Local resilience of almost spanning trees in random graphs

József Balogh* Béla Csaba† Wojciech Samotij‡

April 5, 2010

Abstract

We prove that for fixed integer D and positive reals α and γ, there exists a constant C_0 such that for all p satisfying $p(n) \geq C_0/n$, the random graph $G(n,p)$ asymptotically almost surely contains a copy of every tree with maximum degree at most D and at most $(1-\alpha)n$ vertices, even after we delete a $(1/2-\gamma)$-fraction of the edges incident to each vertex. The proof uses Szemerédi's regularity lemma for sparse graphs and a bipartite variant of the theorem of Friedman and Pippenger on embedding bounded degree trees into expanding graphs.

1 Introduction

The problem of finding sufficient conditions on a graph to contain large trees has been an object of intense study for a long time. An old conjecture of Bollobás [5] stated that if $\gamma > 0$ and n is sufficiently large, then every n-vertex graph G with $\delta(G) \geq (1/2+\gamma)n$ contains all n-vertex trees with bounded degree. In other words, in order to destroy the property of containing all spanning bounded degree trees possessed by the complete graph, one needs to delete at least a $(1/2-\gamma)$-fraction of the edges incident to some vertex. This was proved by Komlós, Sárközy, and Szemerédi [16], and recently improved by Csaba, Levitt, Nagy-György, and Szemerédi [8], who showed that γn can be replaced with $K \log n$, whenever K is a sufficiently large constant. Although the study of the problem of existence of large trees in random graphs has a long history, there is no natural counterpart to the theorem of Komlós et al. [16] in the random setting.

Looking from the edge deletion perspective, the theorem of Komlós et al. [16] can be viewed in the much broader context of *resilience*. Let $\mathcal{P}$ be a graph property and let G be an arbitrary graph from the class $\mathcal{P}$. The resilience of G with respect to the property $\mathcal{P}$ measures how much one has to change G in order to destroy $\mathcal{P}$. Although the notion of

*Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA; and Department of Mathematics, University of Illinois, Urbana, IL 61801, USA. E-mail address: jobal@math.uiuc.edu. This material is based upon work supported by NSF CAREER Grant DMS-0745185 and DMS-0600303, UIUC Campus Research Board Grants 09072 and 08086, and OTKA Grant K76099.

†Department of Mathematics, Western Kentucky University, Bowling Green, KY, 42101, USA. E-mail address: bela.csaba@wku.edu. Partially supported by OTKA K76099.

‡Department of Mathematics, University of Illinois, Urbana, IL, 61801, USA. E-mail address: samotij2@illinois.edu.

resilience, also called *fault tolerance*, has been present in the literature for several years (see, e.g., [1]), only recently it was given a more systematic treatment by Sudakov and Vu [20], who define it as follows.

Definition 1. Let $\mathcal{P}$ be a monotone increasing (decreasing) graph property and let G be an arbitrary graph.

1. The *global resilience* of G with respect to $\mathcal{P}$ is the minimum number r such that by deleting (adding) at most $r \cdot e(G)$ edges from (to) G, one can obtain a graph not in $\mathcal{P}$.
2. The *local resilience* of G with respect to $\mathcal{P}$ is the minimum number r such that by deleting (adding) at most $r \cdot \deg_G(v)$ edges at each vertex v of G, one can obtain a graph not in $\mathcal{P}$.

Using the resilience terminology, one can restate many classic results in graph theory, such as the famous theorem of Turán [21] or the theorem of Dirac [11] giving a sufficient condition for a graph G to contain a Hamiltonian cycle. In this respect, the notion of resilience has proved very useful and initiated a series of generalizations of classic theorems to the more general setting of random and pseudo-random graphs, see [6, 10, 13, 17, 20].

When speaking about global resilience, it is worth to mention a very recent breakthrough result of Conlon and Gowers [7], and, independently, Schacht [19], which resolves a series of open questions concerning global resilience of random discrete structures with respect to certain local properties, culminating previous work of a number of authors (for details, we refer the reader to [19]).

Alon, Krivelevich, and Sudakov [2] proved that for all D and positive α, there exists a constant C such that the random graph $G(n, C/n)$ a.a.s. (asymptotically almost surely) contains all trees with maximum degree at most D and $(1-\alpha)n$ vertices. Recently, using a different argument, Balogh, Csaba, Pei, and Samotij [4] showed that C can be as small as $O\big(d \log d + d/\alpha \log(1/\alpha)\big)$. The main result of this paper generalizes the above theorems – we prove that a.a.s. the local resilience of the random graph $G(n,p)$ with respect to containing almost spanning trees with bounded degree is at least a half, even when pn is only a (large) constant.

Theorem 2. *Let α and γ be positive constants, and assume that $D \geq 2$. There exists a constant C_0 (depending on α, γ, and D) such that for all p satisfying $p(n) \geq C_0/n$, the local resilience of $G(n,p)$ with respect to the property of containing all trees of order at most $(1-\alpha)n$ and maximum degree at most D is almost surely greater than $(1/2-\gamma)$.*

Since the property of containing almost spanning trees is global, it is more meaningful to measure the local resilience of graphs with respect to this property. One can remove from every n-vertex graph G all connected subgraphs of order $(1-\alpha)n$ by deleting at most $2\alpha e \cdot (G)$ edges. This changes when we relax the requirement on the size of the tree. Dellamonica and Kohayakawa [9] provided explicit constructions of graphs that have global resilience with respect to the property of containing large (linear in the order of the graph) trees of bounded degree.

Note that the constant $(1/2-\gamma)$ in the statement of Theorem 2 is best possible, provided that $\alpha < 1/2$. Dellamonica, Kohayakawa, Marciniszyn, and Steger [10] proved that for every positive constant γ, if pn is a large enough constant, one can almost surely find an approximately even bipartition of the vertex set of the random graph $G(n,p)$ such that each vertex v has at most $(1/2+\gamma)\deg(v)$ neighbors in the other partite set. It follows that a.a.s. by deleting at most a $(1/2+\gamma)$-fraction of the edges incident to each vertex, one can turn the random graph $G(n,p)$ into a graph whose largest connected component has about $n/2$ vertices, and hence does not contain any tree with $(1-\alpha)n$ vertices.

A simple argument proves that the constant $(1/2-\gamma)$ is sharp if $D \geq 3$ and $\alpha < (D-2)/(2D-2)$. Recall that one can make an arbitrary graph bipartite by removing at most half the edges at each vertex. It is easy to check that in the random graph $G(n,p)$, where pn is a large enough constant, a.a.s. every bipartite graph obtained in such way has partite sets of approximately even size. Since for all sufficiently large n, there are trees with n vertices and maximum degree D, whose color classes have sizes differing by a factor arbitrarily close to $D-1$, it follows that a.a.s. after deleting at most half the edges at each vertex of $G(n,p)$, the remaining graph cannot contain all trees with maximum degree D of size greater than $D/(2D-2)\cdot n$.

Recall that if $p(n) \leq C/n$, then there is a positive constant α (depending on C) such that almost surely the size of the largest connected component of $G(n,p)$ does not exceed $(1-\alpha)n$. Moreover, if $D(n) \to \infty$ as $n \to \infty$, then a.a.s. $G(n,p)$ contains $o(n/D)$ vertices with degree at least D, and hence it cannot contain all trees with maximum degree D. Therefore, Theorem 2 is in a sense sharp.

Theorem 2 is a direct consequence of a standard uniformity result for the random graph (Lemma 4) and the following more general Theorem 3. The definition of (η,p)-uniform graphs is given at the beginning of Section 2.

Theorem 3. *Let α and γ be positive constants, and assume that $D \geq 2$. There exist η_0 and n_0 (both depending on α, γ, and D) such that the following holds. Let G be an n-vertex (η,p)-uniform graph, with $p > 0$, $\eta < \eta_0$ and $n \geq n_0$. Let G' be a subgraph of G such that $\deg_{G'}(v) \geq (1/2+\gamma)\deg_G(v)$ for each vertex v. Then G' contains all trees with at most $(1-\alpha)n$ vertices and maximum degree at most D.*

The remainder of this paper is organized as follows. In Section 2, we introduce the class of (η,p)-uniform graphs and formulate a version of Szemerédi's regularity lemma for graphs in that class. Moreover, we define a class of bipartite expanders and give a sufficient condition on the expansion parameters which guarantees that all such expanders contain all large bounded degree trees. Graphs from this class, which (as it will turn out in the proof of Theorem 3, in Section 3) are abundant in the graph G', will be used for embedding the bulk of our almost spanning tree. Finally, Section 4 contains a few concluding remarks.

2 Preliminaries and tools

2.1 Uniformity of the random graph

Fix positive constants η and p. We say that an n-vertex graph G is *η-uniform with density p*, or simply *(η,p)-uniform*, if all $A, B \subseteq V(G)$ with $A \cap B = \emptyset$ and $|A|, |B| \geq \eta n$ satisfy

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

and

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

Furthermore, we say that G is *η-upper-uniform with density p*, or simply *(η, p)-upper-uniform*, if only the second inequalities in (1) and (2) hold for all A and B as above. It is not surprising that the random graph $G(n, p)$ is almost surely uniform, provided that its density p is sufficiently large.

Lemma 4. *If $\eta > 0$ and $pn > \frac{8}{\eta^4(1-\eta)}$, then a.a.s. the random graph $G(n,p)$ is (η, p)-uniform.*

Proof. Let $G = G(n,p)$. The probability that a fixed pair of sets A and B violates the η-uniformity condition (1) is

$$P\left(e_G(A,B) - p|A||B| > \eta p|A||B|\right) + P\left(e_G(A,B) - p|A||B| < \eta p|A||B|\right).$$

By a standard Chernoff-type estimate (see, e.g., [3]), this is at most

$$\exp\left(-\frac{(\eta p|A||B|)^2}{2p|A||B|}\right) + \exp\left(-\frac{(\eta p|A||B|)^2}{2p|A||B|} + \frac{(\eta p|A||B|)^3}{2(p|A||B|)^2}\right) \le 2\exp\left(-\frac{\eta^4(1-\eta)pn^2}{2}\right),$$

where the inequality holds since we assumed that $|A|, |B| \ge \eta n$. Similarly, the probability that a fixed set A of size at least ηn violates the η-uniformity condition (2) is at most

$$\exp\left(-\frac{(\eta p\binom{|A|}{2})^2}{2p\binom{|A|}{2}}\right) + \exp\left(-\frac{(\eta p\binom{|A|}{2})^2}{2p\binom{|A|}{2}} + \frac{(\eta p\binom{|A|}{2})^3}{2(p\binom{|A|}{2})^2}\right) \le 2\exp\left(-\frac{\eta^4(1-\eta)pn^2}{8}\right).$$

By our assumption on p and the union bound, the probability that G is not η-uniform is at most $2^n \cdot 2^n \cdot (2e^{-4n}) + 2^n \cdot (2e^{-n}) = o(1)$. $\square$

2.2 Szemerédi's regularity lemma for sparse graphs

Let G be a graph and let $p > 0$. For any two disjoint subsets $A, B \subseteq V(G)$, let us define the *p-density* of the pair (A, B) in G to be the quantity

$$d_{G,p}(A,B) := \frac{e_G(A,B)}{p|A||B|}.$$

Now suppose that $\varepsilon > 0$ and A, B are disjoint sets of vertices in G. We say that the pair (A, B) is *(ε, p)-regular* if all $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \ge \varepsilon|A|$ and $|B'| \ge \varepsilon|B|$ satisfy

$$|d_{G,p}(A', B') - d_{G,p}(A,B)| \le \varepsilon.$$

Finally, we say that a partition $(V_0, \ldots, V_k)$ of $V(G)$ is *(ε, p)-regular* if $|V_0| \le \varepsilon|V(G)|$, $|V_i| = |V_j|$ for all $i, j \in \{1, \ldots, k\}$ and at least $(1-\varepsilon)\binom{k}{2}$ pairs (V_i, V_j) with $1 \le i < j \le k$ are (ε, p)-regular. We may now state a version of Szemerédi's regularity lemma for η-upper-uniform graphs (see, e.g., [15]).

Lemma 5. *For any $\varepsilon > 0$ and $k_0 \ge 1$, there are positive constants η and K_0, with $K_0 \ge k_0$, such that any η-upper-uniform graph G with density $p \in (0,1]$ and at least k_0 vertices admits an (ε, p)-regular partition $(V_0, \ldots, V_k)$ with $k_0 \le k \le K_0$.*

2.3 Embedding trees in bipartite graphs

For an arbitrary graph H and a set $X \subseteq V(H)$, let $N_H(X)$ denote the set of neighbors in H of vertices in X. Extending a path-embedding result of Pósa [18], Friedman and Pippenger [12] proved that all graphs satisfying certain expansion property contain all small trees with bounded maximum degree.

Theorem 6. *Let m and D be positive integers and let H be a non-empty graph. Moreover, assume that every $X \subseteq V(H)$ with $|X| \le 2m$ satisfies $|N_H(X)| \ge (D+1)|X|$. Then H contains every tree with m vertices and maximum degree at most D.*

An apparent limitation of Theorem 6 in our setting is that it can be helpful in finding only relatively small trees. Namely, in a graph of order n, the size of the largest tree whose existence can be guaranteed by Theorem 6 is only about $n/(2D+2)$, where D is the maximum degree of the tree. Building on the ideas developed by Friedman and Pippenger [12], Haxell [14] managed to overcome this problem. Recently, Balogh, Csaba, Pei, and Samotij [4] noted that a simplified version of Haxell's general tree-embedding result suffices for finding almost spanning trees in graphs from certain families of expanders, such as $G(n,p)$ with very low edge probability p. Unfortunately, neither the general result of Haxell [14] nor the simplified statement used in [4] are suitable when the target graph is bipartite – neither of them allows to embed trees of order greater than the number of vertices in the smaller partite set. Luckily, a natural and straightforward modification of the argument of Friedman and Pippenger [12] allows one to prove the following theorem.

Theorem 7. *Let D, m_1, M_1, m_2 and M_2 be positive integers. Assume that H is a non-empty bipartite graph with color classes V_1 and V_2, satisfying the following conditions.*

1. *For every $X \subseteq V_i$ with $0 < |X| \le m_i$, $|N_H(X)| \ge D|X| + 1$ for $i \in \{1,2\}$.*
2. *For every $X \subseteq V_i$ with $m_i < |X| \le 2m_i$, $|N_H(X)| \ge D|X| + M_{3-i}$ for $i \in \{1,2\}$.*

Furthermore, let T be a tree with maximum degree at most D and color classes of sizes M_1 and M_2, respectively, and let v be an arbitrary vertex of T belonging to the first color class. Then every mapping of v to a vertex in V_1 extends to an embedding of T in H.

Note that condition 2 in Theorem 7 implies that $|V_1| \ge M_1 + 2Dm_2$ and $|V_2| \ge M_2 + 2Dm_1$. Therefore, in order to use it for embedding almost spanning trees, one must choose m_1 and m_2 so that $Dm_1 \ll M_2$ and $Dm_2 \ll M_1$. We would also like to remark that the idea of having different types of expansion for sets of different sizes, like conditions 1 and 2 in Theorem 7, was already used by Haxell [14].

Proof of Theorem 7. A tree T will be called (M_1, M_2, D)-*small*, or simply *small*, if $\Delta(T) \le D$ and the two color classes of T, the sets $U_1, U_2 \subseteq V(T)$, satisfy $|U_1| \le M_1$ and $|U_2| \le M_2$. Using induction on the size of T, we will prove that H contains all small trees. First, we need a few definitions. Let f be an embedding of some small tree T into our expanding graph H. The *liability* $B_f(x)$ of a vertex $x \in V(H)$ with respect to f is defined by

$$B_f(x) := \begin{cases} D - \deg_T(v), & x = f(v) \text{ for some } v \in V(T), \\ D, & x \notin f(V(T)). \end{cases}$$

We define the *assets* $\mathcal{A}_f(X)$ of a set $X \subseteq V(H)$ to be the set of neighbors of X that are not used in the embedding, i.e., $\mathcal{A}_f(X) := N_H(X) - f(V(T))$. For every set $X \subseteq V(H)$, define

$A_f(X) := |\mathcal{A}_f(X)|$ and $B_f(X) := \sum_{x\in X} B_f(x)$. The quantity $A_f(X) - B_f(X)$ will be called the *balance* of the set X, denoted $C_f(X)$ .Finally, an embedding f of a small tree T into our graph H will be called *good* if it maps U_1 into V_1 and U_2 into V_2, and moreover, every set $X \subseteq V_1$ of size at most $2m_1$ and every set $X \subseteq V_2$ of size at most $2m_2$ have non-negative balance with respect to f.

In order to prove the existence of a good embedding of an arbitrary small tree into our graph H, it clearly suffices to show that the class of good embeddings satisfies the following two properties.

Property 1. Every embedding of a single-vertex tree into V_1 is good.

Property 2. If T is a small tree and S is its subtree obtained by deleting a leaf and the edge incident to it, then any good embedding of S in H can be extended to a good embedding of T in H.

To prove Property 1, suppose that T is a tree consisting of a single vertex. Let f be an arbitrary embedding of T in V_1. We show that f is good. Fix an arbitrary $i \in \{1, 2\}$ and suppose that $X \subseteq V_i$ and $|X| \le 2m_i$. We can easily assume that $X \neq \emptyset$, since $C_f(\emptyset) = 0$. We have

$$A_f(X) = |N_H(X) - f(V(T))| \ge |N_H(X)| - 1 \ge D|X| + 1 - 1 = B_f(X).$$

Thus every set $X \subseteq V_i$ of size at most $2m_i$ has non-negative balance, and so Property 1 is satisfied.

To prove Property 2, assume that f is a good embedding of a tree S, obtained from a small tree T by removing a leaf v. Without loss of generality we may assume that $v \in U_2$, as the proof for the case $v \in U_1$ is identical. For the sake of brevity, let $U_2' := U_2 - \{v\}$.

Claim 8. *If for some $X \subseteq V_1$ with $|X| \le 2m_1$, $C_f(X) = 0$, then $|X| \le m_1$.*

Proof of Claim 8. Assume that some $X \subseteq V_1$ satisfies $C_f(X) = 0$, but $m_1 < |X| \le 2m_1$. Since T is a small tree, $|U_2'| = |U_2| - 1 \le M_2 - 1$, and thus

$$\begin{aligned} A_f(X) &= |N_H(X) - f(V(S))| = |N_H(X) - f(U_2')| \ge |N_H(X)| - |f(U_2')| \\ &\ge D|X| + M_2 - (M_2 - 1) = D|X| + 1 \ge B_f(X) + 1, \end{aligned}$$

where the second equality follows from the fact that $N_H(X) \subseteq V_2$ and f maps U_1 to V_1. This contradicts the assumption that $A_f(X) = B_f(X)$. □

Claim 9. *If some $X, Y \subseteq V_1$, with $|X| \le |Y| \le 2m_1$, satisfy $C_f(X) = C_f(Y) = 0$, then also $C_f(X \cup Y) = 0$ and $|X \cup Y| \le m_1$.*

Proof of Claim 9. Since B_f is a measure on $V(H)$, clearly

$$B_f(X \cup Y) + B_f(X \cap Y) = B_f(X) + B_f(Y).$$

Moreover, since $\mathcal{A}_f(X \cup Y) = \mathcal{A}_f(X) \cup \mathcal{A}_f(Y)$ and $\mathcal{A}_f(X \cap Y) \subseteq \mathcal{A}_f(X) \cap \mathcal{A}_f(Y)$, we have

$$A_f(X \cup Y) + A_f(X \cap Y) \le |\mathcal{A}_f(X) \cup \mathcal{A}_f(Y)| + |\mathcal{A}_f(X) \cap \mathcal{A}_f(Y)| = A_f(X) + A_f(Y).$$

It follows that $C_f(X \cup Y) \le C_f(X) + C_f(Y) - C_f(X \cap Y) = -C_f(X \cap Y)$. Since $|X \cap Y| \le |Y| \le 2m_1$ and f is good, $C_f(X \cap Y) \ge 0$, and hence $C_f(X \cup Y) \le 0$. By Claim 8, $|X|, |Y| \le m_1$, and so $|X \cup Y| \le 2m_1$. Hence, also $C_f(X \cup Y) \ge 0$, and by Claim 8, $|X \cup Y| \le m_1$. □

Corollary 10. *Suppose $X_1, \dots, X_k \subseteq V_1$ are sets of size at most $2m_1$ having zero balance. Then $C_f(X_1 \cup \dots \cup X_k) = 0$ and $|X_1 \cup \dots \cup X_k| \le m_1$.*

Let w be the only neighbor of v in T. Since $v \in U_2$, clearly $w \in U_1$. Recall that f was a good embedding of S in H and hence $f(w) \in V_1$. Let $Y := \mathcal{A}_f(f(w))$ and note that $Y \subseteq V_2$. We can extend f to an embedding of T by mapping v to any vertex in Y. Suppose that for no $y \in Y$, the extension f_y, defined by

$$f_y(x) = \begin{cases} y & \text{if } x = v, \\ f(x) & \text{if } x \neq v, \end{cases}$$

is good. Since clearly f_y maps U_1 to V_1 and U_2 to V_2, this means that for every $y \in Y$, there is an $i(y) \in \{1, 2\}$ and a set $X_y \subseteq V_{i(y)}$ of size at most $2m_{i(y)}$ with $C_{f_y}(X_y) < 0$. Clearly, for all $y \in Y$, $i(y) = 1$, since for every $X \subseteq V_2$, $\mathcal{A}_f(X) = \mathcal{A}_{f_y}(X)$ and $B_{f_y}(X) \le B_f(X)$. Moreover, for each $y \in Y$, we must have $y \in \mathcal{A}_f(X_y)$, $f(w) \notin X_y$ and $C_f(X_y) = 0$, or otherwise $C_{f_y}(X_y) \ge 0$. Let $X^* := \bigcup_{y \in Y} X_y$. By Corollary 10, $C_f(X^*) = 0$ and $|X^*| \le m_1$. Moreover, if we let $X' := X^* \cup \{f(w)\} \subseteq V_1$, then

$$\mathcal{A}_f(X') = \mathcal{A}_f(X^*) \cup \mathcal{A}_f(f(w)) = \mathcal{A}_f(X^*) \cup Y = \mathcal{A}_f(X^*),$$

since $\mathcal{A}_f(X^*) = \cup_{y \in Y} \mathcal{A}_f(X_y)$ and for all $y \in Y$, $y \in \mathcal{A}_f(X_y)$. Also, since $f(w) \notin X^*$ and $\deg_S(w) = \deg_T(w) - 1 \le D - 1$,

$$B_f(X') = B_f(X^*) + B_f(f(w)) \ge B_f(X^*) + 1.$$

This implies that $C_f(X') < 0$, which is a clear contradiction, since $|X'| \le m_1 + 1 \le 2m_1$ and f was good. Hence, f can be extended to a good embedding of T, and so Property 2 holds. □

Let us now define a class of expanding bipartite graphs that we will use with Theorem 7.

Definition 11. Let $b \ge 2$ and let H be a bipartite graph with color classes V_1 and V_2, where $|V_1| \le |V_2|$. Let q be a positive integer with $q < |V_1|$. We will say that H is a *bipartite (q, b)-expander* if it possesses the following properties.

1. Every subset $X \subseteq V_i$ of size at most q satisfies $|N_H(X)| \ge b|X|$ for $i \in \{1, 2\}$.
2. Every subset $X \subseteq V_i$ of size at least q satisfies $|N_H(X)| \ge |V_{3-i}| - q$ for $i \in \{1, 2\}$.

As an immediate consequence of Theorem 7, we derive the following sufficient condition on the expansion parameters b and q, which guarantees that all sufficiently large bipartite (q, b)-expanders contain every almost spanning tree with bounded maximum degree and color classes of appropriate size.

Corollary 12. *Let $D \ge 2$ and let H be a bipartite graph with color classes V_1 and V_2, where $|V_1| \le |V_2|$. Suppose that H is a bipartite $(q, D+1)$-expander with $0 < q < |V_1|/(2D+1)$. Then H contains all trees with maximum degree at most D and color classes of sizes at most $|V_1| - (2D+1)q$ and $|V_2| - (2D+1)q$ respectively. Furthermore, any such tree can be embedded even if we require that a particular vertex of the tree is mapped to a particular vertex of H, as long as this mapping respects the color classes.*

Proof. It is straightforward to check that H satisfies the assumptions of Theorem 7 with $m_i := q$ and $M_i := |V_i| - (2D+1)q$ for $i \in \{1, 2\}$. □

3 Proof of Theorem 3

We start by defining some constants. Let

$$\delta := \frac{\alpha}{16D^2}, \quad \varepsilon := \min\left\{\frac{\alpha}{64D^3}, \frac{\alpha\gamma\delta}{96}, \frac{\gamma^2}{36}\right\}, \text{ and } \eta_0 := \min\left\{\frac{\varepsilon}{2}, \frac{1}{2K_0}\right\},$$

where K_0 is given by Lemma 5 with $k_0 := \lceil 1/\varepsilon \rceil$. Let G be an (η, p)-uniform n-vertex graph, with $p > 0$, $\eta < \eta_0$ and n larger than some constant n_0 depending on α, γ, and D.[1] We let our adversary remove edges from G, so that no more than $(1/2 - \gamma)\deg_G(v)$ edges incident to every vertex $v \in V(G)$ are deleted. Denote the leftover graph by G'. Clearly, G' is η-upper-uniform with density p. Finally, let T be a tree with at most $(1-\alpha)n$ vertices and maximum degree at most D; without loss of generality we may also assume that T has at least $n/2$ vertices. We will show that $T \subseteq G'$.

3.1 Proof outline

In Section 3.2, we apply Szemerédi's regularity lemma to G', and show that the cluster graph, whose edges are the regular pairs with density bounded away from zero, contains an almost spanning subgraph H'' with minimum degree slightly larger than $|V(H'')|/2$. Such large minimum degree guarantees the existence in H'' of an almost perfect matching M. The tree T will be embedded into G'' – the subgraph of G' induced by the union of the clusters in H''; moreover, most edges of T will be mapped to edges inside the dense, regular pairs in G'' that appear in M.

In Section 3.3, we partition the tree T into a bounded number of small subtrees in such a way that none of these subtrees is adjacent to more than D^3 others and every subtree contains all the children of its root.

In Section 3.4, the vertex set of G'' is partitioned into linear-sized subsets, which are then assigned to subtrees from our partition of T and the edges of T joining those subtrees. Each subtree S is assigned two subsets of the opposite ends of some edge in M, one for each color class of S; both subsets are slightly larger than the color class of S they are assigned to. An edge e joining two subtrees S and S' is assigned a small subset (a 'connecting' set) of a cluster that is adjacent (in H'') to the two clusters that were assigned to the color classes of S and S' which contain the endpoints of e. In Section 3.5, we trim all these subsets so that the pair assigned to every subtree is a bipartite expander, and every 'connecting' set has many neighbors in both sets it 'connects'.

Finally, in Section 3.6, we embed T in G'' in a top-down fashion. The subtree containing the root of T is embedded into the pair of sets assigned to it. For every other subtree, its root is mapped to an appropriate 'connecting' set and the remainder of that subtree is embedded into its pair of sets, which, as we arranged before, induces a bipartite expander in G''.

Our general embedding strategy is somewhat similar to the one in [16]. First, the authors of [16], split the tree T into a constant number of subtrees; secondly, after applying Szemerédi's regularity lemma, they find a perfect matching in the reduced graph; finally, they embed the bulk of T into the 'super-edges' of that perfect matching, in such a way that they can connect all the pieces. However, embedding large trees into regular pairs with

[1] Although we do not give a particular value of n_0, the existence of such a constant will become clear from the proof. The lower bound on n_0 comes mainly from the fact that we apply Szemerédi's regularity lemma to G; additional requirements on the largeness of n_0 are discussed in a footnote at the end of Section 3.4.

vanishing densities is considerably harder than in the constant density case [16], comprising the major difficulty we have to overcome.

3.2 Preparing G'

Since G' is (η, p)-upper-uniform and n is large, we may apply Szemerédi's regularity lemma (Lemma 5) with ε as above and $k_0 := \lceil 1/\varepsilon \rceil$. Let $(V_0, \ldots, V_k)$ be the resulting (ε, p)-regular partition of $V(G')$, and recall that $k \leq K_0$ by the definition of K_0. Define an auxiliary graph H' on the vertex set $\{V_1, \ldots, V_k\}$ as follows. For all i and j with $1 \leq i < j \leq k$, the pair $\{V_i, V_j\}$ will be an edge in H' if and only if the p-density of the pair (V_j, V_j) in G' is at least $\gamma/6$.

Claim 13. *The minimum degree in H' is at least $(1/2 + 2\gamma/3)k$.*

Proof. Fix some $i \in \{1, \ldots, k\}$ and let $V := V(G')$. Since G was η-uniform with density p, $|V_i| \geq n/2k > \eta n$ and $1/k \leq \varepsilon$, then

$$\begin{aligned} e_G(V_i, V - V_0 - V_i) &\geq (1-\eta)p|V_i|(n - |V_0| - |V_i|) \geq (1-\eta)p|V_i|(n - \varepsilon n - n/k) \\ &\geq (1 - \eta - \varepsilon - 1/k)pn|V_i| \geq (1 - 3\varepsilon)pn|V_i|, \end{aligned}$$

and

$$\begin{aligned} \sum_{v \in V_i} \deg_G(v) &= e_G(V_i, V - V_i) + 2e_G(V_i) \leq (1+\eta)p\left[|V_i|(n - |V_i|) + 2\binom{|V_i|}{2}\right] \\ &\leq (1+\eta)p|V_i|(n - |V_i| + |V_i| - 1) \leq (1+\varepsilon)pn|V_i|. \end{aligned}$$

Since our adversary deleted at most $(1/2 - \gamma)\deg_G(v)$ edges at every vertex v, the number of edges of G' that leave the set V_i can be bounded as follows:

$$\begin{aligned} e_{G'}(V_i, V - V_0 - V_i) &\geq e_G(V_i, V - V_0 - V_i) - (1/2 - \gamma) \cdot \sum_{v \in V_i} \deg_G(v) \\ &\geq (1/2 + \gamma - 7\varepsilon/2)pn|V_i|. \end{aligned}$$

Recall that $i \in \{1, \ldots, k\}$ is fixed. The total number of edges in all pairs (V_i, V_j) with $j \in \{1, \ldots, k\} - \{i\}$ whose density is smaller than $\gamma/6$ is at most $\gamma/6 \cdot pn|V_i|$. Moreover, the η-upper-uniformity of G' implies that $e_{G'}(V_i, V_j) \leq (1+\eta)p|V_i||V_j| \leq (1+\varepsilon)p(n/k)|V_i|$ for all $j \neq i$. Therefore,

$$\delta(H') \geq (1/2 + 5\gamma/6 - 7\varepsilon/2)(1+\varepsilon)^{-1}k \geq (1/2 + 5\gamma/6 - 5\varepsilon)k \geq (1/2 + 2\gamma/3)k.$$

□

Now, delete from H' all edges that correspond to pairs (V_i, V_j) that are not (ε, p)-regular in G' and let H'' be the subgraph of H' induced by the set of vertices whose degree in H' after that deletion exceeds $(1/2 + \gamma/2)k$.

Claim 14. *The graph H'' has at least $(1 - \alpha/8)k$ vertices and $\delta(H'') \geq (1/2 + \gamma/3)k$.*

Proof. Since H' contains at most $\varepsilon\binom{k}{2}$ edges corresponding to non-(ε, p)-regular pairs, their deletion lowers the degree sum of H' by no more than εk^2. Since $\delta(H') \geq (1/2 + 2\gamma/3)k$, the degree of at most $(6\varepsilon/\gamma)k$ vertices will fall below the $(1/2+\gamma/2)k$ threshold after the deletion. Recall that $\varepsilon \leq \min\{\gamma^2/36, \alpha\gamma/96\}$, and thus H'' will have at least $(1 - \alpha/8)k$ vertices, and its minimum degree will satisfy $\delta(H'') \geq (1/2 + \gamma/2)k - (6\varepsilon/\gamma)k \geq (1/2 + \gamma/3)k$. □

Let $k' := |V(H'')|$ and let $m' := \lfloor k'/2 \rfloor$. Since $\delta(H'') > k'/2$, H'' contains a matching of size m'. Fix any such matching M and denote its edges by $\{A_1, B_1\}, \dots \{A_{m'}, B_{m'}\}$. Finally, let G'' denote the subgraph of G' induced by the union of all vertices of H'' (which are clusters in G'). Let $n' := |V(G'')|$ and note that $n' \geq (1 - \alpha/8)(1 - \varepsilon)n \geq (1 - \alpha/4)n$.

3.3 Partitioning the tree

Every partition of the vertex set of a tree into connected subsets gives rise to a natural tree structure on the set of parts. Namely, we make two parts adjacent if and only if the subtrees they induce in the original tree are joined by an edge. Let us call this tree the *cluster tree* of our partition. The following general lemma will be crucial in the remainder of the proof.

Lemma 15. *Let t and D be positive integers with $D \geq 2$. Let T be a rooted tree with t vertices and maximum degree at most D. If $\beta \geq 1/t$, then there exists a partition of $V(T)$ into at most $4/\beta$ rooted subtrees of size at most $D^2\beta t$ each such that the maximum degree of the corresponding cluster tree does not exceed D^3 and all children of the root of each subtree belong to the same subtree (the subtree containing that root).*

The proof of Lemma 15 will make use of the following simple statement, whose proof is a straightforward modification of the proof of Proposition 4.2 from [2].

Proposition 16. *Let s and D be positive integers with $D \geq 2$. Let T be a tree with maximum degree at most D and S be a subset of $V(T)$ containing at least $s + 1$ vertices. Then there exists an edge $e \in E(T)$ such that at least one of the two trees obtained from T by deleting e contains at least s and at most $(D - 1)(s - 1) + 1$ vertices from S.*

Proof of Lemma 15. We will construct the required partition in three stages. The first stage will guarantee that the subtrees in our partition are not too large, i.e., they contain no more than $D\beta t$ vertices each. In the second stage we will refine the partition to reduce the maximum degree in the cluster tree to at most D^2. In the third stage we will merge some subtrees to guarantee that each root has children only in its own subtree, and we will do it in such a way that neither the upper bound on the sizes of the subtrees nor the maximum degree of the cluster tree grow more than by a factor of D, and hence in the end they are bounded by $D^2\beta t$ and D^3, respectively.

Stage 1. Start with the trivial partition of $V(T)$ into a single set. We will keep refining it until all parts are small enough, making sure that at all times at most one of the parts is larger than $D\beta t$ and at most one of the parts is smaller than βt. Clearly, our initial partition has that property. Suppose that our partition contains a subtree T' with more than $D\beta t$ vertices. Proposition 16 guarantees that T' contains an edge that splits it into two trees, one of which has at least βt and at most $(D - 1)(\lceil \beta t \rceil - 1) + 1 \leq D\beta t$ vertices. We refine our partition by replacing T' with these two trees. Finally, we iterate this procedure until all parts have at most $D\beta t$ vertices and all but at most one has at least βt vertices. Denote that partition by Π. Clearly, the number of parts is at most $1/\beta + 1$.

Stage 2. Let T_Π be the cluster tree corresponding to the partition Π and let

$$E(\Pi) := \sum_{V \in \Pi} \max\{0, \deg_{T_\Pi}(V) - D^2\}.$$

Clearly,

$$0 \le E(\Pi) \le \sum_{V \in \Pi} \deg_{T_\Pi}(V) = 2(r-1) \le 2/\beta.$$

Suppose that the partition Π does not satisfy our maximum degree requirement, i.e., $\Delta(T_\Pi) > D^2$. Then there must be some $V \in \Pi$ whose degree in T_Π is larger than D^2. Let S be the set of all neighbors of V in T outside of V. Clearly, $|S| = \deg_{T_\Pi}(V)$. Finally, let T' be the subtree of T induced by $V \cup S$. By Proposition 16, T' contains an edge e whose deletion splits T' into two trees, one of which contains at least D and at most $(D-1)^2 + 1 \le D^2 - 1$ vertices from S, see Figure 1. Note that none of the endpoints of e lies in S, or otherwise the two trees would contain 1 and $|S| - 1$ vertices from S respectively, and this is impossible since $1 < D$ and $|S| - 1 > D^2 - 1$. Hence, e partitions V into two connected subsets V' and V''. Let Π' be the partition obtained from Π by replacing V with V' and V''. Note that

$$\deg_{T_{\Pi'}}(V') + \deg_{T_{\Pi'}}(V'') = \deg_{T_\Pi}(V) + 2,$$

and by the choice of e, either $\deg_{T_{\Pi'}}(V')$ or $\deg_{T_{\Pi'}}(V'')$ is at least $D + 1$ and at most D^2. Hence, $E(\Pi') < E(\Pi)$. It follows that by refining our initial partition Π at most $2/\beta$ times, each time increasing the number of parts by one, we will arrive at a partition Π^* with $E(\Pi^*) = 0$. Clearly, the maximum degree of the cluster tree T_{Π^*} is at most D^2, and the number of parts is not greater than $1/\beta + 1 + 2/\beta$, which is at most $4/\beta$.

Stage 3. Root the cluster tree T_{Π^*} at the subset containing the root of the original tree. Order the subsets in Π^* in such a way that all descendants (in the cluster tree) of every subset come later in the ordering (e.g., by performing a breadth-first search on T_{Π^*}). Now, iterate the following procedure until the 'children of the root only in its own tree' condition is satisfied. If there is a subtree whose root has children in other subtrees, merge the first (with respect to our order) such tree with the subtrees containing children of its root and remove all the merged pieces from the ordered list. Clearly, after the procedure terminates, the new partition satisfies the required condition. As a consequence of our 'parent-first' ordering, each tree can be merged only once and hence both the upper bound on the sizes of the parts and the maximum degree of the cluster tree can increase at most D times. Finally, the number of parts in the partition can only become smaller. $\square$

Recall that the number n' of vertices in the graph G'', which is an induced subgraph of G', satisfies $n' \ge (1 - \alpha/4)n$. Root our tree T at an arbitrary vertex. Since T has fewer than n' and more than $n/2$ vertices, by Lemma 15, where we let $\beta := \delta/k'$, there is a partition Π of $V(T)$ into connected subsets $S_1, \ldots, S_\tau$ such that

$$\tau \le 4k'/\delta, \quad \text{and} \quad \max_{1 \le j \le \tau} |S_j| \le D^2 \delta \cdot n'/k', \tag{3}$$

and the maximum degree of the cluster tree T_Π does not exceed D^3. Moreover, assume that the subtrees S_j are ordered in such a way that all descendants (in the cluster tree) of every subtree come later in the ordering. Since each S_j induces a connected bipartite subgraph of T (a subtree of T), it can be uniquely decomposed into two independent sets $S_{j,1}$ and $S_{j,2}$ – the color classes in the unique proper 2-coloring of the tree T restricted to S_j. Let $\mathcal{S}$ be the collection of all these color classes, i.e., $\mathcal{S} := \{S_{j,l} : 1 \le j \le \tau,\ l \in \{1, 2\}\}$. Also, let $\mathcal{E}$ be the set of all pairs $\{S_{j,l}, S_{j',l'}\}$ such that $j \ne j'$ and in T there is an edge joining $S_{j,l}$ and $S_{j',l'}$. Finally, note that the graph obtained from the graph $(\mathcal{S}, \mathcal{E})$ by identifying all pairs $\{S_{j,1}, S_{j,2}\}$ is the cluster tree T_Π. It follows that $|\mathcal{E}| = \tau - 1$. For better visualization, the reader is encouraged to consult Figure 1.

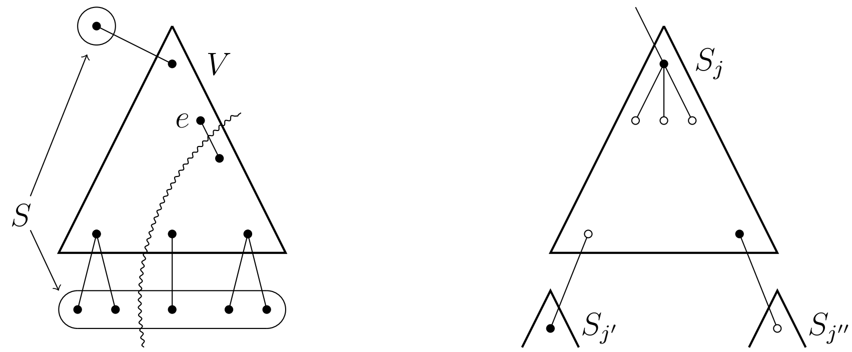

Figure 1: On the left, the edge e splits $T[V \cup S]$ into two trees, partitioning the set S. On the right, a typical subtree S_j; vertices in color classes 1 and 2 are drawn as white and black circles, respectively. Note that all children of the root of S_j belong to S_j. Since $S_{j'}$ and $S_{j''}$ are below S_j in T_Π, we have $j', j'' > j$. Finally, observe that $\{S_{j,1}, S_{j',2}\}, \{S_{j,2}, S_{j'',1}\} \in \mathcal{E}$.

3.4 Planning out the embedding

Recall that M is a maximum matching in the cluster graph H'', and $V(M) = \{A_1, B_1, \ldots, A_{m'}, B_{m'}\}$. Our plan is to embed (most of) the tree T into regular pairs forming M. We will do it piece-by-piece, according to the partition Π. Before we start the actual embedding, we need to lay out a plan in order to make sure that we will never run out of vacant vertices or edges. We start by assigning to each edge in the matching a collection of subtrees of T that we plan to embed in the (ε, p)-regular pair in G'' that is represented by this edge.

Lemma 17. *There is an assignment $\varphi : \mathcal{S} \to V(M)$ with the following two properties.*

1. *For each j, there is an i such that the sets $S_{j,1}$ and $S_{j,2}$ are assigned to two different clusters in the pair $\{A_i, B_i\}$.*

2. *Let $X \in V(M)$ and let $\mathcal{S}(X)$ be the family of sets that φ assigns to X. Define the usage $U(X)$ of the cluster X by*

$$U(X) := \sum_{S \in \mathcal{S}(X)} \left(|S| + 4D^3\varepsilon \cdot \frac{n'}{k'} \right).$$

Then $U(X) \le (1 - \alpha/4) \cdot \frac{n'}{k'}$ for all $X \in V(M)$.

Proof. We can easily construct such a map φ using the following greedy procedure. Start with an empty map φ_0. Assume that $1 \le j \le \tau$ and we have already defined φ_{j-1}. For a cluster $X \in V(M)$, define the usage of X at step $j-1$ as

$$U_{j-1}(X) := \sum_{S \in \mathcal{S}_{j-1}(X)} \left(|S| + 4D^3\varepsilon \cdot \frac{n'}{k'} \right),$$

where $\mathcal{S}_{j-1}(X)$ is the family of sets from $\mathcal{S}$ that φ_{j-1} assigns to X. We claim that there exists an $i \in \{1, \ldots, m'\}$ such that

$$\max\{U_{j-1}(A_i), U_{j-1}(B_i)\} \le (1 - \alpha/4 - 4D^3\varepsilon - D^2\delta) \cdot \frac{n'}{k'}. \tag{4}$$

We postpone the verification of this claim till the end of the proof of Lemma 17. Let $i(j)$ to be the smallest i satisfying (4). Let φ_j be an extension of φ_{j-1} that maps the pair $\{S_{j,1}, S_{j,2}\}$ to $\{A_{i(j)}, B_{i(j)}\}$ in such a way that the smaller of the sets $S_{j,1}, S_{j,2}$ is mapped to the cluster with larger usage U_{j-1} (we break ties arbitrarily). Finally, put $\varphi = \varphi_\tau$.

Note that the way we construct φ guarantees that condition 1 is satisfied. Moreover, since by (3), for all j and l, we have $|S_{j,l}| \leq |S_j| \leq D^2\delta \cdot n'/k'$, it follows that $U_j(X) - U_{j-1}(X) \leq (D^2\delta + 4D^3\varepsilon) \cdot n'/k'$ for every $X \in \{A_{i(j)}, B_{i(j)}\}$ (and clearly $U_j(X) = U_{j-1}(X)$ if $X \notin \{A_{i(j)}, B_{i(j)}\}$), and hence the choice of $i(j)$ at each step guarantees that φ will satisfy condition 2 as well. The only thing we still have to check is that for all j, the index $i(j)$ is well-defined, i.e., inequality (4) is satisfied for some i.

Observe that our strategy of balancing the usage of each pair $\{A_i, B_i\}$ guarantees that for all i and j,

$$\big|U_j(A_i) - U_j(B_i)\big| \leq \max_{j' \leq j} \big||S_{j',1}| - |S_{j',2}|\big| \leq \max_{j' \leq j} |S_{j'}| \leq D^2\delta \cdot \frac{n'}{k'}.$$

Hence, if for some j, inequality (4) is not satisfied for all i, then

$$U_j(X) \geq (1 - \alpha/4 - 4D^3\varepsilon - 2D^2\delta) \cdot \frac{n'}{k'}$$

for all $X \in V(M)$. Recall that $m' = \lfloor k'/2 \rfloor$. It follows that

$$\begin{aligned} \sum_{j' \leq j} |S_{j'}| &= \sum_{X \in V(M)} \left(U_j(X) - 4D^3\varepsilon \cdot \frac{n'}{k'}\right) \geq (1 - \alpha/4 - 8D^3\varepsilon - 2D^2\delta) \cdot \frac{2m'}{k'} \cdot n' \\ &\geq (1 - \alpha/4 - 8D^3\varepsilon - 2D^2\delta - 1/k') \cdot n' \geq (1 - \alpha/2) \cdot n' \geq (1 - 3\alpha/4) \cdot n. \end{aligned}$$

This would be a clear contradiction, since $(1 - \alpha) \cdot n \geq |V(T)| = \sum_j |S_j|$. □

Let φ be a map satisfying both conditions in Lemma 17. Our next step will be planning out connections between all the subtrees in our partition Π, whose locations in the graph G' have already been determined by φ.

Lemma 18. *There is an assignment $\psi : \mathcal{E} \to V(H'')$ with the following two properties.*

1. *For all $e \in \mathcal{E}$, the following holds. Suppose that $e = \{S_{j,l}, S_{j',l'}\}$, where $j < j'$. Then $\psi(e)$ is a common neighbor in H'' of the clusters $\varphi(S_{j,l})$ and $\varphi(S_{j',3-l'})$.*

2. *Every cluster is assigned to at most $6/(\gamma\delta)$ edges in $\mathcal{E}$, i.e., $|\psi^{-1}(X)| \leq 6/(\gamma\delta)$ for all $X \in V(H'')$.*

Proof. We construct such a map greedily, starting from the empty map and extending it one-by-one to the whole set $\mathcal{E}$. Let $e = \{S_{j,l}, S_{j',l'}\} \in \mathcal{E}$, where $j < j'$. Since $\delta(H'') \geq (1/2 + \gamma/3)k \geq (1/2 + \gamma/3)k'$, the clusters $\varphi(S_{j,l})$ and $\varphi(S_{j',3-l'})$ have at least $2\gamma k'/3$ common neighbors. One of them has been used fewer than $|\mathcal{E}|/(2\gamma k'/3) \leq \tau/(2\gamma k'/3) \leq 6/(\gamma\delta)$ times, where the second inequality follows from (3). We let $\psi(e)$ be an arbitrary cluster with that property. □

Now that we have laid out a general plan for the embedding, it is time to assign to each $S_{j,l}$ a particular subset of $V(G'')$, where we will map $S_{j,l}$. We start by choosing in each cluster $X \in V(H'')$ an arbitrary subset $C(X)$ of size $\alpha/8 \cdot n'/k'$. Let e_X be the number of edges in

$\mathcal{E}$ that are assigned to X. We partition $C(X)$ into e_X subsets of equal sizes and label those subsets with elements of $\psi^{-1}(X)$ such that each $e \in \psi^{-1}(X)$ gets its own set $\psi'(e)$ of size at least $\alpha/(4e_X) \cdot n'/k'$, which is at least $\alpha\gamma\delta/48 \cdot n'/k'$.

Next, fix a cluster $X \in V(M)$. For each $S \in \mathcal{S}(X)$, we choose a subset $\varphi'(S)$ of $X - C(X)$ with size $|S| + 4D^3\varepsilon \cdot n'/k'$ such that all these sets are disjoint. By the choice of φ, which satisfies condition 2 in Lemma 17, this is possible. We do this for all clusters in $V(M)$[2].

Finally, note that for all j, $\varphi'(S_{j,1})$ and $\varphi'(S_{j,2})$ are subsets of opposite clusters in an (ε, p)-regular pair in G'' with p-density at least $\gamma/6$ and for each $e = \{S_{j,l}, S_{j',l'}\}$, where $j < j'$, also $\{\varphi'(S_{j,l}), \psi'(e)\}$ and $\{\varphi'(S_{j',3-l'}), \psi'(e)\}$ are pairs of subsets of opposite classes in an (ε, p)-regular pair in G'', whose p-density is at least $\gamma/6$.

3.5 Cleaning up G''

Recall that we have ordered the subtrees in our partition in such a way that all descendants of a tree S in the cluster tree come later in the ordering, i.e., if $S_{j'}$ is a descendant of S_j, then $j' > j$. Our goal in the cleaning-up stage is the following.

Goal. Construct functions $\varphi'' : \mathcal{S} \to \mathcal{P}(V(G''))$ and $\psi'' : \mathcal{E} \to \mathcal{P}(V(G''))$ with the following properties.

1. For all $S \in \mathcal{S}$, $\varphi''(S) \subseteq \varphi'(S)$ and $|\varphi''(S)| \geq |S| + (2D+1)\varepsilon \cdot n'/k'$.
2. For all j, the graph $(\varphi''(S_{j,1}), \varphi''(S_{j,2}))$ is a bipartite $(\varepsilon \cdot n'/k', 2D+2)$-expander.
3. For all $e \in \mathcal{E}$, the following holds. Suppose that $e = \{S_{j_1,l_1}, S_{j_2,l_2}\}$, where $j_1 < j_2$. Then:
   (a) $\psi''(e) \subseteq \psi'(e)$ and $|\psi''(e)| \geq \varepsilon \cdot n'/k'$,
   (b) each vertex in $\varphi''(S_{j_1,l_1})$ has a neighbor in $\psi''(e)$,
   (c) each vertex in $\psi''(e)$ has at least $D+1$ neighbors in $\varphi''(S_{j_2,3-l_2})$.

In the process of achieving our goal, we will extensively use the following two technical lemmas.

Lemma 19. *Let (A, B) be an (ε, p)-regular pair, whose p-density is larger than ε. Suppose that $|A| = |B| = n'/k'$, and $A' \subseteq A$ and $B' \subseteq B$ are sets of size at least $(4D+6)\varepsilon \cdot n'/k'$. Then there are subsets $A'' \subseteq A'$ and $B'' \subseteq B'$ satisfying the following two conditions.*

1. *$|A' - A''| \leq \varepsilon \cdot n'/k'$ and $|B' - B''| \leq \varepsilon \cdot n'/k'$.*
2. *The subgraph (A'', B'') is a bipartite $(\varepsilon \cdot n'/k', 2D+2)$-expander.*

Proof. We will greedily construct such subsets A'' and B''. Before we start, we would like to remark that all neighborhoods are computed in the subgraph (A', B'), and not the graph (A, B) itself. First, let $X := \emptyset$ and $Y := \emptyset$. We will iterate the following procedure. If there is a set $X' \subseteq A' - X$, with $|X'| \leq \varepsilon \cdot n'/k'$, such that $|N(X') - Y| < (2D+2)|X'|$, then

[2] For the sake of clarity of the presentation we tacitly assumed that the numbers $\alpha/(4e_X) \cdot n'/k'$ and $4D^3\varepsilon \cdot n'/k'$ were integers. This is clearly not true in general, but since we assume that n' is large, we can utilize the remaining $\alpha/8 \cdot n'/k'$ unused vertices in each cluster to account for all rounding errors, as the number of sets $\varphi'(S)$ and $\psi'(e)$ is independent of n.

set $X := X \cup X'$. Similarly, if there is a set $Y' \subseteq B' - Y$, with $|Y'| \leq \varepsilon \cdot n'/k'$ such that $|N(Y') - X| < (2D+2)|Y'|$, then set $Y := Y \cup Y'$.

First we show that at all times $|X| \leq \varepsilon \cdot n'/k'$ and $|N(X) - Y| \leq (2D+2)|X|$, and similarly, $|Y| \leq \varepsilon \cdot n'/k'$ and $|N(Y) - X| \leq (2D+2)|Y|$. Certainly, this is true at the beginning of the procedure, since then $|X| = |Y| = |N(X) - Y| = |N(Y) - X| = 0$. Suppose that all four inequalities hold at the beginning of some iteration. Assume that the procedure finds an $X' \subseteq A' - X$ with $|X'| \leq \varepsilon \cdot n'/k'$ and $|N(X') - Y| < (2D+2)|X'|$. Then

$$\begin{aligned} |N(X \cup X') - Y| &= |(N(X) - Y) \cup (N(X') - Y)| \leq |N(X) - Y| + |N(X') - Y| \\ &\leq (2D+2)|X| + (2D+2)|X'| = (2D+2)|X \cup X'|. \end{aligned}$$

Note that in (A, B) there are no edges between $X \cup X'$ and $B' - N(X \cup X')$. Also, since $|X \cup X'| = |X| + |X'| \leq 2\varepsilon \cdot n'/k'$, then

$$|B' - N(X \cup X')| \geq |B'| - |N(X \cup X') - Y| - |Y| \geq |B'| - (4D+5)\varepsilon \cdot n'/k' \geq \varepsilon \cdot n'/k'.$$

Since (A, B) was (ε, p)-regular with p-density larger than ε, and $|A| = |B| = n'/k'$, it must be that $|X \cup X'| < \varepsilon \cdot n'/k'$. A symmetric argument proves the other two inequalities.

Now, put $A'' := A' - X$ and $B'' := B' - Y$. We have already proved that condition 1 holds for this choice of A'' and B''. As for the other condition, the definition of X and Y guarantees that all small subsets of A'' and B'' expand at least $2D+2$ times. It suffices to prove that also large sets expand well enough. Suppose that there is an $X' \subseteq A''$ with $|X'| \geq \varepsilon \cdot n'/k'$ such that $|N(X') \cap B''| < |B''| - \varepsilon \cdot n'/k'$. There are no edges in (A, B) between the sets X' and $B'' - N(X')$, but this is impossible, since (A, B) is (ε, p)-regular with p density larger than ε, and both sets are larger than $\varepsilon \cdot n'/k'$. □

Lemma 20. *Let $b \geq 1$ and let (A, B) be an (ε, p)-regular pair, whose p-density is larger than ε. Suppose that $|A| = |B| = n'/k'$, and $A' \subseteq A$ and $B' \subseteq B$ are sets of size at least $2\varepsilon \cdot n'/k'$ and $b\varepsilon \cdot n'/k'$ respectively. Then there is a subset $A'' \subseteq A'$ such that $|A' - A''| \leq \varepsilon \cdot n'/k'$ and every vertex in A'' has at least b neighbors in B'.*

Proof. Let $X \subseteq A'$ be the set of all vertices in A that have fewer than b neighbors in B and put $A'' = A' - X$. If $|X| \leq \varepsilon \cdot n'/k'$, then there is nothing left to prove. Otherwise, let X' be an arbitrary subset of X of size $\varepsilon \cdot n'/k'$. Clearly, there are no edges in (A, B) between X' and $B' - N(X')$. This is impossible, since (A, B) is (ε, p)-regular with p-density larger than ε, $|X'| \geq \varepsilon \cdot n'/k'$ and by the definition of X, we have $|B' - N(X')| \geq |B'| - (b-1)|X'| \geq \varepsilon \cdot n'/k'$. □

An immediate consequence of Lemma 20 is the following Corollary.

Corollary 21. *Let $d \geq 1$ and let $(A, B_1), \ldots, (A, B_d)$ be (not necessarily distinct) (ε, p)-regular pairs in G''. Suppose that $|A| = |B_1| = \ldots = |B_d| = n'/k'$, $A' \subseteq A$ is a set of size at least $(d+1)\varepsilon \cdot n'/k'$ and $B'_i \subseteq B_i$ are sets of size at least $\varepsilon \cdot n'/k'$ for each $i \in \{1, \ldots, d\}$. Then there is a subset $A'' \subseteq A'$ such that $|A' - A''| \leq d\varepsilon \cdot n'/k'$ and every vertex in A'' has a neighbor in each B'_i.*

We start cleaning up by setting $\varphi'' := \varphi'$ and $\psi'' := \psi'$. Next, we will iteratively, starting with $j := t$ and each time reducing j by one, keep fixing the two functions by making sure that after we have finished step j, the requirements 1, 2 and 3b are met as long as they involve only sets $S_{j',l'}$ with $j' \geq j$ and $l' \in \{1, 2\}$ (i.e., $j_1 \geq j$ in 3), and the requirements 3a and 3c are met as long as they involve sets $\{S_{j_1,l_1}, S_{j_2,l_2}\}$ with $\max\{j_1, j_2\} \geq j$. If we manage to do that, after completing the final step ($j = 1$) our goal will be reached.

Assume that we are at a step j and our functions φ'' and ψ'' satisfy all requirements involving only sets $S_{j',l'}$ with $j' > j$ and requirements 3a and 3c, where $j_2 > j$. Let $\mathcal{D}(S_{j,1})$ and $\mathcal{D}(S_{j,2})$ be the families of the color classes of all the children (in the cluster tree T_Π) of S_j that are adjacent to the color classes $S_{j,1}$ and $S_{j,2}$, respectively. In other words, $S_{j',l'} \in \mathcal{D}(S_{j,l})$ if and only if $S_{j'}$ is a child of S_j in the cluster tree T_Π, and the edge connecting S_j and $S_{j'}$ in T has endpoints in the sets $S_{j,l}$ and $S_{j',l'}$. For each $l \in \{1, 2\}$, the following is true. Fix a set $S_{j',l'} \in \mathcal{D}(S_{j,l})$ and let $e = \{S_{j,l}, S_{j',l'}\}$. Since $S_{j',l'}$ is a descendant of $S_{j,l}$ in the cluster tree, we have $j' > j$, and therefore $|\psi''(e)| \geq \varepsilon \cdot n'/k'$. Since $|\mathcal{D}(S_{j,l})| \leq \Delta(T_\Pi) \leq D^3$, and $|\varphi'(S_{j,l})| \geq |S_{j,l}| + 4D^3\varepsilon \cdot n'/k'$, by Corollary 21, there is a subset $A'_l \subseteq \varphi'(S_{j,l})$ of size at least $|S_{j,l}| + 3D^3\varepsilon \cdot n'/k'$ such that every vertex in A'_l has a neighbor in every $\psi''(\{S_{j,l}, S\})$, for each $S \in \mathcal{D}(S_{j,l})$.

By Lemma 19, there are subsets $A''_1 \subseteq A'_1$ and $A''_2 \subseteq A'_2$ of sizes at least $|S_{j,1}| + 2D^3\varepsilon \cdot n'/k' \geq |S_{j,1}| + (2D + 1)\varepsilon \cdot n'/k'$ and $|S_{j,2}| + 2D^3\varepsilon \cdot n'/k' \geq |S_{j,2}| + (2D + 1)\varepsilon \cdot n'/k'$, respectively, such that the induced graph (A''_1, A''_2) is a bipartite $(\varepsilon \cdot n'/k', 2D + 2)$-expander. We put $\varphi''(S_{j,l}) := A''_l$ for both l.

Finally, let l be such that the set $S_{j,l}$ contains the root of the tree $T[S_j]$. If $j \neq 1$, then $S_{j,l}$ has a unique parent $S \in \mathcal{S}$. Let $e := \{S, S_{j,l}\}$. Clearly, e belongs to $\mathcal{E}$. Since $|\psi'(e)| \geq \alpha\gamma\delta/48 \cdot n'/k' \geq 2\varepsilon \cdot n'/k'$, by Lemma 20 there is a subset $A'' \subseteq \psi'(e)$ of size at least $\varepsilon \cdot n'/k'$ such that every vertex in A'' has at least $D + 1$ neighbors in $\varphi''(S_{j,l})$. We put $\psi''(e) := A''$. If $j = 1$, then S_j is the root of T_Π, and there is nothing to do.

3.6 Embedding T into G''

Finally, we are ready to embed our tree T into G''. We will do the actual embedding in a top-down fashion, starting with S_1 and extending our embedding to all other S_j one-by-one. For each j, the subtree $T[S_j]$ will be embedded into the bipartite expanding graph $(\varphi''(S_{j,1}), \varphi''(S_{j,2}))$ with the small exception that, unless $j = 1$, the root of the tree will be embedded into the appropriate 'connecting' set $\psi''(e)$, where $e \in \mathcal{E}$ represents the edge between $T[S_j]$ and its parent in the cluster tree T_Π.

We start by embedding the first subtree, S_1, into the bipartite graph H_1 induced on the pair $(\varphi''(S_{1,1}), \varphi''(S_{1,2}))$. Since $|\varphi''(S_{1,l})| \geq |S_{1,l}| + 2D^3\varepsilon \cdot n'/k'$ for each $l \in \{1, 2\}$, and H_1 is a bipartite $(\varepsilon \cdot n'/k', D + 1)$-expander, Corollary 12 guarantees that this is possible. Suppose we have already embedded $S_1, \ldots, S_{j-1}$ into T in such a way that for every $j' < j$ and $l' \in \{1, 2\}$, all vertices in $S_{j',l'}$, except the root of $T[S_{j'}]$, are mapped into the set $\varphi''(S_{j',l'})$. Let $l \in \{1, 2\}$ be such that the root of $T[S_j]$, call it r_j, is in $S_{j,l}$. Let p_j be the parent of r_j in the tree T and let j' and l' be such that $p_j \in S_{j',l'}$. Finally, let $e := \{S_{j',l'}, S_{j,l}\}$. The way we defined φ'' and ψ'' guarantees that the image of p_j, which by the definition of Π cannot be the root of its tree and hence has not been mapped to a vertex in one of the 'connecting' sets $\psi''(e)$, is in $\varphi''(S_{j',l'})$ and has a neighbor x in the set $\psi''(e)$, and x has at least $D + 1$ neighbors in $\varphi''(S_{j,3-l})$.

Claim 22. *The graph H_j induced on the pair $(\varphi''(S_{j,l}) \cup \{x\}, \varphi''(S_{j,3-l}))$ is a bipartite $(\varepsilon \cdot n'/k' + 1, D + 1)$-expander.*

Proof. For the sake of brevity, let $A := \varphi''(S_{j,l}) \cup \{x\}$, $B := \varphi''(S_{j,3-l})$ and $q := \varepsilon \cdot n'/k'$. Recall that the graph $(A - \{x\}, B)$ is a bipartite $(q, 2D + 2)$-expander. It is easy to check that H_j satisfies all conditions from Definition 11.

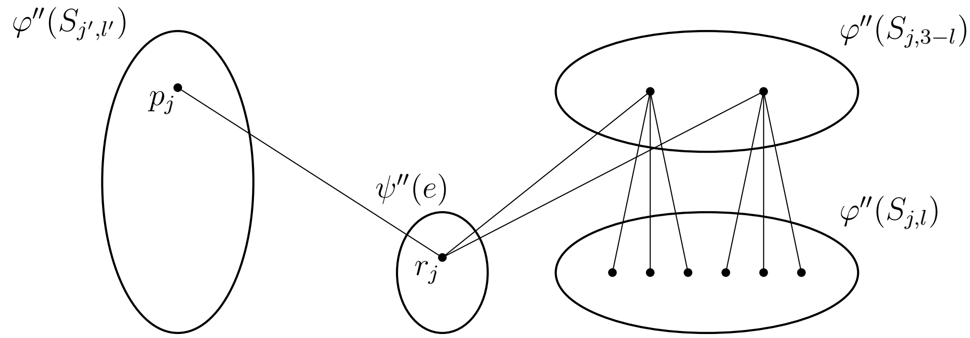

Figure 2: Embedding the subtree S_j into G''

For example, let $X \subseteq A$ be a set of size at most $q+1$. If $x \notin X$, then $|N_{H_j}(X)| \geq (2D+2)\min\{|X|, q\} \geq (D+1)|X|$. If $x \in X$ but $X \neq \{x\}$, then $|N_{H_j}(X)| \geq |N_{H_j}(X - \{x\})| \geq (2D+2)(|X|-1) \geq (D+1)|X|$. Finally, if $X = \{x\}$, then $|N_{H_j}(X)| \geq (D+1)|X|$ by the choice of x.

A similar straightforward case analysis shows that the other conditions from Definition 11 are also satisfied. We omit the details. □

Since $|\varphi''(S_{j,l}) \cup \{x\}| \geq |S_{j,l}| + 2D^3\varepsilon \cdot n'/k'$ and $|\varphi''(S_{j,3-l})| \geq |S_{j,3-l}| + 2D^3\varepsilon \cdot n'/k'$, Corollary 12 says that we can embed $T[S_j]$ in H_j in such a way that r_j is mapped to x. Note that necessarily $S_{j,l} - \{r_j\}$ is mapped to $\varphi''(S_{j,l})$ and $S_{j,3-l}$ is mapped to $\varphi''(S_{j,3-l})$, see Figure 3.6. This completes the proof.

4 Concluding remarks

We would like to mention that our proof of Theorem 3 can be slightly adjusted to yield the following result, which was originally proved by Dellamonica, Kohayakawa, Marciniszyn and Steger [10].

Theorem 23. *Let α and γ be positive constants. There exist η_0 and n_0 (depending on α and γ) with the following property. Let G be an n-vertex (p,η)-uniform graph, with $p > 0$, $\eta < \eta_0$ and $n \geq n_0$. The local resilience of G with respect to having circumference greater than $(1-\alpha)n$ is at least $1/2 - \gamma$.*

In order to prove Theorem 23, one can just follow the proof of Theorem 3 with T being a path of length $(1-\alpha/2)n$, rooted at one of the endpoints. The only difference is that when we partition T and assign the subpaths of T, the sets $S_{j,l}$ in our proof, to the cluster pairs forming our fixed matching M, we make sure that the first and the last subpaths are long enough, e.g., both have $\delta \cdot n'/k'$ vertices, and get assigned to the same cluster pair. After we have embedded T into the graph G', since the first and the last segments of T were mapped to a dense regular pair of clusters in G', we easily find an edge between them, which will close a cycle of length greater than $(1-\alpha)n$.

5 Acknowledgements

Part of this research was done when József Balogh and Wojciech Samotij were supported by the Visiting Scholar Program of the Department of Mathematics at WKU. We would like to thank the anonymous referees for their invaluable comments and suggestions.

References

[1] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi, *Universality and tolerance (extended abstract)*, 41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000), IEEE Comput. Soc. Press, 2000, pp. 14–21.

[2] N. Alon, M. Krivelevich, and B. Sudakov, *Embedding nearly-spanning bounded degree trees*, Combinatorica **27** (2007), 629–644.

[3] N. Alon and J. Spencer, *The probabilistic method. Third edition*, John Wiley & Sons, Inc., 2008.

[4] J. Balogh, B. Csaba, M. Pei, and W. Samotij, *Large bounded degree trees in expanding graphs*, Electronic Journal of Combinatorics **17** (2010), Research paper 6.

[5] B. Bollobás, *Extremal graph theory*, Academic Press, Inc., 1978.

[6] J. Böttcher, Y. Kohayakawa, and A. Taraz, *Almost spanning subgraphs of random graphs after adversarial edge removal*, Electronic Notes in Discrete Mathematics **35** (2009), 335–340.

[7] D. Conlon and T. Gowers, *Combinatorial theorems relative to a random set*, manuscript.

[8] B. Csaba, I. Levitt, J. Nagy-György, and E. Szemerédi, *On embedding trees*, submitted.

[9] D. Dellamonica and Y. Kohayakawa, *An algorithmic Friedman-Pippenger theorem on tree embeddings and applications*, Electronic Journal of Combinatorics **15** (2008), Research paper 127.

[10] D. Dellamonica, Y. Kohayakawa, M. Marciniszyn, and A. Steger, *On the resilience of long cycles in random graphs*, Electronic Journal of Combinatorics **15** (2008), Research Paper 32.

[11] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society (3) **2** (1952), 69–81.

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

[13] A. Frieze and M. Krivelevich, *On two Hamilton cycle problems in random graphs*, Israel Journal of Mathematics **166** (2008), 221–234.

[14] P. Haxell, *Tree embeddings*, Journal of Graph Theory **36** (2001), 121–130.

[15] Y. Kohayakawa, *Szemerédi's Regularity Lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.

[16] J. Komlós, G. Sárközy, and E. Szemerédi, *Proof of a packing conjecture of Bollobás*, Combinatorics, Probability and Computing **4** (1995), 241–255.

[17] M. Krivelevich, C. Lee, and B. Sudakov, *Resilient pancyclicity of random and pseudo-random graphs*, SIAM Journal on Discrete Mathematics **24** (2010), 1–16.

[18] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), 359–364.

[19] M. Schacht, *Extremal results for random discrete structures*, manuscript.

[20] B. Sudakov and V. Vu, *Local resilience of graphs*, Random Structures and Algorithms **33** (2008), 409–433.

[21] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.