Bounded-degree spanning trees in randomly perturbed graphs

Michael Krivelevich, Matthew Kwan, Benny Sudakov

March 16, 2015

Abstract

We show that for any bounded-degree spanning tree T and any fixed dense graph G, a modest random perturbation of G will typically contain a copy of T. This combines the viewpoints of the well-studied problems of embedding trees into fixed dense graphs and into random graphs, and extends a sizeable body of existing research on randomly perturbed graphs. Specifically, we show that there is $c = c(\alpha, \Delta)$ such that if G is a graph on the vertex set $[n]$ with minimum degree at least αn, and T is an n-vertex tree with maximum degree at most Δ, and R is a uniformly random graph on $[n]$ with cn edges, then $T \subseteq G \cup R$ asymptotically almost surely as $n \to \infty$. Our proof depends on a lemma concerning the decomposition of a dense graph into super-regular pairs of uneven sizes, which may be of independent interest.

1 Introduction

A classical theorem of Dirac [6] states that any n-vertex graph ($n \geq 3$) with minimum degree at least $n/2$ has a Hamilton cycle: a cycle that passes through all the vertices of the graph. More recently, there have been many results showing that this kind of minimum degree condition is sufficient to guarantee the existence of different kinds of spanning graphs. For example, [9, Theorem 1′] says that for any Δ and $\gamma > 0$, any n-vertex graph (n sufficiently large) with minimum degree $(1/2 + \gamma)n$ contains every spanning tree which has maximum degree at most Δ. (This has been also been generalized further in two directions: to allow Δ to grow with n and to allow for much more general spanning subgraphs than spanning trees; see [11, 5]).

The constant "1/2" in these Dirac-type theorems is tight: in order to guarantee the existence of these spanning subgraphs we require very strong density conditions. But the picture is very different for a "typical" large graph. If we fix an arbitrarily small $\alpha > 0$ and select a graph uniformly at random among the (labelled) graphs with n vertices and $\alpha\binom{n}{2}$ edges, then the degrees will probably each be about αn. Such a random graph is Hamiltonian with probability $1 - o(1)$ (we say it is Hamiltonian *asymptotically almost surely*, or *a.a.s.*). This follows from a stronger result [17] that gives a *threshold* for Hamiltonicity: a random n-vertex, m-edge graph is Hamiltonian a.a.s. if $m \gg n \log n$, and fails to be Hamiltonian a.a.s. if $m \ll n \log n$. Although the exact threshold for bounded-degree spanning trees is not known, it was proved in [16] that for any Δ and any tree T with maximum degree at most Δ, a random n-vertex graph with $\Delta n(\log n)^5$ edges a.a.s. contains T. Here and from now on, all asymptotics are as $n \to \infty$, and we implicitly round large quantities to integers.

In [3], the authors studied Hamiltonicity in the random graph model that starts with a dense graph and adds m random edges. This model is a natural generalization of the ordinary random graph model where we start with nothing, and offers a "hybrid" perspective combining the extremal and probabilistic settings of the last two paragraphs. It has since been studied in a number of other contexts; see for example [2, 14, 13]. One particularly important motivation for the model is the notion of *smoothed analysis* of algorithms introduced in [18]. This is is a hybrid of worst-case and average-case analysis, studying the performance of algorithms in the more realistic setting of inputs that are "noisy" but not completely random. A property that holds a.a.s. with small m in our random graph model can be said to hold not just for a "globally" typical graph but for the typical graph in every small "neighbourhood" of our space of graphs. This tells us that the graphs which fail to satisfy our property are in some sense "fragile". We highlight that it is generally very easy to transfer results from our model of random perturbation to other natural models, including those that delete as well as add edges. This can be accomplished with standard coupling and conditioning arguments.

The statement of [3, Theorem 1] is that for every $\alpha > 0$ there is $c = c(\alpha)$ such that if we start with a graph with minimum degree at least αn and add cn random edges, then the resulting graph will a.a.s. be Hamiltonian. This saves a logarithmic factor over the usual model where we start with nothing. Note that some dense graphs require a linear number of extra edges to become Hamiltonian (consider the complete bipartite graph with partition sizes $n/3$ and $2n/3$), so the order of magnitude of this result is tight. We can interpret this theorem as quantifying the "fragility" of the few dense graphs that are not Hamiltonian, by determining the amount of random perturbation that is necessary to make a dense graph Hamiltonian.

Let $\mathbb{G}(n,p)$ be the Erdős-Rényi random graph model with vertex set $[n] = \{1,\ldots,n\}$, where each edge is independently present with probability p. (For our purposes this is equivalent to the model that uniformly at random selects a $p\binom{n}{2}$-edge graph on the vertex set $[n]$). In this paper we prove the following theorem, extending the aforementioned result to bounded-degree spanning trees.

Theorem 1. *There is $c = c(\alpha, \Delta)$ such that if G is a graph on the vertex set $[n]$ with minimum degree at least αn, and T is an n-vertex tree with maximum degree at most Δ, and $R \in \mathbb{G}(n, c/n)$, then a.a.s. $T \subseteq G \cup R$.*

A key ingredient of the proof is the following lemma: with the random edges in R alone, we can embed trees that are not too big.

Lemma 2 ([1, Theorem 1.1])**.** *There is $c = c(\varepsilon, \Delta)$ such that $G \in \mathbb{G}(n, c/n)$ a.a.s. contains every tree of maximum degree at most Δ on $(1-\varepsilon)n$ vertices.*

We will split the proof of Theorem 1 into two cases in a similar way to [12]. If our spanning tree T has many leaves, then we remove the leaves and embed the resulting non-spanning tree in R using Lemma 2. To complete this into our spanning tree T, it remains to match the the vertices needing leaves with leftover vertices. This amounts to finding a perfect matching in a certain bipartite graph. The random edges in R provide several almost-perfect matchings on their own; we combine these with the dense graph G to satisfy the conditions of Hall's marriage theorem and guarantee the required perfect matching. The details for this case are given in Section 2.1.

The more difficult case is where T has few leaves, which we attack in Section 2.2. In this case T cannot be very "complicated" and must be a subdivision of a small tree. In particular T must have many long *bare paths*: paths where each vertex has degree exactly two. By removing these bare paths we obtain a small forest which we can embed into R using Lemma 2 (the details are in Section 2.2.2). In order to complete this forest into our spanning tree T, we need to join up distinguished pairs of vertices with disjoint paths of certain lengths. In order to make this task feasible, in Section 2.2.1 we first use Szemerédi's regularity lemma to divide the vertex set into a bounded number of pieces, each of which induces a "super-regular pair" in G (that is, a dense subgraph with edges very well-distributed). In Section 2.2.3 we use R to find most of our desired paths, in such a way that it only remains to join up pairs of vertices within the same super-regular pair. We make some further adjustments with R (Section 2.2.4), after which the super-regularity of the pairs allows us to find the rest of our paths using a tool called the "blow-up lemma" (Section 2.2.5).

2 Proof of Theorem 1

We split the random edges into multiple independent "phases": say $R \supseteq R_1 \cup R_2 \cup R_3 \cup R_4$, where $R_i \in \mathbb{G}(n, c_i/n)$ for some large $c_i = c_i(\alpha, \Delta)$ to be determined (these c_i will in turn determine c).

2.1 Case 1: T has many leaves

For our first case suppose there are at least λn leaves in T (for some $\lambda = \lambda(\alpha) > 0$ to be determined in the second case). Then consider the tree T' with some λn leaves removed. By Lemma 2 we can embed T' into R_1. Let A' be the set of vertices of T' which had a leaf deleted from them, and let B be the set of λn vertices not part of T'. We can assume A' and B are uniformly random disjoint sets of the appropriate size (note $|A'| \geq \lambda n/\Delta$ and $|B| = \lambda n$). For each $b \in B$, the number of neighbours in A' is hypergeometrically distributed with expected value at least $\lambda \alpha n/\Delta$. Let $\beta = \lambda\alpha/(2\Delta)$; by a concentration inequality (for example, see [8, Theorem 2.10]) and the union bound, a.a.s. each $a' \in A'$ has at least βn neighbours in B. Similarly each $b \in B$ a.a.s. has at least βn neighbours in A'. From now on we treat A' and B as fixed sets satisfying these properties.

Let A be a set consisting of ℓ copies of each vertex $a' \in A'$ which needs ℓ extra leaves (so $|A| = \lambda n$). For any graph H on $[n]$, we define a bipartite graph $G_{A,B}(H)$ on the vertex set $A \cup B$, with an edge between $a \in A$ (a copy of $a' \in A'$) and $b \in B$ if there is an edge between a' and b in H. If we can find a perfect matching in $G_{A,B}(G) \cup G_{A,B}(R_2) \subseteq G_{A,B}(G \cup R)$, this gives us a way to extend our embedding of T' to an embedding of T, as desired.

Let $\mathbb{B}_{A,B}(p)$ be the random binomial graph where each of the $|A||B|$ possible edges between A and B are present with probability p. For any $c_{\mathbb{B}}$, we can choose large c_2 such that $1 - c_2/n \leq (1 - c_{\mathbb{B}}/n)^{\Delta}$ for large n. This means $G_{A,B}(R_2)$ stochastically dominates $\mathbb{B}_{A,B}(c_{\mathbb{B}}/n)$, so the following lemma provides the desired perfect matching.

Lemma 3. *Let A and B be n-vertex sets. There is $c_{\mathbb{B}} = c_{\mathbb{B}}(\alpha)$ such that if G is a bipartite graph with bipartition $A \cup B$ and minimum degree at least αn, and if $R \in \mathbb{B}_{A,B}(c_{\mathbb{B}}/n)$, then a.a.s. $G \cup R$ has a perfect matching.*

In order to prove Lemma 3 we need another lemma.

Lemma 4. *Let $\varepsilon > 0$, and let A and B be n-vertex sets. There is $c = c(\varepsilon)$ such that $R \in \mathbb{B}_{A,B}(c/n)$ has a matching of size $(1-\varepsilon)n$.*

Proof. For any subsets A' and B' of size εn, the probability there is no edge between A' and B' is $(1-p)^{(\varepsilon n)^2} \leq e^{-c\varepsilon^2 n}$. There are at most 2^{2n} choices of such A', B', so for large c, by the union bound there is a.a.s. an edge between any such pair of sets. It follows that a maximum-size matching a.a.s. has at least $(1-\varepsilon)n$ edges. $\square$

Proof of Lemma 3. We apply Hall's marriage theorem. If $W \subseteq A$ with $|W| \le \alpha n$ then $|N_G(W)| \ge \alpha n \ge |W|$ by the minimum degree condition on G. Similarly, every vertex outside $N_G(W)$ has at least αn neighbours outside W, so if $|W| \ge (1-\alpha)n$ then $|N_G(W)| = n \ge |W|$. So, consider any W with $\alpha n \le |W| \le (1-\alpha)n$. Split R into phases $R \supseteq R_1 \cup \dots \cup R_r$, with $R_i \in \mathbb{B}_{A,B}(c'/n)$, for some $c' = c'(\alpha)$ and $r = r(\alpha)$ to be determined. By Lemma 4, if c' is large we can find a $(1-\alpha/3)n$-edge matching M_i in each R_i. Let $N_i = N_{M_i}(W)$ be the set of vertices matched with W in each M_i, and note that each $|W| - \alpha n/3 \le |N_i| \le |W|$. Condition on N_1, and condition on the number of vertices $|N_i|$ matched by W in each M_i. Now, for each $i > 1$, note that $N_i \cap N_1$ is hypergeometrically distributed with mean $|N_i||N_1|/n$. Note that

$$|N_i||N_1|/n \le |W|^2/n \le |W|(1-\alpha) \le 2(|W| - \alpha/3) - \alpha/3 \le |N_i| + |N_1| - |W|,$$

so by a concentration inequality for the hypergeometric distribution,

$$\Pr(|N_R(W)| < |W|) \le \prod_{i=2}^{r} \Pr(|N_i \cap N_1| > |N_i| + |N_1| - |W|) = e^{-\Omega(r\alpha^2 n)}.$$

If r is large this probability is $o(2^{-n})$, and since there are at most 2^n choices for W, the union bound gives $|N_R(W)| \ge |W|$ for all required W. $\square$

2.2 Case 2: T has few leaves

Now we address the second case where there are fewer than λn leaves in T.

2.2.1 Partitioning into super-regular pairs

As outlined, we first need to divide G into a bounded number of pairs of clusters with edges well-distributed between them. This partition will inform the way we use R to embed T.

Definition 5. For a disjoint pair of vertex sets (X, Y) in a graph, let its *density* $d(X, Y)$ be the number of edges between X and Y, divided by $|X||Y|$. A pair of vertex sets (V_1, V_2) is said to be *ε-regular* if for any U_1, U_2 with $U_h \subseteq V_h$ and $|U_h| \ge \varepsilon |V_h|$, we have $|d(U_1, U_2) - d(V_1, V_2)| \le \varepsilon$. If alternatively $d(U_1, U_2) \ge \delta$ for all such pairs U_1, U_2 then we say (V_1, V_2) is (ε, δ)-*dense*. Let $\overline{h} = 2 - h$; if (V_1, V_2) is (ε, δ)-dense and moreover each $v \in V_h$ has at least $\delta \left|V_{\overline{h}}\right|$ neighbours in $V_{\overline{h}}$, then we say (V_1, V_2) is (ε, δ)-*super-regular*.

Lemma 6. *For $\alpha, \varepsilon > 0$ with ε sufficiently small relative to α, there are $\delta = \delta(\alpha) > 0$, $M = M(\alpha, \varepsilon)$, $a = a(\alpha, \varepsilon) > 0$ and $b = b(\alpha, \varepsilon) > 0$ such that the following holds. Let G be a graph on $[n]$ with minimum degree at least αn. We can choose $q \le M$ and partition $[n]$ into sets V_i^h ($1 \le i \le q$, $h = 1, 2$) such that each $an \le \left|V_i^h\right| \le bn$ and each pair $\left(V_i^1, V_i^2\right)$ is (ε, δ)-super-regular.*

To prove Lemma 6, we will apply Szemerédi's regularity lemma to obtain a reduced *cluster graph*, then decompose this cluster graph into small stars. In each star T_i, the centre cluster will correspond to V_i^2 and the leaf clusters will be combined to form V_i^2. We will then have to redistribute some of the vertices to ensure super-regularity. Before giving the details of the proof, we give a statement of (a version of) Szemerédi's regularity lemma and some auxiliary lemmas for working with regularity and super-regularity.

Lemma 7 (Szemerédi's regularity lemma, minimum degree form)**.** *For every $\alpha > 0$, and any $\varepsilon > 0$ that is sufficiently small relative to α, there is $\alpha' = \alpha'(\alpha) > 0$ and $K = K(\varepsilon)$ such that the following holds. For any graph G of minimum degree at least $\alpha|G|$, there is a partition of $V(G)$ into clusters $V_0, V_1, \dots V_k$ ($k \leq K$), and a spanning subgraph G' of G, satisfying the following properties. The "exceptional cluster" V_0 has size at most εn, and the other clusters have equal size sn. The minimum degree of G' is at least $\alpha' n$. There are no edges of G' within clusters, and each pair of non-exceptional clusters is ε-regular in G' with density zero or at least α'. Moreover, define the* cluster graph *C as the graph whose vertices are the k non-exceptional clusters V_i, and whose edges are the pairs of clusters between which there is nonzero density in G'. The minimum degree of C is at least $\alpha' k$.*

This version of Szemeredi's regularity lemma follows directly from [15, Proposition 9].

We now give some simple lemmas about regularity, which we will use in the proof of Theorem 1.

Lemma 8. *Suppose $V_1, \dots, V_r$ are clusters of the same sizes such that each $\left(V^1, V_i^2\right)$ is (ε, δ)-dense. Let $V^2 = \bigcup_{i=1}^r V_i^2$; then $\left(V^1, V^2\right)$ is $(\varepsilon, (\delta - \varepsilon)/r)$-dense.*

Proof. Let $U^h \subseteq V^h$ with $\left|U^h\right| \geq \varepsilon\left|V^h\right|$. Let V_i^2 be the cluster which has the largest intersection with U^2, so we have $\left|U^2 \cap V_i^2\right| \geq \varepsilon\left|V^2\right|/r = \varepsilon\left|V_i^2\right|$, and there are therefore at least $(\delta - \varepsilon)\left|U^1\right|\left|U^2 \cap V_i^2\right| \geq ((\delta - \varepsilon)/r)\left|U^1\right|\left|U^2\right|$ edges between U^1 and U^2. □

Lemma 9. *Let $\left(V^1, V^2\right)$ be an (ε, δ)-dense pair, let $0 \leq \gamma \leq 1$, and let $\left(W^1, W^2\right)$ be a pair such that each $\left|V^h \triangle W^h\right| \leq \gamma\left|V^h\right|$. Then $\left(W^1, W^2\right)$ is an (ε', δ')-dense pair, with $\varepsilon' = \varepsilon + 6\sqrt{\gamma/(1+\gamma)}$ and $\delta' = \delta - 4\sqrt{\gamma/(1+\gamma)}$ If moreover $\left(V^1, V^2\right)$ is (ε, δ)-super-regular and each vertex in W^h has at least $\delta\left|W^{\overline{h}}\right|$ neighbours in $W^{\overline{h}}$, then $\left(W^1, W^2\right)$ is an (ε', δ')-super-regular pair.*

Lemma 9 says that regularity-type properties are "robust" with respect to small alterations of the clusters. It is almost the same as [4, Proposition 8], and has exactly the same proof.

Lemma 10. *Every (ε, δ)-dense pair $\left(V^1, V^2\right)$ contains a $(2\varepsilon, \delta - 2\varepsilon)$-super-regular sub-pair $\left(W^1, W^2\right)$, where $\left|W^h\right| \leq (1 - \varepsilon)\left|V^h\right|$.*

Lemma 10 follows easily from the same proof as [4, Proposition 6].

Now we prove Lemma 6. First we need a lemma for our decomposition into small stars.

Lemma 11. *Let G be an n-vertex graph with minimum degree at least αn. Then there is a spanning subgraph Q which is a disjoint union of vertex-disjoint stars, each with at least two vertices and at most $1/\alpha$ leaves.*

Proof. Let Q be a union of such stars with the maximum number of vertices. Suppose there is a vertex v uncovered by Q. If v has a neighbour which is a centre of one of the stars in Q, and that star has fewer than $1/\alpha$ leaves, then we could add v to that star, contradicting maximality. Otherwise, if v has a neighbour w which is a leaf of one of the stars in Q, then we could remove that leaf from its star and create a new 2-vertex star with edge vw, again contradicting maximality. The remaining case is where each of the (at least αn) neighbours of v is a centre of a star with $1/\alpha$ leaves. But these stars would comprise more than n vertices, which is again a contradiction. We conclude that Q covers G, as desired. □

Proof of Lemma 6. Apply our minimum degree form of Szemerédi's regularity lemma, with some ε' to be determined. Let Q be a cover of C by stars $T_1, \ldots, T_q$ of size at most $1/\alpha'$, as guaranteed by Lemma 11. Let W_i^1 be the centre cluster of T_i, and let W_i^2 be the union of the leaf clusters of T_i (for two-vertex stars, arbitrarily choose one vertex as the "leaf" and one as the "centre"). By Lemma 8, with $\delta' = (\alpha')^2/2$ each pair $\left(W_i^1, W_i^2\right)$ is (ε', δ')-dense.

Apply Lemma 10 to obtain sets $V_i^h \subseteq W_i^h$ such that $\left|V_i^h\right| = (1-\varepsilon)\left|W_i^h\right|$ and each $\left(V_i^1, V_i^2\right)$ is $(\varepsilon'', \delta'')$-super-regular, for some $\varepsilon'' = \varepsilon''(\varepsilon') > 0$ and $\delta'' = \delta''(\delta) > 0$. Combining the exceptional set V_0 and all the $V_i^h \backslash W_i^h$, there are at most $2\varepsilon' n$ "bad" vertices not part of a super-regular pair.

Each bad vertex v has at least $(\alpha - 2\varepsilon')n$ neighbours to the clusters V_i^h, and if $\delta'' < \alpha - 2\varepsilon'$ then some cluster must have $\delta'' s(1-\varepsilon')$ neighbours of v. More generally, for small δ'' and ε' there are at least

$$\frac{(\alpha - 2\varepsilon' - \delta'')}{(1-\delta'')s(1-\varepsilon')} \geq \alpha/(2s)$$

clusters which have at least $\delta'' s(1-\varepsilon')$ neighbours of v. Put v in one of these clusters uniformly at random (do this independently for each bad v). By a concentration inequality, after this procedure a.a.s. at most $2(2\varepsilon' n)(2s/\alpha)$ bad vertices have been added to each V_i^h. By Lemma 9, for small ε' each $\left(V_i^1, V_i^2\right)$ is now $(\varepsilon'' f, \delta''/2)$-super-regular, for some $f = f(\alpha)$. With ε'' small relative to ε, we are done. □

We apply Lemma 6 to our graph G, with $\varepsilon = \varepsilon(\delta)$ to be determined. For $\mathbf{i} = (i, h)$, let $V_{\mathbf{i}} = V_i^h$, and let $|V_{\mathbf{i}}| = s_{\mathbf{i}} n$.

2.2.2 Embedding a small subforest of T

Proceeding with the proof outline, we need the fact that T is almost entirely composed of bare paths, as is guaranteed by the following lemma.

Lemma 12 ([12, Lemma 2.1])**.** *Let T be a tree on n vertices with at most ℓ leaves. Then T contains a collection of at least $(n - (2\ell - 2)(k+1))/(k+1)$ vertex-disjoint bare paths of length k each.*

In our case $\ell = \lambda n$. If we choose k large enough and λ small enough, then T contains a collection of $\psi n := n/(2(k-1))$ disjoint bare k-paths. (We also impose that k is odd, for reasons that will become clear later). If we delete the interior vertices of these paths, then we are left with a forest F on $n/2$ vertices.

Now, embed F into R_1. There are ψn "special pairs" of vertices of $F \subseteq R_1$ that need to be connected with k-paths. We call such connecting paths "special paths". For each special pair, arbitrarily choose one vertex to be "x-type" and the other to be "y-type". Let X and Y be the set of x-type and y-type vertices, respectively. By symmetry we can assume

$$X \,\cup\, Y \,\cup\, F\backslash(X \cup Y) \,\cup\, V\backslash F$$

is a uniformly random partition of V into parts of sizes ψn, ψn, $n/2 - 2\psi n$, $n/2$, and that the special pairs correspond to a random bijection between X and Y. Let $X_{\mathbf{i}} = V_{\mathbf{i}} \cap X$ and $Y_{\mathbf{i}} = V_{\mathbf{i}} \cap Y$, and let $Z_{\mathbf{i},\mathbf{j}}$ be the set of special pairs in $X_{\mathbf{i}} \times Y_{\mathbf{j}}$. By a concentration inequality for the hypergeometric distribution and the union bound, a.a.s. each $|V_{\mathbf{i}} \backslash (F \backslash (X \cup Y))| \sim (1/2 + 2\psi)s_{\mathbf{i}} n$. Similarly each $|X_{\mathbf{i}}| \sim |Y_{\mathbf{i}}| \sim \psi s_{\mathbf{i}} n$, and each $|Z_{\mathbf{i},\mathbf{j}}| \sim \psi s_{\mathbf{i}} s_{\mathbf{j}} n$.

Now that we have embedded F, we no longer care about the vertices used to embed $F\backslash(X \cup Y)$; they will never be used again. It is convenient to imagine, for the duration of the proof, that instead of embedding parts of T, we are "removing" vertices from $G \cup R$, gradually making the remaining graph easier to deal with. So, remove from each $V_{\mathbf{i}}$ and $Z_{\mathbf{i},\mathbf{j}}$ all vertices used in $F\backslash(X \cup Y)$. Update n to be the number of vertices not used to embed $F\backslash(X \cup Y)$ (previously $(1/2+2\psi)n$), and update each $s_{\mathbf{i}}$ to still satisfy $s_{\mathbf{i}}n = |V_{\mathbf{i}}|$ (this does not change $s_{\mathbf{i}}$ asymptotically). We now have

$$|Z_{\mathbf{i},\mathbf{j}}| \sim \psi s_{\mathbf{i}} s_{\mathbf{j}} n/(2(1/2+2\psi)) = s_{\mathbf{i}} s_{\mathbf{j}} n/(k+1).$$

Also let $W_{\mathbf{i}} = V_{\mathbf{i}}\backslash(X \cup Y)$.

Here and in future parts of the proof, it is critical that after removing vertices from the $V_{\mathbf{i}}$ we do not destroy super-regularity. This will mostly be guaranteed by making choices randomly and applying the following lemma.

Lemma 13. *Fix any $a > 0$. Let $\left(V^1, V^2\right)$ be an (ε, δ) super-regular pair in a graph G, where each $\left|V^h\right| \geq n$. For any $n_h \geq an$, let W^h be a uniformly random n_h-vertex subset of V^h. Then $\left(W^1, W^2\right)$ is a.a.s. $(\varepsilon, \delta/2)$-super-regular.*

Proof. Since $W^h \subseteq V^h$, by definition $\left(W^1, W^2\right)$ is $(\varepsilon, \delta/2)$-dense. Each vertex in W^h has at least $\delta\left|V^{\overline{h}}\right|$ neighbours in $V^{\overline{h}}$; by a concentration inequality for the hypergeometric distribution and the union bound, at least $\delta\left|W^{\overline{h}}\right|/2$ of those remain in $W^{\overline{h}}$. □

Let $X_{\mathbf{i},\mathbf{j}}$ and $Y_{\mathbf{i},\mathbf{j}}$ be the set of x-type and y-type vertices in pairs in $Z_{\mathbf{i},\mathbf{j}}$, respectively, so that the clusters $X_{\mathbf{i},\mathbf{j}}, Y_{\mathbf{i},\mathbf{j}}, W_{\mathbf{i}}$ (each of size $\Omega(n)$) partition the vertex set. Let $\overline{(i,h)} = \left(i, \overline{h}\right)$; an immediate consequence of Lemma 13 is that each $\left(X_{\mathbf{i},\mathbf{j}}, W_{\overline{\mathbf{i}}}\right)$, $\left(W_{\mathbf{i}}, W_{\overline{\mathbf{i}}}\right)$ and $\left(W_{\mathbf{i}}, Y_{\mathbf{j},\overline{\mathbf{i}}}\right)$ are $(\varepsilon, \delta/2)$-super-regular.

2.2.3 Embedding special paths not compatible with the super-regular partition

Next, we want to eliminate almost all special pairs which are not between clusters of a super-regular pair. This is achieved using the following lemma (which will be used again later).

Lemma 14. *For any $\delta > 0$, $f < 1$, and any integer $k > 2$ there is $c = c(\delta, f, k)$ such that the following holds.*

Let G be a graph on some vertex set $[N]$, together with a vertex partition into $O(1)$ clusters, such that some of the pairs of clusters (U, W) are (ε, δ)-super-regular. Consider any sequence of clusters $X, V_1, \ldots, V_r, Y$, and any partition $t_1 + \cdots + t_r = k-1$ (with each $t_i \in \mathbb{N}$), such that $|V_i| \geq t_i n$ and $|X|, |Y| \geq n$. Suppose that (X, V_1) and (V_r, Y) are (ε, δ)-super-regular pairs, and suppose further that each $x \in X$ is bijectively paired with a vertex $y \in Y$, comprising a special pair (x, y). Let $R \in \mathbb{G}(N, c/N)$.

Then, for any $m \leq fn$, in $G \cup R$ there are m vertex-disjoint special paths, each with t_i vertices in each V_i. These paths can be chosen such that after their removal, each (U, W) that was previously (ε, δ)-super-regular is now $(\varepsilon, \delta/4)$-super-regular.

It is convenient to prove Lemma 14 using a similar, but simpler, version of the lemma.

Lemma 15. *For any $\delta > 0$, $g < 1$ and $k \in \mathbb{N}$, there is $c = c(\delta, p, k)$ such that the following holds. Let G be a graph on the vertex set $[(k+1)n]$, together with a vertex partition $[(k+1)n] = X \cup V_1 \cdots \cup V_{k-1} \cup Y$. Suppose that each $|V_i| = n$, and each $v \in V_0$ (respectively, $v \in V_k$) has at least δn neighbours in V_2 (respectively, in V_{k-1}). Suppose further that each $x \in X$ is bijectively paired with a vertex $y \in Y$, comprising a special pair (x, y). Let $R \in \mathbb{G}(n, c/n)$. Then there are a.a.s. gn vertex-disjoint special paths of the form $xv_1 \dots v_{k-1}y$ ($v_i \in V_i$) in $G \cup R$.*

Proof. It suffices to show that we can a.a.s. find εn such disjoint paths, for any small $\varepsilon = \varepsilon(\delta) > 0$. This is because we can break R into phases $R \supseteq R_1 \cup \cdots \cup R_r$, where r is chosen so that $(1-\varepsilon)^r \leq (1-g)$. Each phase we can cover an ε-fraction of the vertices with suitable k-paths; discard these paths for the next phase and repeat.

By Lemma 4 we can find a matching of $(1 - 1/(2(k-2)))n$ edges between each V_i and V_{i+1} ($1 \leq i < k-1$). This gives us a disjoint union of $n/2$ paths $P_1, \dots, P_{\gamma n}$ in $G \cup R$, where $P_i = v_1^i \dots v_{k-1}^i$ (with $v_j^i \in V_j$). For each j, we can assume that $v_j^1, \dots, v_j^{n/2}$ is a uniformly random ordering of a uniformly random set of $n/2$ elements of V_j,which we can complete into a uniformly random ordering $v_j^1, \dots, v_j^n$ (independently for $2 \leq j \leq k-1$).

Now, let the special pairs be $\big(x^1, y^1\big), \dots, (x^n, y^n)$. Let S_X (respectively S_Y) be the set of indices $i \leq n/2$ such that v_1^i is adjacent to x^i (respectively, v_{k-1}^i is adjacent to y^i). By linearity of expectation, $\mathbb{E}|S_X|, \mathbb{E}|S_Y| \geq \delta n/2$. If we vary the ordering $v_j^1, \dots, v_j^n$ by a transposition, this changes $|S_X|$ by at most 2, so by a bounded differences inequality (for example [7, Lemma B.1]), a.a.s. $|S_X|, |S_Y| \leq \delta n/3$. Conditioning on S_X and $|S_Y|$, a concentration inequality for the hypergeometric distribution shows that a.a.s. $|S_X \cap S_Y| \geq \delta^2 n/10$. The indices in $S_X \cap S_Y$ correspond to εn special paths $x_i, v_1^i, \dots, v_{k-1}^i, y_i$, where $\varepsilon = \delta^2/10$. $\square$

Proof of Lemma 14. Let $V_0 = X$ and $V_k = Y$. For each V_i, set aside a uniformly random subset D_i of size $(|V_i| - t_i m)/2$. Divide each $V_i \backslash D_i$ ($1 \leq i \leq k-1$) randomly into t_i equal-sized pieces and order the resulting $k-1$ clusters as $W_1, \dots, W_{k-1}$ in such a way that $W_1 \subseteq V_1$ and $W_{k-1} \subseteq V_r$. By Lemma 13, (X, W_1) and (W_{k-1}, Y) are $(\varepsilon, \delta/2)$-super-regular. Each $|W_i| \geq n/(2(k-1))$ and also note

$$\begin{aligned}
|V_i \backslash D_i| &= |V_i|/2 + t_i m/2, \\
m &\leq f|V_i|/t_i \\
&\leq 2f|V_i \backslash D_i|/t_i - fm, \\
(1+f)m &\leq 2f|V_i \backslash D_i|/t_i.
\end{aligned}$$

Let $g = 2f/(1+f)$ so each $g|W_i| \geq m$. Apply Lemma 15 to find m special paths in $G \cup R$ with the required intersection with each V_i. After removing these paths, there are $2|D_i|$ vertices left in each V_i. If (V_i, W) was $(\varepsilon, \delta/2)$-super-regular before the deletions (where W does not participate in our special paths), it is trivially still (ε, δ)-dense after the deletions. Moreover, by Lemma 13, (D_i, W) is a.a.s. $(\varepsilon, \delta/2)$-super-regular, so that that (V_i, W) is now $(\varepsilon, \delta/4)$-super-regular, as required. By a similar argument, each (V_i, V_j) that was $(\varepsilon, \delta/2)$-super-regular before the deletions is now $(\varepsilon, \delta/4)$-super-regular after the deletions. $\square$

Now we return to the proof of Theorem 1. For each $\mathbf{i}, \mathbf{j}$ with $\mathbf{i} \neq \mathbf{j}$, and some very small $\xi = \xi(a, \delta) > 0$, we want to find $(1-\xi)s_{\mathbf{i}}s_{\mathbf{j}}n/(k+1)$ special paths connecting special pairs in $Z_{\mathbf{i},\mathbf{j}}$, in such a way that each of these special paths has one interior vertex from $W_{\bar{\mathbf{i}}}$, one interior vertex from $W_{\bar{\mathbf{j}}}$ and the other $k-3$ interior vertices from $W_{\mathbf{i}}$. Provided k is large we can find our desired paths with at most $(2q)^2$ applications of Lemma 14 (splitting R_2 into $(2q)^2$ phases).

Remove these paths, and update each $V_{\mathbf{i}}, W_{\mathbf{i}}, X_{\mathbf{i}}, Y_{\mathbf{i}} X_{\mathbf{i},\mathbf{j}}, Y_{\mathbf{i},\mathbf{j}}, Z_{\mathbf{i},\mathbf{j}}$ accordingly. (but we do not update n or the $s_{\mathbf{i}}$). Each $\left(X_{\mathbf{i},\mathbf{j}}, W_{\bar{\mathbf{i}}}\right)$, $\left(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$ and $\left(W_{\mathbf{i}}, Y_{\mathbf{j},\bar{\mathbf{i}}}\right)$ are (ε, δ')-regular, where $\delta' = (\delta/2)/4^{(2q)^2}$. Let $a' = a'(a)$ satisfy $|W_{\mathbf{i}}| \geq a'n$. (for large k, $a' = a^2/2$ will do).

Now, for every $\mathbf{i}, \mathbf{j}$ with $\mathbf{i} \neq \mathbf{j}$, we have $|Z_{\mathbf{i},\mathbf{j}}| = \xi s_{\mathbf{i}} s_{\mathbf{j}} n/(k+1)$. The special pairs in all such $Z_{\mathbf{i},\mathbf{j}}$ comprise a total of at most $\xi n/(k+1)$ "bad" special pairs. For small ξ this is less than $\delta' a' n/(2(k-1))$ special pairs, which we deal with greedily, as follows. For some such special pair (x, y) with $x \in X_{\mathbf{i}}$ and $y \in Y_{\mathbf{j}}$, choose a length-$(k-3)$ path $xw_1, \ldots, w_{k-3}$, where the w_ℓ alternate between $W_{\mathbf{i}}$ and $W_{\bar{\mathbf{i}}}$ with $w_{k-3} \in W_{\mathbf{i}}$ (recall, k is odd). Then choose a set $U_x \subseteq W_{\bar{\mathbf{i}}}$ of $\delta' a' n/4$ neighbours of w_{k-3} and a disjoint set $U_y \subseteq W_{\bar{\mathbf{j}}}$ of $\delta' a' n/4$ neighbours of y. Now, note that (if c_3 is large), there is a.a.s. an edge in R_3 between any pair of disjoint sets of size $\delta' a' n/4$ (the proof of this is basically identical to the proof of Lemma 4). Such an edge allows us to complete a special path from x to y, which we remove. We can repeat this for each unconnected bad special pair. Note that these special paths comprise less than $\delta' a' n/2$ vertices total, which is so few that each $\left(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$ is $(\varepsilon, \delta'/2)$-super-regular, and each $x \in X_{\mathbf{i}}$ (respectively $y \in Y_{\mathbf{i}}$) has at least $\delta a' n/2$ neighbours in $W_{\bar{\mathbf{i}}}$, throughout the whole process.

Now, update all the relevant variables. We are left with a set of clusters $X_{\mathbf{i}}, W_{\mathbf{i}}, Y_{\mathbf{i}}$ partitioning the vertices of $G \cup R$, such that each $\left(X_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$, $\left(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$ and $\left(W_{\mathbf{i}}, Y_{\bar{\mathbf{i}}}\right)$ is $(\varepsilon, \delta'/2)$-regular, and each special pair is in some $X_{\mathbf{i}} \times Y_{\bar{\mathbf{i}}}$. Moreover, with X (respectively Y) being the set of x-type (respectively y-type) vertices remaining, and W being the set of non-special vertices remaining, we have $(k-1)|X| = (k-1)|Y| = |W|$. For large k each $an \leq |X_{\mathbf{i}}|, |Y_{\mathbf{i}}|, |W_{\mathbf{i}}| \leq bn$ for some a, b, and each $|X_{\mathbf{i}}|, |Y_{\mathbf{i}}| \leq |W_{\mathbf{j}}|$.

2.2.4 Adjusting the relative sizes of the clusters

The next step is to remove paths such that in the graph that remains, we have $|W_{\mathbf{i}}| = (k-1)|X_{\mathbf{i}} \cup Y_{\mathbf{i}}|/2$ for each $\mathbf{i}$. This is easily achievable through repeated application of Lemma 14, but is a bit fiddly to explain.

Suppose $(k-1)|X_{\mathbf{i}} \cup Y_{\mathbf{i}}|/2 + m \geq |W_{\mathbf{i}}|$ and $(k-1)|X_{\mathbf{j}} \cup Y_{\mathbf{j}}|/2 + m \leq |W_{\mathbf{j}}|$, where one of these inequalities is an equality. Use Lemma 14 and R_4 to find $\lfloor m/((k-1)/2 - 1) \rfloor$ special paths connecting special pairs in $X_{\mathbf{i}} \times Y_{\bar{\mathbf{i}}}$, each with $(k-1)/2$ interior vertices in $W_{\bar{\mathbf{i}}}$, one interior vertex in $W_{\mathbf{i}}$, and the remaining $(k-1)/2 - 1$ interior vertices in $W_{\mathbf{j}}$.

Then, let

$$r = ((k-1)/2 - 1)\left\langle \frac{m}{(k-1)/2 - 1} \right\rangle,$$

where $\langle \cdot \rangle$ denotes the fractional part of a real number. With another application of Lemma 14, find a single special path connecting a special pair in $X_{\mathbf{i}} \times Y_{\bar{\mathbf{i}}}$, which has $(k-1)/2$ interior vertices in $W_{\bar{\mathbf{i}}}$, has $(k-1)/2 - 1 - r$ interior vertices in $W_{\mathbf{i}}$, and has r interior vertices in $W_{\mathbf{j}}$. We now have either $(k-1)|X_{\mathbf{i}} \cup Y_{\mathbf{i}}|/2 = |W_{\mathbf{i}}|$ or $(k-1)|X_{\mathbf{j}} \cup Y_{\mathbf{j}}|/2 \leq |W_{\mathbf{j}}|$. After at most $2q$ iterations of this argument, we will have each $(k-1)|X_{\mathbf{i}} \cup Y_{\mathbf{i}}|/2 = |W_{\mathbf{i}}|$ as desired. Each $\left(X_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$, $\left(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$ and $\left(W_{\mathbf{i}}, Y_{\bar{\mathbf{i}}}\right)$ is now (ε, δ'')-super-regular, where $\delta'' = \delta/4^{2q}$.

2.2.5 Completing the embedding with the blow-up lemma

We have now finished with the edges in R; what remains is sufficiently well-structured that we can embed the remaining paths using the blow-up lemma and the super-regularity in G. The idea is illustrated in Figure 1. We first give a statement of the blow-up lemma.

Lemma 16 (Blow-up Lemma [10]). *Let $\delta, \Delta > 0$ and $r \in \mathbb{N}$. There is $\varepsilon = \varepsilon(r, \delta, \Delta)$ such that the following holds. Let C be a graph on the vertex set $[r]$, and let $n_1, \dots, n_r$ be arbitrary positive integers. Let $V_1, \dots, V_r$ be pairwise disjoint sets of sizes $n_1, \dots, n_r$. We construct two graphs on the vertex set $V = \bigcup_{i=1}^{r} V_i$ as follows. The first graph $b(C)$ (the "complete blow-up") is obtained by putting the complete bipartite graph between V_i and V_j whenever $\{i, j\}$ is an edge in C. The second graph $b_{\varepsilon,\delta}(C)$ (a "super-regular blow-up") is obtained by putting edges between each such V_i and V_j such that (V_i, V_j) is an (ε, δ)-super-regular pair. If a graph H with maximum degree bounded by Δ can be embedded into $b(C)$, then it can be embedded into $b_{\varepsilon,\delta}(C)$.*

Now we describe the setup to apply the blow-up lemma. Uniformly at random divide each $W_{\mathbf{i}}$ into sets $A_{\mathbf{i}}$ and $B_{\mathbf{i}}$ of size $(k-1)|X_{\mathbf{i}}|/2$ and $(k-1)|Y_{\mathbf{i}}|/2$ respectively. For $\mathbf{i} = (i, h)$, let $S^0_{\mathbf{i}} = X_{\mathbf{i}}$, $S^1_{\mathbf{i}} = B_{\bar{\mathbf{i}}}$, $S^2_{\mathbf{i}} = A_{\mathbf{i}}$ and $S^3_{\mathbf{i}} = Y_{\bar{\mathbf{i}}}$, so that each $\left(S^0_{\mathbf{i}}, S^2_{\mathbf{i}}, S^3_{\mathbf{i}}, S^3_{\mathbf{i}}\right)$ is a sequence of clusters of sizes $n_{\mathbf{i}}, (k-1)n_{\mathbf{i}}/2, (k-1)n_{\mathbf{i}}/2, n_{\mathbf{i}}$ for some $n_{\mathbf{i}}$, and each $\left(S^{\ell}_{\mathbf{i}}, S^{\ell+1}_{\mathbf{i}}\right)$ is $(\varepsilon, \delta''/2)$-super-regular.

For each $\mathbf{i}$, we construct an auxiliary graph $G_{\mathbf{i}}$ as follows. Start with the subgraph of G induced by $S^0_{\mathbf{i}} \cup S^2_{\mathbf{i}} \cup S^3_{\mathbf{i}} \cup S^3_{\mathbf{i}}$, and remove every edge not between some $S^{\ell}_{\mathbf{i}}$ and $S^{\ell+1}_{\mathbf{i}}$. Contract each special pair to a single vertex, to obtain a super-regular "cluster-cycle" with 3 clusters $S^*_{\mathbf{i}}, S^1_{\mathbf{i}}, S^2_{\mathbf{i}}$ of sizes $n_{\mathbf{i}}, (k-1)n_{\mathbf{i}}/2, (k-1)n_{\mathbf{i}}/2$. This is a (ε, δ'')-super-regular blow-up of a 3-cycle C_3.

The complete blow-up $b(C_3)$ contains $n_{\mathbf{i}}$ disjoint k-cycles (each has a vertex in $S^*_{\mathbf{i}}$ and its other $k-1$ vertices alternate between $S^1_{\mathbf{i}}$ and $S^2_{\mathbf{i}}$). By the blow-up lemma, for small ε the (ε, δ'')-super-regular blow-up $G_{\mathbf{i}}$ of C_3 also contains $n_{\mathbf{i}}$ disjoint k-cycles. Since k is odd, each of these must use at least one vertex from $S^*_{\mathbf{i}}$, but since $|S^*_{\mathbf{i}}| = n_{\mathbf{i}}$, each cycle must then use exactly one vertex from $S^*_{\mathbf{i}}$. Each of these cycles corresponds to a special path running through $S^0_{\mathbf{i}}, S^2_{\mathbf{i}}, S^3_{\mathbf{i}}, S^3_{\mathbf{i}}$, and these special paths complete our embedding of T.

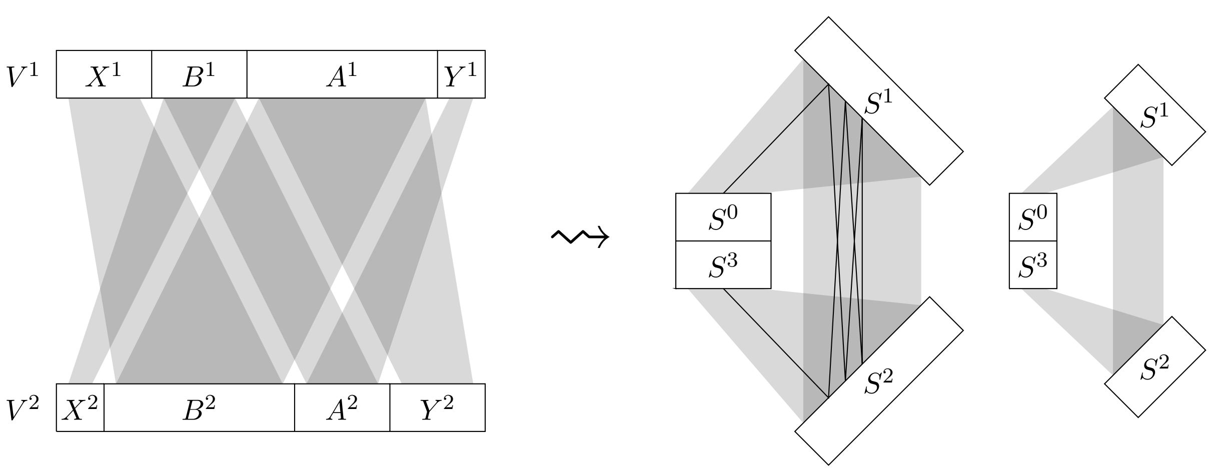

Figure 1. From super-regular pairs to cluster-cycles. An example special path is shown, for $k = 7$.

3 Concluding Remarks

We have proved that any given bounded-degree spanning tree typically appears when a linear number of random edges are added to a dense graph. In the process, we have proved a general lemma concerning the partition of an arbitrary dense graph into uneven superregular pairs.

There are a few interesting questions that remain open. Most prominent is the question of embedding more general kinds of spanning subgraphs into randomly perturbed graphs. It would be particularly interesting if the general result of [5] (concerning arbitrary spanning graphs with bounded degree and low bandwidth) could be adapted to this setting.

It is also possible that our use of Szemerédi's regularity lemma could be avoided, thus drastically improving the constants $c(\alpha)$ and perhaps allowing us to say something about random perturbations of graphs which have slightly sublinear minimum degree.

Acknowledgements: similar stuff (concerning Tel Aviv and ETH) to the last paper?

References

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

[2] T. Bohman, A. Frieze, M. Krivelevich, and R. Martin, *Adding random edges to dense graphs*, Random Structures & Algorithms **24** (2004), no. 2, 105–117.

[3] 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.

[4] J. Böttcher, M. Schacht, and A. Taraz, *Spanning 3-colourable subgraphs of small bandwidth in dense graphs*, Journal of Combinatorial Theory, Series B **98** (2008), no. 4, 752–777.

[5] ———, *Proof of the bandwidth conjecture of Bollobás and Komlós*, Mathematische Annalen **343** (2009), no. 1, 175–205.

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

[7] A. Frieze and B. Pittel, *Perfect matchings in random graphs with prescribed minimal degree*, Mathematics and Computer Science III: Algorithms, Trees, Combinatorics and Probabilities (M. Drmota, P. Flajolet, D. Gardy, and B. Gittenberger, eds.), Trends in Mathematics, Birkhäuser, Basel, 2004, pp. 95–132.

[8] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Cambridge University Press, 2000.

[9] J. Komlós, G. N. Sárközy, and E. Szemerédi, *Proof of a packing conjecture of bollobás*, Combinatorics, Probability and Computing **4** (1995), no. 03, 241–255.

[10] ———, *Blow-up lemma*, Combinatorica **17** (1997), no. 1, 109–123.

[11] ———, *Spanning trees in dense graphs*, Combinatorics, Probability and Computing **10** (2001), 397–416.

[12] M. Krivelevich, *Embedding spanning trees in random graphs*, SIAM Journal on Discrete Mathematics **24** (2010), no. 4, 1495–1500.

[13] M. Krivelevich, M. Kwan, and B. Sudakov, *Cycles and matchings in randomly perturbed digraphs and hypergraphs*, arXiv preprint arXiv:1501.04816 (2015).

[14] M. Krivelevich, B. Sudakov, and P. Tetali, *On smoothed analysis in dense graphs and formulas*, Random Structures & Algorithms **29** (2006), no. 2, 180–193.

[15] D. Kühn, D. Osthus, and A. Taraz, *Large planar subgraphs in dense graphs*, Journal of Combinatorial Theory, Series B **95** (2005), no. 2, 263–282.

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

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

[18] D. A. Spielman and S.-H. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, Journal of the ACM **51** (2004), no. 3, 385–463.