, let C be a cut of T, and let $\mathbb{G}$ be a graph on k vertices. Let $\psi\colon V(T)\setminus C \to V(\mathbb{G})$ be a homomorphism that maps each shrub of T corresponding to C to an edge of $\mathbb{G}$ and that satisfies the walk condition. Then there is a homomorphism $h\colon V(T) \to V(\mathbb{G})$ satisfying*

(1) $|h(\mathrm{N}_T(x))| \leq 2$ *for all vertices $x \in V(T)$ and*
(2) $|\{x \in V(T):\ h(x) \neq \psi(x)\}| \leq 3|C|\Delta^{2k+1}$.

Observe that Property 1 in this lemma asserts that images of neighbours of any vertex in T occupy at most two vertices in $\mathbb{G}$. By assumption, this is clearly true for ψ but we need to make sure that h inherits this feature. Property 2 on the other hand states that h and ψ do not differ much. The assumption that ψ satisfies the walk condition is essential for the construction of the homomorphism h.

Proof of Lemma 20. We start with some definitions. Choose a non-empty shrub corresponding to C in T and call it *shrub* 1. Then choose a cut-vertex $x_0^* \in C$ adjacent to this shrub. We consider x_0^* as the *root* of the tree T. This naturally induces the following partial order $\prec$ on the vertices $V(T)$ of T: For vertices $x, y \in V(T)$ we have $x \prec y$ iff y is a descendant of x in the tree T with root x_0^*. Note that x_0^* is the unique minimal element of $\prec$ and the leaves of T are its maximal elements. Further, for $x \in C$ set $W_x := \{z \in V(T):\ \mathrm{dist}_T(x,z) \leq 2k+1 \ \&\ x \prec z\}$ and let $W = C \cup \bigcup_{x\in C} W_x$. Observe that the bound on the maximal degree of T implies that $|W| \leq 2\Delta^{2k+1}|C| + |C| \leq 3\Delta^{2k+1}|C|$. For $x \in V(T) \setminus W$, we set $h(x) := \phi(x)$. This ensures that Condition 2 is fulfilled. In addition the following fact holds because ψ maps each shrub to an edge of $\mathbb{G}$.

Fact 1. *The mapping h restricted to $V(T) \setminus W$ is a homomorphism. For all vertices $x \in V(T) \setminus W$ all children y of x that are not cut-vertices have the same $h(y)$.*

We shall extend h to the set W. Our strategy is roughly as follows: We start by defining $h(x_0^*)$ for the root cut-vertex x_0^* in a suitable way. Recall that all children of x_0^* are contained in W. Then, we let h map all non-cut-vertex children $y \in \mathrm{N}_T(x_0^*) \setminus C$ of x_0^* to a suitable neighbour of $h(x_0^*)$ in $\mathbb{G}$ and do the following for each of these y. Observe that y is the root of some shrub, which we will call the *shrub of* y. Now, $h(y)$ and $\psi(y)$ might be different. However, we will argue that there is a walk of even length $m \leq 2k$ between $h(y)$ and $\psi(y)$. Then we will define h for all vertices $y' \in W_{x_0^*}$ contained in the shrub of y and with distance at most m from y. More precisely we will use the walk of length m between $h(y)$ and $\psi(y)$ and let h map all y' with distance i to y to the i-th vertex of this walk. All vertices z in the shrub of y for which h is still undefined after these steps are then mapped to $h(z) := \psi(z)$. Once this has been done for all $y \in \mathrm{N}_T(x_0^*) \setminus C$ we proceed in the same way with the next cut-vertex: We choose a cut-vertex x^* with parent x for which $h(x)$ is already defined and proceed similarly for x^* as we did for x_0^*.

We now make the procedure for the extension of h on W precise. Throughout this procedure we will assert the following property for all non-cut vertices y of T such that $h(y)$ is defined.

(8) There is a path of even length in $\mathbb{G}$ between $h(y)$ and $\psi(y)$.

Observe that (8) trivially holds for all $y \in V(T) \setminus W$.

As explained, we start our procedure with the root x_0^* of the tree T. Let x_1 be the root of shrub 1. By definition x_1 is adjacent to x_0^*. Note that, while ψ is not defined on x_0^* it is defined on x_1. Hence we can legitimately set $h(y) = \psi(x_1)$ for all neighbours $y \notin C$ of x_0^* in T and choose $h(x_0^*)$ arbitrarily in $\mathrm{N}_{\mathbb{G}}(\psi(x_1))$. Observe that this is consistent with (8) because for any neighbour $y \notin C$ of x_0^* we have $h(y) = \psi(x_1)$ and $\mathrm{dist}_T(y, x_1) \in \{0, 2\}$. By assumption ψ satisfies the walk condition. Hence there is a walk in $\mathbb{G}$ with even length between $h(y) = \psi(x_1)$ and $\psi(y)$. Let $\mathbb{P}_y = v_0, v_1, \ldots, v_m$ be a walk in $\mathbb{G}$ of minimal but even length with $v_0 = h(y)$ and $v_m = \psi(y)$. As $\mathbb{G}$ has k vertices we have that $m \leq 2k$. For all vertices $z \in W$ that are in the shrub of y and satisfy $\mathrm{dist}_T(y, z) = j$ for some $j \leq m$, we then define $h(z) := v_j$. For the remaining vertices $z \in W$ in the shrub of y we set $h(z) := \psi(z)$. Observe that this is again consistent with (8) and in conjunction with Fact 1 implies the following condition (which we will also guarantee throughout the whole process of defining h).

Fact 2. *Let $x^* \in C$ and $y \notin C$ such that $h(x^*)$ and $h(y)$ are defined. Then the following holds:*

(1) *All children $y' \notin C$ of x^* have the same $h(y')$ and $h(x^*)h(y') \in E(\mathbb{G})$.*
(2) *All children $y' \notin C$ of y have the same $h(y')$ and $h(y)h(y') \in E(\mathbb{G})$.*

In this way we have defined h for all shrubs adjacent to the root x_0^*.

Next we consider any vertex $x^* \in C \cap \mathrm{N}_T(x_0^*)$ and set $h(x^*) := h(x_1)$, where x_1 is as defined above. We let z^* be the parent of x^*, i.e., $z^* = x_0^*$. Then set $h(y) := h(z^*)$ for all children $y \notin C$ of x^*. This is consistent with Fact 2. Afterwards we have the following situation: x^* and $z^* = x_0^*$ are neighbouring cut-vertices and the vertex x_1 is a non-cut-vertex neighbour of x_0^*. Let $y \in \mathrm{N}_T(x^*) \setminus C$. Then we have $\mathrm{dist}_T(x_1, y) = 3$. Because y and x_1 are both non-cut vertices the properties of ψ imply as before that there is a walk in $\mathbb{G}$ of odd length between $\psi(x_1)$ and $\psi(y)$. By the walk condition and the facts that $h(x_1) = \psi(x_1)$ and $h(x_0^*) = h(y)$, we know that in $\mathbb{G}$ there is a walk $\mathbb{P}_y$ of even length $m \leq 2k$ between $h(y)$ and $\psi(y)$. This verifies (8) for y. We thus can define h for the vertices z contained in the shrub of y as above: if $\mathrm{dist}_T(y, z) \leq m$ then we use this path and set $h(z)$ according to $\mathrm{dist}_T(y, z)$ and otherwise we set $h(z) := \psi(z)$. With this we stay consistent with (8) and Fact 2. We then repeat the above procedure for all $x^* \in C \cap \mathrm{N}_T(x_0^*)$ which implies that the next fact holds true.

Fact 3. *All vertices $x \in \mathrm{N}_T(x_0^*)$ have the same $h(x)$.*

Now we are in the following situation.

Fact 4. *The mapping h is defined on all shrubs adjacent to cut vertices x^* with $h(x^*)$ defined. Moreover, for each cut vertex x^* with $h(x^*)$ undefined that has a parent z for which $h(z)$ is defined, then z has a parent z' with $h(z')$ defined and $h(z)h(z')$ is an edge of $\mathbb{G}$.*

As long as h is not defined for all $z \in V(T)$ we then repeat the following. We choose a cut vertex x^* with $h(x^*)$ undefined that is minimal with respect to this property in $\prec$. Denote the parent of x^* by z and let z' be the parent of z. Then, by Fact 4, the mapping h has already been defined for z' and z. Set $h(x^*) := h(z')$ and for all children $y \notin C$ of x^* set $h(y) := h(z)$. Because $h(z')h(z)$ is an edge of $\mathbb{G}$ by Fact 4 this gives the following property for x^* (which we, again, guarantee throughout the definition of h).

Fact 5. *For all cut vertices $x^* \in C$ with $h(x^*)$ defined we have that $h(x^*)h(z)$ is an edge of $\mathbb{G}$, where z is the parent of x^*. Moreover if $x^* \notin \{x_0\} \cup (C \cap \mathrm{N}_T(x_0^*))$, we have that $h(x^*) = h(z')$, where z' is the parent of z.*

Moreover, the definition of $h(y)$ is consistent with (8), i.e. there is a path of even length in $\mathbb{G}$ between $h(y)$ and $\psi(y)$ for all children $y \notin C$ of x^*. Accordingly we can again define h for the vertices in the shrub of y as before, using this path.

This finishes the description of the definition of h. It remains to verify that h is a homomorphism and satisfies Condition 1. For the first part it suffices to check that for any $y \in V(T) \setminus \{x_0^*\}$ with parent x we have $h(y) \in \mathrm{N}_{\mathbb{G}}(h(x))$. If y is a vertex in some shrub then Facts 21 and 22 imply that $h(x)h(y)$ is an edge of $\mathbb{G}$. If y is a cut-vertex, on the other hand, Fact 5 implies that $h(x)h(y)$ is an edge of $\mathbb{G}$. So h is a homomorphism.

Further, by Fact 21 and 2 we get for all vertices x of T that all children $x' \notin C$ of x have the same $h(x')$. By Fact 5, if $x \neq x_0^*$ then all children $x' \in C$ of x and the parent z of x have the same $h(x') = h(z')$. Together wit Fact 3, this implies Property 1. □

Now we are ready to prove Lemma 14.

Proof of Lemma 14. Given $\varepsilon, \mu > 0$ with $\varepsilon \leq \mu/10$ and $k \in \mathbb{N}$ we set α, n_0 and an auxilliary constant $\beta > 0$ such that

$$(9) \qquad \alpha \cdot (2k+1) = \tfrac{1}{2}, \qquad \beta = \varepsilon\mu/(500k^3), \quad \text{and} \quad n_0 = (1500k/(\varepsilon\mu))^4.$$

Let $\mathbb{G}$ be a graph of order k that has an odd connected matching $\mathbb{M}$ of size at least m or a fork system $\mathbb{F}$ of size at least f and ratio r. Let T be a tree satisfying the respective conditions of Case 1 or 6 and let V_1 and V_2 denote the two partition classes of T with $t_1 = |V_1| \geq |V_2| = t_2$. We first construct an S-cut C for T with $S := \beta n \leq \varepsilon \frac{n}{k}$. Lemma 15 asserts that there is such a cut C with

$$(10) \qquad |C| \leq \frac{|V(T)|}{S} \leq \frac{(1-\mu)2k\frac{n}{k}}{\beta n} \overset{(9)}{\leq} \frac{1000k^3}{\varepsilon\mu} \overset{(9)}{\leq} \varepsilon \frac{n}{k}.$$

Let $T_1, \ldots, T_s$ be the shrubs of T corresponding to the cut C. We now distinguish whether we are in Case 1 or 6 of the lemma. In both cases we will construct a mapping ψ that is a homomorphism from $T - C$ to either $\mathbb{M}$ or $\mathbb{F}$ and satisfies the walk condition. After this case distinction the mapping ψ will serve as input for Lemma 20 which we then use to finish this proof.

Case 1: In this case we apply Lemma 16 in order to obtain an assignment of the shrubs to matching edges of $\mathbb{M}$ as follows. Set $a_{i,j} := |V(T_i) \cap V_j|$ for all $i \in [s]$, $j \in [2]$. This implies that $\sum_{i,j} a_{i,j} \le |V(T)| \le t = (1-\mu)2m\frac{n}{k}$ and, because C is an S-cut, that $a_{i,1} + a_{i,2} \le S$ for all $i \in [s]$. Accordingly Lemma 16 produces a mapping $\phi : [s]\times[2] \to V(\mathbb{M})$ satisfying $\phi(i,1)\phi(i,2) \in \mathbb{M}$ and (5).

We now use ϕ to construct a mapping $\psi\colon T \setminus C \to V(\mathbb{M})$. Set $\psi(v) := \phi(a_{i,j})$ for all $v \in V(T_i) \cap V_j$. Note that this definition together with (5) gives

$$|\psi^{-1}(\ell)| \le \frac{t}{2m} + 2S \le (1-\mu)\frac{n}{k} + 2\beta n \tag{11}$$

for all vertices ℓ of $\mathbb{M}$. Each edge of $T - C$ lies in some shrub T_i, $i \in [s]$ and as the mapping ϕ sends each shrub T_i to an edge of $\mathbb{M}$, the mapping ψ is a homomorphism from $T - C$ to $\mathbb{M}$. Moreover, as $\mathbb{M}$ is an odd connected matching, for any pairs of vertices $\ell, \ell' \in V(\mathbb{M})$ there is as well an even as also an odd walk in $\mathbb{G}$ between ℓ and ℓ'. Thus ψ satisfies the walk condition.

Case 6: In this case we apply Lemma 17 in order to obtain an assignment of the shrubs corresponding to C to edges of $\mathbb{F}$. For this application we use parameters $t_1 = |V_1|$, $t_2 = |V_2|$ and $a_{i,j} := |V(T_i) \cap V_j|$ for all $i \in [s]$, $j \in [2]$. It follows that $\sum_i a_{i,1} = t_1$ and $\sum_i a_{i,2} = t_2$. Because C is an S-cut, we further have $a_{i,1} + a_{i,2} \le S$ for all $i \in [s]$. Accordingly Lemma 17 produces a mapping $\phi : [s] \times [2] \to V(\mathbb{F})$ satisfying $\phi(i,1)\phi(i,2) \in \mathbb{F}$ and (6).

Again, we use ϕ to construct the mapping $\psi\colon T \setminus C \to V(\mathbb{F})$ by setting $\psi(v) := \phi(a_{i,j})$ for all $v \in V(T_i) \cap V_j$. By assumption we have $t_1 \le t' = (1-\mu)f\frac{n}{k}$ and $t_2 \le \frac{t'}{r} = (1-\mu)f\frac{n}{rk}$ and hence $t_1 + t_2 \le (1-\mu)f\frac{n}{k}(1+\frac{1}{r})$. Together with (6) this implies for all vertices $\ell_1 \in V_1(\mathbb{F})$ and $\ell_2 \in V_2(\mathbb{F})$ that

$$\begin{aligned} |\psi^{-1}(\ell_1)| &\le \frac{(1-\mu)f\frac{n}{k}}{f} + \sqrt{12(1-\mu)f\tfrac{n}{k}(1+\tfrac{1}{r})Sf} \\ &\le (1-\mu)\tfrac{n}{k} + 2fn\sqrt{6\beta/k}\,, \end{aligned} \tag{12}$$

and similarly

$$|\psi^{-1}(\ell_2)| \le \frac{r(1-\mu)f\frac{n}{rk}}{f} + 2fn\sqrt{6\beta/k} \le (1-\mu)\frac{n}{k} + 2fn\sqrt{6\beta/k}\,. \tag{13}$$

Putting (12) and (13) together, we conclude for any $\ell \in V(\mathbb{F})$ that

$$|\psi^{-1}(\ell)| \le (1-\mu)\tfrac{n}{k} + 2fn\sqrt{6\beta/k} \le (1-\mu)\tfrac{n}{k} + 2n\sqrt{6\beta k}\,. \tag{14}$$

As before it is easy to see that the mapping ψ is a homomorphism from $T-C$ to $\mathbb{F}$. Moreover, as $\mathbb{F}$ is a fork system, there is an even walk between any two vertices $\ell,\ell'\in V_1(\mathbb{F})$ and between any two vertices $\ell,\ell'\in V_2(\mathbb{F})$. Because ψ maps vertices of $V_1(T)$ to $V_1(\mathbb{F})$ and vertices of $V_2(T)$ to vertices of $V_2(\mathbb{F})$, the mapping ψ also satisfies the walk condition in this case.

Applying Lemma 20: In both Cases 1 and 6 we now apply Lemma 20 in order to transform ψ into a homomorphism from the whole tree T to $\mathbb{G}$. With input T, $\Delta:=n^\alpha$, C, $\mathbb{G}$, and ψ this lemma produces a homomorphism $h:V(T)\to V(\mathbb{G})$ satisfying 1 and 2. We claim that h is the desired $(\mu/2,(1-\varepsilon)\frac{n}{k})$-valid assignment.

Indeed, h is a homomorphism and so we have Condition 1 of Definition 11. Condition 2 follows from 1. To check Condition 3 let ℓ be any vertex of $\mathbb{G}$. We need to verify that $|h^{-1}(\ell)|\le(1-\frac{1}{2}\mu)(1-\varepsilon)\frac{n}{k}$. By 2 we have $|h^{-1}(\ell)|\le|\psi^{-1}(\ell)|+3|C|\Delta^{2k+1}$. Because $|C|\le 1000k/(\varepsilon\mu)$ by (10) and $\Delta^{2k+1}=n^{\alpha\cdot(2k+1)}=\sqrt{n}$ by (9) we infer that

$$\begin{aligned}|h^{-1}(\ell)|&\le|\psi^{-1}(\ell)|+\frac{3000k}{\varepsilon\mu}\sqrt{n}\overset{(9)}{\le}|\psi^{-1}(\ell)|+\beta n\\&\overset{(11),(14)}{\le}(1-\mu)\tfrac{n}{k}+\max\left\{2\beta n,2n\sqrt{6\beta k}\right\}+\beta n\overset{(9)}{\le}(1-\tfrac{1}{2}\mu)(1-\varepsilon)\tfrac{n}{k},\end{aligned}$$

where in the last inequality we use that $\varepsilon\le\mu/10$. □

6. Proof of the main embedding lemma

Our proof of Lemma 13 uses a greedy stragety for embedding the vertices of a tree with valid assignment into the given host graph.

Proof of Lemma 13. Let $V_0\dot\cup V_1\dot\cup\ldots\dot\cup V_k$ be an (ε,d)-regular partition of G with reduced graph $\mathbb{G}$ and let T be a tree with $\Delta(T)\le\Delta$ and with a $(\varrho,(1-\varepsilon)\frac{n}{k})$-valid assignment h to $\mathbb{G}$. Further, let C be an S-cut of T, let $T_1,\ldots,T_s$ be the shrubs of T corresponding to C, and assume that

$$(\tfrac{1}{10}d\varrho-10\varepsilon)\tfrac{n}{k}\ge|C|+S+\Delta\,. \tag{15}$$

As last preparation we arbitrarily divide each cluster $V_i=V_i'\dot\cup V_i^*$ into a set V_i' of size $(1-\frac{1}{2}\varrho)|V_i|$, which we will call *embedding space*, and the set of remaining vertices V_i^*, the so-called *connecting space*. Next we will first specify the order in which we embed the vertices of T into G, then describe the actual embedding procedure, and finally justify the correctness of this procedure.

Pick an arbitrary vertex $x_1^* \in C$ as root of T and order the cut vertices $C = \{x_1^*, \dots, x_c^*\}$, $c = |C|$ in such a way that on each $x_1^* - x_i^*$-path in T there are no x_j^* with $j > i$. Similarly, for each $i \in [s]$ let $t(i)$ denote the number of vertices in the shrub T_i and order the vertices $y_1, \dots, y_{t(i)}$ of T_i such that all paths in T_i starting at the root of T_i have solely ascending labels. For embedding T into G we process the cut vertices and shrubs according to these orderings, more precisely we first embed x_1^*, then all shrubs T_i that have x_1^* as parent, one after the other. For embedding T_i we embed its vertices in the order $y_1, \dots, y_{t(i)}$ defined above. Then we embed the next cut vertex x_2^* (which is a child of one of the shrubs embedded already or of x_1^*), then all child shrubs of x_2^*, and so on. Let $x_1, \dots, x_{|V(T)|}$ be the corresponding ordering of $V(T)$.

Before turning to the embedding procedure itself, observe that Property 2 of Definition 11 asserts the following fact. For a vertex x_j of T and for $i \in [k]$ let $\mathrm{N}_i(x_j)$ be the set of neighbours $x_{j'}$ of x_j in T with $j' > j$ and $h(x_{j'}) = i$.

Fact 1. *For all vertices x_j of T at most two sets $\mathrm{N}_i(x_j)$ are non-empty.*

The idea for embedding T into G is as follows. We equip each vertex $x \in V(T)$ with a *candidate set* $V(x) \subseteq V_{h(x)}$ and from which x will choose its image in G. To start with, we set $V(x^*) := V^*_{h(x^*)}$ for all vertices $x^* \in C$ and $V(x) := V'_{h(x)}$ for all other vertices x. Cut vertices will be embedded to vertices in a connecting space and non-cut vertices to vertices in an embedding space. Then we will process the vertices of T in the order $x_1, \dots, x_{|V(T)|}$ defined above and embed them one by one. Whenever we embed a cut vertex x^* to a vertex v in this procedure we will set up so-called *reservoir sets* $R_i \subseteq V_i \cap \mathrm{N}_G(v)$ for all (at most two) clusters V_i such that some child x of x^* is assigned to V_i, i.e., $h(x) = i$. Reservoir sets will be used for embedding the children of cut vertices. We (temporarily) remove the vertices in these reservoir sets from all other candidate sets but put them back after processing all child shrubs of x^*. This will ensure that we have enough space for embedding children of x^*, even after possibly embedding $\Delta - 1$ child shrubs of x^*.

Now let us provide the details of the embedding procedure. Throughout, x^* will denote the cut vertex whose child-shrubs are currently processed. The set U will denote the vertices in G used so far; thus initialize this set to $U := \emptyset$. As indicated above, initialize $V(x^*) := V^*_{h(x^*)}$ for all vertices $x^* \in C$ and $V(x) := V'_{h(x)}$ for $x \in V(T) \setminus C$, and set $R_i := \emptyset$ for all $i \in [k]$. For constructing an embedding $f\colon V(T) \to V(G)$ of T into G, repeat the following steps:

(1) Pick the next vertex x from $x_1, \dots, x_{|V(T)|}$.
(2) Choose a vertex $v \in V(x) \setminus U$ that is typical with respect to $V(y) \setminus U$ for all unembedded $y \in \mathrm{N}_T(x)$, set $f(x) = v$, and $U := U \cup \{v\}$.
(3) For all unembedded $y \in \mathrm{N}_T(x)$ set $V(y) := (V(y) \setminus U) \cap \mathrm{N}_G(v)$.
(4) If $x \in C$ then set $x^* := x$. Further, for all i with $\mathrm{N}_i(x) \setminus C \neq \emptyset$ arbitrarily choose a reservoir set $R_i \subseteq (V_i' \setminus U) \cap \mathrm{N}_G(v)$ of size $5\varepsilon\frac{n}{k} + \Delta$, set $V(y) := R_i$ for all $y \in \mathrm{N}_i(x) \setminus C$, and (temporarily) remove R_i from all other candidate sets in V_i', i.e., set $V(y') := V(y') \setminus R_i$ for all $y' \in V(T) \setminus \mathrm{N}_i(x)$.
(5) After the vertices of all child shrubs of x^* are embedded put the vertices in R_i back to all candidate sets in V_i' for all $i \in [k]$, i.e., $V(y) := V(y) \cup R_i$ for all $y \in V(T) \setminus C$ with $h(y) = i$, and set $R_i := \emptyset$.

Steps 3 and 4 of this procedure guarantee for each vertex y with embedded parent x that the candidate set $V(y)$ is contained in $N_G(f(x))$. Accordingly, if we can argue that in Step 2 we can always choose an image v of x in $V(x)$ (and that we can choose the reservoir sets in Step 4) we indeed obtain an embedding f of T into G. To show this we first collect some observations that will be usefull in the following analysis. The order of $V(T)$ guarantees that all child shrubs of a cut vertex are embedded before the next cut vertex. Notice that this implies the following fact (cf. Step 4 and Step 5).

Fact 2. *For all $i \in [k]$, at any point in the procedure, the reservoir set R_i satisfies $|R_i| = 5\varepsilon\frac{n}{k} + \Delta$ if there is a neighbour x of the current cut-vertex x^* such that $h(x) = i$ and $|R_i| = 0$ otherwise. In addition no reservoir set gets changed before all child shrubs of x^* are embedded.*

Further, since h is a $(\varrho, (1-\varepsilon)\frac{n}{k})$-valid assignment and only cut-vertices are embedded into connecting spaces V_i^*, we always have

$$|V_i' \cap U| \leq (1 - \tfrac{1}{2}\varrho)\tfrac{n}{k} \quad \text{and} \quad |V_i^* \cap U| \leq |C| \quad \text{for all } i \in [k]\,. \tag{16}$$

Now we check that Steps 2 and 4 can always be performed. To this end consider any iteration of the embedding procedure and suppose we are processing vertex x. We distinguish three cases.

Case 1: Assume that x is a cut-vertex. Then we had $V(x) = V_{h(x)}^*$ until the parent x' of x got embedded. In the iteration when x' got embedded then the set $V(x)$ shrunk to a set of size at least $(d-\varepsilon)|V_{h(x)}^* \setminus U|$ in Step 3 because $f(x')$ is typical with respect to $V_{h(x)}^* \setminus U$. No vertices embedded between x' and x (except for possible vertices in C) alter $V(x)$, and so by (16) we have

$$|V(x)\setminus U| \geq (d-\varepsilon)|V^*_{h(x)}| - |C| \geq (d-\varepsilon)\tfrac{1}{2}\varrho\tfrac{n}{k} - |C| \overset{(15)}{>} 4\varepsilon\tfrac{n}{k}$$

when we are about to choose $f(x)$. By Fact 1 at most two of the sets $\mathrm{N}_i(x)$ are non-empty and each of these two sets can contain cut vertices y^* and non-cut vertices y'. We clearly have $V(y^*) = V^*_i$ and $V(y') = V'_i$ and so there are at most 4 different sets $V(y)\setminus U$, each of size at least $\frac{1}{2}\varrho\frac{n}{k} - |C| > \varepsilon\frac{n}{k}$ by (16) and (15), with respect to which we need to choose a typical $f(x)$. By Lemma 2 there are less than $4\varepsilon\frac{n}{k}$ vertices in $V(x)\setminus U$ (which is a subset of V_i) that do not fulfil this requirement. Hence we can choose $f(x)$ whenever $x \in C$. In addition, we can choose the reservoir sets in Step 4 of this iteration: Indeed, let i be such that $\mathrm{N}_i(x)\setminus C \neq \emptyset$ and let $y \in \mathrm{N}_i(x)\setminus C$ be a neighbour of x we wish to embed to V_i. In Step 2, when we choose $f(x)$, then $V(y) = V'_i$ and so $f(x)$ is typical with respect to $V'_i\setminus U$. By Lemma 2 and (16) we thus have in Step 3 of this iteration that

$$|(V'_i\setminus U)\cap \mathrm{N}_G(v)| \geq (d-\varepsilon)|V'_i\setminus U| \geq (d-\varepsilon)\tfrac{1}{2}\varrho\tfrac{n}{k} \overset{(15)}{\geq} 5\varepsilon\tfrac{n}{k} + \Delta.$$

Therefore we can choose R_i in Step 4.

Case 2: Assume that x is not in C but the child of a cut vertex x^*. Then $V(x) = R_{h(x)}$ before x gets embedded. Moreover, due to Step 4, R_i has been removed from all candidate sets besides those of the at most Δ neighbours of x^*. By Fact 2 we have $|R_{h(x)}| = 5\varepsilon\frac{n}{k} + \Delta$ and so we conclude that $|V(x)\setminus U| \geq 5\varepsilon\frac{n}{k} > 4\varepsilon\frac{n}{k}$. As in the previous case, there are at most four different sets $V(y)\setminus U$ for unembedded neighbours y of x, each of size at least $\frac{1}{2}\varrho\frac{n}{k} - |R_{h(y)}| = \frac{1}{2}\varrho\frac{n}{k} - 5\varepsilon\frac{n}{k} - \Delta \geq \varepsilon\frac{n}{k}$ by (15) and (16). Thus Lemma 2 guarantees that there is $v \in V(x)\setminus U$ which is typical with respect to all these sets $V(y)\setminus U$ and hence we can choose $f(x)$ in this case.

Case 3: As third and last case, let x be a vertex of some shrub T_j which is the child of a (non-cut) vertex x' of T_j. Until x' got embedded we had $V(x) = V'_{h(x)}\setminus R_{h(x)}$ and so, $v' = f(x')$ was chosen typical with respect to $V'_{h(x)}\setminus(R_{h(x)}\cup U)$ where U is the set of used vertices in G at the time when x' got embedded. In the corresponding iteration $V(x)$ shrunk to $(V'_{h(x)}\setminus(R_{h(x)}\cup U))\cap \mathrm{N}_G(v')$. This together with (16) implies that immediately after this shrinking we had

$$|V(x)\setminus U| \geq (d-\varepsilon)(\tfrac{1}{2}\varrho\tfrac{n}{k} - |R_{h(x)}|) \geq (d-\varepsilon)(\tfrac{1}{2}\varrho\tfrac{n}{k} - 5\varepsilon\tfrac{n}{k} - \Delta) \overset{(15)}{>} 4\varepsilon\tfrac{n}{k} + |T_j|.$$

By construction only vertices from T_j come between x' and x in the order of $V(T)$ and so when we want to embed x in the procedure above we still have

$|V(x) \setminus U| > 4\varepsilon\frac{n}{k}$ where U now is the set of vertices used until the embedding of x. Similarly as in the other two cases there are at most four different types of candidate sets for non-embedded neighbours of x, all of these have more than $\varepsilon\frac{n}{k}$ vertices and so Lemma 2 allows us to choose an $f(x) \in V(x) \setminus U$ typical with respect to these sets. This concludes the case distinction and hence the proof of correctness of our embedding procedure. □

7. Coloured tripartite graphs are either good or odd

7.1. **Some tools.** In this section we collect some simple but useful propositions. We start with two observations about matchings in η-complete graphs. The first one states that a bipartite η-complete coloured graph contains a reasonably big matching in one of the two colours.

Proposition 21. *Let K be a coloured graph on n vertices and let D and D' be vertex sets of size at least m in K. If $K[D, D']$ is η-complete then $K[D, D']$ contains a matching M either in red or in green of size at least $\frac{m}{2} - \eta n$.*

Proof. Assume without loss of generality that $|D| \leq |D'|$. Colour a vertex $v \in D$ with red if it has more red-neighbours than green-neighbours in $K[D, D']$ and with green otherwise. By the pigeon-hole principle there is a set $X \subseteq D$ of size $\frac{1}{2}|D|$ such that all vertices in X have the same colour, say red. But then each vertex in X has at least $\frac{1}{2}|D'| - \eta n \geq |X| - \eta n$ red-neighbours in D'. Accordingly we can greedily construct a red matching of size at least $|X| - \eta n \geq \frac{m}{2} - \eta n$ between X and D'. □

The next proposition gives a sufficient condition for the existence of an almost perfect matching in a subgraph of $K \in \mathcal{K}_n^\eta$.

Proposition 22. *Let $K \in \mathcal{K}_n^\eta$ have partition classes A, B, and C and let $A' \subseteq A$, $B' \subseteq B$, $C' \subseteq C$ with $|A'| \geq |B'| \geq |C'|$. If $|A'| \leq |B' \cup C'|$ then there is a matching in $K[A', B', C']$ covering at least $|A' \cup B' \cup C'| - 4\eta n - 1$ vertices.*

Proof. Let $x := |B'| - |C'|$ and $y := \lfloor \frac{1}{2}(|A'| - x) \rfloor$. Observe that $x \leq |B'| \leq |A'|$. Hence $y \geq 0$, $x + y \leq \frac{1}{2}(|A'| + x) \leq \frac{1}{2}(|B' \cup C'| + x) = |B'|$, and $y \leq \frac{1}{2}(|A'| - x) \leq \frac{1}{2}(|B' \cup C'| - x) = |C'|$. Choose arbitrary subsets $U_B \subseteq B'$ of size $x + y$, $U_C \subseteq C'$ of size y, set $U := U_B \cup U_C$, $W := B' \setminus U_B$ and $W' := C' \setminus U_C$. Clearly $|W'| = |C'| - y = |B'| - (x + y) = |W|$ and $|A'| - 1 \leq x + 2y = |U| \leq |A'|$. Thus we can choose a subset U' of A' of size $|U|$ that covers all but at most 1 vertex of A' and so that $K[U, U']$ and

$K[W, W']$ are η-complete balanced bipartite subgraphs. A simple greedy algorithm allows us then to find matchings of size at least $|U| - \eta n$ and $|W| - \eta n$ in $K[U, U']$ and $K[W, W']$, respectively. These matchings together form a matching in $K[A', B', C']$ covering at least $|U \cup U' \cup W \cup W'| - 4\eta n \geq |A' \cup B' \cup C'| - 4\eta n - 1$ vertices. $\square$

The following proposition shows that induced subgraphs of η-complete tripartite graphs are connected provided that they are not too small. Moreover, subgraphs that substantially intersect all three partition classes contain a triangle.

Proposition 23. *Let $K \in \mathcal{K}_n^\eta$ be a graph with partition classes A, B, C, and let $A' \subseteq A$, $B' \subseteq B$, $C' \subseteq C$.*

(1) *If $|A'| > 2\eta n$ then every pair of vertices in $B' \cup C'$ has a common neighbour in A'.*
(2) *If $|A'|, |B'| > 2\eta n$ then $K[A', B']$ is connected.*
(3) *If $|A'|, |B'|, |C'| > 2\eta n$ then $K[A', B', C']$ contains a triangle.*

Proof. As $K \in \mathcal{K}_n^\eta$, each vertex in $B' \cup C'$ is adjacent to at least $|A'| - \eta n > |A'|/2$ vertices in A'. Thus every pair of vertices in B' has a common neighbour in A' which gives 1. For the proof of 2 observe that by 1 every pair of vertices in B' has a common neighbour in A'. Since the same holds for pairs of vertices in A' the graph $K[A', B']$ is connected. To see 3 we use 1 again and infer that every pair of vertices in $A' \times B'$ has a common neighbour in C'. As $|A'|, |B'| > 2\eta n$ there is some edge in $A' \times B'$ and thus there is a triangle in $K[A', B', C']$. $\square$

Similar in spirit to 3 of Proposition 23 we can enforce a copy of a cycle of length 5 in a system of η-complete graphs as we show in the next proposition.

Proposition 24. *Let K be a coloured graph on n vertices, let c be a colour, vw be a c-coloured edge of K, and let $D_1, D_2, D_3 \subseteq V(K)$ such that all graphs $K[v, D_1]$, $K[D_1, D_2]$, $K[D_2, D_3]$, and $K[D_3, w]$ are (η, c)-complete bipartite graphs. Set $D := \bigcup_{i \in [3]} D_i \cup \{v, w\}$. If $|D_i| > 2\eta n + 2$ for all $i \in [3]$ then $K[D]$ contains a c-coloured copy of C_5.*

Proof. By Proposition 231 every pair of vertices in $D_1 \cup D_3$ is connected by a path of colour c and length 2 with center in $D_2 \setminus \{v, w\}$. Moreover, v has at least $|D_1| - \eta n \geq 1$ neighbours in D_1 and similarly w has a neighbour in D_3. Hence there is a c-coloured C_5 in $K[D]$. $\square$

7.2. **Non-extremal configurations.** In the proof of Lemma 9 we will use that coloured graphs K from $\mathcal{K}_n^\eta$ have the following property P. Either one colour of K has a big odd connected matching or both colours have big connected matchings whose components are bipartite. Analysing these bipartite configurations will then lead to a proof of Lemma 9. Property P is a consequence of the next lemma, Lemma 25, which states that if all connected matchings in a colour of K are small then the other colour has bigger odd connected matchings.

Lemma 25 (improving lemma)**.** *For every $\eta' > 0$ there are $\eta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Suppose that a coloured graph $K \in \mathcal{K}_n^\eta$ is neither η'-extremal nor $\frac{3}{4}(1-\eta')n$-odd. Let M be a maximum connected matching in K of colour c. If $\eta' n < |M| < (1-\eta')n$ then K has an odd connected matching M' in the other colour satisfying $|M'| > |M|$.*

Proof. Given η' define $\tilde{\eta} := \eta'/3$ and let η be small enough and n_0 large enough such that $(\frac{1}{100}\eta' - 5\eta)n_0 > 1$ (and hence $\eta < \frac{1}{500}\eta'$). For $n \geq n_0$ let $K = (A\dot\cup B\dot\cup C, E)$ be a coloured graph from $\mathcal{K}_n^\eta$ with partition classes A, B, and C that is neither η'-extremal nor $(1-\eta')3n/4$-odd. Suppose $c =$ green and hence that K has a maximum green connected matching M with $\eta' n < |M| < (1-\eta')n$. For $D, D' \in \{A, B, C\}$ with $D \neq D'$ let $M_{DD'} := M \cap (D \times D')$. We call the $M_{DD'}$ the *blocks* of M and say that a block $M_{DD'}$ is *substantial* if $|M_{DD'}| \geq \tilde{\eta} n$. Let R be the set of vertices in K not covered by M. For $D \in \{A, B, C\}$ let $R_D := R \cap D$.

Fact 1. *We have $|R_A| - |M_{BC}| = |R_B| - |M_{CA}| = |R_C| - |M_{AB}| > \eta' n$.*

Indeed, $|R_A| + |M_{AB}| + |M_{AC}| = |R_B| + |M_{AB}| + |M_{BC}|$ and hence $|R_A| - |M_{BC}| = |R_B| - |M_{AC}| = |R_B| - |M_{CA}|$ which proves the first part of this fact. For the second part observe that $|R_A| + 2|M_{AB}| + |R_B| + 2|M_{BC}| + |R_C| + 2|M_{CA}| = 3n$. Hence we conclude from $|M| = |M_{AB}| + |M_{BC}| + |M_{CA}| < (1-\eta')n$ that

$$\begin{aligned} 3(|R_A| - |M_{BC}|) &= (|R_A| - |M_{BC}|) + (|R_B| - |M_{CA}|) + (|R_C| - |M_{AB}|) \\ &= 3(n - |M_{AB}| - |M_{BC}| - |M_{CA}|) > 3\eta' n. \end{aligned}$$

This finished the proof of Fact 1.

In the remainder we assume without loss of generality that $|R_A| \geq |R_B| \geq |R_C|$. By Fact 1 this implies that $|M_{BC}| \geq \frac{1}{3}\eta' n$ since $|M| > \eta' n$ and hence M_{BC} is substantial. Our next main goal is to find a connected matching in red that is bigger than M. For achieving this goal the following fact about red connections between vertices of R will turn out useful.

Fact 2. *There is a vertex $u^* \in R_A$ such that $R - u^*$ is red connected.*

To see this, assume first that there is a vertex $u^* \in R_A$ that has more than $4\eta n$ green-neighbours in M_{BC}. Then more than $2\eta n$ of these neighbours are in, say, $M_{BC} \cap B$. Call this set of vertices B^*. Now let $u \neq u^*$ be any vertex in $R \setminus C$. By the maximality of M the vertex u has no green-neighbours in $M(B^*)$. This implies that u has at least $|M(B^*)| - \eta n > |M(B^*)|/2$ red-neighbours in $M(B^*)$. Thus any two vertices in $R \setminus C$ have a common red-neighbour in $M(B^*)$. A vertex $u \in R_C$ on the other hand has at least $|R_A| - \eta n \geq |M_{BC}| + \eta' n - \eta n > 2\eta n + 1$ neighbours in R_A where the first inequality follows from Fact 1. If at least 2 of these neighbours are red then u is red connected to $R_A - u^*$. Otherwise u has a set U of more than $2\eta n$ green-neighbours in $R_A - u^*$. But then, by the maximality of M, the graph $K[U, M_{BC} \cap B]$ is red. Since $|M_{BC}| \geq \eta' n > \eta n$ the vertex u has a neighbour v in $M_{BC} \cap B$. Since u has a green-neighbour in R_A it follows from the maximality of M that uv is red. Thus u is red connected to U and therefore to all vertices of $(R \setminus C) - u^*$.

If there is no vertex in R_A with more than $4\eta n$ green-neighbours in M_{BC} on the other hand, then any two vertices in R_A obviously have at least $|M_{BC}| - 4\eta n - 2\eta n \geq \frac{1}{3}\eta' n - 6\eta n > 0$ common red-neighbours in $B \cap M_{BC}$. Moreover, by the maximality of M, each vertex $v \in R_C \cup R_B$ is either red connected to R_A or it has only red-neighbours in M_{BC}. Thus v has a common red-neighbour with any vertex in R_A which proves Fact 2 also in this case.

Fact 3. *K has a red connected matching M' with $|M'| \geq |M| + \frac{1}{4}\eta' n$.*

Let uv be an arbitrary edge in M_{BC}. Then, by the maximality of M, one vertex of this edge, say u, has at most one green-neighbour in R_A. By Fact 1 we have $|R_A| \geq |M_{BC}| + \eta' n$ and since u has at most $\eta n < \eta' n$ non-neighbours in R_A it follows that u has at least $|M_{BC}| + 1$ red-neighbours in R_A. Thus, a simple greedy method allows us to construct a red matching M'_{BC} of size $|M_{BC}|$ between $R_A - u^*$ and such vertices u of matching edges in M_{BC}. Let R'_A be the set of vertices in R_A not covered by M'_{BC}. We repeat this process with M_{AC} and M_{AB}, respectively, to obtain red matchings M'_{AC} and M'_{AB} and sets R'_B and R'_C.

By maximality of M, for each vertex $w \in R'_A$ the following is true: either w has no green-neighbour in M_{BC}, or w has no green-neighbour in R'_B. Moreover w has at most ηn non-neighbours. Observe that $|R'_B|, |R'_A| > \eta' n$ by Fact 1 and the set X of vertices in M_{BC} that are not covered by M'_{BC} has size at least $\frac{1}{3}\eta' n$ since $|M_{BC}| = |M'_{BC}| \geq \frac{1}{3}\eta' n$ and each edge of

M'_{BC} uses exactly one vertex from M_{BC}. This implies that we can again use a greedy method to construct a red matching M'_R with edges from $(R'_A - u^*) \times (R'_B \cup X)$ of size at least $\frac{1}{3}\eta' n - \eta n - 1 \geq \frac{1}{4}\eta' n$. Hence we obtain a red matching $M' := M'_{BC} \dot\cup M'_{CA} \dot\cup M'_{AB} \dot\cup M'_R$ of size at least $|M| + \frac{1}{4}\eta' n$. For establishing Fact 3 it remains to show that M' is red connected. This follows from Fact 2 since each edge of M' intersects $R - u^*$.

If the matching M' is odd then the proof of Lemma 25 is complete. Hence assume in the remainder that M' is even. Since M' intersects $R - u^*$ this together with Fact 2 immediately implies the next fact. For simplifying the statement as well as the following arguments we will first delete the vertex u^* from K (and let K denote the resulting graph from now on).

Fact 4. *No odd red cycle in K contains a vertex of R.*

Fact 8 below uses this observation to conclude that K is extremal, contradicting the hypothesis of Lemma 25. To prepare the proof of this fact we first need some auxiliary observations.

Fact 5. *For $\{D, D', D''\} = \{A, B, C\}$, if $M_{DD'}$ is a substantial block then there is a vertex $v^* \in R_{D''}$ such that $K[M_{DD'}, R_{D''} - v^*]$ is red and $K[M_{DD'}]$ is green.*

We first establish the first part of the statement. We may assume that there are vertices $v^* \in R_{D''}$ and $v \in M_{DD'}$ such that $v^* v$ is green (otherwise we are done). Without loss of generality $v \in D$. Let $X = \mathrm{N}(v^*)$. Then, by the maximality of M, all edges between v^* and $X \cap R$ are red. By Fact 4 this implies that all edges between $X \cap R_D$ and $X \cap R_{D'}$ are green. Since $\min\{|X \cap R_D|, |X \cap R_{D'}|\} > \eta n$, this set of edges is not empty. We use the maximality of M to infer that all edges between $M_{DD'}$ and $X \cap (R_D \cup R_{D'})$ are red. Using Fact 4 this in turn implies that edges between $Y := M_{DD'} \cap X$ and v^* are green. By the maximality of M all edges between $M(Y)$ and $R_{D''} - v^*$ are consequently red. We claim that therefore $K[R_{D'} \cap X, R_{D''} - v^*]$ is green. Indeed, assume there was a red edge $ww' \in R_{D'} \cap X \times (R_{D''} - v^*)$. Then w and w' have at least $|M(Y) \cap D| - 2\eta n \geq |M_{DD'}| - 3\eta n \geq \tilde{\eta} n - 3\eta n > 0$ common neighbours w'' in $M(Y) \cap D$. Since edges between $M(Y)$ and $R_{D''} - v^*$ and edges between $M_{DD'}$ and $X \cap R_{D'}$ are red, so are the edges ww'' and $w'w''$ and thus we have a red triangle $ww'w''$ contradicting Fact 4. By Fact 1 we have $|R_{D'} \cap X| \geq \eta' n - \eta n > \eta n$ and so each vertex in $R_{D''} - v^*$ is connected by a green edge to some vertex in $R_{D'} \cap X$. The maximality of M implies that $K[M_{DD'}, R_{D''} - v^*]$ is red as required. For the second part of Fact 5 observe that the fact that $K[M_{DD'}, R_{D''} - v^*]$ is red and

$|R_{D''}| \geq \eta' n > 2\eta n + 1$ imply that each pair of vertices in $M_{DD'}$ has a common red neighbour in $R_{D''} - v^*$ and so by Fact 4 the graph $K[M_{DD'}]$ is green. This establishes Fact 5.

Now we also delete all (at most 3) vertices from R that play the rôle of v^* in Fact 5 (and again keep the names for the resulting sets).

Fact 6. *Suppose that $\{D, D', D''\} = \{A, B, C\}$ and that $M_{DD'}$ is a substantial block. Then for one of the sets D and D', say for D, the graph $K[M_{DD'}, R_D]$ is red and $K[R_{D''}, R_D]$ is green. For the other set D' the following is true. If $v \in R_{D'}$ then $K[v, M_{DD'}]$ and $K[v, R]$ are monochromatic, with distinct colours.*

We start with the first part of this fact and distinguish two cases. First, assume that there is a red edge ww' with $w \in R_{D''}$ and $w' \in R_{D'}$. We will show that in this case $K[M_{DD'}, R_D]$ is red and $K[R_{D''}, R_D]$ is green. Since $M_{DD'}$ is substantial, edges between w and $M_{DD'}$ are red by Fact 5 and hence, owing to Fact 4, edges between $M_{DD'} \cap \mathrm{N}(w)$ and w' are green. Since $K[M_{DD'}]$ is green by Fact 5, since M is maximal, and since each vertex in $M_{DD'} \cap D'$ has some neighbour in $M_{DD'} \cap \mathrm{N}(w')$ this implies that all edges between $M_{DD'}$ and R_D are red. Moreover, edges between $M_{DD'} \cap D'$ and $R_{D''}$ are red by Fact 5 and hence we conclude from Fact 4 that $K[R_{D''}, R_D]$ is green If, on the other hand, there is no red edge between $R_{D''}$ and $R_{D'}$ then the first part of the fact is true with D and D' interchanged: Clearly $K[R_{D''}, R_{D'}]$ is green and by maximality of M we infer that $K[M_{DD'}, R_{D'}]$ is red.

For the second part of the fact suppose that $K[M_{DD'}, R_D]$ is red and $K[R_{D''}, R_D]$ is green. Let $v \in R_{D'}$ and assume first that v has a green neighbour in $M_{DD'}$. The maximality of M then implies that $K[v, R]$ is red and since $K[R_{D''}, M_{DD'}]$ is also red (by Fact 5) we get that $K[v, M_{DD'}]$ is green. Hence it remains to consider the case that $K[v, M_{DD'}]$ is red. By Fact 5 the graph $K[R_{D''}, M_{DD'}]$ is red and so Fact 4 forces the graph $K[v, R_{D''}]$ to be green. To show that also $K[v, R_D]$ is green assume to the contrary that there is a red edge vw with $w \in R_D$. Recall that $K[v, M_{DD'} \cap D]$, $K[M_{DD'} \cap D, R_{D''}]$, $K[R_{D''}, M_{DD'} \cap D']$, and $K[M_{DD'} \cap D', w]$ are red (and clearly η-complete). Since $|M_{DD'} \cap D|, |R_{D''}|, |M_{DD'} \cap D'| \geq \tilde{\eta} n - 1 \geq 2\eta n + 2$ we can apply Proposition 24 to infer that there is a red C_5 touching R which contradicts Fact 4.

Fact 7. *If $M_{DD'}$ and $M_{D'D''}$ are substantial, then $K[M_{DD'}, M_{D'D''}]$ and $K[R_{D''}, R_D]$ are green and $K[M_{DD'} \cup M_{D'D''}, R_{D''} \cup R_D]$ is red. Moreover,* *if $v \in R_{D'}$ then $K[v, M_{DD'} \cup M_{D'D''}]$ and $K[v, R]$ are monochromatic, with distinct colours.*

By Fact 6 every vertex in $R_{D''} \cup R_D$ sends some green edges to R and hence the maximality of M implies that $K[M_{DD'} \cup M_{D''D'}, R_{D''} \cup R_D]$ is red. Since there is no red triangle touching R, the graphs $K[M_{DD'} \cap D, M_{D''D'} \cap D']$, $K[M_{DD'} \cap D', M_{D''D'} \cap D'']$, and $K[R_{D''}, R_D]$ are green. Using Proposition 24 we get similarly as before that also edges in $K[M_{DD'} \cap D, M_{D'D''} \cap D'']$ are green, since otherwise there was a red C_5 touching R. It remains to show the second part of Fact 7. By Fact 6 the graph $K[v, R]$ is monochromatic. Moreover, applying Fact 6 once to $M_{DD'}$ and once to $M_{D''D'}$, we obtain that $K[v, R]$ and $K[v, M_{DD'} \cup M_{D''D'}]$ are monochromatic graphs of distinct colours.

Now we have gathered enough structural information to show that K is extremal.

Fact 8. *K is in spider configuration with parameter $\tilde{\eta}$.*

We first argue that we can assume without loss of generality that

$$C \text{ always plays the rôle of } D' \text{ in Fact 6.} \tag{$*$}$$

Indeed, by Fact 7 this is the case if, besides M_{BC}, the block M_{AC} is substantial. If M_{AC} (and hence also M_{AB}) is not substantial on the other hand then it might be the case that B plays the rôle of D' in Fact 6. Then however we may delete at most $\tilde{\eta}n$ vertices from R_B in order to guarantee $|R_B| \le |R_C|$ and then the following argument still works with B and C interchanged.

To obtain the spider configuration set $A_1 := R_A$, $B_1 := R_B$, let C_1 be the set of those vertices $v \in R_C$ such that $K[v, M_{BC}]$ is red, let $C_C := R_C \setminus C_1$, and define $D_{D'} := M_{DD'} \cap D$ for all $D, D' \in \{A, B, C\}$ with $D \neq D'$. If any of the sets we just defined has less than $\tilde{\eta}n$ vertices delete all vertices in this set. Finally, define A_2, B_2, C_2 as in the definition of the spider configuration (Definition 7). Observe that this together with Fact 6 implies that $K[C_C, M_{BC}]$ is green and $K[C_C, R]$ is red.

Now let $\{X, Y, Z\} = \{A, B, C\}$ arbitrarily. Clearly we have $|X_1 \cup X_2| \ge (1 - 3\tilde{\eta})n \ge (1 - \eta')n$. Moreover $K[X_1, Y_2]$ is η-complete. We next verify that this graph is also red. We distinguish two cases. First assume that $Y \neq C$. In this case $X_1 \subseteq R_X$ and $Y_2 = Y_X \cup Y_Z \subseteq (M_{XY} \cap Y) \cup (M_{YZ} \cap Y)$. We have $Y_Z \neq \emptyset$ only if M_{YZ} is substantial and then Fact 5 implies that $K[R_X, M_{YZ}]$ is red. Similarly $Y_X \neq \emptyset$ only if M_{XY} is substantial. By $(*)$ Fact 6 implies that then $K[R_X, M_{XY}]$ is red if $X \neq C$. By the definition of

C_1 we also get that $K[X_1, M_{XY}]$ is red if $X = C$. Thus all edges between X_1 and Y_2 are red as desired. If $Y = C$ on the other hand then $X_1 \subseteq R_X$ and $Y_2 = Y_X \cup Y_Z \cup C_C \subseteq (M_{XY} \cap Y) \cup (M_{XZ} \cap Y) \cup C_C$. Analogous to the argument in the first case the graphs $K[R_X, M_{YZ}]$ and $K[R_X, M_{XY}]$ are red (since $X \neq C$). As noted above in addition all edges between R and C_C are red and so $K[X_1, Y_2]$ is also red in this case.

We finish the proof of Fact 8 (and hence Lemma 25) by checking that we have a spider configuration with colour $c =$ red. Observe that the graph $K[A_1 \cup B_1 \cup C_1, A_2 \cup B_2 \cup C_2]$ is connected and bipartite. We now verify Conditions 2a–2d of the spider configuration. For Condition 2b assume that $C_C \neq \emptyset$. Fact 7 and the definition of C_2 imply then that M_{AB} is not substantial and hence $|A_B| = 0$. Moreover, since $|R_A| \geq |R_B| \geq |R_C|$ we get the first part of Condition 2a, and $|D'_D| = |D_{D'}|$ is clearly true by definition. By Fact 1 we have $n - |M_{D''D} \cup M_{D''D'}| = |R_{D''}| > |M_{DD'}|$ which implies $n - |D''_2| > |D_{D'}|$ unless $D'' = C$ and $C_C \neq \emptyset$ (if $D'' \neq C$ or $C_C = \emptyset$ then $|M_{DD'}| = |D_{D'}|$). And if $C_C \neq \emptyset$ Condition 2b implies $|D_{D'}| = |A_B| = 0$ and thus we also get $n - |D''_2| > |D_{D'}|$ in this case. This establishes Condition 2a. To see Condition 2c, note that if A_2 is non-empty then either M_{AB} or M_{AC} are substantial. Since in addition M_{BC} is substantial by assumption we conclude from Fact 7 that there is a green triangle connected to M_{BC} and hence to the green matching M. As K is not $\frac{3}{4}(1-\eta')n$-odd this implies $\frac{1}{2}|A_B \cup A_C \cup B_A \cup B_C \cup C_A \cup C_B| \leq |M| < \frac{3}{4}(1-\eta')n$. It remains to verify Condition 2d. Assume, for a contradiction, that $C_1 \neq \emptyset$ and $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$ and $|B_1 \cup C_1| > (1-\eta)\frac{3}{4}n$. As $|R_A| \geq |R_B| \geq |R_C| \geq |C_1|$ and $C_1 \neq \emptyset$ all these sets have size at least $\tilde{\eta}n$ and so $A_1 = R_A$, $B_1 = R_B$ and $C_1 \subseteq R_C$. By Fact 6 and the definition of C_1 the graph $K[A_1, B_1, C_1]$ is (η, green)-complete and thus contains a green triangle by Proposition 233 and is connected by 2 of the same proposition. Observe that this implies that any matching in $K[A_1, B_1, C_1]$ is connected and odd. We will show that $K[A_1, B_1, C_1]$ contains a green matching of size at least $\frac{3}{4}(1-\eta')n$ contradicting the fact that K is not $\frac{3}{4}(1-\eta')n$-odd. We distinguish two cases. If $|A_1| \geq |B_1 \cup C_1|$ an easy greedy algorithm guarantees a green matching of size $|B_1 \cup C_1| - \eta n > (1-3\eta)\frac{3}{4}n \geq \frac{3}{4}(1-\eta')n$ in $K[A_1, B_1 \cup C_1]$. If $|A_1| \leq |B_1 \cup C_1|$ on the other hand there is a green matching covering at least $|A_1 \cup B_1 \cup C_1| - 4\eta n - 1 > (1-4\eta)\frac{3}{2}n - 1 \geq \frac{3}{2}(1-\eta')n$ vertices in $K[A_1, B_1, C_1]$ by Proposition 22. $\square$

We will now use Lemma 25 to prove Lemma 9.

Proof of Lemma 9. Let η' be given and set $\tilde{\eta} := \eta'/15$. Let η_{L25} and n_0 be provided by Lemma 25 for input $\eta'_{\mathrm{L25}} = \tilde{\eta}$ and set $\eta := \min\{\eta_{\mathrm{L25}}, \tilde{\eta}/5\}$. Let $K = (A\dot{\cup}B\dot{\cup}C, E)$ be a non-extremal coloured member of $\mathcal{K}_n^\eta$ with partition classes and assume for a contradiction that K is not $(1-\eta')3n/4$-odd.

Our first step is to show that K has big green and red connected matchings.

Fact 1. *K has even connected matchings M_r and M_g in red and green, respectively, with $|M_r|, |M_b| \geq (1-\tilde{\eta})n$.*

Assume for a contradiction that a maximum matching M in red has size less than $(1-\tilde{\eta})n$. By Lemma 25 applied with $\tilde{\eta}$ we conclude that there is an odd connected matching M' with $|M'| > |M|$. On the other hand K is not $\big((1-\eta')3n/4\big)$-good, hence $|M'| < (1-\eta')3n/4$. Another application of Lemma 25 with $\tilde{\eta} \leq \eta'$ thus provides us with a red connected matching of size bigger than $|M'|$ which contradicts the maximality of M. We conclude that there is a red connected matching M_r, and by symmetry also a green connected matching M_g, of size at least $(1-\tilde{\eta})n$. Clearly, M_r and M_g are even since K is not $\big((1-\eta')3n/4\big)$-good.

Let R be the component of M_r and G be the component of M_g in K. Fact 1 states, that R and G are bipartite. We observe in the following fact that both R and G substantially intersect all three partition classes. For this purpose define $D_r := D \cap V(R)$ and $D_g := D \cap V(G)$, and further $\bar{D}_r := D \setminus D_r$ and $\bar{D}_g := D \setminus D_g$ for all $D \in \{A, B, C\}$.

Fact 2. *For all $D \in \{A, B, C\}$ and $c \in \{r, g\}$ we have $|D_c| \geq 2\tilde{\eta}n$.*

Indeed, assume without loss of generality, that $|A_r| < 2\tilde{\eta}n$ which implies $|\bar{A}_r| > (1-2\tilde{\eta})n$. As $|M_r| \geq (1-\tilde{\eta})n$ it follows that $|B_r| > (1-3\tilde{\eta})n$ and $|C_r| > (1-3\tilde{\eta})n$. By definition all edges between $\bar{A}_r$ and $B_r \cup C_r$ are green and thus K is in pyramid configuration with tunnel, pyramids $(B_r, \bar{A}_r)$ and $(C_r, \emptyset)$, and parameter $3\tilde{\eta} < \eta'$, which is a contradiction.

Next we strengthen the last fact by showing that at most one of the sets $\bar{D}_c$ with $D \in \{A, B, C\}$ and $c \in \{r, g\}$ is significant.

Fact 3. *There is at most one set $D \in \{A, B, C\}$ and colour $c \in \{r, g\}$ such that $|\bar{D}_c| \geq \tilde{\eta}n$.*

If such a D and c exist we assume, without loss of generality, $D = A$ and $c = r$. Hence, for the proof of Fact 3, assume that $|\bar{A}_r| \geq \tilde{\eta}n$. First we show that

$$|\bar{B}_r|, |\bar{C}_r| < \tfrac{\tilde{\eta}}{2}n. \tag{17}$$

Assume for a contradiction and without loss of generality that $|\bar{B}_r| \geq \frac{\tilde{\eta}}{2}n$. By definition, all edges in $E(\bar{A}_r, C_r\dot{\cup}B_r)$ and $E(\bar{B}_r, C_r\dot{\cup}A_r)$ are green. Since $|\bar{A}_r|, |\bar{B}_r| \geq \tilde{\eta}n > 2\eta n$ by assumption and $|A_r|, |B_r| \geq 2\tilde{\eta}n/2 > 2\eta n$ (by Fact 2) we can apply Proposition 232 to infer that the graph with edges $E(\bar{A}_r, C_r\dot{\cup}B_r)$ and $E(\bar{B}_r, C_r\dot{\cup}A_r)$ is connected. As M_r is even we conclude that all edges in $E(A_r, C_r)$, $E(B_r, C_r)$, and $E(A_r, B_r)$ are red. Since $|A_r|, |B_r|, |C_r| \geq \tilde{\eta}n > 2\eta n$ by Fact 2 we infer from Proposition 233 that the graph $K[A_r, B_r, C_r] \subseteq R$ contains a red triangle which contradicts the fact that M_r is even.

Thus it remains to show that $|\bar{D}_g| < \tilde{\eta}n$ for all $D \in \{A, B, C\}$. By (17) and Fact 2 we have $|B_r \cap B_g|, |C_r \cap C_g| > \frac{\tilde{\eta}}{2}n > \eta n$ which implies that there is an edge in $E(B_r \cap B_g, C_r \cap C_g)$. By assumption we also have $|\bar{A}_r| \geq \tilde{\eta}n > 2\eta n$ and thus each pair of vertices in $B_r\dot{\cup}C_r$ has a common neighbour in $\bar{A}_r$ by 1 of Proposition 23. By definition of $\bar{A}_r$ all edges in $E(\bar{A}_r, B_r\dot{\cup}C_r)$ are green, and therefore we conclude that all edges in $E(B_r \cap B_g, C_r \cap C_g)$ are red since otherwise there would be a green triangle connected to M_g. Accordingly $|\bar{A}_g| \leq 2\eta n < \tilde{\eta}n/2$ since otherwise we could equally argue that all edges in $E(B_r \cap B_g, C_r \cap C_g)$ are green, a contradiction. Therefore $|A_g| \geq (1-\tilde{\eta}/2)n$. As $|\bar{A}_r| \geq \tilde{\eta}n$ this implies $|A_g \cap \bar{A}_r| \geq \tilde{\eta}/2n > \eta n$ and from (17) we also get $|B_r \cap \bar{B}_g| \geq \tilde{\eta}/2n > \eta n$. Thus there is an edge in $E(A_g \cap \bar{A}_r, B_r \cap \bar{B}_g)$. However, this edge can neither be red since it connects $\bar{A}_r$ and B_r, nor green since it connects $\bar{B}_g$ and A_g, a contradiction. Therefore $|\bar{B}_g| < \tilde{\eta}n$ and by symmetry also $|\bar{C}_g| < \tilde{\eta}n$ which finishes the proof of Fact 3.

We label the vertices in each of the bipartite graphs R and G according to their bipartition class by 1 and 2. In the remaining part of the proof we examine the distribution of these bipartition classes over the partition classes of K. Let F_{ij} denote the set of vertices in $V(R) \cap V(G)$ with label i in R and label j in G for $i, j \in [2]$. Let further F_{0j} be the set of vertices in $\bar{A}_r \cap V(G)$ that have label j in G for $j \in [2]$. Next we observe that each of the sets F_{ij} with $i, j \in [2]$ is essentially contained in one partition class of K.

Fact 4. *For all $i, j \in [2]$ there is at most one partition class $D \in \{A, B, C\}$ of K with $|F_{ij} \cap D| \geq \tilde{\eta}n$. Moreover $E(F_{0j}, F_{ij}) = \emptyset$.*

To prove the first part of Fact 4 assume for a contradiction that $|F_{ij} \cap A|, |F_{ij} \cap B| \geq \tilde{\eta}n$. Then there would be an edge in $K[A \cap F_{ij}, B \cap F_{ij}]$ since $\tilde{\eta} > \eta$. This contradicts the fact that F_{ij} is independent by definition. For the second part observe that an edge in $E(F_{0j}, F_{ij})$ can neither be red as such an edge would connect vertices from $\bar{A}_r$ to R nor green since $F_{0j} \cup F_{ij}$ lies in one bipartition class j of G.

Fact 5. *There are $X, Y \in \{A, B, C\}$ with $X \neq Y$ and indices $b, b', c, c' \in [2]$ with $bb' \neq cc'$ such that $|F_{bb'} \cap X|, |F_{cc'} \cap Y| \geq (1 - 5\tilde{\eta})n$ and $|F_{0b'}|, |F_{0c'}| \leq \tilde{\eta}n$.*

We divide the proof of this fact into three cases: The first case deals with $\bar{A}_r \neq \emptyset$, the second one with $\bar{A}_r = \emptyset$ and the additional assumption that there are $D \in \{A, B, C\}$ and $ij \neq i'j' \in [2]$ such that $|D \cap F_{ij}|, |D \cap F_{i'j'}| \geq \tilde{\eta}n$. The third and remaining case treats the situation when $\bar{A}_r = \emptyset$ and for each $D \in \{A, B, C\}$ there is at most one index pair (i, j) with $|D \cap F_{ij}| \geq \tilde{\eta}n$.

For the first case, let $j \in [2]$ be such that $F_{0j} \neq \emptyset$. Observe that then the second part of Fact 4 implies that $|F_{1j} \cap (B \cup C)|, |F_{2j} \cap (B \cup C)| < \eta n$. Let $c' = b' \in [2]$ with $c' \neq j$. Then, because Fact 3 implies that $|\bar{B}_r|, |\bar{B}_g|, |\bar{C}_r|, |\bar{C}_g| < \tilde{\eta}n$, we have that $|B \cap (F_{1b'} \cup F_{2b'})| \geq (1 - 4\tilde{\eta})n$ and $|C \cap (F_{1c'} \cup F_{2c'})| \geq (1 - 4\tilde{\eta})n$. Thus there is a $b \in [2]$ such that $|B \cap F_{bb'}| \geq \tilde{\eta}n$. Let $c' \in [2]$ with $c' \neq b'$. The first part of Fact 4 implies that $|C \cap F_{bc'}| < \tilde{\eta}n$, thus $|C \cap F_{cc'}| \geq (1 - 5\tilde{\eta})n \geq \tilde{\eta}n$. By symmetry we also get $|B \cap F_{bb'}| \geq (1 - 5\tilde{\eta})n$. This proves the first part of the statement for the first case. To see the second part, observe that if $F_{0b'} \neq \emptyset$, then $|F_{1b'} \cap (B \cup C)|, |F_{2b'} \cap (B \cup C)| < \eta n$ by Fact 4, a contradiction.

The second part of the second and third cases is straightforward as $F_{0,1}, F_{0,2} \subseteq \bar{A}_r = \emptyset$. To see the first part of the second case let D be as specified above and $\{X, Y\} = \{A, B, C\} \setminus \{D\}$. The first part of Fact 4 implies that $|F_{ij} \cap X|, |F_{i'j'} \cap X|, |F_{ij} \cap Y|, |F_{i'j'} \cap Y| < \tilde{\eta}n$. Thus $|(F_{ij'} \cup F_{i'j}) \cap X| \geq (1 - 2\tilde{\eta})n - 2\tilde{\eta}n$, as $|\bar{X}_r|, |\bar{X}_g| < \tilde{\eta}n$. Without loss of generality, let ij' be such that $|X \cap F_{ij'}| \geq \tilde{\eta}n$. We set $b := i$, $b' := j'$, $c = i'$ and $c' := j$. The rest of the proof is similar to the first case, proving that then $|Y \cap F_{cc'}| \geq (1 - 5\tilde{\eta})n$ and by symmetry that $|X \cap F_{bb'}| \geq (1 - 5\tilde{\eta})n$.

It remains to prove the first part of the third case. For this observe that for all $D \in \{A, B, C\}$ we have that $|D \cap \bigcup_{(i'j') \neq (i,j)} F_{i',j'}| < 3\tilde{\eta}n$, where i, j are as specified in the definition of the third case. Observe also that $|\bar{D}_r|, |\bar{D}_g| < \tilde{\eta}n$. This implies $|D \cap F_{i,j}| \geq (1 - 5\tilde{\eta})n$, as desired. Hence, for $X = B$ and $Y = C$ we obtain indices b, b', c, c' such that $|X \cap F_{bb'}|, |Y \cap F_{i'j'}| \geq (1 - 5\tilde{\eta})n$, with $bb' \neq cc'$ by Fact 5.

This brings us to the last step which shows that K is extremal, a contradiction.

Fact 6. *K is in pyramid configuration with parameter η'.*

Let $X, Y \in \{A, B, C\}$ and $b, b', c, c' \in [2]$ be as in Fact 5. Let $Z \in \{A, B, C\} \setminus \{X, Y\}$. Assume without loss of generality that $b = b' = 1$. Thus Fact 5 states that $|F_{11} \cap X| \geq (1 - 5\tilde{\eta})n$ and $|F_{01}| \leq \tilde{\eta}n$. We distinguish two cases. First, assume that $c' = 2$ and set $\bar{c} := 3 - c$. By Fact 5 this implies $|F_{c2} \cap Y| \geq (1 - 5\tilde{\eta})n$ and $|F_{02}| \leq \tilde{\eta}n$ and thus $|(F_{\bar{c}2} \cup F_{21}) \cap Z| \geq (1 - 5\tilde{\eta})n$ by Fact 4. Moreover $E(F_{11} \cap X, F_{21} \cap Z)$ forms an η-complete red bipartite graph since $F_{11} \cup F_{21}$ is an independent set in G. Similarly $E(F_{c2} \cap Y, F_{\bar{c}2} \cap Z)$ forms an η-complete red bipartite graph. Further, if $c = 2$ then $E(F_{c2} \cap Y, F_{21} \cap Z)$ and $E(F_{11} \cap X, F_{\bar{c}2} \cap Z)$ form η-complete green bipartite graphs (leading to crossings) and if $c = 1$ then $E(F_{11} \cap X, F_{c2} \cap Y)$ forms an η-complete green bipartite graph (leading to a tunel). Therefore, in both subcases, K is in pyramid configuration with parameter $5\tilde{\eta} \leq \eta'$ and pyramids $(F_{11} \cap X, F_{21} \cap Z)$ and $(F_{c2} \cap Y, F_{\bar{c}2} \cap Z)$, unless one of the sets $F_{21} \cap Z$ and $F_{\bar{c}2} \cap Z$ has size at most $10\tilde{\eta}n$. In this case, however, we can simply replace this set by the empty set and still obtain a pyramid configuration with parameter at most $15\tilde{\eta} \leq \eta'$.

In the case $c' = 1$ we have $c = 2$. Fact 5 guarantees that $|F_{21} \cap Y| \geq (1 - 5\tilde{\eta})n$. Since $|F_{01}| \leq \tilde{\eta}n$ we conclude from Fact 4 that $|(F_{12} \cup F_{22} \cup F_{02}) \cap Z| \geq (1 - 5\tilde{\eta})n$. Similarly as before $E(F_{11} \cap X, (F_{12} \cup F_{02}) \cap Z)$ and $E(F_{21} \cap Y, F_{22} \cap Z)$ form η-complete green bipartite graphs and $E(F_{11} \cap X, F_{21} \cap Y)$ forms an η-complete red bipartite graph. Accordingly we also get a pyramid configuration with parameter $5\tilde{\eta} \leq \eta'$ in this case, where the pyramids are $(F_{11} \cap X, (F_{12} \cup F_{02}) \cap Z)$ and $(F_{21} \cap Y, F_{22} \cap Z)$ unless, again, $(F_{12} \cup F_{02}) \cap Z$ or $F_{22} \cap Z$ are too small in which case we proceed as above. □

7.3. **Extremal configurations.** Our aim in this section is to provide a proof of Lemma 10. This proof naturally splits into two cases concerning pyramid and spider configurations, respectively. The former is covered by Proposition 26, the latter by Proposition 27.

Proposition 26. *Lemma 10 is true for pyramid configurations.*

Proof. Given η' set $\eta = \eta'/3$. Let K be a coloured graph from $\mathcal{K}_n^\eta$ that is in pyramid configuration with parameter η and pyramids (D_1, D_1') and (D_2, D_2') such that the requirements of 1 in Definition 7 are met for colours c and c'.

Fact 1. *If the pyramid configuration has crossings then K is $\big((1-\eta')n, (1-\eta')\frac{3}{2}n, 2\big)$-good.*

Indeed, by Proposition 21 there is a matching M of colour either c or c' and size at least $(1-2\eta)\frac{1}{2}n$ in $K[D_1, D_2]$. Note further, that the pyramid configuration with crossings is symmetric with respect to the colours c and c' and hence we may suppose, without loss of generality, that M is of colour c and that $|D'_1| \geq (1-\eta)\frac{1}{2}n$. As $K[D_1, D'_1]$ and $K[D_2, D'_2]$ are (η, c)-complete, there are c-coloured matchings M_1 and M_2 in $K[D_1, D'_1]$ and $K[D_2 \setminus M, D'_2]$, respectively, of size at least $\min\{|D'_1|, |D_1|\} - \eta n$ and $\min\{|D'_2|, |D_2 \setminus M|\} - \eta n$, respectively. This implies

$$|M| + |M_1| + |M_2| \geq (1-3\eta)\frac{3}{2}n = (1-\eta')\frac{3}{2}n.$$

Observe that, depending on the size of M, either $M \cup M_2$ or $M_1 \cup M_2$ is a matching of size at least $(1-3\eta)n = (1-\eta')n$. Now, the union of M, M_1, and M_2 forms a 2-fork system F and since $K[D_1, D'_1]$ and $K[D_2, D'_2]$ are (η, c)-complete the bipartite graph formed by these two graphs and M is connected and has partition classes $D_1 \cup D'_2$ and $D_2 \cup D'_1$. It follows that F has size $|M| + |M_1| + |M_2| \geq (1-\eta')\frac{3}{2}n$.

Fact 2. *If the pyramid configuration has a c'-tunnel and if there is a matching M of colour c' and size at least $(1-\eta')\frac{1}{2}n$ in $K[D_1, D'_1 \cup D'_2]$ or in $K[D_2, D'_1 \cup D'_2]$ then K is $\big((1-\eta')n, (1-\eta')\frac{3}{2}n, 2\big)$-good in colour c'.*

As K has a c'-tunnel, there is a connected matching M' of colour c' and size at least $|D_1| - \eta n \geq (1-\eta')n$ in $K[D_1, D_2]$. We will extend the matching M' (which is a 1-fork-system) to a 2-fork system. Without loss of generality assume that the matching M promised by Fact 2 is in $K[D_1, D'_1 \cup D'_2]$. As $M \cap D_1$ and D_2 are non-negligible the bipartite graph $K[M \cap D_1, D_2]$ is connected by Proposition 232 and thus M is connected. Hence $M \cup M'$ forms a connected 2-fork system centered in D_1 and of size $|M'| + |M| \geq (1-\eta')\frac{3}{2}n$.

Fact 3. *If the pyramid configuration has a c'-tunnel but no crossings and there is no matching of colour c' and size at least $(1-\eta')\frac{1}{2}n$ in $K[D_1, D'_1 \cup D'_2]$ or in $K[D_2, D'_1 \cup D'_2]$ then K is $((1-\eta')n, (1-\eta')\frac{3}{2}n, 3)$-good in colour c.*

To obtain the 3-fork system note that Proposition 21 implies that there are matchings M_1 and M_2 of colour c and sizes at least $(1-\eta')\frac{1}{2}n$ in $K[D_1, D'_1 \cup D'_2]$ and $K[D_2, D'_1 \cup D'_2]$, respectively. The union of M_1 and M_2 forms a 2-fork system F centered in $D'_1 \cup D'_2$ covering at least $(1-\eta')\frac{1}{2}n$ vertices in D_1 and at least $(1-\eta')\frac{1}{2}n$ vertices in D_2. We can assume without loss of generality that $|D'_1| \geq (1-\eta)\frac{1}{2}n \geq (1-\eta')\frac{1}{2}n$. As $K[D_1, D'_1]$ is (η, c)-complete and $|D_1 \setminus F| \leq (1-\eta')n - (1-\eta')\frac{1}{2}n = (1-\eta')\frac{1}{2}n$ we can greedily find a matching between D_1' and $D_1 \setminus F$ covering all but at most ηn vertices of $D_1 \setminus F$. Its union with F forms a 3-fork system F' centered in $D_1' \cup D_2'$ covering at least $(1-\eta')n$ vertices in D_1 and at least $(1-\eta')\frac{1}{2}n$ vertices in D_2, implying that F' has size at least $(1-\eta')\frac{3}{2}n$. The graph $K[D_1, D_1'] \cup K[D_2, D_2']$ clearly contains a matching M of size at least $|D_1' \cup D_2'| - \eta n \geq (1-\eta')n$ in colour c.

Since the pyramid configuration has no crossings there are edges of colour c in $K[D_1, D_2'] \cup K[D_2, D_1']$. Together with the fact that D_1, D_2, D_1', and D_2' are non-negligible, we obtain that the bipartite graphs $K[D_1, D_1' \cup D_2']$ and $K[D_2, D_1' \cup D_2']$ are connected by 2 of Proposition 23. Thus the matching M and the fork system F' are both connected. □

Proposition 27. *Lemma 10 is true for spider configurations.*

Proof. Given η' set $\eta = \eta'/5$ and let K be a coloured graph from $\mathcal{K}_n^\eta$ that is in spider configuration with parameter η, i. e., it satisfies 2 of Definition 7. In this proof we construct only matchings and fork systems of colour c. Observe that these are connected by definition. We distinguish two cases.

Case 1: First assume that $|A_1 \cup B_1 \cup C_1| < (1-\eta)\frac{3}{2}n$. We will show that in this case our configuration contains both a connected matching of size at least $(1-\eta')n$ and a connected 3-fork system of size at least $(1-\eta')\frac{3}{2}n$. We need the following auxiliary observation.

Fact 1. *If* $|A_1 \cup B_1 \cup C_1| < (1-\eta)\frac{3}{2}n$ *then* $A_B = B_A = \emptyset$. *Moreover* $|A_1| + \eta n \geq |B_C| = |C_B|$ *and* $|B_1| + \eta n \geq |A_C| = |C_A|$.

Indeed, by Condition 2c of 2 either $A_2 = \emptyset$ and hence $A_B \subseteq A_2$ is empty or $|A_2 \cup B_2 \cup (C_2 \setminus C_C)| \leq (1-\eta)\frac{3}{2}n$. In the second case we conclude from $|A_1 \cup B_1 \cup C_1| < (1-\eta)\frac{3}{2}n$ that

$$|A_1 \cup A_2| + |B_1 \cup B_2| + |C_1 \cup (C_2 \setminus C_C)| < (1-\eta)3n.$$

As $|A_1 \cup A_2|, |B_1 \cup B_2|, |C_1 \cup C_2| \geq (1-\eta)n$ it follows that $|C_1 \cup (C_2 \setminus C_C)| < (1-\eta)n$ and thus $C_C \neq \emptyset$. By Condition 2b of 2 we get $A_B = \emptyset$. For the second part of the fact observe that Condition 2a of 2 states that $n - |A_2| \geq |B_C| = |C_B|$ and thus we conclude $|A_1| \geq (1-\eta)n - |A_2| \geq |B_C| - \eta n = |C_B| - \eta n$. The inequality $|B_1| \geq |A_C| - \eta n$ is established in the same way.

Fact 2. *If* $|A_1 \cup B_1 \cup C_1| < (1-\eta)\frac{3}{2}n$ *then* K *is* $\big((1-\eta')n, (1-\eta')\frac{3}{2}n), 3\big)$-*good.*

¿From Condition 2a of 2 we infer that $|A_C| < n - |B_2| \leq |B_1| + \eta n$ and Fact 1 implies that $|A_1| + |A_C| = |A_1| + |A_2| \geq (1-\eta)n$ and $|C_1| + |C_2| \geq (1-\eta)n$. We thus conclude from $|A_1 \cup B_1 \cup C_1| < (1-\eta)\frac{3}{2}n$ that

$$|A_C| - \eta n < |B_1| < (1-\eta)\tfrac{3}{2}n - |A_1 \cup C_1| < |A_C| + |C_2| - \eta n.$$

This (together with the fact that $K[B_1, A_C]$ and $K[B_1, C_2]$ are (η, c)-complete) justifies that there is a c-coloured matching M_1 in $K[B_1, A_C \cup C_2]$ covering all vertices of B_1 and all but at most ηn vertices of A_C. Further, by Fact 1 we know that $|B_C| \leq |A_1| + \eta n$ and hence we can find a matching M_2 of colour c in (the (η, c)-complete graph) $K[B_C, A_1]$ covering all but at most ηn vertices of B_C. The matching $M := M_1 \cup M_2$ satisfies

$$|M| \geq |B_1| + |B_C| - \eta n = |B_1| + |B_2| - \eta n \geq (1-\eta)n - \eta n \geq (1-\eta')n,$$

where the equality follows from Fact 1. Next, we extend the matching M to a connected 3-fork system of colour c and size at least $(1-\eta')\frac{3}{2}n$ in the following way. Consider maximal matchings M_3, M_4, and M_5 in $K[B_1, C_A \setminus M_1]$, $K[A_1, C_B \setminus M_1]$ and $K[A_1, C_C \setminus M_1]$, respectively. By Fact 1 we infer that M_3 and M_4 each cover all but at most ηn vertices of $C_A \setminus M_1$ and $C_B \setminus M_1$, respectively. As $|C_C| \leq |C_C \cup C_1| \leq |A_1|$ by Condition 2a of 2 the matching M_5 covers all but at most ηn vertices of C_C.

Then the union $M \cup M_3 \cup M_4 \cup M_5$ is a 3-fork-system F centered in $A_1 \cup B_1$ and covering all but at most $5\eta n$ vertices of $A_C \cup B_C \cup C_A \cup C_B \cup C_C = A_2 \cup B_2 \cup C_2$. Thus F has size at least $(1-\eta)3n - |A_1 \cup B_1 \cup C_1| - 5\eta n \geq (1-\eta')\frac{3}{2}n$.

Case 2: Now we turn to the case $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$. We further divide this case into two subcases, treating $C_1 = \emptyset$ and $C_1 \neq \emptyset$, respectively.

Fact 3. *If* $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$ *and* $C_1 = \emptyset$ *then* K *is* $\big((1-\eta')n, (1-\eta')\frac{3}{2}n, 2\big)$*-good.*

By definition $|C_2| \geq (1-\eta)n - |C_1| = (1-\eta)n$ in this case. Therefore, using the fact that $K[A_1, C_2]$ and $K[B_1, C_2]$ are (η, c)-complete, we can greedily construct a maximal matching M_A in $K[A_1, C_2]$ and a maximal matching M'_B in $K[B_1, C_2 \setminus M_A]$ such that the matching $M := M_A \cup M'_B$ covers C_2 (as $|A_1 \cup B_1| = |A_1 \cup B_1 \cup C_1| > |C_2| + \eta n$) and thus has size at least $(1-\eta)n$. Then we extend M'_B to a maximal matching M_B in $K[B_1, C_2]$. Observe that M_A and M_B cover all but at most ηn vertices of A_1 and B_1, respectively. Thus the 2-fork system $F := M_A \cup M_B$ has size at least $|A_1 \cup B_1| - 2\eta n = |A_1 \cup B_1 \cup C_1| - 2\eta n \geq (1-\eta')\frac{3}{2}n$.

Now consider the subcase when $C_1 \neq \emptyset$.

Fact 4. *If* $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$ *and* $C_1 \neq \emptyset$ *then* $|B_1 \cup C_1| \leq (1-\eta)\frac{3}{4}n$ *and we have* $|C_2| \geq (1-\eta)\frac{1}{4}n$ *and* $|C_1| \leq |B_2| - \eta n$.

The first inequality follows from Condition 2d of 2. Accordingly $|C_2| \geq (1-\eta)n - |C_1| \geq (1-\eta)\frac{1}{4}n$ which establishes the second inequality. For the third inequality we use that $|B_1 \cup B_2| \geq (1-\eta)n$ by definition and so

$$|C_1| \leq (1-\eta)\tfrac{3}{4}n - |B_1| \leq (1-\eta)\tfrac{3}{4}n - (1-\eta)n + |B_2| \leq |B_2| - \eta n.$$

Fact 5. *If* $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$ *and* $C_1 \neq \emptyset$ *then there is a matching* M *of size at least* $(1-\eta)n$ *and colour* c *covering* C_1.

Let M_1 be a maximal matching in $K[C_1, B_2]$. We conclude from Fact 4 that M_1 covers C_1. Let M_2 be a maximal matching in $K[C_2, A_1 \cup B_1]$. As $|C_2| \leq n - |C_1| \leq |A_1 \cup B_1| - \eta n$ the matching M_2 covers C_2. Setting $M := M_1 \cup M_2$, we obtain a matching of size $|M| = |C_1| + |C_2| \geq (1-\eta)n$ as required.

Fact 6. *If* $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$ *and* $C_1 \neq \emptyset$, *then there is a* 3*-fork system of colour* c *and of size at least* $(1-\eta')\frac{3}{2}n$.

Let M be the matching from Fact 5. Clearly, we can greedily construct a 2-fork system F' in the (η, c)-complete graph $K[C_2, (A_1 \cup B_1) \setminus M]$ which either is of size $2|C_2|$ or covers all but at most ηn vertices of $|(A_1 \cup B_1) \setminus M|$. Then $F := M \cup F'$ forms a 3-fork system. If the former case occurs we infer from Fact 4 that F is of size at least $(1-\eta)n + 2|C_2| \geq (1-\eta)\frac{3}{2}n$. In the latter case F covers all but at most ηn vertices of $A_1 \cup B_1 \cup C_1$ and thus has size at least $(1-\eta)\frac{3}{2}n - \eta n \geq (1-2\eta)\frac{3}{2}n$. We conclude that K is $((1-\eta')n, (1-\eta')\frac{3}{2}n, 3)$-good also in the subcase $|A_1 \cup B_1 \cup C_1| \geq (1-\eta)\frac{3}{2}n$ and $C_1 \neq \emptyset$. □

8. Concluding remarks

As noted earlier our proof of Theorem 1 applies to suitably chosen (sparser) subgraphs of $K_{n,n,n}$ as well. More precisely, for any fixed $p \in (0,1)$ the same method can be used to show that asymptotically almost surely $\mathcal{G}_p(n,n,n) \to \mathcal{T}_t^{\Delta}$, where $\mathcal{G}_p(n,n,n)$ is a random tripartite graph with edge probability p and partition classes of size n, and where $t \leq (1-\mu)n/2$ and $\Delta \leq n^{\alpha}$ for a small positive $\alpha = \alpha(\mu, p)$. Indeed, standard methods can be used to show that the following holds asymptotically almost surely for $G = \mathcal{G}_p(n,n,n)$ with partition classes $V_1 \dot\cup V_2 \dot\cup V_3$ and for any $\zeta > 0$:

- G has at most $4pn^2$ edges.
- $e(U, W) \geq p|U||W|/2$ for all $U \subseteq V_i$ and $W \subseteq V_j$, $i \neq j$, with $\min\{|U|, |W|\} > \zeta n$.

The first property guarantees that we obtain a graph with few edges. We claim further that these two properties imply that $G \to \mathcal{T}_k^\Delta$. To see this we proceed as in the proof of Theorem 1 and apply the regularity lemma on the coloured graph G. We then colour an edge in the reduced graph $\mathbb{G}$ by green or red, repectively, if the corresponding cluster pair is regular and has density at least $p/4$ in green or red. Using the two properties from above it is not difficult to verify that $\mathbb{G}$ is a coloured tripartite graph that is η-complete. Hence, from this point on, we can use the strategy described in the proof of Theorem 1, apply our structural lemma, Lemma 8, the assignment lemma, Lemma 14, and the embedding lemma, Lemma 13.

One may ask whether this approach can be pushed even further and consider random tripartite graphs $\mathcal{G}_p(n, n, n)$ with edge probabilities $p(n)$ that tend to zero as n goes to infinity. It is likely that similar methods can be used in this case in conjunction with the regularity method for sparse graphs (see, e.g., [8]).

We close with an extension of Schelp's conjecture that was suggested to us by Jiří Matoušek.

Question 28. *Is it true that for all $\Delta \in \mathbb{N}$ and $\mu > 0$ there is a $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$? If $t \leq (1-\mu)\frac{1}{2}n$ and G is a graph on n vertices with minimum degree $\delta(G) \geq (\frac{2}{3} - \mu)n$ then $G \to \mathcal{T}_t^\Delta$.*

References

1. M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi, *Erdős-Sós conjecture*, In preparation.
2. N. Alon and J. H. Spencer, *The probabilistic method*, John Wiley & Sons, 2000.
3. J. Beck, *On size Ramsey number of paths, trees and circuits. II*, Mathematics of Ramsey theory, Algorithms Combin., vol. 5, Springer, 1990, pp. 34–45.
4. A. Gyárfás, M. Ruszinkó, G. N. Sárközy, and E. Szemerédi, *Tripartite Ramsey numbers for paths*, J. Graph Theory **55** (2007), no. 2, 164–174.
5. A. Gyárfás, G. N. Sárközy, and R. H. Schelp, *Multipartite ramsey numbers for odd cycles*, to appear in J. Graph Theory.
6. P. E. Haxell and Y. Kohayakawa, *The size-Ramsey number of trees*, Israel J. Math. **89** (1995), no. 1-3, 261–274.
7. J. Hladký and D. Piguet, *Loebl-Komlós-Sós conjecture: dense case*, Preprint.
8. Y. Kohayakawa and V. Rödl, *Regular pairs in sparse random graphs. I*, Random Structures Algorithms **22** (2003), no. 4, 359–434.
9. Y. Kohayakawa, V. Rödl, M. Schacht, and E. Szemerédi, *Sparse partition universal graphs for graphs of bounded degree*, Preprint.

10. T. Łuczak, $R(C_n, C_n, C_n) \leq (4+o(1))n$, J. Combin. Theory Ser. B **75** (1999), no. 2, 174–187.
11. F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. **30** (1930), no. 2, 264–286.
12. R.H. Schelp, personal communication (Research Seminar, Alfréd Renyi Institute, Budapest), 2008.
13. E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.
14. Y. Zhao, *Proof of the $(n/2$ - $n/2$ - $n/2)$ Conjecture for large n*, Preprint.

Zentrum Mathematik, Technische Universität München, Boltzmannstrasse 3, D-85747 Garching bei München, Germany
E-mail address: boettche@ma.tum.de

Department of Applied Mathematics, Faculty of mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00, Prague, Czech Republic and Zentrum Mathematik, Technische Universität München, Boltzmannstrasse 3, D-85747 Garching bei München, Germany
E-mail address: honzahladky@googlemail.com

Zentrum Mathematik, Technische Universität München, Boltzmannstrasse 3, D-85747 Garching bei München, Germany
E-mail address: diana@kam.mff.cuni.cz