The approximate Loebl–Komlós–Sós Conjecture I: The sparse decomposition

Jan Hladký [*]   János Komlós[†]   Diana Piguet[‡]
Miklós Simonovits[§]   Maya Stein[¶]   Endre Szemerédi[‖]

Abstract

In a series of four papers we prove the following relaxation of the Loebl–Komlós–Sós Conjecture: For every $\alpha > 0$ there exists a number k_0 such that for every $k > k_0$ every n-vertex graph G with at least $(\frac{1}{2} + \alpha)n$ vertices of degree at least $(1+\alpha)k$ contains each tree T of order k as a subgraph.

The method to prove our result follows a strategy similar to approaches that employ the Szemerédi regularity lemma: we decompose the graph G, find a suitable combinatorial structure inside the decomposition, and then embed the tree T into G using this structure. Since for sparse graphs G, the decomposition given by the regularity lemma is not helpful, we use a more general decomposition technique. We show that each graph can be decomposed into vertices of huge degree, regular pairs (in the sense of the regularity lemma), and two other objects each exhibiting certain expansion properties. In this paper, we introduce this novel decomposition technique. In the three follow-up papers, we find a combinatorial structure suitable inside the decomposition, which we then use for embedding the tree.

Mathematics Subject Classification: 05C35 (primary), 05C05 (secondary).
Keywords: extremal graph theory; Loebl–Komlós–Sós Conjecture; tree embedding; regularity lemma; sparse graph; graph decomposition.

Contents

1 Introduction

1.1 Statement of the problem

This is the first of a series of four papers [HKP$^+$a, HKP$^+$b, HKP$^+$c, HKP$^+$d] in which we provide an approximate solution of the Loebl–Komlós–Sós Conjecture, a problem in extremal graph theory which fits the classical form *Does a certain density condition imposed on a graph guarantee a certain subgraph?* Classical results of this type include Dirac's Theorem which determines the minimum degree threshold for containment of a Hamilton cycle, or Mantel's Theorem which determines the average degree threshold for containment of a triangle. Indeed, most of these extremal problems are formulated in terms of the minimum or average degree of the host graph.

We investigate a density condition which guarantees the containment of *each* tree of order k. The greedy tree-embedding strategy shows that requiring a minimum degree of more than $k - 2$ is sufficient. Further, this bound is best possible as any $(k - 2)$-regular graph avoids the k-vertex star. Erdős and Sós conjectured that one can replace the minimum degree with the average degree, with the same conclusion.

Conjecture 1.1 (Erdős–Sós Conjecture 1963)**.** *Let G be a graph of average degree greater than $k - 2$. Then G contains each tree of order k as a subgraph.*

A solution of the Erdős–Sós Conjecture for all k greater than some absolute constant was announced by Ajtai, Komlós, Simonovits, and Szemerédi in the early 1990's. In a similar spirit, Loebl, Komlós, and Sós conjectured that a *median degree* of $k - 1$ or more is sufficient for containment of any tree of order k. By median degree we mean the degree of a vertex in the middle of the ordered degree sequence.

Conjecture 1.2 (Loebl–Komlós–Sós Conjecture 1995 [EFLS95])**.** *Suppose that G is an n-vertex graph with at least $n/2$ vertices of degree more than $k - 2$. Then G contains each tree of order k.*

We discuss Conjectures 1.1 and 1.2 in detail in Section 1.3. Here, we just state the main result we achieve in our series of four papers, an approximate solution of the Loebl–Komlós–Sós Conjecture.

Theorem 1.3 (Main result [HKP$^+$d])**.** *For every $\alpha > 0$ there exists k_0 such that for any $k > k_0$ we have the following. Each n-vertex graph G with at least $(\frac{1}{2} + \alpha)n$ vertices of degree at least $(1 + \alpha)k$ contains each tree T of order k.*

The proof of this theorem is in [HKP$^+$d]. The first step towards this result is Lemma 3.14, which constitutes the main result of the present paper. It gives a decomposition of the host graph G into several parts which will be later useful for the embedding. See Section 1.5 for a description of the result and its role in the proof of Theorem 1.3. Also see [HPS$^+$15] for a more detailed overview of the proof.

1.2 The regularity lemma and the sparse decomposition

The Szemerédi regularity lemma has been a major tool in extremal graph theory for more than three decades. It provides an approximation of an arbitrary graph by a collection of generalized quasi-random graphs. This allows to represent the graph by a so-called *cluster graph*. Then, instead of solving the original problem, one can solve a modified simpler problem in the cluster graph.

The applicability of the original Szemerédi regularity lemma is, however, limited to *dense graphs*, i.e., graphs that contain a substantial proportion of all possible edges. There is a version of the regularity lemma for sparser graphs by Kohayakawa and Rödl [Koh97] later strengthened by Scott [Sco11], as well as other statements that draw on something from its philosophy (e.g. [EL]). However, these statements provide a picture much less informative than Szemerédi's original result. A regularity type representation of general (possibly sparse) graphs is one of the most important goals of contemporary discrete mathematics. By such a representation we mean an approximation of the input graph by a structure of bounded complexity carrying enough of the important information about the graph.

A central tool in the proof of Theorem 1.3 is a structural decomposition of the graph G. This decomposition — which we call *sparse decomposition* — applies to any graph whose average degree is greater than a constant. The sparse decomposition provides a partition of any graph into vertices of huge degrees and into a bounded degree part. The bounded degree part is further decomposed into dense regular pairs, an edge set with certain expander-like properties, and a vertex set which is expanding in a different way (we shall give a more precise description in Section 1.5). This kind of decomposition was first used by Ajtai, Komlós, Simonovits, and Szemerédi in their yet unpublished work on the Erdős–Sós Conjecture. The main goal of this paper is to present the sparse decomposition, and to show that each graph has such a sparse decomposition: This will be done in Lemma 3.13. Lemma 3.14 provides a sparse decomposition with additional tailor-made features for graphs that fulfil the conditions of Theorem 1.3.

In the case of dense graphs the sparse decomposition produces a Szemerédi regularity partition (as explained in Section 3.8), and thus the decomposition lemma (Lemma 3.13) extends the Szemerédi regularity lemma. But the interesting setting for the decomposition lemma is the field of sparse graphs.

1.3   Loebl–Komlós–Sós Conjecture and Erdős–Sós Conjecture

Let us first introduce some notation. We say that H *embeds* in a graph G and write $H \subseteq G$ if H is a (not necessarily induced) subgraph of G. The associated map $\phi : V(H) \to V(G)$ is called an *embedding of H in G*. More generally, for a graph class $\mathcal{H}$ we write $\mathcal{H} \subseteq G$ if $H \subseteq G$ for every $H \in \mathcal{H}$. Let **trees**(k) be the class of all trees of order k.

Conjecture 1.2 is dominated by two parameters: one quantifies the number of vertices of 'large' degree, and the other tells us how large this degree should actually be. Strengthening either of these bounds sufficiently, the conjecture becomes trivial. Indeed, if we replace $n/2$ with n, then any tree of order k can be embedded greedily. Also, if we replace $k - 2$ with $4k - 4$, then G, being a graph of average degree at least $2k - 2$, has a subgraph G' of minimum degree at least $k - 1$. Again we can greedily embed any tree of order k.

On the other hand, one may ask whether smaller lower bounds would suffice. For the bound $k-2$, this is not the case, since stars of order k require a vertex of degree at least $k - 1$ in the host graph. Another example can be obtained by considering a disjoint union of cliques of order $k - 1$. No tree of order k is contained in such a graph.

For the bound $n/2$, the following example shows that this number cannot be decreased much. First, assume that n is even, and that $n = k$. Let G^* be obtained from the complete graph on n vertices by deleting all edges inside a set of $\frac{n}{2} + 1$ vertices. It is easy to check that G^* does not contain the k-vertex path. In general, G^* does not contain any tree of order k with independence

$$\frac{k}{2} - 1$$

$$\frac{k}{2} + 1$$
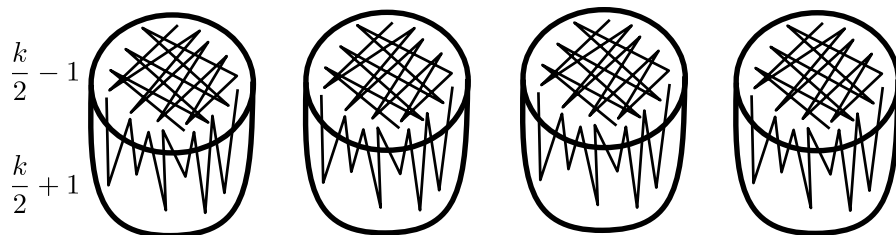
Figure 1.1: An extremal graph for the Loebl–Komlós–Sós Conjecture.

number less than $\frac{k}{2} + 1$. Now, taking the union of several disjoint copies of G^* we obtain examples for other values of n. (And adding a small complete component we can get to *any* value of n.) See Figure 1.1 for an illustration.

However, we do not know of any example attaining the exact bound $n/2$. Thus it might be possible to lower the bound $n/2$ from Conjecture 1.2 to the one attained in our example above:

Conjecture 1.4. *Let $k \in \mathbb{N}$ and let G be a graph on n vertices, with more than $\frac{n}{2} - \lfloor \frac{n}{k} \rfloor - (n \mod k)$ vertices of degree at least $k - 1$. Then* **trees**$(k) \subseteq G$.

It might even be that if n/k is far from integrality, a slightly weaker lower bound on the number of vertices of large degree still works (see [Hla, HP15]).

Several partial results concerning Conjecture 1.2 have been obtained; let us briefly summarize the major ones. Two main directions can be distinguished among those results that prove the conjecture for special classes of graphs: either one places restrictions on the host graph, or on the class of trees to be embedded. Of the latter type is the result by Bazgan, Li, and Woźniak [BLW00], who proved the conjecture for paths. Also, Piguet and Stein [PS08] proved that Conjecture 1.2 is true for trees of diameter at most 5, which improved earlier results of Barr and Johansson [BJ] and Sun [Sun07]. Restrictions on the host graph have led to the following results. Soffer [Sof00] showed that Conjecture 1.2 is true if the host graph has girth at least 7. Dobson [Dob02] proved the conjecture for host graphs whose complement does not contain a $K_{2,3}$. This has been extended by Matsumoto and Sakamoto [MS] who replace the $K_{2,3}$ with a slightly larger graph.

A different approach is to solve the conjecture for special values of k. One such case, known as the Loebl conjecture, or also as the $(n/2$–$n/2$–$n/2)$-Conjecture, is the case $k = n/2$. Ajtai, Komlós, and Szemerédi [AKS95] solved an approximate version of this conjecture, and later Zhao [Zha11] used a refinement of this approach to prove the sharp version of the conjecture for large graphs.

An approximate version of Conjecture 1.2 for dense graphs, that is, for k linear in n, was proved by Piguet and Stein [PS12].

Theorem 1.5 (Piguet–Stein [PS12]). *For any $q > 0$ and $\alpha > 0$ there exists a number n_0 such that for any $n > n_0$ and $k > qn$ the following holds. For each n-vertex graph G with at least $n/2$ vertices of degree at least $(1 + \alpha)k$ we have* **trees**$(k + 1) \subseteq G$.

This result was proved using the regularity method. Adding stability arguments, Hladký and Piguet [HP15], and independently Cooley [Coo09] proved Conjecture 1.2 for large dense graphs.

Theorem 1.6 (Hladký–Piguet [HP15], Cooley [Coo09]). *For any $q > 0$ there exists a number $n_0 = n_0(q)$ such that for any $n > n_0$ and $k > qn$ the following holds. For each n-vertex graph G with at least $n/2$ vertices of degree at least k we have* **trees**$(k + 1) \subseteq G$.
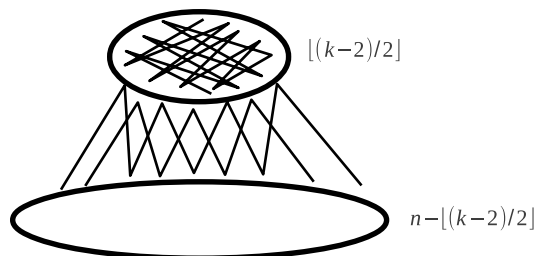
Figure 1.2: An almost extremal graph for the Erdős–Sós Conjecture.

Let us now turn our attention to the Erdős–Sós Conjecture. The Erdős–Sós Conjecture 1.1 is best possible whenever $n(k-2)$ is even. Indeed, in that case it suffices to consider a $(k-2)$-regular graph. This is a graph with average degree exactly $k-2$ which does not contain the star of order k. Even when the star (which in a sense is a pathological tree) is excluded from the considerations, we can — at least when $k-1$ divides n — consider a disjoint union of $\frac{n}{k-1}$ cliques K_{k-1}. This graph contains *no* tree from **trees**(k). There is another important graph with many edges which does not contain for example the path P_k, depicted in Figure 1.2. This graph consists of a set of vertices of size $\lfloor (k-2)/2 \rfloor$ that are connected to all vertices in the graph. This graph has $\frac{1}{2}(k-2)n - O(k^2)$ edges when k is even and $\frac{1}{2}(k-3)n - O(k^2)$ edges otherwise, and therefore gets close to the conjectured bound when $k \ll n$. Apart from the already mentioned announced breakthrough by Ajtai, Komlós, Simonovits, and Szemerédi, work on this conjecture includes [BD96, Hax01, MS, SW97, Woź96].

Both Conjectures 1.2 and Conjecture 1.1 have an important application in Ramsey theory. Each of them implies that the Ramsey number of two trees $T_{k+1} \in \mathbf{trees}(k+1)$, $T_{\ell+1} \in \mathbf{trees}(\ell+1)$ is bounded by $R(T_{k+1}, T_{\ell+1}) \leqslant k + \ell + 1$. Actually more is implied: Any 2-edge-colouring of $K_{k+\ell+1}$ contains either *all* trees in **trees**$(k+1)$ in red, or *all* trees in **trees**$(\ell+1)$ in blue.

The bound $R(T_{k+1}, T_{\ell+1}) \leqslant k + \ell + 1$ is almost tight only for certain types of trees. For example, Gerencsér and Gyárfás [GG67] showed $R(P_k, P_\ell) = \max\{k, \ell\} + \left\lfloor \frac{\min\{k,\ell\}}{2} \right\rfloor - 1$ for paths $P_k \in \mathbf{trees}(k)$, $P_\ell \in \mathbf{trees}(\ell)$. Harary [Har72] showed $R(S_k, S_\ell) = k + \ell - 2 - \varepsilon$ for stars $S_k \in \mathbf{trees}(k)$, $S_\ell \in \mathbf{trees}(\ell)$, where $\varepsilon \in \{0, 1\}$ depends on the parity of k and ℓ. Haxell, Łuczak, and Tingley confirmed asymptotically [HLT02] that the discrepancy of the Ramsey bounds for trees depends on their balancedness, at least when the maximum degrees of the trees considered are moderately bounded.

1.4 Related tree containment problems

Minimum degree conditions for spanning trees. Recall that the tight min-degree condition for containment of a general spanning tree T in an n-vertex graph G is the trivial one, $\mathrm{mindeg}(G) \geqslant n - 1$. However, the only tree which requires this bound is the star. This indicates that this threshold can be lowered substantially if we have a control of $\mathrm{maxdeg}(T)$. Szemerédi and his collaborators [KSS01, CLNGS10] showed that this is indeed the case, and obtained tight min-degree bounds for certain ranges of $\mathrm{maxdeg}(T)$. For example, if $\mathrm{maxdeg}(T) \leqslant n^{o(1)}$, then $\mathrm{mindeg}(G) \geqslant (\frac{1}{2} + o(1))n$ is a sufficient condition. (Note that G may become disconnected close to this bound.)

Trees in random graphs. To complete the picture of research involving tree containment problems we mention two rich and vivid (and also closely connected) areas: trees in random graphs,

and trees in expanding graphs. The former area is centered around the following question: *What is the probability threshold $p = p(n)$ for the Erdős–Rényi random graph $G_{n,p}$ to contain asymptotically almost surely (a.a.s.) each tree/all trees from a given class $\mathcal{F}_n$ of trees?* Note that there is a difference between containing "each tree" and "all trees" (i.e., all trees simultaneously; this is often referred to as *universality*) as the error probabilities for missing individual trees might sum up.

Most research focused on containment of spanning trees, or almost spanning trees. The only well-understood case is when $\mathcal{F}_n = \{P_{k_n}\}$ is a path. The threshold $p = \frac{(1+o(1))\ln n}{n}$ for appearance of a spanning path (i.e., $k_n = n$) was determined by Komlós and Szemerédi [KS83], and independently by Bollobás [Bol84]. Note that this threshold is the same as the threshold for connectedness. We should also mention a previous result of Pósa [Pós76] which determined the order of magnitude of the threshold, $p = \Theta(\frac{\ln n}{n})$. The heart of Pósa's proof, the celebrated rotation-extension technique, is an argument about expanding graphs, and indeed many other results about trees in random graphs exploit the expansion properties of $G_{n,p}$ in the first place.

The threshold for the appearance of almost spanning paths in $G_{n,p}$ was determined by Fernandez de la Vega [FdlV79] and independently by Ajtai, Komlós, and Szemerédi [AKS81]. Their results say that a path of length $(1 - \varepsilon)n$ appears a.a.s. in $G_{n,\frac{C}{n}}$ for $C = C(\varepsilon)$ sufficiently large. This behavior extends to bounded degree trees. Indeed, Alon, Krivelevich, and Sudakov [AKS07] proved that $G_{n,\frac{C}{n}}$ (for a suitable $C = C(\varepsilon, \Delta)$) a.a.s. contains all trees of order $(1 - \varepsilon)n$ with maximum degree at most Δ (the constant C was later improved in [BCPS10]).

Let us now turn to spanning trees in random graphs. It is known [AKS07] that a.a.s. $G_{n,\frac{C\ln n}{n}}$ contains a single spanning tree T with bounded maximum degree and linearly many leaves. This result can be reduced to the main result of [AKS07] regarding almost spanning trees quite easily. The constant C can be taken $C = 1 + o(1)$, as was shown recently by Hefetz, Krivelevich, and Szabó [HKS12]; obviously this is best possible. The same result also applies to trees that contain a path of linear length whose vertices all have degree two. A breakthrough in the area was achieved by Krivelevich [Kri10] who gave an upper bound on the threshold $p = p(n, \Delta)$ for embedding a single spanning tree of a given maximum degree Δ. This bound is essentially tight for $\Delta = n^c$, $c \in (0, 1)$. Even though the argument in [Kri10] is not difficult, it relies on a deep result of Johansson, Kahn and Vu [JKV08] about factors in random graphs. Montgomery [Mona] complemented Krivelevich's result obtaining an almost tight upper bound on $p(n, \Delta)$ in the case when Δ is small. Further, Montgomery [Monb] achieved an essentially optimal bound for containment of some comb-like graphs.

Regarding universality of random graphs with respect to spanning trees, most of the research focused on the subclass of bounded-degree trees. Let us mention papers [JKS12] and [FNP] which improve the upper-bounds for the probability of containing all trees of maximum degree Δ (the results are meaningful for $\Delta < n^c$ for some small value of c).

Trees in expanders.   By an expander graph we mean a graph with a large Cheeger constant, i.e., a graph which satisfies a certain isoperimetric property. As indicated above, random graphs are very good expanders, and this is the main motivation for studying tree containment problems in expanders. Another motivation comes from studying the universality phenomenon. Here the goal is to construct sparse graphs which contain all trees from a given class, and expanders are natural candidates for this. The study of sparse tree-universal graphs is a remarkable area by itself which brings challenges both in probabilistic and explicit constructions. For example, Bhatt, Chung, Leighton, and Rosenberg [BCLR89] give an explicit construction of a graph with only $O_\Delta(n)$ edges

which contains all n-vertex trees with maximum degree at most Δ. The above mentioned paper by Johannsen, Krivelevich, and Samotij [JKS12] shows a number of universality results for expanders, too. For example, they show universality for the class of graphs with a large Cheeger constant that satisfy a certain connectivity condition.

Friedman and Pippenger [FP87] extended Pósa's rotation-extension technique from paths to trees and found many applications (e.g. [HK95, Hax01, BCPS10]). Sudakov and Vondrák [SV10] use tree-indexed random walks to embed trees in $K_{s,t}$-free graphs (this property implies expansion); a similar approach is employed by Benjamini and Schramm [BS97] in the setting of infinite graphs. Tree-indexed random walks are also used (in conjunction with the Regularity Lemma) in the work of Kühn, Mycroft, and Osthus on Sumner's universal tournament conjecture, [KMO11a, KMO11b].

In our proof of Theorem 1.3, embedding trees in expanders play a crucial role, too. However, our notion of expansion is very different from those studied previously. (Actually, we introduce two, very different, notions in Definitions 3.3 and 3.6.)

1.5 Overview of the proof of our main result

This is a very brief overview of the proof. A more thorough overview is given in [HPS$^+$15].

The structure of the proof of our main result (Theorem 1.3) resembles the proof of the dense case, Theorem 1.5. We obtain an approximate representation — called the *sparse decomposition* — of the host graph G from Theorem 1.3. Then we find a suitable combinatorial structure inside the sparse decomposition. Finally, we embed a given tree T into G using this structure.

Here we explain the key ingredients of the proof in more detail. The input graph G has $\Theta(kn)$ edges. Indeed, an easy counting argument gives that $e(G) \geqslant kn/4$. On the other hand, we can assume that $e(G) < kn$, as otherwise G contains a subgraph of minimum degree at least k, and the assertion of Theorem 1.3 follows. Recall that the Szemerédi regularity lemma gives an approximation of dense graphs in which $o(n^2)$ edges are neglected. The sparse decomposition introduced here captures all but at most $o(kn)$ edges. The vertex set of G is partitioned into a set of vertices of degree much larger than k and a set of vertices of degree $O(k)$. Further, the induced graph on the second set is split into regular pairs (in the sense of the Szemerédi regularity lemma) with clusters of sizes $\Theta(k)$ leading to a cluster graph $\mathbf{G}_{\mathrm{reg}}$, and into two additional parts which each have certain (different) expansion properties. The first of these two expanding parts — called G_{exp} — is a subgraph of G that contains no bipartite subgraphs of a density above a certain threshold density (we call such bipartite subgraphs *dense spots*). The second expanding part — called the *avoiding set* $\mathbb{E}$— consists of vertices that lie in many of these dense spots. The vertices of huge degrees, the regular pairs, and the two expanding parts form the sparse decomposition of G. The key ideas behind obtaining this sparse decomposition are given in [HPS$^+$15, Section 3], and full details can be found in Section 3. It is well-known that regular pairs are suitable for embedding small trees. In [HKP$^+$d] we work out techniques for embedding small trees in each of the three remaining parts of the sparse decomposition. A nontechnical description of these techniques is given in Section 3.5 (for $\mathbb{E}$) and Section 3.6 (for G_{exp}). It is a bit difficult to describe precisely the way the huge degree vertices are utilized. At this moment it suffices to say that it is easy to extend a partial embedding of a k-vertex tree from a vertex u mapped to a huge-degree vertex x to the children of u. Of course, for such an extension alone, $\deg(x) \geqslant k - 1$ would have been sufficient. So, the fact that the degree of x is much larger than k is used (together with other properties) to accommodate these children so that it is possible to continue even with subsequent extensions.

Tree-embedding results in the dense setting (e.g. Theorem 1.5) rely on finding a (connected) matching structure in the cluster graph. Indeed, this allows for distributing different parts of the tree in the matching edges. In analogy, in the second paper of this series [HKP+b] we find a structure based on the sparse decomposition. This *rough structure* utilizes all the concepts suitable for embedding trees described above: huge degree vertices, the avoiding set $\mathbb{E}$, the graph $G_{\exp}$, and dense regular pairs. Somewhat surprisingly, the dense regular pairs do not come only from $\mathbf{G}_{\mathrm{reg}}$. Let us make this more precise. An initial matching structure is found in $\mathbf{G}_{\mathrm{reg}}$ and this structure is enhanced using other parts of G to yield further regular pairs, referred to in this context as the *regularized matching*. One may ask what the role of $\mathbf{G}_{\mathrm{reg}}$ is. The answer is that either we can take directly a sufficiently large matching in $\mathbf{G}_{\mathrm{reg}}$, or the lack of any such matching in $\mathbf{G}_{\mathrm{reg}}$ gives us information about a compensating enhancement in a form of a regularized matching based on other parts of the decomposition. A simplified version of this rough structure is given as Lemma 7 in [HPS+15].

However, the rough structure is not immediately suitable for embedding T, and we shall further refine it in the third paper of this series [HKP+c]. We will show that in the setting of Theorem 1.3, we can always find one of ten *configurations*, denoted by ($\diamond\mathbf{1}$)–($\diamond\mathbf{10}$), in the host graph G. Obtaining these configurations from the rough structure is based on pigeonhole-type arguments such as: if there are many edges between two sets, and few "kinds" of edges, then many of the edges are of the same kind. The different kinds of edges come from the sparse decomposition (and allow for different kinds of embedding techniques). Just "homogenizing" the situation by restricting to one particular kind is not enough, we also need to employ certain "cleaning lemmas". A simplest such lemma would be that a graph with many edges contains a subgraph with a large minimum degree; the latter property evidently being more directly applicable for a sequential embedding of a tree. The actual cleaning lemmas we use are complex extensions of this simple idea.

Finally, in [HKP+d], we show how to embed the tree T. This is done by first establishing some elementary embedding lemmas for small subtrees, and then combine these for each of the cases ($\diamond\mathbf{1}$)–($\diamond\mathbf{10}$) to yield an embedding of the entire tree T.

A scheme of the proof of Theorem 1.3 is given in Figure 1.3.

2 Notation and preliminaries

2.1 General notation

All graphs considered in this paper are finite, undirected, without multiple edges, and without self-loops. We write $V(G)$ and $E(G)$ for the vertex set and edge set of a graph G, respectively. Further, $v(G) = |V(G)|$ is the order of G, and $e(G) = |E(G)|$ is its number of edges. If $X, Y \subseteq V(G)$ are two, not necessarily disjoint, sets of vertices we write $e(X)$ for the number of edges induced by X, and $e(X, Y)$ for the number of ordered pairs $(x, y) \in X \times Y$ such that $xy \in E(G)$. In particular, note that $2e(X) = e(X, X)$.

For a graph G, a vertex $v \in V(G)$ and a set $U \subseteq V(G)$, we write $\deg(v)$ and $\deg(v, U)$ for the degree of v, and for the number of neighbours of v in U, respectively. We write $\mathrm{mindeg}(G)$ for the minimum degree of G, $\mathrm{mindeg}(U) := \min\{\deg(u) : u \in U\}$, and $\mathrm{mindeg}(V_1, V_2) = \min\{\deg(u, V_2) : u \in V_1\}$ for two sets $V_1, V_2 \subseteq V(G)$. Note that for us, the minimum degree of a graph on zero vertices is ∞. Similar notation is used for the maximum degree, denoted by $\mathrm{maxdeg}(G)$. The neighbourhood of a vertex v is denoted by $\mathrm{N}(v)$. We set $\mathrm{N}(U) := \bigcup_{u \in U} \mathrm{N}(u)$. The symbol $-$ is

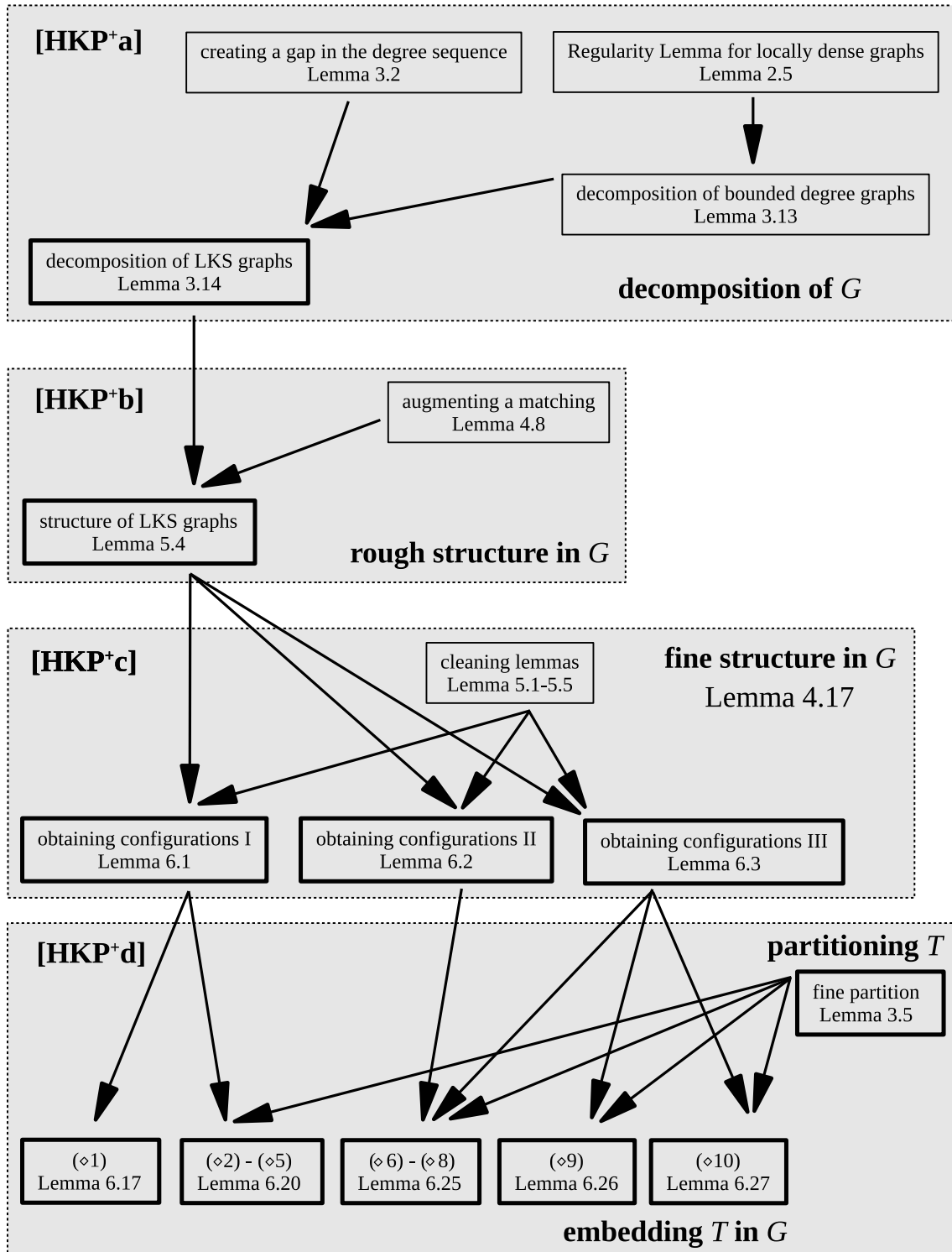**Figure 1.3:** Structure of the proof of Theorem 1.3, including parts from [HKP$^+$b, HKP$^+$c, HKP$^+$d].

used for two graph operations: if $U \subseteq V(G)$ is a vertex set then $G - U$ is the subgraph of G induced by the set $V(G) \setminus U$. If $H \subseteq G$ is a subgraph of G then the graph $G - H$ is defined on the vertex set $V(G)$ and corresponds to deletion of edges of H from G. Any graph with zero edges is called *empty*.

A family $\mathcal{A}$ of pairwise disjoint subsets of $V(G)$ is an *ℓ-ensemble in G* if $|A| \geqslant \ell$ for each $A \in \mathcal{A}$. The set $\{1, 2, \ldots, n\}$ of the first n positive integers is denoted by $[n]$.

Suppose that we have a nonempty set A, and $\mathcal{X}$ and $\mathcal{Y}$ each partition A. Then $\mathcal{X} \boxplus \mathcal{Y}$ denotes the coarsest common refinement of $\mathcal{X}$ and $\mathcal{Y}$, i.e.,

$$\mathcal{X} \boxplus \mathcal{Y} := \{X \cap Y \ : \ X \in \mathcal{X}, Y \in \mathcal{Y}\} \setminus \{\emptyset\} \ .$$

We frequently employ indexing by many indices. We write superscript indices in parentheses (such as $a^{(3)}$), as opposed to notation of powers (such as a^3). We use sometimes subscript to refer to parameters appearing in a fact/lemma/theorem. For example, $\alpha_{\text{T1.3}}$ refers to the parameter α from Theorem 1.3. We omit rounding symbols when this does not affect the correctness of the arguments. In overviews we use the symbol $\ll$ equivalently to the $o(\cdot)$ symbol.

In Table 2.1 we indicate our notation system (with an outlook to [HKP$^+$b]–[HKP$^+$d]).

Table 2.1: Specific notation used in the series.

lower case Greek letters	small positive constants ($\ll 1$)
	ϕ reserved for embedding; $\phi : V(T) \to V(G)$
upper case Greek letters	large positive constants ($\gg 1$)
one-letter bold	sets of clusters
bold (e.g., $\mathbf{trees}(k), \mathbf{LKS}(n, k, \eta)$)	classes of graphs
blackboard bold (e.g., $\mathbb{H}, \mathbb{E}, \mathbb{S}_{\eta,k}(G), \mathbb{XA}$)	distinguished vertex sets, except for
	$\mathbb{N}$ which denotes the set $\{1, 2, \ldots\}$
script (e.g., $\mathcal{A}, \mathcal{D}, \mathcal{N}$)	families (of vertex sets, "dense spots", and regular pairs)
∇ ("nabla")	reserved for "sparse decomposition"

Lemma 2.1. *For all $\ell, n \in \mathbb{N}$, every n-vertex graph G contains a (possibly empty) subgraph G' such that $\mathrm{mindeg}(G') \geqslant \ell$ and $e(G') \geqslant e(G) - (\ell - 1)n$.*

Proof. We construct the graph G' by sequentially removing vertices of degree less than ℓ from the graph G. In each step we remove at most $\ell - 1$ edges. Thus the statement follows. $\square$

2.2 Regular pairs

In this section we introduce the notion of regular pairs which is central for Szemerédi's regularity lemma and its extension, discussed in Section 2.3. We also list some simple properties of regular pairs.

Given a graph H and a pair (U, W) of disjoint sets $U, W \subseteq V(H)$ the *density of the pair (U, W)* is defined as

$$\mathrm{d}(U, W) := \frac{e(U, W)}{|U||W|} \ .$$

For a given $\varepsilon > 0$, a pair (U, W) of disjoint sets $U, W \subseteq V(H)$ is called an *ε-regular pair* if $|\mathrm{d}(U, W) - \mathrm{d}(U', W')| < \varepsilon$ for every $U' \subseteq U$, $W' \subseteq W$ with $|U'| \geqslant \varepsilon|U|$, $|W'| \geqslant \varepsilon|W|$. If the pair (U, W) is not ε-regular, then we call it *ε-irregular*.

We give a useful and well-known property of regular pairs.

Fact 2.2. *Suppose that (U, W) is an ε-regular pair of density d. Let $U' \subseteq W, W' \subseteq W$ be sets of vertices with $|U'| \geqslant \alpha|U|$, $|W'| \geqslant \alpha|W|$, where $\alpha > \varepsilon$. Then the pair (U', W') is a $2\varepsilon/\alpha$-regular pair of density at least $d - \varepsilon$.*

The following fact states a simple relation between the density of a (not necessarily regular) pair and the densities of its subpairs.

Fact 2.3. *Let $H = (U, W; E)$ be a bipartite graph of $d(U, W) \geqslant \alpha$. Suppose that the sets U and W are partitioned into sets $\{U_i\}_{i \in I}$ and $\{W_j\}_{j \in J}$, respectively. Then at most $\beta e(H)/\alpha$ edges of H belong to a pair (U_i, W_j) with $d(U_i, W_j) \leqslant \beta$.*

Proof. Trivially, we have

$$\sum_{i \in I, j \in J} \frac{|U_i||W_j|}{|U||W|} = 1 \,. \tag{2.1}$$

Consider a pair (U_i, W_j) of $d(U_i, W_j) \leqslant \beta$. Then

$$e(U_i, W_j) \leqslant \beta|U_i||W_j| = \frac{\beta}{\alpha} \frac{|U_i||W_j|}{|U||W|} \alpha|U||W| \leqslant \frac{\beta}{\alpha} \frac{|U_i||W_j|}{|U||W|} e(U, W) \,.$$

Summing over all such pairs (U_i, W_j) and using (2.1) yields the statement. □

2.3   Regularizing locally dense graphs

The regularity lemma [Sze78] has proved to be a powerful tool for attacking graph embedding problems; see [KO09] for a survey. We first state the lemma in its original form.

Lemma 2.4 (Regularity lemma)**.** *For all $\varepsilon > 0$ and $\ell \in \mathbb{N}$ there exist $n_0, M \in \mathbb{N}$ such that for every $n \geqslant n_0$ the following holds. Let G be an n-vertex graph whose vertex set is pre-partitioned into sets $V_1, \ldots, V_{\ell'}$, $\ell' \leqslant \ell$. Then there exists a partition $\{U_0, U_1, \ldots, U_p\}$ of $V(G)$, $\ell < p < M$, with the following properties.*

(1) For every $i, j \in [p]$ we have $|U_i| = |U_j|$, and $|U_0| < \varepsilon n$.

(2) For every $i \in [p]$ and every $j \in [\ell']$ either $U_i \cap V_j = \emptyset$ or $U_i \subseteq V_j$.

(3) All but at most εp^2 pairs (U_i, U_j), $i, j \in [p]$, $i \neq j$, are ε-regular.

Property (3) of Lemma 2.4 is often called *ε-regularity* of the partition $\{U_0, U_1, \ldots, U_p\}$. For us, it is more convenient to introduce this notion in the bipartite context (in Definition 2.6).

We shall use Lemma 2.4 for auxiliary purposes only as it is helpful only in the setting of dense graphs (i.e., graphs which have n vertices and $\Omega(n^2)$ edges). This is not necessarily the case in Theorem 1.3. For this reason, we give a version of the regularity lemma — Lemma 2.5 below — which allows us to regularize even sparse graphs.

More precisely, suppose that we have an n-vertex graph H whose edges lie in bipartite graphs $H[W_i, W_j]$, where $\{W_1, \ldots, W_\ell\}$ is an ensemble of sets of individual sizes $\Theta(k)$. Although ℓ may be unbounded, for a fixed $i \in [\ell]$ there are only a bounded number (independent of k), say m, of indices $j \in [\ell]$ such that $H[W_i, W_j]$ is non-empty. See Figure 2.1 for an example. Lemma 2.5 then
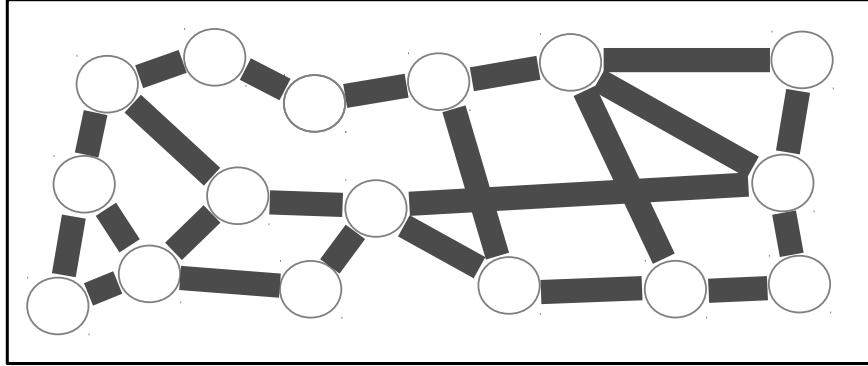
Figure 2.1: A locally dense graph as in Lemma 2.5. The sets $W_1, \ldots, W_\ell$ are depicted with grey circles. Even though there is a large number of them, each W_i is linked to only boundedly many other W_j's (at most four, in this example). Lemma 2.5 allows us to regularize all the bipartite graphs using the same system of partitions of the sets W_i.

allows us to regularize (in the sense of the regularity lemma, Lemma 2.4) all the bipartite graphs $G[W_i, W_j]$ using the same partition $\{W_i^{(0)} \dot\cup W_i^{(1)} \dot\cup \ldots \dot\cup W_i^{(p_i)} = W_i\}_{i=1}^{\ell}$. Note that as $|W_i| = \Theta(k)$ for all $i \in [\ell]$ then H has at most

$$\Theta(k^2) \cdot m \cdot \ell \leqslant \Theta(k^2) \cdot m \cdot \frac{n}{\Theta(k)} = \Theta(kn)$$

edges. Thus, when $k \ll n$, this is a regularization of a sparse graph. This "sparse regularity lemma" is very different to that of Kohayakawa and Rödl (see e.g. [Koh97]). Indeed, the Kohayakawa–Rödl regularity lemma only deals with graphs which have no local condensation of edges, e.g., subgraphs of random graphs.[1] Consequently, the resulting regular pairs are of density $o(1)$. In contrast, Lemma 2.5 provides us with regular pairs of density $\Theta(1)$, but, on the other hand, is useful only for graphs which are locally dense.

Lemma 2.5 (Regularity lemma for locally dense graphs)**.** *For all $m, z \in \mathbb{N}$ and $\varepsilon > 0$ there exists $q_{\mathrm{MAXCL}} \in \mathbb{N}$ such that the following is true. Suppose H and F are two graphs, $V(F) = [\ell]$ for some $\ell \in \mathbb{N}$, and $\mathrm{maxdeg}(F) \leqslant m$. Suppose that $\mathcal{Z} = \{Z_1, \ldots, Z_z\}$ is a partition of $V(H)$. Let $\{W_1, \ldots, W_\ell\}$ be a q_{MAXCL}-ensemble in H, such that for all $i, j \in [\ell]$ we have*

$$2|W_i| \geqslant |W_j| \,. \tag{2.2}$$

Then for each $i \in [\ell]$ there exists a partition $W_i^{(0)}, W_i^{(1)}, \ldots, W_i^{(p_i)}$ of the set W_i such that for all $i, j \in [\ell]$ we have

(a) $1/\varepsilon \leqslant p_i \leqslant q_{\mathrm{MAXCL}}$,

(b) $|W_i^{(i')}| = |W_j^{(j')}|$ for each $i' \in [p_i]$, $j' \in [p_j]$,

(c) for each $i' \in [p_i]$ there exists $x \in [z]$ such that $W_i^{(i')} \subseteq Z_x$,

---

[1]There is a recent refinement of the Kohayakawa–Rödl regularity lemma, due to Scott [Sco11]. Scott's regularity lemma gets around the no-condensation condition, which proves helpful in some situations, e.g. [AKSV14]; still, the main features remain.

(d) $\sum_i |W_i^{(0)}| < \varepsilon \sum_i |W_i|$, *and*

(e) at most $\varepsilon |\mathcal{Y}|$ pairs $\left(W_i^{(i')}, W_j^{(j')} \right) \in \mathcal{Y}$ form an ε-irregular pair in H, where

$$\mathcal{Y} := \left\{ \left(W_i^{(i')}, W_j^{(j')} \right) \; : \; ij \in E(F), i' \in [p_i], j' \in [p_j] \right\} \; .$$

We use Lemma 2.5 in the proof of Lemma 3.13. Lemma 3.13 is in turn the main tool in the proof of our main structural decomposition of the graph $G_{\text{T1.3}}$, Lemma 3.14. In the proof of Lemma 3.14 we decompose the input graph into several parts with very different properties, and one of these parts is a locally dense graph which can be then regularized by Lemma 3.13. A similar regularity lemma is used in [AKSS].

The proof of Lemma 2.5 is similar to the proof of the standard regularity lemma (Lemma 2.4), as given for example in [Sze78]. The key notion is that of the index (a.k.a. the mean square density) which we recall now. For us, it is convenient to work in the category of bipartite graphs.

Definition 2.6. *Suppose that $\mathcal{X} = \{X_0, X_1, \ldots, X_\ell\}$ and $\mathcal{Y} = \{Y_0, Y_1, \ldots, Y_p\}$ are partitions of a set X, and of Y with distinctive sets X_0 and Y_0 which we call* garbage clusters. *We use the symbol $\circ$ to indicate a new partition in which the garbage cluster is broken into singletons, e.g., $\mathcal{X}^\circ = \{X_1, \ldots, X_\ell\} \cup \{\{x\} : x \in X_0\}$. We say that $\mathcal{X}$ refines $\mathcal{Y}$ up to the garbage cluster if $\mathcal{X}^\circ$ refines $\mathcal{Y}^\circ$.*

Suppose that $G = (A, B; E)$ is a bipartite graph. Let $\mathcal{A} = \{A_0, A_1, \ldots, A_s\}$ and $\mathcal{B} = \{B_0, B_1, \ldots, B_t\}$ be partitions of A and B, with garbage clusters A_0 and B_0. We say that the pair $(\mathcal{A}, \mathcal{B})$ is an ε-regular partition of G if at most εst pairs (A_i, B_j), $i \in [s], j \in [t]$ are irregular. Otherwise, $(\mathcal{A}, \mathcal{B})$ is ε-irregular.

We then define the index *of $(\mathcal{A}, \mathcal{B})$ by*

$$\text{ind}(\mathcal{A}, \mathcal{B}) = \frac{1}{(|A| + |B|)^2} \cdot \sum_{X \in \mathcal{A}^\circ, Y \in \mathcal{B}^\circ} \frac{e(X, Y)^2}{|X||Y|} \; .$$

Clearly, $\text{ind}(\mathcal{A}, \mathcal{B}) \in [0, 1]$. Here is another fundamental property of the index.

Fact 2.7 (Bipartite version of Lemma 7.4.2 in [Die05])**.** *Suppose that $G = (A, B; E)$ is a bipartite graph. Let $\mathcal{A}$ and $\mathcal{A}'$ be partitions of A with given garbage clusters. Let $\mathcal{B}$ and $\mathcal{B}'$ be partitions of B with given garbage clusters. Suppose that $\mathcal{A}'$ refines $\mathcal{A}$ and $\mathcal{B}'$ refines $\mathcal{B}$ up to garbage clusters. Then $\text{ind}(\mathcal{A}', \mathcal{B}') \geqslant \text{ind}(\mathcal{A}, \mathcal{B})$.*

Lemma 2.8 (Index Pumping Lemma; bipartite version of Lemma 7.4.4. in [Die05])**.** *Let $\varepsilon \in (0, \frac{1}{4})$ and $p, q \in \mathbb{N}$. Let G be a bipartite graph $G = (A, B; E)$, with $\frac{|A|}{2} \leqslant |B| \leqslant 2|A|$. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are partitions of vertex sets A and B with distinctive garbage clusters A_0 and B_0. Suppose further that*

(a) $p \leqslant |\mathcal{A}|, |\mathcal{B}| \leqslant q$,

(b) $|A_0| < \varepsilon|A|$, $|B_0| < \varepsilon|B|$, and

(c) all the sets in $\mathcal{A} \cup \mathcal{B} \setminus \{A_0, B_0\}$ have the same size.

If $(\mathcal{A}, \mathcal{B})$ *is not an ε-regular partition of G then there exist partitions $\mathcal{A}'$ and $\mathcal{B}'$ of A and B with garbage clusters A_0' and B_0' such that*

(i) $p + 1 \leqslant |\mathcal{A}'|, |\mathcal{B}'| \leqslant 2q16^q$, *and*

(ii) $|A_0'| \leqslant |A_0| + \frac{|A|}{2^p}$, $|B_0'| \leqslant |B_0| + \frac{|B|}{2^p}$,

(iii) *all the sets in $\mathcal{A}' \cup \mathcal{B}' \setminus \{A_0', B_0'\}$ have the same size,*

(iv) *the partitions $\mathcal{A}'$ and $\mathcal{B}'$ refine $\mathcal{A}$ and $\mathcal{B}$ up to garbage clusters, and*

(v) $\operatorname{ind}(\mathcal{A}', \mathcal{B}') \geqslant \operatorname{ind}(\mathcal{A}, \mathcal{B}) + \frac{\varepsilon^5}{3691}$.

We note that by stating a version for bipartite graphs we had to adjust numerical values compared to [Die05]. Recall that the proof of Lemma 2.8 has two independent steps: first the partitions $\mathcal{A}$ and $\mathcal{B}$ are suitably refined (so that the index increases) and then these new partitions are further refined (up to garbage clusters) so that the non-garbage sets have the same size. The latter step does not decrease the index by Fact 2.7 but may possibly increase the sizes of the garbage clusters. Thus, we can state a version of Lemma 2.8 in which refinements are performed simultaneously on a number of bipartite graphs (referred to as $(G_i)_i$ in the corollary below), and in addition refines further partitions (referred to as $(C_j)_j$ below) on which no regularization is imposed.

Corollary 2.9. *Let $\varepsilon \in (0, \frac{1}{4})$ and $p, q \in \mathbb{N}$. Let G_i $i \in I$ be bipartite graphs $G_i = (A_i, B_i; E_i)$. Let C_j, $j \in J$ be sets of vertices. Suppose that all the sets A_i, B_i, and C_j are mutually disjoint. Suppose further that for each $i \in I$ and $j \in J$, $\max\{\frac{|A_i|}{2}, \frac{|B_i|}{2}\} \leqslant |C_j| \leqslant \min\{2|A_i|, 2|B_i|\}$. Suppose that $\mathcal{A}_i$ and $\mathcal{B}_i$ are partitions of A_i and B_i with garbage clusters A_{0i} and B_{0i}, and that $\mathcal{C}_j$ are partitions of C_j with garbage clusters C_{0j}. Suppose further that for each $i \in I$, and $j \in J$,*

(a) $p \leqslant |\mathcal{A}_i|, |\mathcal{B}_i|, |\mathcal{C}_j| \leqslant q$,

(b) $|A_{0i}| < \varepsilon|A_i|$, $|B_{0i}| < \varepsilon|B_i|$, $|C_{0j}| < \varepsilon|C_j|$, *and*

(c) *all the sets in $\bigcup_{m \in I}(\mathcal{A}_m \cup \mathcal{B}_m \setminus \{A_{0m}, B_{0m}\}) \cup \bigcup_{n \in J}(\mathcal{C}_n \setminus \{C_{0n}\})$ have the same size.*

If all partitions $(\mathcal{A}_i, \mathcal{B}_i)$, $i \in I$ are ε-irregular then there exist partitions $\mathcal{A}_i'$ and $\mathcal{B}_i'$ of A and B with garbage clusters A_{0i}' and B_{0i}', and partitions $\mathcal{C}_j'$ of C_j with garbage clusters C_{0j}' such that for each $i \in I$ and $j \in J$,

(i) $p + 1 \leqslant |\mathcal{A}_i'|, |\mathcal{B}_i'|, |\mathcal{C}_j'| \leqslant 2q \cdot 16^q$, *and*

(ii) $|A_{0i}'| \leqslant |A_{0i}| + \frac{|A_i|}{2^p}$, $|B_{0i}'| \leqslant |B_{0i}| + \frac{|B_i|}{2^p}$, *and* $|C_{0j}'| \leqslant |C_{0j}| + \frac{|C_j|}{2^p}$,

(iii) *all the sets in $\bigcup_{m \in I} \mathcal{A}_m' \cup \mathcal{B}_m' \setminus \{A_{0m}', B_{0m}'\} \cup \bigcup_{n \in J} \mathcal{C}_n' \setminus \{C_{0n}'\}$ have the same size,*

(iv) *the partitions $\mathcal{A}_i'$, $\mathcal{B}_i'$ and $\mathcal{C}_j'$ refine $\mathcal{A}_i$, $\mathcal{B}_i$ and $\mathcal{C}_j$ up to garbage clusters, and*

(v) $\operatorname{ind}(\mathcal{A}_i', \mathcal{B}_i') \geqslant \operatorname{ind}(\mathcal{A}_i, \mathcal{B}_i) + \frac{\varepsilon^5}{3691}$.

We are now in a position when we can prove Lemma 2.5. But before, let us describe how a more naive approach fails. For each edge $ij \in E(F)$ consider a regularization of the bipartite graph $H[W_i, W_j]$, let $\{U_{i,j}^{(i')}\}_{i' \in [q_{i,j}]}$ be the partition of W_i into clusters, and let $\{U_{j,i}^{(j')}\}_{j' \in [q_{j,i}]}$ be the partition of W_j into clusters such that almost all pairs $(U_{i,j}^{(i')}, U_{j,i}^{(j')}) \subseteq (W_i, W_j)$ form an ε'-regular pair (for some ε' of our taste). We would now be done if the partition $\{U_{i,j}^{(i')}\}_{i' \in [q_{i,j}]}$ of W_i was independent of the choice of the edge ij. This however need not be the case. The natural next step would therefore be to consider the common refinement

$$\underset{j:ij\in E(F)}{\boxplus} \left\{ U^{(i')i,j} \right\}_{i' \in [q_{ij}]}$$

of all the obtained partitions of W_i. The pairs obtained in this way lack however any regularity properties as they are too small. Indeed, it is a notorious drawback of the regularity lemma that the number of clusters in the partition is enormous as a function of the regularity parameter. In our setting, this means that $q_{i,j} \gg \frac{1}{\varepsilon'}$. Thus a typical cluster $U_{i,j_1}^{(i'_1)}$ occupies on average only a $\frac{1}{q_{i,j_1}}$-fraction of the cluster $U_{i,j_2}^{(i'_2)}$, and thus already the set $U_{i,j_1}^{(i'_1)} \cap U_{i,j_2}^{(i'_2)} \subseteq U_{i,j_2}^{(i'_2)}$ is not substantial (in the sense of the regularity). The same issue arises when regularizing multicoloured graphs (cf. [KS96, Theorem 1.18]). The solution is to impel the regularizations to happen in a synchronized way.

Proof of Lemma 2.5. Without loss of generality, assume that $\varepsilon < 1$. Set $\tilde{\varepsilon} = \varepsilon/8$. The number q_{MAXCL} can be taken by considering the function $q \mapsto 2q \cdot 16^q$ with initial value $\lceil \frac{4z}{\tilde{\varepsilon}} \rceil$ and iterating it $\lceil \frac{3691(m+1)}{\tilde{\varepsilon}^6} \rceil$-many times.

For each $i \in [\ell]$ consider an arbitrary initial partition $\mathcal{W}_i = \{W_i^{(0)}, W_i^{(1)}, \ldots, W_i^{(p_i)}\}$ of W_i such that for the garbage cluster we have $|W_i^{(0)}| \leqslant \tilde{\varepsilon}|W_i|$, all the non-garbage clusters $W_i^{(i')}$ are disjoint subsets of some set Z_r, $r \in [z]$. Further, we make all the non-garbage clusters (coming from all the sets W_i) have the same size. It is clear that this can be achieved, and that we can further impose that

$$1 + \frac{1}{\varepsilon} \leqslant p_i \leqslant \frac{4z}{\varepsilon} \ . \tag{2.3}$$

By Vizing's Theorem we can cover the edges of F by non-empty disjoint matchings $M_1, \ldots, M_Q$, $Q \leqslant m + 1$. For each $q \in [Q]$ we shall introduce a variable ind_q,

$$\text{ind}_q = \frac{1}{|M_q|} \sum_{xy \in M_q} \text{ind}(\mathcal{W}_x, \mathcal{W}_y) \ .$$

We shall now keep refining the partitions $(\mathcal{W}_z)_{z \in V(F)}$ in steps $\ell = 1, 2, \ldots$, as follows. Suppose that for some $q \in [Q]$, for the matching

$$M_q' = \{xy \in M_q : \text{the partition } (\mathcal{W}_x, \mathcal{W}_y) \text{ of the bipartite graph } H[W_x, W_y] \text{ is } \tilde{\varepsilon}\text{-irregular}\}$$

we have $|M_q'| \geqslant \tilde{\varepsilon}|M_q|$. We apply Corollary 2.9 with the following setting. The partitions $\{(\mathcal{W}_x, \mathcal{W}_y)\}_{xy \in M_q'}$ play the role of $(\mathcal{A}_i, \mathcal{B}_i)$'s and the partitions $\{\mathcal{W}_x\}_{x \in V(F) \setminus V(M_q')}$ play the role of $\mathcal{C}_j$'s. Further, in step ℓ we set $p_{\text{C2.9}} = \frac{1}{\varepsilon} + \ell$. Corollary 2.9 says that for the modified partitions (which we still denote the same), the index along each edge of M_q' increased by at least $\frac{\tilde{\varepsilon}^5}{3691}$. Combined with Fact 2.7 and with the fact that $|M_q'| \geqslant \tilde{\varepsilon}|M_q|$, we get that ind_q increased by at least

$\frac{\tilde{\varepsilon}^6}{3691}$, and none other index ind_t decreased. Observe also that the lower-bound in Corollary 2.9(i) makes it possible to apply Corollary 2.9 with $p_{C2.9}$ increased by one in a next step.

Since $\sum_{t=1}^{Q} \mathrm{ind}_t \leqslant |Q| \leqslant m+1$, we conclude that after at most $\frac{3691(m+1)}{\tilde{\varepsilon}^6}$ steps, for each $q \in [Q]$ the number of bipartite graphs $H[W_x, W_y]$, $xy \in M_q$ that are partitioned $\tilde{\varepsilon}$-irregularly is less than $\tilde{\varepsilon}|M_q|$. In particular, among all the bipartite graphs $(H[W_x, W_y])_{xy \in E(F)}$, at most $\tilde{\varepsilon} \cdot e(F)$ are partitioned $\tilde{\varepsilon}$-irregularly. We claim that this system of partitions satisfies Properties (a)–(e) as in the statement of the lemma.

Each irregular pair counted in (e) is a pair contained either in $\tilde{\varepsilon}$-regularly partitioned or an $\tilde{\varepsilon}$-irregularly partitioned bipartite graph $H[W_x, W_y]$. It follows from above that the number of irregular pairs of each of these two types is upper-bounded by $\frac{\varepsilon}{2}|\mathcal{Y}|$. For the bound (d), recall that initially we had $|W_i^{(0)}| \leqslant \frac{\varepsilon}{2}|W_i|$. During each application of Corollary 2.9, the garbage sets $W_i^{(0)}$ could have grown by at most $\frac{|W_i|}{2^p}$, for $p = \frac{1}{\varepsilon}+1, \frac{1}{\varepsilon}+2, \ldots$. Thus, at the end of the process, we have $|W_i^{(0)}| \leqslant (\frac{\varepsilon}{2} + \sum_{p > \frac{1}{\varepsilon}} 2^{-p})|W_i| \leqslant \varepsilon|W_i|$, as needed. The other assertions of the lemma are clear. $\qquad \square$

Usually after applying the regularity lemma to some graph G, one bounds the number of edges which correspond to irregular pairs, to regular, but sparse pairs, or are incident with the exceptional sets U_0. We shall do the same for the setting of Lemma 2.5.

Lemma 2.10. *In the situation of Lemma 2.5, suppose that $\mathrm{maxdeg}(H) \leqslant \Omega k$ and $e(H) \leqslant kn$, and that each edge $xy \in E(H)$ is captured by some edge $ij \in E(F)$, i.e., $x \in W_i$, $y \in W_j$. Moreover suppose that*

$$\mathrm{d}(W_i, W_j) \geqslant \gamma \text{ if } ij \in E(F). \tag{2.4}$$

Then all but at most $(\frac{4\varepsilon}{\gamma} + \varepsilon\Omega + \gamma)nk$ edges of H belong to regular pairs $(W_{i'}^{(i)}, W_{j'}^{(j)})$, $i, j \neq 0$, of density at least γ^2.

Proof. Set $w := \min\{|W_i| : i \in V(F)\}$. By (2.4), each edge of F represents at least γw^2 edges of H. Since $e(H) \leqslant kn$ it follows that $e(F) \leqslant kn/(\gamma w^2)$. Thus, by the assumption (2.2), $\sum_{AB \in E(F)} |A||B| \leqslant e(F)(2w)^2 \leqslant \frac{4kn}{\gamma}$. Using (e) of Lemma 2.5 we get that the number of edges of H contained in ε-irregular pairs from $\mathcal{Y}$ is at most

$$\frac{4\varepsilon nk}{\gamma} . \tag{2.5}$$

Write E_1 for the set of edges of H which are incident with a vertex in $\bigcup_{i \in [\ell]} W_i^{(0)}$. Then by (d) of Lemma 2.5, and since $\mathrm{maxdeg}(H) \leqslant \Omega k$,

$$|E_1| \leqslant \varepsilon\Omega nk . \tag{2.6}$$

Let E_2 be the set of those edges of H which belong to ε-regular pairs $(W_i^{(i')}, W_j^{(j')})$ with $ij \in E(F), i' \in [p_i], j' \in [p_j]$ of density at most γ^2. We claim that

$$|E_2| \leqslant \gamma kn . \tag{2.7}$$

Indeed, because of (2.4) and by Fact 2.3 (with $\alpha_{F2.3} := \gamma$ and $\beta_{F2.3} := \gamma^2$), for each $ij \in E(F)$ there are at most $\gamma e_H(W_i, W_j)$ edges contained in the bipartite graphs $H[W_i^{(i')}, W_j^{(j')}]$, $i' \in [p_i], j' \in [p_j]$, with $\mathrm{d}_H(W_i^{(i')}, W_j^{(j')}) \leqslant \gamma^2$. Since $\sum_{ij \in E(F)} e_H(W_i, W_j) \leqslant kn$, the validity of (2.7) follows. Combining (2.5), (2.6), and (2.7) we finish the proof. $\qquad \square$

2.4 LKS graphs

Write **LKS**(n, k, α) for the class of all n-vertex graphs with at least $(\frac{1}{2} + \alpha)n$ vertices of degrees at least $(1 + \alpha)k$. With this notation, Conjecture 1.2 states that every graph in **LKS**$(n, k, 0)$ contains every tree from **trees**$(k + 1)$.

Given a graph G, denote by $\mathbb{S}_{\eta,k}(G)$ the set of those vertices of G that have degree less than $(1 + \eta)k$ and by $\mathbb{L}_{\eta,k}(G)$ the set of those vertices of G that have degree at least $(1 + \eta)k$.[2] Thus the sizes of the sets $\mathbb{S}_{\eta,k}(G)$ and $\mathbb{L}_{\eta,k}(G)$ are what specifies the membership to **LKS**(n, k, η).

Define **LKSmin**(n, k, η) as the set of all graphs $G \in$ **LKS**(n, k, η) that are edge-minimal with respect to the membership in **LKS**(n, k, η). In order to prove Theorem 1.3 it suffices to restrict our attention to graphs from **LKSmin**(n, k, η), and this is why we introduce the class. Let us collect some properties of graphs in **LKSmin**(n, k, η).

Fact 2.11. *For any graph $G \in$ **LKSmin**(n, k, η) the following is true.*

1. *$\mathbb{S}_{\eta,k}(G)$ is an independent set.*

2. *All the neighbours of every vertex $v \in V(G)$ with $\deg(v) > \lceil (1 + \eta)k \rceil$ have degree exactly $\lceil (1 + \eta)k \rceil$.*

3. *$|\mathbb{L}_{\eta,k}(G)| \leqslant \lceil (1/2 + \eta)n \rceil + 1$.*

Observe that every edge in a graph $G \in$ **LKSmin**(n, k, η) is incident to at least one vertex of degree exactly $\lceil (1 + \eta)k \rceil$. This gives the following inequality.

$$e(G) \leqslant \lceil (1 + \eta)k \rceil |\mathbb{L}_{\eta,k}(G)| \overset{\text{F2.11(3.)}}{\leqslant} \lceil (1 + \eta)k \rceil \left(\left\lceil \left(\frac{1}{2} + \eta \right) n \right\rceil + 1 \right) < kn . \tag{2.8}$$

(The last inequality is valid under the additional mild assumption that, say, $\eta < \frac{1}{20}$ and $n > k > 20$. This can be assumed throughout the paper.)

Definition 2.12. *Let **LKSsmall**(n, k, η) be the class of those graphs $G \in$ **LKS**(n, k, η) for which we have the following three properties:*

1. *All the neighbours of every vertex $v \in V(G)$ with $\deg(v) > \lceil (1 + 2\eta)k \rceil$ have degrees at most $\lceil (1 + 2\eta)k \rceil$.*

2. *All the neighbours of every vertex of $\mathbb{S}_{\eta,k}(G)$ have degree exactly $\lceil (1 + \eta)k \rceil$.*

3. *We have $e(G) \leqslant kn$.*

Observe that the graphs from **LKSsmall**(n, k, η) also satisfy 1., and a quantitatively somewhat weaker version of 2. of Fact 2.11. This suggests that in some sense **LKSsmall**(n, k, η) is a good approximation of **LKSmin**(n, k, η).

As said, we will prove Theorem 1.3 only for graphs from **LKSmin**(n, k, η). However, it turns out that the structure of **LKSmin**(n, k, η) is too rigid. In particular, **LKSmin**(n, k, η) is not closed under discarding a small amount of edges during our cleaning procedures. This is why the class **LKSsmall**(n, k, η) comes into play: starting with a graph in **LKSmin**(n, k, η) we perform some initial cleaning and obtain a graph which lies in **LKSsmall**$(n, k, \eta/2)$. We then heavily use its structural properties from Definition 2.12 throughout the proof.

---

[2] "$\mathbb{S}$" stands for "small", and "$\mathbb{L}$" for "large".

3 Decomposing sparse graphs

In this section, we work out a structural decomposition of a possibly sparse graph which is suitable for embedding trees. Our motivation comes from the success of the regularity method in the setting of dense graphs (see [KO09]). The main technical result of this section, the "decomposition lemma", Lemma 3.13, provides such a decomposition. Roughly speaking, each graph of a moderate maximum degree can be decomposed into regular pairs, and two different expanding parts.

We then combine Lemma 3.13 with a lemma on creating a gap in the degree sequence (Lemma 3.2) to get a decomposition lemma for graphs from $\mathbf{LKS}(n, k, \eta)$, Lemma 3.14. Lemma 3.14 asserts that each graph from $\mathbf{LKS}(n, k, \eta)$ can be decomposed into vertices of degree much larger than k, regular pairs, and expanding parts. Further we give a non-LKS-specific version of Lemma 3.14 in Lemma 3.15, which asserts that *each* graph with average degree bigger than an absolute constant has a sparse decomposition. Such a decomposition lemma was used by Ajtai, Komlós, Simonovits and Szemerédi in their work on the Erdős–Sós conjecture and we expect that it will find applications in other tree embedding problems, and possibly elsewhere.

3.1 Creating a gap in the degree sequence

The goal of this section is to show that any graph $G \in \mathbf{LKSmin}(n, k, \eta)$ has a subgraph $G' \in \mathbf{LKSsmall}(n, k, \eta/2)$ which has a gap in its degree sequence. Note that G' then contains almost all the edges of G. This is formulated in Lemma 3.2. Before stating and proving it, we illustrate our proof technique on a simpler version of Lemma 3.2 that applies to all graphs. This simpler lemma will not be used except in the proof of Lemma 3.15 which serves also for illustration only.

Lemma 3.1. *Let $(\Omega_i)_{i \in \mathbb{N}}$ be a sequence of positive numbers with $\frac{\Omega_j}{\Omega_{j+1}} \leqslant \frac{\eta}{2}$ for all $j \in \mathbb{N}$. Let G be a graph of order n with average degree k. Then there is an index $i^* \leqslant \frac{4}{\eta}$ and a spanning subgraph $G' \subseteq G$ with $e(G') \geqslant e(G) - \eta k n$ and with the property that G' contains no vertex with degree in the interval $[\Omega_i k, \Omega_{i+1} k)$.*

Proof. Set $R := \lfloor 4\eta^{-1} \rfloor$. For $i \in [R]$ and any graph $H \subseteq G$ define the sets $X_i(H) := \{v \in V(H) : \deg_H(v) \in [\Omega_i k, \Omega_{i+1} k)\}$ and for $i = R + 1$ set $X_i(H) := \{v \in V(H) : \deg_H(v) \in [\Omega_i k, \infty)\}$. As

$$\sum_{i \in [R]} \sum_{v \in X_i(G) \cup X_{i+1}(G)} \deg(v) \leqslant 4e(G) \ ,$$

by averaging we find an index $i^* \in [R]$ such that

$$\sum_{v \in X_{i^*}(G) \cup X_{i^*+1}(G)} \deg(v) \leqslant \frac{4e(G)}{R} = \frac{2kn}{R}.$$

Let $G_0 \subseteq G$ be obtained from G by deleting all the edges incident with $X_{i^*}(G) \cup X_{i^*+1}(G)$. In particular,

$$e(G_0) \geqslant e(G) - \eta k n/2 \ . \tag{3.1}$$

We continue successively deleting edges as follows. If in some step $j = 1, 2, \dots$ the set $X_{i^*}(G_{j-1})$ is non-empty, we take an arbitrary vertex $v_j \in X_{i^*}(G_{j-1})$ and obtain a new graph G_j from G_{j-1} by deleting all the (at most $\Omega_{i^*+1} k$ many) edges incident with v_j. Obviously, this procedure will

terminate eventually. Let G' denote the final graph. Clearly, G' has the desired gap in the degree sequence. It therefore suffices to upper bound $e(G) - e(G')$.

Observe that for any vertex v_j above we have $v_j \in \bigcup_{i=i^*+2}^{R+1} X_i(G)$. Thus,

$$e(G') - e(G_0) \leqslant \Omega_{i^*+1}k \left| \bigcup_{i=i^*+2}^{R+1} X_i(G) \right| \leqslant \Omega_{i^*+1}k \cdot \frac{2e(G)}{\Omega_{i^*+2}k} \leqslant \frac{\eta k n}{2} \ .$$

Combined with (3.1) we get the statement. $\qquad\square$

Lemma 3.2. *Let $\eta \in (0,1)$, $G \in \mathbf{LKSmin}(n,k,\eta)$ and let $(\Omega_i)_{i \in \mathbb{N}}$ be a sequence of positive numbers with $\Omega_1 > 2$ and $\Omega_j/\Omega_{j+1} \leqslant \eta^2/100$ for all $j \in \mathbb{N}$. Then there exist an index $i^* \leqslant 100\eta^{-2}$ and a subgraph $G' \subseteq G$ such that*

(i) $G' \in \mathbf{LKSsmall}(n,k,\eta/2)$, and

(ii) no vertex $v \in V(G')$ has degree $\deg_{G'}(v) \in [\Omega_{i^*}k, \Omega_{i^*+1}k)$.

Proof. Set $R := \lfloor 100\eta^{-2} \rfloor$. For $i \in [R]$ and any graph $H \subseteq G$ define the sets $X_i(H) := \{v \in V(H) : \deg_H(v) \in [\Omega_i k, \Omega_{i+1}k)\}$ and for $i = R+1$ set $X_i(H) := \{v \in V(H) : \deg_H(v) \in [\Omega_i k, \infty)\}$. As

$$\sum_{i \in [R]} \sum_{v \in X_i(G) \cup X_{i+1}(G)} \deg(v) \leqslant 4e(G) \ ,$$

by averaging we find an index $i^* \in [R]$ such that

$$\sum_{v \in X_{i^*}(G) \cup X_{i^*+1}(G)} \deg(v) \leqslant \frac{4e(G)}{R}. \tag{3.2}$$

Let E_0 be the set of all the edges incident with $X_{i^*}(G) \cup X_{i^*+1}(G)$. Now, starting with $G_0 := G - E_0$, inductively define graphs $G_j \subsetneq G_{j-1}$ for $j \geqslant 1$ using any of the following two types of edge deletions:

(T1) If there is a vertex $v_j \in X_{i^*}(G_{j-1})$ then we choose an edge e_j incident with v_j, and set $G_j := G_{j-1} - e_j$.

(T2) If there is an edge $e_j = u_j v_j$ of G_{j-1} with $u_j \in \mathbb{S}_{\eta/2,k}(G_{j-1})$ and $v_j \in \bigcup_{i=i^*+1}^{R+1} X_i(G_{j-1})$ then set $G_j := G_{j-1} - e_j$.

Since we keep deleting edges, the procedure stops at some point, say at step j^*, when neither of (T1), (T2) is applicable. Note that the resulting graph G_{j^*} already has Property (ii).

Let $E_1 \subseteq E(G)$ be the set of those edges deleted by applying (T1). We shall estimate the size of E_1. First, observe that

$$\left| \bigcup_{i=i^*+2}^{R+1} X_i(G) \right| \leqslant \frac{2e(G)}{\Omega_{i^*+2}k} \ .$$

Moreover, each vertex of $\bigcup_{i=i^*+2}^{R+1} X_i(G)$ appears at most $(\Omega_{i^*+1} - \Omega_{i^*})k < \Omega_{i^*+1}k$ times as the vertex v_j in the deletions of type (T1). Consequently,

$$|E_1| \leqslant \Omega_{i^*+1} \left| \bigcup_{i=i^*+2}^{R+1} X_i(G) \right| k \leqslant \frac{2\Omega_{i^*+1}e(G)}{\Omega_{i^*+2}} \ . \tag{3.3}$$

Consider an arbitrary vertex $w \in \mathbb{L}_{\eta,k}(G) \cap \mathbb{S}_{\eta/2,k}(G_{j^*})$ and the interval of those $(j-1)$'s for which $w \in \mathbb{L}_{\eta/2,k}(G_{j-1}) \cap \mathbb{S}_{\eta,k}(G_{j-1})$. In such a step the vertex w cannot play the role of the vertices u_j or v_j in (T2). So, each vertex from $\mathbb{L}_{\eta,k}(G) \cap \mathbb{S}_{\eta/2,k}(G_{j^*})$ is incident with at least $\eta k/2$ edges from the set $E_0 \cup E_1$. Therefore, by the definition of E_0, by (3.2), and by (3.3),

$$\left|\mathbb{L}_{\eta,k}(G) \cap \mathbb{S}_{\eta/2,k}(G_{j^*})\right| \leqslant \frac{2 \cdot |E_0 \cup E_1|}{\eta k/2} \leqslant \left(\frac{4}{R} + \frac{2\Omega_{i^*+1}}{\Omega_{i^*+2}}\right) \cdot \frac{4e(G)}{\eta k} \overset{(2.8)}{\leqslant} \frac{\eta n}{2} \ .$$

Thus

$$|\mathbb{L}_{\eta/2,k}(G_{j^*})| \geqslant |\mathbb{L}_{\eta,k}(G)| - |\mathbb{L}_{\eta,k}(G) \cap \mathbb{S}_{\eta/2,k}(G_{j^*})| \geqslant (1/2 + \eta/2)n \ ,$$

and consequently, $G_{j^*} \in \mathbf{LKS}(n, k, \eta/2)$.

Last, we obtain the graph G' by successively deleting any edge from G_{j^*} which connects a vertex from $\mathbb{S}_{\eta/2,k}(G_{j^*})$ with a vertex whose degree is not exactly $\lceil(1+\frac{\eta}{2})k\rceil$. This does not affect the already obtained Property (ii), since we could not apply (T2) to G_{j^*}. We claim that for the resulting graph G' we have $G' \in \mathbf{LKSsmall}(n, k, \eta/2)$. Indeed, $\mathbb{L}_{\eta/2,k}(G') = \mathbb{L}_{\eta/2,k}(G_{j^*})$, and thus $G' \in \mathbf{LKS}(n, k, \eta/2)$. Property 2 of Definition 2.12 follows from the last step of the construction of G'. To see Property 1 of Definition 2.12 we use Fact 2.11(2) for G (which by assumption is in $\mathbf{LKSmin}(n, k, \eta)$). $\qquad\square$

3.2   Decomposition of graphs with moderate maximum degree

First we introduce some useful notions. We start with dense spots which indicate an accumulation of edges in a sparse graph.

Definition 3.3 ((m,γ)-dense spot, (m,γ)-nowhere-dense). *An (m,γ)-dense spot in a graph G is a non-empty bipartite subgraph $D = (U, W; F)$ of G with $\mathrm{d}(D) > \gamma$ and $\mathrm{mindeg}(D) > m$. We call G (m,γ)-nowhere-dense if it does not contain any (m,γ)-dense spot.*

We remark that dense spots as bipartite graphs do not have a specified orientation, that is, we view $(U, W; F)$ and $(W, U; F)$ as the same object.

Fact 3.4. *Let $(U, W; F)$ be a $(\gamma k, \gamma)$-dense spot in a graph G of maximum degree at most Ωk. Then $\max\{|U|, |W|\} \leqslant \frac{\Omega}{\gamma}k$.*

Proof. It suffices to observe that

$$\gamma|U||W| \leqslant e(U, W) \leqslant \mathrm{maxdeg}(G) \cdot \min\{|U|, |W|\} \leqslant \Omega k \cdot \min\{|U|, |W|\}.$$

$\qquad\square$

The next fact asserts that in a bounded degree graph there cannot be too many edge-disjoint dense spots containing a given vertex.

Fact 3.5. *Let H be a graph of maximum degree at most Ωk, let $v \in V(H)$, and let $\mathcal{D}$ be a family of edge-disjoint $(\gamma k, \gamma)$-dense spots in H. Then less than $\frac{\Omega}{\gamma}$ dense spots from $\mathcal{D}$ contain v.*

Proof. This follows as v sends more than γk edges to each dense spot from $\mathcal{D}$ it is incident with, the dense spots $\mathcal{D}$ are edge-disjoint, and $\deg(v) \leqslant \Omega k$. $\qquad\square$

Our second definition of this section might seem less intuitive at first sight. It describes a property for finding dense spots outside some "forbidden" set U, which in later applications will be the set of vertices already used for a partial embedding of a tree $T \in \mathbf{trees}(k)$ from Theorem 1.3 during our sequential embedding procedure. In Section 3.5 we give a non-technical description of this embedding technique. Informally, a set $\mathbb{E}$ of vertices is avoiding if for each set U of size $\Theta(k)$ and each vertex $v \in \mathbb{E}$ there is a dense spot containing v that is almost disjoint from U.

Definition 3.6 $((\Lambda, \varepsilon, \gamma, k)$**-avoiding set**$)$**.** *Suppose that G is a graph and $\mathcal{D}$ is a family of dense spots in G. A set $\mathbb{E} \subseteq \bigcup_{D \in \mathcal{D}} V(D)$ is $(\Lambda, \varepsilon, \gamma, k)$-avoiding with respect to $\mathcal{D}$ if for every $U \subseteq V(G)$ with $|U| \leqslant \Lambda k$ the following holds for all but at most εk vertices $v \in \mathbb{E}$. There is a dense spot $D \in \mathcal{D}$ with $|U \cap V(D)| \leqslant \gamma^2 k$ that contains v.*

Note that a subset of a $(\Lambda, \varepsilon, \gamma, k)$-avoiding set is also $(\Lambda, \varepsilon, \gamma, k)$-avoiding.

We now come to the main concepts of this section, the bounded and the sparse decompositions. These notions in a way correspond to the partition structure from the regularity lemma, although naturally more complex since we deal with (possibly) sparse graphs here. Lemma 3.13 is then a corresponding regularization result.

Definition 3.7 $((k, \Lambda, \gamma, \varepsilon, \nu, \rho)$**-bounded decomposition**$)$**.** *Suppose that $k \in \mathbb{N}$ and $\varepsilon, \gamma, \nu, \rho > 0$ and $\Lambda > 2$. Let $\mathcal{V} = \{V_1, V_2, \ldots, V_s\}$ be a partition of the vertex set of a graph G. We say that $(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ is a $(k, \Lambda, \gamma, \varepsilon, \nu, \rho)$-bounded decomposition of G with respect to $\mathcal{V}$ if the following properties are satisfied:*

1. *G_{exp} is a $(\gamma k, \gamma)$-nowhere-dense subgraph of G with $\mathrm{mindeg}(G_{\mathrm{exp}}) > \rho k$.*

2. *The elements of $\mathbf{V}$ are disjoint subsets of $V(G)$.*

3. *G_{reg} is a subgraph of $G - G_{\mathrm{exp}}$ on the vertex set $\bigcup \mathbf{V}$. For each edge $xy \in E(G_{\mathrm{reg}})$ there are distinct $C_x \ni x$ and $C_y \ni y$ from $\mathbf{V}$, and $G[C_x, C_y] = G_{\mathrm{reg}}[C_x, C_y]$. Furthermore, $G[C_x, C_y]$ forms an ε-regular pair of density at least γ^2.*

4. *We have $\nu k \leqslant |C| = |C'| \leqslant \varepsilon k$ for all $C, C' \in \mathbf{V}$.*

5. *$\mathcal{D}$ is a family of edge-disjoint $(\gamma k, \gamma)$-dense spots in $G - G_{\mathrm{exp}}$. For each $D = (U, W; F) \in \mathcal{D}$ all the edges of $G[U, W]$ are covered by $\mathcal{D}$ (but not necessarily by D).*

6. *If G_{reg} contains at least one edge between $C_1, C_2 \in \mathbf{V}$ then there exists a dense spot $D = (U, W; F) \in \mathcal{D}$ for which $C_1 \subseteq U$ and $C_2 \subseteq W$.*

7. *For each $C \in \mathbf{V}$ there is a $V \in \mathcal{V}$ so that either $C \subseteq V \cap V(G_{\mathrm{exp}})$ or $C \subseteq V \setminus V(G_{\mathrm{exp}})$. For each $C \in \mathbf{V}$ and $D = (U, W; F) \in \mathcal{D}$ we have $C \cap U, C \cap W \in \{\emptyset, C\}$.*

8. *$\mathbb{E}$ is a $(\Lambda, \varepsilon, \gamma, k)$-avoiding subset of $V(G) \setminus \bigcup \mathbf{V}$ with respect to dense spots $\mathcal{D}$.*

We say that the bounded decomposition $(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ respects the avoiding threshold b if for each $C \in \mathbf{V}$ we either have $\mathrm{maxdeg}_G(C, \mathbb{E}) \leqslant b$, or $\mathrm{mindeg}_G(C, \mathbb{E}) > b$.

Here "exp" in G_{exp} stands for "expander" and "reg" in G_{reg} stands for "regular(ity)".

The members of $\mathbf{V}$ are called *clusters*. Define the *cluster graph* $\mathbf{G}_{\mathrm{reg}}$ as the graph on the vertex set $\mathbf{V}$ that has an edge $C_1 C_2$ for each pair (C_1, C_2) which has density at least γ^2 in the graph G_{reg}.

Property 7 tells us that the clusters may be prepartitioned, just as it is the case in the classic regularity lemma. When in Lemma 3.14 below we classify the graph G from Theorem 1.3 we shall use the prepartition into (roughly) $\mathbb{S}_{\alpha_{\mathrm{T1.3}},k}(G)$ and $\mathbb{L}_{\alpha_{\mathrm{T1.3}},k}(G)$.

As said above, the notion of bounded decomposition is needed for our regularity lemma type decomposition given in Lemma 3.13. It turns out that such a decomposition is possible only when the graph is of moderate maximum degree. On the other hand, Lemma 3.1 tells us that the vertex set of any graph can be decomposed into vertices of enormous degree and moderate degree. The graph induced by the latter type of vertices then admits the decomposition from Lemma 3.13. Thus, it makes sense to enhance the structure of bounded decomposition by vertices of unbounded degree. This is done in the next definition.

Definition 3.8 $((k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$**-sparse decomposition**)**.** *Suppose that $k \in \mathbb{N}$ and $\varepsilon, \gamma, \nu, \rho > 0$ and $\Lambda, \Omega^*, \Omega^{**} > 2$. Let $\mathcal{V} = \{V_1, V_2, \ldots, V_s\}$ be a partition of the vertex set of a graph G. We say that $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ is a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition of G with respect to $V_1, V_2, \ldots, V_s$ if the following holds.*

1. *$\mathbb{H} \subseteq V(G)$, $\mathrm{mindeg}_G(\mathbb{H}) \geqslant \Omega^{**}k$, $\mathrm{maxdeg}_H(V(G) \setminus \mathbb{H}) \leqslant \Omega^*k$, where H is spanned by the edges of $\bigcup \mathcal{D}$, G_{exp}, and edges incident with $\mathbb{H}$,*

2. *$(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ is a $(k, \Lambda, \gamma, \varepsilon, \nu, \rho)$-bounded decomposition of $G - \mathbb{H}$ with respect to $V_1 \setminus \mathbb{H}, V_2 \setminus \mathbb{H}, \ldots, V_s \setminus \mathbb{H}$.*

If the parameters do not matter, we call ∇ simply a *sparse decomposition*, and similarly we speak about a *bounded decomposition*.

Definition 3.9 (**captured edges**)**.** *In the situation of Definition 3.8, we refer to the edges in $E(G_{\mathrm{reg}}) \cup E(G_{\mathrm{exp}}) \cup E_G(\mathbb{H}, V(G)) \cup E_{G_{\mathcal{D}}}(\mathbb{E}, \mathbb{E} \cup \bigcup \mathbf{V})$ as* captured *by the sparse decomposition. We write G_∇ for the subgraph of G on the same vertex set which consists of the captured edges. Likewise, the captured edges of a bounded decomposition $(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ of a graph G are those in $E(G_{\mathrm{reg}}) \cup E(G_{\mathrm{exp}}) \cup E_{G_{\mathcal{D}}}(\mathbb{E}, \mathbb{E} \cup \bigcup \mathbf{V})$.*

Throughout the paper we write $G_{\mathcal{D}}$ for the subgraph of G which consists of the edges contained in $\mathcal{D}$. We now include an easy fact about the relation of $G_{\mathcal{D}}$ and G_{reg}.

Fact 3.10. *Let $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ be a sparse decomposition of a graph G. Then each edge $xy \in E(G_{\mathcal{D}})$ with $x, y \in \bigcup \mathbf{V}$ is either contained in G_{reg}, or is not captured.*

Proof. Indeed, suppose that $xy \in E(G_{\mathcal{D}})$, $x, y \in \bigcup \mathbf{V}$, and $xy \notin E(G_{\mathrm{reg}})$. Property 2 of Definition 3.8 says that $x, y \notin \mathbb{H}$. Further, by Property 8 of Definition 3.7, we have $x, y \notin \mathbb{E}$. Last, Property 5 of Definition 3.7 implies that $xy \notin E(G_{\mathrm{exp}})$. Hence xy is not captured, as desired. $\square$

We now give a bound on the number of clusters reachable through edges of the dense spots from a fixed vertex outside $\mathbb{H}$.

Fact 3.11. *Let $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ be a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition of a graph G. Let $x \in V(G) \setminus \mathbb{H}$. Assume that $\mathbf{V} \neq \emptyset$, and let $\mathfrak{c}$ be the size of each of the members of $\mathbf{V}$. Then there are less than*

$$\frac{2(\Omega^*)^2 k}{\gamma^2 \mathfrak{c}} \leqslant \frac{2(\Omega^*)^2}{\gamma^2 \nu}$$

clusters $C \in \mathbf{V}$ with $\deg_{G_{\mathcal{D}}}(x, C) > 0$.

Proof. Property 1 of Definition 3.8 says that $\deg_{G_{\mathcal{D}}}(x) \leqslant \Omega^* k$. For each $D \in \mathcal{D}$ with $x \in V(D)$ we have that $\deg_D(x) > \gamma k$, since D is a $(\gamma k, \gamma)$-dense spot. By Fact 3.5

$$|\{D \in \mathcal{D} : \deg_D(x) > 0\}| < \frac{\Omega^*}{\gamma}. \tag{3.4}$$

Furthermore, by Fact 3.4, and using Properties 4 and 6 of Definition 3.7, we see that for a fixed $D \in \mathcal{D}$, we have

$$|\{C \in \mathbf{V} \: : \: C \subseteq V(D)\}| \leqslant \frac{2\Omega^* k}{\gamma} \cdot \frac{1}{\mathfrak{c}} \leqslant \frac{2\Omega^*}{\gamma\nu} \ .$$

Together with (3.4) this gives that the number of clusters $C \in \mathbf{V}$ with $\deg_{G_{\mathcal{D}}}(x, C) > 0$ is less than

$$\frac{\Omega^*}{\gamma} \cdot \frac{2\Omega^* k}{\gamma\mathfrak{c}} \leqslant \frac{\Omega^*}{\gamma} \cdot \frac{2\Omega^*}{\gamma\nu} \ ,$$

as desired. $\qquad\square$

As a last step before we state the main result of this section we show that the cluster graph $\mathbf{G}_{\mathrm{reg}}$ corresponding to a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ has bounded degree.

Fact 3.12. *Let $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ be a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition of a graph G, and let $\mathbf{G}_{\mathrm{reg}}$ be the corresponding cluster graph. Let $\mathfrak{c}$ be the size of each cluster in $\mathbf{V}$. Then $\mathrm{maxdeg}(\mathbf{G}_{\mathrm{reg}}) \leqslant \frac{\Omega^* k}{\gamma^2 \mathfrak{c}} \leqslant \frac{\Omega^*}{\gamma^2 \nu}$.*

Proof. Let $C \in \mathbf{V}$. Then by the definition of $\mathbf{G}_{\mathrm{reg}}$ and by Property 3 of Definition 3.7 we have $\deg_{\mathbf{G}_{\mathrm{reg}}}(C) \leqslant \sum_{C' \in \mathrm{N}_{\mathbf{G}_{\mathrm{reg}}}(C)} \frac{e_{G_{\mathrm{reg}}}(C, C')}{\gamma^2 |C||C'|} = \sum_{C' \in \mathrm{N}_{\mathbf{G}_{\mathrm{reg}}}(C)} \frac{e_{G_{\mathrm{reg}}}(C, C')}{\gamma^2 |C|\mathfrak{c}}$. Since the maximum degree in G_{reg} is upper-bounded by $\Omega^* k$ (c.f. Property 1 of Definition 3.8), we get

$$\deg_{\mathbf{G}_{\mathrm{reg}}}(C) \leqslant \sum_{C' \in \mathrm{N}_{\mathbf{G}_{\mathrm{reg}}}(C)} \frac{e_{G_{\mathrm{reg}}}(C, C')}{\gamma^2 |C|\mathfrak{c}} \leqslant \frac{\Omega^* k |C|}{\gamma^2 |C|\mathfrak{c}} \overset{\mathrm{D3.7(4)}}{\leqslant} \frac{\Omega^*}{\gamma^2 \nu} \ ,$$

as desired. $\qquad\square$

We now state the most important lemma of this section. It says that any graph of bounded degree has a bounded decomposition which captures almost all its edges. This lemma can be considered as a sort of regularity lemma for sparse graphs.

Lemma 3.13 (Decomposition lemma)**.** *For each $\Lambda, \Omega, s \in \mathbb{N}$ and each $\gamma, \varepsilon, \rho > 0$ there exist $k_0 \in \mathbb{N}$, $\nu > 0$ such that for every $k \geqslant k_0$ and every n-vertex graph G with $e(G) \leqslant kn$, $\mathrm{maxdeg}(G) \leqslant \Omega k$, and with a given partition $\mathcal{V}$ of its vertex set into at most s sets, the following holds for each $b > 0$. There exists a $(k, \Lambda, \gamma, \varepsilon, \nu, \rho)$-bounded decomposition $(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ with respect to $\mathcal{V}$, which captures all but at most $(\frac{4\varepsilon}{\gamma} + \varepsilon\Omega + \gamma + \rho)kn$ edges of G and respects avoiding threshold b. Furthermore, we have*

$$|E(\mathcal{D}) \setminus (E(G_{\mathrm{reg}}) \cup E_{G_{\mathcal{D}}}[\mathbb{E}, \mathbb{E} \cup \bigcup \mathbf{V}])| \leqslant (\frac{4\varepsilon}{\gamma} + \varepsilon\Omega + \gamma)kn \ . \tag{3.5}$$

A proof of Lemma 3.13 is given in Section 3.7.

3.3   Decomposition of LKS graphs

Lemma 3.2 and Lemma 3.13 enable us to decompose graphs in $\mathbf{LKS}(n, k, \eta)$ in a particular manner.

Lemma 3.14. *For every $\eta, \Lambda, \gamma, \varepsilon, \rho \in (0, 1)$ there are $\nu > 0$ and $k_0 \in \mathbb{N}$ such that for every $k > k_0$ and for every number $b > 0$ the following holds. For every sequence $(\Omega_j)_{j \in \mathbb{N}}$ of positive numbers with $\Omega_1 > 2$, $\Omega_j / \Omega_{j+1} \leqslant \eta^2/100$ for all $j \in \mathbb{N}$ and for every $G \in \mathbf{LKS}(n, k, \eta)$ there are an index i and a subgraph G' of G with the following properties:*

(a) $G' \in \mathbf{LKSsmall}(n, k, \eta/2)$,

(b) $i \leqslant 100\eta^{-2}$,

(c) G' *has a* $(k, \Omega_{i+1}, \Omega_i, \Lambda, \gamma, \varepsilon, \nu, \rho)$*-sparse decomposition* $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G'_{\text{reg}}, G'_{\text{exp}}, \mathbb{E})$ *with respect to the partition* $\{V_1, V_2\} := \{\mathbb{S}_{\eta/2, k}(G'), \mathbb{L}_{\eta/2, k}(G')\}$*, and with respect to avoiding threshold b,*

(d) $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G'_{\text{reg}}, G'_{\text{exp}}, \mathbb{E})$ *captures all but at most* $(\frac{4\varepsilon}{\gamma} + \varepsilon\Omega_{\lfloor 100\eta^{-2}\rfloor} + \gamma + \rho)kn$ *edges of G', and*

(e) $|E(\mathcal{D}) \setminus (E(G'_{\text{reg}}) \cup E_{G'}[\mathbb{E}, \mathbb{E} \cup \bigcup \mathbf{V}])| \leqslant (\frac{4\varepsilon}{\gamma} + \varepsilon\Omega_{\lfloor 100\eta^{-2}\rfloor} + \gamma)kn$.

Proof. Let ν and k_0 be given by Lemma 3.13 for input parameters $\Omega_{\text{L3.13}} := \Omega_{\lfloor 100\eta^{-2}\rfloor}$, $\Lambda_{\text{L3.13}} := \Lambda, \gamma_{\text{L3.13}} := \gamma, \varepsilon_{\text{L3.13}} := \varepsilon, \rho_{\text{L3.13}} := \rho, b_{\text{L3.13}} := b$, and $s_{\text{L3.13}} := 2$. Now, given G, let us consider a subgraph $\tilde{G}$ of G such that $\tilde{G} \in \mathbf{LKSmin}(n, k, \eta)$. Lemma 3.2 applied to the sequence $(\Omega_j)_j$ and $\tilde{G}$ yields a graph $G' \in \mathbf{LKSsmall}(n, k, \eta/2)$ and an index $i \leqslant 100\eta^{-2}$. We set $\mathbb{H} := \{v \in V(G) : \deg_{G'}(v) \geqslant \Omega_{i+1}k\}$.

Observe that by (2.8), $e(G') < kn$. Let $(\mathbb{H}, \mathcal{D}, G'_{\text{reg}}, G'_{\text{exp}}, \mathbb{E})$ be the $(k, \Lambda, \gamma, \varepsilon, \nu, \rho)$-bounded decomposition of the graph $G' - \mathbb{H}$ with respect to $\{\mathbb{S}_{\eta/2, k}(G'), \mathbb{L}_{\eta/2, k}(G') \setminus \mathbb{H}\}$ that is given by Lemma 3.13. Clearly, $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G'_{\text{reg}}, G'_{\text{exp}}, \mathbb{E})$ is a $(k, \Omega_{i+1}, \Omega_i, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition of G' capturing at least as many edges as promised in the lemma. $\qquad\square$

The process of embedding a given tree $T_{\text{T1.3}} \in \mathbf{trees}(k)$ into $G_{\text{T1.3}}$ is based on the sparse decomposition $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\text{reg}}, G_{\text{exp}}, \mathbb{E})$ of $G_{\text{T1.3}}$ given by Lemma 3.14 and is much more complex than in approaches based on the standard regularity lemma. The embedding ingredient in the classic (dense) regularity method inheres in blow-up lemma type statements which roughly tell that regular pairs of positive density in some sense behave like complete bipartite graphs. In our setting, in addition to regular pairs we shall use three other components of ∇: the vertices of huge degree $\mathbb{H}$, the nowhere-dense graph G_{exp}, and the avoiding set $\mathbb{E}$. Each of these components requires a different strategy for embedding (parts of) $T_{\text{T1.3}}$. Let us mention that rather major technicalities arise when combining these strategies.

These strategies are described precisely and in detail in [HKP$^+$d]. An informal account on the role of $\mathbb{E}$ is given in Section 3.5. We discuss the use of G_{exp} in Section 3.6. Only very little can be said about the set $\mathbb{H}$ at an intuitive level: these vertices have huge degrees but are very unstructured otherwise. If only $o(kn)$ edges are incident with $\mathbb{H}$ then we can neglect them. If, on the other hand, there are $\Omega(kn)$ edges incident with $\mathbb{H}$, then we have no choice but to use them for our embedding. Very roughly speaking, in that case we find sets $\mathbb{H}' \subseteq \mathbb{H}$ and $V' \subseteq V(G) \setminus \mathbb{H}$ such that still $\text{mindeg}(\mathbb{H}', V') \gg k$, and $\text{mindeg}(V', \mathbb{H}') = \Omega(k)$, and then use $\mathbb{H}'$ and V' in our embedding.

Last, let us note that when $G_{\text{T1.3}}$ is close to the extremal graph (depicted in Figure 1.1) then all the structure in $G_{\text{T1.3}}$ captured by Lemma 3.14 accumulates in the cluster graph G'_{reg}, i.e., $\mathbb{H}$, G'_{exp} and $\mathbb{E}$ are all almost empty. For that reason, when some of $\mathbb{H}$, G'_{exp} or $\mathbb{E}$ is substantial we gain some extra aid. In comparison, one of the almost extremal graphs for the Erdős–Sós Conjecture 1.1 has a substantial $\mathbb{H}$-component (see Figure 1.2).

3.4   Decomposition of general graphs

A version of Lemma 3.14 can be formulated for general graphs. To illustrate this, we present below a generic lemma of this type, which will not be used in the proof of the main theorem.

Lemma 3.15. *For every $\eta, \Lambda, \gamma, \varepsilon, \rho > 0$ there are numbers $\nu > 0$ and $k_0 \in \mathbb{N}$ such that for every sequence $(\Omega_j)_{j \in \mathbb{N}}$ of positive numbers with $\frac{\Omega_j}{\Omega_{j+1}} \leqslant \frac{\eta}{4}$ the following holds. Suppose that G is a graph of order n with average degree $k > k_0$. Then there is an index $i \leqslant \frac{4}{\eta}$, such that G has a $(k, \Omega_{i+1}, \Omega_i, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\text{reg}}, G_{\text{exp}}, \mathbb{E})$ that captures all but at most*

$$(\eta + \frac{4\varepsilon}{\gamma} + \varepsilon \Omega_{\lfloor 4\eta^{-1} \rfloor} + \gamma + \rho)kn \tag{3.6}$$

edges.

The proof follows the same strategy as that of Lemma 3.14.

Proof outline. By Lemma 3.1 there exists a spanning sugraph G' of G with $e(G) - e(G') < \eta kn$, and an index $i \leqslant \frac{4}{\eta}$ such that the assertion of Lemma 3.2(ii) holds. The bounded-degree part can then be decomposed using Lemma 3.13, yielding the desired sparse decomposition. $\qquad\square$

This decomposition could be used to attack other problems; probably with a version of Lemma 3.15 tailored to a particular setting similarly as we did in Lemma 3.14. However, our feeling is that such a decomposition lemma is limited in applications to tree-containment problems. The reason is that two of the features of the sparse decomposition, the nowhere-dense graph G_{exp} and the avoiding set $\mathbb{E}$, seem to be useful only for embedding trees. See Section 3.5 and Section 3.6 for a discussion of the respective embedding strategies.

3.5   The role of the avoiding set $\mathbb{E}$

Let us explain the role of the avoiding set $\mathbb{E}$ in Lemma 3.13. As said above, our aim in Lemma 3.13 will be to locally regularize parts of the input graph G. Of course, first we try to regularize as large a part of the G as possible. The avoiding set arises as a result of the impossibility to regularize certain parts of the graph. Indeed, it is one of the most surprising steps in our proof of Theorem 1.3 that the set $\mathbb{E}$ is initially defined as — very loosely speaking — "those vertices where the regularity lemma fails to work properly", and only then we prove that $\mathbb{E}$ actually satisfies the useful conditions of Definition 3.6.

We now sketch how to utilize avoiding sets for the purpose of embedding trees. In our proof of Theorem 1.3 we preprocess the tree $T = T_{\text{T1.3}} \in \mathbf{trees}(k)$ by choosing several cut-vertices so that the tree decomposes into small components, called *shrubs*. We cut T so that the order of each shrub is at most τk, where $\tau > 0$ is a small constant. Then we sequentially embed those shrubs. Thus embedding techniques for embedding a single shrub are the building blocks of our embedding

machinery; and $\mathbb{E}$ is one of the enviroments which provides us with such a technique. Let us discuss here the simpler case of embedding end shrubs (i.e. shrubs incident to a single cut-vertex). More precisely, we show how to extend a partial embedding of a tree by one end-shrub. To this end, let us suppose that ϕ is a partial embedding of a tree T, and $v \in V(T)$ is its *active vertex*, i.e., a vertex which is embedded, but not all its children are. We write $U \subseteq V(G)$ for the current image of ϕ. Let $T' \subseteq T$ be an end-shrub which is not embedded yet, and suppose $u \in V(T')$ is adjacent to v. We have $v(T') \leqslant \tau k$.

We now show how to extend the partial embedding ϕ to T', assuming that $\deg_G\left(\phi(v), \mathbb{E} \setminus U\right) \geqslant \gamma k$ for some $(1, \varepsilon, \gamma, k)$-avoiding set $\mathbb{E}$ (where $\tau \ll \varepsilon \ll \gamma \ll 1$). Let X be the set of at most εk exceptional vertices from Definition 3.6 corresponding to the set U. We now embed T' into G, starting by embedding u in a vertex of $\mathbb{E} \setminus (U \cup X)$ in the neighbourhood of $\phi(v)$. By Definition 3.6, there is a dense spot $D = (A_D, B_D; F) \in \mathcal{D}$ such that $\phi(u) \in V(D)$ and $|U \cap V(D)| \leqslant \gamma^2 k$. As D is a dense spot, we have $\deg_G(\phi(u), V(D)) > \gamma k$. We can greedily embed T' into D using the minimum degree in D. See Figure 3.1 for an illustration, and [HKP$^+$d, Lemma 6.4] for a precise formulation.
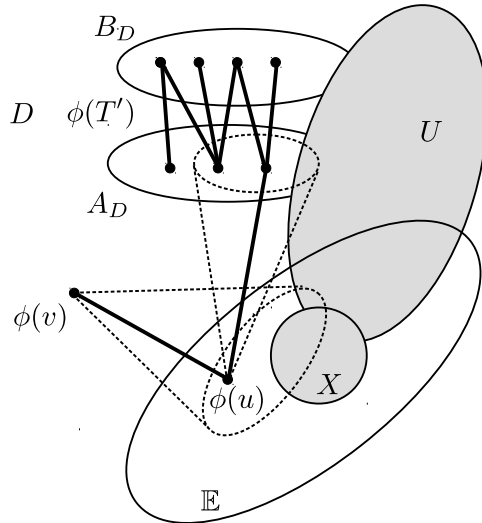
Figure 3.1: Embedding using the set $\mathbb{E}$.

We indeed use the avoiding set for embedding shrubs of T as above. The major simplification we made in the exposition is that we only discussed the case when T' is an end shrub. To cover embedding of an internal shrub T' as well (i.e. a shrub that is incident to more than one cut-vertex), one needs to have a more detailed control over the embedding, i.e., one must be able to extend the embedding of T' to the neighbouring cut-vertices, in such a way that one can then continue the embedding.

Last, let us remark, that unlike our baby-example above, we use an $(\Lambda, \varepsilon, \gamma, k)$-avoiding set with $\Lambda \gg 1$. This is because in the actual proof one has to avoid more vertices than just the current image of the embedding.

3.6 The role of the nowhere-dense graph G_{exp}

In this section we shall give some intuition on how the $(\gamma k, \gamma)$-nowhere-dense graph G_{exp} from the $(k, \Omega^{**}, \Omega^{*}, \Lambda, \gamma, \varepsilon', \nu, \rho)$-sparse decomposition[3] $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\text{reg}}, G_{\text{exp}}, \mathbb{E})$ of a graph G is useful for embedding a given tree $T \in \mathbf{trees}(k)$. We start out with the rather simple case when T is a path. We then point out an issue with this approach for trees with many branching vertices and show how to overcome this problem.

Embedding a path in G_{exp}. Assume we are given a path $T = u_1 u_2 \cdots u_k \in \mathbf{trees}(k)$ and we wish to embed it into G_{exp}. The idea is to apply a one-step look-ahead strategy. We first embed u_1 in an arbitrary vertex $v \in V(G_{\text{exp}})$. Then, we extend our embedding ϕ_ℓ of the path $u_1 \cdots u_\ell$ in G_{exp} in step ℓ by embedding $u_{\ell+1}$ in a (yet unused) neighbour w of the image of the *active* vertex u_ℓ, requiring that

$$\deg_{G_{\text{exp}}} \left(w, \phi_\ell(u_1 \cdots u_\ell) \right) < \sqrt{\gamma} k . \tag{3.7}$$

Let us argue that such a vertex w exists using induction on ℓ. First, observe that Property 1 of Definition 3.7 implies that $\phi_\ell(u_\ell)$ has at least ρk neighbours. By (3.7) applied to $\ell-1$, at most $\sqrt{\gamma} k$ of these neighbours lie inside $\phi_\ell(u_1 \cdots u_{\ell-1})$; this property is also trivially satisfied when $\ell = 1$. Further, an easy calculation shows that at most $16\sqrt{\gamma} k$ of them have degree more than $\sqrt{\gamma} k$ in G_{exp} into the set $\phi_\ell(u_1 \cdots u_\ell)$, otherwise we would get a contradiction to G_{exp} being $(\gamma k, \gamma)$-nowhere-dense. Since we assumed $\rho > 17\sqrt{\gamma}$ we can find a vertex w satisfying (3.7) and thus embed all of T.

Embedding trees with many branching points. We certainly cannot hope that a nonempty graph G_{exp} alone will provide us with embeddings of all trees $T \in \mathbf{trees}(k)$ from Theorem 1.3. For instance, if T is a star, then we need in G a vertex of degree $k - 1$, which G_{exp} might not have. The structure of LKS graphs allows to deal with embedding high-degree vertices. However, even without any vertex of large degree in our tree, the method described above might not always work, as we show next.

Consider a binary tree $T \in \mathbf{trees}(k)$, rooted at its central vertex r. Now if we try to embed T sequentially as above we will arrive at a moment when there are many (as many as $\log_2 k$) active vertices; regardless in which order we embed.[4] Now, the neighbourhoods of the images of the active vertices cannot be controlled much, i.e., they may be intersecting considerably. Hence, when embedding children of active vertices we might block available space in the neighbourhoods of other active vertices. See Figure 3.2 for an illustration.

To rescue the situation we partition T so that the first q levels of T from the root r form the set of the cut-vertices W. All other vertices make up the end shrubs $T_1^*, \ldots, T_h^*$. That is, $|W| = 2^q - 1$, and $h = 2^{q+1} - 2$.

We first embed the few cut-vertices W. As ρk will be much larger 2^q, following a strategy similar to the one above we ensure that all cut-vertices get correctly embedded. The next step is to make the transitions at the q-th level from embedding cut-vertices to embedding shrubs $T_1^*, \ldots, T_h^*$. But since this step requires to exploit the structure of LKS graphs, we skip the details in the high-level

---

[3]We shall assume that $17\sqrt{\gamma} < \rho$; this will be the setting of the sparse decomposition we shall work with in the proof of Theorem 1.3.

[4]The only requirement on the ordering is that in each moment the embedded part of the tree forms a connected subgraph; in particular we may use the depth-first and the breadth-first orders.
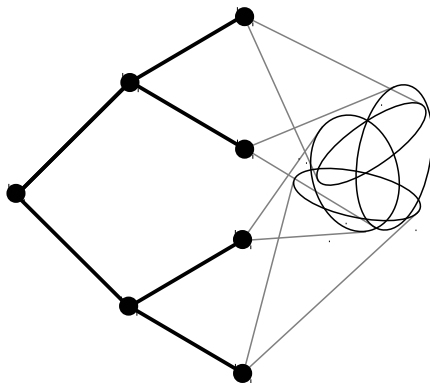
Figure 3.2: Embedded part of the binary tree in bold. The neighbourhoods of active vertices may overlap.

overview here. For the sake of this simplified example, let us assume that all the cut-vertices are embedded in a set L with

$$\mathrm{mindeg}_{G_{\exp}}(L) \geqslant \delta k \ , \tag{3.8}$$

(where $\rho \ll \delta < 1$ is a small constant).

For the point we wish to make here, it is more relevant to see how to complete the last part of our embedding, that is, how to embed a tree T_i^* whose root r_i is already embedded in a vertex $\phi(r_i) \in V(G_{\exp})$. Let $\mathrm{im}_i := \mathrm{im}(\phi)$ be the current (partial) image of ϕ. Further, we assume that throughout the entire process we have

$$\deg_{G_{\exp}}(\phi(r_i), V(G_{\exp}) \setminus \mathrm{im}) \gtrapprox \delta k/2 \ , \tag{3.9}$$

where im is the image at that moment (and in particular, also at the end of the process). We explain how to achieve this property at the end.

We emphasize that at this moment we are working exclusively with the tree T_i^*, i.e., any other tree T_j^* is either completely embedded, or will be embedded only after we finish the embedding of T_i^*. Suppose we are about to embed a vertex $v \in V(T_i^*)$ whose ancestor $v' \in V(T_i^*) \cup W$ is already embedded in $V(G_{\exp})$. We choose for the image of v any (yet unused) vertex w in the neighbourhood of $\varphi(v')$, requiring that

$$\deg_{G_{\exp}}(w, \mathrm{im}_i) < \rho k/100 \ . \tag{3.10}$$

This condition is very similar to our path-embedding procedure above, and can be proved in exactly the same way, using the fact that $G_{\exp}$ is $(\gamma k, \gamma)$-nowhere-dense. When $v' \in W$ is a cut-vertex, we need to combine this argument with (3.9).

Note that during our embedding $|\mathrm{im}(\phi) \setminus \mathrm{im}_i|$ will grow. However, $|\mathrm{im}(\phi) \setminus \mathrm{im}_i|$ is at most $v(T_i^*)$, which is much smaller than ρk. Thus, for every vertex $v'' \in V(T_i^*)$, when its time to be embedded comes, we still have a small degree to the partial image of the tree. Therefore v'' can be embedded on a vertex w that satisfies $\deg_{G_{\exp}}(w, \mathrm{im}(\phi)) < \rho k/50$ similarly as in (3.10).

Note that the trick here was to keep on working on one subtree T_i^*, whose size is small enough to be negligible in comparison to the degree of the vertices in $G_{\exp}$. So, by avoiding the vertices that have a considerable degree into im_i, we actually also avoid those vertices that have a considerable

degree into $\mathrm{im}(\phi)$. Breaking up the tree into tiny shrubs was thus the key to successfully embedding it.

Let us now explain how to achieve (3.9). Instead of just embedding the tree T_i^* we shall also reserve an equal amount of vertices in G_{exp} that are touched only exceptionally. More precisely, in a given step, instead of extending the embedding from a vertex to its two children, we first find four candidate vertices, and we randomly select two of them to host these children and insert the remaining two into a reserve set R. Condition (3.10) is replaced by $\deg_{G_{\mathrm{exp}}}(w, \mathrm{im}_i \cup R) < \rho k/100$. This allows us to avoid not only im_i but also R when extending the embedding of T_i^*. The only time the set R may be used to host a vertex v of some tree $T_1^*, \ldots, T_h^*$ is when v is the root of such a tree. Since the choice for the inclusion of vertices to R was random, with high probability we have

$$\deg_{G_{\mathrm{exp}}}(\phi(r_i), \mathrm{im}_i) \approx \deg_{G_{\mathrm{exp}}}(\phi(r_i), R) \pm h ,$$

where the $\pm h$ term amounts to the roots for which the random choice is not used. Recall that $h \ll \rho k$. This together with (3.8) establishes (3.9).

3.7 Proof of the decomposition lemma

This subsection is devoted to the proof of the decomposition lemma (Lemma 3.13). In the proof, we start by extracting the edges of as many $(\gamma k, k)$-dense spots from G as possible; these together with the incident vertices will form the auxiliary graph $G_{\mathcal{D}}$. Most of the remaining edges will form the edge set of the graph G_{exp}. Next, we consider the intersections of the dense spots captured in $G_{\mathcal{D}}$. We apply the regularity lemma for locally dense graphs (Lemma 2.5) to the subgraph of $G_{\mathcal{D}}$ that is spanned by the large intersections, and thus obtain G_{reg}. The other part of $V(G_{\mathcal{D}})$ will be taken as the $(\Lambda, \varepsilon, \gamma, k)$-avoiding set $\mathbb{E}$.

Setting up the parameters. We start by setting

$$\tilde{\nu} := \varepsilon \cdot 3^{-\frac{\Omega \Lambda}{\gamma^3}}.$$

Let q_{MAXCL} be given by Lemma 2.5 for input parameters

$$m_{\mathrm{L2.5}} := \frac{\Omega}{\gamma \tilde{\nu}} \quad , \quad z_{\mathrm{L2.5}} := 4s \quad \text{and} \quad \varepsilon_{\mathrm{L2.5}} := \varepsilon . \tag{3.11}$$

Define an auxiliary parameter $q := \max\{q_{\mathrm{MAXCL}}, \varepsilon^{-1}\}$ and choose the output parameters of Lemma 3.13 as

$$k_0 := \left\lceil \frac{q_{\mathrm{MAXCL}}}{\tilde{\nu}} \right\rceil \quad \text{and} \quad \nu := \frac{\tilde{\nu}}{q}.$$

Defining $\mathcal{D}$ and G_{exp}. Given a graph G, take a family $\mathcal{D}$ of edge-disjoint $(\gamma k, \gamma)$-dense spots such that the resulting graph $G_{\mathcal{D}} \subseteq G$ (which contains those vertices and edges that are contained in $\bigcup \mathcal{D}$) has the maximum number of edges.

Then by Lemma 2.1 there exists a graph $G_{\mathrm{exp}} \subseteq G - G_{\mathcal{D}}$ with $\mathrm{mindeg}(G_{\mathrm{exp}}) > \rho k$ and such that

$$|E(G) \setminus (E(G_{\mathrm{exp}}) \cup E(G_{\mathcal{D}}))| \leqslant \rho k n . \tag{3.12}$$

This choice of $\mathcal{D}$ and G_{exp} already satisfies Properties 5 and 1 of Definition 3.7.

Preparing for an application of the regularity lemma. Let

$$\mathcal{X} := \boxplus_D \{U, W, V(G) \setminus V(D)\} ,\tag{3.13}$$

where the latter partition refinement ranges over all $D = (U, W; F) \in \mathcal{D}$. Let $\mathcal{B} := \{X \in \mathcal{X} : X \subseteq V(G_{\mathcal{D}})\}$, $\tilde{\mathcal{B}} := \{B \in \mathcal{B} : |B| > 2\tilde{\nu}k\}$, and $\tilde{\mathcal{C}} := \mathcal{B} \setminus \tilde{\mathcal{B}}$. Furthermore let $\tilde{B} := \bigcup_{B \in \tilde{\mathcal{B}}} B$ and $\mathbb{E} := \bigcup_{C \in \tilde{\mathcal{C}}} C$.

Now, partition each set $B \in \tilde{\mathcal{B}}$ into $c_B := \lceil |B|/2\tilde{\nu}k \rceil$ subsets $B_1, \ldots, B_{c_B}$ of cardinalities differing by at most one, and let $\mathcal{B}'$ be the set containing all the sets B_i (for all $B \in \tilde{\mathcal{B}}$). Then for each $B \in \mathcal{B}'$ we have that

$$\tilde{\nu}k \leqslant |B| \leqslant 2\tilde{\nu}k \leqslant \varepsilon k .\tag{3.14}$$

Construct a graph H on $\mathcal{B}'$ by making two vertices $A_1, A_2 \in \mathcal{B}'$ adjacent in H if

(A) there is a dense spot $D = (U, W; F) \in \mathcal{D}$ such that $A_1 \subseteq U$ and $A_2 \subseteq W$, and

(B) $\mathrm{d}_G(A_1, A_2) \geqslant \gamma$.

Note that it follows from the way $\mathcal{D}$ was chosen that if $A_1 A_2 \in E(H)$ then $G[A_1, A_2] = G_{\mathcal{D}}[A_1, A_2]$. On the other hand note that we do not necessarily have $G[A_1, A_2] = D[A_1, A_2]$ for the dense spot D appearing in (A); just because there may be several such dense spots D.

By the assumption of Lemma 3.13, $\mathrm{maxdeg}(G) \leqslant \Omega k$. So, for each $B \in \mathcal{B}'$ we have $e_G(B, \tilde{B} \setminus B) \leqslant \Omega k |B|$. On the other hand, (3.14) and (B) imply that $\gamma\tilde{\nu}k|B| \deg_H(B) \leqslant e_G(B, \tilde{B} \setminus B)$. We conclude that

$$\mathrm{maxdeg}(H) \leqslant \frac{\Omega}{\gamma\tilde{\nu}} = m_{\mathrm{L}2.5} .\tag{3.15}$$

Regularising the dense spots in $\tilde{B}$. We apply Lemma 2.5 with parameters $m_{\mathrm{L}2.5}, z_{\mathrm{L}2.5}$ and $\varepsilon_{\mathrm{L}2.5}$ as defined by (3.11) to the graphs $H_{\mathrm{L}2.5} := G_{\mathcal{D}}$ and $F_{\mathrm{L}2.5} := H$, together with the ensemble $\mathcal{B}'$ in the role of the sets W_i, and partition of $V(G_{\mathcal{D}})$ induced by

$$\mathcal{Z}_{\mathrm{L}2.5} := \mathcal{V} \boxplus \{V(G_{\mathrm{exp}}), V(G) \setminus V(G_{\mathrm{exp}})\} \boxplus \{V_{\rightsquigarrow\mathbb{E}}, V(G) \setminus V_{\rightsquigarrow\mathbb{E}}\} ,$$

where $V_{\rightsquigarrow\mathbb{E}} := \{v \in V(G) : \deg(v, \mathbb{E}) > b\}$.

Observe that $\mathcal{B}'$ is an $(\tilde{\nu}k)$-ensemble satisfying condition (2.2) of Lemma 2.5, by (3.14), by the choice of k_0, and by (3.15). Thus we obtain integers $\{p_A\}_{A \in \mathcal{B}'}$ and a family $\mathbf{V} = \{W_A^{(1)}, \ldots, W_A^{(p_A)}\}_{A \in \mathcal{B}'}$ and a set $W_0 := \bigcup_{A \in \mathcal{B}'} W_A^{(0)}$ such that, in particular, we have the following.

(I) We have $\varepsilon^{-1} \leqslant p_A \leqslant q_{\mathrm{MAXCL}}$ for all $A \in \mathcal{B}'$.

(II) We have $|W_A^{(x)}| = |W_B^{(y)}|$ for any $A, B \in \mathcal{B}'$ and for any $x \in [p_A]$, $y \in [p_B]$.

(III) For any $A \in \mathcal{B}'$ and any $a \in [p_A]$, there is a set $V \in \mathcal{V}$ for which $W_A^{(a)} \subseteq V$. We either have that $W_A^{(a)} \subseteq V(G_{\mathrm{exp}})$, or $W_A^{(a)} \cap V(G_{\mathrm{exp}}) = \emptyset$ and $W_A^{(a)} \subseteq V_{\rightsquigarrow\mathbb{E}}$, or $W_A^{(a)} \cap V_{\rightsquigarrow\mathbb{E}} = \emptyset$.

(IV) $\sum_{e \in E(H)} |\mathrm{irreg}(e)| \leqslant \varepsilon \sum_{AB \in E(H)} |A||B|$, where $\mathrm{irreg}(AB)$ is the set of all edges of the graph G contained in an ε-irregular pair $(W_A^{(x)}, W_B^{(y)})$, with $x \in [p_A]$, $y \in [p_B]$, $AB \in E(H)$.

Let G_{reg} be obtained from $G_{\mathcal{D}}$ by erasing all vertices in W_0, and all edges that lie in pairs $(W_A^{(x)}, W_B^{(y)})$ which are irregular or of density at most γ^2. Then Properties 2, 3, 6 and 7 of Definition 3.7 are satisfied. Further, Lemma 2.10 implies (3.5). Together with (3.12) we obtain that the number of edges that are not captured by $(\mathbf{V}, \mathcal{D}, G_{\text{reg}}, G_{\text{exp}}, \mathbb{E})$ is at most $(\frac{4\varepsilon}{\gamma} + \varepsilon\Omega + \gamma + \rho)kn$.

Note that Properties (I), (II) and (3.14) imply that for all $A \in \mathcal{B}'$ and for any $a \in [p_A]$ we have that

$$\varepsilon k \geqslant |A| \geqslant |W_A^{(a)}| \geqslant \frac{\tilde{\nu}k}{q_{\text{MAXCL}}} \geqslant \frac{\tilde{\nu}k}{q} = \nu k.$$

Thus also Property 4 of Definition 3.7 holds.

The refinement in (3.13) guarantees that the bounded decomposition we have constructed respects the avoiding threshold b.

So, it only remains to see Property 8 of Definition 3.7.

The avoiding property of $\mathbb{E}$. In order to see Property 8 of Definition 3.7, we have to show that $\mathbb{E}$ is $(\Lambda, \varepsilon, \gamma, k)$-avoiding with respect to $\mathcal{D}$. For this, let $\bar{U} \subseteq V(G)$ be such that $|\bar{U}| \leqslant \Lambda k$. Let X be the set of those vertices $v \in \mathbb{E}$ that are not contained in any dense spot $D \in \mathcal{D}$ for which $|\bar{U} \cap V(D)| \leqslant \gamma^2 k$. Our aim is to see that $|X| \leqslant \varepsilon k$.

Let $\mathcal{D}_X \subseteq \mathcal{D}$ be the set of all dense spots D with $X \cap V(D) \neq \emptyset$. Setting $\mathcal{A} := \{A \in \tilde{\mathcal{C}} : A \cap X \neq \emptyset\}$, the definition of $\mathbb{E}$ trivially implies that $\frac{|X|}{2\tilde{\nu}k} \leqslant |\mathcal{A}|$. Now, by the definition of $\mathcal{B}$, we know that there are at most $3^{|\mathcal{D}_X|}$ sets $A \in \mathcal{A}$. Indeed, for each $D = (U, W; F) \in \mathcal{D}_X$, either A is a subset of U, or of W, or of $V(G) \setminus V(D)$. Thus,

$$3^{|\mathcal{D}_X|} \geqslant |\mathcal{A}| \geqslant \frac{|X|}{\tilde{\nu}k} . \tag{3.16}$$

By Fact 3.5, each vertex of $V(G)$ lies in at most Ω/γ of the $(\gamma k, \gamma)$-dense spots from $\mathcal{D}$. Hence

$$\frac{\Omega}{\gamma}|\bar{U}| \geqslant \sum_{D \in \mathcal{D}_X} |V(D) \cap \bar{U}| \geqslant |\mathcal{D}_X|\gamma^2 k \overset{(3.16)}{\geqslant} \log_3\left(\frac{|X|}{\tilde{\nu}k}\right)\gamma^2 k ,$$

where the second inequality holds by the definition of X. Thus

$$|X| \leqslant 3^{\frac{\Omega\Lambda}{\gamma^3}} \cdot \tilde{\nu}k = \varepsilon k ,$$

as desired. This finishes the proof of Lemma 3.13.

3.8 Sparse decomposition of dense graphs

Let us explain our remark above that in the setting of a dense graph G, Lemmas 3.14 and 3.15 produce a regularity partition in the usual sense. So, suppose that G is an n-vertex graph and has at least an^2 edges. This needs to be understood with the usual quantification "$a > 0$ is fixed and n is large".

Recall that when we inquire a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition, the parameters satisfy $\Omega^{**}, \Omega^*, \Lambda \gg 1 \gg \gamma, \varepsilon, \nu, \rho > 0$. The interplay between the parameters is quite complicated, and we do not give it here in full (see [HKP$^+$d, p. 65] for details). We justify with "parameter choice" any further relation we assume between them. Also, let us note that while our exact choice

of parameters made in [HKP$^+$d] are tailored for proving Theorem 1.3, we expect these relations to be satisfied in any application of Lemma 3.15, at least on the loose level we make use of them in this section.

First, we argue that it makes sense to set k linear in n, i.e., $k = cn$ for some c depending on a only. Indeed, having $k \gg n$ would allow that all edges of G are uncaptured in (3.6), which would make the lemma worthless. On the other hand, with $k \ll n$ we would have all vertices from $Q = \{v \in V(G) : \deg(v) > \sqrt{a}n\}$ ending up in the huge-degree set $\mathbb{H}$ for which the sparse decomposition provides no structural information. Since $|Q| \geqslant \sqrt{a}n$, that would be a big loss of information, and thus often undesirable.

So, suppose now that $k = cn$, and suppose that $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ is a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition of the dense graph G. Since $\Omega^{**}k = \Omega^{**}cn > n$, we have that $\mathbb{H} = \emptyset$. Next, we argue that G_{exp} contains no vertices. Suppose on the contrary that it does. Then the minimum degree condition in Property 1 of Definition 3.7 tells us that G_{exp} has at least ρk vertices of degrees at least ρk each. Thus, $e(G_{\mathrm{exp}}) \geqslant \rho^2 k/2 = c^2 \rho^2 n^2/2$. Since G_{exp} has at most n vertices, and since $c\rho \gg \gamma$ (parameter choice), we get that G_{exp} contains at least one $(\gamma k, \gamma)$-dense spot, a contradiction to G_{exp} being nowhere-dense. Last, we claim, that $|\mathbb{E}| \leqslant \varepsilon k$. To this end, consider the set $U_{\mathrm{D3.6}} = V(G)$. We have $|U_{\mathrm{D3.6}}| = n \leqslant \Lambda cn$, and thus the condition in Definition 3.6 applies. But there cannot exist any $(\gamma k, \gamma)$-dense spot as asserted in Definition 3.6 since for such a dense spot D we would have $|V(D)| < \gamma^2 k$, contradicting its required minimum degree condition. Thus, we conclude that all the vertices $v \in \mathbb{E}$ are exceptional in the sense of Definition 3.6, leading to the desired bound on $|\mathbb{E}|$.

To summarize, in the sparse decomposition $(\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$, we have that $\mathbb{H}, G_{\mathrm{exp}}, \mathbb{E}$ are empty or almost empty. Thus, according to Definition 3.9, all the captured edges lie in the regularized graph G_{reg}. Property 4 of Definition 3.7 tells us that the clusters have size at least $\nu k = (\nu c)n$, that is, linear in the order of G. Further, this property tells us that these clusters are of the same size. We conclude that G_{reg} is a regularization of G in the sense of the original regularity lemma.

3.9 Algorithmic aspects of the decomposition lemma

Let us look back at the proof of the decomposition lemma (Lemma 3.13) and observe that we can get a bounded decomposition of any bounded-degree graph algorithmically in quasipolynomial time (in the order of the graph). Note that this in turn provides efficiently a sparse decomposition of any graph, since the initial step of splitting the graph into huge versus bounded degree vertices (cf. Lemma 3.2) can be done in polynomial time.

There are only two steps in the proof of Lemma 3.13 which need to be done algorithmically: the extraction of dense spots, and the simultaneous regularization of some dense pairs.

It will be more convenient to work with a relaxation of the notion of dense spots. We call a graph H (d, ℓ)-*thick* if $v(H) \geqslant \ell$, and $e(H) \geqslant dv(H)^2$. The notion of thick graphs is a relaxation of dense spots, where the minimum degree condition is replaced by imposing a lower bound on the order, and the bipartiteness requirement is dropped. It can be verified that in our proof it is not important that the dense spots $\mathcal{D}$ and the nowhere-dense graph G_{exp} are parametrized by the same constants, i.e., the entire proof would go through even if the spots in $\mathcal{D}$ were $(\gamma k, \gamma)$-dense, and G_{exp} were $(\beta k, \beta)$-nowhere-dense for some $\beta \gg \gamma$. Each $(\beta k, \beta)$-thick graph gives (algorithmically) a $(\beta k/4, \beta/4)$-dense spot, and thus it is enough to extract thick graphs.

For the extraction of thick graphs we would need to efficiently answer the following: Given a

number $\beta > 0$, find a number $\gamma > 0$ such that for an input number h and an N-vertex graph we can localize in G a (γ, h)-thick graph if it contains a (β, h)-thick graph, and output NO otherwise.[5] Employing techniques from a deep paper of Arora, Frieze and Kaplan [AFK02], one can solve this problem in quasipolynomial time $O(N^{c \cdot \log N})$. This was communicated to us by Maxim Sviridenko. On the negative side, a truly polynomial algorithm seems to be out of reach, as Alon, Arora, Manokaran, Moshkovitz, and Weinstein [AAM+] reduced the problem to the notorious hidden clique problem, whose tractability has been open for twenty years.

Theorem 3.16 ([AAM+])**.** *If there is no polynomial time algorithm for solving the clique problem for a planted clique of size $n^{1/3}$, then for any $\varepsilon \in (0, 1)$ and $\delta > 0$ there is no polynomial time algorithm that distinguishes between a graph G on N vertices containing a clique of size $\kappa = N^\varepsilon$ and a graph G' on N vertices in which the densest subgraph on κ vertices has density at most δ.[6]*

Of course, Theorem 3.16 leaves some hope for a polynomial time algorithm when $h = N^{o(1)}$ (which corresponds to $k_{\text{L3.13}} = n_{\text{L3.13}}^{o(1)}$).

The regularity lemma can be made algorithmic [ADL+94]. The algorithm from [ADL+94] is based on index pumping-up, and thus applies even to the locally dense setting of Lemma 2.5.

It will turn out that the extraction of dense spots is the only obstruction to a polynomial time algorithm for Theorem 1.3. In [HKP+d], we sketch a truly polynomial time algorithm which avoids this step. It seems that the method sketched there is generally applicable for problems which employ sparse decompositions.

4 Acknowledgements

---

[5]We could additionally assume that $\text{maxdeg}(G) \leqslant O(h)$ due to the previous step of removing the set $\mathbb{H}$ of huge degree vertices.

[6]The result as stated in [AAM+] covers only the range $\varepsilon \in (\frac{1}{3}, 1)$. However there is a simple reduction by taking many disjoint copies of the general range to the restricted one.

Symbol index

General index

References

[AAM⁺] N. Alon, S. Arora, R. Manokaran, D. Moshkovitz, and O. Weinstein. Inapproximability of densest κ-subgraph from average case hardness. Available via http://people.csail.mit.edu/dmoshkov/papers/.

[ADL⁺94] N. Alon, R. Duke, H. Lefmann, V. Rödl, and R. Yuster. The algorithmic aspects of the regularity lemma. *J. Algorithms*, 16:80–109, 1994.

[AFK02] S. Arora, A. Frieze, and H. Kaplan. A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. *Math. Program.*, 92(1):1–36, 2002.

[AKS81] M. Ajtai, J. Komlós, and E. Szemerédi. The longest path in a random graph. *Combinatorica*, 1(1):1–12, 1981.

[AKS95] M. Ajtai, J. Komlós, and E. Szemerédi. On a conjecture of Loebl. In *Graph theory, combinatorics, and algorithms, Vol. 1, 2 (Kalamazoo, MI, 1992)*, Wiley-Intersci. Publ., pages 1135–1146. Wiley, New York, 1995.

[AKS07] N. Alon, M. Krivelevich, and B. Sudakov. Embedding nearly-spanning bounded degree trees. *Combinatorica*, 27(6):629–644, 2007.

[AKSS] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. Erdős-Sós conjecture. In preparation.

[AKSV14] P. Allen, P. Keevash, B. Sudakov, and J. Verstraete. Turán numbers of bipartite graphs plus an odd cycle. *J. Combin. Theory Ser. B*, 106:134–162, 2014.

[BCLR89] S. N. Bhatt, F. R. K. Chung, F. T. Leighton, and A. L. Rosenberg. Universal graphs for bounded-degree trees and planar graphs. *SIAM J. Discrete Math.*, 2(2):145–155, 1989.

[BCPS10] J. Balogh, B. Csaba, M. Pei, and W. Samotij. Large bounded degree trees in expanding graphs. *Electron. J. Combin.*, 17(1):Research Paper 6, 9, 2010.

[BD96] S. Brandt and E. Dobson. The Erdős–Sós conjecture for graphs of girth 5. *Discr. Math.*, 150:411–414, 1996.

[BJ] O. Barr and R. Johansson. Another Note on the Loebl–Komlós–Sós Conjecture. Research reports no. 22, (1997), Umeå University, Sweden.

[BLW00] C. Bazgan, H. Li, and M. Woźniak. On the Loebl-Komlós-Sós conjecture. *J. Graph Theory*, 34(4):269–276, 2000.

[Bol84] B. Bollobás. The evolution of sparse graphs. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57. Academic Press, London, 1984.

[BS97] I. Benjamini and O. Schramm. Every graph with a positive Cheeger constant contains a tree with a positive Cheeger constant. *Geom. Funct. Anal.*, 7(3):403–419, 1997.

REFERENCES

[CLNGS10] B. Csaba, I. Levitt, J. Nagy-György, and E. Szemeredi. Tight bounds for embedding bounded degree trees. In *Fete of combinatorics and computer science*, volume 20 of *Bolyai Soc. Math. Stud.*, pages 95–137. János Bolyai Math. Soc., Budapest, 2010.

[Coo09] O. Cooley. Proof of the Loebl-Komlós-Sós conjecture for large, dense graphs. *Discrete Math.*, 309(21):6190–6228, 2009.

[Die05] R. Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 2005.

[Dob02] E. Dobson. Constructing trees in graphs whose complement has no $K_{2,s}$. *Combin. Probab. Comput.*, 11(4):343–347, 2002.

[EFLS95] P. Erdős, Z. Füredi, M. Loebl, and V. T. Sós. Discrepancy of trees. *Studia Sci. Math. Hungar.*, 30(1-2):47–57, 1995.

[EL] G. Elek and G. Lippner. An analogue of the Szemerédi regularity lemma for bounded degree graphs. Manuscript (arXiv:0809.2879).

[FdlV79] W. Fernandez de la Vega. Long paths in random graphs. *Studia Sci. Math. Hungar.*, 14(4):335–340, 1979.

[FNP] A. Ferber, R. Nenadov, and U. Peter. Universality of random graphs and rainbow embedding. to appear in Random Structures Algorithms, arXiv:1311.7063.

[FP87] J. Friedman and N. Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7(1):71–76, 1987.

[GG67] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 10:167–170, 1967.

[Har72] F. Harary. Recent results on generalized Ramsey theory for graphs. In *Graph theory and applications (Proc. Conf., Western Michigan Univ., Kalamazoo, Mich., 1972; dedicated to the memory of J. W. T. Youngs)*, pages 125–138. Lecture Notes in Math., Vol. 303. Springer, Berlin, 1972.

[Hax01] P. E. Haxell. Tree embeddings. *J. Graph Theory*, 36(3):121–130, 2001.

[HK95] P. E. Haxell and Y. Kohayakawa. The size-Ramsey number of trees. *Israel J. Math.*, 89(1-3):261–274, 1995.

[HKP⁺a] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi. The approximate Loebl–Komlós–Sós Conjecture I: The sparse decomposition. Manuscript (arXiv:1408.3858).

[HKP⁺b] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi. The approximate Loebl–Komlós–Sós Conjecture II: The rough structure of LKS graphs. Manuscript (arXiv:1408.3871).

[HKP⁺c] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi. The approximate Loebl–Komlós–Sós Conjecture III: The finer structure of LKS graphs. Manuscript (arXiv:1408.3866).

REFERENCES

[HKP⁺d]   J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi. The approximate Loebl–Komlós–Sós Conjecture IV: Embedding techniques and the proof of the main result. Manuscript (arXiv:1408.3870).

[HKS12]   D. Hefetz, M. Krivelevich, and T. Szabó. Sharp threshold for the appearance of certain spanning trees in random graphs. *Random Structures and Algorithms*, 41:391–412, 2012.

[Hla]   J. Hladký. Szemerédi Regularity Lemma and its applications in combinatorics. MSc. Thesis, Charles University in Prague, 2008, available via http://homepages.warwick.ac.uk/~maskai/papers.html.

[HLT02]   P. E. Haxell, T. Luczak, and P. W. Tingley. Ramsey numbers for trees of small maximum degree. *Combinatorica*, 22(2):287–320, 2002. Special issue: Paul Erdős and his mathematics.

[HP15]   J. Hladký and D. Piguet. Loebl–Komlós–Sós Conjecture: dense case. *J. Combin. Theory Ser. B*, 2015. http://dx.doi.org/10.1016/j.jctb.2015.07.004.

[HPS⁺15]   J. Hladký, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi. The approximate Loebl–Komlós–Sós conjecture and embedding trees in sparse graphs. *Electron. Res. Ann. Math. Sci.*, 22:1–11, 2015.

[JKS12]   D. Johannsen, M. Krivelevich, and W. Samotij. Expanders are universal for the class of all spanning trees. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 1539–1551. SIAM, 2012.

[JKV08]   A. Johansson, J. Kahn, and V. Vu. Factors in random graphs. *Random Structures Algorithms*, 33(1):1–28, 2008.

[KMO11a]   D. Kühn, R. Mycroft, and D. Osthus. An approximate version of Sumner's universal tournament conjecture. *J. Combin. Theory Ser. B*, 101(6):415–447, 2011.

[KMO11b]   D. Kühn, R. Mycroft, and D. Osthus. A proof of Sumner's universal tournament conjecture for large tournaments. *Proc. Lond. Math. Soc. (3)*, 102(4):731–766, 2011.

[KO09]   D. Kühn and D. Osthus. Embedding large subgraphs into dense graphs. In *Surveys in combinatorics 2009*, volume 365 of *London Math. Soc. Lecture Note Ser.*, pages 137–167. Cambridge Univ. Press, Cambridge, 2009.

[Koh97]   Y. Kohayakawa. Szemerédi's regularity lemma for sparse graphs. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 216–230. Springer, Berlin, 1997.

[Kri10]   M. Krivelevich. Embedding spanning trees in random graphs. *SIAM J. Discrete Math.*, 24(4):1495–1500, 2010.

[KS83]   J. Komlós and E. Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Math.*, 43(1):55–63, 1983.

REFERENCES

[KS96]   J. Komlós and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 295–352. János Bolyai Math. Soc., Budapest, 1996.

[KSS01]   J. Komlós, G. N. Sárközy, and E. Szemerédi. Spanning trees in dense graphs. *Combin. Probab. Comput.*, 10(5):397–416, 2001.

[Mona]   R. Montgomery. Embedding bounded degree spanning trees in random graphs. Manuscript (arXiv:1405.6559).

[Monb]   R. Montgomery. Sharp threshold for embedding combs and other spanning trees in random graphs. Manuscript (arXiv:1405.6560).

[MS]   M. Matsumoto and S. Sakamoto. Constructing trees in graphs whose complement has no one-leaved $K_{2,s}$. Preprint 2009.

[Pós76]   L. Pósa. Hamiltonian circuits in random graphs. *Discrete Math.*, 14(4):359–364, 1976.

[PS08]   D. Piguet and M. J. Stein. Loebl-Komlós-Sós conjecture for trees of diameter 5. *Electron. J. Combin.*, 15(1):Research Paper 106, 11 pp. (electronic), 2008.

[PS12]   D. Piguet and M. J. Stein. An approximate version of the Loebl-Komlós-Sós conjecture. *J. Combin. Theory Ser. B*, 102(1):102–125, 2012.

[Sco11]   A. Scott. Szemerédi's regularity lemma for matrices and sparse graphs. *Combin. Probab. Comput.*, 20(3):455–466, 2011.

[Sof00]   S. N. Soffer. The Komlós-Sós conjecture for graphs of girth 7. *Discrete Math.*, 214(1-3):279–283, 2000.

[Sun07]   L. Sun. On the Loebl-Komlós-Sós conjecture. *Australas. J. Combin.*, 37:271–275, 2007.

[SV10]   B. Sudakov and J. Vondrák. A randomized embedding algorithm for trees. *Combinatorica*, 30(4):445–470, 2010.

[SW97]   J.-F. Saclé and M. Woźniak. A note on the Erdős–Sós conjecture for graphs without C_4. *J. Combin. Theory (Series B)*, 70(2):229–234, 1997.

[Sze78]   E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[Woź96]   M. Woźniak. On the Erdős–Sós conjecture. *J. Graph Theory*, 21(2):229–234, 1996.

[Zha11]   Y. Zhao. Proof of the $(n/2 - n/2 - n/2)$ conjecture for large n. *Electron. J. Combin.*, 18(1):Paper 27, 61, 2011.