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 αn, and T is an n-vertex tree with maximum degree at most Δ, then if we add cn uniformly random edges to G, the resulting graph will contain T 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 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 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 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 [18] 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 [17] 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. 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 [19]. 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.

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 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 $[n] = \{1, \ldots, n\}$, 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. $T \subseteq G \cup R$.*

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

At several points in the proof we will prove that (with high probability) there exist certain substructures in the random edges R, and we will assume that these substructures (or some corresponding vertex sets) are themselves uniformly random. A simple way to see why this kind of reasoning is valid, is to notice that if we apply a uniformly random vertex permutation to a random graph $R \in \mathbb{G}(n, p)$, then by symmetry the resulting distribution is still $\mathbb{G}(n, p)$. So we can imagine that we are finding substructures in an auxiliary random graph, then randomly mapping these structures to our vertex set.

2.1 Case 1: T has many leaves

For our first case suppose there are at least λn leaves in T (for some fixed $\lambda = \lambda(\alpha) > 0$ to be determined in the second case). Then consider the tree T' with some λn leaves removed. By Lemma 2 we can 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 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 $a \in A'$, 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 βn neighbours in B, and 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 $|A| = \lambda n$). 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,

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

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 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 $(1-c/n)^{(\varepsilon n)^2} \le e^{-c\varepsilon^2 n}$. 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. □

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 $\alpha n \le |W| \le (1-\alpha)n$ (in which case we must have $\alpha \le 1/2$). 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 $|W| - \alpha n/8 \le |N_i| \le |W|$. 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 $|N_i||N_1|/n$. Since $|W|(n-|W|) \ge \alpha n(1-\alpha)n \ge \alpha n^2/2$, we have $|W|^2 \le |W|n - \alpha n^2/2$ and

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

So, by a concentration inequality,

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

and therefore

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

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. □

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 $|U_h| \geq \varepsilon|V_h|$, 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 (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 $\delta=\delta(\alpha)>0$ and $M=M(\alpha,\varepsilon)$, such that the following holds. Let G be a graph on $[n]$ with minimum degree at least αn. 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 $\left|V_i^h\right|=\Theta(n)$ and each pair $\left(V_i^1,V_i^2\right)$ is (ε,δ)-super-regular.*

We emphasize that in Lemma 6 we do *not* guarantee that the sets V_i^h are the same size. We will however see in the proof that the variation in the sizes of the clusters only depends on α and not ε (that is, the ratio between the largest and smallest cluster is bounded by a function of α).

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 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].

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

Lemma 8. *Suppose $V_1^2,\dots,V_r^2$ are clusters 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 (ε,δ')-dense, where $\delta'=\delta\min_i\left|V_i^2\right|/\left|V^2\right|$. In particular if each V_i^2 has the same size then $\delta'=\delta/r$. If moreover each each $\left(V^1,V_i^2\right)$ is (ε,δ)-super-regular, then $\left(V^1,V^2\right)$ is (ε,δ')-super-regular.*

Proof. Let $U^h \subseteq V^h$ with $\left|U^h\right| \geq \varepsilon\left|V^h\right|$. Choose i so that $\left|U^2 \cap V_i^2\right|/\left|V_i^2\right|$ is maximized, so that $\left|U^2 \cap V_i^2\right| \geq \varepsilon\left|V_i^2\right|$. By our choice of i,

$$\left|U^2\right|=\sum_{j=1}^r \frac{\left|U^2 \cap V_j^2\right|}{\left|V_j^2\right|}\left|V_j^2\right| \leq \frac{\left|U^2 \cap V_i^2\right|}{\left|V_i^2\right|}\sum_{j=1}^r\left|V_j^2\right|=\frac{\left|U^2 \cap V_i^2\right|}{\left|V_i^2\right|}\left|V^2\right|,$$

so $\left|U^2 \cap V_i^2\right| \geq \left|U^2\right| \min_j \left|V_j^2\right| / \left|V^2\right|$. It follows that there are at least $\delta\left|U^1\right|\left|U^2 \cap V_i^2\right| \geq \delta'\left|U^1\right|\left|U^2\right|$ edges between U^1 and U^2, and $\left(V^1, V^2\right)$ is (ε, δ')-dense.

The claim about super-regularity is straightforward. Note that each $v \in V^2$ has $\delta\left|V^1\right| \geq \delta'\left|V^1\right|$ neighbours in V^1 by the super-regularity of each $\left(V^1, V_i^2\right)$. Each $v \in V^1$ has $\delta\left|V_i^2\right|$ neighbours in each V_i^1, for a total of $\delta\left|V^2\right| \geq \delta'\left|V^2\right|$ neighbours in V^2. $\square$

Lemma 9. *Let $\left(V^1, V^2\right)$ be an (ε, δ)-dense pair with ε sufficiently small , and suppose we have $W^h \supseteq V^h$ with $\left|W^h\right| \leq (1 + f\varepsilon)\left|V^h\right|$. Then $\left(W^1, W^2\right)$ is an (ε', δ')-dense pair, with $\varepsilon' = \max\{2f, 1 + f\}\varepsilon$ and $\delta' = \min\{1/4, 1/(1 + f\varepsilon)\}\delta$. If moreover $\left(V^1, V^2\right)$ is (ε, δ)-super-regular and each vertex in $W^h \backslash V^h$ has at least $\delta'\left|V^{\overline{h}}\right|$ neighbours in $V^{\overline{h}}$, then $\left(W^1, W^2\right)$ is an (ε', δ')-super-regular pair.*

Proof. Suppose $U^h \subseteq W^h$ with $\left|U^h\right| \geq \varepsilon'\left|W^h\right|$. Then $\left|U^h \cap V^h\right| \geq \varepsilon'\left|W^h\right| - f\varepsilon\left|V^h\right| \geq \varepsilon\left|V^h\right|$ so there are at least $\delta\left|U^1 \cap V^1\right|\left|U^2 \cap V^2\right|$ edges between U^1 and U^2. But note that $\left|U^h \cap V^h\right| \geq \left|U^h\right| - f\varepsilon\left|W^h\right| \geq (1 - (f\varepsilon)/\varepsilon')\left|U^h\right| \geq (1/2)\left|U^h\right|$, so $d\left(U^1, U^2\right) \geq \delta/4$ and $\left(W^1, W^2\right)$ is (ε', δ')-dense.

Now consider the second part of the lemma where $\left(V^1, V^2\right)$ is (ε, δ)-super-regular. By super-regularity and the assumption in the lemma, each $v \in W^h$ has $\delta\left|V^{\overline{h}}\right| \geq \delta'\left|W^{\overline{h}}\right|$ neighbours in $W^{\overline{h}}$, proving that $\left(W^1, W^2\right)$ is (ε', δ')-super-regular. $\square$

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

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 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 $1 + \lfloor 1/\alpha \rfloor$ 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 αn) 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 ε' to be determined (which we will repeatedly assume is sufficiently small). Let Q be a cover of C by stars $T_1, \ldots, T_q$ of size at most $1 + 1/\alpha'$, 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"). Each edge of the cluster graph C corresponds to an $(\varepsilon', \alpha'/2)$-dense pair, and each W_i^h is the union of at most $1/\alpha'$ clusters V_i. By Lemma 8, with $\delta' = (\alpha')^2/2$ each pair $\left(W_i^1, W_i^2\right)$ is (ε', δ')-dense in G' (therefore G).

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 $(2\varepsilon', \delta'')$-super-regular, for $\delta'' = \delta'/2$. Combining the exceptional cluster V_0 and all the $W_i^h \backslash V_i^h$, 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 $\delta''\left|V_i^h\right|$ such neighbours contain at most $\delta'' n$ of these neighbours. Each V_i^h has size at most sn/α', so for small ε' 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 $\delta''\left|V_i^h\right|$ neighbours of v. Pick one such V_i^h uniformly at random, and put v in $V_i^{\overline{h}}$ (do this independently for each bad v). By a concentration inequality, after this procedure a.a.s. at most $2(2\varepsilon' n)/\left((\alpha')^2/(2s)\right) \leq \left(8\varepsilon'/(\alpha')^2\right)\left|W_i^h\right| \leq \left(9\varepsilon'/(\alpha')^2\right)\left|V_i^h\right|$ 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 $(O(\varepsilon'), \delta''/4)$-super-regular. With $\delta = \delta''/4$ and ε' small relative to ε, we are done. □

We apply Lemma 6 to our graph G, with $\varepsilon_1 = \varepsilon_1(\delta)$ 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 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 $n/(2(k-1))$ disjoint bare k-paths. (The definite value of k will be determined later, and will depend on ε through the $s_{\mathbf{i}}$). 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 $n/(2(k-1))$ "special pairs" of "special 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 $n/(2(k-1))$, $n/(2(k-1))$, $n/2 - 2n/(2(k-1))$ and $n/2$ respectively. We can also assume that the special pairs correspond to a uniformly 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 $X_{\mathbf{i}} \times Y_{\mathbf{j}}$. Let $X_{\mathbf{i},\mathbf{j}} \subseteq X_{\mathbf{i}}$ and $Y_{\mathbf{i},\mathbf{j}} \subseteq Y_{\mathbf{j}}$ be the set of x-type and y-type vertices in pairs in $Z_{\mathbf{i},\mathbf{j}}$, respectively. Also let $W_{\mathbf{i}} = V_{\mathbf{i}} \backslash F$ be the set of "free" vertices remaining in $V_{\mathbf{i}}$.

By a concentration inequality for the hypergeometric distribution and the union bound, a.a.s. each $|W_{\mathbf{i}}| \sim s_{\mathbf{i}}n/2$. Similarly each $|X_{\mathbf{i}}| \sim |Y_{\mathbf{i}}| \sim s_{\mathbf{i}}n/(2(k-1))$, and each $|X_{\mathbf{i},\mathbf{j}}| = |Y_{\mathbf{i},\mathbf{j}}| = |Z_{\mathbf{i},\mathbf{j}}| \sim s_{\mathbf{i}}s_{\mathbf{j}}n/(2(k-1))$. If we condition on the sizes of each $W_{\mathbf{i}}$, $X_{\mathbf{i},\mathbf{j}}$, $Y_{\mathbf{j},\mathbf{i}}$, each is a uniformly random subset of $V_{\mathbf{i}}$ 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 ε and δ). We will ensure this by appealing to the following lemma.

Lemma 13. *Fix $\delta, \varepsilon' > 0$. There is $\varepsilon = \varepsilon(\varepsilon') > 0$ such that the following holds. Suppose $\left(V^1, V^2\right)$ is (ε, δ)-super-regular, let n_h satisfy $\left|V^h\right| \geq n_h = \Omega(n)$ (not necessarily uniformly over ε and δ) and for each h let $W^h \subseteq V^h$ be a uniformly random subset of size n_h. Then $\left(W^1, W^2\right)$ is a.a.s. $(\varepsilon', \delta/2)$-super-regular.*

Proof. The fact that $\left(W^1, W^2\right)$ is a.a.s. $(\varepsilon', \delta/2)$-dense for suitable ε, follows from two applications of [7, Theorem 3.6]. (The notion of "lower regularity" in that paper essentially corresponds to our notion of (ε, δ)-density).

For each $v \in W^h$, the number of neighbours of $v \in W^{\overline{h}}$ is hypergeometrically distributed, with expected value at least $\delta n_{\overline{h}}$. Since $n_{\overline{h}} = \Omega(n)$, a concentration inequality and the union bound immediately tells us that a.a.s. each such v has $\delta n_{\overline{h}}/2$ such neighbours. This proves that $\left(W^1, W^2\right)$ is a.a.s. $(\varepsilon', \delta/2)$-super-regular. $\square$

Let $\overline{(i,h)} = \left(i, \overline{h}\right)$. It is an immediate consequence of Lemma 13 that each of $\left(X_{\mathbf{i},\mathbf{j}}, W_{\overline{\mathbf{i}}}\right)$, $\left(W_{\mathbf{i}}, W_{\overline{\mathbf{i}}}\right)$ and $\left(W_{\mathbf{i}}, Y_{\mathbf{j},\overline{\mathbf{i}}}\right)$ are $(\varepsilon_2, \delta_2)$-super-regular, where $\delta_2 = \delta_1/2$ and ε_2 can be made arbitrarily small by choice of ε_1.

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 $V_{\mathbf{i}}$ 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 $(1/2 + 1/(k-1))n$). We now have

$$|V_{\mathbf{i}}| \sim s_{\mathbf{i}}n, \ |W_{\mathbf{i}}| \sim \frac{k-1}{k+1}s_{\mathbf{i}}n, \ |X_{\mathbf{i}}|, |Y_{\mathbf{i}}| \sim \frac{1}{k+1}s_{\mathbf{i}}n, \ |Z_{\mathbf{i},\mathbf{j}}| \sim \frac{1}{k+1}s_{\mathbf{i}}s_{\mathbf{j}}n$$

for each $\mathbf{i}, \mathbf{j}$.

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

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

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

Let G be a graph on some vertex set $[N]$ ($n \geq \gamma N$), together with a vertex partition into clusters. Consider any sequence of clusters $X, V_1, \ldots, V_r, Y$, and any integer partition $t_1 + \cdots + t_r = k - 1$

(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 $\delta_1 < \delta$ 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 "δ_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 $v_j^i \in V_j$. 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 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 $\left(x^1, y^1\right), \dots, (x^n, y^n)$. 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). Since the situation is δ_1-good, $\mathbb{E}|I_X|, \mathbb{E}|I_Y| \geq \delta' n/2$ by linearity of expectation. 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 (see for example [16, Section 16.2 (McDiarmid's Inequality)]), 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 $\gamma(\delta_1) = \delta_1^2/10 < (\delta_1/3)^2$. 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 and their vertices. The vertices removed in W_1 and W_{k-1} comprise uniformly random subsets of their size, so again we can use concentration and the union bound to see that the situation is still a.a.s. δ_2-good (for any $\delta_2 < \delta_1$) after these removals. We can therefore repeat the previous argument using R_2 (and a slightly lower "goodness parameter"), to find more special paths. In fact for any sequence $\delta_1 > \delta_2 > \dots > \delta_r$, we can repeat the argument r times, where before iteration i we are δ_i-good. Let $\delta_r = \delta/2$ (say $\delta_i = \delta(1 - i/(2r))$), so there are at most

$$n \prod_{i=1}^{r} (1 - \gamma(\delta_i)) \leq (1 - \gamma(\delta/2))^r n$$

vertices remaining in X and in Y after all the r iterations. If we choose large enough r this is fewer than ξn remaining vertices, as required. $\square$

Now we return to the proof of Theorem 1. Recall that we have partitioned the vertex set V into sets of non-special vertices $W_{\mathbf{i}}$ and sets of special vertices $X_{\mathbf{i}}, Y_{\mathbf{i}}$. We denoted the set of special pairs in $X_{\mathbf{i}} \times Y_{\mathbf{i}}$ by $Z_{\mathbf{i},\mathbf{j}}$, and we had

$$|W_{\mathbf{i}}| \sim \frac{k-1}{k+1} s_{\mathbf{i}} n, \;\; |X_{\mathbf{i}}|, |Y_{\mathbf{i}}| \sim \frac{1}{k+1} s_{\mathbf{i}} n, \;\; |Z_{\mathbf{i},\mathbf{j}}| \sim \frac{1}{k+1} s_{\mathbf{i}} s_{\mathbf{j}} n.$$

For each $\mathbf{i},\mathbf{j}$ with $\mathbf{j}\neq\bar{\mathbf{i}}$, and some very small $\xi=\xi(\varepsilon,\delta)>0$, we want to find $(1-\xi)s_{\mathbf{i}}s_{\mathbf{j}}n/(k+1)$ special paths connecting special pairs in $Z_{\mathbf{i},\mathbf{j}}$. It is convenient to do this in such a way that each path connecting a special pair in $Z_{\mathbf{i},\mathbf{j}}$ contains mostly vertices from $W_{\mathbf{i}}$. In particular we would like each such special path to have one interior vertex from $W_{\bar{\mathbf{i}}}$, one interior vertex from $W_{\bar{\mathbf{j}}}$ and the other $k-3$ interior vertices from $W_{\mathbf{i}}$. This would remove a total of

$$(1-\xi)\big(2(1-s_{\mathbf{i}})s_{\bar{\mathbf{i}}}+(k-3)s_{\mathbf{i}}\big(1-s_{\bar{\mathbf{i}}}\big)\big)n/(k+1)$$

vertices from each $W_{\mathbf{i}}$. For large k this expression is less than $s_{\mathbf{i}}\big(1-s_{\bar{\mathbf{i}}}/2\big)n$, so if k is large enough then at least $\big(s_{\mathbf{i}}s_{\bar{\mathbf{i}}}/3\big)n$ vertices would remain in each $W_{\mathbf{i}}$. We can therefore find our desired paths with at most $(2q)^2$ applications of Lemma 14 (recall that q is the number of cluster-pairs in Lemma 6). To be precise, as in the proof of Lemma 14, we split R_2 into $(2q)^2$ phases and use Lemma 14 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 (leaving $\Theta(n)$ vertices) from each $V_{\mathbf{i}},W_{\mathbf{i}},X_{\mathbf{i},\mathbf{j}},Y_{\mathbf{i},\mathbf{j}}$. By Lemma 13, each $\big(X_{\mathbf{i},\mathbf{j}},W_{\bar{\mathbf{i}}}\big)$, $\big(W_{\mathbf{i}},W_{\bar{\mathbf{i}}}\big)$ and $\Big(W_{\mathbf{i}},Y_{\mathbf{j},\bar{\mathbf{i}}}\Big)$ are (ε_3,δ_3)-super-regular, where $\delta_3=\delta_2/2$ and ε_3 can be arbitrarily small respective to δ_3. Let $a=\min_{\mathbf{i}}s_{\mathbf{i}}s_{\bar{\mathbf{i}}}/3$, so $|W_{\mathbf{i}}|\geq an$ for all $\mathbf{i}$.

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. Let $\xi=\delta_3\varepsilon_3 a/2$; this will mean there are so few bad special pairs that we can greedily remove corresponding special paths without destroying super-regularity, as follows. For some bad special pair (x,y) with $x\in X_{\mathbf{i}}$ and $y\in Y_{\mathbf{j}}$, greedily build a length-$(k-3)$ path $xw_1,\ldots,w_{k-3}$ in G, where the w_ℓ alternate between $W_{\mathbf{i}}$ and $W_{\bar{\mathbf{i}}}$. Then choose a set $U_x\subseteq W_{\bar{\mathbf{i}}}$ of $\delta_3 an/4$ neighbours of w_{k-3} and a disjoint set $U_y\subseteq W_{\bar{\mathbf{j}}}$ of $\delta_3 an/4$ 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 $\delta_3 an/4$ (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 bad special pair. Note that these special paths comprise fewer than $\delta_3 an/2$ vertices total, which means that each vertex in $X_{\mathbf{i}},Y_{\mathbf{i}},W_{\mathbf{i}}$ has at least $\delta_3 an/2$ neighbours in $W_{\bar{\mathbf{i}}}$ throughout the whole process. Also, after this process each $\big(X_{\mathbf{i}},W_{\bar{\mathbf{i}}}\big)$, $\big(W_{\mathbf{i}},W_{\bar{\mathbf{i}}}\big)$ and $\big(W_{\mathbf{i}},Y_{\bar{\mathbf{i}}}\big)$ is $(2\varepsilon_3,\delta_3/2)$-regular.

Now, it is convenient to drop the distinction between x-type and y-type vertices: let $Z_{\mathbf{i}}=X_{\mathbf{i}}\cup Y_{\mathbf{i}}$. At the beginning of this section we had $|X_{\mathbf{i}}|\sim|Y_{\mathbf{i}}|$, and this is still true since we have deleted asymptotically equally many vertices from $X_{\mathbf{i}}$ and $Y_{\mathbf{i}}$. So, by Lemma 8, each $\big(Z_{\mathbf{i}},W_{\bar{\mathbf{i}}}\big)$ is $(2\varepsilon_3,\delta_3/5)$-regular.

In summary, the situation now is that we have a set of clusters $Z_{\mathbf{i}},W_{\mathbf{i}}$ of size $\Theta(n)$ partitioning the vertex set V, such that each $\big(Z_{\mathbf{i}},W_{\bar{\mathbf{i}}}\big)$, and $\big(W_{\mathbf{i}},W_{\bar{\mathbf{i}}}\big)$ is $(2\varepsilon_3,\delta_3/5)$-regular, and each special pair has one element each from $Z_{\mathbf{i}}$ and $Z_{\bar{\mathbf{i}}}$. Moreover, with Z and W being the sets of special and non-special vertices remaining, we have $(k-1)|Z|=2|W|$ (this is invariant through removals of special paths). Finally, for large k each $|W_{\mathbf{i}}|\geq|Z_{\mathbf{i}}|+\Theta(n)$ (recall that $|W_{\mathbf{i}}|$ about $s_{\mathbf{i}}s_{\bar{\mathbf{i}}}n$, and $|Z_{\mathbf{i}}|$ is about $(s_{\mathbf{i}}/k)n)$.

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 $2|W_{\mathbf{i}}|=(k-1)|Z_{\mathbf{i}}|$ for each $\mathbf{i}$. This is easily achievable through repeated application of Lemma 14, but is a bit fiddly to explain. From now on we impose that k is odd.

Let $D_{\mathbf{i}} = (k-1)|Z_{\mathbf{i}}| - 2|W_{\mathbf{i}}|$ be the "deficit" for $\mathbf{i}$. Since $(k-1)|Z| = 2|W|$, we have $\sum_{\mathbf{i}} D_{\mathbf{i}} = 0$. If all the deficits are not already zero, then let $D_{\mathbf{i}}$ be the maximum positive deficit. We will remove special paths between $Z_{\mathbf{i}}$ and $Z_{\bar{\mathbf{i}}}$ in such a way that $D_{\mathbf{i}}$ becomes zero, no nonpositive deficit becomes positive, and such that we still have each $|W_{\mathbf{i}}| \geq |Z_{\mathbf{i}}| + \Theta(n)$. After at most $2q$ iterations of this argument, all the deficits will be zero, as desired.

In each iteration we will remove a total of $\lfloor D_{\mathbf{i}}/(k-3) \rfloor + 1$ special paths. Each of the first $\lfloor D_{\mathbf{i}}/(k-3) \rfloor$ paths will have one interior vertex in $W_{\mathbf{i}}$ and in $W_{\bar{\mathbf{i}}}$, and the other $k-3$ interior vertices coming from $W_{\mathbf{i}}$ with $\mathbf{i} \neq \bar{\mathbf{i}}$ (we will explain how to choose these $W_{\mathbf{i}}$ shortly). Next, let r be the remainder after dividing $D_{\mathbf{i}}$ by $k-3$. Our final special path will have $(k-1)/2 - r/2$ interior vertices in $W_{\mathbf{i}}$ and in $W_{\bar{\mathbf{i}}}$ (recall k is odd), and the other r interior vertices coming from some $W_{\mathbf{i}}$ with $\mathbf{i} \neq \bar{\mathbf{i}}$. In all these special paths, we choose the as-yet undetermined $W_{\mathbf{i}}$ arbitrarily, so long as this includes no more than $-\left(D_{\bar{\mathbf{i}}} - D_{\mathbf{i}}\right)/2$ (≥ 0) vertices from $W_{\bar{\mathbf{i}}}$, and no more than $-D_{\mathbf{i}}/2$ vertices from each of the other $W_{\mathbf{i}}$ which have negative deficit. It is easy to see that removing all these paths will have the desired effects on the deficits.

We find all our special paths with $O(1)$ applications of Lemma 14, using the random edges in R_4. We make a few observations to show why this is valid. First, note that we need $|Z_{\mathbf{i}}|, \left|Z_{\bar{\mathbf{i}}}\right| = \lfloor D_{\mathbf{i}}/(k-3) \rfloor + 1 + \Theta(n)$ and $|W_{\mathbf{i}}|, \left|W_{\bar{\mathbf{i}}}\right| = \lfloor D_{\mathbf{i}}/(k-3) \rfloor + (k-1)/2 - r/2 + \Theta(n)$. But this holds for large n because $|W_{\mathbf{i}}|, \left|W_{\bar{\mathbf{i}}}\right| \geq |Z_{\mathbf{i}}| = \left|Z_{\bar{\mathbf{i}}}\right|$ and

$$((k-1)-2)|Z_{\mathbf{i}}| \geq (k-1)|Z_{\mathbf{i}}| - 2|W_{\mathbf{i}}| + \Theta(n) = D_{\mathbf{i}} + \Theta(n).$$

We also need there to be enough vertices in the appropriate $W_{\mathbf{i}}$ to fill the undetermined vertices in the special paths (we need $(k-3)\lfloor D_{\mathbf{i}}/(k-3) \rfloor + r = D_{\mathbf{i}}$ such). But this is a consequence of the fact that $\sum_{\mathbf{j}} D_{\mathbf{j}} = 0$. To be precise, the sum of the negative deficits is at least as negative as $-D_{\mathbf{i}}$, so our number of available vertices is

$$\sum_{\mathbf{j}} (-D_{\mathbf{j}}/2) - \left(D_{\bar{\mathbf{i}}} - D_{\mathbf{i}}\right)/2 \geq D_{\mathbf{i}},$$

where the sum is over all $\mathbf{j} \neq \mathbf{i}, \bar{\mathbf{i}}$ with $D_{\mathbf{j}}$ negative.

After we choose the number of vertices to take from each $W_{\mathbf{i}}$, the removed vertices comprise uniformly random subsets of certain sizes from each $Z_{\mathbf{i}}$ and $W_{\mathbf{i}}$. By Lemma 13 each $\left(Z_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$ and $\left(W_{\mathbf{i}}, W_{\bar{\mathbf{i}}}\right)$ is now $(\varepsilon_4, \delta_4)$-super-regular, where $\delta_4 = \delta_3/10$ and ε_4 can be arbitrarily small compared to δ_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 15 (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 $n_1, \ldots, n_r \in \mathbb{N}^+$. Let $V_1, \ldots, V_r$ be pairwise disjoint sets of sizes $n_1, \ldots, n_r$. 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.*

Now, 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. For each $i \in [q]$, let $S_i^0 = Z_{(i,1)}$, $S_i^1 = W_{(i,2)}$, $S_i^2 = W_{(i,1)}$ and $S_i^3 = Z_{(i,2)}$. Let $n_i = \left|Z_{(i,1)}\right| = \left|Z_{(i,2)}\right|$. We construct an auxiliary graph G_i as follows. Start with the subgraph of G induced by $S_i^0 \cup S_i^2 \cup S_i^3 \cup S_i^3$, and remove every edge not between some S_i^ℓ and $S_i^{\ell+1}$. Contract each special pair (consisting of a vertex in S_i^0 and in S_i^3) to a single vertex. This yields a super-regular "cluster-cycle" with 3 clusters S_i^*, S_i^1 and $S_{\mathbf{i}}^2$ of sizes n_i, $(k-1)n_i/2$ and $(k-1)n_i/2$. This is a (ε_5,δ_5)-super-regular blow-up of a 3-cycle C_3.

The corresponding 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 ε the (ε,δ'')-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^2, S_i^3, S_i^3, and these special paths complete our embedding of T.

THE FOLLOWING PICTURE NEEDS TO BE UPDATED

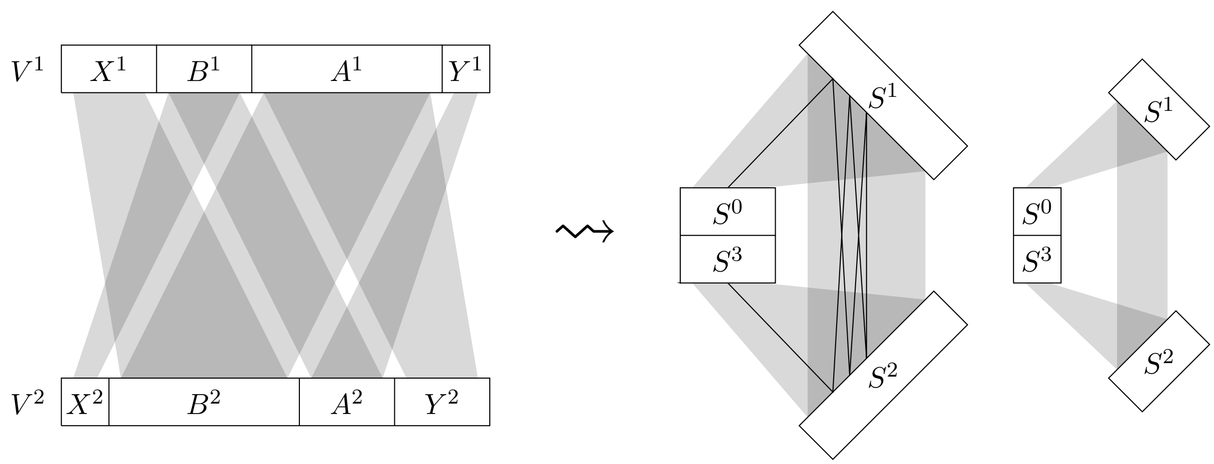

Figure 1. From super-regular pairs to cluster-cycles. An example special path is shown, for $k = 7$.

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] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, *Small subsets inherit sparse ε-regularity*, Journal of Combinatorial Theory, Series B **97** (2007), no. 1, 34–56.

[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] M. Molloy and B. Reed, *Graph colouring and the probabilistic method*, vol. 23, Springer Science & Business Media, 2002.

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

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

[19] 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.