Universality for bounded degree spanning trees in randomly perturbed graphs

LSE Research Online URL for this paper: http://eprints.lse.ac.uk/100041/
Version: Published Version

Article:
Böttcher, Julia, Han, Jie, Kohayakawa, Yoshiharu, Montgomery, Richard, Parczyk, Olaf and Person, Yury (2019) Universality for bounded degree spanning trees in randomly perturbed graphs. Random Structures and Algorithms. ISSN 1042-9832
https://doi.org/10.1002/rsa.20850

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/
Universality for bounded degree spanning trees in randomly perturbed graphs

Julia Böttcher\textsuperscript{1} | Jie Han\textsuperscript{2} | Yoshiharu Kohayakawa\textsuperscript{3} | Richard Montgomery\textsuperscript{4} | Olaf Parczyk\textsuperscript{5} | Yury Person\textsuperscript{5}

\textsuperscript{1}Department of Mathematics, London School of Economics, London, UK
\textsuperscript{2}Department of Mathematics, University of Rhode Island, Kingston, Rhode Island
\textsuperscript{3}Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil
\textsuperscript{4}School of Mathematics, University of Birmingham, Birmingham, UK
\textsuperscript{5}Institut für Mathematik, Technische Universität Ilmenau, Ilmenau, Germany

Correspondence
Julia Böttcher, Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK.
Email: j.boettcher@lse.ac.uk

Funding information
This research was supported by the EPSRC; EP/R00532X/1, FAPESP; 2014/18641-5, 2013/03447-6, 2013/03447-6, CNPq; 310974/2013-5, 311412/2018-1, 459335/2014-6, DFG; PE 2299/1-1, CAPES-DAAD Probral; 430/15, 57350402, DAAD 57391197.

Abstract
We solve a problem of Krivelevich, Kwan and Sudakov 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 $G_n$ on $n$ vertices with $\delta(G_n) \geq \alpha n$ for $\alpha > 0$ and we add to it the binomial random graph $G(n, C/n)$, then with high probability the graph $G_n \cup G(n, C/n)$ contains copies of all spanning trees with maximum degree at most $\Delta$ simultaneously, where $C$ depends only on $\alpha$ and $\Delta$.

KEYWORDS
perturbed graphs, random graphs, spanning trees, universality

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 the family $(F_i)_{i \geq 1}$) such that $\delta(G_n) \geq \alpha n$ implies $F_n \subseteq G_n$. A prime example is Dirac’s theorem [10] stating that $\delta(G_n) \geq n/2$ ensures that $G_n$ is Hamiltonian if $n \geq 3$. 

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

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 $F$. 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) $F$ is a monotone property, there exists a threshold function $\hat{p} = \hat{p}(n)$ such that, when $p = \hat{p}(n)$, then $\lim_{n \to \infty} P[F_n \subseteq G(n, p)] = 0$, whereas, if $p = o(\hat{p}(n))$, then $\lim_{n \to \infty} P[F_n \subseteq G(n, p)] = 1$. When the conclusion of the latter case holds, we say that $G(n, p)$ contains $F$ asymptotically almost surely (a.a.s.). For example, a famous result of Koršunov [16] and Pósa [23] 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 $G_\alpha$ on $n$ vertices with $\delta(G_\alpha) \geq \alpha n$. Here, $\alpha$ can be arbitrarily small and hence $G_\alpha$ can be far from containing any Hamilton cycle. They proved that, after adding $m = C(\alpha)n$ edges uniformly at random to $G_\alpha$, the new graph $G$ becomes Hamiltonian a.a.s., where $C(\alpha)$ is a constant that depends only on $\alpha$. Letting $G_\alpha$ be the complete unbalanced bipartite graph $K_{\alpha n, (1-\alpha)n}$, one sees that the addition of linearly many edges to $G_\alpha$ is necessary for this result to hold in general. Furthermore, clearly, the conditions on $\delta(G_\alpha)$ and on $p = m/n^2$ 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 binomial $G(n, p)$ model, which is known to be essentially equivalent when $p = m/n^2$ (see, eg, [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 $G_\alpha$ corresponds to a small random perturbation in the structure of $G_\alpha$. 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 $C = C(\epsilon, \Delta)$, the random graph $G(n, C/n)$ alone a.a.s. contains any tree with at most $(1 - \epsilon)n$ vertices and maximum degree at most \Delta, where the bounds on $C = C(\epsilon, \Delta)$ 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 [22] showed that for each constant $\Delta$ and every sequence of trees $T_n$ with maximum degree $\Delta$, the threshold in $G(n, p)$ for a copy of $T_n$ to appear is $(\log n)/n$ (see also [21]). However, Krivelevich, Kwan and Sudakov [18] 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 $\Delta$ when $p = C(\Delta, \alpha)/n$.

In the concluding remarks of [18], 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 $\Delta$ 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 $\tau(n, \Delta)$ the family of all trees of maximum degree at most $\Delta$ 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 $G_\alpha$ is an $n$-vertex graph with $\delta(G_\alpha) \geq \alpha n$, then the randomly perturbed graph $G_\alpha \cup G(n, D/n)$ is a.a.s. $\tau(n, \Delta)$-universal.
This result is asymptotically optimal for $0 < \alpha < 1/2$, as with $G_a$ the complete unbalanced bipartite graph $K_{an,(1-\alpha)n}$ we need a linear number of edges from $G(n,p)$ already for the perfect matching. For $\alpha > 1/2$ this follows from a more general result, applied to $G_a$ alone, due to Komlós, Sárközy, and Szemerédi [15], for trees with maximum degree up to $n/\log n$. Kim and Joos [14] have succeeded in transferring this result to the perturbed model.

Theorem 1 is an immediate consequence of a technical theorem, Theorem 2, which states that the union of $G_a$ with any reasonably expanding graph $G$ is $T(n, \Delta)$-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 and the proof of Lemma 5 in Section 3.2).

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 [17], and tight Hamilton cycles in hypergraphs [11]. Most of the mentioned results exhibit the following phenomenon: in the presence of a dense graph $G_a$, a smaller edge probability than in $G(n,p)$ alone suffices. The only exception to this rule so far are $F$-factors for certain nonstrictly balanced graphs $F$ covered in [4]. Moreover, some variations of such results when $\alpha$ is at least some positive constant $c$ (which depends on other parameters of the problems at hand) were considered in [5,6,20].

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

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

**Theorem 2** (Main technical result). For any $\alpha > 0$ and integers $C \geq 2$ and $\Delta \geq 1$, there exist $\epsilon > 0$, $D_0$ and $n_0$ such that the following holds for any $D \geq D_0$ and $n \geq n_0$. Suppose $G \in \mathcal{G}(n,D/n,\epsilon,C)$ and $G_a$ are $n$-vertex graphs on the same vertex set and $\delta(G_a) \geq an$. Then $H := G_a \cup G$ is $T(n, \Delta)$-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 $G \in \mathcal{G}(n,D/n,\epsilon,C)$. We embed an arbitrary $T \in \mathcal{T}(n, \Delta)$ into $H := G_a \cup G$ in three phases. In the first phase, we find a subtree $T_1$ of $T$ (see Lemma 3) of small linear size, say $\beta n$ 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 \( N_T(w) \subseteq V(T_1) \) such 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). Alternatively, calling them switching sets would emphasize that each of them can only be used once.

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] (see Corollary 6), which ensures such almost spanning embeddings exist 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 \( T_i' = T_i \subseteq \cdots \subseteq T_{2\varepsilon n} = T \) is the sequence of subtrees of \( T \) that we encounter in this process. Suppose we already have the embedding \( g_{i-1} : V(T_{i-1}') \to V(H) \), and we wish to extend it to \( g_i : V(T_i') \to V(H) \) by defining the image of the leaf \( b \in V(T_i') \setminus V(T_{i-1}') \). Given some vertex \( v \) of \( H \) available for embedding \( b \) (that is, \( v \not\in g_{i-1}(V(T_{i-1}')) \)), 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 \( c \in B_{T,H}(g_{i-1}(u), v) \). We then let \( g_i(b) = g_{i-1}(c) \) and we let \( g_i(c) = v \). This defines a valid embedding \( g_i : V(T_i) \to V(H) \). (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 [18]. 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 \( \Delta \), 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 \( \beta, \varepsilon > 0 \) and let \( n, \Delta \) be positive integers such that \( \Delta \beta + 2\varepsilon < 1 \). Then, for any \( T \in \mathcal{T}(n, \Delta) \), there exist subtrees \( T_1 \subseteq T' \subseteq T \) such that

\[
(a) \quad \beta n \leq |T_1| \leq \Delta \beta n,
\]

\[
(b) \quad \varepsilon n \leq |T'| \leq (1-\varepsilon) n.
\]
Let |⟨v⟩s| ≥ βn. If |Cvi| > βn, then similarly there is a child v2 of v1 such that |Cv2| ≥ (Δβn−1)/(Δ−1) ≥ βn. Repeating this argument gives a desired v′ such that βn ≤ |Cv′| ≤ Δβn.

Let T be a tree. Given vertices x1, . . . , xn of T, let ⟨x1, . . . , xn⟩T be the minimal subtree of T that contains the vertices x1, . . . , xn, which is just the subtree of T obtained from the union of the vertex sets of all the paths between xi, xj, i ≠ j, in T. For two distinct vertices x, y of T, we write distT(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 ⊆ V(T) such that x /∈ Y, let distT(x, Y) := miny∈Y distT(x, y).

The following lemma provides us with vertices x1, . . . , xs in a tree T which cover T well, but are not too close. In particular, this gives us a collection of stars xi ∪ N[T(xi)] 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 ∈ N and vertices x1, . . . , xs ∈ V(T) such that

(a) for any 2 ≤ i ≤ s, distT(xi, ⟨x1, . . . , xi−1⟩T) = 5,
(b) |T|/(5Δ4) ≤ s ≤ (|T| + 4)/5, and
(c) distT(x, ⟨x1, . . . , x⟩T) ≤ 4 for all vertices x ∈ V(T).

Proof We start with picking x1 arbitrarily. We greedily pick the vertices x2, . . . , xs in V(T) sequentially as long as there is a vertex x i such that distT(xi, ⟨x1, . . . , xi−1⟩T) = 5. Note that for any 2 ≤ i ≤ s, |⟨x1, . . . , xi⟩T \ ⟨x1, . . . , xi−1⟩T| = 5, so we inductively get that |⟨x1, . . . , x⟩T| = 5s − 4. This implies s ≤ (|T| + 4)/5. Since T is connected, the maximality of s implies that distT(x, ⟨x1, . . . , x⟩T) ≤ 4 for all vertices x ∈ V(T). Thus we have |T| ≤ (5s − 4)Δ4, 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) ⊆ V. For v ∈ V, let

$$B_{T,G}(v) := \{w ∈ V(T) : N_T(w) ⊆ N_G(v)\}.$$  

For distinct vertices u and v ∈ 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 neighbor of the vertex embedded to \( u \).

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

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

**Proof of Lemma 5** First we choose the parameters \( D_0 \) and \( \varepsilon \) as follows:

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

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 \( 2 \leq i \leq s \), \( \text{dist}_{T_1}(x_i, \langle x_1, \ldots, x_{i-1} \rangle_{T_1}) = 5 \), and \( |T_1|/(5\Delta^4) \leq s \leq (|T_1| + 4)/5 \). Our embedding of \( T_1 \) consists of three steps. First we iteratively 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 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 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.

The following claim states that we can pick disjoint stars with \( \Delta \) 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 \( \Delta \) 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 \( N_{G_a}(u) \) and their leaves in \( N_{G_a}(v) \).

**Proof of the Claim** We randomly and sequentially pick \( s \) stars \( S_1, \ldots, S_s \) with \( \Delta \) 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 \( \tilde{x}_i \in N_{G_a}(u) \) and \( R_i \subseteq N_{G_a}(v) \), where \( \tilde{x}_i \) is the center of \( S_i \) and \( R_i \) is the set of leaves of \( S_i \). Since \( \delta(G_a) \geq an \), \( |T_1| \leq an/(2\Delta) \) and the existing stars cover at most

\[
(\Delta + 1)s \leq (\Delta + 1)\left( \frac{|T_1|}{5} + 4 \right) \leq (\Delta + 1)\left( \frac{an}{10\Delta} + 4 \right) \leq an/4
\]
vertices, there are at least \( 3an/4 \) vertices available in both \( U := N_{G_a}(u) \setminus \bigcup_{i \in \{1, \ldots, s\}} V(S_i) \) and \( W := N_{G_a}(v) \setminus \bigcup_{i \in \{1, \ldots, s\}} V(S_i) \).

Since \( G \in G(n, D/n, \varepsilon, C) \) and \( 3\alpha/4 \geq \varepsilon \), \( e(U,W) \geq D|U||W|/(Cn) \geq 3aD|U|/(4C) \). 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

\[
\sum_{u \in U} \left( \frac{\deg_{\Delta}(u)}{\Delta} \right) \geq |U| \left( \sum_{u \in U} \frac{\deg_{\Delta}(u)}{\Delta}/|U| \right) \geq |U| \left( \frac{3aD/(4C)}{\Delta} \right).
\]
Since $\Delta(G) \leq CD$, the total number of $K_{1,\Delta}$-stars in $G$ is at most $n(CD/\Delta)$. This allows us to obtain the following lower bound on $\mathbb{E}(Y_i^{u,v} | Y_1^{u,v}, \ldots, Y_{i-1}^{u,v})$:

$$\mathbb{E}(Y_i^{u,v} | Y_1^{u,v}, \ldots, Y_{i-1}^{u,v}) \geq \frac{|U|}{n} \left(\frac{3aD/(4C)}{\Delta} \right) \geq 2^{-\Delta-1}a^{\Delta+1}C^{-2\Delta}.$$ 

Let $p := 2^{-\Delta-1}a^{\Delta+1}C^{-2\Delta}$ and

$$x := sp \geq \frac{|T_i|}{5\Delta^4}p \geq \frac{a\Delta}{2\Delta^2 \cdot 5\Delta^4} \cdot \frac{a^{\Delta+1}}{2\Delta+1C^{2\Delta}} \geq 4(\Delta + 3)\epsilon n,$$

by the choice of $\epsilon$ in (1). Thus, by Lemma 2.2 (the sequential dependence lemma) from [1] with $\delta = 1/2$, or a simple coupling argument, we get

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

Thus by the union bound, we conclude that there is a choice of $S_1, \ldots, S_s$ such that, for each $u, v \in V$, $Y_1^{u,v} + \cdots + Y_s^{u,v} \geq 2(\Delta + 3)\epsilon n$, that is, 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 $\Delta$ leaves) on vertices $\{x_1\} \cup N_{T_1}(x_1)$ to an arbitrary partial subtree of $S_i$, with $x_i$ mapped to the center $\tilde{x}_i$. This gives us an embedding of the forest of stars $T[\{x_1\} \cup N_{T_1}(x_1) \cup \cdots \cup \{x_s\} \cup N_{T_1}(x_s)]$.

Next we extend our 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 $\langle x_1, \ldots, x_s \rangle$. Suppose we have connected the first $i - 1$ stars, that is, we have an embedding of $\langle x_1, \ldots, x_{i-1} \rangle_{T_1}$, and now we will connect it to $\tilde{x}_i$, the image of $x_i$. Recall that $\text{dist}_{T_1}(x_i, \langle x_1, \ldots, x_{i-1} \rangle_{T_1}) = 5$ 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 $H = G \cup G_a$. 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 $G_a$ which is not covered by the current partial forest as $g(y_4)$. This is possible because $\delta(G_a) \geq an$ and $|T_1| \leq an/(2\Delta)$. Note that, using $G_a$, there are at least $an/2$ choices for the image of $y_2$ and at least $an/2$ choices for the image of $y_3$, so, as $G \in \mathcal{G}(n, D/n, \epsilon, C)$, we can pick $\tilde{y}_2$ and $\tilde{y}_3$ so that $\tilde{y}_2\tilde{y}_3$ is an edge of $G$. Thus, the sequence $\tilde{x}_i, g(y_1), \tilde{y}_2, \tilde{y}_3, g(y_4), g(z)$ forms a path in $H$. Define $g(y_i) = \tilde{y}_i$ 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 $\delta(G_a) \geq an$ and $|T_1| \leq an/(2\Delta)$. Let $g$ be the resulting embedding function and $\tilde{T}_1 = g(T_1)$.

By the claim for any $u, v \in V$, there are at least $2(\Delta + 3)\epsilon n$ stars from $S_1, \ldots, S_s$ such that their centers are in $N_{G_a}(u)$ and their leaves are in $N_{G_a}(v)$. Since these stars are subtrees of $\tilde{T}_1$, we conclude that $|B_{\tilde{T}_1, H}(u, v)| \geq 2(\Delta + 3)\epsilon n$ 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 corollary of a tree embedding result of Haxell [12] (this is her Theorem 1 with $\epsilon' = 1$ and each $d_i = \Delta$). We note that it was first observed by Balogh, Csaba, Pei, and Samotij [3] that this is applicable in sparse random graphs. For a graph $G$ and vertex set $X \subseteq V(G)$, we let $N_G(X) := \bigcup_{x \in X} N_G(x)$. 


Corollary 6  Let $T$ be a tree with $t$ edges and maximum degree at most $\Delta$. Suppose $k \geq 1$ is an integer and $G$ is a graph satisfying the following two conditions:

(i) $|N_G(X)| \geq \Delta |X| + 1$ for every $X \subseteq V(G)$ with $1 \leq |X| \leq 2k$,

(ii) $|N_G(X)| \geq \Delta |X| + t + 1$ for every $X \subseteq V(G)$ with $k < |X| \leq 2k + 1$.

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 6. We then use the reservoir sets $B_{T,H}(u,v)$ to extend the embedding to cover the last few vertices.

Although we risk being somewhat repetitive, with the relevant definitions at hand, we are able to say more precisely how the sets $B_{T,H}(u,v)$ will help us to embed these last few vertices. Suppose we have a partial embedding $g : T' \rightarrow H$ of our tree $T$ into the host graph $H$, such that $T' \subseteq T$ is connected and let $\tilde{T}' = g(T')$. 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 neighbor in $V(T')$. Assume that this neighbor $a$ of $b$ in $T$ has been embedded to, 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 $B_{\tilde{T}',H}(u,v) \neq \emptyset$, then we can proceed with the embedding as follows. Let $w \in B_{\tilde{T}',H}(u,v)$ and note that, by the definition of $B_{\tilde{T}',H}(u,v)$, we have $w \in V(\tilde{T}')$. Let $c = g^{-1}(w)$, and let $g'(c) = v$ for any $x \in V(T') \setminus \{c\}$. Using the definition of $B_{\tilde{T}',H}(u,v)$, 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 $B_{\tilde{T}',H}(u,v)$ will be large enough to proceed greedily.

Proof  Given $\alpha, \Delta$ and $\varepsilon$, 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 $D \geq D_0$ and $n \geq n_0$, $G \in \mathcal{G}(n,D/n,\varepsilon,C)$ and that $G_a$ is an $n$-vertex graph on $V(G)$ with $\delta(G_a) \geq an$, 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 $an/(2\Delta^2) \leq |T_1| \leq an/(4\Delta)$ and $|T\setminus T'| = 2\varepsilon n$. We apply Lemma 5 and obtain an embedding $g$ of $T_1$ in $H := G_a \cup G$ such that $|B_{\tilde{T}_1,H}(u,v)| \geq 2(\Delta + 3)\varepsilon'n$ for any $u,v \in V$, where $\tilde{T}_1 = g(T_1)$. Let $ab \in E(T)$ be the unique edge between $T_1$ and $T\setminus T_1$ such that $a \in V(T_1)$, and let $\tilde{a} = g(a)$. Define $T'' := T\setminus\{T_1[\{a\}]\}$ and $H' := H\backslash(V(\tilde{T}_1)[\{\tilde{a}\}])$.

We want to apply Corollary 6 to find an embedding $g'$ of $T''$ in $H'$, with $g'(a) = \tilde{a}$. So we need to verify the assumptions of Corollary 6 with $k = \varepsilon n - 1$. First, note that by $\delta(G_a) \geq an$ and $|T_1| \leq an/(2\Delta)$, we know that $\delta(H') \geq an - |T_1| \geq an/2 \geq \Delta \cdot 2k + 1$. Thus, condition (i) of Corollary 6 holds for sets on at most $2k$ vertices. Second, we claim that for any set $X \subseteq V(H')$ of size at least $k + 1 = \varepsilon n$ we have $|V(H') \setminus N_{H'}(X)| < \varepsilon n$. Indeed, since $G \in \mathcal{G}(n,D/n,\varepsilon,C)$ and both $X$ and $V(H') \setminus N_{H'}(X)$ are subsets of $V(H)$, if $|V(H') \setminus N_{H'}(X)| \geq \varepsilon n$ then there is an edge in $H$, and hence $H'$, between $X$ and $V(H') \setminus N_{H'}(X)$, 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 $\varepsilon' \geq 2\Delta\varepsilon$,

$|N(X)| \geq |H'| - \varepsilon n = |T''| + (2\varepsilon' - \varepsilon)n > |T''| + \Delta \cdot (2k + 1)$.
Thus, we can apply Corollary 6 and obtain the embedding $g' \circ T''$ into $H'$. Combine $g$ and $g'$ to obtain an embedding $g_0$ of $T'$ in $H$, and write $\tilde{T'} = g_0(T')$.

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 $w \in V(\tilde{T}_1) \setminus \{a\}$ we have $N_{\tilde{T}_1}(w) = N_{\tilde{T}}(w)$, and so $|B_{\tilde{T},H}(u, v)| \geq |B_{\tilde{T}_1,H}(u, v)| - 1 \geq 2(\Delta + 3)\epsilon'n - 1$ for any $u, v \in V$.

It remains to embed the $2\epsilon'n$ vertices in $V(T' \setminus T')$ to $H$. We achieve this using $B_{\tilde{T}_i,H}(u, v)$ 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 $T' := T'_0 \subseteq T'_1 \subseteq \ldots \subseteq T'_{2\epsilon'n} = T$. We claim that we can extend the embedding inductively while keeping $|B_{\tilde{T}_i,H}(u, v)| \geq |B_{\tilde{T}_i-1,H}(u, v)| - (\Delta + 3)$ for every $i \in [2\epsilon'n]$, where each $\tilde{T}_i'$ is the image of $T_i'$ in $H$. Indeed, fix some index $i \in [2\epsilon'n]$ and now we need to attach the vertex $b_i \in V(T'_i \setminus T'_{i-1})$, whose parent $a_i \in T'_{i-1}$ has been embedded to $\tilde{a}_i$. Pick any vertex $v'$ in $V(H) \setminus \tilde{V}_{i-1}$.

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

we can pick $w \in B_{\tilde{T}_{i-1},H}(\tilde{a}_i, v')$ and let $c = g_i^{-1}(w)$. 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 $\tilde{T}_i' = g_i(T_i')$. Note that $N_{\tilde{T}_i}(x) = N_{\tilde{T}_{i-1}}(x)$ for all but at most $\Delta + 3$ vertices $x$ in $V(\tilde{T}_{i-1})$: the vertices $\tilde{a}_i, v', w$ and the neighbors of $w$ in $T_{i-1}$—because they are the vertices that are incident to the edges in $E(\tilde{T}_i') \setminus E(\tilde{T}_{i-1})$. Thus, we have $|B_{\tilde{T}_i,H}(u, v)| \geq |B_{\tilde{T}_{i-1},H}(u, v)| - (\Delta + 3)$, 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 7** For any $\epsilon > 0$ and $C \geq 2$ there exists $D_0$ such that the following holds for any $D \geq D_0$. The random graph $G(n, D/n)$ a.a.s. contains some graph $G \in \mathcal{G}(n, D/n, \epsilon, C)$.

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

$$3D|U||W|/(4n) \leq e_H(U, W) \leq 5D|U||W|/(4n)$$

is at least $1 - 2^{n}e^{-De^2n/4800} = 1 - \omega(1)$. 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 it satisfies $e_H(A, V(H)) > 5D|A|/4$. Thus, by the property in (2), we have that $|A| < \epsilon n / 10$.

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

$$e_G(U, W) \geq \frac{3D}{4n} \cdot |U \setminus A| \cdot |W \setminus A| \geq \frac{3D}{4n} \cdot (9|U|/10) \cdot (9|W|/10) \geq \frac{1D}{c \cdot n}|U||W|.$$ 

Thus, $G$ is in $\mathcal{G}(n, D/n, \epsilon, C)$, as required.
Proof of Theorem 1  Given \( \alpha \) and \( \Delta \), let \( \varepsilon, D_0 \) and \( n_0 \) be given by Theorem 2 on inputting \( \alpha, \Delta \) and \( C = 2 \). We choose \( D'_0 \geq D_0 \) so that Proposition 7 with \( \varepsilon \) and \( C \) is applicable for \( D \geq D'_0 \). Since a.a.s. the random graph \( G(n, D/n) \) contains a graph from \( G(n, D/n, \varepsilon, C) \) we have, by Theorem 2, that \( G_\alpha \cup G(n, D/n) \) is a.a.s. \( T(n, \Delta) \)-universal.

\[ \]

6  | CONCLUDING REMARKS

A graph \( G \) is called an \((n, d, \lambda)\)-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 \( \lambda \). There is extensive literature on the properties of \((n, d, \lambda)\)-graphs, see, for example, a survey of Krivelevich and Sudakov [19]. It is known that \((n, d, \lambda)\)-graphs \( G \) satisfy the so-called expander mixing lemma, that is, for all vertex subsets \( A, B \subseteq V(G) \), we have

\[
|e(A, B) - \frac{d}{n} |A||B| | \leq \lambda \sqrt{|A||B|}.
\]

Our main technical result, Theorem 2, easily implies that, for any \( \alpha \) and \( \Delta \), there is some sufficiently small \( \varepsilon \) such that, for any sufficiently large \( d \) and \( \lambda \leq \varepsilon d/2 \), any union of \( G_\alpha \), a graph on \( n \) vertices with minimum degree at least \( \alpha n \), with an \((n, d, \lambda)\)-graph is \( T(n, \Delta) \)-universal.

ACKNOWLEDGMENTS

J.B. is partially supported by EPSRC (EP/R00532X/1). J.H. is supported by FAPESP (2014/18641-5, 2013/03447-6). Y.K. is partially supported by FAPESP (2013/03447-6) and CNPq (310974/2013-5, 311412/2018-1, 459335/2014-6). O.P. and Y.P. 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, 57350402). FAPESP is the São Paulo Research Foundation. CNPq is the National Council for Scientific and Technological Development of Brazil.

REFERENCES


**How to cite this article:** Böttcher J, Han J, Kohayakawa Y, Montgomery R, Parczyk O, Person Y. Universality for bounded degree spanning trees in randomly perturbed graphs. *Random Struct Alg*. 2019;1–11. [https://doi.org/10.1002/rsa.20850](https://doi.org/10.1002/rsa.20850)