UNIVERSALITY FOR BOUNDED DEGREE SPANNING TREES IN RANDOMLY PERTURBED GRAPHS

JULIA BÖTTCHER, JIE HAN, YOSHIHARU KOHAYAKAWA, RICHARD MONTGOMERY, OLAF PARCZYK, AND YURY PERSON

Abstract. We solve a problem of Krivelevich, Kwan and Sudakov [16] concerning the threshold for the containment of all bounded degree spanning trees in the model of randomly perturbed dense graphs. More precisely, we show that, if we start with a dense graph G_α on n vertices with $\delta(G_\alpha) \geqslant \alpha n$ for $\alpha > 0$ and we add to it the binomial random graph $G(n, C/n)$, then with high probability the graph $G_\alpha \cup G(n, C/n)$ contains copies of all spanning trees with maximum degree at most Δ simultaneously, where C depends only on α and Δ.

§1. Introduction

Many problems from extremal graph theory concern Dirac-type questions. These ask for asymptotically optimal conditions on the minimum degree $\delta(G_n)$ for an n-vertex graph G_n to contain a given spanning graph F_n. Typically there exists a constant $\alpha > 0$ (depending on F_n) so that $\delta(G_n) \geqslant \alpha n$ implies $F_n \subseteq G_n$. A prime example is Dirac's theorem [10] stating that $\delta(G_n) \geqslant n/2$ ensures that G_n is Hamiltonian if $n \geqslant 3$.

On the other hand, a large branch of the theory of random graphs studies when random graphs typically contain a copy of a given spanning structure F_n. Let $G(n, p)$ be the n-vertex binomial random graph, where each of the $\binom{n}{2}$ possible edges is present independently at random with probability $p = p(n)$. A classical result of Bollobás and Thomason [8] states that every nontrivial monotone property has a threshold in $G(n, p)$. Since containing a copy of (a sequence of graphs) F_n is a monotone property, there exists a *threshold function* $\hat{p} = \hat{p}(n) \colon \mathbb{N} \to [0, 1]$ such that, if $p = o(\hat{p})$, then $\lim_{n\to\infty} \mathbb{P}[F_n \subseteq G(n, p)] = 0$, whereas, if $p = \omega(\hat{p})$, then $\lim_{n\to\infty} \mathbb{P}[F_n \subseteq G(n, p)] = 1$. When the conclusion of the latter case holds, we say that $G(n, p)$ contains F_n *asymptotically almost surely* (a.a.s.). For example, a famous result of Koršunov [14] and Pósa [21] asserts that the threshold for Hamiltonicity in $G(n, p)$ is $(\log n)/n$.

Bohman, Frieze and Martin discovered the following phenomenon in [7]. Given a fixed $\alpha > 0$, they started with a graph G_α on n vertices with $\delta(G_\alpha) \geqslant \alpha n$. Here, α can be arbitrarily small and hence G_α can be far from containing any Hamilton cycle. They proved that, after adding $m = C(\alpha)n$ edges uniformly at random to G_α, the new graph G becomes Hamiltonian a.a.s., where $C(\alpha)$ is a constant that depends only on α. Letting G_α be the complete unbalanced bipartite

Date: 2018/04/01, 8:47am

JB is partially supported by EPSRC (EP/R00532X/1). JH is supported by FAPESP (2014/18641-5, 2013/03447-6). YK is partially supported by FAPESP (2013/03447-6) and CNPq (310974/2013-5, 459335/2014-6). OP and YP are supported by DFG grant PE 2299/1-1. The cooperation of the authors was supported by a joint CAPES-DAAD PROBRAL project (Proj. no. 430/15, 57143515).

graph $K_{\alpha n, (1-\alpha)n}$, one sees that the addition of linearly many edges to G_α is necessary for this result to hold in general. Furthermore, clearly, the conditions on $\delta(G_\alpha)$ and on $p = m/\binom{n}{2}$ in this result are weaker than in the corresponding Dirac-type problem and the threshold problem, respectively. More precisely, the probability p turns out to be smaller by a factor of $\Theta(\log n)$. Here, we have switched from choosing m edges uniformly at random to the the binomial $G(n,p)$ model, which is known to be essentially equivalent when $p = m/\binom{n}{2}$ (see, e.g., [13]).

The model $G_\alpha \cup G(n,p)$ is known as the *randomly perturbed graph model*. Typically $p = o(1)$, so an 'addition' of $G(n,p)$ to the dense graph G_α corresponds to a small random perturbation in the structure of G_α. This model and its related generalizations to hypergraphs and digraphs sparked a great deal of research in recent years.

In this paper we are concerned with spanning trees in randomly perturbed graphs. For *almost spanning* trees it was shown by Alon, Krivelevich and Sudakov [2] that, for some constant $C = C(\varepsilon, \Delta)$, the random graph $G(n, C/n)$ alone a.a.s. contains any tree with at most $(1-\varepsilon)n$ vertices and maximum degree at most Δ, where the bounds on $C = C(\varepsilon, \Delta)$ have subsequently been improved [3]. Since the random graph $G(n, C/n)$ a.a.s. contains isolated vertices, it obviously does not contain spanning trees. The problem of determining the threshold of bounded degree spanning trees attracted much attention. Recently, Montgomery [20] showed that for each constant Δ and every sequence of trees T_n with maximum degree Δ, the threshold in $G(n,p)$ for a copy of T_n to appear is $(\log n)/n$ (see also [19]). However, Krivelevich, Kwan and Sudakov [16] showed that, again, a smaller probability suffices in the randomly perturbed graph model. They proved that $G_\alpha \cup G(n,p)$ a.a.s. contains a given spanning tree T_n with maximum degree at most Δ when $p = C(\Delta, \alpha)/n$.

In the concluding remarks of [16], Krivelevich, Kwan and Sudakov raised the question of whether $G_\alpha \cup G(n, D/n)$ contains all spanning trees of maximum degree at most Δ simultaneously, for some constant $D = D(\Delta, \alpha)$. The purpose of this paper is to answer their question in the affirmative. For stating our result we need some notation. For a family $\mathcal{F}$ of graphs, we say that a graph G is *$\mathcal{F}$-universal* if G contains a copy of every graph F from $\mathcal{F}$. We denote by $\mathcal{T}(n, \Delta)$ the family of all trees of maximum degree at most Δ on n vertices.

Theorem 1. *For each $\alpha > 0$ and $\Delta \in \mathbb{N}$, there exists a constant $D = D(\Delta, \alpha)$ such that the following holds. If G_α is an n-vertex graph with $\delta(G_\alpha) \geqslant \alpha n$, then the randomly perturbed graph $G_\alpha \cup G(n, D/n)$ is a.a.s. $\mathcal{T}(n, \Delta)$-universal.*

This theorem is an immediate consequence of a technical theorem, Theorem 2, which states that the union of G_α with any reasonably expanding graph G is $\mathcal{T}(n, \Delta)$-universal. The proof of Theorem 2 relies on the use of reservoir sets resembling those introduced in [9] as part of the so-called assisted absorption method. The novelty in our proof is that we construct these reservoir sets using expanding graphs rather than random graphs, which is not possible with the techniques from [9] (see also the discussion in Section 2).

Before we turn to the details of our embedding technique, we mention further results concerning randomly perturbed graphs. Further spanning structures whose appearance in randomly perturbed graphs has been studied are F-factors (for fixed graphs F) [4], squares of Hamilton cycles and copies of general bounded degree spanning graphs [9], perfect matchings and loose Hamilton cycles in uniform hypergraphs [15], and tight Hamilton cycles in hypergraphs [11]. Most of the mentioned results exhibit the following phenomenon: in the presence of a dense graph G_α, a smaller edge probability than in $G(n,p)$ alone suffices. The only exception to this rule so far are F-factors for certain non-strictly-balanced graphs F covered in [4]. Moreover, some variations of such results when α is at least some positive constant c (which depends on other parameters of the problems at hand) were considered in [5, 6, 18].

§2. Notation, main technical result, and proof overview

We will use standard graph theoretic notation throughout. In the following, we briefly recap most of the relevant terminology. Given graphs G and H, write $|G| = |V(G)|$ and $G \setminus H = G[V(G) \setminus V(H)]$, that is, the induced subgraph of G on $V(G) \setminus V(H)$. Throughout this note we omit floors and ceilings. For two not necessarily disjoint sets U and W of vertices of a graph G we write $e(U,W)$ for the number of edges with one endpoint in U and the other in W, where we count edges that lie in $U \cap W$ twice.

We say that an n-vertex graph G is an (n,p,ε,C)*-graph* if $\Delta(G) \leqslant Cpn$ and, for any U, $W \subseteq V(G)$ such that $|U|, |W| \geqslant \varepsilon n$, we have $e(U,W) \geqslant (p/C)|U||W|$. We further denote the family of (n,p,ε,C)-graphs by $\mathcal{G}(n,p,\varepsilon,C)$. Intuitively, the graphs from $\mathcal{G}(n,p,\varepsilon,C)$ are graphs with a certain degree bound which are expanding for vertex subsets of linear size.

Our main technical result states that perturbing graphs G_α with minimum degree at least αn by graphs $G \in \mathcal{G}(n, D/n, \varepsilon, C)$ results in $\mathcal{T}(n,\Delta)$-universal graphs.

Theorem 2 (Main technical result)**.** *For any $\alpha > 0$ and integers $C \geqslant 2$ and $\Delta \geqslant 1$, there exist $\varepsilon > 0$, D_0 and n_0 such that the following holds for any $D \geqslant D_0$ and $n \geqslant n_0$. Suppose $G \in \mathcal{G}(n, D/n, \varepsilon, C)$ and G_α are n-vertex graphs on the same vertex set and $\delta(G_\alpha) \geqslant \alpha n$. Then $H := G_\alpha \cup G$ is $\mathcal{T}(n,\Delta)$-universal.*

We will show in Section 5 that this result implies Theorem 1. In the remainder of this section, we give a brief outline of our proof of Theorem 2.

2.1. **Proof overview.** Let $G \in \mathcal{G}(n, D/n, \varepsilon, C)$. We embed an arbitrary $T \in \mathcal{T}(n,\Delta)$ into $H := G_\alpha \cup G$ in three phases. In the first phase, we find a subtree T_1 of T (see Lemma 3) of small linear size, say βn with $\beta \ll \alpha$, and we embed this subtree T_1 into H using a randomized algorithm (see Lemma 5). In doing so, we can show that there is some such embedding in which, for any given pair of vertices u, $v \in V(H)$, there are at least $3\Delta\varepsilon n$ vertices $w \in V(T_1)$ with $N_T(w) \subseteq V(T_1)$ so that w is embedded into $N_H(u)$ and $N_T(w)$ is embedded into $N_H(v)$ – a fact which will turn out to be crucial later. We denote by $B_{T,H}(u,v)$ such a set of vertices w, and refer to such sets $B_{T,H}(u,v)$ as *reservoir sets* (see Section 3.2 for the formal definition).

In the second phase, we extend the tree T_1 to an almost spanning subtree T' of T with $|T \setminus T'| = 2\varepsilon n$. For this purpose we use a theorem of Haxell [12] (Theorem 7), which ensures such almost spanning embeddings given sufficient expansion in the host graph H.

Finally, in the third phase, we complete our embedding using a greedy approach and the reservoir sets $B_{T,H}(u,v)$ for the following *swapping* trick: since T' is a subtree of T, we can extend it by consecutively appending degree-1 vertices and thus growing the tree T' into T. Suppose $T' = T'_0 \subseteq \cdots \subseteq T'_{2\varepsilon n} = T$ is the sequence of subtrees of T that we encounter in this process. Suppose we already have the embedding $g_{i-1}\colon V(T'_{i-1}) \to V(H)$, and we wish to extend it to $g_i\colon V(T'_i) \to V(H)$ by defining the image of the leaf $b \in V(T'_i) \smallsetminus V(T'_{i-1})$. Given some vertex v of H available for embedding b (that is, $v \notin g_{i-1}(V(T'_{i-1})))$, if there is an edge in H from v to $g_{i-1}(u)$, where u is the parent of b in T_i, then we simply embed b onto v (that is, we let $g_i(b) = v$). On the other hand, if there is no edge in H from v to $g_{i-1}(u)$, we proceed as follows. We will set things up so that, by counting, we will be able to show that there is some $c \in V(T_{i-1})$ such that $g_{i-1}(c) \in B_{T,H}(g_{i-1}(u), v)$. We then let $g_i(b) = g_{i-1}(c)$ and we let $g_i(c) = v$. This defines a valid embedding $g_i\colon V(T_i) \to V(H)$. (We remark that we said that we would *extend* g_{i-1} to g_i; as it will be clear by now, this is not strictly speaking correct, as we may alter g_{i-1} slightly before extending it to g_i.)

As mentioned earlier, the reservoir sets used in our proof are similar to those introduced in the setting of randomly perturbed graphs in [9]. In that work, the reservoir sets are used to prove a general result about spanning structures in randomly perturbed graphs, which can be easily applied to consider the appearance of various different single spanning structures. In particular, this gives a short proof of the appearance of any single bounded degree spanning tree in this model, a problem that was first solved in [16]. The argument from [9] does not work for universality statements. However, here we show that the reservoirs can be found and the swapping trick employed in the completely deterministic setting by embedding the first part of the tree in a randomized way.

§3. Auxiliary Lemmas

The lemmas provided in this section will be used in the proof of Theorem 2. We start in Section 3.1 with two lemmas for partitioning the tree T we want to embed. We then explain how we obtain good reservoir sets by embedding a subtree T_1 of T randomly in Section 3.2. Finally, in Section 3.3 we provide the tools to extend this embedding to an almost spanning subgraph of T.

3.1. Tree partitioning lemmas.

Recall that $\mathcal{T}(n, \Delta)$ is the collection of all trees on n vertices with maximum degree at most Δ, and that a graph G on n vertices is said to be $\mathcal{T}(n, \Delta)$-universal if G contains a copy of T for every $T \in \mathcal{T}(n, \Delta)$.

The main assertion of the following lemma is that we can find in any bounded degree tree T a subtree T_1 of roughly any desired size so that removing T_1 from T leaves a tree. We will use this lemma to find a small linear sized subtree T_1, which we embed in our first phase.

Lemma 3. *Let β, $\varepsilon > 0$ and let n, Δ be positive integers such that $\Delta\beta + 2\varepsilon < 1$ and $\beta < 1/\Delta$. Then, for any $T \in \mathcal{T}(n, \Delta)$, there exist subtrees $T_1 \subseteq T' \subseteq T$ such that*

(*a*) $\beta n \leqslant |T_1| \leqslant \Delta\beta n$,
(*b*) $e(T_1, T \smallsetminus T_1) = 1$, *and*
(*c*) $|T \smallsetminus T'| = 2\varepsilon n$.

Proof. Fix any vertex v of T as the root and, for each $w \in V(T)$, write C_w for the branch (subtree) of T consisting of w and all of its descendants. By $\beta < 1/\Delta$, $\deg(v) \leqslant \Delta$ and averaging, there is $v_1 \in N_T(v)$ such that $|C_{v_1}| \geqslant \beta n$. If $|C_{v_1}| > \Delta\beta n$, then similarly there is a child v_2 of v_1 such that $|C_{v_2}| \geqslant (\Delta\beta n - 1)/(\Delta - 1) \geqslant \beta n$. Repeating this argument gives a desired v' such that $\beta n \leqslant |C_{v'}| \leqslant \Delta\beta n$. Let $T_1 := C_{v'}$. Note that (*b*) holds by the definition of T_1. Finally, let T' be an (arbitrary) subtree of T such that $T_1 \subseteq T' \subseteq T$ and $|T \smallsetminus T'| = 2\varepsilon n$. □

Let T be a tree. Given vertices $x_1, \ldots, x_m$ of T, let $\langle x_1, \ldots, x_m \rangle_T$ be the minimal subtree of T that contains the vertices $x_1, \ldots, x_m$, which is just the subtree of T obtained from the union of the vertex sets of all the paths between x_i, x_j, $i \neq j$, in T. For two distinct vertices x, y of T, we write $\mathrm{dist}_T(x, y)$ for their distance in T, namely, the length of the (unique) path on T connecting x and y. Given a vertex x of T and a vertex set $Y \subseteq V(T)$ such that $x \notin Y$, let $\mathrm{dist}_T(x, Y) := \min_{y \in Y} \mathrm{dist}_T(x, y)$.

The following lemma provides us with vertices $x_1, \ldots, x_s$ in a tree T which cover T well, but are not too close. In particular, this gives us a collection of stars $x_i \cup N_T(x_i)$ which are far enough apart that they are relatively independent.

Lemma 4. *For any tree T with maximum degree at most Δ, there exist $s \in \mathbb{N}$ and vertices $x_1, \ldots, x_s \in V(T)$ such that*

(*a*) *for any $2 \leqslant i \leqslant s$, $\mathrm{dist}_T(x_i, \langle x_1, \ldots, x_{i-1} \rangle_T) = 5$,*
(*b*) *$|T|/(5\Delta^4) \leqslant s \leqslant (|T| + 4)/5$, and*
(*c*) *$\mathrm{dist}_T(x, \langle x_1, \ldots, x_s \rangle_T) \leqslant 4$ for all vertices $x \in V(T)$.*

Proof. We start with picking x_1 arbitrarily. We greedily pick the vertices $x_2, \ldots, x_s$ in $V(T)$ sequentially as long as there is a vertex x_i such that $\mathrm{dist}_T(x_i, \langle x_1, \ldots, x_{i-1} \rangle_T) = 5$. Note that for any $2 \leqslant i \leqslant s$, $|\langle x_1, \ldots, x_i \rangle_T \smallsetminus \langle x_1, \ldots, x_{i-1} \rangle_T| = 5$, so we inductively get that $|\langle x_1, \ldots, x_s \rangle_T| = 5s - 4$. This implies $s \leqslant (|T| + 4)/5$. Since T is connected, the maximality of s implies that $\mathrm{dist}_T(x, \langle x_1, \ldots, x_s \rangle_T) \leqslant 4$ for all vertices $x \in V(T)$. Thus we have $|T| \leqslant (5s - 4)\Delta^4$, which implies (*b*). □

3.2. **A randomized embedding – controlling reservoir sets.** In the following, we define formally the reservoir sets $B_{T,H}(u, v)$, already mentioned in the proof overview given in Section 2.1, and show that we can force them to be suitably large. These reservoir sets will be helpful when finishing the embedding of T, since they will allow us to alter locally partial embeddings that we construct sequentially. We warn the reader that, for technical convenience, the sets $B_{T,H}(u, v)$ are defined here in a slightly different manner in comparison with the informal definition given earlier in Section 2.1. Let V be a set of n vertices. Let G be a graph on V and let T be a tree with $V(T) \subseteq V$ (T is not necessarily a subgraph of G). For $v \in V$, let

$$B_{T,G}(v) := \left\{ w \in V(T) : N_T(w) \subseteq N_G(v) \right\}.$$

For distinct vertices u and $v \in V$, we define their *reservoir set* $B_{T,G}(u, v)$ as follows:

$$B_{T,G}(u, v) := B_{T,G}(v) \cap N_G(u).$$

Recall that the idea is that we can free up any $w \in B_{T,G}(u,v)$ used already in the embedding, by moving the vertex embedded to w to v. This then allows us to use w for embedding any unembedded neighbour of the vertex embedded to u.

Our next lemma shows that we can embed the linear sized subtree T_1 of T into $H = G \cup G_\alpha$ using a randomized algorithm, such that we get large reservoir sets.

Lemma 5. *For any $\alpha > 0$ and integers $C \geqslant 2$ and $\Delta \geqslant 1$, there exist $\varepsilon > 0$, D_0 and n_0, such that the following holds for $D \geqslant D_0$ and $n \geqslant n_0$. Suppose $G \in \mathcal{G}(n, D/n, \varepsilon, C)$ and G_α is an n-vertex graph such that $\delta(G_\alpha) \geqslant \alpha n$ and $V(G) = V(G_\alpha)$. Then, for any tree T_1 such that $\Delta(T_1) \leqslant \Delta$ and $\alpha n/(2\Delta^2) \leqslant |T_1| \leqslant \alpha n/(2\Delta)$, there is an embedding g of T_1 into $H := G \cup G_\alpha$ such that $|B_{\widetilde{T_1},H}(u,v)| \geqslant 2(\Delta+3)\varepsilon n$ for any u and $v \in V$, where $\widetilde{T_1} = g(T_1)$.*

In the proof of this lemma we use the following concentration result.

Lemma 6 (Sequential dependence lemma, Lemma 2.2 from [1])**.** *Let Ω be a finite probability space and $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_n$ be each a partition of Ω. For each $i \in [n]$ let Y_i be a Bernoulli random variable on Ω which is constant on each part of $\mathcal{F}_i$ ($\mathcal{F}_i$-measurable), and let p_i be a real-valued random variable on Ω which is constant on each part of $\mathcal{F}_{i-1}$. Let x be a real number, $\delta \in (0, 3/2)$, and $X = Y_1 + \cdots + Y_n$. If $\sum_{i=1}^n p_i \geqslant x$ holds almost surely, and $\mathbb{E}[Y_i \mid \mathcal{F}_{i-1}] \geqslant p_i$ holds almost surely for all $i \in [n]$, then*

$$\Pr\big(X < (1-\delta)x\big) < e^{-\delta^2 x/3}.$$

Proof of Lemma 5. First we choose the parameters D_0 and ε as follows:

$$D_0 := 2C\Delta/\alpha \quad \text{and} \quad \varepsilon := \alpha^{\Delta+2} C^{-2\Delta} 2^{-\Delta-8} \Delta^{-7}, \tag{1}$$

and then we choose n_0 large enough.

We apply Lemma 4 to T_1 and obtain $s \in \mathbb{N}$ and vertices $x_1, \ldots, x_s \in V(T_1)$ such that, for any $2 \leqslant i \leqslant s$, $\mathrm{dist}_{T_1}(x_i, \langle x_1, \ldots, x_{i-1} \rangle_{T_1}) = 5$, and $|T_1|/(5\Delta^4) \leqslant s \leqslant (|T_1| + 4)/5$. Our embedding of T_1 consists of three steps. First we embed the disjoint stars with centers at $x_1, \ldots, x_s$ uniformly at random into stars in H (using only the edges of G) whose vertices have not yet been used as images. Next we connect these stars and obtain an embedding of a subtree of T_1 as the union of the stars and $\langle x_1, \ldots, x_s \rangle_{T_1}$. At last we embed the rest of the vertices of T_1 greedily, which will be possible using G_α as $|T_1| \leqslant \alpha n/(2\Delta)$ and $\delta(G_\alpha) \geqslant \alpha n$.

The following claim states that we can pick disjoint stars with Δ leaves (that is, copies of $K_{1,\Delta}$) in G, within which we will later embed the stars in T_1 with centers at $x_1, \ldots, x_s$.

Claim. *There is a choice of disjoint stars $S_1, \ldots, S_s$ with Δ leaves in G such that, for each $u, v \in V$ there are at least $2(\Delta+3)\varepsilon n$ stars among $S_1, \ldots, S_s$ with their centers in $N_{G_\alpha}(u)$ and their leaves in $N_{G_\alpha}(v)$.*

Proof of the Claim. We randomly and sequentially pick s stars $S_1, \ldots, S_s$ with Δ leaves from G, where each star S_i is picked uniformly at random from the copies of $K_{1,\Delta}$ which are disjoint from $S_1, \ldots, S_{i-1}$ (we show below that this is indeed possible).

For $u, v \in V$, $i \in [s]$, let $Y_i^{u,v}$ be the Bernoulli random variable for the event that $\widetilde{x}_i \in N_{G_\alpha}(u)$ and $R_i \subseteq N_{G_\alpha}(v)$, where $\widetilde{x}_i$ is the center of S_i and R_i is the set of leaves of S_i. Since $\delta(G_\alpha) \geqslant \alpha n$, $|T_1| \leqslant \alpha n/(2\Delta)$ and the existing stars cover at most

$$(\Delta+1)s \leqslant (\Delta+1)\left(\frac{|T_1|}{5}+4\right) \leqslant (\Delta+1)\left(\frac{\alpha n}{10\Delta}+4\right) \leqslant \alpha n/4$$

vertices, there are at least $3\alpha n/4$ vertices available in both $U := N_{G_\alpha}(u) \setminus \bigcup_{j\in[i-1]} V(S_j)$ and $W := N_{G_\alpha}(v) \setminus \bigcup_{j\in[i-1]} V(S_j)$.

Since $G \in \mathcal{G}(n, D/n, \varepsilon, C)$ and $3\alpha/4 \geqslant \varepsilon$, $e(U, W) \geqslant D|U||W|/(Cn) \geqslant 3\alpha D|U|/(4C)$. By the convexity of the binomial function, the number of $K_{1,\Delta}$-stars with center in U and leaves in W is at least

$$\sum_{u\in U}\binom{\deg_W(u)}{\Delta} \geqslant |U|\binom{\sum_{u\in U}\deg_W(u)/|U|}{\Delta} \geqslant |U|\binom{3\alpha D/(4C)}{\Delta}.$$

Since $\Delta(G) \leqslant CD$, the total number of $K_{1,\Delta}$-stars in G is at most $n\binom{CD}{\Delta}$. This allows us to obtain the following lower bound on $\mathbb{E}(Y_i^{u,v} \mid Y_1^{u,v}, \ldots, Y_{i-1}^{u,v})$:

$$\mathbb{E}(Y_i^{u,v} \mid Y_1^{u,v}, \ldots, Y_{i-1}^{u,v}) \geqslant \frac{|U|}{n}\frac{\binom{3\alpha D/(4C)}{\Delta}}{\binom{CD}{\Delta}} \geqslant 2^{-\Delta-1}\alpha^{\Delta+1}C^{-2\Delta}.$$

Let $p_i := 2^{-\Delta-1}\alpha^{\Delta+1}C^{-2\Delta}$ and

$$x := sp_i \geqslant \frac{|T_1|}{5\Delta^4}p_i \geqslant \frac{\alpha n}{2\Delta^2\cdot 5\Delta^4}\cdot\frac{\alpha^{\Delta+1}}{2^{\Delta+1}C^{2\Delta}} \geqslant 4(\Delta+3)\varepsilon n,$$

by the choice of ε in (1). Thus, by Lemma 6 with $\delta = 1/2$, we get

$$\mathbb{P}\big(Y_1^{u,v} + \cdots + Y_s^{u,v} < 2(\Delta+3)\varepsilon n\big) \leqslant \mathbb{P}\big(Y_1^{u,v} + \cdots + Y_s^{u,v} < x/2\big) < e^{-x/12} \leqslant e^{-\varepsilon n}.$$

Thus by the union bound, we conclude that there is a choice of $S_1, \ldots, S_s$ such that, for each $u, v \in V$, $Y_1^{u,v} + \cdots + Y_s^{u,v} \geqslant 2(\Delta+3)\varepsilon n$, i.e., the claim holds. □

Now let $S_1, \ldots, S_s$ be as given by the claim. Define the embedding g of the stars in T_1 on vertices $\{x_1\} \cup N_{T_1}(x_1) \cup \cdots \cup \{x_s\} \cup N_{T_1}(x_s)$ by mapping the star (which does not necessarily have Δ leaves) on vertices $\{x_i\} \cup N_{T_1}(x_i)$ to an arbitrary subset of S_i, with x_i mapped to the center $\widetilde{x}_i$. This gives us an embedding of the (partial) forest of stars $T[\{x_i\} \cup N_{T_1}(x_i) \cup \cdots \cup \{x_s\} \cup N_{T_1}(x_s)]$.

Next we extend our partial forest by connecting these stars according to the order $x_1, \ldots, x_s$, and obtain an embedding of a subtree of T_1 which is the union of the stars and $\langle x_1, \ldots, x_s\rangle_{T_1}$. Suppose we have connected the first $i-1$ stars, i.e., we have an embedding of $\langle x_1, \ldots, x_{i-1}\rangle_{T_1}$, and now we will connect it to $\widetilde{x}_i$, the image of x_i. Recall that $\mathrm{dist}_{T_1}(x_i, \langle x_1, \ldots, x_{i-1}\rangle_{T_1}) = 5$ and thus let the path to be embedded be $x_i, y_1, y_2, y_3, y_4, z$. Note that x_i, z, y_1 are already embedded in $H = G \cup G_\alpha$. Moreover, if $z \in \{x_1, \ldots, x_{i-1}\}$, then y_4 has already been embedded; otherwise, fix a neighbor of $g(z)$ in G_α which is not covered by the current partial forest as $g(y_4)$. This is possible because $\delta(G_\alpha) \geqslant \alpha n$ and $|T_1| \leqslant \alpha n/(2\Delta)$. Note that, using G_α, there are at least $\alpha n/2$ choices for the image of y_2 and at least $\alpha n/2$ choices for the image of y_3, so, as $G \in \mathcal{G}(n, D/n, \varepsilon, C)$, we can pick $\widetilde{y}_2$ and $\widetilde{y}_3$ so that $\widetilde{y}_2\widetilde{y}_3$ is an edge of G. Thus, the sequence $\widetilde{x}_i, g(y_1), \widetilde{y}_2, \widetilde{y}_3, g(y_4), g(z)$ forms a path in H. Define $g(y_i) = \widetilde{y}_i$ for $i = 2, 3$. When finished, this completes the second step of the embedding.

For the last step, note that since the partial tree that has been embedded is connected, we can finish the embedding of T_1 by iteratively attaching leaves to the partial embedding. This is always possible because $\delta(G_\alpha) \geqslant \alpha n$ and $|T_1| \leqslant \alpha n/(2\Delta)$. Let g be the resulting embedding function and $\widetilde{T_1} = g(T_1)$.

By the claim for any $u, v \in V$, there are at least $2(\Delta+3)\varepsilon n$ stars from $S_1, \ldots, S_s$ such that their centers are in $N_{G_\alpha}(u)$ and their leaves are in $N_{G_\alpha}(v)$. Since these stars are subtrees of $\widetilde{T_1}$, we conclude that $|B_{\widetilde{T_1},H}(u,v)| \geqslant 2(\Delta+3)\varepsilon n$ for any $u, v \in V$, as required. □

3.3. **Almost spanning tree embeddings.** To extend T_1 to the almost spanning tree T', we will use the following tree embedding result of Haxell [12]. For a graph G and vertex set $X \subseteq V(G)$, we let $N_G(X) := \bigcup_{x \in X} N_G(x)$.

Theorem 7. *Let T be a tree with t edges and maximum degree d. Let $\varnothing = T_0 \subseteq T_1 \subseteq T_2 \subseteq \cdots \subseteq T_\ell \subseteq T$ be a sequence of subtrees of T such that T can be obtained by attaching new leaves to T_l. Let $d = d_1 \geqslant \cdots \geqslant d_\ell \geqslant 1$ be a sequence of integers such that for each $i \in [\ell]$ and each $v \in V(T)$ we have $\deg_T(v) - \deg_{T_{i-1}}(v) \leqslant d_i$. Let $t_i = |E(T_i)|$. Suppose $k \geqslant 1$ is an integer and G is a graph satisfying the following $\ell + 2$ conditions:*

- *(0) $|N_G(X)| \geqslant d|X| + 1$ for every $X \subseteq V(G)$ with $1 \leqslant |X| \leqslant 2k$,*
- *(i) $|N_G(X)| \geqslant d_i|X| + t_i + 1$ for every $X \subseteq V(G)$ with $k < |X| \leqslant 2k$, and for each $i \in [\ell]$,*
- *(ℓ+1) $|N_G(X)| \geqslant t + 1$ for every $X \subseteq V(G)$ with $|X| = 2k + 1$.*

Then G contains T as a subgraph. Moreover, for any vertex x_0 of T_1 and any $y \in V(G)$, there exists an embedding f of T into G such that $f(x_0) = y$.

Let us formulate the consequence of this embedding result we will use in our proof, which easily follows by taking $\ell = 1$ and each $d_i = \Delta$.

Corollary 8. *Let T be a tree with t edges and maximum degree at most Δ. Suppose $k \geqslant 1$ is an integer and G is a graph satisfying the following two conditions:*

- *(i) $|N_G(X)| \geqslant \Delta|X| + 1$ for every $X \subseteq V(G)$ with $1 \leqslant |X| \leqslant 2k$,*
- *(ii) $|N_G(X)| \geqslant \Delta|X| + t + 1$ for every $X \subseteq V(G)$ with $k < |X| \leqslant 2k + 1$.*

Then G contains T as a subgraph. Moreover, for any vertex x_0 of T and any $y \in V(G)$, there exists an embedding f of T into G such that $f(x_0) = y$.

§4. Main technical result

In this section we prove our main technical result, Theorem 2. Given $T \in \mathcal{T}(n, \Delta)$ we will use Lemma 3 to obtain a subtree T_1 of T of small linear size, which we embed with the help of Lemma 5 and then extend to the embedding of an almost spanning subtree of T using Corollary 8. We then use the reservoir sets $B_{T,H}(u,v)$ to extend the embedding to cover the last few vertices.

The sets $B_{T,H}(u,v)$ will help us to embed these last few vertices in the following way. Suppose we have a partial embedding $g : T' \to H$ of our tree T into the host graph H, such that $T' \subseteq T$ is connected and let $\widetilde{T}' = g(T')$. Since T' is a subtree in T we can extend it vertex by vertex by connecting T' with some new vertex $b \in V(T \setminus T')$, which has one neighbour in $V(T')$. Assume that this neighbour a of b in T has been embedded to u, but none of the unused vertices is connected to u in H so that we cannot simply embed b to one of the unused vertices. Instead, if there exists an unused vertex v such that $B_{\widetilde{T'},H}(u,v) \neq \varnothing$, then we can proceed with the embedding as follows. Let $w \in B_{\widetilde{T'},H}(u,v)$ and note that, by the definition of $B_{\widetilde{T'},H}(u,v)$, we have $w \in V(\widetilde{T}')$. Let $c = g^{-1}(w)$, and let $g'(x) = g(x)$, for any $x \in V(T') \smallsetminus \{c\}$, $g'(c) = v$ and $g'(b) = w$. Using the definition of $B_{\widetilde{T'},H}(u,v)$, this gives a partial embedding g' into H with one more leaf, b, embedded. We will show that we only need this procedure to embed the last $2\varepsilon n$ vertices of T, and, for any $u,v \in V$, by the property guaranteed by Lemma 5, the reservoir sets $|B_{\widetilde{T'},H}(u,v)|$ will be large enough to proceed greedily.

Proof. Given α, C and Δ, set $\varepsilon' = \alpha^{\Delta+2}C^{-2\Delta}2^{-\Delta-8}\Delta^{-7}$, a constant small enough that by taking D_0 and n_0 to be large we can use the conclusion of Lemma 5 with $\varepsilon = \varepsilon'$ (cf. (1)). Set $\varepsilon := \min\{\alpha/(3\Delta), \varepsilon'/(2\Delta)\}$. Suppose then that $D \geqslant D_0$ and $n \geqslant n_0$, $G \in \mathcal{G}(n, D/n, \varepsilon, C)$ and that G_α is an n-vertex graph on $V(G)$ with $\delta(G_\alpha) \geqslant \alpha n$, and let $T \in \mathcal{T}(n, \Delta)$.

By Lemma 3 with $\beta = \alpha/(2\Delta^2)$, there exist subtrees $T_1 \subseteq T' \subseteq T$ so that $\alpha n/(2\Delta)^2 \leqslant |T_1| \leqslant \alpha n/(4\Delta)$, $e(T_1, T \smallsetminus T_1) = 1$ and $|T \smallsetminus T'| = 2\varepsilon' n$. We apply Lemma 5 and obtain an embedding g of T_1 in $H := G_\alpha \cup G$ such that $|B_{\widetilde{T}_1,H}(u,v)| \geqslant 2(\Delta+3)\varepsilon' n$ for any $u,v \in V$, where $\widetilde{T}_1 = g(T_1)$. Let $ab \in E(T)$ be the unique edge between T_1 and $T \smallsetminus T_1$ such that $a \in V(T_1)$, and let $\widetilde{a} = g(a)$. Define $T'' := T' \smallsetminus (T_1 \smallsetminus \{a\})$ and $H' := H \smallsetminus (V(\widetilde{T}_1) \smallsetminus \{\widetilde{a}\})$.

We want to apply Corollary 8 to find an embedding g' of T'' in H', with $g'(a) = \widetilde{a}$. So we need to verify the assumptions of Corollary 8 with $k = \varepsilon n - 1$. Firstly, note that by $\delta(G_\alpha) \geqslant \alpha n$ and $|T_1| \leqslant \alpha n/(2\Delta)$, we know that $\delta(H') \geqslant \alpha n - |T_1| \geqslant \alpha n/2 \geqslant \Delta \cdot 2k + 1$. Thus, condition (i) of Corollary 8 holds for sets on at most $2k$ vertices. Secondly, we claim that for any set $X \subseteq V(H')$ of size at least $k+1 = \varepsilon n$ we have $|V(H') \smallsetminus N_{H'}(X)| < \varepsilon n$. Indeed, since $G \in \mathcal{G}(n, D/n, \varepsilon, C)$ and both X and $V(H') \smallsetminus N_{H'}(X)$ are subsets of $V(H)$, if $|V(H') \smallsetminus N_{H'}(X)| \geqslant \varepsilon n$ then there is an edge in H, and hence H', between X and $V(H') \smallsetminus N_{H'}(X)$, a contradiction. Thus, since $|T_1| - 1 = |T'| - |T''| = |H| - |H'|$ and $|H| - |T'| = 2\varepsilon' n$, we have $|H'| - |T''| = |H| - |T'| = 2\varepsilon' n$, and thus, as $\varepsilon' \geqslant 2\Delta\varepsilon$,

$$|N(X)| \geqslant |H'| - \varepsilon n = |T''| + (2\varepsilon' - \varepsilon)n > |T''| + \Delta \cdot (2k+1).$$

Thus, we can apply Corollary 8 and obtain the embedding g' of T'' into H'. Combine g and g' to obtain an embedding g_0 of T' in H, and write $\widetilde{T}' = g_0(T')$.

For any $u,v,w \in V$ and any two trees S and S', observe that if $N_S(w) = N_{S'}(w)$ and $w \in B_{S,H}(u,v)$, then $w \in B_{S',H}(u,v)$. Since, by construction, for any vertex $w \in V(\widetilde{T}_1) \smallsetminus \{\widetilde{a}\}$ we have $N_{\widetilde{T}_1}(w) = N_{\widetilde{T}'}(w)$, and so $|B_{\widetilde{T}',H}(u,v)| \geqslant |B_{\widetilde{T}_1,H}(u,v)| - 1 \geqslant 2(\Delta+3)\varepsilon' n - 1$ for any $u,v \in V$.

It remains to embed the $2\varepsilon' n$ vertices in $V(T \smallsetminus T')$ to H. We achieve this using $B_{\widetilde{T'},H}(u,v)$ as explained at the beginning of this section. More precisely, since T' is connected, we can obtain T from T' by iteratively attaching one new leaf at a time, say using the sequence $T' := T'_0 \subseteq T'_1 \subseteq \cdots \subseteq T'_{2\varepsilon n} = T$. We claim that we can extend the embedding inductively while keeping $|B_{\widetilde{T}'_i,H}(u,v)| \geqslant |B_{\widetilde{T}'_{i-1},H}(u,v)| - (\Delta+3)$ for every $i \in [2\varepsilon' n]$, where each $\widetilde{T}'_i$ is the image of T'_i in H. Indeed, fix some index $i \in [2\varepsilon' n]$ and now we need to attach the vertex $b_i \in V(T'_i \smallsetminus T'_{i-1})$, whose parent $a_i \in T'_{i-1}$ has been embedded to $\widetilde{a}_i$. Pick any vertex v' in $V(H) \setminus V(\widetilde{T}'_{i-1})$. Since

$$|B_{\widetilde{T}'_{i-1},H}(\widetilde{a}_i, v')| \geqslant |B_{\widetilde{T}',H}(\widetilde{a}_i, v')| - (i-1)(\Delta+3) > 2(\Delta+3)\varepsilon' n - 1 - (i-1)(\Delta+3) > 0,$$

we can pick $w \in B_{\widetilde{T}'_{i-1},H}(\widetilde{a}_i, v')$ and let $c = g_{i-1}^{-1}(w)$. We now swap c out of the current embedding and use its previous image w to embed b_i, and embed c to v' instead. Precisely, define the new embedding g_i by $g_i(x) = g_{i-1}(x)$ for any $x \in V(T'_{i-1}) \setminus \{c\}$, $g_i(c) = v'$ and $g_i(b_i) = w$. Let $\widetilde{T}'_i = g_i(T'_i)$. Note that $N_{\widetilde{T}'_i}(x) = N_{\widetilde{T}'_{i-1}}(x)$ for all but at most $\Delta + 3$ vertices x in $V(\widetilde{T}_{i-1})$: the vertices $\widetilde{a}_i, v', w$ and the neighbors of w in $\widetilde{T}_{i-1}$ – because they are the vertices that are incident to the edges in $E(\widetilde{T}'_i) \setminus E(\widetilde{T}_{i-1})$. Thus, we have $|B_{\widetilde{T}'_i,H}(u, v)| \geqslant |B_{\widetilde{T}'_{i-1},H}(u, v)| - (\Delta + 3)$, for any $u, v \in V$, and we are done. ☐

§5. Tree universality in randomly perturbed dense graphs

In this section, we show how Theorem 2 implies Theorem 1, using the following simple proposition.

Proposition 9. *For any $\varepsilon > 0$ and $C \geqslant 2$ there exists D_0 such that the following holds for any $D \geqslant D_0$. The random graph $G(n, D/n)$ a.a.s. contains some graph $G \in \mathcal{G}(n, D/n, \varepsilon, C)$.*

Proof. Choose D_0 such that $D_0 \geqslant 10^4\varepsilon^{-2}$. Let $D \geqslant D_0$ and $H := G(n, D/n)$. Note that, by an application of Lemma 6 (or a simple Chernoff bound), the probability that, for all $U, W \subseteq V(H)$, with $|U|, |W| \geqslant \varepsilon n/10$, we have

$$3D|U||W|/(4n) \leqslant e_H(U, W) \leqslant 5D|U||W|/(4n) \tag{2}$$

is at most $2^{2n}e^{-D\varepsilon^2 n/4800} = 1 - o(1)$. Assume then that the property in (2) holds. We will show that there are few vertices with high degree in H.

Let $A \subseteq V(H)$ be the set of vertices with degree exceeding $5D/4$ in H, and note that $e_H(A, V(H)) > 5D|A|/4$. Thus, by the property in (2), we have that $|A| < \varepsilon n/10$.

If we delete all the edges incident to vertices of degree larger than $CD \geqslant 5D/4$ from H then we are left with a graph H' of maximum degree at most CD satisfying that for any two sets U and W of size at least εn, we have

$$e_{H'}(U, W) \geqslant \tfrac{3}{4}\tfrac{D}{n} \cdot |U \setminus A| \cdot |W \setminus A| \geqslant \tfrac{3}{4}\tfrac{D}{n} \cdot (9|U|/10) \cdot (9|W|/10) \geqslant \tfrac{1}{C}\tfrac{D}{n}|U||W|.$$

Thus, H' is in $\mathcal{G}(n, D/n, \varepsilon, C)$, as required. ☐

Proof of Theorem 1. Given α and Δ, let ε, D_0 and n_0 be given by Theorem 2 on inputting α, Δ and $C = 2$. We choose $D'_0 \geqslant D_0$ so that Proposition 9 with ε and C is applicable for $D \geqslant D'_0$. Since a.a.s. the random graph $G(n, D/n)$ contains a graph from $\mathcal{G}(n, D/n, \varepsilon, C)$ we have, by Theorem 2, that $G_\alpha \cup G(n, D/n)$ is a.a.s. $\mathcal{T}(n, \Delta)$-universal. ☐

§6. Concluding remarks

A graph G is called an (n, d, λ)-graph if $|G| = n$, G is d-regular and the second largest eigenvalue of the adjacency matrix of G in absolute value is at most λ. There is extensive literature on the properties of (n, d, λ)-graphs, see, e.g., a survey of Krivelevich and Sudakov [17]. It is known that (n,d,λ)-graphs G satisfy the so-called expander mixing lemma, that is, for all vertex subsets A, $B \subseteq V(G)$, we have

$$\left|e(A,B) - \tfrac{d}{n}|A||B|\right| \leqslant \lambda\sqrt{|A||B|}.$$

Our main technical result, Theorem 2, easily implies that, for any α and Δ, there is some sufficiently small ε such that, for any sufficiently large d and $\lambda \leqslant \varepsilon d/2$, any union of G_α, a graph on n vertices with minimum degree at least αn, with an (n,d,λ)-graph is $\mathcal{T}(n,\Delta)$-universal.

References

[1] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, 2016. arXiv:1612.00622. ↑6

[2] N. Alon, M. Krivelevich, and B. Sudakov, *Embedding nearly-spanning bounded degree trees*, Combinatorica **27** (2007), no. 6, 629–644. ↑2

[3] J. Balogh, B. Csaba, M. Pei, and W. Samotij, *Large bounded degree trees in expanding graphs*, Electronic Journal of Combinatorics **17** (2010), no. 1, R6. ↑2

[4] J. Balogh, A. Treglown, and A. Z. Wagner, *Tilings in randomly perturbed dense graphs*, 2017. arXiv:1708.09243. ↑2, 3

[5] W. Bedenknecht, J. Han, Y. Kohayakawa, and G. O. Mota, *Powers of tight Hamilton cycles in random perturbed hypergraphs*, in preparation. ↑3

[6] P. Bennett, A. Dudek, and A. M. Frieze, *Adding random edges to create the square of a Hamilton cycle*, 2017. arXiv:1710.02716. ↑3

[7] T. Bohman, A. Frieze, and R. Martin, *How many random edges make a dense graph Hamiltonian?*, Random Structures & Algorithms **22** (2003), no. 1, 33–42. ↑1

[8] B. Bollobás and A. G. Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38. ↑1

[9] J. Böttcher, R. H. Montgomery, O. Parczyk, and Y. Person, *Embedding spanning bounded degree subgraphs in randomly perturbed graphs*, in preparation. ↑2, 4

[10] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society **3** (1952), no. 1, 69–81. ↑1

[11] J. Han and Y. Zhao, *Embedding Hamilton ℓ-cycles in randomly perturbed hypergraphs*, in preparation. ↑3

[12] P. Haxell, *Tree embeddings*, Journal of Graph Theory **36** (2001), no. 3, 121–130. ↑3, 8

[13] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. ↑2

[14] A. D. Koršunov, *Solution of a problem of P. Erdős and A. Rényi on Hamiltonian cycles in undirected graphs*, Doklady Akademii Nauk SSSR **228** (1976), no. 3, 529–532. ↑1

[15] M. Krivelevich, M. Kwan, and B. Sudakov, *Cycles and matchings in randomly perturbed digraphs and hypergraphs*, Combinatorics, Probability and Computing **25** (2016), no. 6, 909–927. ↑3

[16] ———, *Bounded-degree spanning trees in randomly perturbed graphs*, SIAM Journal on Discrete Mathematics **31** (2017), no. 1, 155–171. ↑1, 2, 4

[17] M. Krivelevich and B. Sudakov, *Pseudo-random graphs*, More sets, graphs and numbers, 2006, pp. 199–262. ↑10

[18] A. McDowell and R. Mycroft, *Hamilton ℓ-cycles in randomly perturbed hypergraphs*, in preparation. ↑3

[19] R. Montgomery, *Embedding bounded degree spanning trees in random graphs*, 2014. arXiv:1405.6559v2. ↑2

[20] ———, *Spanning trees in random graphs*, 2016. Submitted. ↑2

[21] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), no. 4, 359–364. ↑1

Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK

Email addresses: `j.boettcher@lse.ac.uk`

Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508-090 São Paulo, Brazil

Email addresses: `{jhan|yoshi}@ime.usp.br`

Trinity College, Cambridge, CB2 1TQ, UK

Email addresses: `r.h.montgomery@dpmms.cam.ac.uk`

Institut für Mathematik, Goethe Universität, Robert-Mayer-Str. 6-10 60325 Frankfurt am Main, Germany

Email addresses: `{parczyk|person}@math.uni-frankfurt.de`