

# Local resilience of almost spanning trees in random graphs

József Balogh\*      Béla Csaba†      Wojciech Samotij‡

October 2, 2009

## Abstract

We prove that for fixed integer  $D \geq 2$  and positive reals  $\alpha, \gamma > 0$ , there exists a constant  $C_0 = C_0(\alpha, \gamma, D)$ , such that for all  $p = p(n) \geq C_0/n$ , the random graph  $G(n, p)$  will almost surely contain a copy of every tree with maximum degree at most  $D$  and at most  $(1 - \alpha)n$  vertices, even after we delete  $(1/2 - \gamma)$  fraction of 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 [9] 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 [4] 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 edges incident to some vertex. This was proved by Komlós, Sárközy and Szemerédi [13], and recently improved by Csaba, Levitt, Nagy-György and Szemerédi [6], who showed that  $\gamma n$  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. [13] in the random setting.

Looking from the edge deletion perspective, the theorem of Komlós et al. can be viewed in a 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  $\mathcal{P}$ . This natural concept, especially suited for

---

\*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.

monotone graph properties, was recently formalized by Sudakov and Vu [16], 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 [17] or the theorem of Dirac [8] 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 a more general setting of random and pseudo-random graphs, see [5, 7, 10, 14, 16].

Alon, Krivelevich and Sudakov [1] proved that for all  $D \geq 2$  and  $\alpha > 0$ , there exists a constant  $C = C(\alpha, D)$ , such that the random graph  $G(n, C/n)$  almost surely contains all trees with maximum degree at most  $D$  on  $(1 - \alpha)n$  vertices. Recently, using a different argument, Balogh, Csaba, Pei and Samotij [3] showed that  $C$  can be as small as  $O(d \log d + d/\alpha \log(1/\alpha))$ . The main result of this paper generalizes the above theorem – we prove that almost surely 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  $\alpha$  and  $\gamma$  be positive constants, and  $D \geq 2$ . There exists a constant  $C_0 = C_0(\alpha, \gamma, D)$ , such that for all  $p = p(n) \geq C_0/n$ , 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  $(1/2 - \gamma)$ .*

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 [7] proved that for every positive constant  $\gamma$ , 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 with high probability by deleting at most  $(1/2 + \gamma)$  fraction of edges incident to each vertex, the random graph  $G(n, p)$  can be made 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 much simpler argument than the one used in [7] proves that the  $(1/2 - \gamma)$  constant 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, almost surely 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 with high probability, 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  $p = p(n) \leq C/n$ , then there is a constant  $\alpha = \alpha(C) > 0$ , such that with high probability the size of the largest connected component of  $G(n, p)$  does not exceed  $(1 - \alpha)n$ . Moreover, if  $\omega = \omega(n) \rightarrow \infty$  as  $n \rightarrow \infty$ , almost surely  $G(n, p)$  contains  $o(n/\omega)$  vertices with degree at least  $\omega$ , and hence it cannot contain all trees with maximum degree  $\omega$ . 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 a  $(\eta, p)$ -uniform graph is given at the beginning of Section 2.

**Theorem 3.** *Let  $\alpha$  and  $\gamma$  be positive constants, and  $D \geq 2$ . There exist  $\eta_0 = \eta_0(\alpha, \gamma, D)$  and  $n_0 = n_0(\alpha, \gamma, D)$ , such that the following holds. Let  $G$  be an  $n$ -vertex  $(\eta, p)$ -uniform graph, with  $p > 0$ ,  $\eta < \eta_0$  and  $n \geq n_0$ . Let  $G'$  be a subgraph of  $G$ , such that for each vertex  $v$ ,  $\deg_{G'}(v) \geq (1/2 + \gamma) \deg_G(v)$ . 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  $(\eta, p)$ -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. Using the tools developed in Section 2, we prove Theorem 3 in Section 3. Finally, Section 4 contains a few concluding remarks.

## 2 Preliminaries and tools

### 2.1 Uniformity of the random graph

Fix positive constants  $\eta > 0$  and  $p > 0$ . We say that an  $n$ -vertex graph  $G$  is  $\eta$ -uniform with density  $p$ , or simply  $(\eta, p)$ -uniform, if, for all  $A, B \subseteq V(G)$  with  $A \cap B = \emptyset$  and  $|A|, |B| \geq \eta n$ , we have

$$(1 - \eta)p|A||B| \leq e_G(A, B) \leq (1 + \eta)p|A||B| \quad (1)$$

and

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

Furthermore, we say that  $G$  is  $\eta$ -upper-uniform with density  $p$ , or simply  $(\eta, p)$ -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 large enough.

**Lemma 4.** *Let  $\eta > 0$  and  $pn > \frac{8}{\eta^4(1-\eta)}$ . Then with high probability the random graph  $G(n, p)$  is  $(\eta, p)$ -uniform.*

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

$$\mathbb{P}(e_G(A, B) - p|A||B| > \eta p|A||B|) + \mathbb{P}(e_G(A, B) - p|A||B| < \eta p|A||B|).$$

By a standard Chernoff-type estimate (see, e.g., [2]), this is at most

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

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

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

By our assumption on  $p$  and the union bound, the probability that  $G$  is not  $\eta$ -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  $p > 0$  be a positive constant. 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  $\varepsilon > 0$  and  $A, B$  are disjoint sets of vertices in  $G$ . We say that the pair  $(A, B)$  is  $(\varepsilon, p)$ -regular if for all  $A' \subseteq A$ ,  $B' \subseteq B$  with  $|A'| \geq \varepsilon|A|$  and  $|B'| \geq \varepsilon|B|$ , we have

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

Finally, we say that a partition  $\Pi = (V_0, \dots, V_k)$  of  $V(G)$  is  $(\varepsilon, p)$ -regular if  $|V_0| \leq \varepsilon|V(G)|$ ,  $|V_i| = |V_j|$  for all  $i, j \in \{1, \dots, k\}$  and at least  $(1 - \varepsilon)\binom{k}{2}$  pairs  $(V_i, V_j)$  with  $1 \leq i < j \leq k$  are  $(\varepsilon, p)$ -regular. We may now state a version of Szemerédi's regularity lemma for  $\eta$ -upper-uniform graphs (see, e.g., [12]).

**Lemma 5.** *For any  $\varepsilon > 0$  and  $k_0 \geq 1$ , there are constants  $\eta = \eta(\varepsilon, k_0) > 0$  and  $K_0 = K_0(\varepsilon, k_0) \geq k_0$  such that any  $\eta$ -upper-uniform graph  $G$  with density  $0 < p \leq 1$  admits an  $(\varepsilon, p)$ -regular partition  $\Pi = (V_0, \dots, V_k)$  with  $k_0 \leq k \leq K_0$ .*

## 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 [15], Friedman and Pippenger [9] proved that all graphs satisfying certain expansion properties contain all small trees with bounded maximum degree.

**Theorem 6** ([9]). *Let  $m$  and  $D$  be positive integers and let  $H$  be a non-empty graph. Moreover, assume that every  $X \subseteq V(H)$  with  $|X| \leq 2m$  satisfies  $|N_H(X)| \geq (D + 1)|X|$ . Then  $H$  contains every tree with  $m$  vertices and maximum degree at most  $D$ .*

An apparent shortcoming of Theorem 6 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 [9], Haxell [11] managed to overcome this problem. Recently, Balogh, Csaba, Pei and Samotij [3] noted that a simplified version of Haxell’s general tree-embedding result allows one to find 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 nor the simplified statement used in [3] do not work very well 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 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  $0 < |X| \leq m_i$ ,  $|N_H(X)| \geq D|X| + 1$  for  $i \in \{1, 2\}$ .*
2. *For every  $X \subseteq V_i$  with  $m_i < |X| \leq 2m_i$ ,  $|N_H(X)| \geq D|X| + M_{3-i}$  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  $H$ .*

*Proof.* A tree  $T$  will be called  $(M_1, M_2, D)$ -small, or simply *small*, if  $\Delta(T) \leq D$  and the two color classes of  $T$ , the sets  $U_1, U_2 \subseteq V(T)$ , satisfy  $|U_1| \leq M_1$  and  $|U_2| \leq M_2$ . Let  $f$  be an embedding of a small tree  $T$  into our expanding graph  $H$ . The *liability* of a vertex  $x \in V(H)$  with respect to  $f$  is the quantity

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

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  $C_f(X) := A_f(X) - B_f(X)$  will be called the *balance* of the set  $X$ . 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  $|X| \leq 2m_i$ . We can easily assume that  $X \neq \emptyset$ , since  $C_f(\emptyset) = 0$ . We have

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

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  $|X| \leq 2m_1$ ,  $C_f(X) = 0$ , then  $|X| \leq m_1$ .*

*Proof of Claim 8.* Assume that some  $X \subseteq V_1$  satisfies  $C_f(X) = 0$ , but  $m_1 < |X| \leq 2m_1$ . Since  $T$  is a small tree,  $|U'_2| = |U_2| - 1 \leq M_2 - 1$ , and thus

$$\begin{aligned} A_f(X) &= |N_H(X) - f(V(S))| = |N_H(X) - f(U'_2)| \geq |N_H(X)| - |f(U'_2)| \\ &\geq D|X| + M_2 - (M_2 - 1) = D|X| + 1 \geq B_f(X) + 1, \end{aligned}$$

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  $A_f(X) = B_f(X)$ .  $\square$

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

*Proof of Claim 9.* Since  $B_f(\cdot)$  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

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

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

**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  $|X_1 \cup \dots \cup X_k| \leq m_1$ .*  $\square$

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  $f(w) \in V_1$ . We can extend  $f$  to an embedding of  $T$  by mapping  $v$  to any vertex in  $Y := \mathcal{A}_f(f(w)) \subseteq V_2$ . Suppose that for no  $y \in Y$ , the extension  $f_y := f \cup \{(v, y)\}$  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  $B_{f_y}(X) \leq B_f(X)$ . 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  $C_{f_y}(X_y) \geq 0$ . By Corollary 10,  $X^* := \cup_{y \in Y} X_y$  has zero balance and size at most  $m_1$ . Moreover, if we let  $X' := X^* \cup \{f(w)\} \subseteq V_1$ ,

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

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  $\deg_S(w) = \deg_T(w) - 1 \leq D - 1$ ,

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

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

Let us now define a class of expanding bipartite graphs that seems to be the most adequate and convenient in our further considerations.

**Definition 11.** Let  $b \geq 2$  and let  $H$  be a bipartite graph with color classes  $V_1$  and  $V_2$ , where  $|V_1| \leq |V_2|$ . Let  $0 < q < |V_1|$ . 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  $|N_H(X)| \geq b|X|$  for  $i \in \{1, 2\}$ .
2. Every subset  $X \subseteq V_i$  of size at least  $q$  satisfies  $|N_H(X)| \geq |V_{3-i}| - q$  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  $D \geq 2$  and let  $H$  be a bipartite graph with color classes  $V_1$  and  $V_2$ , where  $|V_1| \leq |V_2|$ . Suppose that  $H$  is a bipartite  $(q, D+1)$ -expander for some  $0 < q < |V_1|/(2D+1)$ . 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  $m_i := q \geq 1$  and  $M_i := |V_i| - (2D+1)q$  for  $i \in \{1, 2\}$ .  $\square$

### 3 Proof of Theorem 3

We start by defining a few constants. Let

$$\delta := \frac{\alpha}{16D^2}, \quad \varepsilon := \min \left\{ \frac{\alpha}{64D^3}, \frac{\alpha\gamma\delta}{48}, \frac{\gamma^2}{36} \right\}, \quad \text{and} \quad \eta_0 := \min \left\{ \frac{\varepsilon}{2}, \frac{1}{2K_0} \right\},$$

where  $K_0 = K_0(\varepsilon, \lceil 1/\varepsilon \rceil)$  is given by Lemma 5. Let  $G$  be an  $(\eta, p)$ -uniform  $n$ -vertex graph, with  $p > 0$ ,  $\eta < \eta_0$  and  $n$  larger than some constant  $n_0 = n_0(\alpha, \gamma, D)$ .<sup>1</sup> 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  $\eta$ -upper-uniform with density  $p$ . Finally, let  $T$  be a tree with at most  $(1 - \alpha)n$  vertices and maximum degree at most  $D$ . We will show that  $T \subseteq G'$ .

<sup>1</sup>Although we do not give a particular value of  $n_0$ , the existence of such a constant will become clear from the proof.

### 3.1 Preparing $G'$

Since  $G'$  is  $(\eta, p)$ -upper-uniform, we may apply Szemerédi's regularity lemma (Lemma 5) with  $\varepsilon$  as above and  $k_0 := \lceil 1/\varepsilon \rceil$ . Let  $(V_0, \dots, V_k)$  be the resulting  $(\varepsilon, p)$ -regular partition of  $V(G')$ . Define an auxiliary graph  $H'$  on the vertex set  $\{V_1, \dots, V_k\}$  as follows. For  $1 \leq i < j \leq k$ , the pair  $\{V_i, V_j\}$  will be an edge in  $H'$  if and only if the  $p$ -density of the pair  $(V_i, V_j)$  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  $i \in \{1, \dots, k\}$  and let  $V := V(G')$ . Since  $G$  was  $\eta$ -uniform with density  $p$ ,  $|V_i| \geq n/2k > \eta n$  and  $1/k \leq \varepsilon$ , we have

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

and

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

Hence,

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

Furthermore, the total number of edges in all pairs  $(V_i, V_j)$  whose density is smaller than  $\gamma/6$  is at most  $\gamma/6 \cdot pn|V_i|$ . Moreover, the  $\eta$ -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

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

Now, delete from  $H'$  all edges that correspond to pairs  $(V_i, V_j)$  that are not  $(\varepsilon, p)$ -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- $(\varepsilon, p)$ -regular pairs, their deletion lowers the degree sum of  $H'$  by no more than  $\varepsilon k^2$ . 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  $\varepsilon \leq \min\{\gamma^2/36, \alpha\gamma/48\}$ , and thus  $H''$  will have at least  $(1 - \alpha/8)k$  vertices, and its minimum degree,  $\delta(H'') \geq (1/2 + \gamma/2)k - (6\varepsilon/\gamma)k \geq (1/2 + \gamma/3)k$ .  $\square$

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  $n' \geq (1 - \alpha/8)(1 - \varepsilon)n \geq (1 - \alpha/4)n$ .

## 3.2 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  $D \geq 2$  be positive integers. Let  $T$  be a rooted tree with  $t$  vertices and maximum degree at most  $D$ . Then for every  $\beta \geq 1/t$ , 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 descendants of the root of each subtree belong to that subtree.*

The proof of Lemma 15 will make use of the following simple proposition, whose proof is a straightforward modification of the proof of Proposition 4.2 from [1].

**Proposition 16.** *Let  $s$  and  $D \geq 2$  be positive integers. 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; 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 descendants only in its own tree.

*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  $\beta t$ . 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  $\beta t$  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  $\beta t$  vertices. Denote that partition by  $\Pi := \{V_1, \dots, V_r\}$ . Clearly, the number of parts,  $r \leq 1/\beta + 1$ .

*Stage 2.* Let  $T_\Pi$  be the cluster tree corresponding to the partition  $\Pi$  and note that

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

is the total excess degree in the cluster tree. Clearly,

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

Suppose that the partition  $\Pi$  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  $T_\Pi$  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  $(D - 1)^2 + 1 \leq D^2 - 1$

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

*Stage 3.* Root the cluster tree  $T_{\Pi^*}$  at the subset containing the root of the original tree. Order the subsets in  $\Pi^*$  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  $T_{\Pi^*}$ ). Now, iterate the following procedure until the descendence condition is satisfied. If there is a subtree whose root has descendants in other subtrees, merge the first (with respect to this order) such tree with the descendants of its root and remove all the merged pieces from the ordered list. Clearly, after the procedure terminates, the new partition satisfies the descendence condition. As a consequence of our 'parent-first' ordering, each tree can be merged only once and hence both 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 of vertices in the graph  $G''$ , which is an induced subgraph of  $G'$ ,  $n' = |V(G'')| \geq (1 - \alpha/4)n$ . Root our tree  $T$  at an arbitrary vertex. Since  $T$  has fewer than  $n'$  vertices, by Lemma 15 there is a partition  $\Pi$  of  $V(T)$  into  $t$  connected subsets of size at most  $D^2\delta \cdot n'/k'$ , with  $t \leq 4k'/\delta$ , such that the maximum degree of the cluster tree  $T_{\Pi}$  does not exceed  $D^3$ . Let  $S_1, \dots, S_t$  be the sets in that partition, ordered in such a way that all descendants (in the cluster tree) of every tree come later in the ordering. Since each  $S_j$  induces a connected bipartite graph (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  $T[S_j]$ . Let  $\mathcal{S}$  be the collection of all these color classes, i.e.,  $\mathcal{S} := \{S_{j,l} : 1 \leq j \leq t, l \in \{1, 2\}\}$ . Also, let  $\mathcal{E}$  be the set of all pairs  $\{S_{j,l}, S_{j',l'}\}$  such that  $j \neq j'$  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  $T_{\Pi}$ . It follows that  $|\mathcal{E}| = t - 1$ .

### 3.3 Planning out the embedding

Recall that  $M$  is a maximum matching in the cluster graph  $H''$ , and  $V(M) = \{A_1, B_1, \dots, A_{m'}, B_{m'}\}$ . 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  $\Pi$ . 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  $(\varepsilon, p)$ -regular pair in  $G''$  that is represented by this edge.

**Lemma 17.** *There is an assignment  $\varphi : \mathcal{S} \rightarrow V(M)$  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  $\varphi$  assigns to  $X$ . Then the usage of the cluster  $X$ ,

$$U(X) := \sum_{S \in \mathcal{S}(X)} \left( |S| + 4D^3\varepsilon \cdot \frac{n'}{k'} \right) \leq (1 - \alpha/4) \cdot \frac{n'}{k'}.$$

*Proof.* We can easily construct such a map  $\varphi$  using the following greedy procedure. Start with an empty map  $\varphi_0$ . Assume that  $1 \leq j \leq t$  and we have already defined  $\varphi_{j-1}$ . 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  $\mathcal{S}_{j-1}(X) := \varphi_{j-1}^{-1}(X)$  is the family of sets from  $\mathcal{S}$  that  $\varphi_{j-1}$  assigns to  $X$ . Let  $i(j)$  be the smallest  $i$  such that

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

Let  $\varphi_j$  be an extension of  $\varphi_{j-1}$  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  $\varphi = \varphi_t$ .

Note that the way we construct  $\varphi$  guarantees that condition 1 is satisfied. Moreover, since by our choice of  $\Pi$ , 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  $\varphi$  will satisfy condition 2 as well. The only thing we have to check is that for all  $j$ , the index  $i(j)$  is well-defined, i.e., inequality (3) 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  $j$ ,

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

Hence, if for some  $j$ , inequality (3) is not satisfied for all  $i$ , we must have

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

for all  $X \in V(M)$ . 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  $(1 - \alpha) \cdot n \geq |V(T)| = \sum_j |S_j|$ .  $\square$

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

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

1. Let  $e = \{S_{j,l}, S_{j',l'}\} \in \mathcal{E}$ , 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 less than  $|\mathcal{E}|/(2\gamma k'/3) \leq t/(2\gamma k'/3) \leq 6/(\gamma\delta)$  times. We let  $\psi(e)$  be an arbitrary cluster with that property.  $\square$

Now that we have laid out a plan of the embedding, it is time to make room for the tree. We start by choosing in each cluster  $X \in V(H'')$  an arbitrary subset  $C(X)$  of size  $\alpha/4 \cdot n'/k'$ . 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 size 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  $\alpha/(4e_X) \cdot n'/k' \geq \alpha\gamma\delta/24 \cdot n'/k'$ .

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

Finally, note that for all  $j$ ,  $\varphi'(S_{j,1})$  and  $\varphi'(S_{j,2})$  are subsets of opposite clusters in an  $(\varepsilon, p)$ -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  $(\varepsilon, p)$ -regular pair in  $G''$ , whose  $p$ -density is at least  $\gamma/6$ .

### 3.4 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  $\varphi'' : \mathcal{S} \rightarrow \mathcal{P}(V(M))$  and  $\psi'' : \mathcal{E} \rightarrow \mathcal{P}(V(G''))$  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  $e \in \mathcal{E}$ ,  $\psi''(e) \subseteq \psi'(e)$  and  $|\psi''(e)| \geq \varepsilon \cdot n'/k'$ .
4. For all  $e \in \mathcal{E}$  the following holds. Suppose  $e = \{S_{j,l}, S_{j',l'}\}$ , where  $j < j'$ . Then each vertex in  $\varphi''(S_{j,l})$  has a neighbor in  $\psi''(e)$  and each vertex in  $\psi''(e)$  has at least  $D + 1$  neighbors in  $\varphi''(S_{j',3-l'})$ .

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

**Lemma 19.** *Let  $(A, B)$  be an  $(\varepsilon, p)$ -regular pair, whose  $p$ -density is larger than  $\varepsilon$ . 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 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  $(\varepsilon, p)$ -regular with  $p$ -density larger than  $\varepsilon$ , 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  $|N(X') \cap B''| < |B''| - \varepsilon \cdot n'/k'$ . It means that there are no edges in  $(A, B)$  between the sets  $X'$  and  $B'' - N(X')$ . This is impossible, since  $(A, B)$  is  $(\varepsilon, p)$ -regular with  $p$  density larger than  $\varepsilon$ , and both sets are larger than  $\varepsilon \cdot n'/k'$ .  $\square$

**Lemma 20.** *Let  $b \geq 1$  and let  $(A, B)$  be an  $(\varepsilon, p)$ -regular pair, whose  $p$ -density is larger than  $\varepsilon$ . 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  $|X| \leq \varepsilon \cdot n'/k'$ , 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  $(\varepsilon, p)$ -regular with  $p$ -density larger than  $\varepsilon$ ,  $|X'| \geq \varepsilon \cdot n'/k'$  and by the definition of  $X$ ,  $|B' - N(X')| \geq |B'| - (b - 1)|X'| \geq \varepsilon \cdot n'/k'$ .  $\square$

An immediate consequence of Lemma 20 is the following Corollary.

**Corollary 21.** *Let  $d \geq 1$  and let  $(A, B_1), \dots, (A, B_d)$  be (not necessarily distinct)  $(\varepsilon, p)$ -regular pairs in  $G''$ . Suppose that  $|A| = |B_1| = \dots = |B_d| = n'/k'$ ,  $A' \subseteq A$  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  $i \in \{1, \dots, d\}$ . 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  $j$ , all requirements are met as long as they involve only sets  $S_{j',l'}$  with  $j' \geq j$  and  $l' \in \{1, 2\}$ . 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  $\varphi''$  and  $\psi''$  satisfy all requirements involving only sets  $S_{j',l'}$  with  $j' > j$ . Let  $\mathcal{D}(S_{j,1})$  and  $\mathcal{D}(S_{j,2})$  be all the immediate descendants (in the cluster tree  $T_\Pi$ ) of  $S_{j,1}$  and  $S_{j,2}$  respectively. 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  $|S_{j,1}| + 2D^3\varepsilon \cdot n'/k'$  and  $|S_{j,2}| + 2D^3\varepsilon \cdot n'/k'$  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 there is a unique set  $S \in \mathcal{S}$  such that  $e := \{S, S_{j,l}\}$  belongs to  $\mathcal{E}$ . Since  $|\psi'(e)| \geq \alpha\gamma\delta/24 \cdot n'/k' \geq 2\varepsilon \cdot n'/k'$ , 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  $\psi''(e) := A''$ .

### 3.5 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  $T_\Pi$ .

We start by embedding the first subtree,  $S_1$ , into the graph  $H_1 := (\varphi''(S_{1,1}), \varphi''(S_{1,2}))$ . 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  $S_1, \dots, S_{j-1}$  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  $\varphi''$  and  $\psi''$  guarantees that the image of  $p_j$ , which by the definition of  $\Pi$  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  $H_j := (\varphi''(S_{j,l}) \cup \{x\}, \varphi''(S_{j,3-l}))$  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.

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.  $\square$

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  $\varphi''(S_{j,3-l})$ . 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, Marcinişyn and Steger [7].

**Theorem 23.** *Let  $\alpha$  and  $\gamma$  be positive constants. There exist  $\eta_0 = \eta_0(\alpha, \gamma)$  and  $n_0 = n_0(\alpha, \gamma)$  with the following property. Let  $G$  be an  $n$ -vertex  $(p, \eta)$ -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 just follows 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.

## References

- [1] N. Alon, M. Krivelevich, and B. Sudakov, *Embedding nearly-spanning bounded degree trees*, *Combinatorica* **27** (2007), 629–644.
- [2] N. Alon and J. Spencer, *The probabilistic method. Third edition*, John Wiley & Sons, Inc., 2008.

- [3] J. Balogh, B. Csaba, M. Pei, and W. Samotij, *Large bounded degree trees in expanding graphs*, submitted.
- [4] B. Bollobás, *Extremal graph theory*, Academic Press, Inc., 1978.
- [5] J. Böttcher, Y. Kohayakawa, and A. Taraz, *Almost spanning subgraphs of random graphs after adversarial edge removal*, in preparation.
- [6] B. Csaba, I. Levitt, J. Nagy-György, and E. Szemerédi, *On embedding trees*, submitted.
- [7] 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.
- [8] G. A. Dirac, *Some theorems on abstract graphs*, *Proceedings of the London Mathematical Society* (3) **2** (1952), 69–81.
- [9] J. Friedman and N. Pippenger, *Expanding graphs contain all small trees*, *Combinatorica* **7** (1987), 71–76.
- [10] A. Frieze and M. Krivelevich, *On two Hamilton cycle problems in random graphs*, *Israel Journal of Mathematics* **166** (2008), 221–234.
- [11] P. Haxell, *Tree embeddings*, *Journal of Graph Theory* **36** (2001), 121–130.
- [12] Y. Kohayakawa, *Szemerédi’s Regularity Lemma for sparse graphs*, *Foundations of computational mathematics* (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.
- [13] J. Komlós, G. Sárközy, and E. Szemerédi, *Proof of a packing conjecture of Bollobás*, *Combinatorics, Probability and Computing* **4** (1995), 241–255.
- [14] M. Krivelevich, C. Lee, and B. Sudakov, *Resilient pancyclicity of random and pseudo-random graphs*, submitted.
- [15] L. Pósa, *Hamiltonian circuits in random graphs*, *Discrete Mathematics* **14** (1976), 359–364.
- [16] B. Sudakov and V. Vu, *Local resilience of graphs*, *Random Structures and Algorithms* **33** (2008), 409–433.
- [17] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, *Mat. Fiz. Lapok* **48** (1941), 436–452.