RAMSEY GOODNESS OF TREES IN RANDOM GRAPHS

PEDRO ARAÚJO, LUIZ MOREIRA, AND MATÍAS PAVEZ-SIGNÉ

Abstract. For a graph G, we write $G \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$ if every red-blue colouring of the edges of G contains either a blue K_{r+1}, or red copies of every tree with n edges and maximum degree at most D. In 1977, Chvátal proved that, for any integers $r, n, D \geqslant 2$, $K_N \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$ if and only if $N \geqslant rn + 1$.

In this paper we prove that there exists $C > 0$ such that, with high probability, if $N \geqslant rn + C/p$ then

$$G(N,p) \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$$

for any $p \gg N^{-2/(r+2)}$. The bound on N is best possible up to the value of C. We also prove an approximate random analogue of the Erdős-Sós conjecture for trees with linear size and bounded maximum degree. That is, for every for $\varrho \in (0,1)$, every subgraph of $G(N,p)$ with at least $\big(\varrho + o(1)\big)p\binom{N}{2}$ edges contains, with high probability, every tree $T \in \mathcal{T}(\varrho N, D)$, provided that $p \gg 1/N$.

1. Introduction

Ever since the seminal work of Erdős and Rényi [11], the study of the binomial random graph have played a central role in combinatorics. In this paper we will study the Ramsey properties of the Erdős–Rényi random graph, continuing a line of research that was initiated in the 1980s by Frankl and Rödl [12] and Łuczak, Ruciński, and Voigt [26]. Let us write $G \to (H_1, H_2)$ to denote that every blue-red colouring of the edges of G contains either a blue copy of H_1 or a red copy of H_2 (if $H_1 = H_2$ then we write $G \to H$). An important early breakthrough by Rödl and Ruciński [31, 32] established the following threshold result for fixed non-acyclic graphs H:

$$\lim_{n\to\infty} \mathbb{P}\big(G(N,p) \to H\big) = \begin{cases} 1 & \text{if } p \gg N^{-1/m_2(H)}, \\ 0 & \text{if } p \ll N^{-1/m_2(H)}, \end{cases}$$

where $m_2(H) = \max\big\{\frac{e(H')-1}{v(H')-2} : H' \subseteq H \text{ with } v(H') \geqslant 3\big\}$. A corresponding result for hypergraphs was obtained by Friedgut, Rödl and Schacht [13], and independently by Conlon and Gowers [9], and the 1-statement of an asymmetric version (conjectured by Kreuter and Kohayakawa [21] in 1997) was recently proved by Mousset, Nenadov, and Samotij [30].

This research of PA and LM was partially supported by CNPq, and the research of MPS was partially supported by CONICYT Doctoral Fellowship 21171132 and the project UCH-1566 "Consolidación de la internacionalización de la investigación y postgrado de la Universidad de Chile".

Ramsey properties of random graphs involving sparse graphs have also attracted significant attention in recent years. To give just two examples, Letzter [27] proved that if $\varepsilon > 0$ and $pn \to \infty$, then

$$G\big((3/2+\varepsilon)n, p\big) \to P_n$$

with high probability (the constant 3/2 is best possible), and Kohayakawa, Mota and Schacht [22] proved that $\left(\frac{\log N}{N}\right)^{1/2}$ is the threshold for the event that for any two-colouring of the edges of $G(N,p)$, there exist two monochromatic trees that partition the vertex set.

In this paper we will be interested in the problem of extending to the setting of sparse random graphs a theorem of Chvátal [8] from 1977, which states that if $r \in \mathbb{N}$, and T is a tree with n edges, then

$$K_N \to (K_{r+1}, T) \qquad \Leftrightarrow \qquad N \geqslant rn+1. \tag{1}$$

The necessity of the lower bound on N is easy to see, and (as was first observed by Burr [6]) holds in significantly greater generality. To be precise, if H is a connected graph, F is a graph with $\sigma(F) \leqslant |H|$, where $\sigma(F)$ is the minimum size of a colour class in a proper $\chi(F)$-colouring of F, and $N < \big(\chi(F)-1\big)\big(|H|-1\big) + \sigma(F)$, then

$$K_N \not\to (F, H).$$

Indeed, to see this it suffices to consider $\chi(F)-1$ disjoint red cliques of size $|H|-1$, and one additional disjoint red clique of size $\sigma(F)-1$. A (connected) graph H is said to be *Ramsey F-good* (or just *F-good*) if $K_N \to (F,H)$ whenever $N \geqslant \big(\chi(F)-1\big)\big(|H|-1\big) + \sigma(F)$. The systematic study of Ramsey goodness was initiated by Burr and Erdős [7] in 1983.

As far as we are aware, the problem of Ramsey goodness in random graphs was first studied only very recently, by the second author [29], who considered the case in which F is a clique and H is a path. The main results of [29] identified two different thresholds for the event that $G(N,p) \to (K_{r+1}, P_n)$, for different values of N. More precisely, it was proved there that if $p \gg n^{-2/(r+2)}$ and $t \gg 1/p$, then

$$G\big(rn+t, p\big) \to \big(K_{r+1}, P_n\big),$$

while if $p \gg n^{-2/(r+1)}$ and $t = \Omega(n)$, then

$$G\big(rn+t, p\big) \to \big(K_{r+1}, P_n\big)$$

in both cases with high probability as $n \to \infty$. These results are sharp in the sense that whp $G(rn+t,p) \not\to (K_{r+1}, P_n)$ in three different settings. First, if $p \in (0,1)$ and $t \ll 1/p$, then one can partition $V(G(N,p)) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that $|V_0| = t$ and $e(V_0, V_r) = 0$. This is possible since, with high probability, sets of size $o(1/p)$ have $o(n)$ external neighbours in $G(N,p)$. Then we can colour the edges in red if and only if they have both endpoints inside parts without creating a blue K_{r+1} or any red component with more than n vertices. Second, for $n^{-2/(r+1)} \ll p \ll n^{-2/(r+2)}$, one can show that there are values of $t \gg 1/p$ such that $G\big(rn+t,p\big) \nrightarrow \big(K_{r+1},P_n\big)$. Finally, if $p \ll n^{-2/(r+1)}$ and $t = O(n)$, then, with high probability, $G(N,p)$ has $o(n)$ copies of K_{r+1}, whose edges can be all coloured in red without creating any red copy of P_n, see [29] for the details.

Our main theorems generalise the results of [29] from paths to arbitrary bounded degree trees. Let us denote by $\mathcal{T}(n,D)$ the class of all trees with n edges and maximum degree at most D. Let us write $G \to (K_{r+1},\mathcal{T}(n,D))$ to denote that $G \to (K_{r+1},T)$ for every $T \in \mathcal{T}(n,D)$.

Theorem 1.1. *For each $r,D \geqslant 2$, there exist $C,C' > 0$ such that the following holds. If*

$$p \geqslant C'N^{-2/(r+2)} \qquad \text{and} \qquad N \geqslant rn + C/p,$$

then $G(N,p) \to \big(K_{r+1},\mathcal{T}(n,D)\big)$ with high probability as $n \to \infty$.

As mentioned above, it follows from the results of [29] that the bound on N is sharp up to the value of C, and the bound on p is sharp up to a the value of C'. For smaller values of p we obtain the following bound.

Theorem 1.2. *For every $r,D \geqslant 2$ and $\varepsilon > 0$ there exists $C' > 0$ such that the following holds. If*

$$p \geqslant C'N^{-2/(r+1)} \qquad \text{and} \qquad N \geqslant rn + \varepsilon n,$$

then $G(N,p) \to \big(K_{r+1},\mathcal{T}(n,D)\big)$ with high probability as $n \to \infty$.

We will prove Theorem 1.2 by iteratively applying a theorem due to Haxell [17] to find either red copies of every tree in $\mathcal{T}(n,D)$, or $r+1$ large disjoint sets with only blue edges between them. The result will then follow by a straightforward application of the Janson inequalities. The proof of Theorem 1.1 is significantly more challenging, and is based on a stability argument. One of the key steps is to prove that the random graph not only contains all large bounded degree trees, but is also resilient with respect to this property.

Resilience is a measure of how much one has to perturb a graph in order to destroy a given property of it (see e.g. [5] for a discussion of resilience in the random graph) and it is a convenient way of phrasing extremal problems in general settings. For example, in the context of bounded degree trees, a classical result of Komlós, Sárközy and Szemerédi [24] says that every n-vertex graph G with $\delta(G) \geqslant (1/2 + o(1))n$ contains all trees in $\mathcal{T}(n,D)$, for n large enough. In other words, one can say that if an adversary deletes a $(1/2 - o(1))$-proportion of the edges incident at each vertex of K_n, the resulting graph still contains all trees in $\mathcal{T}(n,D)$. Balogh, Csaba and Samotij [1] proved that the same happens a.a.s. in the random graph for the class of almost spanning trees with bounded degree, provided that $p \gg 1/n$. That is, any subgraph of $G(n,p)$ obtained by deleting at most a $(1/2-o(1))$-proportion of the edges incident to each vertex of $G(n,p)$ contains all trees in $\mathcal{T}((1-o(1))n,D)$ with high probability.

In [1], the authors developed tools for embedding trees in "well-behaved" sparse bipartite graphs. We combine these tools with the approach of Besomi, Stein and the third author [4] to the Erdős–Sós Conjecture[1], for bounded degree trees and dense host graphs, to obtain the following "global" resilience result.

Theorem 1.3. *For every $D \geqslant 2$ and $\delta, \varrho \in (0,1)$, there exists $C > 0$ such that if $p \geqslant C/N$, then $G = G(N,p)$, with high probability, has the following property. Every subgraph $G' \subseteq G$ with $e(G') \geqslant (\varrho + \delta)\, e(G)$ is $T \in \mathcal{T}(\varrho N, D)$-universal.*

Theorem 1.3 will follow by a stronger result, in which $G(N,p)$ can be replaced by pseudo-random graphs. More precisely, we ask that the number of edges between any disjoint pair of sets is roughly what one would expect in $G(N,p)$.

In terms of resilience, Theorem 1.3 says that if $pN \gg 1$, then a.a.s. one can delete a $(1 - \varrho - o(1))$-proportion of the edges of $G(N,p)$ so that the resulting graph still contains all trees in $\mathcal{T}(\varrho N, D)$. This result can be viewed as an approximate random analogue of the Erdős–Sós conjecture for bounded degree trees of linear size. We point out that Theorem 1.3 is sharp in the following senses: the value of p is best possible, up to a constant factor, since the largest connected component of $G(N,p)$ is sublinear when $p \ll 1/N$. Moreover, for an integer $r \geqslant 2$ and $\varrho = 1/r$, the constant ϱ cannot be improved. Indeed, one can partition the vertex set in $r+1$ parts, one with at most r vertices and the others in a balanced way and thus with fewer than N/r vertices. If the edges between parts are deleted, then a.a.s. we get a subgraph $G' \subseteq G(N,p)$ which has $(1/r - o(1))e(G(N,p))$ edges but every connected component of G' has less than N/r vertices.

The proof of Theorem 1.3 relies on the so-called *regularity method* for random graphs. Let $G' \subseteq G(N,p)$ be a graph with $e(G) \geqslant (\varrho + \delta)e(G(N,p))$. Using the sparse regularity lemma one finds a regular partition of $V(G')$ such that its corresponding reduced graph R has edge density at least $\varrho + \delta/2$. Let k be the number of clusters in R. By a standard argument, we can find an induced subgraph $R' \subseteq R$ with average degree at least $(\varrho + \delta/2)k$ and minimum degree at least $(\varrho + \delta/2)k/2$.

We will use the following structure of R'. We consider a cluster $X \subset V(R')$ with maximum degree in R', in particular we have $|N_{R'}(X)| \geqslant (\varrho + \delta/2)k$. We can then partition $N_{R'}(X)$ into a matching $\mathcal{M}$ and an independent set $\mathcal{Y}$ so that every cluster in $\mathcal{Y}$ has a large degree outside $N_{R'}(X)$. Let $\mathcal{H}$ be the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z} = N_{R'}(\mathcal{Y}) \setminus (X \cup N_{R'}(X))$. We point out that since $|V(\mathcal{M})| + |\mathcal{Y}| = |N_{R'}(X)| \geqslant (\varrho + \delta/2)k$, then this structure has enough space in order to embed any tree of size ϱn.

[1] The Erdős–Sós Conjecture [10] from 1964 says that, given $k \in \mathbb{N}$, every graph G with average degree greater than k contains all trees with $k+1$ edges. In particular, it says that if $k = \varrho n$ for some $n \in \mathbb{N}$, then every graph G with $e(G) > \varrho n^2/2 \approx \varrho\binom{n}{2}$ edges contains each tree with $\varrho n + 1$ edges.

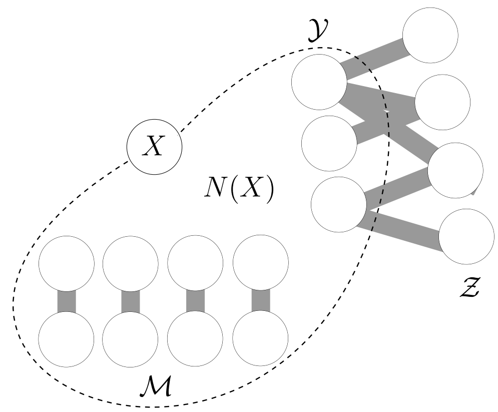

FIGURE 1. Structure in the reduced graph

First we consider the case of a path P with ϱn edges. We first cut P into a constant number of very small subpaths of odd length. We embed $P = P_1 \dots P_t$ sequentially path-by-path, in such a way that the embedding of P remains connected at each step. Starting with P_1, we embed the starting point of P in X. In general, the starting point of each subpath P_i is embedded into X, and the rest of P_i is embedded into some edge from R which is adjacent to X. Since $\mathcal{H}$ is bipartite and the number of vertices of P_i is odd, the last vertex of P_i can be embedded into a vertex having a large neighbourhood in X. This allows us to continue with the embedding of P_{i+1}, and so on.

The proof for general trees $T \in \mathcal{T}(\varrho n, D)$ follows the same general strategy. We first split T into small rooted subtrees, so that the roots of which are at even distance from each other. The embedding of T follows a breadth first search order by embedding the small subtrees into some edge from $\mathcal{M}$ or $\mathcal{H}$. Notice that we may use the clusters of $\mathcal{M}$ in a balanced way because X is connected to both sides of each edge of $\mathcal{M}$, and thus we may use almost all the vertices covered by $\mathcal{M}$. The main problem that appears is that the bipartition of the subtrees might be unbalanced. This may be problematic because the strategy of embed the roots in X imply that the vertices of T that are embedded in $\mathcal{Y}$ are all in the same colour class, in which case might be impossible to use up almost all vertices in $\mathcal{Y}$, as we can run out of space in $\mathcal{Z}$. We solve this problem by assigning trees to $\mathcal{Y}$ so that we always use up more vertices in $\mathcal{Y}$ than in $\mathcal{Z}$. Therefore, if a vertex of $Y \in \mathcal{Y}$ had no neighbours with spare room to embed a subtree, this would imply that we would have filled at least $2|N_{\mathcal{H}}(Y)|$ clusters of $\mathcal{H}$. The minimum degree of R' is then enough to guarantee that we can go on with the aforementioned strategy.

The remainder of the paper is organised as follows. In Section 2 we give an outline of the proof of Theorem 1.1. In Section 3 we state a series of results regarding tree embeddings in expander graphs, and then we prove Theorem 1.2 in Section 4. In Section 5 we recall the sparse regularity lemma and some basic facts about the random graph. We prove Theorem 1.3 in Section 6, and then, putting everything together, we prove Theorem 1.1 in Section 7. Finally, we leave some remarks and open questions to Section 8.

2. Sketch of proof of Theorem 1.1 assuming Theorem 1.3

In this section, we outline the proof of Theorem 1.1 by dividing it in three pieces. First consider a *typical* outcome of $G = G(N, p)$ and any given blue-red colouring of its edges with no blue copies of K_{r+1} or red copies of every tree in $\mathcal{T}(n, D)$. In particular, by Theorem 1.3 we have that $e(G_R) \leqslant (1/r + o(1))e(G)$ and consequently that

$$e(G_B) \geqslant \left(1 - \frac{1}{r} - o(1)\right) e(G). \tag{2}$$

In other words, in this scenario the blue graph has asymptotically the same number of edges as the intersection of the Turán graph with G.

By a result of Conlon and Gowers [9] and of Schacht [34], with high probability, every subgraph G with as many edges as in (2) either contains a K_{r+1} or it is almost r-partite. Since we assumed that G_B is K_{r+1}-free, then there exists a partition $V(G) = V_1' \cup \cdots \cup V_r'$ with $o(pN^2)$ blue edges within the parts.

This part of the argument is captured by Proposition 7.2 and by its statement it is possible to conclude Theorem 1.1 if $N \geqslant rn + o(n)$, since at least one of the V_i' would have more than $(1 + o(1))n$ vertices. However, in order to prove Theorem 1.1 we need to push further this stability argument, which is the second part of the proof.

We first define V_0 as the set of vertices with $\Omega(pN)$ blue neighbours inside the part to which it belongs, together with those vertices with $o(pN)$ neighbours in any of the parts. One can show that there are at most $o(N)$ of the first kind and at most $O(1/p)$ of the second. By setting $V_i = V_i' \setminus V_0$, we get a new partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$, with $|V_0| = o(N)$ and such that for each $i \in [r]$ every vertex $v \in V_i$ has $\Omega(pN)$ red neighbours and only few blue neighbours, both in V_i.

We may show that the red graphs induced in each part is an expander graph, which roughly means that every set has a large red external neighbourhood. In particular, we show that these red graphs satisfy the hypothesis of a theorem of Haxell [17], which turns to imply that for every bounded degree tree T of size $(1 - o(1))|V_i|$ and for every $v \in T$ and $u \in V_i$ there exists an embedding of T in $G_R[V_i]$ that maps v to u. This already implies that $|V_i| = (1 + o(1))n$ for every $i \in [r]$ and also that there are no red edges between different parts. To see why the second property is true, we point out that Haxell's Theorem allows us to map a particular vertex of the tree into a particular vertex of the host graph. Suppose that there exists a red edge between different parts, say V_1 and V_2. We can split the tree in two subtrees and then, by using Haxell's Theorem, we may embed the tree by mapping one part of the tree into V_1 and the other part into V_2 and then use the red edge connecting V_1 and V_2 to complete the embedding of the tree.

At this point, if some vertex in V_0 has $\Omega(pN)$ blue neighbours in each of the other r parts, then the fact that all edges between them are blue would yield a blue copy of K_{r+1}, by Janson's inequality. On the other hand, if a vertex in V_0 has $\Omega(pN)$ red neighbours in more than one part, then, in a similar way as mentioned above, Haxell's Theorem would yield all trees of $\mathcal{T}(n, D)$ in red. Therefore, for every vertex $v \in V_0$ (with $\Omega(pN)$ neighbours in each part) there is exactly one $i \in [r]$ in which v has $\Omega(pN)$ red neighbours.

By repeating this argument, we may reallocate vertices from V_0 into some of the V_i's without losing the expansion properties of the red graphs. More precisely, all but $O(1/p)$ vertices from V_0 are left and thus we have to deal with the problem of embedding trees inside expander graphs of size $n + \Omega(1/p)$, namely the largest of the V_i's (let us say V_1). This is the final aspect of the proof of Theorem 1.1.

For trees with less than $n/(\log^3 n)$ leaves, we use a result of Montgomery [28] that in fact yields embeddings of this class of spanning trees in expander graphs. In the case of trees with many leaves, previous results do not fit in our context. However, we may use an intermediate result in the proof of Haxell's Theorem, regarding how to extend an embedding of a tree by adding a new leaf. In fact, it yields not only the versatility of choosing a vertex from which the embedding starts, but also gives sufficient conditions to extend an embedding of a tree T' by adding a leaf. The first condition is that the host graph has "good" enough expansion properties. The second is that the embedding T' was done in a "good" way, which roughly means that the expansion of the host graph is not concentrated in the image of T' by the embedding.

However, this strategy reaches the following barrier in our context. There might be disjoint sets of sizes $\omega(1/p)$ and $n/(\log^3 n)$, respectively, with no edges of $G(N, p)$ between them. To see why this is an impediment, let $T \in \mathcal{T}(n, D)$ be a tree with at least $n/(\log^3 n)$ leaves and let $T' \subseteq T$ be the tree obtained by removing the leaves from T. If there exists an embedding of T' in $G_R[V_1]$, then we can extend it to an embedding of T if and only if one can guarantee a Hall-type condition in the bipartite graph induced by the image of the parents of the leaves in V_1 and the unused vertices. However, since this graph might have $\omega(1/p)$ isolated vertices and we have only $O(1/p)$ "extra" vertices, then we need some additional work.

We deal with this problem beforehand in the proof of Theorem 3.4 in Section 3. The basic idea is to choose a random set $R \subseteq V_1$ of size roughly $n/(\log^3 n)$. We prove that there exists a realisation of R such that for every set $X \subseteq V_1$ of size $\Omega(1/p)$ and for every set $Y \subseteq R$ of size $n/(\log^3 n)$ there is at least one edge between them. With some additional work, we show that we cannot only embed T' in H, but we may even require that the parents of the leaves are embedded in R and then we can apply Hall's Theorem to finish the proof.

3. Trees in expanders

For a graph H and $X \subseteq V(H)$ we will denote by $N(X)$ the external neighbours of X, that is, the set of vertices in $V(H)\backslash X$ having at least one neighbour in X. In this section, we study the family of graphs H in which every subset of $V(H)$ has a large external neighbourhood and we call this family of graphs *expanders*. The notion of expander graphs has a plentiful number of applications in combinatorics and it is particularly useful for embedding trees. To see why, notice that if we have found a copy of a tree T' in H and we want to extend this embedding to a tree $T \supseteq T'$, we only need to look for the external neighbours of the image of T' in H. Following this idea, and extending Pósa's rotation-extension technique, Friedman and Pippenger [14] proved that if, for some integers m and D, a graph H satisfies

$$|N(X)| \geqslant (D+1)|X| \text{ for all } X \subseteq V(H) \text{ with } 1 \leqslant |X| \leqslant 2m,$$

then H contains all trees with m vertices and maximum degree D. A limitation of this result is that it only works for trees of size at most $|V(H)|/(2D+2)$. In a successful attempt to overcome this, Haxell [17] considered different expansion notions for sets of different sizes and proved the following theorem.

Theorem 3.1. *Let $D, m, t \in \mathbb{N}$ and let H be a graph with the following properties:*

(1) $|N(X)| \geqslant D|X| + 1$, for all $X \subseteq V(H)$ with $1 \leqslant |X| \leqslant m$.

(2) $|N(X)| \geqslant t + D|X| + 1$, for all $X \subseteq V(H)$ with $m \leqslant |X| \leqslant 2m$.

Then H is $\mathcal{T}(t, D)$-universal. Moreover, given $v \in V(H)$, $T \in \mathcal{T}(t, D)$ and $u \in V(T)$, there exists an embedding of T that maps u to v.

A different and convenient way of phrasing property (2) of Theorem 3.1 is the following. Let H be a graph such that every pair of disjoint sets $X, Y \subseteq V(H)$, with $|X| = m_1$ and $|Y| = m_2$, satisfies $e(X,Y) > 0$. Then for every $Z \subseteq V(H)$, with $m_1 \leqslant |Z| \leqslant 2m_1$, there are at most $m_2 - 1$ vertices in the non-neighbourhood of Z, since between these two sets there are no edges. By discounting the non-neighbours and the vertices in Z, we get that

$$|N(Z)| \geqslant |V(H)| - |Z| - m_2 + 1. \tag{3}$$

Therefore, when $|V(H)| - m_2 \geqslant t + 2(D+1)m_1$ we recover property (2). The main result of this section considers the case where m_1 and m_2 have different orders of magnitude, which leads us to the following definition.

Definition 3.2. *Let D, m_1, m_2 be integers. A graph H is a (m_1, m_2, D)-expander if*

(i) $|N(X)| \geqslant D|X| + 1$ for all $X \subseteq V(H)$ with $1 \leqslant |X| \leqslant m_1$, and

(ii) $e(X,Y) > 0$ for all disjoint sets $X, Y \subseteq V(H)$ with $|X| = m_1$ and $|Y| = m_2$.

Moreover, if only property (ii) *holds, then we say that H is a weak (m_1, m_2)-expander. We will often omit D when it is clear from context.*

As is usual with tree embedding problems, we deal separately with trees with many or few leaves. Using the Absorption Method, Montgomery [28] proved the following.

Theorem 3.3. *Let n be sufficiently large and D an integer, and set $d = D\log^4 n/20$. If H is a $(n/2d, n/2d, d)$-expander on n vertices, then H contains every tree $T \in \mathcal{T}(n, D)$ with at most n/d leaves.*

We remark that although Theorem 3.3 is not stated explicitly in [28], it follows directly from Montgomery's proof, see [28, Section 4.2], where he only used the fact that that $G(n, p)$ is an expander as in Theorem 3.3. The main result of this section deals with the case of (non-spanning) trees with many leaves.

Theorem 3.4. *Let m_1, m_2, n, D be positive integers satisfying $m_1 \leqslant m_2$, $6m_1 \log n \leqslant m_2$ and $16Dm_2 \leqslant n$, and assume that n is sufficiently large. Let H be a graph with n vertices such that H is*

(1) a weak $(m_1, n/32D)$-expander, and

(2) a weak (m_2, m_2)-expander.

Then H contains every tree $T \in \mathcal{T}(n - m_1, D)$ with at least $24Dm_2$ leaves.

The proof of Theorem 3.1 in [17] relies on a clever inductive argument in order to embed all vertices but the leaves, and then on a Hall-type theorem to finish the embedding. However, the hypothesis of Theorem 3.4 does not enable a straightforward modification of this proof, for the following reason. Given a tree T, let $L \subseteq V(T)$ be the set of leaves of T and let $P = N(L)$ be their parents. Note that if $T \in \mathcal{T}(n-m_1, D)$ is a tree with $|L| \geqslant 24Dm_2$ leaves, then $|P| \geqslant 24m_2$. Suppose we want to embed T. Since the image of P in an embedding of $T \setminus L$ could have $m_2 - 1$ non-neighbours, we would only expect to find trees of size $n - m_2 + 1$.

We address this obstacle by finding a set $W \subseteq V(H)$ with $\Theta(m_2)$ vertices with the property that every subset $X \subseteq W$ with $|X| = m_2$ has less than m_1 non-neighbours in H. We prove that such set exists by selecting one at random. We then manage to find an embedding $f : V(T) \setminus L \to V(H)$ such that $f(P) \subseteq W$. In this case we would already have, for $X \subseteq f(P)$ with $|X| \geqslant m_2$, that

$$|N(X) \setminus f(V(T) \setminus L)| \geqslant n - |V(T) \setminus L| - m_1 + 1 > |L|.$$

Nevertheless, is also necessary to guarantee that small subsets of $f(P)$ have enough neighbours in the set of unused vertices. Motivated by this, we need to define a *good* embedding of a tree. Basically, we say that an embedding of a tree T is good if the set of used vertices satisfies a Hall-type condition (for small sets) to the set of unused vertices.

Definition 3.5. *Let m be a positive integer, let $T \in \mathcal{T}(n, D)$ and let $H = (V_1 \cup V_2; E)$ be a bipartite graph. We say that an embedding $f : V(T) \to V(H)$ is m-good in H if for every*

$i \in \{1,2\}$ and $X \subseteq V_i$, with $1 \leqslant |X| \leqslant m$, we have

$$|N_H(X) \setminus f(V(T))| \geqslant \sum_{v \in f^{-1}(X)} \big(D - d_T(v)\big) + D|X \setminus f(V(T))|.$$

In the previous definition we considered H as being bipartite for technical reasons. More specifically, since we want to embed the set P into a set W, then we have to alternate the embedding of vertices of T between W and $V(H) \setminus W$, so it is easier to consider H as being a bipartite graph. The next lemma gives sufficient conditions to extend good embeddings and it was proved in [1] as the induction step[2] in the proof of a bipartite analogue of Theorem 3.1 (see Theorem 6.6).

Lemma 3.6. *Let m, n, D be positive integers, let $T \in \mathcal{T}(n, D)$ and let $H = (V_1 \cup V_2; E)$ be a bipartite graph. Suppose that there exists a m-good embedding f in H. Moreover, assume that for every $i \in \{1,2\}$ and $X \subseteq V_i$, with $m \leqslant |X| \leqslant 2m$, we have*

$$|N_H(X) \setminus f(V(T))| \geqslant 2Dm + 2. \tag{4}$$

Then for every vertex $v \in T$, with $d_T(v) < D$, there exists a m-good embedding of the tree obtained by adding a leaf to v in T.

Now we only need to define bipartite expanders, which first appeared in [1]. We want such graphs to have two properties: that the embedding of any single vertex tree into any vertex is good, for the base case of the induction; and that condition (4) is satisfied whenever the tree is small enough.

Definition 3.7. *Let $D \geqslant 2$ and let H be a bipartite graph with colour classes V_1 and V_2, where $|V_1| \leqslant |V_2|$. Let m be a positive integer with $m < |V_1|$. We say that H is a bipartite (m, D)-expander if the following two properties hold.*

(1) For $i \in \{1,2\}$, every set $X \subseteq V_i$ of size at most m satisfies $|N_H(X)| \geqslant D|X|$.
(2) For every pair of sets $X_1 \subseteq V_1$ and $X_2 \subseteq V_2$, each of size at least m, we have $e(X_1, X_2) > 0$.

Note that property (2) implies that for every $X \subseteq V_i$, with $|X| \geqslant m$, we have that

$$|N(X)| \geqslant |V_{3-1}| - m + 1.$$

This will guarantee that (4) holds for any embedding of trees with small enough colour classes. Now we are ready to go through one of the main aspects of the proof of Theorem 3.4, which is the embedding of the parents of the leaves into a specific set.

[2]Under the hypothesis Theorem 7, they state that good embeddings can be extended as "Property 2" in page 6. Moreover, the only place where they use the size of neighbours of sets with more than m vertices is in the proof of Claim 8. One can check that (4) is enough to get the same proof.

Lemma 3.8. *Let m, D be positive integers and let $T \in \mathcal{T}(n, D)$. Let $U_1 \cup U_2$ be any partition of one the colour classes of T and let U_3 be the other colour class. Let H be a graph with disjoint subsets $V_1, V_2, V_3 \subseteq V(H)$ such that $|V_i| \geqslant |U_i| + 3Dm$ for $i \in \{1, 2, 3\}$. If $H[V_1, V_3]$, $H[V_2, V_3]$ and $H[V_1 \cup V_2, V_3]$ are bipartite (m, D)-expanders, then there exists a m-good embedding $f : V(T) \to V(H)$ such that $f(U_i) \subseteq V_i$ for $i \in \{1, 2, 3\}$.*

The strategy of the proof of Lemma 3.8 is to iteratively apply Lemma 3.6 in order to extend the embedding of a subtree by adding a leaf at each step. Since we will alternate between mapping vertices into V_1, V_2 and V_3, we will need to keep track that the embeddings are "good" in the graphs $H[V_1, V_3]$, $H[V_2, V_3]$ and $H[V_1 \cup V_2, V_3]$, respectively. This will guarantee that at any stage of the embedding, subsets of $V_1 \cup V_2$ have enough neighbours in the unused vertices of V_3, and that subsets of V_3 have enough neighbours in the unused vertices of both V_1 and V_2. To keep track of this at each step of the embedding, we introduce the following notion.

For the proof of Lemma 3.8 we say that an embedding f of a subtree $S \subseteq T$ is *great* if

(I) $U_i \cap V(S)$ is mapped to V_i, for $i \in \{1, 2, 3\}$.

(II) f is m-good in $H[V_1 \cup V_2, V_3]$ and in $H[V_i, V_3] \cup H[f(V(S))]$, for $i \in \{1, 2\}$.

Proof of Lemma 3.8. The proof follows an inductive argument. Indeed, we start by showing that there exists a great embedding of any single vertex subtree $S \subseteq T$

Claim 3.9. *Let $S \subseteq T$ be a single vertex subtree. If $f : V(S) \to V(H)$ is an embedding which satisfies property* (I)*, then f is great.*

Proof of Claim. Let us prove only that f is m-good in $H[V_1, V_3]$, as the other case are completely analogous. Since $H[V_1, V_3]$ is a bipartite (m, D)-expander, then for $X \subseteq V_1$, with $1 \leqslant |X| \leqslant m$, we have

$$|(N(X) \cap V_3) \setminus f(V(S))| \geqslant D|X| + 1 - 1 = D|X|,$$

which is larger then the required lower bound in the definition of m-good. Since we have the same bound for $X \subseteq V_3$, it follows that f is m-good in $H[V_1, V_3]$. □

Now that we have proved the base case, we prove if $f : V(S) \to V(H)$ is great, then for any tree $S \subseteq T' \subseteq T$, obtained by adding a leaf to S, there exists a great embedding $f' : V(T') \to V(H)$ that extends f. Let v be this leaf. We deal separately with the cases of $v \in U_3$ or $v \in U_1 \cup U_2$. If $v \in U_3$, then we use Lemma 3.6 to extend f to a m-good embedding in $H[V_1 \cup V_2, V_3]$. Indeed, since $H[V_1 \cup V_2, V_3]$ is a (m, D)-expander, then for $X \subseteq V_1 \cup V_2$ (and analogously for $X \subseteq V_3$), with $m \leqslant |X| \leqslant 2m$, we have that

$$\begin{aligned} |(N(X) \cap V_3) \setminus f(V(S))| &\geqslant |N(X)| - |U_3| \\ &\geqslant |V_3| - m - |U_3| \qquad (5) \\ &\geqslant 3Dm - m \geqslant 2Dm + 2, \end{aligned}$$

and thus we get a m-good embedding f' of T' in $H[V_1 \cup V_2, V_3]$. We argue that f' is m-good in $H[V_i, V_3]$, for $i \in \{1, 2\}$. Indeed, for sets $X \subseteq V_i$, we already have the lower bound on $|(N(X) \cap V_3) \setminus f(V(S))|$, since f' is m-good in $H[V_1 \cup V_2, V_3]$. For sets $X \subseteq V_3$, we have nothing to prove, since f was great and we did not use any additional vertices from either V_1 or V_2. Moreover, note that the required lower bound on the definition of a m-good embedding gets weaker as we embed more vertices.

The case when $v \in U_1$ (or analogously in U_2) is similar, but we apply Lemma 3.6 to f in the bipartite graph $H[V_1, V_3] \cup H[f(V(S))]$, together with the same calculation as in (5), to get a m-good embedding f'. Note that the embedding of v does not take any vertex from V_2 or V_3. This guarantees that f' is m-good in $H[V_2, V_3] \cup H[f(V(S))]$. Moreover, for $H[V_1 \cup V_2, V_3]$, we only need to guarantee the lower bound on $N(X) \setminus f(V(T'))$ for $X \subseteq V_3$ with $1 \leqslant |X| \leqslant m$. We finish the proof by pointing out that since f' is m-good in $H[V_1, V_3] \cup H[f(V(S))]$ and since

$$|(N(X) \cap (V_1 \cup V_2)) \setminus f(V(S))| \geqslant |(N(X) \cap V_1) \setminus f(V(S))|,$$

then f' is m-good in $H[V_1 \cup V_2, V_3]$. $\square$

Now we are ready to prove Theorem 3.4.

Proof of Theorem 3.4. Let L be a set of $12Dm_2$ leaves of T in the same colour class and let U_1 be the parents of L in T. Note that $12m_2 \leqslant |U_1| \leqslant 12Dm_2$. Now we choose, uniformly at random, a set $W \subseteq V$ with $r = |U_1| + 4Dm_2 \leqslant 16Dm_2 \leqslant h$ vertices. For each set $X \subseteq V$ with m_1 vertices, let $Z_X = \{y \in W \setminus X : d(y, X) = 0\}$. Since H is a $(m_1, h/72D)$-expander, then

$$\mathbb{E}|Z_X| \leqslant \frac{r}{h} \cdot \frac{h}{32D} \leqslant \frac{m_2}{2}.$$

By standard tail bounds for the hypergeometric distribution (see Theorem 2.10 in [19]), we have

$$\mathbb{P}(|Z_X| \geqslant m_2) \leqslant \exp\left(-\frac{m_2}{6}\right).$$

Denoting by Z the number of sets $X \in \binom{V}{m_1}$ such that $|Z_X| \geqslant m_2$, we have

$$\mathbb{E}[Z] \leqslant h^{m_1} \exp(-m_2/6) < 1,$$

since $m_2 > 6m_1 \log h$. This implies that there is a realisation of W, denoted by W_1, such that every set $X \subseteq V$ of size m_1 has less than m_2 non-neighbours in W_1. Set $T' = T \setminus L$ and suppose that one of the colour classes of T' is $U_1 \cup U_2$ and the other is U_3. We take sets W_2 and W_3 with $|W_i| = |U_i| + 4Dm_2$ for $i \in \{2, 3\}$, which is possible since in this case we have

$$|W_1| + |W_2| + |W_3| = |T'| + 12Dm_2 = h - |L| + 12Dm_2 = h.$$

Claim 3.10. *For $i \in \{1, 2, 3\}$ there exist $V_i \subseteq W_i$, with $|W_i \setminus V_i| \leqslant 2m_2$, such that the graphs $H[V_1 \cup V_2, V_3]$, $H[V_1, V_3]$ and $H[V_2, V_3]$ are bipartite (m_2, D)-expanders.*

Proof of Claim 3.10. Since H is a weak (m_2, m_2)-expander, then the second property of the bipartite expansion is already satisfied for all three bipartite graphs. We will find the sets V_i's iteratively. So, set $X_i = \emptyset$ and $W_i := V_i$ for $i \in \{1, 2, 3\}$ and while there exist

- $X \subseteq W_3$ with $|X| \leqslant m$ and $|N(X) \cap W_i| \leqslant D|X|$ for some $i \in \{1, 2\}$,
  set $X_i := X_i \cup X$ and $W_3 := W_3 \setminus X$, and
- $X \subseteq W_1 \cup W_2$ with $|X| \leqslant m$ and $|N(X) \cap W_3| \leqslant D|X|$,
  set $X_3 := X_3 \cup X$ and $W_i := W_i \setminus X$ for $i \in \{1, 2\}$.

First, we show that at each step in this algorithm we have for $i \in \{1, 2, 3\}$ that $|X_i| \leqslant m_2$ and that

$$|N(X_3) \cap W_3| \leqslant D|X| \quad \text{and} \quad |N(X_i) \cap W_i| \leqslant D|X|$$

for $i \in \{1, 2\}$. Indeed, if this is satisfied for some X_1, X_2, X_3 and there exists $X \subseteq V_1 \cup V_2$ (or analogously for $X \subseteq V_3$) with $|N(X) \cap W_3| \leqslant D|X|$, then clearly we have that

$$\begin{aligned} |N(X_3 \cup X) \cap V_3| &\leqslant |N(X_3) \cap W_3| + |N(X) \cap W_3| \\ &\leqslant D|X| + D|X| = D|X_3 \cup X|. \end{aligned}$$

If we had that $|X| \geqslant m_2$, then by the weak (m_2, m_2)-expansion of H, then X would have less than m_2 non-neighbours in V_3 and therefore we would have that

$$|N(X) \cap V_3| \geqslant |V_3| - m_2 \geqslant 2Dm_2 \geqslant D|X| + 1.$$

This finishes the proof of claim since $|X_1 \cup X_2|, |X_3| \leqslant 2m_2$. $\square$

Since $H[V_1 \cup V_2, V_3]$, $H[V_1, V_3]$ and $H[V_2, V_3]$ are bipartite (m_2, D)-expanders and since

$$|V_i| \geqslant |U_i| + 4Dm_2 - 2m_2 \geqslant |U_i| + 3Dm_2,$$

then we get a m_2 good embedding T' such that U_i is mapped to V_i, for $i \in \{1, 2, 3\}$. Let f be this embedding and denote by $A = f(P)$ the image of the parents of L and $B = V \setminus f(V(T'))$. To embed L, we will use the following well-known generalisation of Hall's Theorem in the bipartite graph $H[A, B]$

For a sequence of non-negative integers $(d_a : a \in A)$, $H[A, B]$ contains a forest F such that $d_F(a) = A$ for $a \in A$ and $d_F(b) \leqslant 1$ for $b \in B$ if and only if

$$|N(X) \cap B| \geqslant \sum_{x \in X} d_x \text{ for all } X \subseteq A. \tag{6}$$

Note that (6) is satisfied for $X \subseteq A$ with $|X| \leqslant m_2$, since f is m_2-good. Moreover, since $A \subseteq U_1$, then by the choice of U_1 every $X \subseteq A$, with $|X| \geqslant m_2$, has less than m_1 non-neighbours and therefore

$$|N(X) \cap B| \geqslant |B| - m_1 \geqslant |L|,$$

which finishes the proof.

$\square$

4. Proof of Theorem 1.2

The proof of Theorem 1.2 follows by applying Proposition 4.2 $r+1$ times. For an appropriate choice of m_1 and m_2 there will be two possibilities. If the red graph is a weak (m_1, m_2)-expander, then we show that it is $\mathcal{T}(n, D)$-universal, using Theorem 3.1. Otherwise it will contain two disjoint sets of size m_1 and m_2 with all edges in between coloured in blue. We repeat this argument r times in the induced graph on the set with m_2 vertices. At the end, if the red graph is not $\mathcal{T}(n, D)$-universal, then we get $r+1$ disjoint sets, each of size m_1, with all the edges in between coloured in blue. This reasoning is made precise in the proof of the following lemma.

Lemma 4.1. *Let n, m, r, D be positive integers and let H be a graph on $N = rn + 10Drm$ vertices. Then one of the following holds:*

(1) H is $\mathcal{T}(n, D)$-universal.
(2) There are disjoint sets $U_1, \ldots, U_{r+1} \subseteq V(H)$, each of size m, such that $e(V_i, V_j) = 0$ for $1 \leqslant i < j \leqslant r+1$.

Before proving Lemma 4.1 we deal with the technical part of showing the weak expanders are almost expanders.

Proposition 4.2. *Let D, m_1, m_2 be integers and let $H = (V, E)$ be a graph with $|V| \geqslant m_2 + (2D+2)m_1$. If H is a weak (m_1, m_2)-expander, then there exists a set $V' \subseteq V$, with $|V \setminus V'| \leqslant m_1$, such that $H[V']$ is a (m_1, m_2)-expander.*

Proof. Take a maximal set $Z \subseteq V$ with $|Z| < m_1$ and $|N(Z)| \leqslant D|Y|$, and set $V' = V \setminus Z$. We will prove that for any $X \subseteq V'$ with $|N(X) \cap V'| \leqslant D|X|$ we have that $|X| > m_1$, which shows that $H[V']$ is a (m_1, m_2)-expander. For such X we have that

$$|N(Z \cup X)| < D|Z \cup X|,$$

since we are only counting external neighbours of $Z \cup X$. By the maximality of Z, we conclude that $|Z \cup X| \geqslant m_1$. Since H is a weak (m_1, m_2)-expander, then there are less than m_2 non-neighbours of $Z \cup X$ in V'. Therefore

$$\begin{aligned}
D|X| &\geqslant |N(X) \cap V'| \\
&\geqslant |N(Z \cup X) \cap V'| - |N(Z) \cap V| \\
&> |V'| - m_2 - |X| - D|Z| \\
&\geqslant |V| - m_2 - (D+1)m_1 - |X| \geqslant (D+1)m_1 - |X|,
\end{aligned}$$

which implies that $|X| > m_1$ and finishes the proof. $\square$

Now we move to the proof of Lemma 4.1.

Proof of Lemma 4.1. We assume that H is not $\mathcal{T}(n,D)$-universal and set $V_0 = V(H)$. We will prove that for $s \in [r]$ there exist disjoint sets U_s, V_s, with

$$|U_s| = m \quad \text{and} \quad |V_s| = (r-s)n + (r-s+1)5Dm,$$

such that $e(U_s, V_s) = 0$ and $U_s, V_s \subseteq V_{s-1}$. Indeed, if this is true, we set $U_{r+1} = V_r$ and get that $e(U_i, U_j) = 0$ for every $1 \leqslant i < j \leqslant r+1$, which is what we want to prove.

Suppose we have sets $V_0, U_1, V_1, \cdots, U_s, V_s$ as above for $s \in [r]$, or just V_0 for $s = 0$. Let $m_s = (r-s-1)n + (r-s)5Dm$. We show that if $H[V_s]$ were a weak (m, m_s)-expander, then it would be $\mathcal{T}(n,D)$-universal, which we assumed not to be true. To prove that, first we check that

$$|V_s| - m_2 \geqslant n + 5Dm.$$

In particular, $|V_s| \geqslant (D+2)m + m_s$, which is the requirement to apply Proposition 4.2. Therefore, there exists $V_s' \subseteq V_s$ such that $|V_s \setminus V_s'| \leqslant m$ and $H[V_s']$ is (m, m_s)-expander. As reasoned in (3), for sets $X \subseteq V_{r-1}$, with $m \leqslant |X| \leqslant 2m$, the (m, m_s)-expansion implies that

$$\begin{aligned}
|N(X) \cap V_s'| &\geqslant |V_s'| - m_2 - |X| + 1 \\
&\geqslant |V_s| - m - m_2 - 2m + 1 \\
&\geqslant n + 5Dm - 3m + 1 \\
&\geqslant n + 2Dm + 1 \geqslant n + D|X| + 1.
\end{aligned}$$

The above inequality and the first property of (m, m_s)-expansion imply that $H[V_i']$ is $\mathcal{T}(n,D)$-universal, by Theorem 3.1. $\square$

Lemma 4.1 reduces the proof of Theorem 1.2 to finding the minimum value m such that every collection of $r+1$ disjoint m-sets, with high probability, span in $G(N,p)$ a copy of K_{r+1} with one vertex in each set, which we will call a *canonical copy.* To do this we have the following lemma, whose proof is a standard application of Janson's inequality and therefore we omit it.

Lemma 4.3. *Let $r \geqslant 1$ and let $G = G(N,p)$, with $p \gg N^{-2/(r+1)}$. Fix a disjoint collection $V_1, \ldots, V_{r+1} \subseteq V(G)$, with $|V_i| = m_i$. Then the probability that $V_1, \cdots, V_{r+1}$ spans a canonical copy of K_{r+1} is at least*

$$1 - \exp\left(-\Omega\left(p^{\binom{r+1}{2}} \prod_{i=1}^{r+1} m_i\right)\right)$$

In particular, there exists a constant $C > 0$ such that if an integer m satisfies

$$m^{r+1} p^{\binom{r+1}{2}} \geqslant C \log \binom{N}{m}, \tag{7}$$

then with high probability there exists a canonical copy of K_{r+1} in every collection of $r+1$ disjoint m-sets.

Now we may state a stronger version of Theorem 1.2, with $t = O(m)$ and m satisfying (7). Note that when $t = \Omega(N)$, (7) is equivalent to say that $p \geqslant CN^{-2/(r+1)}$, for some $C > 0$.

Theorem 4.4. *For every $r, D \geqslant 2$ and for every $p = p(n)$ and m satisfying* (7)*, if*

$$N \geqslant rn + 10Drm,$$

then $G(N,p) \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$ with high probability.

Proof. Let $G = G(N,p)$, where $N = rn + 10Drm$, and consider the even in which every collection of $r+1$ disjoint sets of size m span a canonical copy of K_{r+1}. By Lemma 4.3 and the hypothesis on m, this happens with high probability. Let G_R, $G_B \subseteq G$ be the red and blue graphs in a given edge colouring of G. By Lemma 4.1, if G_R is not $\mathcal{T}(n,D)$-universal, then there are disjoint sets $U_1, \ldots, U_{r+1}$ of size m such that $e_R(U_i, U_j) = 0$ for all $1 \leqslant i < j \leqslant r+1$. In other words, all the edges in between these sets are coloured blue, which spans a blue copy of K_{r+1}, by the choice of m. □

5. Regularity and facts about the random graph

In this section we state some tools needed for the proof of Theorem 1.1 and Theorem 1.3.

5.1. **The sparse random Erdős–Simonovits stability theorem.** The following result is one of a series of random analogues of extremal results proved, independently, by Conlon and Gowers [9] and by Schacht [34].

Theorem 5.1. *For every $r \geqslant 2$ and $\varepsilon > 0$, there exist positive numbers C' and δ such that if $p \geqslant C'N^{-2/(r+2)}$ then a.a.s. the following holds. Every K_{r+1}-free subgraph G of $G(N,p)$ with*

$$e(G) \geqslant \left(1 - \frac{1}{r} - \delta\right) p \binom{N}{2}$$

can be made r-partite by removing at most $\varepsilon p N^2$ edges.

5.2. **Sparse regularity.** The proof of Theorem 1.3 relies on a sparse version of the Szemerédi Regularity lemma. In order to state this result we need some basic definitions.

Definition 5.2. *Let $\eta, p \in (0,1)$. We say that an n-vertex graph G is (η,p)-uniform, if all disjoint sets $A, B \subseteq V(G)$ with $|A|, |B| \geqslant \eta n$ satisfy*

$$(1-\eta)p|A||B| \leqslant e_G(A,B) \leqslant (1+\eta)p|A||B| \tag{8}$$

and

$$(1-\eta)p\binom{|A|}{2} \leqslant e_G(A) \leqslant (1+\eta)p\binom{|A|}{2}. \tag{9}$$

Furthermore, we say that G is (η,p)-upper-uniform if (possibly) only the upper bounds in (8) *and* (9) *hold for all $A, B \subseteq V(G)$ as above.*

Let G be a graph and let $p \in (0,1)$. Given two disjoint sets $A, B \subseteq V(G)$, we define the p-density of the pair (A, B) by

$$d_p(A,B) = \frac{e(A,B)}{p|A||B|}.$$

Given $\varepsilon > 0$, we say that (A, B) is (ε, p)-regular if for all $A' \subseteq A$ and $B' \subseteq B$, with $|A'| \geqslant \varepsilon|A|$ and $|B'| \geqslant \varepsilon|B|$, we have

$$|d_p(A', B') - d_p(A, B)| \leqslant \varepsilon.$$

Now we state some standard results regarding properties of regular pairs (we refer to the survey [15] for the proofs).

Lemma 5.3. *Given $\alpha > \varepsilon > 0$, let G be a graph and let $A, B \subseteq V(G)$ be disjoint sets such that (A, B) is (ε, p)-regular with $d_p(A, B) = d > 0$. Then the following are true.*

(1) For any $A' \subseteq A$ with $|A'| \geqslant \alpha|A|$ and $B' \subseteq B$ with $|B'| \geqslant \alpha|B|$, the pair (A', B') is $(\varepsilon/\alpha, p)$-regular with p-density at least $d - \varepsilon$.

(2) There are at most $\varepsilon|A|$ vertices in A with less then $(d - \varepsilon)p|B|$ neighbours in B.

A partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$ is said to be (ε, p)-regular if

(1) $|V_0| \leqslant \varepsilon|V(G)|$,
(2) $|V_i| = |V_j|$ for all $i, j \in [k]$, and
(3) all but at most εk^2 pairs (V_i, V_j) are (ε, p)-regular.

We may now state a sparse version of Szemerédi's regularity lemma, due to Kohayakawa and Rödl [20, 23] .

Theorem 5.4. *Given $\varepsilon > 0$ and $k_0 \in \mathbb{N}$, there are $\eta > 0$ and $K_0 \geqslant k_0$ such that the following holds. Let G be an η-upper-uniform graph on $n \geqslant k_0$ vertices and let $p \in (0,1)$, then G admits an (ε, p)-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$ with $k_0 \leqslant k \leqslant K_0$.*

Let G be a graph that admits an (ε, p)-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$. Let $d \in (0,1)$. The (ε, p, d)-reduced graph R, with respect to this (ε, p)-regular partition of G, is the graph with vertex set $V(R) = \{V_i : i \in [k]\}$, called clusters, such that V_iV_j is an edge if and only if (V_i, V_j) is an (ε, p)-regular pair with $d_p(V_i, V_j) \geqslant d$. Next proposition establishes that the edge density of R is roughly the same as in G. Since its proof is fairly standard in the applications of the Regularity Lemma, we omit it.

Proposition 5.5. *Let $\varepsilon, \eta, p, d \in (0,1)$ and let $k \in \mathbb{N}$ such that $k \geqslant 1/\varepsilon$. Let G be an (η, p)-upper uniform graph on n vertices that admits an (ε, p)-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$, and let R be the (ε, p, d)-reduced graph of G with respect to this partition. Then*

$$e(R) \geqslant \frac{e(G)}{(1+\eta)p}\left(\frac{k}{n}\right)^2 - \frac{6\varepsilon + d}{1+\eta}k^2.$$

5.3. **Facts about the random graph.** We state three lemmas concerning properties of $G(N,p)$ and we omit their proofs. The first two follow by fairly simple applications of Chernoff's bound and the third by Janson's inequality.

Lemma 5.6. *For every $\eta > 0$ there exists $C > 0$ such that if $p \geqslant C/N$ then a.a.s. $G(N,p)$ is (η, p)-uniform.*

In particular, since any spanning subgraph of an (η, p)-uniform graph is (η, p)-upper-uniform, then, with high probability, every spanning subgraph of $G(N,p)$ is (η, p)-upper-uniform, as long as $p \geqslant C/N$.

Lemma 5.7. *For every $\gamma > 0$, $G = G(N,p)$ a.a.s satisfies the following properties.*

(*i*) *For every set $U \subseteq V$ with $|U| \geqslant \gamma N$, there are at most $64/\gamma p$ vertices in V with less than $\gamma p N/8$ neighbours in U.*

(*ii*) *For every $c > 0$, there exists $0 < c' < 1$ such that G is a weak $(c/p, c'N)$-expander. Moreover, $c' \to 0$ as $c \to \infty$.*

Lemma 5.8. *For every $\gamma > 0$ there exists $C' > 0$ such that if $p \geqslant C'N^{-2/(r+2)}$, then $G = G(N,p)$ with high probability has the following property. For every $v \in V(G)$ and any r disjoint sets $W_1, \ldots, W_r \subseteq N(v)$, with $|W_i| \geqslant \gamma p N$ for each $i \in [r]$, there exists a copy of K_{r+1} containing v and one vertex in each W_i, for $i \in [r]$.*

6. Global Resilience of Large Trees

This section is devoted to prove the global resilience of trees of linear size and bounded maximum degree in $G(N,p)$. Actually, we will prove the following stronger result.

Theorem 6.1. *Let $\delta, \varrho \in (0,1)$ and $D \geqslant 2$. There are positive constants n_0, η_0 and C_0 such that for all $\eta \leqslant \eta_0$ and $n \geqslant n_0$ the following holds. Let G be a (η, p)-uniform graph on n vertices and let $p \in [0,1]$ with $pn \geqslant C_0$. Then every subgraph $G' \subseteq G$ with $e(G') \geqslant (\varrho + \delta)\, e(G)$ is $\mathcal{T}(\varrho n, D)$-universal.*

It turns out that Theorem 1.3 easily follows from Theorem 6.1. Indeed, given $\delta, \varrho \in (0,1)$ and $D \geqslant 2$, by Lemma 5.6 we know that $G(N,p)$ is, with high probability, (η_0, p)-uniform for $p \geqslant C/N$ and therefore, by Theorem 6.1, any subgraph $G' \subseteq G(N,p)$ with $e(G') \geqslant (\varrho + \delta) e(G(N,p))$ is $\mathcal{T}(\varrho N, D)$-universal.

6.1. **Outline of the proof.** Let G be an (η, p)-uniform graph and let $G' \subseteq G$ be a subgraph of G such that $e(G') \geqslant (\varrho + \delta) e(G)$. Since we obtained G' by removing edges from G, is clear that G' is (η, p)-upper uniform, and therefore, by the regularity lemma (Theorem 5.4), we know that $V(G')$ admits an (ε, p)-regular partition. We will work on the reduced graph R' of G' in order to find a good structure into which any given bounded degree tree can be embedded. Let k be the number of vertices of R'. Since R' inherit the edge density of G', we can show that the average degree of R' satisfies $d(R') \geqslant (\varrho + \delta/3)k$, and thus, by a standard argument, we can find a subgraph $R \subseteq R'$ such that $d(R) \geqslant (\varrho + \delta/3)k$ and has minimum degree $\delta(R) \geqslant (\varrho + \delta/3)k/2$. Let $X \in V(R)$ be a vertex of degree at least the average. Note that $N(X)$ is larger than the size of the tree (scaled by the size of the clusters), and so our plan will be to use the neighbourhood of X in order to embed every tree of size ϱn and bounded maximum degree. We can prove that the neighbourhood of X can be partitioned into a large matching $\mathcal{M}$ and an independent set $\mathcal{Y}$. If we denote by $\mathcal{H}$ the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z} = N(\mathcal{Y}) \setminus (X \cup N(X))$, then by the minimum degree of R we can prove that $\mathcal{Y}$ has large minimum degree in $\mathcal{H}$, as long as $\mathcal{M}$ is not larger than $(\varrho + \delta/16)k$.

Given a tree $T \in \mathcal{T}(\varrho n, D)$, our goal is to embed T using the structure that we have found in the neighbour of X. To do so, we first need to cut the tree into very small subtrees and then locate every such subtree into some edge of the reduced graph. If $\mathcal{M}$ is large enough, then we will locate each subtree into an edge of the matching, using both clusters of the edge in a balanced way. Otherwise, we will first locate subtrees into edges from $\mathcal{H}$, until a large proportion of $\mathcal{Y} \cup \mathcal{Z}$ is used. The leftover subtrees can be located into $\mathcal{M}$, always using both clusters from each edge in a balanced way. In any case, once we have located the subtrees, we will use an embedding technique due to Balogh, Csaba and Samotij [1], in order to embed each of this subtrees into the (ε, p)-regular pair that was assigned to this subtree. The role of X here is to connect the embedding, meaning that X will be used in order to go from one edge to another in $\mathcal{M} \cup \mathcal{H}$.

6.2. **Cutting up a tree.** Now we show how to cut a given tree T into a constant number of tiny rooted subtrees, such that the root of each of this subtrees is at even distance from the root of T. The following lemma, proved by Balogh, Csaba and Samotij [1], gives a partition of the tree into a constant number of subtrees such that each subtree has few vertices and is adjacent to a bounded number of others subtrees.

Lemma 6.2. *Let $D \geqslant 2$ and let (T, r) be a rooted tree with maximum degree at most D. If $\beta \geqslant 1/|V(T)|$, then there exists a family of $t \leqslant 4/\beta$ disjoint rooted subtrees $(T_i, r_i)_{i \in [t]}$ such that $V(T) = V(T_1) \cup \cdots \cup V(T_t)$ and for each $i \in [t]$ we have*

(1) $|V(T_i)| \leqslant D^2 \beta |V(T)|$,
(2) T_i is connected (by an edge) to at most D^3 others subtrees, and
(3) T_i is rooted at r_i and all the children of r_i belong to T_i.

Given a tree T, let $(T_i, r_i)_{i \in [t]}$ be the family given by Lemma 6.2. We may define an auxiliary graph T_Π, called *cluster tree*, with vertex set $V(T_\Pi) = [t]$ and edge set

$$E(T_\Pi) = \{ij \mid \ T_i \text{ and } T_j \text{ are adjacent in } T\}.$$

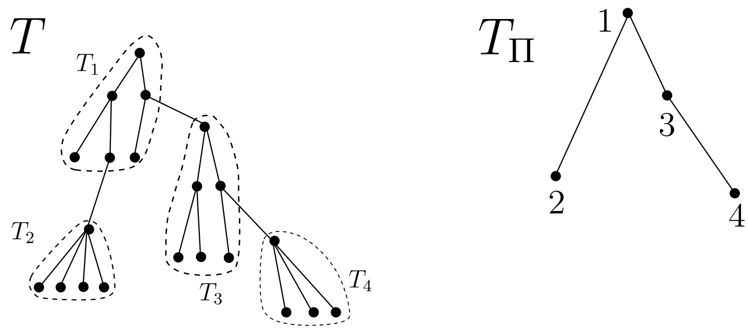

Figure 2. Cluster tree

Now we need to refine the partition given by Lemma 6.2 in order to impose that the root of each subtree is at even distance from the root of T.

Proposition 6.3. *Let $D \geqslant 2$ and let (T, r) be a rooted tree with maximum degree at most D. If $\beta \geqslant 1/|V(T)|$, then there exists a family of $t \leqslant 4D/\beta$ disjoint rooted subtrees $(T_i, r_i)_{i\in[t]}$ such that $V(T) = V(T_1) \cup \cdots \cup V(T_t)$ and for each $i \in [t]$ we have*

(1) $|V(T_i)| \leqslant D^4\beta|V(T)|$,
(2) T_i is rooted at r_i and the distance from r_i to r is even,
(3) all the children of r_i belong to T_i, and
(4) the corresponding cluster tree has maximum degree at most D^4.

Proof. Starting with the partition given by Lemma 6.2, we will refine this partition as we run a breadth first search on (T, r). Suppose that in this search we have reached a vertex v, which is the root of a subtree in the current partition, such that v and all roots before v are at even distance from each other in the current partition.

If there is a root u of some subtree in the current partition, which is at odd distance from v and such that the subtree pending from v is adjacent to u, then we may update the partition by splitting the tree pending from u (each neighbour of u is now the root of a subtree) and adding u to the subtree pending from v. Note that after this splitting, the root of each tree that is adjacent to the tree pending from v is at even distance from all the previous roots.

At the end of this process, each subtree of the original partition is split into at most D parts and hence we end up with at most $4D/\beta$ rooted subtrees. For the same reason, the maximum degree of the cluster tree cannot go higher than D^4. Moreover, the size of each subtree grows by at most D^3 when the roots are added, so at the end of the process each subtree has size at most $D^2\beta|V(T)| + D^3 \leqslant D^4\beta|V(T)|$. $\square$

6.3. **Structure in the reduced graph.** In this subsection, we will follow a strategy inspired in the approach of Besomi, Stein and the third author [4] to the Erdős–Sós conjecture for bounded degree trees and dense host graphs. We will prove that if H is an (η,p)-upper-uniform graph with $2e(H) \geqslant (\varrho+\delta/2)pn^2$, then H has an (ε,p,d)-reduced graph R with a useful substructure. That is, R contains a cluster X of large degree such that its neighbourhood can be partitioned as $N(X) = V(\mathcal{M}) \cup \mathcal{Y}$, where $\mathcal{M}$ is a matching and $\mathcal{Y}$ is an independent set. Moreover, if $\mathcal{H}$ denotes the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z} = N(\mathcal{Y}) \setminus (X \cup N(X))$, then every cluster in $\mathcal{Y}$ has large degree in $\mathcal{H}$.

We need the following lemma (see [3] for a proof).

Lemma 6.4. *Given a graph F, there exists an independent set I, a matching M and a family of triangles Γ, such that $V(F) = I \cup V(M) \cup V(\Gamma)$. Moreover, we may write $V(M) = M_1 \cup M_2$, where each edge $e \in M$ is of the form $e = v_1v_2$ with $v_i \in M_i$ for $i \in \{1,2\}$, so that $N(I) \subseteq M_1$.*

Proposition 6.5. *Let $\varepsilon,\delta,\varrho \in (0,1)$ and let $d = \delta/100$. There exist $n_0, K_0 \in \mathbb{N}$ and $n_0 > 0$ such that for all $0 < \eta \leqslant \eta_0$, $p \in (0,1)$ and $n \geqslant n_0$, the following holds. Let H be an (η,p)-upper uniform graph on n vertices such that $2e(H) \geqslant (\varrho+\delta/2)pn^2$. Then H admits an (ε,p)-regular partition $V(H) = V_0 \cup V_1 \cup \cdots \cup V_k$, with $1/\varepsilon \leqslant k \leqslant K_0$, such that if R is the (ε,p,d)-reduced graph with respect to this partition, then R contains a cluster X, a matching $\mathcal{M}$ and a bipartite subgraph $\mathcal{H}$, with vertex set $V(\mathcal{H}) = \mathcal{Y} \cup \mathcal{Z}$, satisfying the following properties:*

(a) $N(X) = V(\mathcal{M}) \cup \mathcal{Y}$ and $V(\mathcal{M}) \cap \mathcal{Y} = \emptyset$;

(b) $|V(\mathcal{M})| + |\mathcal{Y}| \geqslant (\varrho+\delta/3)\,k$; and

(c) for all $Y \in \mathcal{Y}$ we have

$$|N_{\mathcal{H}}(Y)| \geqslant \left(\varrho + \frac{\delta}{4}\right)\frac{k}{2} - \frac{|V(\mathcal{M})|}{2}.$$

Proof. Given $\varepsilon' = \min\{\varepsilon/5, \delta/1000\}$ and $k_0 = 1/\varepsilon'$, let η_0, n_0' and K_0' be the outputs of the regularity lemma (Theorem 5.4) with parameters ε' and k_0. Setting $n_0 = n_0'$ and $\eta_0 = \min\{\eta_0', \delta/1000\}$, let H be an (η,p)-upper uniform graph on $n \geqslant n_0$ vertices and $0 < \eta \leqslant \eta_0$. Then H admits an (ε',p)-regular partition $V(H) = V_0' \cup V_1' \cup \cdots \cup V_\ell'$, with $1/\varepsilon' \leqslant \ell \leqslant K_0$, and let us denote by R' the $(\varepsilon',p,2d)$-reduced graph of H with respect to this regular partition. By Proposition 5.5 and the bound on $e(H)$ we have

$$e(R') \geqslant (1+\eta)^{-1}\Big(\varrho + \frac{\delta}{2}\Big)\frac{\ell^2}{2} - (1+\eta)^{-1}(6\varepsilon' + 2d)\ell^2 \geqslant \Big(\varrho + \frac{\delta}{3}\Big)\frac{\ell^2}{2}. \tag{10}$$

Note that (10) implies that the average degree of R' is at least $(\varrho+\delta/3)\ell$. Thus, by successively removing vertices of low degree, we may find a subgraph $R_0 \subseteq R'$ such that

$$d(R_0) \geqslant \Big(\varrho + \frac{\delta}{3}\Big)\ell \quad \text{and} \quad \delta(R_0) \geqslant \Big(\varrho + \frac{\delta}{3}\Big)\frac{\ell}{2}.$$

In particular, this implies that there exists a cluster $X' \in V(R_0)$ with degree at least $(\varrho+\delta/3)\ell$ in R_0. Applying Lemma 6.4 to $N_{R_0}(X')$, we find an independent set I, a matching $\mathcal{M}'$ and a collection of triangles Γ that partition $N_{R_0}(X') = I \cup V(\mathcal{M}') \cup V(\Gamma)$, and moreover, by writing $V(\mathcal{M}') = M_1 \cup M_2$ we have that $N_{R_0}(I) \subseteq M_1$. Note that the minimum degree on R_0 implies that for all $Y \in I$ we have

$$|N_{R_0}(Y) \setminus (X' \cup N_{R_0}(X))| \geqslant \left(\varrho + \frac{\delta}{3}\right)\frac{\ell}{2} - 1 - \frac{|V(\mathcal{M})|}{2} \geqslant \left(\varrho + \frac{\delta}{4}\right)\frac{\ell}{2} - \frac{|V(\mathcal{M})|}{2}. \tag{11}$$

If there are no triangles in this decomposition, then we would finish the proof by setting $\mathcal{M} = \mathcal{M}'$ and $\mathcal{H}$ as the bipartite graph induced by I and $N_{R'}(I) \setminus (X \cup N_{R'}(X))$. If is not the case, for each $i \in [\ell]$ we may arbitrarily partition $V_i = V_{i,0} \cup V_{i,1} \cup V_{i,2}$ so that $|V_{0,i}| \leqslant 1$ and $|V_{i,1}| = |V_{i,2}|$. Notting that $|V_{i,1}| = |V_{i,2}| \geqslant |V_i|/3$ for every $i \in [\ell]$, because of Lemma 5.3, for each $V_iV_j \in E(R')$ and $a, b \in \{1, 2\}$ the pair $(V_{i,a}, V_{j,b})$ is (ε, p)-regular with density at least d. Moreover, by setting $V_0 = V_0' \cup V_{1,0} \cup \cdots \cup V_{\ell,0}$ we conclude that $V(H) = V_0 \cup V_{1,2} \cup V_{2,2} \cup \cdots \cup V_{\ell,1} \cup V_{\ell,2}$ is an (ε, p)-regular partition with $2\ell + 1$ parts. Let R be the (ε, p, d)-reduced graph of H with respect to this partition, and let $k = 2\ell$ be the number of vertices of R (note that R is a blow-up of R'). We set X as one of the clusters coming from X', and $\mathcal{Y}$ as the set of all the $V_{i,a}$ such that $V_i' \in I$ and $a \in \{1, 2\}$. Now note that each triangle in Γ can be decomposed as three disjoint edges in R. Then we set

$$\mathcal{M} = \bigcup_{V_iV_j \in \mathcal{M}'} \{V_{i,1}V_{j,1}, V_{i,2}V_{j,2}\} \cup \bigcup_{V_aV_bV_c \in \Gamma} \{V_{a,1}V_{b,1}, V_{b,2}V_{c,1}, V_{c,2}V_{a,2}\}$$

and $\mathcal{Z} = N_R(\mathcal{Y}) \setminus (X \cup N_R(X))$. Letting $\mathcal{H}$ as the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z}$, is clear that X, $\mathcal{M}$ and $\mathcal{H}$ satisfy (a) and (b), (c) follows from (11).

$\square$

6.4. **Proof of Theorem 6.1.** In this subsection we put everything togheter in order to prove Theorem 6.1. As we mentioned in the sketch of the proof, the idea is to use the structure given by Proposition 6.5, that is, the cluster X, the matching $\mathcal{M}$ and the bipartite graph $\mathcal{H}$. To do so, we first need to cut the tree into a family $(T_i, r_i)_{i \in [t]}$ of tiny subtrees such that the root of all the subtrees are in the same color class (see Proposition 6.3). The main idea of the proof is to first assign each T_i to some edge of $\mathcal{M} \cup \mathcal{H}$. After this, we may remove some bad vertices from each cluster that is used, and thus each subtree T_i can be assigned to a pair $(Y_{i,1}, Y_{i,2})$ which induces a bipartite expander graph and that connects well with a large subset of X (see Claim 6.8). Finally, by using an embedding tool due to Balogh, Csaba and Samotij [1], we can embed each subtree into the pair that was assigned to that tree.

The following lemma, proved in [1], gives sufficient expansion conditions for a bipartite graph to contain all trees of a given size. This is the bipartite version of Theorem 3.1, and is useful because it is sensitive to the unbalance of the tree in question.

Lemma 6.6. *Let $D \geqslant 2$ and let H be a bipartite graph with colour classes V_1 and V_2, where $|V_1| \leqslant |V_2|$. Suppose that H is a bipartite $(m, D+1)$-expander with $0 < m < |V_1|/(2D+1)$. Then H contains all trees with maximum degree at most D and colour classes of sizes at most $|V_1| - (2D+1)m$ and $|V_2| - (2D+1)m$ respectively. Furthermore, any such tree can be* *embeddeded even if we require that a particular vertex of the tree is mapped to a particular vertex of H, as long as this mapping respect the colour classes.*

Although is not true that (ε, p)-regular pairs are bipartite expanders (since they can have isolated vertices), any large subgraphs of an (ε, p)-regular pairs contains an almost spanning subgraph which is a bipartite expander. The proof of the following result is similar as the proof of Proposition 4.2 and it was proved in [1].

Lemma 6.7. *Let (A, B) be an (ε, p)-regular pair such that $d_p(A, B) > \varepsilon$. Suppose that $|A| = |B| = m$ and let $A' \subseteq A$ and $B' \subseteq B$ be sets of size at least $(4D+6)\varepsilon m$. Then there are subsets $A'' \subseteq A'$ and $B'' \subseteq B'$ such that*

(a) $|A' \setminus A''| \leqslant \varepsilon m$ and $|B' \setminus B''| \leqslant \varepsilon m$, and

(b) the subgraph induced by (A'', B'') is a bipartite $(\varepsilon m, 2D+2)$-expander.

Now we are ready to prove Theorem 6.1.

Proof of Theorem 6.1. Let n_0', K_0 and η_0 be the outputs of Proposition 6.5 with inputs δ, ϱ and $\varepsilon = \delta^4/(2^{28}D^6)$. We set

$$\beta = \frac{\delta^2}{2^{12}kD^4} \quad \text{and} \quad C_0 = \frac{2^{17}10^2 D^5 K_0^2}{\delta^3}, \tag{12}$$

and let $n_0 = \max\{n_0', \beta^{-1}\}$ and $n \geqslant n_0$. Given $p \geqslant C_0/n$ and $0 < \eta \leqslant \eta_0$, let G be an (η, p)-uniform graph on n vertices and let $G' \subseteq G$ be a subgraph with

$$2e(G') \geqslant (\varrho + \delta)2e(G) \geqslant (1-\eta)(\varrho+\delta)pn^2 \geqslant \left(\varrho + \frac{\delta}{2}\right)pn^2.$$

Since G' is (η, p)-upper uniform, by Proposition 6.5 we may find an (ε, p)-regular partition $V(G') = V_0 \cup V_1 \cup \cdots \cup V_k$, with $1/\varepsilon \leqslant k \leqslant K_0$, such that the $(\varepsilon, p, \delta/100)$-reduced graph R, with respect to this partition, contains a cluster X, a matching $\mathcal{M}$ and a bipartite subgraph $\mathcal{H}$, with vertex set $V(\mathcal{H}) = \mathcal{Y} \cup \mathcal{Z}$, satisfying the conclusions of Proposition 6.5.

Let $T \in \mathcal{T}(\varrho n, D)$ be given. We consider the bipartition of T that assigns colour 1 to the smaller partition class of T and colour 2 to the larger one, and then we choose an arbitrary vertex r in colour 1 as the root of T. We apply Proposition 6.3 to (T, r), with parameter β, obtaining a family $(T_i, r_i)_{i\in[t]}$ of $t \leqslant 4D/\beta$ rooted trees, each of size at most $D^4\beta\varrho n$. Furthermore, each root r_i is at even distance from r and therefore every root has colour 1. For $i \in [t]$, let us write $T_{i,j}$ for the set of vertices of T_i having colour $j \in \{1, 2\}$.

Let m denote the size of the clusters and observe that $m \geqslant (1-\varepsilon)n/k$. The heart of the proof is the following claim.

Claim 6.8. *For each $i \in [t]$, there are sets $(Y_{i,1}, Y_{i,2})$ and $W_i \subseteq X$ such that the following holds.*

(1) *There is an edge* $V_{i,1}V_{i,2} \in \mathcal{M} \cup E(\mathcal{H})$ *such that* $Y_{i,1} \subseteq V_{i,1}$ *and* $Y_{i,2} \subseteq V_{i,2}$*. Moreover, if* $V_{i,1}V_{i,2} \in E(\mathcal{H})$ *then* $V_{i,2} \in \mathcal{Y}$*.*
(2) *For* $\ell \neq i$ *and* $j, j' \in \{1, 2\}$*,* $Y_{i,j} \cap Y_{\ell,j'} = \emptyset$*.*
(3) *For* $j \in \{1, 2\}$*,* $|Y_{i,j}| \geqslant |T_{i,j}| + 13D\varepsilon m$*.*
(4) $G'[Y_{i,1}, Y_{i,2}]$ *is a bipartite* $(\varepsilon m, 2D + 2)$*-expander.*
(5) *Every vertex of* $Y_{i,2}$ *has at least* $\delta pm/(200)$ *neighbours in* W_i*.*
(6) *If* T_ℓ *is a child of* T_i *in the cluster tree, then every vertex of* W_i *has at least* $D + 1$ *neighbours in* $Y_{\ell,2}$*.*

Before proving Claim 6.8, let us show how to use it in order to finish the proof of Theorem 6.1. Assume that we have ordered $[t]$ so that if T_i is below T_ℓ, with respect to the root of T, then $i \leqslant \ell$. Starting with the subtree containing r, we will embed $(T_i)_{i\in[t]}$ following this ordering. Let us denote by φ the partial embedding of T. For every embedded subtree (T_i, r_i) we will ensure that

(a) $\varphi(r_i) \in W_s$ for some $s \leqslant i$, and
(b) $\varphi(T_{i,j} \setminus \{r_i\}) \subseteq Y_{i,j}$ for $j \in \{1, 2\}$.

Suppose we are about to embed a subtree T_ℓ which is a child of some subtree T_i that was already embedded satisfying (a) and (b). Let $v_i \in V(T_i)$ be the parent of r_ℓ and note that v_i is embedded into some vertex $\varphi(v_i) \in Y_{i,2}$ (since v_i is adjacent to r_ℓ and every root has colour 1).

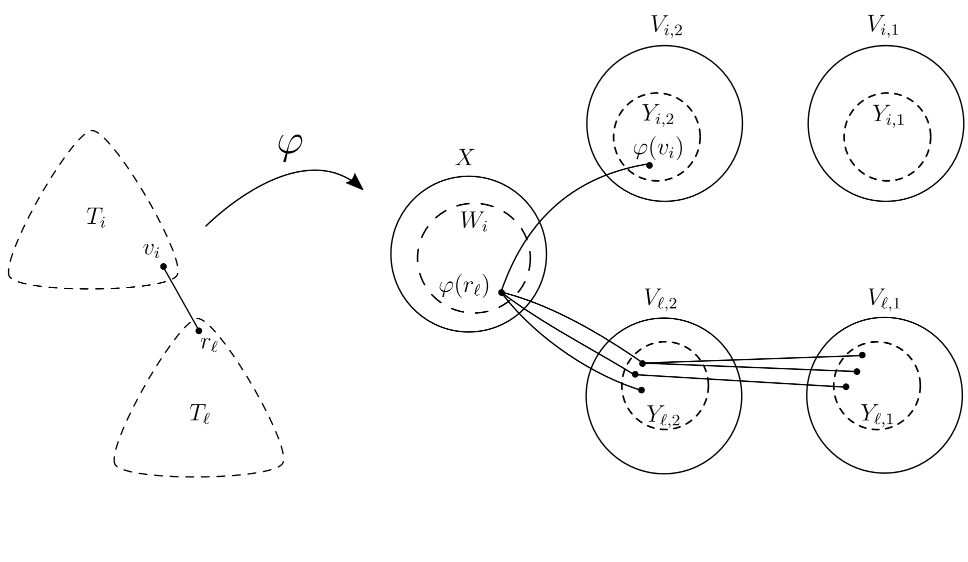

FIGURE 3. Embedding of T_ℓ

Then, because of Claim 6.8 (5)

$$|N_{G'}(\varphi(v_i)) \cap W_i| \geqslant \frac{\delta}{200}pm \geqslant (1-\varepsilon)\frac{\delta C_0}{200k} \geqslant \frac{8D}{\beta} \geqslant 2t$$

and therefore at least one neighbour of $\varphi(v_i)$ has not been used during the embedding. We choose any unused vertex $w_\ell \in W_i \cap N_{G'}(\varphi(v_i))$ and set $\varphi(r_\ell) = w_\ell$ (when we embed T_1, we choose any vertex vetex $w_1 \in W_1$ as the image of $r_1 = r$). By Claim 6.8 (4) we know that $G'[Y_{i,1}, Y_{i,2}]$ is a bipartite $(\varepsilon m, 2D+2)$-expander, we will prove now that

$$G'[Y_{\ell,1} \cup \{w_\ell\}, Y_{\ell,2}] \text{ is a bipartite } (\varepsilon m + 1, D+1)\text{-expander.}$$

Indeed, since $G'[Y_{i,1}, Y_{i,2}]$ is a bipartite $(\varepsilon m, 2D+2)$-expander is easy to see that the expansion conditions hold for every set $X \subseteq Y_{\ell,1} \cup Y_{\ell,2}$. Let $X' \subseteq Y_{\ell,1}$ non-empty and let us consider $X = X' \cup \{w_\ell\}$. If $|X'| \leqslant \varepsilon m$ then we have

$$|N_{G'}(X) \cap Y_{\ell,2}| \geqslant (2D+2)|X'| \geqslant (D+1)|X|,$$

where the first inequality follows because $G'[Y_{\ell,1}, Y_{\ell,2}]$ is bipartite $(\varepsilon m, 2D+2)$-expander. Similarly, if $|X'| \geqslant \varepsilon m$ then we have

$$|N_{G'}(X) \cap Y_{\ell,2}| \geqslant |N_{G'}(X') \cap Y_{\ell,2}| \geqslant |Y_{\ell,2}| - (\varepsilon m + 1).$$

Finally, if $X = \{w_\ell\}$ then by Claim 6.8 (6) we know that $|N_{G'}(w_\ell) \cap Y_{\ell,2}| \geqslant D+1$, and therefore $G'[Y_{\ell,1} \cup \{w_\ell\}, Y_{\ell,2}]$ is a bipartite $(\varepsilon m + 1, D+1)$-expander.

To complete the embedding of T_ℓ, note that because of Claim 6.8 (3) we have

$$|Y_{\ell,j}| - (2D+1)(\varepsilon m + 1) \geqslant |T_{\ell,j}| + 13D\varepsilon m - 6D\varepsilon m \geqslant |T_{\ell,j}|$$

for $j \in \{1,2\}$. Thus, using Lemma 6.6 we may extend φ to T_ℓ, embedding T_ℓ into $(Y_{\ell,1} \cup \{w_\ell\}, Y_{\ell,2})$ so that $\varphi(T_{\ell,j} \setminus \{r_\ell\}) \subseteq Y_{\ell,j}$ for $j \in \{1,2\}$ and w_ℓ is fixed as the image of r_ℓ (we remark that Claim 6.8 (2) allows us to ensure that at every step of the embedding we are using unused vertices). $\square$

Proof of Claim 6.8. Let σ be a permutation on $[t]$ such that for all $1 \leqslant i < j \leqslant t$ we have

$$|T_{\sigma(i),2}| - |T_{\sigma(i),1}| \geqslant |T_{\sigma(j),2}| - |T_{\sigma(j),1}|.$$

Recall that we chose colour 2 for the larger partition class of $V(T)$. Therefore, for every $\ell \in [t]$ we have

$$\sum_{i=1}^{\ell} (|T_{\sigma(i),2}| - |T_{\sigma(i),1}|) \geqslant 0. \tag{13}$$

The proof of Claim 6.8 will be done in two stages. In the first stage, for each $i \in [t]$ the subtree T_i will be assigned to a pair of sets $(X_{i,1}, X_{i,2})$, contained in some edge from $\mathcal{M} \cup E(\mathcal{H})$, such that $|X_{i,j}| = |T_{i,j}| + 16D\varepsilon m$ for $j \in \{1,2\}$. In the second stage, we will remove some vertices from each set in order to find the sets $W_i \subseteq X$ and $Y_{i,j} \subseteq X_{i,j}$ satisfying the properties $(1)-(6)$ from Claim 6.8.

Stage 1 (Assignation): In this stage we will prove that for each $i \in [t]$, there exist an edge $V_{i,1}V_{i,2} \in \mathcal{M} \cup E(\mathcal{H})$ and sets $X_{i,j} \subseteq V_{i,j}$, for $j \in \{1,2\}$, such that

(A) $X_{i,j} \cap X_{\ell,j'} = \emptyset$ if $\{i,j\} \neq \{\ell,j'\}$;
(B) $|X_{i,j}| = |T_{i,j}| + 16D\varepsilon m$; and
(C) if $(V_{i,1}, V_{i,2}) \in E(\mathcal{H})$ then $V_{i,2} \in \mathcal{Y}$.

The assignment will be done in two steps following the order given by σ. At step 1 we assign trees to edges from $\mathcal{H}$ until we use a large proportion of $\mathcal{Y} \cup \mathcal{Z}$, and at step 2 we will use edges from $\mathcal{M}$ ensuring that the clusters from each edge of $\mathcal{M}$ are used in a balanced way.

Step 1: We will assume that $|\mathcal{M}| \leqslant (\varrho + \delta/16)k$, as otherwise we just skip this step. Let us set $Q = (\varrho + \delta/4)k - |V(\mathcal{M})|$ and note that we have

$$|\mathcal{Y}| \geqslant Q \geqslant \frac{\delta}{16}k \quad \text{and} \quad d_{\mathcal{H}}(Y) \geqslant Q/2 \text{ for all } Y \in \mathcal{Y}.$$

We will choose sets in $\mathcal{Y} \cup \mathcal{Z}$ until we have assigned at least $(1-\delta/16)Qm$ vertices to $\mathcal{Y} \cup \mathcal{Z}$. Following the order of σ, assume that we have made the assignation up to some $0 \leqslant \ell \leqslant t-1$ and we are about to assign the tree $T_{\sigma(\ell+1)}$. Suppose that there are $Y \in \mathcal{Y}$ such that

$$\sum_{X_{\sigma(i),2} \subseteq Y} |X_{\sigma(i),2}| \leqslant m - (D^4\beta n + 16D\varepsilon m), \tag{14}$$

and $Z \in N_{\mathcal{H}}(Y)$ with

$$\sum_{X_{\sigma(i),1} \subseteq Z} |X_{\sigma(i),1}| \leqslant m - (D^4\beta n + 16D\varepsilon m). \tag{15}$$

Since $|T_{\sigma(\ell+1)}| \leqslant D^4\beta\varrho n$, we can select sets $X_{\sigma(\ell+1),1} \subseteq Z$ and $X_{\sigma(\ell+1),2} \subseteq Y$, disjoints from the previously chosen sets, such that $|X_{\sigma(\ell+1),j}| = |T_{\sigma(\ell+1),j}| + 16D\varepsilon m$ for $j \in \{1,2\}$. So, if there is no $Y \in \mathcal{Y}$ satisfying (14), then we have

$$\begin{aligned}
\sum_{i=1}^{\ell} |T_{\sigma(i)}| \geqslant \sum_{i=1}^{\ell} |T_{\sigma(i),2}| = \sum_{i=1}^{\ell} \big(|X_{\sigma(i),2}| - 16D\varepsilon m\big) \\
&\geqslant |\mathcal{Y}|m - t \cdot 16D\varepsilon m - k \cdot (D^4\beta n + 16D\varepsilon m) \\
&\geqslant |\mathcal{Y}|m - \frac{\delta^2}{16^2}km \\
&\geqslant \Big(1 - \frac{\delta}{16}\Big)Qm.
\end{aligned}$$

This means that we have already used enough vertices from $\mathcal{Y} \cup \mathcal{Z}$. On the other hand, if every Y satisfying (14) has no neighbours satisfying (15), we may use (13) to deduce

$$\begin{aligned}\sum_{i=1}^{\ell} |T_\sigma(i)| \geqslant 2\sum_{i=1}^{\ell} |T_{\sigma(i),1}| &= 2\sum_{i=1}^{\ell} \big(|X_{\sigma(i),1}| - 16D\varepsilon m\big) \\ &\geqslant 2d_{\mathcal{H}}(Y)m - t\cdot 32D\varepsilon m - k\cdot 2(D^4\beta n + 16D\varepsilon m) \\ &\geqslant Qm - \frac{\delta^2}{16^2}km \\ &\geqslant \Big(1-\frac{\delta}{16}\Big)Qm.\end{aligned}$$

This means that if at step $\ell+1 \in [t]$ we could not find a pair (Y,Z) satisfying (14) and (15), then we have used vertices at least $(1-\delta/16)Qm$ vertices from $\mathcal{Y} \cup \mathcal{Z}$ at step ℓ.

Step 2: Let $0 \leqslant \ell_0 \leqslant t$ be such that $T_{\sigma(1)}, \dots, T_{\sigma(\ell_0)}$ have been assigned to $\mathcal{Y} \cup \mathcal{Z}$, satisfying (A),(B) and (C), and

$$\Big(1-\frac{\delta}{16}\Big)Qm \leqslant \sum_{i=1}^{\ell_0} |T_{\sigma(i)}| \leqslant \Big(1-\frac{\delta}{16}\Big)Qm + D^4\beta\varrho n. \tag{16}$$

Assume that $\ell_0 < t$, otherwise we are done. For $\ell_0+1 \leqslant i \leqslant t$ we will assign each $T_{\sigma(i)}$ to some edge $AB \in \mathcal{M}$. At each step we will ensure that for every edge $AB \in \mathcal{M}$ we have

$$\left| \sum_{X_{\sigma(i),j} \subseteq A} |X_{\sigma(i),j}| - \sum_{X_{\sigma(i),j} \subseteq B} |X_{\sigma(i),j}| \right| \leqslant D^4\beta\varrho n. \tag{17}$$

Suppose we are about to assign a subtree $T_{\sigma(\ell)}$, for some $\ell \geqslant \ell_0+1$, and that (17) holds at step $i = \ell-1$ (note that (17) holds trivially at step ℓ_0). Suppose that there is an edge $AB \in \mathcal{M}$ such that

$$\max\Big\{ \sum_{X_{\sigma(i),j} \subseteq A} |X_{\sigma(i),j}|, \sum_{X_{\sigma(i),j} \subseteq B} |X_{\sigma(i),j}| \Big\} \leqslant m - (D^4\beta\varrho n + 16D\varepsilon m). \tag{18}$$

Assuming that $\sum_{X_{\sigma(i),j} \subseteq A} |X_{\sigma(i),j}| \leqslant \sum_{X_{\sigma(i'),j'} \subseteq B} |X_{\sigma(i'),j'}|$, we let $j^\star = \operatorname*{argmax}_{j\in\{1,2\}} |T_{\sigma(\ell),j}|$ and then we may take sets

- $X_{\sigma(\ell),j^\star} \subseteq A$ with $|X_{\sigma(\ell),j^\star}| = |T_{\sigma(\ell),j^\star}| + 16D\varepsilon m$, and
- $X_{\sigma(\ell),3-j^\star} \subseteq B$ with $|X_{\sigma(\ell),3-j^\star}| = |T_{\sigma(\ell),3-j^\star}| + 16D\varepsilon m$.

disjoints from the previously chosen sets. Note that we have assigned the larger colour class of $T_{\sigma(\ell)}$ to the less occupied cluster in $\{A,B\}$. Furthermore, since (17) holds at step $\ell-1$ and as $|T_{\sigma(\ell)}| \leqslant D^4\beta\varrho n$, the assignment of $T_{\sigma(\ell)}$ implies that (17) holds at step ℓ. So suppose that (18) does not hold at step $\ell-1$ for any $AB \in \mathcal{M}$. Then we have

$$\begin{aligned}\sum_{i=\ell_0+1}^{\ell-1} |T_{\sigma(i)}| &\geqslant |V(\mathcal{M})|m - t\cdot 32D\varepsilon m - k\cdot(3D^4\beta\varrho n + 32D\varepsilon m)\\ &\geqslant |V(\mathcal{M})|m - \frac{\delta}{16}km\end{aligned}$$

that together with (16) yields

$$\begin{aligned}\sum_{i=1}^{\ell-1} |T_{\sigma(i)}| &\geqslant \Big(1-\frac{\delta}{16}\Big)Qm + |V(\mathcal{M})|m - \frac{\delta}{16}km\\ &\geqslant \Big(1-\frac{\delta}{16}\Big)\Big(\varrho+\frac{\delta}{4}\Big)km - \frac{\delta}{16}km\\ &\geqslant \Big(\varrho+\frac{\delta}{8}\Big)km\\ &\geqslant \Big(\varrho+\frac{\delta}{16}\Big)n,\end{aligned}$$

which is impossible since $|T| = \varrho n$. This implies that we can make the assignation for each $\ell \in [t]$.

Stage 2 (Cleaning): Assume that the cluster tree is ordered according to a BFS starting from the subtree which the root of T. Starting with a leaf of the cluster tree, suppose that we have found the sets $Y_{i,j}$ satisfying properties $(1)-(6)$ for all subtrees T_i below T_ℓ in the order of the cluster tree. Let us define

$$W_\ell := \{v \in X : d(v, Y_{i,2}) \geqslant D+1 \text{ for all } i \text{ such that } T_i \text{ is a child of } T_\ell\},$$

we want to prove that W_ℓ has a reasonable size. Given a child T_i of T_ℓ in the cluster tree, we have that

$$|Y_{i,2}| \geqslant |T_{i,j}| + 13D\varepsilon m \geqslant (D+1)\varepsilon m$$

and therefore, since $(X, V_{i,2})$ is (ε, p)-regular, by Lemma 5.3 there are at most $(D+1)\varepsilon m$ vertices in X with less than $D+1$ neighbours in $Y_{i,2}$. Since the auxiliary tree has maximum degree D^4, then W_ℓ has at least

$$|X| - (D+1)D^4\varepsilon|X| \geqslant \frac{m}{2}$$

vertices. Now, since $(X, V_{\ell,2})$ is (ε, p)-regular, then by Lemma 5.3 the pair $(W_\ell, V_{\ell,2})$ is $(2\varepsilon, p)$-regular with p-density at least $\delta/(100) - \varepsilon$. By Lemma 5.3 there are at most $2\varepsilon m$ vertices of $V_{\ell,2}$ with less than

$$\left(\frac{\delta}{100} - 3\varepsilon\right) p|W_\ell| \geqslant \frac{\delta}{200}pm$$

neighbours in W_ℓ. We remove each such vertex from $X_{\ell,2}$ thus obtaining a set $X'_{\ell,2}$ such that every vertex in $X'_{\ell,2}$ has at least $\delta pm/200$ neighbours in W_ℓ. Now, we need to find an expander subgraph of $(X_{\ell,1}, X'_{\ell,2})$. Since $(V_{\ell,1}, V_{\ell,2})$ is (ε, p)-regular with $d_p(V_{\ell,1}, V_{\ell,2}) \geqslant \delta/100$ and

$$|X_{\ell,1}|, |X'_{\ell,2}| \geqslant 16D\varepsilon m - 2\varepsilon m \geqslant (4D+6)\varepsilon m,$$

we may use Lemma 6.7 to obtain a pair $(Y_{\ell,1}, Y_{\ell,2})$, with $Y_{\ell,1} \subseteq X_{\ell,1}$ and $Y_{\ell,2} \subseteq X'_{\ell,2}$, such that $G'[Y_{\ell,1}, Y_{\ell,2}]$ is bipartite $(\varepsilon m, 2D+2)$-expander and satisfies $|Y_{\ell,j}| \geqslant |X_{\ell,j}| - 3\varepsilon m \geqslant |T_{\ell,j}| + 13D\varepsilon m$ for $j \in \{1,2\}$. □

7. Proof of Theorem 1.1

The proof of Theorem 1.1 follows from the following stability result for the Ramsey problem of cliques and trees in $G(N,p)$.

Theorem 7.1. *For every $r, D \geqslant 2$ there exist $\delta, C, C' > 0$ such that if $N \geqslant (1-\delta)rn$ and $p \geqslant (C'/N)^{2/(r+2)}$, then $G = G(N,p)$ has, with high probability, the following property. For every blue-red colouring of $E(G)$, at least one of the following holds:*

a) G contains a blue copy of K_{r+1}.

b) G contains a red copy of every $T \in \mathcal{T}(n,D)$.

c) There exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$, with $|V_0| \leqslant C/p$ and $|V_i| \leqslant n + C/p$ for each $i \in [r]$, such that all edges of $G[V_i, V_j]$ are coloured in blue for each $1 \leqslant i < j \leqslant r$.

To see why Theorem 7.1 implies Theorem 1.1, notice that (c) cannot occur if $N > rn + (r+1)C/p$. Moreover, this stability result is sharp in the same directions of Theorem 1.1, however, it might be possible to improve the value of δ up to $1/D$.

From now on we will always assume that the edges of $G = G(N,p)$ are blue-red coloured without having any blue copy of K_{r+1} or being $\mathcal{T}(n,D)$-universal in red. Before proving Theorem 7.1, we will prove an intermediate stability result. More precisely, we will prove that there exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that $|V_0| = o(N)$, and $|V_i| \leqslant (1+o(1))n$ and $e_B(V_i) = o(pN^2)$ for each $i \in [r]$. Indeed, we will prove this by combining Theorem 1.3 and 5.1 to get a partition with $o(pN^2)$ blue edges within the parts. Since G is (η, p)-uniform, for any given $\eta > 0$ and $p \gg 1/N$, we can say that the red graph, induced by each part given by Theorem 1.1, is indeed a weak $(\eta N, \eta N)$-expander. Therefore, if for some $i \in [r]$ and $\alpha \gg \eta$ we have $|V_i| > (1+\alpha)n$, then Proposition 4.2 and Theorem 3.1 would imply that G_R is $\mathcal{T}(n,D)$-universal

At this point, we will update the partition by removing $o(N)$ "bad" vertices: those with low degree in any part and those with high blue degree inside the part it belongs to. We will call *good* any r-partition in which each part has roughly $(1-1/D)$n vertices and the vertices have the "correct" red and blue degrees in each part. More precisely, we want that each vertex has $\Omega(pN)$ red neighbours and $o(pN)$ blue neighbours inside the part it belongs, and has $\Omega(pN)$ blue neighbours in each other part.

For some specific parameters, we show that a *good* partition has no red edges between different parts. This is true because every tree $T \in \mathcal{T}(n, D)$ has a cut edge $e \in E(T)$ such that the trees T_1 and T_2, obtained by removing e, both have size at most $n - n/D$. Therefore, any red edge between different parts would yield a copy of T, since Theorem 3.1 allows us to choose the starting point of the embeddings.

The proof follows now by relocating all but $O(1/p)$ of the leftover vertices to a specific part, so that this new partition is still *good*. For some small $\varepsilon > 0$, if some leftover vertex has εpN blue neighbours in each part, then a simple application of Lemma 5.8 would yield a blue copy of K_{r+1}. On the other hand, if a leftover vertex has less than εpN blue neighbours and more than $\Omega(pN)$ red neighbours in one part, then this vertex can be relocated to this part and thus getting a new *good* partition. Therefore, there exists a *good* partition with at most $O(1/p)$ leftover vertices, those with low degree in at least one of the parts. Finally, since $K_{r+1} \nsubseteq G_B$ we may apply Proposition 7.6, which follows from the tools developed in Section 3, to conclude that each $G_R[V_i]$ is $\mathcal{T}(|V_i| - O(1/p), D)$-universal, which finishes the proof.

To make this whole argument precise, we begin by showing the following intermediate stability result.

Proposition 7.2. *For every $\alpha, \varepsilon > 0$ and integers $r, D \geqslant 2$, there exist $C', \delta > 0$ such that if $N \geqslant (1-\delta)rn$ and $p \geqslant C'N^{-2/(r+2)}$, then $G = G(N,p)$ has, with high probability, the following property. For every blue-red colouring of $E(G)$, at least one of the following holds:*

a) G contains a blue copy of K_{r+1}.

b) G contains a red copy of any $T \in \mathcal{T}(n, D)$.

c) There exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that $|V_0| \leqslant \alpha n$ and for each $i \in [r]$ we have that $||V_i| - n| \leqslant \alpha n$ and that $e_B(V_i) \leqslant \varepsilon pN^2$.

Proof. Without loss of generality, we ask that ε is small enough for calculations. For ε and r, we get C' and δ' from Theorem 5.1. Let $\delta = \alpha/(2r^2)$, $\varrho = 1/r + 2\delta$, $N \geqslant (1-\delta)rn$ and $p \geqslant C'N^{-2/(r+2)}$. Since $p \gg 1/N$, then Theorem 1.3 implies that, with high probability, if $e(G_R) \geqslant (\varrho + \delta')e(G)$ then G_R contains all trees with maximum degree D and

$$\varrho N \geqslant \left(\frac{1}{r} + 2\delta\right)(1-\delta)rn \geqslant n$$

edges. Assuming otherwise, we have

$$e(G_B) \geqslant \left(1 - \frac{1}{r} - \delta'\right)e(G).$$

Theorem 5.1 implies that, with high probability, all K_{r+1}-free subgraphs with this many edges are εpN^2-close to being r-partite. Therefore, we may assume that there exists a partition $V(G) = W_1 \cup \cdots \cup W_r$ such that $e_B(W_i) \leqslant \varepsilon pN^2$ for each $i \in [r]$. Since $p \gg 1/N$, we may also rule out the event in which G is not (η, p)-uniform for some $\eta \ll \alpha$.

Claim 7.3. *In the events considered above, for each $i \in [r]$ the following holds. If $|W_i| \geqslant N/2r$, then there exists $V_i \subseteq W_i$, with $|W_i \setminus V_i| \leqslant \eta N$, such that $G_R[V_i]$ is a $(\eta N, \eta N)$-expander.*

Proof of Claim 7.3. We prove first that $G_R[W_i]$ is a weak $(\eta N, \eta N)$-expander. Since G is (η, p)-uniform, then for every pair of disjoint sets $X, Y \subseteq V(G)$, with $|X|, |Y| \geqslant \eta N$, we have

$$\begin{aligned} e_R(X,Y) &= e(X,Y) - e_B(X,Y) \\ &\geqslant \frac{p}{2}|X||Y| - \varepsilon p N^2 > 0, \end{aligned}$$

as long as $2\varepsilon < \eta^2$. Since $|W_i| \geqslant (D+3)\eta N$, provided η is small enough, we may apply Proposition 3.6 to find $V_i \subseteq W_i$, with $|W_i \setminus V_i| \leqslant \eta N$, such that $G_R[V_i]$ is an $(\eta N, \eta N)$-expander. $\square$

For each $i \in [r]$ such that $|W_i| \geqslant N/2r$, by Claim 7.3 we know that $G_R[V_i]$ is $(\eta N, \eta N)$-expander and thus for all $X \subseteq V_i$, with $\eta N \leqslant |X| \leqslant 2\eta N$, we have (by discounting the non-neighbours and vetices from X) that

$$\begin{aligned} |N_R(X) \cap V_i| &\geqslant |V_i| - \eta N - |X| + 1 \\ &\geqslant (|V_i| - \eta N - (D+1)|X|) + D|X| + 1 \\ &\geqslant (|V_i| - 3D\eta N) + D|X| + 1. \end{aligned}$$

Suppose that V_1 is the largest of the V_i's and notice that $|W_1| \geqslant |V_1| \geqslant N/r - \eta N \geqslant N/2r$, for η small enough. Therefore, if $G_R[V_1]$ is not $\mathcal{T}(n, D)$-universal, then Theorem 3.1 implies that

$$|V_i| \leqslant |V_1| \leqslant n + 3D\eta N$$

for all $i \in [r]$. Set $V_0 = V(G) \setminus (V_1 \cup \cdots \cup V_r)$ and choose η small enough so that

$$|V_0| \leqslant \frac{\alpha n}{2r} \quad \text{and} \quad |V_i| \leqslant \left(1 + \frac{\alpha}{r}\right) n$$

for each $i \in [r]$. We finish the proof by pointing out that the upperbound on the size of each part and the lower bound on N implies that $|V_i| \geqslant (1-\alpha)n$. To see this, suppose wlog that $|V_r| < (1-\alpha)n$. Then there exists $j \in [r-1]$ such that

$$\begin{aligned} |V_j| &\geqslant \frac{N - |V_r| - |V_0|}{r-1} \\ &> \frac{1}{r-1}\left((1-\delta)rn - (1-\alpha)n - \frac{\alpha n}{2r}\right) \\ &\geqslant \left(1 + \frac{\alpha}{r}\right) n, \end{aligned}$$

which is a contradiction and thus $||V_i| - n| \leqslant \alpha n$ for all $i \in [r]$. $\square$

Now we begin to push the stability even further. Until now we have used Proposition 3.6 to remove small sets from weak expanders so that the remaining graph is an actual expander. However, we cannot spare these extra vertices in the process of relocating vertices from V_0 to one of the other parts. We then prove that if a set induces a graph with high minimun red degree and with roughly the expected codegree, then it has the property (i) of expansion.

Lemma 7.4. *For every $\gamma, C > 0$ there exists $\gamma' > 0$ such that the following is true for all $p \in (0,1)$ and N with $pN \gg \log N$. Let G be a N-vertex graph such that for all $u, v \in V(G)$ we have $d(u) \geqslant \gamma pN$ and $|N(u) \cap N(v)| \leqslant 2p^2 N \log N$. Then for every $X \subseteq V(G)$ with $|X| \leqslant C/p$ we have*

$$|N(X)| \geqslant \frac{\gamma' pN}{\log N}|X|.$$

Proof. For $X \subseteq V(G)$ with $|X| \leqslant C/p$, take a subset $X' \subseteq X$ with $|X'| \leqslant \gamma/(4p \log N)$. By an inclusion-exclusion argument, we can bound the size of $N(X)$ by

$$\begin{aligned}
|N(X)| &\geqslant |N(X')| - |X| \\
&\geqslant \sum_{u \in X'} |N(u)| - \sum_{v,w \in \binom{X'}{2}} |N(v) \cap N(w)| - |X| \\
&\geqslant \gamma pN|X'| - |X'|^2 \cdot (2p^2 N \log N) - |X| \\
&\geqslant \gamma pN|X'| - \frac{\gamma pN}{2}|X'| - |X| \\
&\geqslant \Omega\left(\frac{pN}{\log N}\right)|X|,
\end{aligned}$$

where in the last inequality we used that $pN \gg \log N$. □

The last ingredient before we move to the proof of Theorem 7.1 is to show that if two parts, say V_1 and V_2, are "good" enough, then we have $e_R(V_1, V_2) = 0$. More precisely, we require them to be large enough and that they induce graphs with high minimun degree in red and low maximum degree in blue. This motivates the following definition.

Definition 7.5. *Let $\varepsilon > 0$ and let $r, D \geqslant 2$ be integers. For a blue-red colored N-vertex graph G we say that a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ is ε-good if for every $i \in [r]$*

a) $|V_i| \geqslant (1 - 1/2D)N/r$,
b) $d_R(v, V_i) \geqslant pN/32r$ for every $v \in V_i$, and
c) $d_B(v, V_i) \leqslant \varepsilon pN$ for every $v \in V_i$.

We prove now that for any ε-good partition of $V(G(N,p))$ we have that $e_R(V_i, V_j) = 0$ for all $1 \leqslant i < j \leqslant r$. First, we prove that the graphs $G_R[V_i]$ are expanders for each $i \in [r]$. Thus, by Haxel's theorem (Theorem 3.1) we can embed any tree of size $(1 - o(1))n$ into any of the V_i's. Suppose there is a red edge between V_i and V_j. We may split any given tree $T \in \mathcal{T}(n, D)$ into two trees T_1 and T_2, connected by an edge and both having at most $(1 - 1/D)n$ vertices. Thus we may embed T_1 into V_i and T_2 into V_j, and complete the embedding of T by using the red edge between V_i and V_j.

With this new information we may guarantee that the red graphs induced by the V_i's have even stronger expansion properties. In fact, for each $i \in [r]$ we may show that every pair of large disjoint subsets of V_i always have at least one red edge in between. Indeed, if for some $i \in [r]$ there exist a pair of disjoint sets $X, Y \subseteq V_i$ with

$$|X| = |Y| \geqslant \frac{\log N}{N^{2/(r+1)}}$$

and no red edges in between, then, with high probability, X, Y and the remaining V_j's would span a canonical blue-copy of K_{r+1}.

Proposition 7.6. *For integers $r, D \geqslant 2$ there exist $\delta, \varepsilon, C, C' > 0$ such that the following holds for $N \geqslant (1 - \delta)rn$ and $p \geqslant C' N^{-2/r+2}$. With high probability, for every blue-red colouring of the edges of $G = G(N, p)$ that admits a ε-good partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$, at least one of the following holds:*

a) G contains a blue copy of K_{r+1}.
b) G contains a red copy $T \in \mathcal{T}(n, D)$.
c) For every $1 \leqslant i < j \leqslant r$ we have $e_R(V_i, V_j) = 0$. Moreover, for each $i \in [r]$ the graph $G_R[V_i]$ is $\mathcal{T}(|V_i| - C/p, D)$-universal

Proof. Assume that neither (a) nor (b) hold. For $\alpha = 1/32D$, we take C from Lemma 5.7 so that, with high probability, G is a weak $(C/p, \alpha N/4r)$-expander, and set $\varepsilon = \alpha/(6CD)$. Moreover, there exists a constant C' such that if $p \geqslant C' N^{-1/2}$, then, with high probability, every pair of vertices in G has at most $2p^2 N \log N$ common neighbours. Finally, because of the first property of the ε-good partition, we deduce that $N \leqslant 2r|V_i|$.

Our first goal is to prove that each V_i satisfies the hypothesis of Theorem 3.1 in order to show that $G_R[V_i]$ is $\mathcal{T}((1 - 1/D)n, D)$-universal and thus deduce the first part of (c). For $i \in [r]$, we apply Lemma 7.4 to $G_R[V_i]$, with parameters $\gamma = 1/32r$ and C, so that for every $X \subseteq V_i$, with $1 \leqslant |X| \leqslant C/p$, we have

$$|N_R(X) \cap V_i| = \Omega\left(\frac{pN}{\log N}\right)|X| \geqslant D|X| + 1. \tag{19}$$

For $X \subseteq V_i$ with $C/p \leqslant |X| \leqslant 2C/p$, we use that G is a weak $(C/p, \alpha N/4r)$-expander to deduce that

$$
\begin{aligned}
|N_R(X) \cap V_i| &\geqslant |V_i| - \frac{\alpha N}{4r} - \varepsilon p N|X| - |X| \\
&\geqslant \left(1 - \frac{\alpha}{2}\right)|V_i| - 2\varepsilon p N|X| \\
&\geqslant \left(1 - \frac{\alpha}{2}\right)|V_i| - 3\varepsilon p N|X| + D|X| + 1 \\
&\geqslant (1 - \alpha)|V_i| + D|X| + 1,
\end{aligned} \tag{20}
$$

where in the third inequality we used that $pN \gg 1$. Since $\alpha \leqslant 1/D$, then

$$
(1-\alpha)|V_i| \geqslant \left(1 - \frac{1}{D}\right)\left(1 - \frac{1}{2D}\right)\frac{N}{r} \geqslant \left(1 - \frac{1}{D}\right)n,
$$

and thus we may use Theorem 3.1 on each $G_R[V_i]$ in order to find trees of size $(1-1/D)n$ and maximum degree at most D.

Let $T \in \mathcal{T}(n, D)$ be given. By a lemma of Krivelevich [25], there exists a cut edge $u_1u_2 \in E(T)$ which splits T into two trees T_1 and T_2, both with at least n/D vertices and, consequently, at most $(1-1/D)n$ vertices. Suppose that exists a red edge v_1v_2 between two different parts, say $v_1 \in V_1$ and $v_2 \in V_2$. By Theorem 3.1, we may find an embedding of T_i in $G_R[V_i]$ that maps u_i to v_i, for $i \in \{1, 2\}$, and thus, together with the red edge v_1v_2, yields an embedding of T. Therefore, there are no red edges between different parts. Now we move to prove the second part of (c).

Set $d = D\log^4 n/20$. We will show now that $G_R[V_i]$ is a $(|V_i|/2d, |V_i|/2d, d)$-expander for each $i \in [r]$. Indeed, for $X \subseteq V_i$ with $1 \leqslant |X| \leqslant C/p$, by (19) we get $|N_R(X) \cap V_i| \geqslant d|X| + 1$. For $C/p \leqslant |X| \leqslant |V_i|/d$, by (20) we have that

$$
|N_R(X) \cap V_i| \geqslant (1-\alpha)|V_i| \geqslant d|X| + 1,
$$

since $\alpha < 1/2$. To show the second expansion property, suppose that there existed a pair of disjoint sets $X, Y \subseteq V_i$ with $|X| = |Y| = |V_i|/2d = \Omega(N/\log^4 N)$ such that $e_R(X, Y) = 0$. In particular, all the edges between X and Y would be coloured in blue, and therefore by Lemma 4.3, with high probability there would be a copy of K_{r+1} with one vertex in each of the sets X, Y and the V_j's, for $j \in [r]$ with $j \neq i$ (these sets are polinomially larger than the required size to apply Janson's inequality). Since we assumed that $K_{r+1} \nsubseteq G_B$, then we do have the required expansion properties in order to apply Theorem 3.3, and therefore $G_R[V_i]$ contains all trees in $\mathcal{T}(|V_i|, D)$ with at most $|V_i|/d$ leaves.

For trees with at least $|V_i|/d$ leaves, we already have that $G_R[V_i]$ is a weak $(|V_i|/d, |V_i|/d)$-expander, so we only have to show that it is also a weak $(C/p, |V_i|/32D)$-expander. But since $\alpha \leqslant 1/(32D)$, this is already guaranteed by (20). So we apply Theorem 3.4 to $G_R[V_i]$ to conclude that it is $\mathcal{T}(|V_i| - C/p, D)$-universal. $\square$

Now we are ready to prove Theorem 7.1. We are left to prove only that there exists a ε-good partition with $V_0 \leqslant C/p$.

Proof of Theorem 7.1. We apply Proposition 7.6, with parameters r and D, to get $\delta_1, \varepsilon, C, C_1'$ and we define $\alpha \leqslant 1/6D$ later so that it is small enough for calculations. Then we apply Proposition 7.2, with parameters $\varepsilon^2/4$ and α, to get C_2' and δ_2. Moreover let C_3' be given by Lemma 5.8, and set $C_4' = 10^5 r^2$. Now we set $\delta = \min\{\delta_1, \delta_2\}$ and $C' = \max\{C_1', C_2', C_3', C_4'\}$ and we consider $N \geqslant (1-\delta)rn$ and $p \geqslant C' N^{-2/(r+2)}$.

By Proposition 7.2, with high probability, if $K_{r+1} \nsubseteq G_B$ and if G_R is not $\mathcal{T}(n, D)$-universal, then there exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that

$$||V_i| - n| \leqslant \alpha n \quad \text{and} \quad |V_i| \leqslant \alpha n,$$

and $e_B(V_i) \leqslant \varepsilon^2 p N^2/4$ for each $i \in [r]$. We now define a new partition by first removing from each V_i the “bad” vertices. Let

$$B = \{v \in V(G) : d_B(v, V_i') \geqslant \varepsilon p N/2 \text{ for some } i \in [r]\}$$

and notice that $|V_i \setminus B| \geqslant |V_i| - \varepsilon N/2$ for each $i \in [r]$. Now, we define

$$B' = \left\{ v \in V(G) : d(v, V_i' \setminus B) \leqslant \frac{pN}{16r} \text{ for some } i \in [r] \right\}.$$

We show that the partition defined by $W_0 = V(G) \setminus (\bigcup_{i \in [r]} W_i)$ and $W_i = V_i \setminus (B \cup B')$, for $i \in [r]$, is $\varepsilon/2$-good. First we provide a lower bound on the size of each W_i. Since $e_B(V_i) \leqslant \varepsilon^2 p N^2/4$, then $|B \cap V_i| \leqslant \varepsilon N/2$ and we argue that

$$|V_i \setminus B| \geqslant |V_i| - \frac{\varepsilon N}{2} \geqslant (1 - 2\alpha)\frac{N}{r}. \tag{21}$$

Indeed, since $|V_0| \leqslant \alpha n$ and $|V_i| \leqslant (1+\alpha)n$, then $n \geqslant N/(1 + 2r\alpha)$. Therefore, (21) follows by choosing α sufficiently small and because $\varepsilon \leqslant \alpha/r$. In particular, since $|V_i \setminus B| \geqslant N/2r$, then by Lemma 5.7, there are at most $128r/p$ vertices of G with less than $pN/16r$ neighbours in $V_i \setminus B$. Thus we have

$$\begin{aligned} |V_i \setminus (B \cup B')| &\geqslant (1 - 2\alpha)\frac{N}{r} - \frac{128r^2}{p} \\ &\geqslant (1 - 3\alpha)\frac{N}{r} \\ &\geqslant \left(1 - \frac{1}{2D}\right)\frac{N}{r}. \end{aligned}$$

Now we basically have proved that this partition is $\varepsilon/2$-good. We already have the upper bound on the maximum blue degree within each part, and note that for each $i \in [r]$ and $u \in W_i$ we have that

$$d_R(u, W_i) \geqslant \frac{pN}{16r} - \frac{\varepsilon p N}{2} - \frac{128r^2}{p} \geqslant \frac{pN}{32r},$$

since $\varepsilon \leqslant 1/20r$ and $pN \geqslant C_4/p$. To finish the proof, take a maximal $\varepsilon/2$-good partition $U_0 \cup U_1 \cup \cdots \cup U_r$ such that $W_i \subseteq U_i$ for every $i \in [r]$. We prove that if $U_0 \nsubseteq B$, then this partition would not be maximal. Suppose there exists $u \in U_0 \setminus B$. If $d_B(u, U_i) \geqslant \varepsilon pN/2$ for all $i \in [r]$, then we would get blue copy of K_{r+1} containing u, by Lemma 5.8, which we assumed not to exist.

Then there must exist some $i \in [r]$ such that $d_R(u, U_i) \geqslant pN/32r$, in which case we update $U_i := U_i \cup \{u\}$. The new partition $V(G) = U_0 \cup U_1 \cup \cdots \cup U_r$ is ε-good, since it adds at most $1 \ll pN$ vertex to each blue neighborhood. Moreover, Proposition 7.6 implies that $e_R(U_i, U_j) = 0$ for every $1 \leqslant i < j \leqslant r$. Finally, we use Lemma 5.8 again to get that the maximum blue degree inside parts is actually $\varepsilon pN/2$, which makes this partition $\varepsilon/2$-good. This contradicts the maximality of the partition and thus finish the proof.

□

8. Open questions and concluding remarks

8.1. **Gap between Theorem 1.1 and Theorem 1.2.** The problem that we have considered in this paper is the following: Determine the smallest $t = t(n)$ such that if $N \geqslant rn + t$ then $G(N, p) \to (K_{r+1}, \mathcal{T}(n, D))$.

Recall that Theorem 1.1 says that $t \geqslant C/p$ suffices for $p \gg N^{-2/(r+2)}$ and, moreover, for this p the value of t is optimal up to the constant term. On the other hand, in the regime $p \ll N^{-2/(r+1)}$, for any constant $C > 0$ and $N = Cn$, with high probability, $G(N, p) \not\to (K_{r+1}, \mathcal{T}(n, D))$. It would be interesting to understand what happens with t in the regime $N^{-2/(r+1)} \ll p \ll N^{-2/(r+2)}$.

Let us show that $t \gg 1/p$ is not enough for p in this regime. To illustrate this, let us restrict to the case of triangles. Set $t = o(1/p^3N)$, let $G = G(N, p)$ and let $T \subseteq V(G)$ be a set of size t. Note that the expected number of triangles in G, with exactly one vertex in T, is roughly $p^3tN^2 = o(N)$. Therefore, if $N \leqslant 2n + t$, then there exists a partition $V(G) = T \cup V_1 \cup V_2$, with $|V_i| = n$ for $i \in \{1, 2\}$, such that there is no triangle in G with one vertex on each set. In this case, if all the edges between different parts are coloured with blue and all the edges within each part is coloured with red, then we do not create any blue triangles or a red tree with n edges. Moreover, notice that

$$\frac{1}{p^3N} \gg \frac{1}{p}$$

if and only if $p \ll N^{-1/2}$, which makes the transition in this behavior consistent with the regimes of Theorems 1.1 and Theorem 1.2. A similar argument shows the same behavior of t for every clique of fixed size.

Problem 1. *For $N^{-2/(r+1)} \ll p \ll N^{-2/(r+2)}$, determine t such that if $N \geqslant rn + t$, then $G(N, p) \to (K_{r+1}, \mathcal{T}(n, D))$ with high probability.*

8.2. **Other Ramsey-good pairs.** Is it possible to extend some of our results by replacing the complete graph for any fixed graph H with chromatic number at least 3. Indeed, given $\varepsilon > 0$, we can prove that if $N \geqslant (\chi(H) - 1 + \varepsilon)n$ and $p \gg N^{-1/m_2(H)}$, then $G(N, p) \to (H, \mathcal{T}(n, d))$. The proof of this result relies on Theorem 1.3 and the general form of the Erdős–Simonovits stability theorem in $G(N, p)$ due to Samotij [33]. It would be interesting to extend Theorem 1.1 to general graphs.

Problem 2. *Given a graph H with $\chi(H) \geqslant 3$, determines t and the regime of p such that if $N \geqslant (\chi(H) - 1)n + t$ then $G(N, p) \to (H, \mathcal{T}(n, D))$ with high probability.*

8.3. **More about Ramsey results for trees.** An important consequence of the Erdős–Sós conjecture is that for all $s, t \in \mathbb{N}$ one has $r(T_s, T_t) \leqslant s + t$, where T_s and T_t denote a tree with s and t vertices, respectively. Therefore, is not surprising that from Theorem 1.3 one can deduce a similar result for random graphs.

Theorem 8.1. *Let $s, D \geqslant 2$ and let $\varepsilon > 0$. Then there exists $C > 0$ such that the following holds. If $p \geqslant C/n$, then*

$$G((1 + \varepsilon)sn, p) \to (\mathcal{T}(n, D), \dots, \mathcal{T}(n, D))$$

with high probability as $n \to \infty$.

A very interesting consequence of Theorem 8.1 is an upper bound for the multicolour size Ramsey number of bounded degree trees. Given a graph F and an integer $s \geqslant 2$, the s-colour size Ramsey number $\hat{r}_s(F)$ of F is the smallest integer m so that there exists a graph G with m edges such that every s-colouring of $E(G)$ yields a monochromatic copy of F.

In the case of trees, it was conjecture in 1983 by Beck [2] that $\hat{r}_2(T) = O(Dn)$ for any fixed tree $T \in \mathcal{T}(n, D)$. This conjecture was settled in 1995 by Haxell and Kohayakawa [18]. For $s \geqslant 2$, one can deduce from Theorem 8.1 that

$$\hat{r}_s(\mathcal{T}(n, D)) = O(n). \tag{22}$$

This upper bound was already known since it can be deduced from a result due to Han, Jenssen, Kohayakawa, Mota and Roberts [16] on the size Ramsey number for power of paths. We point out that the constant that we can get in (22) is of tower type depending on s and D. Therefore, a very interesting open question is to find the dependence of (22) in terms of s and D.

Acknowledgement. The authors would like to thank Robert Morris for indispensable help with the manuscript, and to Hiệp Hàn and Maya Stein for many useful discussions about the topics of this paper. We also would like to thank the reader for their time.

References

[1] J. Balogh, B. Csaba, and W. Samotij. Local resilience of almost spanning trees in random graphs. *Random Structures & Algorithms*, 38(1-2):121–139, 2011.

[2] J. Beck. On size Ramsey number of paths, trees, and circuits. i. *J. Graph Theory*, 7(1):115–129, 1983.

[3] G. Besomi, M. Pavez-Signé, and M. Stein. Degree conditions for embedding trees. *SIAM J. Discrete Math.*, 33(3):1521–1555, 2019.

[4] G. Besomi, M. Pavez-Signé, and M. Stein. On the Erdős-Sós conjecture for trees with bounded degree. *arXiv e-prints*, 2019.

[5] J. Böttcher. Large-scale structures in random graphs. *Surveys in Combinatorics*, 440:87, 2017.

[6] S. Burr. Ramsey numbers involving graphs with long suspended paths. *J. Lond. Math. Soc.*, 2(3):405–413, 1981.

[7] S. A. Burr and P. Erdős. Generalizations of a Ramsey-theoretic result of Chvátal. *J. Graph Theory*, 7(1):39–51, 1983.

[8] V. Chvátal. Tree-complete graph Ramsey numbers. *J. Graph Theory*, 1(1):93–93, 1977.

[9] D. Conlon and W. T. Gowers. Combinatorial theorems in sparse random sets. *Ann. of Math.*, pages 367–454, 2016.

[10] P. Erdős. Extremal problems in graph theory. In *Theory of graphs and its applications, Proc. Sympos. Smolenice*, pages 29–36, 1964.

[11] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[12] P. Frankl and V. Rödl. Large triangle-free subgraphs in graphs withoutk 4. *Graphs Combin.*, 2(1):135–144, 1986.

[13] E. Friedgut, V. Rödl, and M. Schacht. Ramsey properties of random discrete structures. *Random Structures & Algorithms*, 37(4):407–436, 2010.

[14] J. Friedman and N. Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7(1):71–76, 1987.

[15] S. Gerke and A. Steger. *The sparse regularity lemma and its applications*, page 227–258. Surveys in Combinatorics 2005 (London Mathematical Society Lecture Note Series). Cambridge University Press, 2005.

[16] J. Han, M. Jenssen, Y. Kohayakawa, G. O. Mota, and B. Roberts. The multicolour size-Ramsey number of powers of paths. *arXiv preprint arXiv:1811.00844*, 2018.

[17] P. E. Haxell. Tree embeddings. *J. Graph Theory*, 36(3):121–130, 2001.

[18] P. E. Haxell and Y. Kohayakawa. The size-Ramsey number of trees. *Israel J. Math.*, 89(1-3):261–274, 1995.

[19] S. Janson, T. Luczak, and A. Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[20] Y. Kohayakawa. Szemerédi's regularity lemma for sparse graphs. In *Found. Comput. Math.*, pages 216–230. Springer, 1997.

[21] Y. Kohayakawa and B. Kreuter. Threshold functions for asymmetric Ramsey properties involving cycles. *Random Structures & Algorithms*, 11(3):245–276, 1997.

[22] Y. Kohayakawa, G. O. Mota, and M. Schacht. Monochromatic trees in random graphs. *Math. Proc. Philos. Soc.*, 166(1):191–208, 2019.

[23] Y. Kohayakawa and V. Rödl. Szemerédi's regularity lemma and quasi-randomness. In *Recent advances in algorithms and combinatorics*, pages 289–351. Springer, 2003.

[24] J. Komlós, G. N. Sárközy, and E. Szemerédi. Proof of a packing conjecture of Bollobás. *Combin. Probab. Comput.*, 4(3):241–255, 1995.

[25] M. Krivelevich. Embedding spanning trees in random graphs. *SIAM J. Discrete Math.*, 24(4):1495–1500, 2010.

[26] T. Ł uczak, A. Ruciński, and B. Voigt. Ramsey properties of random graphs. *J. Combin. Theory Ser. B*, 56(1):55–68, 1992.

[27] S. Letzter. Path Ramsey number for random graphs. *Combin. Probab. Comput.*, 25(4):612–622, 2016.

[28] R. Montgomery. Embedding bounded degree spanning trees in random graphs. *arXiv preprint arXiv:1405.6559*, 2014.

[29] L. Moreira. Ramsey goodness of paths in random graphs. *arXiv preprint*, 2019.

[30] F. Mousset, R. Nenadov, and W. Samotij. Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties. *arXiv preprint arXiv:1808.05070*, 2018.

[31] V. Rödl and A. Rucinski. Lower bounds on probability thresholds for Ramsey properties. *Combinatorics, Paul Erdős is eighty*, 1:317–346, 1993.

[32] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

[33] W. Samotij. Stability results for random discrete structures. *Random Structures & Algorithms*, 44(3):269–289, 2014.

[34] M. Schacht. Extremal results for random discrete structures. *Ann. of Math.*, pages 333–365, 2016.

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brasil
E-mail address: `pedroc|luizfm@impa.br`

Departamento de Ingeniería Matemática, Universidad de Chile, Beauchef 851, Santiago, Chile
E-mail address: `mpavez@dim.uchile.cl`