

# Bounded-degree spanning trees in randomly perturbed graphs

Michael Krivelevich, Matthew Kwan, Benny Sudakov

## 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 an  $n$ -vertex graph with minimum degree at least  $\alpha n$ , and  $T$  is an  $n$ -vertex tree with maximum degree at most  $\Delta$ , then if we add  $cn$  uniformly random edges to  $G$ , the resulting graph will contain  $T$  asymptotically almost surely (as  $n \rightarrow \infty$ ). 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  $\Delta$  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  $\Delta$ . (This has also been generalized further in two directions: to allow  $\Delta$  to grow with  $n$  and to allow for much more general spanning subgraphs than spanning trees; see [11, 5]).

The constant “1/2” in the above Dirac-type theorems is tight: in order to guarantee the existence of these spanning subgraphs we require very strong density conditions. But the situation 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  $\alpha n$ . 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  $\Delta$  and any tree  $T$  with maximum degree at most  $\Delta$ , 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  $n \rightarrow \infty$ , 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. The model has since been studied in a number

of other contexts; see for example [2, 14, 13]. 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). We also note that a particularly important motivation is the notion of *smoothed analysis* of algorithms introduced in [18]. This 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.

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  $\alpha n$  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 for example the complete bipartite graph with partition sizes  $n/3$  and  $2n/3$ ), so the order of magnitude of this result is tight.

Let  $\mathbb{G}(n, p)$  be the Erdős-Rényi random graph model with vertex set  $[n] = \{1, \dots, n\}$ , 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  $\alpha n$ , and  $T$  is an  $n$ -vertex tree with maximum degree at most  $\Delta$ , and  $R \in \mathbb{G}(n, c/n)$ , then a.a.s.  $T \subseteq G \cup R$ .*

One of the ingredients of the proof is the following lemma, due to Alon and two of the authors ([1, Theorem 1.1]). It shows that the random edges in  $R$  are already enough to embed almost-spanning trees.

**Lemma 2.** *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  $\Delta$  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 later (these  $c_i$  will in turn determine  $c$ ). Let  $V = [n]$  be the common vertex set of  $G$ ,  $R$  and the  $R_i$ .

### 2.1 Case 1: $T$ has many leaves

For our first case suppose there are at least  $\lambda n$  leaves in  $T$  (for some fixed  $\lambda = \lambda(\alpha) > 0$  to be determined in the second case). Then consider the tree  $T'$  with some  $\lambda n$  leaves removed. By Lemma 2 we can a.a.s. embed  $T'$  into  $R_1$ . Let  $A' \subseteq V$  be the set of vertices of  $T'$  which had a leaf deleted from them, and let  $B \subseteq V$  be the set of  $\lambda n$  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  $a \in A'$ , the random variable  $|N_G(a) \cap B|$  is hypergeometrically distributed with expected value  $\lambda \alpha n$ , and similarly each  $|N_G(b) \cap A'|$  ( $b \in B$ ) is hypergeometric with expectation 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, in  $G$  each  $a' \in A'$  a.a.s. has at least  $\beta n$  neighbours in  $B$ , and each  $b \in B$  a.a.s. has at least  $\beta n$  neighbours in  $A'$ . From now on we treat  $A'$  and  $B$  as fixed sets satisfying these properties.

Let  $A$  be a set consisting of  $\ell$  copies of each vertex  $a' \in A'$  which needs  $\ell$  extra leaves (so  $|A| = \lambda n$ ). Define a bipartite graph  $\Gamma_{A,B}(G)$  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  $G$ . Similarly define  $\Gamma_{A,B}(R_2)$  and  $\Gamma_{A,B}(G \cup R)$  with edges deriving from  $R_2$  and  $G \cup R$ . If we can find a perfect matching in  $\Gamma_{A,B}(G) \cup \Gamma_{A,B}(R_2) \subseteq \Gamma_{A,B}(G \cup R)$ , 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 bipartite graph distribution 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 for any  $a' \in A$  and  $b \in B$ , and for large  $n$ ,

$$\begin{aligned} \Pr(\text{all copies of } a' \text{ are adjacent to } b \text{ in } \Gamma_{A,B}(R_2)) &= c_2/n \\ &\geq 1 - (1 - c_{\mathbb{B}}/n)^{\Delta} = \Pr(\text{some copy of } a' \text{ is adjacent to } b \text{ in } \mathbb{B}_{A,B}(c_{\mathbb{B}}/n)). \end{aligned}$$

This means we can couple  $\Gamma_{A,B}(R_2)$  and  $G_{\mathbb{B}} \in \mathbb{B}_{A,B}(c_{\mathbb{B}}/n)$  in such a way that  $\Gamma_{A,B}(R_2) \supseteq G_{\mathbb{B}}$  (the distribution of  $\Gamma_{A,B}(R_2)$  *stochastically dominates*  $\mathbb{B}_{A,B}(c_{\mathbb{B}}/n)$ ). The following lemma then 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  $\alpha n$ , 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  $\varepsilon n$ , the probability there is no edge between  $A'$  and  $B'$  is  $(1 - c/n)^{(\varepsilon n)^2} \leq e^{-c\varepsilon^2 n}$ . 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  $|W| \leq \alpha n$  then  $|N_G(W)| \geq \alpha n \geq |W|$  by the minimum degree condition on  $G$ . Similarly, every vertex outside  $N_G(W)$  has at least  $\alpha n$  neighbours outside  $W$ , so if  $|W| \geq (1 - \alpha)n$  then  $|N_G(W)| = n \geq |W|$ . So, consider any  $W$  with  $\alpha n \leq |W| \leq (1 - \alpha)n$  (in which case we must have  $\alpha \leq 1/2$ ). 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  $(1 - \alpha/8)n$ -edge matching  $M_i$  in each  $R_i$ . Let  $N_i = N_{M_i}(W)$  contain the vertices matched with  $W$  in each  $M_i$ , and note that each  $N_i$  is a random set satisfying  $|W| - \alpha n/8 \leq |N_i| \leq |W|$ . 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  $|N_i||N_1|/n$ . Since  $|W|(n - |W|) \geq \alpha n(1 - \alpha)n \geq \alpha n^2/2$ , we have  $|W|^2 \leq |W|n - \alpha n^2/2$  and

$$|N_i||N_1|/n \leq |W|^2/n \leq |W| - \alpha n/2 = 2(|W| - \alpha/8) - |W| - \alpha/4 \leq |N_i| + |N_1| - |W| - \alpha n/4.$$

So, by a concentration inequality,

$$\Pr(|N_i \cup N_1| < |W|) = \Pr(|N_i \cap N_1| > |N_i| + |N_1| - |W|) = e^{-\Omega(\alpha^2 n)}$$

and therefore

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

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  $|N_R(W)| \geq |W|$  for all required  $W$ .  $\square$

## 2.2 Case 2: $T$ has few leaves

Now we address the second case where there are fewer than  $\lambda n$  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  $\varepsilon$ -*regular* if for any  $U_1, U_2$  with  $U_h \subseteq V_h$  and  $|U_h| \geq \varepsilon|V_h|$ , we have  $|d(U_1, U_2) - d(V_1, V_2)| \leq \varepsilon$ . If alternatively  $d(U_1, U_2) \geq \delta$  for all such pairs  $U_1, U_2$  then we say  $(V_1, V_2)$  is  $(\varepsilon, \delta)$ -*dense*. Let  $\bar{h} = 2 - h$ ; if  $(V_1, V_2)$  is  $(\varepsilon, \delta)$ -dense and moreover each  $v \in V_h$  has at least  $\delta|V_{\bar{h}}|$  neighbours in  $V_{\bar{h}}$ , then we say  $(V_1, V_2)$  is  $(\varepsilon, \delta)$ -*super-regular*.

**Lemma 6.** *For  $\alpha, \varepsilon > 0$  with  $\varepsilon$  sufficiently small relative to  $\alpha$ , 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  $\alpha n$ . We can choose  $q \leq M$  and partition  $[n]$  into sets  $V_i^h$  ( $1 \leq i \leq q$ ,  $h = 1, 2$ ) such that each  $|V_i^h| \leq bn$  and each pair  $(V_i^1, V_i^2)$  is  $(\varepsilon, \delta)$ -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  $V_i^1$  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  $\alpha$ , 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  $\varepsilon n$ , 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  $\varepsilon$ -regular in  $G'$  with density zero or at least  $\alpha'$ . 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 Szemerédi’s regularity lemma follows directly from [15, Proposition 9].

Before we proceed to the proof of Lemma 6, we give some simple lemmas about  $(\varepsilon, \delta)$ -denseness. Note that  $(\varepsilon, \delta)$ -denseness is basically a one-sided version of  $\varepsilon$ -regularity that is more convenient in proofs about super-regularity. Here and in later sections, most of the theorems about  $(\varepsilon, \delta)$ -dense pairs correspond to analogous theorems for  $\varepsilon$ -regular pairs with density about  $\delta$ .

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

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

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

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  $\alpha n$ . Then there is a spanning subgraph  $Q$  which is a union of vertex-disjoint stars, each with at least two and at most  $1 + 1/\alpha$  vertices.*

*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  $1 + 1/\alpha$  vertices, then we could add  $v$  to that star, contradicting maximality. (Here we allow either vertex of a 2-vertex star to be considered the “centre”). 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  $\alpha n$ ) neighbours of  $v$  is a centre of a star with  $1 + \lfloor 1/\alpha \rfloor > 1/\alpha$  vertices. But these stars would comprise more than  $n$  vertices, which is again a contradiction. We conclude that  $Q$  covers  $G$ , as desired.  $\square$

*Proof of Lemma 6.* Apply our minimum degree form of Szemerédi’s regularity lemma, with some  $\varepsilon'$  to be determined (which we will repeatedly assume is sufficiently small). Let  $Q$  be a cover of  $C$  by stars  $T_1, \dots, T_q$  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  $(W_i^1, W_i^2)$  is  $(\varepsilon', \delta')$ -dense in  $G'$  (therefore  $G$ ).

Apply Lemma 10 to obtain sets  $V_i^h \subseteq W_i^h$  such that  $|V_i^h| = (1 - \varepsilon')|W_i^h|$  and each  $(V_i^1, V_i^2)$  is  $(2\varepsilon', \delta'')$ -super-regular, for  $\delta'' = \delta'/2$ . Combining the exceptional cluster  $V_0$  and all the  $W_i^h \setminus V_i^h$ , there are at most  $2\varepsilon'n$  “bad” vertices not part of a super-regular pair.

Each bad vertex  $v$  has at least  $\alpha'n - 2\varepsilon'n$  neighbours to the clusters  $V_i^h$ . The clusters  $V_i^h$  which have fewer than  $\delta''|V_i^h|$  such neighbours contain at most  $\delta''n$  of these neighbours. Each  $V_i^h$  has size at most  $sn/\alpha'$ , so for small  $\varepsilon'$  there are at least

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

clusters  $V_i^h$  which have at least  $\delta''s(1 - \varepsilon')n$  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  $2(2\varepsilon'n)/((\alpha')^2/(2s)) \leq (8\varepsilon'/(\alpha')^2)|W_i^h| \leq (9\varepsilon'/(\alpha')^2)|V_i^h|$  bad vertices have been added to each  $V_i^h$ . By Lemma 9, for small  $\varepsilon'$  each  $(V_i^1, V_i^2)$  is now  $(O(\sqrt{\varepsilon'}), \delta''/2)$ -super-regular. With  $\delta = \delta''/2$ ,  $a = s/2$  and  $b = 2s/\alpha'$ , and  $\varepsilon'$  small relative to  $\varepsilon$ , we are done.  $\square$

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 $T$

Recall that a *bare path* in  $T$  is a path  $P$  such that every vertex of  $P$  has degree exactly two in  $T$ . 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  $\ell$  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  $\lambda$  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  $\psi n$  “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

$$X \cup Y \cup F \setminus (X \cup Y) \cup V \setminus F$$

is a uniformly random partition of  $V$  into parts of sizes  $\psi n$ ,  $\psi n$ ,  $n/2 - 2\psi n$ ,  $n/2$ , and that the special pairs correspond to a uniformly random bijection between  $X$  and  $Y$ . Let  $X_i = V_i \cap X$  and  $Y_i = V_i \cap Y$ , and let  $Z_{i,j}$  be the set of special pairs in  $X_i \times Y_j$ . Let  $X_{i,j} \subseteq X_i$  and  $Y_{i,j} \subseteq Y_j$  be the set of  $x$ -type and  $y$ -type vertices in pairs in  $Z_{i,j}$ , respectively. Also let  $W_i = V_i \setminus F$  be the set of “free” vertices remaining in  $V_i$ .

By a concentration inequality for the hypergeometric distribution and the union bound, a.a.s. each  $|W_i| \sim s_i n/2$ . Similarly each  $|X_i| \sim |Y_i| \sim \psi s_i n$ , and each  $|X_{i,j}| = |Y_{i,j}| = |Z_{i,j}| \sim \psi s_i s_j n$ . If we condition on the sizes of each  $W_i$ ,  $X_{i,j}$ ,  $Y_{j,i}$ , each is a uniformly random subset of  $V_i$  of its size.

Here and in future parts of the proof, it is critical that these kinds of random subsets of our super-regular pairs maintain super-regularity (albeit with slightly weaker  $\varepsilon$  and  $\delta$ ). We will ensure this by appealing to the following lemma.

**Lemma 13.** *Fix  $\delta, \varepsilon > 0$ . Suppose  $(V^1, V^2)$  is  $(\varepsilon, \delta)$ -super-regular, let  $n_h$  satisfy  $|V^h| \geq n_h = \Omega(n)$  (the implied constant can depend on  $\varepsilon$  and  $\delta$ ) and for each  $h$  let  $W^h \subseteq V^h$  be a uniformly random subset of size  $n_h$ . Then  $(W^1, W^2)$  is a.a.s.  $(2\sqrt{\varepsilon}, \delta/2)$ -super-regular.*

To prove Lemma 13 we will apply a bipartite version of Szemerédi’s regularity lemma on the edges between  $V^1$  and  $V^2$ . The following formulation can be found in [19].

**Lemma 14** (Szemerédi’s regularity lemma, bipartite version). *For any  $\varepsilon > 0$ , there is  $K = K(\varepsilon)$  such that the following holds. For any bipartite graph  $G$  with bipartition  $V^1 \cup V^2$ , there are partitions of each  $V^h$  into clusters  $V_0^h, V_1^h, \dots, V_k^h$  ( $k \leq K$ ), such that the “exceptional clusters”  $V_0^h$  has size at most  $\varepsilon |V^h|$ , the other clusters  $V_i^h$  in each part have equal size  $s_h |V^h|$ , and all but  $\varepsilon k_1 k_2$  of the pairs of non-exceptional clusters are  $\varepsilon$ -regular.*

We will also need a linear relaxation of the notion of an  $(\varepsilon, \delta)$ -dense pair.

**Lemma 15.** *If  $(V^1, V^2)$  is  $(\varepsilon, \delta)$ -dense then for any functions  $\pi^h : V^h \rightarrow [0, 1]$ , we have*

$$\sum_{v \in V^1} \sum_{u \in V^2} \pi^1(v) \pi^2(u) \mathbf{1}_{v \sim u} \geq \delta \sum_{v \in V^1} \pi^1(v) \sum_{u \in V^2} \pi^2(u) - \varepsilon \delta |V^1| |V^2|.$$

*Proof.* Let  $f : [0, 1]^{V^1 \times V^2} \rightarrow \mathbb{R}$  be defined by

$$f(\pi) = \sum_{v \in V^1} \sum_{u \in V^2} \pi^1(v) \pi^2(u) \mathbf{1}_{v \sim u} - \delta \sum_{v \in V^1} \pi^1(v) \sum_{u \in V^2} \pi^2(u)$$

Note that  $f$  is linear in each  $\pi^h$  so it attains its extrema at vertices of the hypercube  $[0, 1]^{V^1 \times V^2}$ . But if each  $\pi^h(v) \in [0, 1]$  then  $\pi^1$  and  $\pi^2$  just correspond to subsets  $U^h \subseteq V^h$ . We just need to show that  $f(\pi) \geq -\varepsilon\delta|V^1||V^2|$  for each such  $\pi$ . If each  $|U^h| \geq \varepsilon|V^h|$  then  $(\varepsilon, \delta)$ -density implies  $f(\pi) = \varepsilon|U^1||U^2| - \delta|U^1||U^2| \geq 0$ . Otherwise  $|U^1||U^2| \leq \varepsilon|V^1||V^2|$  so trivially  $f(\pi) \geq -\varepsilon\delta|V^1||V^2|$ .  $\square$

*Proof of Lemma 13.* First, note that for each  $v \in W^h$ , the number of neighbours of  $v \in W^h$  is hypergeometrically distributed, with expected value at least  $\delta n_h$ . Since  $n_h = \Omega(n)$ , a concentration inequality and the union bound immediately tells us that a.a.s. each such  $v$  has  $\delta n_h/2$  such neighbours.

It remains to show that  $(W^1, W^2)$  is a.a.s.  $(2\sqrt{\varepsilon}, \delta/2)$ -dense. Apply Lemma 14 to the bipartite graph between  $V^1$  and  $V^2$ , with some regularity parameter  $\varepsilon'$  to be determined. By a concentration inequality, a.a.s. for all  $i \geq 1$  we have  $|W^h \cap V_i^h| \leq (1 + \gamma)s_h n_h$  for some  $\gamma \rightarrow 0$ . Now, consider any  $U^h \subseteq W^h$  with  $|U^h| \geq 2\sqrt{\varepsilon}n_h$ , and let  $U_i^h = U^h \cap V_i^h$ . Let  $I^h = \{i \geq 1 : |U_i^h| \geq \varepsilon'|V_i^h|\}$  be the set of cluster indices with non-negligible intersection with  $U^h$ . We have

$$\begin{aligned} e(U^1, U^2) &\geq \sum_{i \in I^1} \sum_{j \in I^2} d(U_i^1, U_j^2) |U_i^1| |U_j^2| \\ &\geq \sum_{i \in I^1} \sum_{j \in I^2} (d(V_i^1, V_j^2) - \varepsilon') |U_i^1| |U_j^2| - k^2 \varepsilon' ((1 + \gamma)s_1 n_1) ((1 + \gamma)s_2 n_2) \\ &\geq \sum_{i \in I^1} \sum_{j \in I^2} d(V_i^1, V_j^2) |U_i^1| |U_j^2| - \varepsilon' |U^1| |U^2| - \varepsilon' n_1 n_2 - o(n^2) \\ &\geq \sum_{i=1}^k \sum_{j=1}^k d(V_i^1, V_j^2) |U_i^1| |U_j^2| - \varepsilon' |V^1| |V^2| - \varepsilon' |U^1| |U^2| - \varepsilon' n_1 n_2 - o(n^2). \end{aligned}$$

Now, we define  $\pi^1, \pi^2$  as follows. If  $v$  is in some exceptional cluster  $V_0^h$  then let  $\pi^h(v) = 0$ . Otherwise if  $v \in V_i^h$  let  $\pi^h(v) = |U_i^h| / ((1 + \gamma)s_1 n_1) \in [0, 1]$ . By Lemma 15,

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^k d(V_i^1, V_j^2) |U_i^1| |U_j^2| &\geq \left( \frac{(1 + \gamma)n_1}{|V^1|} \right) \left( \frac{(1 + \gamma)n_2}{|V^2|} \right) \sum_{v \in V^1} \sum_{u \in V^2} \pi^1(v) \pi^2(u) \mathbf{1}_{v \sim u} \\ &\geq \left( \frac{(1 + \gamma)n_1}{|V^1|} \right) \left( \frac{(1 + \gamma)n_2}{|V^2|} \right) \left( \delta \sum_{v \in V^1} \pi^1(v) \sum_{u \in V^2} \pi^2(u) - \varepsilon\delta|V^1||V^2| \right) \\ &\geq \delta(|U^1| - \varepsilon'|V^1|)(|U^2| - \varepsilon'|V^2|) - (1 + \gamma)^2 n_1 n_2 \varepsilon\delta \\ &\geq \delta|U^1||U^2| - 3\varepsilon'|V^1||V^2| - \varepsilon\delta n_1 n_2 - o(n^2). \end{aligned}$$

It follows that

$$e(U^1, U^2) \geq \delta|U^1||U^2| - 6\varepsilon'|V^1||V^2| - (\delta/4)|U^1||U^2| - o(n^2).$$

Let  $\varepsilon' = \varepsilon\delta n_1 n_2 / (6|V^1||V^2|)$ , so that  $e(U^1, U^2) \geq (\delta/2)|U^1||U^2|$  as required.  $\square$

Let  $\overline{(i, h)} = (i, \bar{h})$ . It is an immediate consequence of Lemma 13 that each of  $(X_{i, \bar{j}}, W_{\bar{i}}^-)$ ,  $(W_i, W_{\bar{i}})$  and  $(W_i, Y_{\bar{j}, \bar{i}})$  are  $(\varepsilon', \delta')$ -super-regular, where  $\varepsilon' = 2\sqrt{\varepsilon}$  and  $\delta' = \delta/2$ .



Now that we have embedded  $F$ , we no longer care about the vertices used to embed  $F \setminus (X \cup Y)$ ; 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  $V_i$  all vertices used in  $F \setminus (X \cup Y)$ . Update  $n$  to be the number of vertices not used to embed  $F \setminus (X \cup Y)$  (previously  $(1/2 + 2\psi)n$ ), and update each  $s_i$  to still satisfy  $s_i n = |V_i|$  (this does not change  $s_i$  asymptotically). We now have

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

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

*Let  $G$  be a graph on some vertex set  $[N]$ , together with a vertex partition into clusters. Consider any sequence of clusters  $X, V_1, \dots, V_r, Y$ , and any integer partition  $t_1 + \dots + t_r = k - 1$  (each  $t_i \in \mathbb{N}^+$ ), such that  $|V_i| \geq t_i n$  and  $|X|, |Y| = n$ . Suppose that each  $x \in X$  (respectively,  $y \in Y$ ) has at least  $\delta|V_1|$  neighbours in  $V_1$  (respectively,  $\delta|V_r|$  neighbours in  $V_r$ ).*

*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, in  $G \cup R$  we can find  $(1 - \xi)n$  vertex-disjoint special paths, each with  $t_i$  vertices in each  $V_i$ . These paths depend on  $R$  so are random; in fact we can assume that the vertices used in each  $V_i$ ,  $X$  and  $Y$  are independently uniformly random subsets of their respective sizes.*

*Proof.* Uniformly at random choose  $t_i$  disjoint  $n$ -vertex subsets of each  $V_i$ . This gives a total of  $k - 1$  such subsets; arrange these into a sequence  $W_1, \dots, W_{k-1}$  in such a way that  $W_1 \subseteq V_1$  and  $W_{k-1} \subseteq V_r$ . By concentration and the union bound, for  $\delta_1 = \delta/2$  each  $x \in X$  (respectively,  $y \in Y$ ) has at least  $\delta_1|W_1|$  neighbours in  $W_1$  (respectively,  $\delta_1|W_{k-1}|$  neighbours in  $W_{k-1}$ ). To summarize these minimum degree conditions, we say that the situation is “ $\delta_1$ -good”.

Now, break  $R$  into phases  $R \supseteq R_1 \cup \dots \cup R_r$ , for some large  $r = r(\delta)$  to be determined. By Lemma 4, in  $R_1$  we can find a matching of  $(1 - 1/(2(k - 2)))n$  edges between each  $W_i$  and  $W_{i+1}$  ( $1 \leq i < k - 1$ ). This gives us a disjoint union of  $n/2$  length- $(k - 2)$  paths  $P_1, \dots, P_{n/2}$ ; we write  $P_i = v_1^i \dots v_{k-1}^i$  with  $v_j^i \in W_j$ . 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  $W_j$ , which we can complete into a uniformly random ordering  $v_j^1, \dots, v_j^n$  (independently for  $1 \leq j \leq k - 1$ ). Now, let the special pairs be  $(x^1, y^1), \dots, (x^n, y^n)$ . Let  $I_X$  (respectively  $I_Y$ ) 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,  $\mathbb{E}|I_X|, \mathbb{E}|I_Y| \geq (\delta_1/2)n$ . If we vary the ordering  $v_1^1, \dots, v_1^n$  (respectively  $v_{k-1}^1, \dots, v_{k-1}^n$ ) by a transposition, this changes  $|I_X|$  (respectively  $|I_Y|$ ) by at most 2. So, by a bounded-differences concentration inequality (for example [7, Lemma B.1]), a.a.s. we have  $|I_X|, |I_Y| \geq (\delta_1/3)n$ . Conditioning on  $I_X$  and  $I_Y$ , a concentration inequality for the hypergeometric distribution shows that a.a.s.  $|I_X \cap I_Y| \geq \gamma(\delta_1)n$ , where  $\gamma(\delta_1) = \delta_1^2/10$ . The indices in  $I_X \cap I_Y$  correspond to  $\gamma(\delta_1)n$  special paths  $x^i, v_1^i, \dots, v_{k-1}^i, y^i$ .

Remove these special paths and their vertices in  $W_i$ . Let  $\delta_i = \delta_1/(1 + (i-1)/r)$ ; by concentration and the union bound the situation is still a.a.s.  $\delta_2$ -good after these removals. We can therefore repeat the previous argument using  $R_2$  to find more special paths. After repeating the argument all  $r$  times (before iteration  $i$  we are  $\delta_i$ -good), there are at most

$$n \prod_{i=0}^{r-1} (1 - \gamma(\delta_i)) \leq (1 - \gamma(\delta/4))^r n$$

vertices remaining in  $X$  and in  $Y$ . If we choose large enough  $r$  this is fewer than  $\xi n$  remaining vertices, as required.  $\square$

Now we return to the proof of Theorem 1. For each  $\mathbf{i}, \mathbf{j}$  with  $\mathbf{j} \neq \bar{\mathbf{i}}$ , 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 16. To be precise, as in the proof of Lemma 16, we split  $R_2$  into  $(2q)^2$  phases and use Lemma 16 to remove paths in each phase sequentially. By concentration and the union bound, after each phase we still have a suitable minimum degree condition for this to be possible.

As well as removing these paths, remove their corresponding vertices from  $V_{\mathbf{i}}, W_{\mathbf{i}}, X_{\mathbf{i}}, Y_{\mathbf{i}}X_{\mathbf{i}, \mathbf{j}}, Y_{\mathbf{i}, \mathbf{j}}, Z_{\mathbf{i}, \mathbf{j}}$ . We have removed a uniformly random subset of a certain size from each  $V_{\mathbf{i}}, W_{\mathbf{i}}, X_{\mathbf{i}, \mathbf{j}}, Y_{\mathbf{i}, \mathbf{j}}$ . By Lemma 13, each  $(X_{\mathbf{i}, \mathbf{j}}, W_{\bar{\mathbf{i}}})$ ,  $(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}})$  and  $(W_{\mathbf{i}}, Y_{\mathbf{j}, \bar{\mathbf{i}}})$  are  $(\varepsilon'', \delta'')$ -super-regular, where  $\varepsilon'' = 2\sqrt{\varepsilon'}$  and  $\delta'' = \delta'/2$ . 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  $\mathbf{j} \neq \bar{\mathbf{i}}$ , 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  $\xi$  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  $xw_1, \dots, w_{k-3}$  in  $G$ , where the  $w_\ell$  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  $(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}})$  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  $(X_{\mathbf{i}}, W_{\bar{\mathbf{i}}})$ ,  $(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}})$  and  $(W_{\mathbf{i}}, Y_{\bar{\mathbf{i}}})$  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  $a''n \leq |X_{\mathbf{i}}|, |Y_{\mathbf{i}}|, |W_{\mathbf{i}}| \leq b''n$  for some  $a'', b'' > 0$ , and each  $|X_{\mathbf{i}}|, |Y_{\mathbf{i}}| \leq |W_{\mathbf{i}}|$ .

## 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 16, but is a bit fiddly

to explain.

Suppose  $(k-1)|X_i \cup Y_i|/2 + m \geq |W_i|$  and  $(k-1)|X_j \cup Y_j|/2 + m \leq |W_j|$  for some  $m > 0$ , where one of these inequalities is an equality. Use Lemma 16 and  $R_4$  to find  $\lfloor m/((k-1)/2 - 1) \rfloor$  special paths connecting special pairs in  $X_i \times Y_i$ , each with  $(k-1)/2$  interior vertices in  $W_i$ , one interior vertex in  $W_j$ , and the remaining  $(k-1)/2 - 1$  interior vertices in  $W_j$ .

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 16, find a single special path connecting a special pair in  $X_i \times Y_i$ , which has  $(k-1)/2$  interior vertices in  $W_i$ , has  $(k-1)/2 - r$  interior vertices in  $W_i$ , and has  $r$  interior vertices in  $W_j$ . We now have either  $(k-1)|X_i \cup Y_i|/2 = |W_i|$  or  $(k-1)|X_j \cup Y_j|/2 = |W_j|$ . After at most  $2q$  iterations of this argument, we will have each  $(k-1)|X_i \cup Y_i|/2 = |W_i|$  as desired. Each  $(X_i, W_i)$ ,  $(W_i, W_i)$  and  $(W_i, Y_i)$  is now  $(\varepsilon''', \delta''')$ -super-regular, where  $\varepsilon''' = 2\sqrt{\varepsilon''}$  and  $\delta''' = \delta''/4$ .

### 2.2.5 Completing the embedding with the blow-up lemma

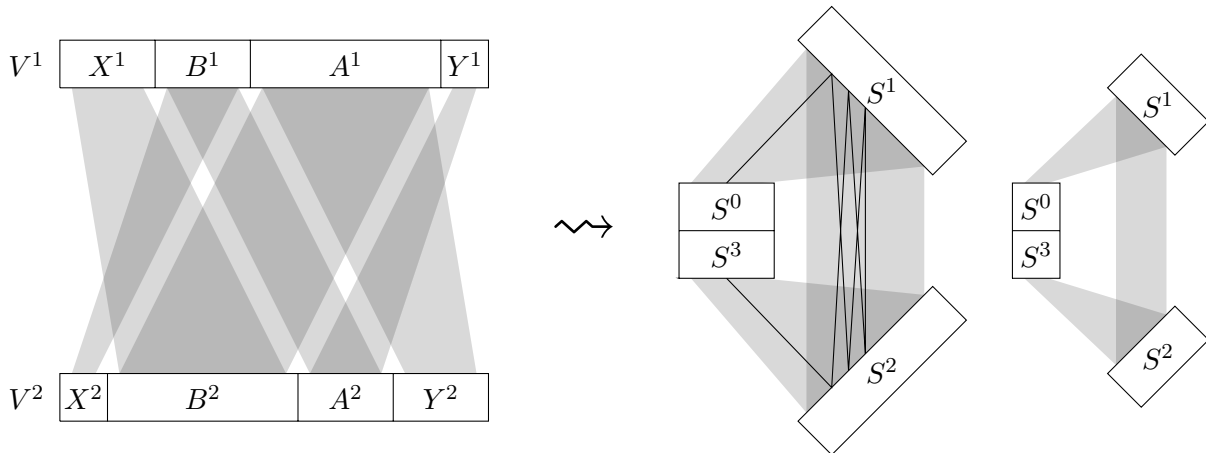
We have now finished with the edges in  $R$ ; what remains is sufficiently well-structured that we can embed the remaining paths using the blow-up lemma and the super-regularity in  $G$ . The idea is illustrated in Figure 1. We first give a statement of the blow-up lemma.

**Lemma 17** (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  $(\varepsilon, \delta)$ -super-regular pair. If a graph  $H$  with maximum degree bounded by  $\Delta$  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  $W_i$  into sets  $A_i$  and  $B_i$  of size  $(k-1)|X_i|/2$  and  $(k-1)|Y_i|/2$  respectively. For  $\mathbf{i} = (i, h)$ , let  $S_i^0 = X_i$ ,  $S_i^1 = B_i$ ,  $S_i^2 = A_i$  and  $S_i^3 = Y_i$ , so that each  $(S_i^0, S_i^1, S_i^2, S_i^3)$  is a sequence of clusters of sizes  $n_i, (k-1)n_i/2, (k-1)n_i/2, n_i$  for some  $n_i$ , and each  $(S_i^\ell, S_i^{\ell+1})$  is  $(\varepsilon, \delta''/2)$ -super-regular.

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

The complete blow-up  $b(C_3)$  contains  $n_i$  disjoint  $k$ -cycles (each has a vertex in  $S_i^*$  and its other  $k-1$  vertices alternate between  $S_i^1$  and  $S_i^2$ ). By the blow-up lemma, for small  $\varepsilon$  the  $(\varepsilon, \delta'')$ -super-regular blow-up  $G_i$  of  $C_3$  also contains  $n_i$  disjoint  $k$ -cycles. Since  $k$  is odd, each of these must use at least one vertex from  $S_i^*$ , but since  $|S_i^*| = n_i$ , each cycle must then use exactly one vertex from  $S_i^*$ . Each of these cycles corresponds to a special path running through  $S_i^0, S_i^1, S_i^2, S_i^3$ , 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. 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.

**Acknowledgement.** Parts of this work were carried out when the first author visited the Institute for Mathematical Research (FIM) of ETH Zurich, and also when the third author visited the School of Mathematical Sciences of Tel Aviv University, Israel. We would like to thank both institutions for their hospitality and for creating a stimulating research environment.

### 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.
- [19] T. Tao, *Szemerédi’s regularity lemma revisited*, Contributions to Discrete Mathematics **1** (2006), no. 1, 8–28.