UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

TREE EMBEDDINGS IN DENSE GRAPHS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
CIENCIAS DE LA INGENIERÍA, MENCIÓN MATEMÁTICAS APLICADAS
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

GUIDO ANDRÉS BESOMI ORMAZÁBAL

PROFESORA GUÍA:
MAYA JAKOBINE STEIN

MIEMBROS DE LA COMISIÓN:
HIỆP HÀN
JUAN PEYPOUQUET URBANEJA
JOSÉ SOTO SAN MARTÍN

Este trabajo ha sido parcialmente financiado por CMM Conicyt PIA AFB170001

SANTIAGO DE CHILE
2018

RESUMEN DE LA MEMORIA PARA OPTAR AL TÍTULO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN MATEMÁTICAS APLICADAS
POR: GUIDO ANDRÉS BESOMI ORMAZÁBAL
FECHA: 2018
PROF. GUÍA: MAYA JAKOBINE STEIN

TREE EMBEDDINGS IN DENSE GRAPHS

En 1995 Komlós, Sárközy y Szemerédi probaron que para cualquier $\delta > 0$ y cualquier entero positivo Δ, todo grafo G de orden n, con n suficientemente grande, que satisfaga $\delta(G) \geq (1+\delta)\frac{n}{2}$, contiene como subgrafo a todo árbol de n vértices y grado máximo acotado por Δ. En esta memoria se presentan dos posibles generalizaciones de este resultado, estableciendo condiciones suficientes para el *embedding* de árboles de orden k en grafos con grado mínimo al menos $(1+\delta)\frac{k}{2}$, donde k es lineal en el orden del grafo anfitrión.

En 1963 Erdős y Sós conjeturaron que, dado un entero k, un grafo G con grado promedio mayor que $k-1$ debería contener todos los árboles en k aristas como subgrafos. Como consecuencia de uno de los resultados principales de esta memoria, se demuestra una versión parcial de la conjetura de Erdős-Sós.

Siguiendo la linea del *embedding* de árboles en grafos con condiciones de grado mínimo, Havet, Reed, Stein y Wood conjeturaron el 2016 que todo grafo con grado mínimo al menos $\lfloor \frac{2k}{3} \rfloor$ y grado máximo al menos k contiene todo árbol con k aristas como subgrafo. Las técnicas aquí desarrolladas permiten, adicionalmente, probar una versión parcial de esta conjetura.

ABSTRACT OF THE REPORT TO QUALIFY TO THE DEGREE OF
MASTER OF SCIENCE IN ENGINEERING, MENTION APPLIED MATHEMATICS
BY: GUIDO ANDRÉS BESOMI ORMAZÁBAL
DATE: 2018
GUIDE: MAYA JAKOBINE STEIN

TREE EMBEDDINGS IN DENSE GRAPHS

In 1995 Komlós, Sárközy and Szemerédi proved that for any $\delta > 0$ and any positive integer Δ, every graph G of order n, with n sufficiently large, that satisfies $\delta(G) \geq (1+\delta)\frac{n}{2}$, contains every tree on n vertices and maximum degree bounded by Δ as a subgraph. In this thesis we present two possible generalizations of this result, establishing sufficient conditions for the embedding of trees of order k in graphs with minimum degree at least $(1+\delta)\frac{k}{2}$, where k is linear in the order of the host graph.

In 1963 Erdős and Sós conjectured that, given an integer k, a graph G with average degree greater than $k-1$ should contain any tree on k edges as a subgraph. As a consequence of one of the main results of this tesis, we prove a partial approximated version of the Erdős-Sós conjecture.

Following the line of minimum degree conditions Havet, Reed, Stein and Wood conjectured in 2016 that every graph with minimum degree at least $\lfloor \frac{2k}{3} \rfloor$ and maximum degree at least k contains every tree on k edges as a subgraph. The techniques here developed allow, additionally, to prove a partial approximated version of this conjecture.

... the atlas is a manifold. This is a typical mathematician's use of the word "is", and should not be confused with the normal use.

Timothy Gowers

Agradecimientos

En primer lugar quisiera agradecer a mis papás, Carmen y Guido, por el infinito esfuerzo y cariño que han dedicado a mis hermanas y a mí desde que tengo memoria, y por su comprensión y apoyo incondicional. No existen palabras, y posiblemente tampoco acciones, que compensen de alguna manera todo lo que han hecho por nosotros. Agradezco también a mis hermanas, Tuti, Jose e Isi, con las que he convivido desde siempre, pero cuya compañía ha sido particularmente importante estos años lejos de la casa.

Agradezco a mis amigos, que hicieron de mi paso por la universidad un período valiosísimo y muy bonito también. Cofré, la Rosario, Lara, la Camila y Piero. En esos días en que ir a la universidad se hacía difícil, igual valía la pena subirse a una micro o al metro con tal de llegar a almorzar juntos y tener una de esas conversaciones interminables. Agradezco también a la Amelia, que tiene una fortaleza única y ha sido un apoyo fundamental en mi proceso académico y en mi vida en general. No puedo dejar de mencionar también al Tomi, que ha estado siempre presente, en buenos y malos momentos, y con quien comparto innumerables historias.

Agradezco a los profesores del DIM, que con su labor construyen día a día un espacio valioso para la creación y la difusión de las matemáticas. Agradezco también a los funcionarios del DIM, por su esfuerzo constante y dedicación, en especial a Karen y Silvia, quienes en más de una ocasión me sacaron de un apuro.

Finalmente agradezco a la profesora Maya Stein, por haberme guíado en este largo pero muy bonito proceso y con quien ha sido un privilegio trabajar este último año. Agradezco también a los profesores Hiệp Hàn y José Soto por sus correcciones y por su participación en mi comisión evaluadora.

Contents

Introduction

In 1907 Mantel proved that an n-vertex graph with no triangle as a subgraph contains at most $\frac{n^2}{4}$ edges. Many years later, in 1941, Turán [45] published a generalization of Mantel's theorem, determining the maximum number of edges that a graph could have without containing a clique of a given size as a subgraph. In light of subsequent events this result can be seen as the starting point of a prolific branch of graph theory known as extremal graph theory. Generally speaking, the primary question motivating extremal graph theory is how and when can a global property force some local structure in a graph. Global properties can vary from the number of edges to the chromatic number of a graph while a local structure could be anything from a given subgraph, such as a clique or a tree, to the more general notions of minor and immersion. The present thesis is motivated by the questions in extremal theory concerning trees, specifically the Erdős-Sós conjecture and some of the posterior works in tree embeddings.

In 1963 Erdős and Sós [15] conjectured that, given an integer k, a graph G with average degree greater than $k-1$ should contain any tree on k edges as a subgraph. The conjecture is trivially true for stars and the work published by Erdős and Gallai [13] in 1959 shows that it is also true for paths of length k. Besides, when the average degree condition is replaced by a minimum degree condition, i.e, when G satisfies $\delta(G) > k-1$, the embedding can be easily performed by means of a greedy argument. To see that the conjecture is best possible it is enough to consider the complete graph on k vertices, K_k, which does not contain any tree on k edges and has average degree $k-1$. Many efforts have been made around this conjecture, resulting in several partial results. The conjecture has been proved to be true for graphs of order $k+1$, $k+2$, $k+3$, $k+4$ and $k+5$ by Zhou [49] in 1984, by Slater, Teo and Yap [41] in 1985, by Woźniak [46] in 1996, by Tiner [44] in 2010 and by Yuan and Zhang [47] in 2014, respectively. A generalization of these results was obtained by Görlich and Żak [19] in 2016, proving that the conjecture holds for graphs of order $k+c$, where c is any given constant and k is sufficiently large. Other types of restrictions have also been made on the class of host graphs. In 1996 Brandt and Dobson [7] proved the conjecture for graphs of girth at least five. Later, in 1997, Saclé and Woźniak [40] improved the result of Brandt and Dobson by showing that the conjecture holds for graphs not containing a C_4 as a subgraph. Balasubramanian and Dobson [2] showed the conjecture for graphs not containing a $K_{2,s}$, with $s < \frac{k}{12} + 1$. In 2013 Eaton and Tiner [11] proved it for graphs not containing a path of length $k+4$. Interestingly, similar restrictions over the complement of the host graph also help to prove the conjecture. Li, Liu and Wang [35] showed the conjecture for the graphs whose complement is of girth at least five. This result was later improved by Dobson [9] showing it for graphs with a complement not containing a $K_{2,4}$. Restrictions on the class of trees have also proved to be helpful. Eaton and Tiner [10] successfully embedded the trees having a vertex with at least $\lceil \frac{k}{2} \rceil - 2$ leaf-neighbors and Woźniak [46] did the same with the trees of diameter at most four having at most one vertex of degree greater than two. Ajtai, Komlós, Simonovits and Szemerédi announced a proof of the conjecture for sufficiently large host graphs in the 1990's, but it has not been published to the date.

There are other conditions besides the average degree that are worth exploring. As we mentioned above, if the minimum degree of a graph G is at least k, then any tree on k edges can be easily embedded into G using a greedy strategy. Besides, it is known that every graph with average degree greater than k contains a subgraph with minimum degree at least $\frac{k}{2}$. Thus, it is natural to think of minimum degree conditions as an alternative to the average degree.

In 1995 Komlós, Sárközy and Szemerédi [30] proved that for any $\delta > 0$ and any positive integer Δ, every graph G of order n, with n sufficiently large, that satisfies $\delta(G) \geq (1+\delta)\frac{n}{2}$, contains every tree on n vertices and maximum degree bounded by Δ as a subgraph. This result gave an affirmative answer to a conjecture of Bollobás [6]. In 2001 the same authors [31] improved their result by showing that the bound on the degree of the trees can be replaced by a function of the form $c\frac{n}{\log n}$, where c is a constant, and this is best possible. A different improvement was made by Csaba, Levitt, Nagy-György and Szemerédi [8] in 2010, showing that, when considering trees with maximum degree bounded by a constant, it is enough to ask for the host graph to have minimum degree at least $\frac{n}{2} + c\log n$.

Following the line of minimum degree conditions Havet, Reed, Stein and Wood [20] conjectured in 2016 that every graph with minimum degree at least $\lfloor \frac{2k}{3} \rfloor$ and maximum degree at least k contains every tree on k edges as a subgraph. We will refer to this conjecture as the $\frac{2}{3}$ conjecture. The maximum degree condition is obviously needed since without this condition there would be no way of embedding the star on k edges. Also, there is no way of embedding the tree $T_{k,3}$ formed by a vertex connected to the central vertices of three stars of order $\frac{k}{3}$ into the graph formed by two copies of $K_{\frac{2k}{3}-1}$ and a universal vertex seeing both cliques; hence the bound $\frac{2k}{3}$ on the minimum degree is tight. The authors provide two results that support their conjecture. First they show that the conjecture holds when the bound on the maximum degree is replaced by a function which is exponential on k. Secondly they prove that there is a constant $\gamma > 0$ such that any graph G with $\delta(G) \geq (1-\gamma)k$ and $\Delta(G) \geq k$ contains every tree on k edges. Reed and Stein [39] have recently proved an approximated version of the conjecture for spanning trees.

Another interesting related conjecture that have received considerable attention is the Loebl-Komlós-Sós conjecture [12], which considers a median degree condition. It states that an n-vertex graph G with at least $\frac{n}{2}$ vertices of degree at least k contains every tree on k edges as a subgraph. Several efforts have been made around this conjecture too. Soffer [42] proved in 2000 that the conjecture was true for graphs of girth at least seven. Piguet and Stein [38] proved in 2012 an approximated version for the dense case. In 2017 Hladký and Piguet [27] improved the result from [38] by showing the exact case for dense host graphs. Zhao [48] proved the case where $k = n$ for large n. In a series of four papers [23–26] Hladký, Komlós, Piguet, Simonovits, Stein and Szemerédi proved a relaxation of the Loebl-Komlós-Sós conjecture: for every $\alpha > 0$ there exists $k_0 \in \mathbb{N}$ such that for every $k \geq k_0$, every n-vertex graph 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.

Other approaches around tree embeddings involve considering expansion properties [1, 3, 18, 21, 28], random host graphs [4, 16, 17, 22, 28, 29, 32, 36, 37] and random perturbations of deterministic graphs [33].

The main problem studied in this thesis can be seen as a generalization of the result of Komlós, Sárközy and Szemerédi [30], but the tools developed to treat this problem will also have implications on the $\frac{2}{3}$ conjecture and the Erdős-Sós conjecture. We are interested in the embedding of trees on k edges into graphs with minimum degree at least $(1+\delta)\frac{k}{2}$, where δ is any given positive constant and k is greater than a constant k_0 depending on δ. The sole hypothesis of high minimum degree is not sufficient, because if the minimum degree is below k, the host graph might simply be too small; thus, an additional condition is needed. A first idea would be to impose a restriction on the order of the host graph, forcing it to be larger than the trees we want to embed, in addition to being connected. However, the graph formed by two copies of $K_{\frac{k}{2}}$ connected by an edge does not contain $T_{k,3}$ as a subgraph. This example can be generalized to show that no bound on the order of the host graph is sufficient to ensure the appearance of every tree on k edges as a subgraph. Instead, we consider two types of maximum degree conditions. Firstly we ask for a vertex of high degree, specifically one with at least $(1+\delta)2k$ neighbors. In Section 3.2.2 we present an example showing that the bounds $\frac{k}{2}$ and $2k$ for the minimum and the maximum degree, respectively, are tight when no other assumptions are made. Secondly we forget about the high degree vertex and we ask for a small portion of vertices having degree at least $(1+\delta)k$. Whatever the case may be, we will only work with dense host graphs, i.e., host graphs whose order is linear in the size of the trees; this allows us to make use of an important tool in extremal graph theory known as the Regularity Lemma [43]. Also, the maximum degree of the trees will be restricted. The approximative constant δ, the density of the host graph and the bound on the degree of the trees are necessary for our techniques to work, but we have no evidence that these restrictions cannot be relaxed or even discarded. All the results here presented come from a joint work with Matías Pavez and Maya Stein. The results from Chapter 2 and Chapter 3 are also part of [5].

The work is organized as follows. In the first chapter we give the basic notions and definitions necessary to the understanding of the thesis, as well as any previous result used in the proofs here developed. We will devote a section of Chapter 1 to the Regularity Lemma and to some simple but useful results derived from it. In Chapter 2 we present a lemma and a proposition relative to the cutting of the trees, which will be of particular help for the results presented in Chapter 3 and may be of independent interest. Chapter 3 is dedicated to the proof of five theorems concerning the embedding of trees in graphs with a minimum degree condition and a vertex of high degree, one of which is a partial approximated version of the $\frac{2}{3}$ conjecture. In Chapter 4 we present a single result for which, despite being similar to the theorems in Chapter 3, we occupy a somewhat different technique. Finally, Chapter 5 is devoted to the Erdős-Sós conjecture. In this rather short chapter we make use of the result from Chapter 4 to derive a partial approximated version of the Erdős-Sós conjecture.

Results overview

As we mentioned in the Introduction, we will be considering graphs subject to a minimum degree condition to which we will add a maximum degree hypothesis. In Chapter 3 we study the first type of maximum degree hypothesis. Here the graphs will have a certain minimum degree and a vertex of high degree. The next two theorems are the main results from Chapter 3.

Theorem 3.2.5 Let $\delta \in (0,1)$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$ and maximum degree at least $(1+\delta)2k$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{90}}$, then T is a subgraph of G.

Theorem 3.2.8 Let $\delta \in (0,1)$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{2k}{3}$ and maximum degree at least $(1+\delta)k$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{66}}$, then T is a subgraph of G.

As you can see, Theorem 3.2.5 generalizes, in one possible direction, the result of Komlós, Sárközy and Szemerédi [30], while Theorem 3.2.8 provides support to the $\frac{2}{3}$ conjecture.

In order to simplify the proofs of the theorems from Chapter 3 we first give a general lemma, Lemma 3.2.4, describing a variety of possible structures in the host graph which enable us to perform the embedding. Thus, all we have to do is to show that at least one of the configurations from Lemma 3.2.4 appears in our host graph.

In Chapter 4 we study the second type of maximum degree hypothesis. Here we will be asking for a small portion of vertices of degree at least $(1+\delta)k$. Theorem 4.0.6 is the main result from this chapter and it can also be considered a generalization of the result of Komlós, Sárközy and Szemerédi [30].

Theorem 4.0.6 Let $\delta \in (0,1)$. There exist $n_0 = n_0(\delta) \in \mathbb{N}$ and $c = c(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph with n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$ and $|\{v \in V(G) : \deg(v) \geq (1+\delta)k\}| \geq \delta n$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most ck, then T is a subgraph of G.

As we mentioned earlier, we will present in Chapter 5 a partial approximated version of the Erdős-Sós conjecture. A well-known result in graph theory establish, for any given graph G, the existence of a subgraph $H \subset G$ maintaining the average degree of G, but with minimum degree at least $\frac{1}{2}\mathrm{d}(G)$. Thus, Theorem 5.0.1 will be a direct consequence of Theorem 4.0.6 and of the density of the host graph: a graph H with average degree linear in $|V(H)|$ must contain a small portion of vertices of high degree, i.e., of degree close to the average degree.

Theorem 5.0.1 Let $\delta \in (0,1)$. There exist $n_0 = n_0(\delta) \in \mathbb{N}$ and $c = c(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which satisfies

$$\mathrm{d}(G) > k - 1 + \delta k,$$

with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most ck, then T is a subgraph of G.

Variants of Theorem 3.2.5

In Chapter 3 we also present some variants of Theorem 3.2.5, which consider additional or alternative hypothesis. For instance, when considering trees with maximum degree bounded by a constant, we can lower the bound on the maximum degree of the host graph in Theorem 3.2.5.

Theorem 3.2.7 Let $\delta \in (0,1)$ and $\Delta \geq 2$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$ and maximum degree at least $(1+\delta)2\frac{(\Delta-1)}{\Delta}k$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most Δ, then T is a subgraph of G.

Also, in Section 3.2.3, we define a special class of graphs $\mathcal{N}_{k,\delta}$ and we show that the maximum degree condition in Theorem 3.2.5 is only needed for graphs that belong to this class, in the rest a vertex with degree at least $(1+\delta)\frac{4k}{3}$ is sufficient.

Theorem 3.2.9 Let $\delta \in (0,1)$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$, maximum degree at least $(1+\delta)\frac{4k}{3}$ and does not belong to $\mathcal{N}_{k,\delta}$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{90}}$, then T is a subgraph of G.

Another way of relaxing the maximum degree condition in Theorem 3.2.5 is by asking for a vertex with a large first and second neighborhood.

Theorem 3.2.10 Let $\delta \in (0,1)$. There exist $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph with n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$, with $n \geq k \geq \delta n$. Suppose there is a vertex $x \in V(G)$ with $|N(x)| \geq (1+\delta)\frac{4k}{3}$ and $|N_2(x)| \geq (1+\delta)\frac{4k}{3}$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{90}}$, then T is a subgraph of G.

In each of these settings we find the embedding of T by showing that at least one of the structures described in Lemma 3.2.4 appears in the host graph.

Chapter 1

Preliminaries

1.1 Basic notation

We begin with some standard graph theory notation. Given $\ell \in \mathbb{N}$, we write $[\ell] = \{1, \dots, \ell\}$. Let H be a graph, we denote by $V(H)$ and $E(H)$ the set of vertices and edges respectively, and we write $v(H) = |V(H)|$ and $\mathrm{e}(H) = |E(H)|$ for their sizes. Let $X, Y \subseteq V(H)$, we write $E_H(X, Y)$ for the set of edges $xy \in E(H)$ with $x \in X$ and $y \in Y$ and $\mathrm{e}_H(x, y)$ for the number of such edges. Given $x \in V(H)$ we write $N_H(x)$ for its set of neighbors, $\deg_H(x)$ for its cardinality and $E_H(x)$ for the set of edges incident to x. For $S \subseteq V(H)$ we write $N_H(x, S) = N_H(x) \cap S$ for the set of neighbors of x in S and $\deg_H(x, S)$ for its cardinality. If the underlying graph is clear we omit the subscript. The second neighborhood of x will be denoted by $N_2(x) := (\bigcup_{x \in N(x)} N(y)) \setminus N(x)$. The i-th neighborhood can be defined in a similar way. We write $\delta(H) := \min_{v \in V(H)} \deg(v)$, $\mathrm{d}(H) := \frac{\sum_{v \in V(H)} \deg(v)}{|V(H)|}$ and $\Delta(H) := \max_{v \in V(H)} \deg(v)$ for the minimum, average and maximum degree of H, respectively. Given a set $U \subset V(H)$ we write $H[U]$ for the graph induced in H by the vertices in U. We denote by $\mathcal{C}(H)$ the family of components of H. Let T be a tree and let $\mathcal{F}$ be a forest. The set of leaves of T will be denoted by $\ell(T)$, the same for $\mathcal{F}$. If T and $\mathcal{F}$ are rooted, $r(T)$ will stand for the root of T and $r(\mathcal{F})$ for set of roots of $\mathcal{F}$. Let $\mathcal{T}$ be a family of disjoint trees, then $\bigcup \mathcal{T}$ is the forest induced by the union of the trees in $\mathcal{T}$.

Given two vertices $x, y \in V(H)$ we denote the distance between x and y, that is, the length of the shortest path connecting x and y, by $\mathrm{dist}_H(x, y)$. The diameter of H is denoted by $\mathrm{diam}(H)$ and is defined as $\mathrm{diam}(H) := \max_{x,y \in V(H)} \mathrm{dist}_H(x, y)$.

We say that a graph H embeds in a graph G if there is an injective function $\phi : V(H) \to V(G)$ preserving adjacency, that is, if uv is an edge in H, then $\phi(u)\phi(v)$ is an edge in G.

A proper 2-coloring c of a graph G is a map $c : V(G) \to \{0, 1\}$ such that for any $uv \in E(G)$, $c(u) \neq c(v)$. For such a coloring we define $c_0 := \{v \in V(G) : c(v) = 0\}$ and $c_1 := \{v \in V(G) : c(v) = 1\}$. Throughout the rest of the thesis we will assume that $|c_0| \geq |c_1|$. For the sake of clarity, when we are talking about a tree T we will refer to its color classes as $A(T)$

and $B(T)$ and we will also assume that $|A(T)| \geq |B(T)|$.

Let us mention here a well-known fact in graph theory that we will need when working on the Erdős-Sós conjecture.

Fact 1.1.1 Let $t > 0$ and let H be a graph with $\mathrm{d}(H) \geq t$, then there exists a subgraph $H' \subset H$ with $\mathrm{d}(H') \geq t$ and $\delta(H') \geq \frac{t}{2}$.

In the next section we introduce the notion of regularity and the famous Regularity Lemma due to Endre Szemerédi.

1.2 Regularity Lemma

We begin by presenting the concept of regular pair.

Definition 1.2.1 Let $H = (A, B; E)$ be a bipartite graph with density $\mathrm{d}(A, B) := \frac{\mathrm{e}(A,B)}{|A||B|}$. Let $\varepsilon > 0$. We say that the pair (A, B) is ε-regular if for any $X \subseteq A$ and $Y \subseteq B$, with $|X| > \varepsilon|A|$ and $|Y| > \varepsilon|B|$, we have that

$$|\mathrm{d}(X, Y) - \mathrm{d}(A, B)| < \varepsilon.$$

It turns out that regular pairs behave, in many ways, like random bipartite graphs with the same edge density. Given (A, B) an ε-regular pair with density d, we say that a subset $X \subseteq A$ is significant if $|X| > \varepsilon|A|$, and similar for subsets of B. A vertex $x \in A$ is called typical to a significant set $Y \subseteq B$ if $\deg(x, Y) > (d - \varepsilon)|Y|$. The next fact states that in a regular pair almost every vertex is typical to any given significant set, and also that regularity is inherited by subpairs.

Fact 1.2.2 Let (A, B) be an ε-regular pair with density d. The following holds:

1. For any significant $Y \subset B$, all but at most $\varepsilon|A|$ vertices from A are typical to Y.
2. Let $\delta \in (0, 1)$. For any subsets $X \subset A$ and $Y \subset B$, with $|X| \geq \delta|A|$ and $|Y| \geq \delta|B|$, the pair (X, Y) is $\frac{2\varepsilon}{\delta}$-regular with density between $d - \varepsilon$ and $d + \varepsilon$.

The Regularity Lemma states that the vertex set of any large graph can be partitioned into a bounded number of clusters, such that the graph induced by almost any pair of those clusters is ε-regular, for a given $\varepsilon > 0$. Let us now state the Regularity Lemma in a precise form.

Theorem 1.2.3 (Regularity Lemma) Let $\varepsilon > 0$ and let $m_0 \in \mathbb{N}$. There exists n_0 and M_0, depending only on ε and m_0, such that the following is true. For every graph G on $n \geq n_0$ vertices there exists a vertex partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_\ell$, where $m_0 \leq \ell \leq M_0$, such that

1. $|V_0| \leq \varepsilon n$,

2. $|V_1| = |V_2| = \cdots = |V_\ell|$ and

3. all but at most $\varepsilon\ell^2$ of the pairs (V_i, V_j) are ε-regular.

Once the Regularity Lemma is applied to a large graph G it is more comfortable to work with a subgraph $G_d \subseteq G$, usually called *pure graph*, that is simpler to analyze and, furthermore, it approximates very well G in the sense that $|E(G)| - |E(G_d)| < \alpha|V(G)|^2$, where α is a constant that can be arbitrarily lowered. The pure graph is obtained by deleting the vertices in V_0, the edges inside the clusters, the edges of the pairs of clusters that are not regular and the edges of those pairs that have a density below some given threshold $d \geq 0$. Let us say $n := |V(G)|$ and $s := |V_i|$. During this deletion process we throw away

$$|E(G - G_d)| < \varepsilon n^2 + \ell\frac{s^2}{2} + \varepsilon\ell^2 s^2 + \frac{\ell^2}{2}ds^2 \leq \left(\frac{1}{2m_0} + 2\varepsilon + \frac{d}{2}\right)n^2$$

edges and, therefore, we could say that $\alpha = \frac{1}{2m_0} + 2\varepsilon + \frac{d}{2}$, which are all parameters that can be chosen. We say then that the pure graph admits an (ε, d)-upper regular partition, more precisely: a vertex partition $V(G_d) = V_1 \cup \cdots \cup V_\ell$ is called (ε, d)-upper regular if

1. $|V_1| = |V_2| = \cdots = |V_\ell|$,

2. V_i is independent for all $i \in [\ell]$, and

3. for all $1 \leq i < j \leq \ell$, (V_i, V_j) is ε-regular with density either $\mathrm{d}(V_i, V_j) > \mathrm{d}$ or $\mathrm{d}(V_i, V_j) = 0$.

We state now another version of the lemma, which will be specially useful for us. This slightly different version was published by Kühn and Osthus [34] in 2009 and has the additional advantage of providing a bound for the degree of each vertex, which will be of particular help when working with graphs subject to a minimum degree condition.

Proposition 1.2.4 Let $\varepsilon > 0$ and let $m_0 \in \mathbb{N}$. There exists N_0, M_0, depending only on ε and m_0, such that the following holds. Let $d \in [0, 1]$ and let G be any graph on $n \geq N_0$ vertices. There exists a subgraph $G' \subset G$, with $|V(G) \setminus V(G')| \leq \varepsilon n$ and $\deg_{G'}(x) \geq \deg_G(x) - (d + \varepsilon)n$ for all $x \in V(G')$, such that G' admits an (ε, d)-upper regular partition $V(G') = V_1 \cup \cdots \cup V_\ell$, where $m_0 \leq \ell \leq M_0$.

We have already seen that in a regular pair most of the vertices are typical to a given significant set. The next lemma generalizes that result for more than one significant set.

Definition 1.2.5 Let $\varepsilon > 0$ and let $d > 0$. Let G be a graph that admits an (ε, d)-upper regular partition. Let $\mathcal{C}$ be a family of significant sets and consider $v \in V(G)$. We define $\mathcal{T}_v(\mathcal{C}) := \{C \in \mathcal{C} : v \text{ is typical to } C\}$.

Lemma 1.2.6 Let $\varepsilon \in (0, 1)$ and let $d > 0$. Let G be a graph that admits a (ε, d)-upper regular partition. Let $\mathcal{Y}$ be a family of significant sets. For any cluster C,

$$|\mathcal{T}_v(\mathcal{Y})| \geq (1 - \sqrt{\varepsilon})|\mathcal{Y}|$$

for at least $(1 - \sqrt{\varepsilon})|C|$ vertices $v \in C$.

Proof. Suppose that there is a cluster C and a set $\overline{C} \subset C$ with $|\overline{C}| \geq \sqrt{\varepsilon}|C|$ such that $|\mathcal{T}_v(\mathcal{Y})| < (1-\sqrt{\varepsilon})|\mathcal{Y}|$, for each $v \in \overline{C}$. Then,

$$\begin{aligned}\sum_{Y\in\mathcal{Y}} |\{v \in C : v \text{ is not typical to } Y\}| &= \sum_{v\in C} |\{Y \in \mathcal{Y} : v \text{ is not typical to } Y\}| \\ &\geq \sum_{v\in\overline{C}} |\{Y \in \mathcal{Y} : v \text{ is not typical to } Y\}| \\ &\geq |\overline{C}|\sqrt{\varepsilon}|\mathcal{Y}| \\ &> \varepsilon|C||\mathcal{Y}|\end{aligned}$$

This means that there is a cluster $Y \in \mathcal{Y}$ such that more than $\varepsilon|C|$ vertices in C are not typical to Y, a contradiction. □

Let G be a graph admitting an (ε, d)-upper regular partition for some $\varepsilon > 0$ and some $d > 0$. We will say that two clusters are adjacent if they share common edges. Using this definition of adjacency we can define a graph G_r called the *reduced graph*, whose vertex set is the family of clusters in G and where two clusters are connected by an edge if they are adjacent in G. We will speak about some cluster structures such as cluster-matchings, cluster-triangles, cluster-paths, cluster-walks and cluster-cycles; each of these structures is defined as the corresponding object in the reduced graph. For a family of clusters $\mathcal{C}$ we will write $V(\mathcal{C}) := \bigcup_{C\in\mathcal{C}} C$, also for a cluster-matching M we will write $V(M) := \bigcup_{(A,B)\in M}(A \cup B)$ and similarly for other cluster structures. Sometimes we will speak about the weighted degree of a cluster C which we define as follows,

$$\overline{\deg}(C) := \frac{\mathrm{e}(C, V(G) \setminus C)}{|C|}.$$

Also, we define the weighted degree of a cluster C towards a family of clusters or towards a cluster structure $\mathcal{C}$ as

$$\overline{\deg}_{\mathcal{C}}(C) := \frac{\mathrm{e}(C, V(\mathcal{C}))}{|C|}.$$

1.3 Trees

Let T be a rooted tree. We define a partial order $\preceq$ on $V(T)$ in the following way. We say that $x \preceq y$ if and only if y lies in the unique path from x to $r(T)$, if $x \preceq y$ we will say that x is below y. Given $x \in V(T)$, we say that $v \in V(T)$ is a child of x if $xv \in E(T)$ and v is below x. Given a vertex $x \in V(T)$, we define the tree induced by x, denoted by $T(x)$, as the subtree of T with vertex set $V(T(x)) = \{v : v \preceq x\}$. For $i \geq 0$ we may define $L_i \subseteq V(T)$, the i-th level of T, as those vertices that are at distance i from the root.

The following lemma gives an idea of why regularity is so useful for the task of embedding trees. It ensures that we can embed small trees in sufficiently large regular pairs.

Lemma 1.3.1 Let (A, B) be a ε-regular pair with $\mathrm{d}(A, B) > d$ and size $|A| = |B| = s$. Let $X \subset A, Y \subset B$ be subsets with more than αs vertices, where $\alpha = \frac{\beta+4\varepsilon}{d-\varepsilon}$, and let T be a tree on βs vertices, where $0 < \beta \leq \varepsilon$. Then T can be embedded into $X \cup Y$ by a function $\phi : T \to X \cup Y$ such that all but εs vertices can be chosen as the root of T, and for any other $x \in V(T)$, $\phi(x)$ is chosen from a set of more than $3\varepsilon s + 1$ vertices.

Proof. Let T be a tree on βs vertices rooted at $r \in V(T)$, we will construct an embedding $\phi : V(T) \to X \cup Y$ sequentially through the levels of T. We embed r into a typical vertex of $X \cup Y$, let suppose that $\phi(r) \in X$ so it will be typical to Y. Assume that $L_0, \ldots, L_i$ are already embedded, where i is without loss of generality even, in a way such that $\deg(\phi(x), Y') \geq (d - \varepsilon)|Y'|$ for all $x \in L_i$, where X' and Y' are the set of unoccupied vertices in X and Y respectively. Let v be any vertex in L_{i+1} and let $x \in L_i$ be its father. We have to select a vertex $u \in N(\phi(x), Y')$ which is typical to X'. Since at most εs vertices from Y' are not typical to X' and at most $|L_{i+1}| - 1$ vertices of the neighborhood of $\phi(x)$ in Y' may be occupied, then $\phi(y)$ can be chosen from at least

$$|N(\phi(x), Y')| - (|L_{i+1})| - 1) - \varepsilon s \geq (d - \varepsilon)|Y'| - |L_{i+1}| - \varepsilon s + 1$$

vertices. Notice that

$$(d-\varepsilon)|Y'|-|L_{i+1}| \geq (d-\varepsilon)(|Y|-\sum_{t \leq i} |L_t|)-|L_{i+1}| > (d-\varepsilon)|Y|-\sum_{t \leq i+1} |L_t| \geq (d-\varepsilon)|Y|-|V(T)|.$$

Since $\alpha = \frac{\beta + 4\varepsilon}{d - \varepsilon}$, it implies that $(d - \varepsilon)|Y| \geq \beta s + 4\varepsilon s \geq |V(T)| + 4\varepsilon s$ and thus the number of choices of $\phi(y)$ is at least $3\varepsilon s + 1$. $\square$

Remark 1.3.2 It is important that $\phi(y)$ can be chosen in at least $3\varepsilon s + 1$ ways, because for some $y \in V(T)$ we will need to choose $\phi(y)$ not only typical to X or Y, but also typical to other two clusters D_1 and D_2 in order to continue the embedding of the tree.

We now state a simple lemma on tree cutting and then a numeric lemma that will simplify the treatment of the pieces resulting from the cut.

Lemma 1.3.3 [38] Let T be a tree on t edges. There is a vertex $z \in V(T)$ such that every component of $T - z$ has $\lceil \frac{t}{2} \rceil$ or fewer vertices.

Proof. Choose any vertex $x \in V(T)$ to be the root of T. Let z be the vertex minimal with respect to the order given by the root x such that $|V(T(z))| > \frac{t}{2}$. Note that each component below z covers at most $\lfloor \frac{t}{2} \rfloor$ vertices. From the choice of z we also have that $|V(T - T(z))| \leq \lceil \frac{t}{2} \rceil$. $\square$

Lemma 1.3.4 Let $t \in \mathbb{N}$ and let $(a_i)_{i=1}^m$ a sequence of positive integers with $m \geq 1$ such that $0 < a_i \leq t/2$, for each $i = 1 \ldots, m$, and $\sum_{i=1}^m a_i \leq t$. Then:

1. There is a partition $\{I_1, I_2, I_3\}$ of the set $[m]$ such that $\sum_{i \in I_k} a_i \leq \frac{t}{2}$, for $k = 1, 2, 3$, and $\sum_{i \in I_3} a_i \leq \sum_{i \in I_2} a_i \leq \sum_{i \in I_1} a_i$.
2. There is a partition $\{J_1, J_2\}$ of the set $[m]$ such that $\sum_{i \in J_1} a_i \leq \frac{2t}{3}$, $\sum_{i \in J_2} a_i \leq \frac{t}{2}$ and $\sum_{i \in J_2} a_i \leq \sum_{i \in J_1} a_i$.

Proof. We first pick a set $I_1 \subset [m]$ with $\sum_{i\in I_1} a_i \leq \frac{t}{2}$ that maximizes the sum. From $[m]\setminus I_1$ we extract a second set I_2 with $\sum_{i\in I_2} a_i \leq \frac{t}{2}$ that maximizes the sum. The choice of I_1, I_2 ensures that for $I_3 := [m] \setminus (I_1 \cup I_2)$ also holds that $\sum_{i\in I_3} a_i \leq \frac{t}{2}$. These sets fulfill condition (1). Notice that I_2 and I_3 may be empty.

Let us call $S_k := \sum_{i\in I_k} a_i \leq \frac{t}{2}$ and $S := S_1 + S_2 + S_3$. Observe that $S_1 \geq \min\{S, \frac{t}{3}\}$ and, therefore, $S_2 + S_3 \leq \frac{2t}{3}$. Thus, if $I_3 \neq \varnothing$ we set $J_1 := I_2 \cup I_3$ and $J_2 := I_1$. When $I_3 = \varnothing$ we just set $J_1 := I_1$ and $J_2 := I_2$, which trivially satisfy point (2). □

Remark 1.3.5 Observe that I_3 can have at most one element, otherwise, due to the maximality of I_1 and I_2, there would be $j, k \in I_3$ such that $a_j + \sum_{i\in I_1} a_i > \frac{t}{2}$ and $a_k + \sum_{i\in I_2} a_i > \frac{t}{2}$, which contradicts the fact that $\sum_{i=1}^{m} a_i \leq t$.

Lemma 1.3.4 tells us that after cutting a tree with Lemma 1.3.3 (1) we can form groups of trees with the components of $T - z$ having some control over the sizes of the groups. In the next chapter we study the "balance" of the resulting forest.

Chapter 2

Results on tree cutting

After cutting a tree with Lemma 1.3.3 we would like to say something about the "balance" of the resulting forest, and for this we resort to the concept of vertex coloring. Recall that for a proper 2-coloring c of a graph G, $c_0 = \{v \in V(G) : c(v) = 0\}$ and $c_1 = \{v \in V(G) : c(v) = 1\}$. Remember also that $|c_0| \geq |c_1|$. When we are talking about a tree T we refer to its color classes as $A(T)$ and $B(T)$ and we also assume that $|A(T)| \geq |B(T)|$.

Lemma 2.0.1 Let T be a tree on $t+1$ vertices. There exists $z \in V(T)$ and a proper 2-coloring $c : V(T-z) \to \{0,1\}$ of $T-z$ such that $|c_0| \leq \frac{3t}{4}$ and $|c_1| \leq \frac{t}{2}$.

Proof. Let us apply Lemma 1.3.3 to obtain a cut vertex z and a forest $T-z$ with components $\{T_i\}_i^m$ such that $|T_i| \leq \frac{t}{2}$, for every i. To simplify things we will use Lemma 1.3.4 to group the components of $T-z$: setting $a_i := |T_i|$ we get three sets I_1, I_2 and I_3 such that the forests $F_j := \bigcup_{i \in I_j} T_i$, with $j = 1,2,3$, cover less than $\frac{t}{2}$ vertices each. For $j = 1,2,3$ consider an arbitrary 2-coloring c^j of the forest F_j.

Suppose first that the heavier color class of F_1, c_0^1, satisfy $|c_0^1| > \frac{3}{4}|F_1|$. Define $c_0 := c_0^1 \cup c_1^2 \cup c_1^3$ and $c_1 := V(T-z) \setminus c_0 = c_1^1 \cup c_0^2 \cup c_0^3$. This yields:

$$|c_0| \leq |c_0^1| + |c_1^2| + |c_1^3| \leq |F_1| + \frac{|F_2|}{2} + \frac{|F_3|}{2} = \frac{|F_1|}{2} + \frac{|T-z|}{2} \leq \frac{t}{4} + \frac{t}{2} = \frac{3t}{4}.$$

Moreover,

$$|c_1| \leq t - |c_0^1| \leq t - \frac{3|F_1|}{4} \leq t - \frac{t}{4} = \frac{3t}{4}$$

where the last inequality comes from the fact that $|F_1| \geq \frac{t}{3}$. These two bounds ensure that $\max\{|c_0|, |c_1|\} \leq \frac{3t}{4}$; renaming the color classes if necessary we get the result.

We can now assume that $|c_0^1| \leq \frac{3|F_1|}{4}$. Define $c_0 := c_0^1 \cup c_1^2 \cup c_0^3$ and $c_1 := V(T-z) \setminus c_0 = c_1^1 \cup c_0^2 \cup c_1^3$:

$$|c_0| \leq \frac{3|F_1|}{4} + \frac{|F_2|}{2} + |F_3| = \frac{2|T-z|}{4} + \frac{|F_1| + 2|F_3|}{4} \leq \frac{t}{2} + \frac{|F_1| + |F_2| + |F_3|}{4} = \frac{3t}{4}.$$

For c_1 we have:

$$|c_1| \leq \frac{|F_1|}{2} + |F_2| + \frac{|F_3|}{2} = \frac{t}{2} + \frac{|F_2|}{2} \leq \frac{3t}{4}.$$

Again we obtain $\max\{|c_0|, |c_1|\} \leq \frac{3t}{4}$. □

Lemma 2.0.1 states that we can always cut a tree in a somewhat balanced manner. If the cut is made in a way that all the resulting trees cover at most half of the total vertices, then the fraction $\frac{3}{4}$ in Lemma 2.0.1 is not replaceable by a smaller one: consider the tree obtained from the union of a path of order $\frac{t}{2}+2$ and a star of order $\frac{t}{2}$ such that the central vertex of the star is one of the end vertices of the path and this is the only vertex they share; cutting this tree would result in a forest composed by a path of order $\frac{t}{2}$ and a star of order $\frac{t}{2}$; one of the color classes of this forest will always contain $\frac{3t}{4}-1$ vertices. This last example shows that if we want to cut our tree so that the resulting forest is even more balanced, then we need to cut it in a different fashion. The following proposition deals with this problem.

Definition 2.0.2 Given a graph G and a proper 2-coloring of its vertex set $c : V(G) \to \{0,1\}$ we define the imbalance of c as $\sigma(c) := |c_0| - |c_1|$. For a tree T we will use $\sigma(T)$ to denote the imbalance of its unique 2-coloring, i.e., $\sigma(T) := |A(T)| - |B(T)|$.

Proposition 2.0.3 Let T be a tree on $t+1$ vertices. There exists $z \in V(T)$ and a proper 2-coloring $c : V(T-z) \to \{0,1\}$ of $T-z$ such that $|c_0| \leq \frac{2t}{3}$ and $|c_1| \leq \frac{t}{2}$.

Proof. Consider $z_0 \in V(T)$ and $c : V(T-z_0) \to \{0,1\}$ from Lemma 2.0.1, this means that the heavier color class induced by c, c_0, contains at most $\frac{3t}{4}$ vertices. If there were a 2-coloring of $T-z_0$ with both color classes containing less than $\frac{2k}{3}$ vertices, we would be done, therefore we will assume that the heavier color class of any 2-coloring of $T-z_0$ covers at least $\frac{2t}{3}$ vertices, in particular we have

$$|c_0| \geq \frac{2t}{3} \text{ and } |c_1| \leq \frac{t}{3}. \tag{2.1}$$

Let us call $\mathcal{R} \subseteq \mathcal{C}(T-z_0) = \{T_i\}_{i\in I}$ the collection of components of $T-z_0$ with its heavier color class contained in c_0, i.e., $T_i \in \mathcal{R}$ if and only if $A(T_i) \subseteq c_0$, and let $R \subset I$ be the set of its indices. For the sake of order we state a few facts before proving the result.

Fact 2.0.4 For every $P \subseteq R$ either $\sum_{i\in P} \sigma(T_i) < \frac{t}{12}$ or $\sum_{i\in P} \sigma(T_i) > \frac{t}{3}$.

If this were not true, we could invert the colors in the trees contained in a subfamily $\{T_i\}_{i\in P} \subseteq \mathcal{R}$ with $\frac{t}{12} \leq \sum_{i\in P} \sigma(T_i) \leq \frac{t}{3}$; this would yield a coloring c' such that $|c'_0| = |c_0| - \sum_{i\in P} \sigma(T_i) \leq \frac{3t}{4} - \frac{t}{12} = \frac{2t}{3}$ and $|c'_1| = |c_1| + \sum_{i\in P} \sigma(T_i) \leq \frac{2t}{3}$, where we have used (2.1). This contradicts our first assumption.

Obviously, since c is an unbalanced coloring, it must happen that $\sum_{i\in R} \sigma(T_i) > \frac{t}{3}$. From Fact 2.0.4 we can also derive the following,

$$\text{there is no } P \subseteq R \text{ such that } \sigma(T_i) < \frac{t}{12} \text{ for every } i \in P \text{ and } \sum_{i\in P} \sigma(T_i) \geq \frac{t}{12}, \tag{2.2}$$

because we could easily find a set $Q \subseteq P$ with $\frac{t}{12} \leq \sum_{i\in Q} \sigma(T_i) \leq \frac{t}{3}$. In particular, we deduce that there is a tree T_{i_1} in $\mathcal{R}$ with $\sigma(T_{i_1}) > \frac{t}{3}$. Notice that this unbalanced tree must be unique, for if there were two such trees we could switch color classes in one of them and obtain a contradiction to the initial assumption. We will assume that T_1 is the unique component in $T - z_0$ of great imbalance. Thus, $\sigma(T_i) < \frac{t}{12}$ for every $i \in R \setminus \{1\}$, and making use of (2.2) we have

$$\sum_{i \in R \setminus \{1\}} \sigma(T_i) \le \frac{t}{12}. \tag{2.3}$$

We are now ready to state our second fact.

Fact 2.0.5 $\sum_{i \in I \setminus \{1\}} \sigma(T_i) \le \frac{t}{6}$.

If $\sum_{i \in I \setminus \{1\}} \sigma(T_i) > \frac{t}{6}$, we can switch colors in every tree of the family $\mathcal{R} \setminus \{T_1\}$ to obtain a 2-coloring c' of $T - z_0$ with $|c'_1| < \frac{t}{3} + \frac{t}{12} < \frac{t}{2}$, where we have used (2.1) and (2.3). Besides,

$$|c'_1| = |B(T_1)| + \sum_{i \in I \setminus \{1\}} |A(T_i)| \ge \sum_{i \in I \setminus \{1\}} \left(\frac{|T_i|}{2} + \frac{\sigma(T_i)}{2} \right) > \frac{t - |T_1|}{2} + \frac{t}{12} \ge \frac{t}{4} + \frac{t}{12} = \frac{t}{3},$$

which implies $\frac{t}{2} < |c'_0| < \frac{2t}{3}$, a contradiction.

Now, apply Lemma 1.3.3 to obtain $z_1 \in V(T_1)$ such that every component of $T_1 - z_1$ covers at most $\frac{|T_1|-1}{2} < \frac{t}{4}$ vertices. We call these components $\{T_{1,i}\}_{i \in I_1}$. Let us call T_{z_0} the component of $T - z_1$ that contains z_0 and let us say that $T_{1,1}$ is the unique component of $T_1 - z_1$ that is contained in T_{z_0} ($T_{1,1}$ may be empty). Use Lemma 1.3.4 (2) to group the components $\{T_{1,j}\}_{j>1}$ into two forests F^A and F^B fulfilling

$$|F^A| \le \frac{2(|T_1| - 1)}{3} \le \frac{t}{3} \quad \text{and} \quad |F^B| \le \frac{|T_1| - 1}{2} \le \frac{t}{4}. \tag{2.4}$$

For $i = A, B$ consider c^i a proper 2-coloring of F^i that maximizes $\sigma(c^i)$. We state one final fact before completing our proof.

Fact 2.0.6 $\sigma(c^A) + \sigma(c^B) > \frac{t}{12}$.

This comes simply from the fact that $\sigma(T_1) < |T_{1,1}| + \sigma(c^A) + \sigma(c^B)$, from which we obtain $\sigma(c^A) + \sigma(c^B) > \sigma(T_1) - |T_{1,1}| \ge \frac{t}{12}$. Now we consider and treat separately two possible cases.

Case 1: $\sigma(T_{z_0}) \le \frac{t}{3}$.

If $\sigma(c^A) \ge \sigma(c^B)$, define $c'_0 := A(T_{z_0}) \cup c^A_1 \cup c^B_0$ and $c'_1 := (T - z_1) \setminus c'_0 = B(T_{z_0}) \cup c^A_0 \cup c^B_1$. This implies

$$|c'_0| = \frac{|T_{z_0}|}{2} + \frac{\sigma(T_{z_0})}{2} + \frac{|F^A|}{2} - \frac{\sigma(c^A)}{2} + \frac{|F^B|}{2} + \frac{\sigma(c^B)}{2} \le \frac{t}{2} + \frac{\sigma(T_{z_0})}{2} \le \frac{2t}{3}$$

where we have used the bound on the imbalance of T_{z_0}. Also, by (2.4),

$$|c'_1| < \frac{|T_{z_0}|}{2} + |F^A| + \frac{|F^B|}{2} = \frac{t}{2} + \frac{|F^A|}{2} \le \frac{2t}{3}.$$

If $\sigma(c^B) \ge \sigma(c^A)$, define $c'_0 := A(T_{z_0}) \cup c^A_0 \cup c^B_1$ and $c'_1 := (T - z_1) \setminus c'_0$. Again we obtain $|c'_0| \le \frac{2t}{3}$ and $|c'_1| \le \frac{2t}{3}$. Thus, c' fulfills the balance condition we are looking for.

Case 2: $\sigma(T_{z_0}) > \frac{t}{3}$.

This time we define $c_0' := A(T_{z_0}) \cup c_1^A \cup c_1^B$ and $c_1' := B(T_{z_0}) \cup c_0^A \cup c_0^B$. Recalling Facts 2.0.5 and 2.0.6 we get

$$|c_0'| = \frac{t}{2} + \frac{\sigma(T_{z_0})}{2} - \frac{\sigma(c^A) + \sigma(c^B)}{2} \le \frac{t}{2} + \frac{\sigma(T_{1,1}) + \sum_{i \in I \setminus \{1\}} \sigma(T_i) + 1}{2} - \frac{t}{24} \le \frac{t}{2} + \frac{t}{8} + \frac{t}{12} - \frac{t}{24} = \frac{2t}{3}$$

and, by (2.4),

$$|c_1'| < \frac{|T_{z_0}|}{2} - \frac{\sigma(T_{z_0})}{2} + |F^A| + |F^B| < \frac{t}{2} + \frac{|F^A| + |F^A|}{2} - \frac{t}{6} < \frac{2t}{3},$$

which again yields $\max\{c_0', c_1'\} \le \frac{2t}{3}$, finding with this the wanted coloring and completing the proof. $\square$

Observe that the bound $\frac{2t}{3}$ given by Proposition 2.0.3 is almost best possible, in the sense that the fraction $\frac{2}{3}$ is not replaceable by a smaller one. To see this it is enough to consider the tree consisting of a vertex connected to the central vertices of three stars of order $\frac{t}{3}$.

Chapter 3

Embedding of trees in graphs with minimum degree condition

In this chapter we study the embedding of trees of bounded degree into graphs subject to a minimum degree condition and which contain a vertex of high degree. We will first establish some results working on connected regularized graphs to then move on to the general case, where the regularized graph might not be connected. In this case the high degree vertex will play a fundamental role, as it will act as a link between different components of the host graph.

3.1 Embedding of trees in large connected graphs

The results of this section work for large connected graphs that admit a regular partition. These propositions are meant to be used once the Regularity Lemma has been applied to our host graph. We treat separately the bipartite case and the non-bipartite case. Then, in Proposition 3.1.14, we show how to improve the bound on the degree of the trees.

3.1.1 Cutting the trees into small pieces

As we saw in Lemma 1.3.1 we can easily map small trees into regular pairs. Thus, when we want to embed a tree T into a regularized graph, it is useful to cut down T into small subtrees. We present here a proposition showing that any sufficiently large tree can be decomposed into a family of small subtrees connected by few vertices.

Proposition 3.1.1 Let $\beta \in (0,1)$. There exists $t_0 = t_0(\beta) \in \mathbb{N}$ such that for all $t \geq t_0$ the following holds. Let T be a rooted tree on $t+1$ vertices. There exist a set of *seeds* $S \subset V(T)$ and a family $\mathcal{P}$ of disjoint rooted trees which will be called *pieces*, such that

(i) $r(T) \in S$;

(ii) $\mathcal{P}$ consists of the components of $T - S$;

(iii) each piece in $\mathcal{P}$ has at most βt vertices; and

(iv) $|S| < \frac{1}{\beta} + 2$.

Proof. We iteratively construct the set S, starting with $T^0 := T$ and $S^0 := \varnothing$. In step $i+1$, let s_{i+1} be the minimal vertex of T^i (minimal with respect to the order defined in 1.3) such that

$$|T^i(s_{i+1})| > \beta t.$$

Note that by the minimality of s_{i+1} the trees in $T^i(s_{i+1}) - s_{i+1}$ each cover at most βt vertices. Obtain S_{i+1} by adding s_{i+1} to S^i and set $T^{\mathrm{i}+1} = T^i - T^i(s_{i+1})$. If at some step j there is no vertex s_{j+1} with $|T^j(s_{j+1})| > \beta t$, then $|T^j| \leq \beta t$, and we end the process. We set $S := S^j \cup \{r(T)\}$ and $\mathcal{P} := \mathcal{C}(T - S)$.

Properties (i)–(iii) clearly hold. For (iv) observe that $|T^{i+1}| < |T^i| - \beta t$. Hence,

$$0 \leq |T^m| \leq |T^0| - j \cdot \beta t$$

which in turn implies that $|S| = j + 1 \leq \frac{|T|}{\beta t} + 1 < \frac{1}{\beta} + 2$. □

3.1.2 Bipartite case

We said that a graph with minimum degree at least k contains every tree on k edges as a subgraph, and that this can be proved with a simple greedy argument. In the case where the host graph is bipartite the minimum degree condition can be relaxed and the greedy argument will continue to work. Consider a bipartite graph $G = X \cup Y$. If $\deg(x) \geq \lfloor \frac{k}{2} \rfloor$ for all $x \in X$, and $\deg(y) \geq k$ for all $y \in Y$, then each tree T with k edges is a subgraph of G. This weakened hypothesis works because the lighter color class of T, $B(T)$, contains at most $\lfloor \frac{k}{2} \rfloor$ vertices and, if the embedding is done in a way that $\phi(B(T)) \subset Y$, the minimum degree of X will always suffice to find enough unoccupied vertices in Y.

The next proposition shows that when the trees have bounded degree and G admits a regular partition, it suffices to ask for a minimum degree greater than $\frac{k}{2}$ in only one of bipartition classes, subject to the size of that class being sufficiently large.

Proposition 3.1.2 Let $\varepsilon \in (0, 10^{-4})$ and let $d \in \mathbb{N}$. There exists $k_0 = k_0(\varepsilon) \in \mathbb{N}$ such that for all $k \geq k_0$ the following holds. If $G = (X, Y; E)$ is a connected bipartite graph with an $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition, and corresponding reduced graph G_r, such that

(i) $\mathrm{diam}(G_r) \leq d$;

(ii) $\deg(x) \geq (1 + 100\sqrt{\varepsilon})\frac{k}{2}$, for all $x \in X$; and

(iii) $|X| \geq (1 + 100\sqrt{\varepsilon})k$,

then G contains any tree T with k edges and $\Delta(T) \leq k^{\frac{1}{d+1}}$ as a subgraph.

Proof. Let $X = X_1 \cup \cdots \cup X_s$ and $Y = Y_1 \cup \cdots \cup Y_t$ be the $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition of G, and m denote the cardinality of each cluster. For each $i \in [s]$, we arbitrarily partition X_i into sets $X_{i,S}$, $X_{i,L}$, $X_{i,C}$; and for each $j \in [t]$ we arbitrarily partition Y_j into $Y_{j,S}$, $Y_{j,L}$, $Y_{j,C}$, such that

$$|X_{i,S}| = |X_{i,L}| = |Y_{j,S}| = |Y_{j,L}| = \lceil 10\sqrt{\varepsilon} m \rceil.$$

The letters S, L and C stand for seeds, links and clusters, respectively (sets $X_{i,C}$ and $Y_{j,C}$ contain the bulk of the clusters). We call these subsets the L-, S- or C-slice of the corresponding cluster.

Note that, by Fact 1.2.2, for every (X_i, Y_j) with positive density, each of the pairs $(X_{i,K}, Y_{j,K'})$, with $K, K' \in \{S, L, C\}$, is $\frac{\sqrt{\varepsilon}}{5}$-regular with density greater than $4\sqrt{\varepsilon}$.

Root T at any vertex $r(T)$. By Proposition 3.1.1, with parameters $\beta = \frac{\varepsilon}{s+t}$, we obtain a decomposition of T into a collection of pieces $\mathcal{P}$, each of order at most βk, and a family of seeds S of size at most $\frac{2}{\beta}$. Order the elements from $S \cup \mathcal{P}$ in a way that the first element is $r(T)$, and the parent of each element is either an earlier seed or belongs to an earlier piece.

Our plan is to embed the elements from $S \cup \mathcal{P}$ in this order. Seeds will go to appropriate slices $X_{i,S}$ or $Y_{j,S}$, with $r(T)$ going to a cluster $X_i \subseteq X$ if $r(T)$ belongs to the heavier bipartition class of T, and going to a cluster from Y otherwise.

Pieces from $\mathcal{P}$ will go into slices $(X_{i,C}, Y_{j,C})$ of appropriate pairs (X_i, Y_j), and into L-slices of other clusters. More precisely, for each piece $P \in \mathcal{P}$ we will find a pair (X_i, Y_j) such that there is enough space left in $(X_{i,C}, Y_{j,C})$ to accommodate P. At this point, the parent of P is already embedded into some cluster Z, so we need to embed part of P into a path $ZZ_0Z_1Z_2 \ldots Z_h$ that connects Z with the pair (X_i, Y_j). Because of the bounded degree of T, and since the diameter of G is also bounded, this path can be chosen short enough to ensure that the levels of P that are embedded into this path only contain a small fraction of k. So we can use the L-slices of the clusters Z_ℓ for these levels. The remaining levels of P will be embedded into the free space of $(X_{i,C}, Y_{j,C})$.

Let us make this sketch more precise. During the embedding procedure, we will write $X'_{i,C}$ and $Y'_{j,C}$ for the set of unoccupied vertices of $X_{i,C}$ and $Y_{j,C}$ respectively. We will say that a pair (X_i, Y_j) is *good* if $\mathrm{d}(X_i, Y_j) > \sqrt{\varepsilon}$ and $\min\{|X'_{i,C}|, |Y'_{j,C}|\} \geq 5\sqrt{\varepsilon} m$. Hence we will be able to apply Lemma 1.3.1 to any good pair and any piece belonging to $\mathcal{P}$.

The embedding $\phi : V(T) \to V(G)$ will be constructed iteratively, following the embedding order of $S \cup \mathcal{P}$ chosen above. Employing the strategy explained above, we make sure that at every step, the following conditions will be satisfied:

(A) Each vertex is embedded into a neighbor of the image of its already embedded parent;

(B) each $s \in S$ is embedded into the S-slice of some cluster;

(C) for each $P \in \mathcal{P}$, the first (up to d) levels are embedded into the L-slices of some clusters, and the rest goes into the C-slices; and

(D) every $v \in V(T)$ is mapped into a vertex that is typical towards both the S-slice and the

L-slice of some adjacent cluster.

Since the set S has constant size, and since we do not particularly care[1] into which cluster a seed goes, as long as it goes to the S-slice, it is clearly possible to embed a seed s, when its time comes, satisfying conditions (A), (B) and (D).

So assume we are about to embed a piece $P \in \mathcal{P}$. The parent of the root $r(P)$ of P is already embedded into some vertex that is typical with respect to the L-slice of some cluster Z_0. In order to be able to embed P according to our plan, it suffices to ensure that

(I) there exists some good pair (X_i, Y_j);

(II) there is a path $Z_0Z_1Z_2 \dots Z_h$ of length $h \le d$ from Z_0 to one of X_i, Y_j;

(III) the union of the first h levels of P is small enough to fit into the free space in the L-slices of $\{Z_0, Z_1, Z_2, \dots, Z_{h-1}\}$.

For the moment, assume (I) holds. Note that then (II) holds because of condition (i) of Proposition 3.1.2. Let us prove (III).

Using (C) for already embedded pieces P', and using the fact that, for any such piece P', the number of vertices in their first d levels is bounded by $2(\Delta(T)-1)^{d-1}$ (except if $\Delta(T) \le 2$, in which case this number is bounded by d), we have that the total number of occupied vertices in L-slices is at most

$$|S| \cdot \Delta(T) \cdot 2(\Delta(T)-1)^{d-1} \le \frac{4}{\beta} \cdot k^{\frac{d}{d+1}} < \varepsilon m$$

for k sufficiently large. In particular, each L-slice of a cluster Z_ℓ has at least $\lceil 9\sqrt{\varepsilon} m \rceil$ unused vertices. This is enough to ensure that we can embed each vertex of the first h levels of P into the L-slices of the clusters Z_0, Z_1, $Z_2, \dots, Z_{h-1}$ in a way that (A) and (D) hold. This proves (III). Finally, repeatedly apply Lemma 1.3.1 to embed the trees induced by the remaining levels of P into $(X'_{i,C}, Y'_{j,C})$ in a way that (A) and (D) hold.

Let us now prove (I). We first note that there exists some cluster X_i such that $|\phi^{-1}(X_{i,C})| < |X_{i,C}| - 5\sqrt{\varepsilon} m$. Indeed, otherwise we have used at least

$$(1-21\sqrt{\varepsilon})|X| - 5\sqrt{\varepsilon}|X| \ge (1-26\sqrt{\varepsilon})(1+100\sqrt{\varepsilon})k > (1+2\sqrt{\varepsilon})k > k+1$$

vertices from X already, a contradiction, since $|T| = k+1$.

Next, we claim there exists some cluster Y_j such that (X_i, Y_j) is good. If this was not the case, then we have used at least

$$(1-26\sqrt{\varepsilon})|N(X_i)| \ge (1-26\sqrt{\varepsilon})(1+100\sqrt{\varepsilon})\frac{k}{2} > (1+2\sqrt{\varepsilon})\frac{k}{2} > \frac{k+1}{2}$$

vertices of Y already, a contradiction, as we placed the root $r(T)$ of T in a way that guaranteed we would embed the smaller bipartition class of T into Y. □

[1] except if it is the root, see above

Remark 3.1.3 It is easy to see that instead of conditions (ii) and (iii) from Proposition 3.1.2 we could use the weaker requirement that there is a set $\mathcal{C}$ of clusters in X such that $\deg(x) \geq (1+100\sqrt{\varepsilon})\frac{k}{2}$, for all $x \in V(\mathcal{C})$, and $|V(\mathcal{C})| \geq (1+100\sqrt{\varepsilon})k$.

Remark 3.1.4 Another useful hypothesis can be added to Proposition 3.1.2, implying only small modifications to the proof. Consider an arbitrary set $U \subset V(G)$ and any tree T such that $|U|+|T| \leq k+1$, $|U \cap A|+c_0(T) \leq k$ and $|U \cap B|+c_1(T) \leq \frac{k}{2}$. If T has maximum degree bounded by $k^{\frac{1}{d+1}}$, then T can be embedded into G, avoiding U, i.e., $\phi(V(T)) \subset V(G) \setminus U$.

Moreover, observe that repeatedly applying Proposition 3.1.2 together with Remark 3.1.4, we can actually embed a forest instead of a tree. A (k_1, k_2, c)-*forest* $\mathcal{F}$ is a forest with at most k_1+k_2 vertices and maximum degree bounded by $(k_1+k_2-1)^{\frac{1}{c}}$ that admits a proper vertex coloring with $c_0(\mathcal{F}) \leq k_1$ and $c_1(\mathcal{F}) \leq k_2$.

Corollary 3.1.5 Let $\varepsilon \in (0, 10^{-4})$ and let $d \in \mathbb{N}$. There exists $k_0 = k_0(\varepsilon) \in \mathbb{N}$ such that for all $k_1, k_2 \geq k_0$ the following holds. If $G = (X, Y; E)$ is a connected bipartite graph with an $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition with clusters of size m, and corresponding reduced graph G_r, such that

(i) $\mathrm{diam}(G_r) \leq d$;

(ii) $\deg(x) \geq (1+100\sqrt{\varepsilon})k_2$, for all $x \in X$; and

(iii) $|X| \geq (1+100\sqrt{\varepsilon})k_1$,

then G contains any $(k_1, k_2, d+1)$-forest as a subgraph. Moreover, there is a set X' with $|X'| > (1-2\varepsilon)|X|$ such that the roots that must be mapped into X can be chosen from X', and similarly for the roots that must be mapped into Y.

Let us now quickly discuss the bound on the size of bipartition class X from Proposition 3.1.2. This bound is close to best possible, which can be seen by considering the tree T on an even number of levels, where every vertex in an even level (including the root) has degree $\Delta \geq 3$, and every vertex in an odd level, except for the leaves, has degree 2. Setting $k := |T|-1$, we can calculate that there are $\frac{(\Delta-1)}{\Delta}k+1$ vertices in odd levels. For any $0 < \delta \leq \frac{1}{3}$ consider the complete bipartite graph with bipartition classes of sizes $\frac{(\Delta-1)}{\Delta}k$ and $\lceil (1+\delta)\frac{k}{2} \rceil$. This graph has minimum degree $\lceil (1+\delta)\frac{k}{2} \rceil$, but does not contain T.

It is easy to see that the tree from the previous paragraph is the most unbalanced tree whose maximum degree is bounded by Δ. Therefore, one can show the following improvement of Proposition 3.1.2.

Corollary 3.1.6 Let $\varepsilon \in (0, 10^{-4})$, $d \in \mathbb{N}$ and $\Delta \geq 2$. There exists $k_0 = k_0(\varepsilon) \in \mathbb{N}$ such that for all $k \geq k_0$ the following holds. If $G = (A, B; E)$ is a connected bipartite graph with an $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition, and corresponding reduced graph G_r, such that

(i) $\mathrm{diam}(G_r) \leq d$;

(ii) $\deg(x) \geq (1 + 100\sqrt{\varepsilon})\frac{k}{2}$ for all $x \in X$;

(iii) $|X| \geq (1 + 100\sqrt{\varepsilon})\frac{(\Delta-1)}{\Delta}k$,

then G contains any tree T with k edges and $\Delta(T) \leq \Delta$ as a subgraph.

3.1.3 Non-bipartite case

When the graphs we are considering are non-bipartite, the presence of a cluster matching of a certain size will be enough to ensure that a given tree can be embedded. This case is very similar to Proposition 3.1.2, so we will omit some details. When we speak about a sequence of clusters $W_1W_2 \dots W_\ell$ from X to Y, this means that $W_1 = X$, $W_\ell = Y$ and, for each $i \in [\ell - 1]$, W_i is adjacent to W_{i+1}.

Proposition 3.1.7 Let $\varepsilon \in (0, 10^{-4})$ and let $d \in \mathbb{N}$. There exists $k_0 = k_0(\varepsilon) \in \mathbb{N}$ such that for all $k \geq k_0$ the following holds. Let G be a connected non-bipartite graph with an $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition and corresponding reduced graph G_r. If $\mathrm{diam}(G_r) \leq d$, and G_r has a matching M with $|V(M)| \geq (1 + 100\sqrt{\varepsilon})k$, then any tree T with k edges and $\Delta(T) \leq k^{\frac{1}{4d+2}}$ is a subgraph of G.

Proof. Let $V(G) = V_1 \cup \dots \cup V_\ell$ be the $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition of G. For each $i \in [\ell]$, we partition V_i into sets $V_{i,S}$, $V_{i,L}$, $V_{i,C}$ in the same way we did it in Proposition 3.1.2. Also, consider the decomposition of T into $\mathcal{T}$ and S given by Proposition 3.1.1 with $\beta = \frac{\varepsilon}{\ell}$. We order $S \cup \mathcal{P}$ in the same way as in the proof of Proposition 3.1.2.

The embedding $\phi : V(T) \to V(G)$ will be constructed iteratively, following the order of $S \cup \mathcal{P}$. We make sure that at every step, the following conditions will be satisfied:

(A) Each vertex is embedded into a neighbor of the image of its already embedded parent;

(B) each $s \in S$ is embedded into the S-slice of some cluster;

(C) for each $P \in \mathcal{P}$, the first (up to $4d + 1$) levels are embedded into the L-slices of some clusters, and the rest goes into the C-slices;

(D) every $v \in V(T)$ is mapped into a vertex that is typical towards both the S-slice and the L-slice of some adjacent cluster; and

(E) for each pair $(V_i, V_j) \in M$, $||\phi^{-1}(V_{i,C})| - |\phi^{-1}(V_{j,C})|| \leq \varepsilon m$.

We already know that it is no problem to embed a seed s, when its time comes, satisfying conditions (A), (B) and (D). So assume we are about to embed a piece $P \in \mathcal{P}$. The parent of the root $r(P)$ of P is already embedded into some vertex that is typical with respect to the L-slice of some cluster Z_0. In order to be able to embed P, it suffices to ensure that

(I) there exists some good pair (V_i, V_j);

(II) there is a sequence $Y = Y_0Y_1 \dots Y_h$ from Z_0 to V_i and there is a sequence $W = W_0W_1 \dots W_h$ from Z_0 to Y_j such that $h \leq 4d+1$;

(III) the union of the first h levels of P is small enough to fit into the free space in the L-slices of $\{Y_0, Y_1, \dots, Y_{h-1}\}$ or into the free space in the L-slices of $\{W_0, W_1, \dots, W_{h-1}\}$.

For the moment, assume (I) holds. Now, let $C = C_0C_1 \dots C_pC_0$ be a minimal odd cycle in the reduced graph. Since C is minimally odd, the shortest path between two clusters in C is the shortest arc in the cycle, hence the length of C is at most $2d+1$. Let $R = Z_0R_1 \dots R_sC_0$ be a shortest path going from Z_0 to C_0 and let $Q = C_0Q_1 \dots Q_tV_i$ be a shortest path going from C_0 to V_i. As $\operatorname{diam}(G_r) \leq d$, the lengths of R and Q are at most d. Thus, the sequence $Y = Z_0R_1 \dots R_sC_0C_1 \dots C_pC_0Q_1 \dots Q_tV_i$ has length at most $4d+1$, i.e., the number of clusters in Y is at most $4d+2$. Now, the sequence $W' = Z_0R_1 \dots R_sC_0Q_1 \dots Q_tV_iV_j$ has the same parity of Y and has at least two clusters less than Y, therefore, we can concatenate V_iV_j at the end of W' as many times as necessary to form a sequence $W = W'V_iV_j \dots V_iV_j$ with the same length as Y. As Y goes from Z_0 to V_i and W goes from Z_0 to V_j, condition (II) holds.

Using the same reasoning as in Proposition 3.1.2 we can prove that the total number of occupied vertices in L-slices is at most

$$|S| \cdot \Delta(T) \cdot 2(\Delta(T)-1)^{4d} \;\leq\; \frac{4}{\beta} \cdot k^{\frac{4d+1}{4d+2}} \;<\; \varepsilon m$$

for k sufficiently large. Which proves (III). In particular, the L-slice of each cluster has at least $\lceil 9\sqrt{\varepsilon}m \rceil$ unused vertices and, therefore, we can embed each vertex of the first h levels of P into the L-slices of the first h clusters of any the two sequences.

Now, in order to ensure that condition (E) holds at the end of this step, we need to be careful with the election of the sequence through which the first h levels of P will be embedded. We know that, regardless of the choice of the sequence, the set $\bigcup_{i \geq h} L_i(P)$ has to be mapped into $(V_{i,C}, V_{j,C})$, so all we have to determine is which of the sets $\bigcup_{i \geq \lceil \frac{h}{2} \rceil} L_{2i}(P)$ and $\bigcup_{i \geq \lfloor \frac{h}{2} \rfloor} L_{2i+1}(P)$ has more vertices and which of the slices $V_{i,C}$ and $V_{j,C}$ has less vertices assigned so far. Let us assume that $|\bigcup_{i \geq \lceil \frac{h}{2} \rceil} L_{2i}(P)| \geq |\bigcup_{i \geq \lfloor \frac{h}{2} \rfloor} L_{2i+1}(P)|$ and that $|\phi(V(T)) \cap V_{i,C}| \geq |\phi(V(T)) \cap V_{j,C}|$, then

$$\text{we need to embed } \bigcup_{i \geq \lceil \frac{h}{2} \rceil} L_{2i}(P) \text{ into } V_{j,C} \text{ and } \bigcup_{i \geq \lfloor \frac{h}{2} \rfloor} L_{2i+1}(P) \text{ into } V_{i,C}. \tag{3.1}$$

Suppose also that h is even. We can embed each vertex of the first h levels of P into the L-slices of the clusters W_0, W_1, $W_2, \dots, W_{h-1}$ in a way that (A) and (D) hold. Finally, we repeatedly apply Lemma 1.3.1 to embed the trees induced by the remaining levels of P into $(X'_{i,C}, Y'_{j,C})$ in a way that (A), (D) and (3.1) hold. This and the fact that $|P| \leq \beta k \leq \varepsilon m$ ensures that condition (E) holds. If h were odd, then we use Y instead of W.

Let us now prove (I). Suppose there is no good pair in M. This together with (E) imply that the number embedded vertices is at least

$$\sum_{(U,V) \in M} (|U_{i,C}| - 6\sqrt{\varepsilon}m + |V_{i,C}| - 6\sqrt{\varepsilon}m) \geq (1 - 33\sqrt{\varepsilon})(1 + 100\sqrt{\varepsilon})k > k+1,$$

a contradiction, since $|T| = k+1$. □

Remark 3.1.8 When $d = 1$ we can actually embed trees with maximum degree bounded by ck, where c is a sufficiently small constant, without modifying our proof.

Remark 3.1.9 In the non-bipartite case we can add an hypothesis similar to the one in Remark 3.1.4. Consider an arbitrary set $U \subset V(G)$ and any tree T such that $|U|+|T| \leq k+1$ and such that U is balanced in M, i.e., $||U \cap C| - |U \cap D|| < \varepsilon m$. If T has maximum degree bounded by $k^{\frac{1}{4d+2}}$, then T can be embedded into G, avoiding U.

Repeatedly applying Proposition 3.1.7 together with Remark 3.1.9 we can embed a forest instead of a tree.

Corollary 3.1.10 Let $\varepsilon \in (0, 10^{-4})$ and let $d \in \mathbb{N}$. There exists $k_0 = k_0(\varepsilon) \in \mathbb{N}$ such that for all $k \geq k_0$ the following holds. Let G be a connected non-bipartite graph with an $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition and corresponding reduced graph G_r. If $\operatorname{diam}(G_r) \leq d$ and G_r has a matching M with $|V(M)| \geq (1 + 100\sqrt{\varepsilon})k$, then any forest $\mathcal{F}$ on at most $k + 1$ vertices that satisfies $\Delta(\mathcal{F}) \leq k^{\frac{1}{4d+2}}$ is a subgraph of G. Moreover, there is a set V' with $|V'| > (1 - 2\varepsilon)|V(M)| \cdot m$ such that the roots of $\mathcal{F}$ can be chosen from V'.

The next lemma shows that a bound on the minimum degree can be used to prove the existence of a cluster matching of a certain size, which will allow us to use Proposition 3.1.7 and Corollary 3.1.10 in graphs with a minimum degree condition.

Lemma 3.1.11 Let $\varepsilon \in (0, 1)$. There exists $t_0 = t_0(\varepsilon) \in \mathbb{N}$ such that for all $t \geq t_0$ the following holds. Let H be a connected graph on $n \geq (1 + 100\sqrt{\varepsilon})2t$ vertices. If H admits an $(\varepsilon, 6\sqrt{\varepsilon})$-upper regular partition with M_0 clusters and has minimum degree at least $(1 + 100\sqrt{\varepsilon})t$, then there exist a set $V_0 \subset V(G)$ with $|V_0| \leq M_0$, a $(6\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition of the graph $G - V_0$ with at most $2M_0$ clusters, a cluster matching M in $G - V_0$ and an independent family of clusters $\mathcal{C}$ in $G - V_0$ such that,

- $V(M) \cap V(\mathcal{C}) = \varnothing$,
- $V(M) \cup V(\mathcal{C}) \supseteq V(G) \setminus V_0$,
- a pair in M is seen in at most one side by the clusters in $\mathcal{C}$, and
- $|V(M)| \geq \min\{n, (1 + 100\sqrt{\varepsilon})2t\} - M_0$.

Proof. Let us choose a pair (M_1, Γ), where M_1 is a cluster matching and Γ is a family of disjoint triangles, disjoint from M_1, such that $|V(M_1) \cup V(\Gamma)| = |V(M_1)| + |V(\Gamma)|$ is maximized. Let $\mathcal{C}_1$ be the family of clusters not covered by (M_1, Γ). Note that $\mathcal{C}_1$ may be empty, but we will momentarily suppose that it is not. Due to the maximality of (M_1, Γ), $\mathcal{C}_1$ must be an independent family. Let $C \in \mathcal{C}_1$. If C sees a vertex in some triangle $XYZ \in \Gamma$, say C is adjacent to X, then $(M_1 \cup \{(C, X), (Y, Z)\}, \Gamma \setminus \{XYZ\})$ contradicts the maximality of (M_1, Γ). Suppose that C sees both sides of a pair $(A, B) \in M_1$, meaning that (C, A) and (C, B) are ε-regular with positive density, then ABC is a triangle and $(M_1 \setminus \{(A, B)\}, \Gamma \cup \{ABC\})$ cover more vertices than (M_1, Γ). Finally, let D be another cluster in $\mathcal{C}_1$ and suppose that there is a pair $(A, B) \in M_1$ such that C sees one side of it and D sees the other side, namely (C, A) and (D, B) are ε-regular with positive density, then $((M_1 \setminus \{(A, B)\}) \cup \{(C, A), (D, B)\}, \Gamma)$ contradicts again the maximality of (M_1, Γ). Therefore, $\mathcal{C}_1$ is an independent family not seeing Γ and seeing at most one side of each pair in M_1.

If $\mathcal{C}_1 = \varnothing$, then $|V(M_1) \cup V(\Gamma)| = n$. If $\mathcal{C}_1 \neq \varnothing$, consider again the cluster $C \in \mathcal{C}_1$. As $\delta(H) \geq (1+\delta)t$ and C only sees one side of the matching M_1, and nothing else, we have that C sees at least $(1+\delta)t$ vertices in M_1 and therefore $|V(M_1)| \geq (1+\delta)2t$.

Now, if Γ is empty, we are done. Thus, suppose that there is at least one triangle. For each cluster X we arbitrarily choose two subsets $X^1, X^2 \subset X$ such that $|X^1| = |X^2| = \lfloor \frac{|X|}{2} \rfloor$. Define $V_0 := \bigcup_X X \setminus (X^1 \cup X^2)$ and observe that it contains either zero or M_0 vertices (recall that all clusters have the same size). Thanks to second point of Fact 1.2.2, partition $V(G) \setminus V_0 = \bigcup_X X^1 \cup X^2$ is $(6\varepsilon, 5\sqrt{\varepsilon})$-upper regular. Finally we define

$$M := \bigcup_{(A,B)\in M_1} \{(A^1, B^1), (A^2, B^2)\} \cup \bigcup_{XYZ\in\Gamma} \{(X^1, Y^1), (Y^2, Z^1), (Z^2, X^2)\}$$

and

$$\mathcal{C} := \bigcup_{C\in\mathcal{C}_1} \{C^1, C^2\}$$

As the original clusters are independent sets and $\mathcal{C}_1$ does not see Γ, M and $\mathcal{C}$ inherit the properties of $\mathcal{C}_1$ and M_1. □

As we have seen, the embedding propositions in this section require a bound on the diameter of the host graph. We state here a result from [14] that will be useful to obtain this type of bounds.

Theorem 3.1.12 (Erdős, Pach, Pollach and Tuza [14]) Let G be connected graph with n vertices and with minimum degree $\delta \geq 2$. Then $\mathrm{diam}(G) \leq \lfloor \frac{3n}{\delta+1} \rfloor - 1$.

3.1.4 Embedding constant degree trees

In this section we shall improve the upper bound on the degree of the trees. We will consider those trees whose maximum degree is bounded above by a function of the form $k^{\frac{1}{c}}$, where $c > 1$ is a constant we will specify later. The difference between this result and those from the previous sections is that the constant c does not depend on the diameter of the graph, and therefore, it does not depend on the constant of approximation δ.

We begin by presenting a lemma which can be seen as a variant of Theorem 3.1.12.

Lemma 3.1.13 Let $p, q \in \mathbb{N}$ and let G be a connected graph with $\delta(G) \geq p$. For every vertex $v \in V(G)$

$$\left| \bigcup_{i=0}^{3q+1} N_i(v) \right| \geq \min\{(q+1)(p+1), |V(G)|\}.$$

PROOF. Let $v \in V(G)$. Notice that if $N_i(v) = \varnothing$ for some $i = 1, \ldots, 3q+1$, then, as G is connected, $V(G) \subseteq \bigcup_{j=0}^{i-1} N_j(v)$; thus $|\bigcup_{i=0}^{3q+1} N_i(v)| = |V(G)|$. Therefore, we assume that

$N_i(v) \neq \varnothing$ for every $i = 1, \ldots, 3q+1$. Now pick vertices $v_{3j} \in N_{3j}(v)$ for $j = 1, \ldots, q$ and observe that $N(v_{3j}) \subseteq N_{3j-1}(v) \cup N_{3j}(v) \cup N_{3j+1}(v)$, for each j; this tells us that $|N_{3j-1}(v) \cup N_{3j}(v) \cup N_{3j+1}(v)| \geq p+1$. We also know that $|N_0(v) \cup N_1(v)| = |N(v)| + 1 \geq p+1$. Putting all this together yields

$$\left| \bigcup_{i=0}^{3q+1} N_i(v) \right| \geq (q+1)(p+1).$$

□

The next proposition is this section's main result. In order to prove it we will resort first to a strategy similar to the one used in Propositions 3.1.2 and 3.1.7; if this strategy fails, we will have found a good structure in the host graph and, thus, we will forget about the embedding made in the first manner and make use of this structure to do the embedding in a different way. When we speak about the distance between two clusters X and Y we are referring to the length of the shortest cluster-path that has X and Y as endpoints.

Proposition 3.1.14 Let $\alpha \in [\frac{1}{2}, 1)$ and let $\varepsilon \in (0, 10^{-4})$. There exists $k_0 = k_0(\delta) \in \mathbb{N}$ such that for all $k \geq k_0$ the following holds. Let G be a connected graph with $\delta(G) \geq (1+100\sqrt{\varepsilon})\alpha k$ which admits an $(\varepsilon, 5\sqrt{\varepsilon})$-upper regular partition, then G contains each tree T with k edges and maximum degree at most $k^{\frac{1}{c}}$ as a subgraph, where $c = 24\lceil \frac{2}{\alpha} \rceil - 6$.

Proof. Given $\alpha \in [\frac{1}{2}, 1)$ we define

$$d_1 := 3\lceil \frac{2}{\alpha} \rceil - 2, \ \ d_2 := 2(\mathrm{d}_1 + 1) \ \text{ and } \ c := 4d_2 + 2.$$

Note that if $|V(G)| < (1 + 100\sqrt{\varepsilon})2k$, then, by Theorem 3.1.12, $\mathrm{diam}(G) \leq \lfloor \frac{6}{\alpha} \rfloor - 1 \leq d_2$ and, therefore, we can apply either Proposition 3.1.2 or Proposition 3.1.7 to conclude. Thus, from now on we will assume that

$$|V(G)| \geq (1 + 100\sqrt{\varepsilon})2k. \tag{3.2}$$

Now, let T be a tree on k edges with $\Delta(T) \leq k^{\frac{1}{c}}$. Root T at any vertex. Partition T using Proposition 3.1.1 with $\beta \ll \varepsilon$, thus obtaining a set of special vertices S and a family of pieces $\mathcal{T}$. We first try to emulate the embedding scheme used in Proposition 3.1.2.

Partition each cluster X into three sets X_C, X_S, X_L where $|X_S| = |X_L| = \lceil 10\sqrt{\varepsilon}|X| \rceil$. We are going to do the embedding in $|S|$ steps. In each step j we consider a vertex $s_j \in S$ not embedded yet, but whose father u_j is already embedded, except for the step $j = 1$, in which we will embed the root of T into any cluster of our choice. We can assume that $\phi(u_j)$ is typical towards the S-slice of some adjacent cluster Q. Embed s_j in Q_S, choosing $\phi(s_j)$ typical to R_L and to R_S, where R is any neighbor of Q. Now, let (W, Z) be any good pair such that $\mathrm{dist}(Q, W) \leq \mathrm{d}_1$. Find a shortest cluster-path P from R to W and suppose that $P = X_0X_1 \ldots X_{l-1}X_l$, where $X_0 = R$ and $X_l = W$. Clearly, the length of P is at most $d_1 + 1$. Consider a piece T' hanging from s_j and let us embed the first l levels of T' into P, mapping the vertices from $L_i(T)$ into unoccupied vertices from $(X_i)_L$ that are typical towards $(X_{i+1})_L$ and towards $(X_{i+1})_S$, for each $i = 0, \ldots, l-1$. For the root of T' we take the additional consideration of mapping it into the neighborhood of $\phi(s_j)$ in $(X_0)_L$. Finish the embedding of T' into the unoccupied vertices of (W_C, Z_C) using Lemma 1.3.1, mapping the vertices in $L_l(T)$ into W_C and picking all vertices typical towards the L-slice and the S-slice of some adjacent cluster. Repeat this procedure for every non embedded piece adjacent to s_j and then move to the next special vertex.

If every step of the process is successful, T is satisfactorily embedded into G. It might also happen that the embedding can not be completed, because we are not able to find a good pair at close range. Consider the special vertex s^* where the process stopped and the cluster C^* to which s^* was assigned. Let us define $H := G[\bigcup_{i=0}^{d_1} V(N_i(C^*))]$ and $\mathcal{S} := \{C \text{ cluster in } H : |C \setminus C'| \geq 5\sqrt{\varepsilon}|C|\}$. We know that, as the assignment could not be finished, $\mathcal{S}$ is an independent set of clusters. Now, pick any vertex $x \in C^*$ and observe that $\bigcup_{i=0}^{d_1} N_i(x) \subset V(H)$; this implies, thanks to Lemma 3.1.13 applied with $q = \lceil \frac{2}{\alpha} \rceil - 1$ and $p = (1 + 100\sqrt{\varepsilon})\alpha k$, that $|V(H)| > (1 + 100\sqrt{\varepsilon})2k$. We also know that, as $|T| \leq k + 1$, $|V(\mathcal{S})| \geq (1 + 100\sqrt{\varepsilon})k$. Now, consider $H' := G[\bigcup_{i=0}^{d_1+1} V(N_i(C^*))]$; there are two possible cases.

Case 1: H' is non-bipartite

We can use $\mathcal{S}$ to construct a cluster-matching M in H' with $|V(M)| \geq (1 + 100\sqrt{\varepsilon})k$: consider a maximal cluster-matching such that for every pair $(C, D) \in M$, either $C \in \mathcal{S}$ or $D \in \mathcal{S}$ and suppose that $|V(M)| < (1 + 100\sqrt{\varepsilon})k$; since $|V(\mathcal{S})| \geq (1 + 100\sqrt{\varepsilon})k$, there exists $x \in V(\mathcal{S}) \setminus V(M)$; observe that $\deg_{H'}(v) \geq (1+100\sqrt{\varepsilon})\frac{k}{2}$ for each $v \in V(\mathcal{S})$; this implies that x has a neighbor $y \in V(\mathcal{S}) \setminus V(M)$, thus contradicting the independence of S. Observing that $\mathrm{d}iam(H') \leq 2(d_1 + 1) = d_2$ and applying Proposition 3.1.7 to H' we can conclude.

Case 2: H' is bipartite

Let us say $H' = (A, B)$. As $|V(H)| \geq (1+100\sqrt{\varepsilon})2k$, we have that, without loss of generality, $|A \cap H| \geq (1 + 100\sqrt{\varepsilon})k$. Finally observe that $\deg_{H'}(x) \geq (1 + 100\sqrt{\varepsilon})\frac{k}{2}$ for each $x \in V(H)$. Applying Proposition 3.1.2 with Remark 3.1.3 to H' we get the result. □

3.2 Main results

We are ready to establish our main results. The first part of this section is dedicated to understand some global structures of the host graph in which we are able to perform the embedding. In the second section we prove the main theorems of this chapter, while in the third section we study some alternative conditions on the host graph with which the maximum degree condition of 3.2.5 can be relaxed.

3.2.1 A general embedding lemma

We present here a lemma containing a series of configurations in which we are able to perform the embedding of a bounded degree tree. This result can be seen as a compendium of favorable scenarios. Before stating the lemma we need three simple but useful definitions.

Definition 3.2.1 Let $n \in \mathbb{N}$ and let $\theta \in (0,1)$. Given an n-vertex graph G and a vertex x, let $\mathcal{G}_x(\theta)$ denote the class of all graphs obtained by deleting at most θn^2 edges from G, none which may be incident with x, and by deleting at most θn vertices from G, none which may be x.

Definition 3.2.2 Let $n \in \mathbb{N}$ and let $\theta \in (0,1)$. A vertex x of an n-vertex graph H θ-sees a set $U \subseteq V(H)$ if it has at least $\theta|U|$ neighbors in U.

Definition 3.2.3 Let $k \in \mathbb{N}$ and let $\theta \in (0,1)$. A non-bipartite connected graph G is said to be (k,θ)-small if $|V(G)| < (1+\theta)k$. A bipartite graph $H = (A, B)$ is said to be (k,θ)-small if $\max\{|A|,|B|\} < (1+\theta)k$. When a connected graph is not (k,θ)-small, we will say that it is (k,θ)-large.

Lemma 3.2.4 Let $\alpha \in [\frac{1}{2}, 1)$ and let $\delta \in (0,1)$. There exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be an n-vertex graph of minimum degree at least $(1+\delta)\alpha k$, and let T be a k-edge tree with maximum degree bounded by $k^{\frac{1}{c}}$, where $n \geq k \geq \delta n$ and $c = 24\lceil\frac{2}{\alpha}\rceil - 6$. Then T embeds in G if there is an $x \in V(G)$ such that at least one of the following conditions holds for each $G' \in \mathcal{G}_x(\frac{\delta^3}{100})$:

(I) $G' - x$ has a $(k, \frac{\delta}{100})$-large non-bipartite component; or

(II) $G' - x$ has a $(|A(T)|, \frac{\delta}{100})$-large bipartite component; or

(III) $G' - x$ has a $(\frac{2k}{3}, \frac{\delta}{100})$-large bipartite component such that, in G', vertex x $\frac{\delta^2}{100}$-sees both sides of the bipartition; or

(IV) x $\frac{\delta^2}{100}$-sees two components C_1, C_2 of $G' - x$ in a way that one of the following holds:

  (a) x sends an edge to a third component C_3; or

  (b) there is $i \in \{1,2\}$ such that C_i is non-bipartite and $(\frac{2k}{3}, \frac{\delta}{100})$-large; or

  (c) there is $i \in \{1,2\}$ such that C_i is bipartite and x sees both sides of the bipartition; or

  (d) there is $i \in \{1,2\}$ such that C_i is bipartite with parts A and B, $\min\{|A|,|B|\} \geq (1+\frac{\delta}{100})\frac{2k}{3}$ and x sees only one side of the bipartition.

  (e) C_1 and C_2 are bipartite with parts A_1, B_1 and A_2, B_2 respectively, $\min\{|A_1|,|B_2|\} \geq (1+\frac{\delta}{100})\frac{2k}{3}$ and x sees only A_1 and A_2.

Proof. Given δ, we suitably choose ε and d such that

$$0 < \varepsilon \ll d \ll \delta. \tag{3.3}$$

Let N_0, M_0 be given by Proposition 1.2.4 for input ε and $m_0 := \frac{1}{\varepsilon}$. Let k_0 be the maximum of the outputs k_0 and t_0 of Proposition 3.1.14, Corollary 3.1.5, Corollary 3.1.10 and Lemma 3.1.11, and set $n_0 := \max\{N_0, \delta^{-1}k_0\}$.

Now given an n-vertex graph G of minimum degree at least $(1+\delta)\alpha k$, with $n \geq k \geq \delta n$ and $n \geq n_0$, choose a vertex $x \in V(G)$ as in Lemma 3.2.4. Applying Proposition 1.2.4 to $G-x$ yields a subgraph $G_d \subseteq G-x$ that admits an (ε, d)-upper regular partition, with at most M_0 clusters, and having minimum degree

$$\delta(G_d) \geq (1+\frac{\delta}{2})\alpha k. \tag{3.4}$$

Let T be a k-edge tree with maximum degree bounded by $k^{\frac{1}{c}}$. Our goal is to find an embedding of T in $G' := G_d + x + E_G(x, G_d) \in \mathcal{G}_x(\frac{\delta^3}{100})$. Note that G' must fulfill one of the conditions from Lemma 3.2.4.

Let $\mathcal{C}$ be the collection of connected components in G_d, that is, $\mathcal{C} := \mathcal{C}(G_d)$. If $\mathcal{C}$ contains a $(k, \frac{\delta}{100})$-large non-bipartite component or a $(|A(T)|, \frac{\delta}{100})$-large bipartite component, then we can, recalling (4.7) and (3.4), conclude by Proposition 3.1.14. So we can discard scenarios (I) and (II) from Lemma 3.2.4. Thus, by Theorem 3.1.12, for any component C of the reduced graph G_r, we have that

$$\operatorname{diam}(C) \leq \frac{3|V(C)|}{(1+\frac{\delta}{100})\alpha k} \leq \frac{6}{\alpha}. \tag{3.5}$$

This implies that $4\operatorname{diam}(C)+2 \leq c$ for each component C of G_r, which means that we could eventually apply Corollaries 3.1.5 and 3.1.10 to the elements of $\mathcal{C}$.

In order to embed T into G' under scenarios (III) and (IV), we use the results from Section **??**. Applying Proposition 2.0.3 to T we obtain a vertex $z_0 \in V(T)$ and a proper 2-coloring $c: V(T-z_0) \to \{0,1\}$ of $T-z_0$ with $|c_1| \leq |c_0|$ such that $|c_0| \leq \frac{2k}{3}$ and $|c_1| \leq \frac{k}{2}$. Also, let $z_1 \in V(T)$ be the vertex given by Lemma 1.3.3 and let $\mathcal{T}$ be the set of connected components of $T-z_1$. Then $\mathcal{T}$ is a family of at most $\Delta(T)$ rooted trees whose roots are neighbors of z_1 in T, and $|V(T')| \leq \lceil \frac{k}{2} \rceil$ for every $T' \in \mathcal{T}$. Applying Lemma 1.3.4 (i) to $\mathcal{T}$ we obtain three disjoint families of trees $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{F}_3$, where $\mathcal{F}_3$ could be empty, such that $|V(\bigcup \mathcal{F}_i)| \leq \lceil \frac{k}{2} \rceil$ for each $i = 1,2,3$ and $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 = \mathcal{T}$. For later use, let us record here that

$$|\mathcal{F}_1| + |\mathcal{F}_2| + |\mathcal{F}_3| \leq \Delta(T). \tag{3.6}$$

Similarly, applying Lemma 1.3.4 (ii) to $\mathcal{T}$ we obtain two disjoint families of trees $\mathcal{J}_1$ and $\mathcal{J}_2$ such that $|V(\bigcup \mathcal{J}_2)| \leq |V(\bigcup \mathcal{J}_1)| \leq \frac{2k}{3}$ and $\mathcal{J}_1 \cup \mathcal{J}_2 = \mathcal{T}$.

We split the remainder of the proof into six cases, according to which of the conditions (III), (IVa), (IVb), (IVc), (IVd) or (IVe) holds. Depending on the case we will make use of partition $\{\mathcal{F}_i\}$ or $\{\mathcal{J}_i\}$ or we will just use the cutvertex z_0 and the forest $T-z_0$.

Case 1 (scenario (III)**):** $G'-x$ has a $(\frac{2k}{3}, \frac{\delta}{100})$-large bipartite component C such that x $\frac{\delta^2}{100}$-sees both sides of the bipartition.

We map z_0 into x. Recalling (3.4), (3.5) and the fact that $T-z_0$ is a $(\frac{2k}{3}, \frac{k}{2}, c)$-forest we can apply Corollary 3.1.10 to embed $T-z_0$ into C, picking the images of the roots of $T-z_0$ as neighbors of x. (Note that, as x $\frac{\delta^2}{100}$-sees C_1, the neighborhood of x contains enough typical vertices.)

Case 2 (scenario (IVa)**):** The vertex x $\frac{\delta^2}{100}$-sees two components C_1, C_2 of $G'-x$ and sends an edge to a third component C_3.

We embed z_1 into x, and then proceed to embed the roots of the trees from $\mathcal{F}_i$ into $N_{G_d}(x, C_i)$, for each $i = 1, 2, 3$. (This is possible because, due to Remark 1.3.5, the set $\mathcal{F}_3$ contains at most one tree and thus, there is at most one root to embed into C_3. Furthermore, by (3.6), there are at most $\Delta(T) \leq \frac{\delta}{100}|C_i|$ roots to be embedded into C_i, for $i = 1, 2$.)

Finally, because of the minimum degree in G_d, we can greedily embed the remaining vertices of each forest $\bigcup \mathcal{F}_i$ into component C_i. This finishes Case 2.

Case 3 (scenario (IVb)**):** The vertex x $\frac{\delta^2}{100}$-sees two components C_1, C_2 of $G' - x$, one of these components is non-bipartite and $(\frac{2k}{3}, \frac{\delta}{100})$-large.

Let us assume that C_1 is the $(\frac{2k}{3}, \frac{\delta}{100})$-large non-bipartite component. We map z_1 into x, and then embed $\bigcup \mathcal{J}_2$ greedily into C_2. We can make use of Corollary 3.1.10, whose conditions hold by (3.4), (3.5) and the fact that $|V(\bigcup \mathcal{J}_1)| \leq \frac{2k}{3}$, to complete the embedding of $\bigcup \mathcal{J}_1$ into C_1.

Case 4 (scenario (IVc)**):** The vertex x $\frac{\delta^2}{100}$-sees two components C_1, C_2 of $G' - x$, one of these components is bipartite and x sees both sides of the bipartition.

Let us assume that C_1 is the bipartite component of which x sees both sides, namely A and B. Note that, as x $\frac{\delta^2}{100}$-sees C_1, we can assume that x $\frac{\delta^2}{100}$-sees A. We map z_1 into x and then embed $\bigcup \mathcal{F}_1$ greedily into C_2. For the remaining forests observe that for any proper 2-coloring of $\bigcup \mathcal{F}_2$ and $\bigcup \mathcal{F}_3$, the larger color class of $\bigcup \mathcal{F}_i$ and the smaller color class of $\bigcup \mathcal{F}_{5-i}$, for $i = 2, 3$, add up to at most

$$|\bigcup \mathcal{F}_i| + \frac{|\bigcup \mathcal{F}_{5-i}|}{2} \leq \frac{|\bigcup \mathcal{F}_1| + |\bigcup \mathcal{F}_2)| + |\bigcup \mathcal{F}_3|}{2} = \frac{k}{2}. \tag{3.7}$$

As x has at least one neighbor $v \in B$, we can map the root of $\bigcup \mathcal{F}_3$ into v and then greedily embed the rest of $\bigcup \mathcal{F}_3$ into C_1. Now, we can make use of Corollary 3.1.5 together with Remark 3.1.4, whose condition holds by (3.7), to complete the embedding of $\bigcup \mathcal{F}_2$ into C_1, avoiding $\phi(V(\bigcup \mathcal{F}_3))$. This finishes Case 4.

Case 5 (scenario (IVd)**):** The vertex x $\frac{\delta^2}{100}$-sees two components C_1, C_2 of $G' - x$, one of them is bipartite with parts A and B, $\min\{|A|, |B|\} \geq (1 + \frac{\delta}{100})\frac{2k}{3}$ and x sees only one side of the bipartition.

Let us assume that C_1 is the bipartite component with parts A and B containing at least $(1 + \frac{\delta}{100})\frac{2k}{3}$ vertices each and that x only sees the set A. Note that x actually $\frac{\delta^2}{100}$-sees A. We map z_1 into x and then embed $\bigcup \mathcal{J}_2$ greedily into C_2. Applying Corollary 3.1.10 we can embed $\bigcup \mathcal{J}_1$ into C_1 in a way that the images of its roots are neighbors of x. Note that this time the roots of $\bigcup \mathcal{J}_2$ are forced to be mapped into A, but this is not a problem since both sides of C_1 have enough space for the whole forest to fit in.

Case 6 (scenario (IVe)**):** The vertex x $\frac{\delta^2}{100}$-sees two components C_1, C_2 of $G' - x$, both of them are bipartite with parts A_1, B_1 and A_2, B_2 respectively, $\min\{|A_1|, |B_2|\} \geq (1 + \frac{\delta}{100})\frac{2k}{3}$ and x sees only A_1 and A_2.

Note that x $\frac{\delta^2}{100}$-sees A_1 and A_2. Consider the coloring c that T induces in $\bigcup \mathcal{J}_1$. If the roots of the trees in $\mathcal{J}_1$ are contained in c_0, the heavier color class of c, then we embed $\bigcup \mathcal{J}_1$ into C_1, otherwise we embed $\bigcup \mathcal{J}_1$ into C_2. In any case we make use of Corollary 3.1.10 and we take care of mapping the roots into neighbors of x. We greedily embed $\bigcup \mathcal{J}_1$ into the remaining component. Finally, we map z_1 into x. This completes the proof of Lemma 3.2.4. □

3.2.2 Minimum degree theorems

The results of this section will be a consequence of Lemma 3.2.4. Using the minimum degree conditions and the high degree vertex we will be able to show that at least one of the configurations from Lemma 3.2.4 appears in our host graph. We begin with Theorem 3.2.5.

Theorem 3.2.5 Let $\delta \in (0,1)$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$ and maximum degree at least $(1+\delta)2k$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{90}}$, then T is a subgraph of G.

Proof. Given $\delta \in (0,1)$ and $\alpha = \frac{1}{2}$ consider n_0 from Lemma 3.2.4. Let G be an n-vertex graph with minimum degree at least $(1+\delta)\frac{k}{2}$ and maximum degree at least $(1+\delta)2k$, where $n \geq k \geq \delta n$ and $n \geq n_0$. Let $x \in V(G)$ be a vertex of degree at least $(1+\delta)2k$. We are going to show that for any removal of at most $\frac{\delta^3}{100}n^2$ edges not incident with x and of at most $\frac{\delta^3}{100}n$ vertices different from x, the resulting graph G' fulfills at least one of conditions (I)- (IV) from Lemma 3.2.4.

Let $\mathcal{C}$ be the collection of connected components in $G' - x$, that is, $\mathcal{C} := \mathcal{C}(G' - x)$. We will assume that

$$\text{all components in } \mathcal{C} \text{ are } (k, \frac{\delta}{100})\text{-small,} \tag{3.8}$$

otherwise we are done. First note that, as G' misses at most $\frac{\delta^3}{100}$ vertices from G,

$$\deg_{G'}(x) \geq (1+\frac{\delta}{2})2k. \tag{3.9}$$

Suppose now that x does not $\frac{\delta^2}{100}$-see any component in $\mathcal{C}$. This would mean that

$$2\delta n \leq (1+\frac{\delta}{2})2k \leq \deg_{G'}(x) = \sum_{C \in \mathcal{C}} \deg_{G'}(x, C) \leq \frac{\delta^2}{100}n.$$

Therefore, there is a component $C_1 \in \mathcal{C}$ receiving more than $\frac{\delta^2}{100}|C_1|$ edges from x. By (3.8), vertex x can have at most $2(1+\frac{\delta}{100})k$ neighbors in C_1. So by (3.9), there are at least $(\frac{\delta}{2} - \frac{\delta}{100})k > \frac{\delta}{4}k \geq \frac{\delta^2}{4}n$ neighbors of x outside C_1. Following the same reasoning as before, there must be a second component C_2 receiving at least $\frac{\delta^2}{100}|C_2|$ edges from x. We can assume that x has not neighbors outside $C_1 \cup C_2$, otherwise condition (IVa) from Lemma 3.2.4 holds. By (3.9) we can assume that

$$\deg_{G'}(x, C_1) \geq (1+\frac{\delta}{2})k.$$

In particular, we can again employ (3.8) to see that C_1 is bipartite, and, letting A and B denote the color classes of C_1,

$$\min\{\deg_{G'}(x,A), \deg_{G'}(x,B)\} \;\geq\; \frac{\delta}{4}k \;>\; \frac{\delta^2}{100}\cdot\max\{|A|,|B|\}.$$

Therefore, condition (IVc) from Lemma 3.2.4 holds and the proof is finished. $\square$

The following example shows that in general, the bound $(1+\delta)2k$ on the maximum degree cannot be replaced by a bound of the form ck with $c<2$.

Example 3.2.6 Let $\ell \in \mathbb{N}$ be odd and let $k \in \mathbb{N}$ be divisible by l. Let $H_{k,\ell}$ be the graph consisting of two copies of $K_{(\ell-1)(\frac{k}{\ell}-1),\frac{k}{2}}$, which we call H_1 and H_2, and a vertex x that is adjacent to all the vertices in the parts of size $(\ell-1)(\frac{k}{\ell}-1)$.
Consider the tree $T_{k,\ell}$ formed by ℓ stars of size $\frac{k}{\ell}$ and an additional vertex v connected to the centers of the stars. Notice that we cannot embed $T_{k,\ell}$ in $H_{k,\ell}$ by mapping v into x, since there would be too many stars assigned to one of the H_i.
Therefore, if we want to embed $T_{k,\ell}$ into $H_{k,\ell}$, the vertex v must be mapped into one of the bipartite graphs, say into H_1. But then, we have to embed at least $\ell-1$ stars into H_1, with the leaves of these stars going to the same side as v, which is impossible. Thus $T_{k,\ell} \not\subseteq H_{k,\ell}$.

When the considered trees have its maximum degree bounded by a constant, the maximum degree of the host graph can be lowered by a linear factor. Theorem 3.2.7 illustrates this possibility.

Theorem 3.2.7 Let $\delta \in (0,1)$ and $\Delta \geq 2$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$ and maximum degree at least $(1+\delta)2\frac{(\Delta-1)}{\Delta}k$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most Δ, then T is a subgraph of G.

Proof. Given $\delta \in (0,1)$ and $\alpha = \frac{1}{2}$ consider n_0 from Lemma 3.2.4. Let G be an n-vertex graph as in Theorem 3.2.7, where $n \geq n_0$. Let $x \in V(G)$ be a vertex of degree at least $2(1+\delta)\frac{(\Delta-1)}{\Delta}k$ and consider the graph G' resulting of the arbitrary removal of at most $\frac{\delta^3}{100}n^2$ edges not incident with x and of at most $\frac{\delta^3}{100}n$ vertices different from x. Let $\mathcal{C}$ be the set of components in $G'-x$.

An important thing to take into account is that, given $\Delta \geq 2$, each tree T on k edges with maximum degree at most Δ will satisfy

$$|A(T)| \leq \frac{(\Delta-1)}{\Delta}k. \tag{3.10}$$

We can discard scenarios (I) and (II) and therefore assume that

$$\text{all non-bipartite components in } \mathcal{C} \text{ are } (k, \frac{\delta}{100})\text{-small,} \tag{3.11}$$

and, by (3.10),

$$\text{all bipartite components in } \mathcal{C} \text{ are } (\frac{(\Delta-1)}{\Delta}k, \frac{\delta}{100})\text{-small.} \tag{3.12}$$

As we removed few vertices from G, vertex x has at least $2(1+\frac{\delta}{2})\frac{(\Delta-1)}{\Delta}k$ neighbors in G'. This together with (3.11) and (3.12) imply that there are components $C_1, C_2 \in \mathcal{C}$ such that

$$\deg_{G'}(x, C_i) \geq \frac{\delta^2}{100}|C_i|, \text{ for } i = 1, 2.$$

We will suppose that x does not see other components, otherwise G' satisfies condition (IVa) from Lemma 3.2.4 and we are done. Thus, we can assume, by symmetry, that

$$\deg_{G'}(x, C_1) \geq (1+\frac{\delta}{2})\frac{(\Delta-1)}{\Delta}k. \tag{3.13}$$

If C_1 is non-bipartite, G' satisfies condition (IVb) from Lemma 3.2.4. If C_1 is bipartite with parts A and B, we can employ again (3.12) together with (3.13) to conclude that $\min\{\deg_{G'}(x, A), \deg_{G'}(x, B)\} \geq \frac{\delta^2}{100}\max\{|A|, |B|\}$, which says that G' satisfies condition (IVc) from Lemma 3.2.4 and concludes the proof. $\square$

An analogue version of Theorem 3.2.5 can be proved for graphs with minimum degree above $\frac{2k}{3}$, only this time a vertex of degree greater than k will do. Theorem 3.2.8 certainly helps to support the $\frac{2}{3}$ conjecture.

Theorem 3.2.8 Let $\delta \in (0, 1)$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{2k}{3}$ and maximum degree at least $(1+\delta)k$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{66}}$, then T is a subgraph of G.

Proof. Given $\delta \in (0, 1)$ and $\alpha = \frac{2}{3}$ consider n_0 from Lemma 3.2.4. Let G be an n-vertex graph as in Theorem 3.2.8, where $n \geq n_0$. Let $x \in V(G)$ be a vertex of degree at least $(1+\delta)k$ and consider the graph G' resulting of the arbitrary removal of at most $\frac{\delta^3}{100}n^2$ edges not incident with x and of at most $\frac{\delta^3}{100}n$ vertices different from x. Note that $\deg_{G'}(x) \geq (1+\frac{\delta}{2})k$. Let $\mathcal{C}$ be the set of components in $G' - x$.

Due to the degree of x, there must be a component $C_1 \in \mathcal{C}$ such that x $\frac{\delta^2}{100}$-sees C_1. We can assume that C_1 is $(k, \frac{\delta}{100})$-small. Thus, if x has more than $(1+\frac{\delta}{100})k$ neighbors in C_1, C_1 must be bipartite and x must see at least a $\frac{\delta^2}{100}$-portion of both sides of the bipartition, namely A and B. Note that there is a vertex in C_1 having degree at least $(1+\frac{\delta}{100})\frac{2k}{3}$, otherwise, due to the minimum degree in G, the number of removed edges is at least

$$\frac{1}{2}\left(|C_1| \cdot \frac{\delta}{2} \cdot \frac{2k}{3}\right) \geq \frac{\delta^3}{6}n^2,$$

which is a contradiction. Therefore, $\max\{|A|, |B|\} \geq (1+\frac{\delta}{100})\frac{2k}{3}$ and G' satisfies condition (III) from Lemma 3.2.4. If x has less than $(1+\frac{\delta}{100})k$ neighbors in C_1, then there must be a second component $C_2 \in \mathcal{C}$ in which x has at least $\frac{\delta^2}{100}|C_2|$ neighbors. We can assume that x does not send edges to any other component. Also, we can suppose that $\deg_{G'}(x, C_1) \geq (1+\frac{\delta}{2})\frac{k}{2}$. Following the same reasoning as before we can conclude that there is a vertex in C_1 having at least $(1+\frac{\delta}{100})\frac{2k}{3}$ neighbors and, therefore, $|C_1| \geq (1+\frac{\delta}{100})\frac{2k}{3}$. If C_1 is non-bipartite, then G' satisfies condition (IVb) from Lemma 3.2.4. So we suppose C_1

is bipartite and, furthermore, we assume vertex x does not see both sides of the bipartition, otherwise condition (IVc) from Lemma 3.2.4 holds. Let us call A and B each side of C_1 and let us assume, by symmetry, that $\deg_{G'}(x, A) \geq (1 + \frac{\delta}{4})\frac{k}{2}$. Again we can find a vertex in A having more than $(1 + \frac{\delta}{100})\frac{2k}{3}$ neighbors in B, which in turn implies that there is a vertex in B having more than $(1+\frac{\delta}{100})\frac{2k}{3}$ neighbors in A. This means that G' satisfies condition (IVd) from Lemma 3.2.4, which competes the proof. □

Note that the bound $\frac{2k}{3}$ on the minimum degree of Theorem 3.2.8 is best possible. Consider a graph G formed by two copies of $K_{\frac{2k}{3}-1}$ and a universal vertex seeing both cliques. It is easy to see that the tree $T_{k,3}$ is not a subgraph of G.

3.2.3 Maximum degree $\frac{4k}{3}$

As we saw in Example 3.2.6, there is no way of lowering the bound on the maximum degree of the host graph when no further assumptions are made. Nevertheless, graphs with the structure of $H_{k,\ell}$ might be the only obstructions to embed a tree on k edges into a graph of maximum degree greater than $(1+\delta)\frac{4k}{3}$ and minimum degree greater than $(1+\delta)\frac{k}{2}$. We now show that for bounded degree trees, and dense host graphs, this is the case.

A graph G will be called (k, δ)-good if $\delta(G) \geq (1 + \delta)\frac{k}{2}$ and $\Delta(G) \geq (1 + \delta)\frac{4k}{3}$. Set

$$L(G) := \left\{ x \in V(G) : \deg(x) \geq (1 + \frac{\delta}{100})\frac{4k}{3} \right\}.$$

For $x \in L(G)$ we say that the pair (G, x) satisfies **property** $(\star)$ if the following three conditions hold.

(a) Every component of $G - x$ is $(k, \frac{\delta}{100})$-small;

(b) x sees exactly two of the components, C_1 and C_2;

(c) For $i = 1, 2$, if C_i is non-bipartite, then C_i is $(\frac{2k}{3}, \frac{\delta}{100})$-small, and if C_i is bipartite, then x sees at most one side of the bipartition.

We define $\mathcal{N}_{k,\delta}$ as the class of all (k, δ)-good graphs G such that for all $x \in L(G)$, there is $G' \in \mathcal{G}_x(\frac{\delta^3}{100})$ such that the pair (G', x) satisfies property $(\star)$.

Theorem 3.2.9 Let $\delta \in (0, 1)$. There exists $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$, maximum degree at least $(1 + \delta)\frac{4k}{3}$ and does not belong to $\mathcal{N}_{k,\delta}$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{90}}$, then T is a subgraph of G.

Proof. Given $\delta \in (0, 1)$ and $\alpha = \frac{1}{2}$ consider n_0 from Lemma 3.2.4. Let G be an n-vertex graph as in Theorem 3.2.9, where $n \geq n_0$. Let $x \in L(G)$ be one of the vertices that make G not belong to $\mathcal{N}_{k,\delta}$ and consider the graph G' resulting of the arbitrary removal of at most

$\frac{\delta^3}{100}n^2$ edges not incident with x and of at most $\frac{\delta^3}{100}n$ vertices different from x. Note that $\deg_{G'}(x) \geq (1+\frac{\delta}{2})k$. Let $\mathcal{C}$ be the set of components in $G'-x$.

We know the pair (G',x) does not satisfy property $(\star)$, but if there is a $(k, \frac{delta}{100})$-large component in $\mathcal{C}$, we are done, se we can suppose condition (a) holds.

If condition (b) does not hold for G' and x, then either x sees only one component C_1 or it sees at least three components. The first case leads to scenario (III) from Lemma 3.2.4, as x has at least $(1+\frac{\delta}{2})\frac{4k}{3}$ neighbors in C_1 and C_1 is $(k, \frac{\delta}{100})$-small. For the second case: if x has more than $(1+\frac{\delta}{4})k$ neighbors in one component, then we are essentially in the first case, so we can conclude, as in previous proofs, that x $\frac{\delta^2}{100}$-sees at least two components, which together with the fact that x sends an edge to a third component leads to scenario (IVa) from Lemma 3.2.4.

Finally we assume condition (b) holds and, therefore, condition (c) does not. We can assume x $\frac{\delta^2}{100}$-sees both C_1 and C_1, otherwise we get (III) from Lemma 3.2.4. Now, either one of the components is non-bipartite and $(\frac{2k}{3}, \frac{\delta}{100})$-large or one of the components is bipartite and x sees both sides of the bipartition, which are scenarios (IVb) and (IVc) from Lemma 3.2.4, respectively. This concludes the proof. $\square$

A different way of avoiding Example 3.2.6 is to impose some conditions over the size of the second neighborhood of a vertex already belonging to $L(G)$. We explore this approach in the next theorem.

Theorem 3.2.10 Let $\delta \in (0,1)$. There exist $n_0 = n_0(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph with n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$, with $n \geq k \geq \delta n$. Suppose there is a vertex $x \in V(G)$ with $|N(x)| \geq (1+\delta)\frac{4k}{3}$ and $|N_2(x)| \geq (1+\delta)\frac{4k}{3}$. If T is a tree with k edges and maximum degree at most $k^{\frac{1}{90}}$, then T is a subgraph of G.

Proof. Given $\delta \in (0,1)$ and $\alpha = \frac{1}{2}$ consider n_0 from Lemma 3.2.4. Let G be an n-vertex graph as in Theorem 3.2.10, where $n \geq n_0$. Let $x \in V(G)$ be a vertex with $|N(x)|, |N_2(x)| \geq (1+\delta)\frac{4k}{3}$ and consider the graph G' resulting of the arbitrary removal of at most $\frac{\delta^3}{100}n^2$ edges not incident with x and of at most $\frac{\delta^3}{100}n$ vertices different from x. Note that the first and the second neighborhood of x in G' contain at least $(1+\frac{\delta}{2})\frac{4k}{3}$ vertices each. Let $\mathcal{C}$ be the set of components in $G'-x$.

We assume that every component in $\mathcal{C}$ is $(k, \frac{\delta}{100})$-small and that

$$G' \text{ does not satisfy condition (III) from Lemma 3.2.4,} \tag{3.14}$$

otherwise we are done. This implies that x $\frac{\delta^2}{100}$-sees at least two components $C_1, C_2 \in \mathcal{C}$. If x sends and edge to a third component, the result follows, so we can assume that x only sees C_1 and C_2 and, therefore,

$$\text{the second neighborhood of } x \text{ in } G' \text{ is entirely contained in } C_1 \cup C_2. \tag{3.15}$$

Also, by symmetry, we can suppose that $\deg_{G'}(x, C_1) \geq (1 + \frac{\delta}{2})\frac{2k}{3}$. We will additionally assume that

$$G' \text{ does not meet conditions (IVb) or (IVc) from Lemma 3.2.4.} \tag{3.16}$$

This means that C_1 is bipartite with parts A_1, B_1 and that x sees only one side of C_1, namely A_1. This means that $|A_1| \geq (1 + \frac{\delta}{100})\frac{2k}{3}$. Now, if $|B_1| \geq (1 + \frac{\delta}{100})\frac{2k}{3}$, we are in scenario (IVd) from Lemma 3.2.4, so we can assume, by (3.15), that most of the second neighborhood of x is contained in C_2, which by the first part of (3.16) implies that C_2 is bipartite with parts A_2 and B_2. Moreover, by (3.14) x sees only one side of C_2, namely A_2. This implies that the second neighborhood of x in C_2 is entirely contained in B_2, which means $|B_2| \geq (1 + \frac{\delta}{100})\frac{2k}{3}$. Thus, G' satisfies condition (IVe) from Lemma 3.2.4, which concludes our proof. $\square$

Chapter 4

Embedding of trees with linear degree

This chapter is devoted to a single result, Theorem 4.0.6. This theorem is the result of exploring an alternative condition to the high degree vertex of Theorem 3.2.5. Instead of requiring a vertex of degree greater than $(1+\delta)2k$, we ask for a few vertices of degree at least $(1+\delta)k$. This new condition will allow us to enlarge the family of embeddable trees, by replacing the bound on the maximum of degree of the trees by a linear function.

4.0.1 Cutting the trees into small pieces

The following proposition is an alternative version of Proposition 3.1.1, where all the pieces have an odd number of levels and all the seeds are at even distance from the root.

Proposition 4.0.1 Let $\beta \in (0,1)$. There exist $t_0 = t_0(\beta) \in \mathbb{N}$ and $c = c(\beta) \in \mathbb{N}$ such that for all $t \geq t_0$ the following holds. Let T be a rooted tree on $t+1$ vertices and maximum degree at most ct. There exist a set of *seeds* $S \subset V(T)$ and a family $\mathcal{T}$ of disjoint rooted trees which we call *pieces* that satisfy

(i) $r(T) \in S$,

(ii) $\mathcal{T}$ consists of the components of $T - S$,

(iii) each $s \in S$ is at even distance from $r(T)$,

(iv) each piece in $\mathcal{T}$ has at most βt vertices, and

(v) $|S| < \beta t$.

Proof. Given $\beta \in (0,1)$, set $c < \frac{\beta^2}{2}$ and $t_0 > \frac{6}{\beta^2}$. Let T be a rooted tree on $t+1$ vertices with $\Delta(T) \leq ct$, where $t \geq t_0$. Applying proposition 3.1.1 to T we obtain a set of seeds S' and a family of pieces $\mathcal{P}$. Let us call S_A the set of vertices in S' that are at odd distance from the root and S_B those which are at even distance from the root. Consider the set $\overline{S_B}$ of vertices in $T - S'$ that are adjacent to a vertex in S_A and set

$$S := S_B \cup \overline{S_B}.$$

Notice that $\overline{S_B}$ can be decomposed into two perhaps intersecting sets,

$$X := \{x \in T - S' : x \text{ is father of some } a \in S_A\}$$

and

$$Y := \{x \in T - S' : \text{some } a \in S_A \text{ is father of } x\}.$$

As a vertex in T can have only one father, we have that $|X| \leq |S_A| < \frac{1}{\beta} + 2$. Besides, $|Y| < |S_A|\Delta(T) < \frac{\beta t}{2}$. Therefore, we get

$$|\overline{S_B}| < \frac{1}{\beta} + 2 + \frac{\beta t}{2},$$

which in turn implies $|S| < \beta t$. Finally we can define our family of pieces as $\mathcal{T} := \mathcal{C}(T - S)$. As each $P \in \mathcal{T}$ is subgraph of some $P' \in \mathcal{P}$, point (4) remains true, concluding with this the decomposition of our tree. □

In order to embed T, we will need to ensure some balance condition over its pieces. The next numeric lemma will later allow us to find a subfamily of pieces having a given size and which maintains the global imbalance of T.

Lemma 4.0.2 Let I be a finite set and let $\lambda > 0$. Let $\{a_i\}_{i \in I}$ and $\{b_i\}_{i \in I}$ be two sequences of positive real numbers such that $a_i + b_i \leq \lambda$, for each $i \in I$. We call $\Sigma_a := \sum_{i \in I} a_i$ and $\Sigma_b := \sum_{i \in I} b_i$. For any $M \geq 0$ there is a set $J \subseteq I$ such that

$$\min\{M - \lambda, \Sigma_a + \Sigma_b\} \leq \sum_{i \in J}(a_i + b_i) \leq M \quad \text{and} \quad \frac{\sum_{i \in J} a_i}{\sum_{i \in J} b_i} \geq \frac{\Sigma_a}{\Sigma_b}.$$

Proof. Let us define a total order $\preceq$ in I such that $i \preceq j$ implies that $\frac{b_i}{a_i} \leq \frac{b_j}{a_j}$. From the definition of this order we get, for each $l \in I$,

$$\frac{b_{l+1}}{a_{l+1}} \geq \frac{\sum_{i \preceq l} b_i}{\sum_{i \preceq l} a_i}.$$

Multiplying both sides by $a_{l+1} \sum_{i \preceq l} a_i$, adding $\sum_{i \preceq l} a_i \sum_{i \preceq l} b_i$ and dividing by $\sum_{i \preceq l} b_i \sum_{i \preceq l+1} b_i$ we obtain

$$\frac{\sum_{i \preceq l} a_i}{\sum_{i \preceq l} b_i} \geq \frac{\sum_{i \preceq l+1} a_i}{\sum_{i \preceq l+1} b_i}$$

and thus

$$\frac{\sum_{i \preceq l} a_i}{\sum_{i \preceq l} b_i} \geq \frac{\sum_{i \in I} a_i}{\sum_{i \in I} b_i} = \frac{\Sigma_a}{\Sigma_a}.$$

Let i^* be the maximal element in I such that $\sum_{i \preceq i^*}(a_i + b_i) \leq M$. Defining $J := \{i \in I : i \preceq i^*\}$ we get the result. □

Remark 4.0.3 Note that consequently,

$$\Sigma_a + \Sigma_b - M \leq \sum_{i \in I \setminus J} (a_i + b_i) \leq \max\{\Sigma_a + \Sigma_b - M + \lambda, 0\}.$$

4.0.2 A structural lemma

We know from Lemma 3.1.11 that the minimum degree condition ensures the existence of a matching M and a set of clusters $\mathcal{C}$ seeing only one side of M. We will use M and $\mathcal{C}$ together with the minimum degree condition to derive a specific structure, which will be of special help to perform the embedding.

Lemma 4.0.4 Let $\varepsilon \in (0, 10^{-4})$. There exists k_0 such that for all $k \geq k_0$ the following holds. Let G be a connected graph with n vertices, where $n \geq k \geq 110\sqrt[4]{\varepsilon}n$, which satisfies $\delta(G) \geq (1+110\sqrt[4]{\varepsilon})\frac{k}{2}$ and $|\{v \in V(G) : \deg(v) \geq (1+110\sqrt[4]{\varepsilon})k\}| > 2\varepsilon n$. Suppose that G admits an $(\frac{\varepsilon}{6}, 2\sqrt[4]{\varepsilon})$-upper regular partition with $M_0' > \frac{1}{100\sqrt{\varepsilon}}$ clusters of size $m' := \frac{n}{M_0}$. There exist a set $V_0 \subset V(G)$ with $|V_0| \leq M_0'$, an $(\varepsilon, \sqrt{\varepsilon})$-upper regular partition of the graph $G - V_0$ with at most $M_0 = 2M_0'$ clusters of size $m = \lfloor \frac{m'}{2} \rfloor$, a cluster C_S, two cluster-matchings M_W, M_V and a bipartite graph $H = (X, Y; E)$ subgraph of G, satisfying the following properties,

(I) $C_S \cap (V(M_W) \cup V(M_V) \cup V(H)) = \varnothing$,
(II) $V(M_W) \cap V(M_V) = \varnothing$ and $V(M_W) \cap V(H) = \varnothing$,
(III) C_S sees both sides of each pair in M_W,
(IV) C_S sees exactly one cluster of each pair in M_V,
(V) C_S sees every cluster in X,
(VI) $X \cap V(M_V) = \varnothing$,
(VII) for every $v \in X$, $\deg_H(v) \geq \max\{0, \frac{1}{2}((1+90\sqrt[4]{\varepsilon})k - |V(M_W)|)\}$, and
(VIII) $|V(M_W)| + \frac{1}{2}|V(M_V)| + |X| \geq (1+80\sqrt[4]{\varepsilon})k$.

Proof. Applying Lemma 3.1.11 to G we obtain a set V_0, an $(\varepsilon, \sqrt{\varepsilon})$-upper regular partition on $G - V_0$, a cluster-matching M_1 and an independent family of clusters $\mathcal{C}$ such that,

- $V(M_1) \cap V(\mathcal{C}) = \varnothing$,
- $V(M_1) \cup V(\mathcal{C}) = V(G)$,
- a pair in M_1 is seen in at most one side by the clusters in $\mathcal{C}$, and
- $|V(M_1)| \geq (1+100\sqrt{\varepsilon})k$.

Note that as $|V_0| \leq M_0' < \varepsilon n$, for n sufficiently large, there are still at least $\varepsilon n \geq \varepsilon|V(G-V_0)|$ vertices of degree greater than $(1+100\sqrt[4]{\varepsilon})k$ in $G - V_0$. Also, $\delta(G-V_0) \geq (1+100\sqrt[4]{\varepsilon})\frac{k}{2}$. Now, observe that there is a cluster C_S with $|\{v \in C_S : \deg(v) \geq (1+100\sqrt[4]{\varepsilon})k\}| > \varepsilon m$ and, therefore, $\overline{\deg}(C_S) \geq (1+90\sqrt[4]{\varepsilon})k$.

We will denote by $\mathcal{C}_1$ the set of clusters seen by some cluster in $\mathcal{C}$. If C_S belongs to a pair in M_1, then we denote by $M_1(C_S)$ its neighbor in the matching.

Suppose first that $\overline{\deg}_{M_1}(C_S) \geq (1+90\sqrt{\varepsilon})k$. We can either define $M := M_1 \setminus \{(C_S, M_1(C_S))\}$, if $C_S \subset V(M_1)$, or $M := M_1$, if $C_S \in \mathcal{C}$. Since $|C_S| + |M_1(C_S)| = 2m < 2\sqrt[4]{\varepsilon}k$, we have that $\overline{\deg}_M(C_S) > (1+80\sqrt{\varepsilon})k$. Setting $X, Y, E := \varnothing$ we get the result.

We can now assume that $\overline{\deg}_{M_1}(C_S) < (1+90\sqrt{\varepsilon})k$. Let us call M_1' and $\mathcal{C}'$ the set of pairs in M_1 and the set of clusters in $\mathcal{C}$, respectively, receiving edges from C_S. The set of pairs in M_1'

receiving edges from C_S in both sides will be called M_W^1. The set of pairs in $M_1' \setminus M_W^1$ seen by C_S in the side lying in $\mathcal{C}_1$ will be called M_U. We also define $M_V^1 := M_1' \setminus (M_W^1 \cup M_U)$.

Now, consider the bipartite graph $N := (V(\mathcal{C}') \cup (V(\mathcal{C}_1) \cap V(M_U)), E_G(V(\mathcal{C}'), V(\mathcal{C}_1) \cap V(M_U)))$ and let M_W^2 be a maximal cluster-matching in N. Note that the pairs in M_W^2 receive edges from C_S in both sides. We denote by M_V^2 the set of pairs in M_U not intersecting a pair in M_W^2. Notice that the vertices in $V(M_V^2) \cap V(\mathcal{C}_1)$ are not seen by the vertices in $V(\mathcal{C}') \setminus V(M_W^2)$, since M_W^2 is maximal. Set $M_W := M_W^1 \cup M_W^2$ and $M_V := M_V^1 \cup M_V^2$. Finally define $X := V(\mathcal{C}') \setminus V(M_W^2)$, $Y := V(\mathcal{C}_1) \setminus V(M_W^1 \cup M_U)$, $E := E_G(X, Y)$ and set $H := (X, Y; E)$.

From the construction, points (II)-(VI) are met. Additionally, due to the minimum degree of G and as the vertices in X do not see the vertices in $V(M_V^2)$, we have that for every $v \in X$, $\deg_H(v) \geq \max\{0, \frac{1}{2}((1 + 100\sqrt[4]{\varepsilon})k - |V(M_W)|)\}$. Also, as the neighborhood of C_S is contained in $V(M_W) \cup V(M_V) \cup X$, we have that $|V(M_W)| + \frac{1}{2}|V(M_V)| + |X| \geq (1 + 90\sqrt[4]{\varepsilon})k$. Removing C_S and $M_1(C_S)$, if necessary, from our structures, we get points (I), (VII) and (VIII). □

4.0.3 Main result

Like in Chapter 3, we first prove a preliminary lemma that is thought to be employed once the host graph has already been regularized. Is in this lemma where the embedding procedure is done, and Theorem 4.0.6 will only be a rather direct consequence of it.

Proposition 4.0.5 Let $\varepsilon \in (0, 10^{-8})$. There exists $k_0 = k_0(\varepsilon) \in \mathbb{N}$ such that for all $k \geq k_0$ the following holds. Let G be a graph with n vertices, where $n \geq k \geq 110\sqrt[4]{\varepsilon}n$, which satisfies $\delta(G) \geq (1 + 110\sqrt[4]{\varepsilon})\frac{k}{2}$ and $|\{v \in V(G) : \deg(v) \geq (1 + 110\sqrt[4]{\varepsilon})k\}| > 2\varepsilon n$. Suppose that G admits an $(\frac{\varepsilon}{6}, 2\sqrt[4]{\varepsilon})$-upper regular partition with $M_0' > \frac{1}{100\sqrt{\varepsilon}}$ clusters of size $m' := \frac{n}{M_0}$. There exists $c = c(\varepsilon)$ such that any tree T with k edges and $\Delta(T) \leq ck$ is a subgraph of G.

Proof. Set $\beta < \frac{\sqrt[4]{\varepsilon}}{100M_0}$ and consider t_0, c given by Proposition 4.0.1 and k_0 given by 4.0.4. Let $k \geq \max\{t_0, k_0\}$, let G be a graph as in Proposition 4.0.5 and let T be a tree on k edges and maximum degree bounded by ck. Apply Lemma 4.0.4 to G to obtain a new regular partition over $G - V_0$, where V_0 is of constant size, a cluster C_S, two cluster-matchings M_V, M_W and a bipartite graph $H = (X, Y; E) \subset G$ satisfying properties (I)-(VIII). We will denote by $\mathcal{C}_V$ the set of clusters in M_V that are seen by C_S. Also, we will denote by m the size of the clusters in the new partition, where $m = \lfloor \frac{m'}{2} \rfloor$.

Recalling our convention, $|A(T)| \geq |B(T)|$. Let us pick a vertex $r \in B(T)$ and root T at it. Proposition 4.0.1 allows us to decompose T into a set of seeds S and a family of small pieces $\mathcal{T}$, which for our convenience will be indexed by a set I. Note that $S \subseteq B(T)$.

Our intention is to embed each seed $s \in S$ into the cluster C_S to then map each piece hanging from s, that is, each piece whose root is a child of s, either into an edge of M_W or M_V or into the bipartite graph H. To do this in an orderly manner we will partition $\mathcal{T}$ into two families, one of which will be embedded into $V(M_W)$ while the other one will be embedded into $V(M_V) \cup V(H)$. We obviously need each family to have an appropriate size, but also we will require a certain balance condition to hold over the family of pieces assigned to $V(M_V) \cup V(H)$, since the total amount of vertices going into $Y \cup (V(M_V) \setminus V(\mathcal{C}_V))$ cannot exceed a certain number for our technique to work. Lemma 4.0.2 will provide a partition of $\mathcal{T}$ with such characteristics. Let us precise here that this partition is only needed if M_W is not empty and $\mathcal{T}$ cannot be entirely embedded into $V(M_W)$, so, to place ourselves in the more general setting, we will suppose that this is the case, i.e., we assume that

$$M_W \neq \varnothing \quad \text{and} \quad |V(\mathcal{T})| > (1 - 10\sqrt[4]{\varepsilon})|V(M_W)|. \tag{4.1}$$

Condition (VIII) from Lemma 4.0.4, together with assumption (4.1), ensures that $V(M_V) \cup X$ is not empty either.

Let us partition $\mathcal{T}$. For each $i \in I$ we define $a_i := |A(T) \cap V(T_i)|$ and $b_i := |B(T) \cap V(T_i)|$. Applying Lemma 4.0.2 to the sequences $\{a_i\}_{i \in I}$ and $\{b_i\}_{i \in I}$ with parameters $\lambda = \beta k$ and $M = |V(\mathcal{T})| - (1 - 11\sqrt[4]{\varepsilon})|V(M_W)|$ we obtain a set $J \subseteq I$ such that

$$\sum_{i \in J} |V(T_i)| \leq |V(\mathcal{T})| - (1 - 11\sqrt[4]{\varepsilon})|V(M_W)| \leq (1 - 10\sqrt[4]{\varepsilon})(\frac{1}{2}|V(M_V)| + |X|), \tag{4.2}$$

where the second inequality comes from condition (VIII) in Lemma 4.0.4. The set J also satisfies

$$\frac{\sum_{i \in J} a_i}{\sum_{i \in J} b_i} \geq \frac{\sum_{i \in I} a_i}{\sum_{i \in I} b_i} = \frac{|A(T) \cap V(T - S)|}{|B(T) \cap V(T - S))|} > \frac{|A(T)| - \beta k}{|B(T)|} \geq 1 - \frac{\beta k}{|B(T)|},$$

thus implying that J fulfills the following useful inequality,

$$\sum_{i \in J} a_i \geq \sum_{i \in J} b_i - \beta k. \tag{4.3}$$

Recalling Remark 4.0.3 we also obtain

$$\sum_{i \in I \setminus J} |V(T_i)| \leq (1 - 11\sqrt[4]{\varepsilon})|V(M_W)| + \beta k \leq (1 - 10\sqrt[4]{\varepsilon})|V(M_W)|. \tag{4.4}$$

As we said before, our intention is, for each $i \in J$, to embed T_i into either a pair in M_V or a pair of adjacent clusters in H, while the trees with index in $I \setminus J$ will be embedded in M_W.

We will define our embedding in an iterative process consisting of $|S|$ steps. In each step j we embed a seed $s_j \in S$ into C_S, then we embed each one of the pieces below s_j into $V(M_V) \cup V(H)$ or $V(M_W)$ depending on the index of the piece. We go through S in an order such that the embedding remains connected in each step, starting with the root of T, i.e, $s_1 = r(T)$.

Suppose that we are in the step $j \in \{1, \dots, |S|\}$. We call V_j to the set of vertices in $V(T)$ already mapped into G. The map, which is so far only defined for the vertices in V_j, will be called ϕ. For any set of vertices R, we call $U_j(R)$ to the set of unoccupied vertices in R, i.e., $U_j(R) = R \setminus \phi(V_j)$. Consider any seed $s_j \in S$ not embedded yet, but whose father is already embedded. The set of pieces hanging from s_j will be called $\mathcal{T}_j$. A cluster C with

$|U_j(C)| \geq 7\sqrt[4]{\varepsilon}m$ will be a good cluster. A pair of clusters (C, D) will be a good pair if C and D are adjacent and they are both good clusters. The clusters and the pairs which are not good will be bad clusters and bad pairs.

Five conditions hold at the beginning of each step $j \in \{1, \dots, |S|\}$,

(i) if $j > 1$, the father of s_j is embedded into a vertex typical towards C_S,
(ii) for every cluster C, $|U_j(C)| > 5\sqrt[4]{\varepsilon}m$,
(iii) if $V(T_i) \subseteq V_j$ with $i \in J$, then $\phi(A(T) \cap V(T_i)) \subset (V(\mathcal{C}_V) \cup X)$ and $\phi(B(T) \cap V(T_i)) \subset ((V(M_V) \setminus V(\mathcal{C}_V)) \cup Y)$,
(iv) if $V(T_i) \subseteq V_j$ with $i \in I \setminus J$, then $\phi(V(T_i)) \subset V(M_W)$, and
(v) for a pair $(C, D) \in M_W$, $||\phi(V_j) \cap C| - |\phi(V_j) \cap D|| < \beta k$.

Now, let us partition the family $\{U_j(C) : C \text{ is a cluster adjacent to } C_S\}$ into three families $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$, where $D \in \mathcal{Y}_1$ if $D \subset V(M_W)$, $D \in \mathcal{Y}_2$ if $D \subset V(M_V)$ and $D \in \mathcal{Y}_3$ if $D \subset X$. Condition (ii) ensures that every set in $\mathcal{Y}_i$ is significant, for each $i = 1, 2, 3$. Applying Lemma 1.2.6 to C_S and $\mathcal{Y}_i$, for each $i = 1, 2, 3$, we find a set $C'_S \subset C_S$ with $|C'_S| \geq (1 - 3\sqrt{\varepsilon})m$ such that for every $v \in C'_S$,

$$|\mathcal{T}_v(\mathcal{Y}_i)| \geq (1 - \sqrt{\varepsilon})|\mathcal{Y}_i|, \text{ for each } i = 1, 2, 3. \tag{4.5}$$

Let $f \in V(T)$ be the father of s_j. From condition (i) we know that $\deg_{C_S}(\phi(f)) \geq (\sqrt[4]{\varepsilon} - \varepsilon)m$. This means that $|N_{C_S}(\phi(f)) \cap C'_S| \geq (\sqrt[4]{\varepsilon} - 4\sqrt{\varepsilon})m > \beta k + 1 > |S| + 1$. Therefore, $|N_{U_j(C_S)}(\phi(f)) \cap C'_S| \geq 1$. Pick a vertex $v_j \in N_{U_j(C_S)}(\phi(f)) \cap C'_S$ and set $\phi(s_j) = v_j$.

Now that s_j is embedded into v_j we proceed to embed the trees in $\mathcal{T}_j$ one by one, adopting different strategies depending on the index of the tree, but first we will save some space for the roots of the trees in $\mathcal{T}_j$. Consider a cluster C such that v_j is typical towards $U_j(C)$. Condition (ii) allows us to ensure that $|N_{U_j(C)}(v_j)| \geq 5\sqrt{\varepsilon}m > 2\varepsilon m + ck$, since $c < \beta$. We arbitrarily choose $\lceil 2\varepsilon m + ck \rceil$ vertices in $N_{U_j(C)}(v_j)$ and call the resulting set C_R. We will only embed in C_R the sons of s_j which are roots of trees in $\mathcal{T}_j$. The size of C_R gives us the freedom to apply Lemma 1.3.1 choosing the root of the tree to be typical towards C_S, in case that the root is father of a seed, for every tree in $\mathcal{T}_j$. Let $T_i \in \mathcal{T}_j$ be the next tree we want to embed.

Case 1: $i \in J$.

If $|\phi(V_j) \cap V(M_V)| < (1 - 10\sqrt[4]{\varepsilon})\frac{1}{2}|V(M_V)|$, then there must be more than $\sqrt{\varepsilon}|M_V|$ good pairs in M_V, otherwise

$$\begin{aligned}
(1 + 10\sqrt[4]{\varepsilon})\frac{1}{2}|V(M_V)| &\leq \sum_{(C,D) \in M_V} (|U_j(C)| + |U_j(D)|) \\
&= \sum_{\substack{(C,D) \in M_V \\ (C,D) \text{ good}}} (|U_j(C)| + |U_j(D)|) + \sum_{\substack{(C,D) \in M_V \\ (C,D) \text{ bad}}} (|U_j(C)| + |U_j(D)|) \\
&\leq \sqrt{\varepsilon}|M_V|2m + |M_V|(1 + 7\sqrt[4]{\varepsilon})m
\end{aligned}$$

$$< (1 + 9\sqrt[4]{\varepsilon})\frac{1}{2}|V(M_V)|$$

which is a contradiction. Therefore, using (4.5), we know that there must be a good pair $(C, D) \in M_V$ such that v_j is typical to $U_j(D)$. Map T_i into $(U_j(C), U_j(D))$ avoiding C_R and D_R. The root of T_i must be embedded into D_R, and thus all the vertices in $A(T) \cap V(T_i)$ will also be mapped into $D \in \mathcal{C}_V$. We also take care of embedding the vertices in $V(T_i)$ that are fathers of seeds into vertices typical towards C_S.

If $|\phi(V_j) \cap V(M_V)| > (1 - 10\sqrt[4]{\varepsilon})\frac{1}{2}|V(M_V)|$, then we have to consider a good pair in H to do the embedding of T_i. Let us call $\mathcal{C}_X$ the set of clusters contained in X. Since $\sum_{i\in J} |V(T_i)| \leq (1 - 10\sqrt[4]{\varepsilon})(\frac{1}{2}|V(M_V)| + |X|)$, we have that $|\phi(V_j) \cap X| < (1 - 10\sqrt[4]{\varepsilon})|X|$. Hence, there must be more than $\sqrt{\varepsilon}|\mathcal{C}_X|$ good clusters in $\mathcal{C}_X$, otherwise

$$\begin{aligned}
10\sqrt[4]{\varepsilon}|X| &< \sum_{C\in\mathcal{C}_X} |U_j(C)| \\
&= \sum_{\substack{C\in\mathcal{C}_X \\ C \text{ good}}} |U_j(C)| + \sum_{\substack{C\in\mathcal{C}_X \\ C \text{ bad}}} |U_j(C)| \\
&\leq \sqrt{\varepsilon}|\mathcal{C}_X|m + |\mathcal{C}_X|7\sqrt[4]{\varepsilon}m \\
&< 8\sqrt[4]{\varepsilon}|X|
\end{aligned}$$

Thus, there is a good cluster $C \in \mathcal{C}_X$ such that v_j is typical towards $U_j(C)$. Now, we know that $\deg_H(v) \geq \max\{0, \frac{1}{2}((1 + 90\sqrt[4]{\varepsilon})k - |V(M_W)|)\}$, for every $v \in X$. We will assume that the minimum degree in H of the vertices in X is greater than zero, otherwise $|V(M_W)| \geq (1 + 90\sqrt[4]{\varepsilon})k$ and, therefore, $k \leq (1 - 10\sqrt[4]{\varepsilon})|V(M_W)|$, contradicting assumption (4.1). Thus, the neighborhood of C covers at least $\frac{1}{2}((1 + 90\sqrt[4]{\varepsilon})k - |V(M_W)|) > 0$ vertices in Y.

Making use of condition (iii), property (4.3) and the fact that $\sum_{i\in J} |V(T_i)| \leq |V(\mathcal{T})| - (1 - 11\sqrt[4]{\varepsilon})|V(M_W)|$, we get

$$\begin{aligned}
|\phi(V_j) \cap Y| &\leq \frac{1}{2}\left[\sum_{\mathrm{i}\in J} |V(T_{\mathrm{i}})| + \beta k\right] \\
&< \frac{1}{2}\left[|V(\mathcal{T})| - (1 - 11\sqrt[4]{\varepsilon})|V(M_W)| + \beta k\right] \qquad (4.6) \\
&< \frac{1}{2}\left[k - (1 - 11\sqrt[4]{\varepsilon})|V(M_W)| + \beta k\right]
\end{aligned}$$

Suppose that there are not good clusters in the neighborhood of C in H. Then, due to the minimum degree of the vertices of C in H,

$$\begin{aligned}
|\phi(V_j) \cap Y| &> (1 - 7\sqrt[4]{\varepsilon})|\{v \in D \,:\, D \text{ is a cluster contained in } Y \text{ and } D \text{ is adjacent to } C\}| \\
&\geq (1 - 7\sqrt[4]{\varepsilon})\frac{1}{2}\left[(1 + 90\sqrt[4]{\varepsilon})k - |V(M_W)|\right]
\end{aligned}$$

Joining this inequality with (4.6) we obtain

$$\begin{aligned}
4\sqrt[4]{\varepsilon}|V(M_W)| &> (83\sqrt[4]{\varepsilon} - 630\sqrt[2]{\varepsilon} - \beta)k \\
&> 10\sqrt[4]{\varepsilon}|V(\mathcal{T})|
\end{aligned}$$

$$
\begin{aligned}
&> 10\sqrt[4]{\varepsilon}(1 - 10\sqrt[4]{\varepsilon})|V(M_W)| \\
&> 4\sqrt[4]{\varepsilon}|V(M_W)|
\end{aligned}
$$

where we have used the inequality in (4.1). Thus, there is a good cluster $D \subset Y$ neighbor of C. Map T_i into $(U_j(C), U_j(D))$ avoiding C_R and D_R. The root of T_i must be mapped into C_R. We take care of embedding the vertices in $V(T_i)$ that are fathers of seeds into vertices typical towards C_S.

Case 2: $i \in I \setminus J$.

From (4.4) we know that $|\phi(V_j) \cap V(M_W)| < (1 - 10\sqrt[4]{\varepsilon})|V(M_W)|$ and, therefore, there must be more than $2\sqrt{\varepsilon}|M_W|$ good pairs in M_W, otherwise using condition (v) we obtain,

$$
\begin{aligned}
10\sqrt[4]{\varepsilon}|V(M_W)| &\geq \sum_{(C,D)\in M_W} (|U_j(C)| + |U_j(D)|) \\
&= \sum_{\substack{(C,D)\in M_W \\ (C,D) \text{ good}}} (|U_j(C)| + |U_j(D)|) + \sum_{\substack{(C,D)\in M_W \\ (C,D) \text{ bad}}} (|U_j(C)| + |U_j(D)|) \\
&\leq 2\sqrt{\varepsilon}|M_W|2m + |M_W|(7\sqrt[4]{\varepsilon}m + 7\sqrt[4]{\varepsilon}m + \beta k) \\
&< 8\sqrt[4]{\varepsilon}|V(M_W)|
\end{aligned}
$$

Thus, we know that there must be a good pair $(C, D) \in M_W$ such that v_j is typical to $U_j(C)$ and to $U_j(D)$. We are going to map T_i into $(U_j(C), U_j(D))$ avoiding C_R and D_R. If $|U_j(C)| > |U_j(D)|$, then we chose the side for the root in a way that $A(T_i)$ is mapped into C and $B(T_i)$ is mapped into D. If $|U_j(C)| \leq |U_j(D)|$, then we map $A(T_i)$ into D and $B(T_i)$ into C. This ensures that condition (v) remains true. The root of T_i must be mapped into the R-slice of the cluster it is assigned. We also take care of embedding the vertices in $V(T_i)$ that are fathers of seeds into vertices typical towards C_S. This completes the proof. □

Theorem 4.0.6 Let $\delta \in (0, 1)$. There exist $n_0 = n_0(\delta) \in \mathbb{N}$ and $c = c(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph with n vertices which has minimum degree at least $(1+\delta)\frac{k}{2}$ and $|\{v \in V(G) : \deg(v) \geq (1+\delta)k\}| \geq \delta n$, with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most ck, then T is a subgraph of G.

Proof. Given δ, we suitably choose ε and d such that

$$0 < \varepsilon \ll d \ll \delta. \tag{4.7}$$

Let N_0, M_0' be given by Proposition 1.2.4 for input ε and $m_0 := \frac{1}{\varepsilon}$. Let k_0 be the output of Proposition 4.0.5, and set $n_0 := \max\{N_0, \delta^{-1}k_0\}$.

Now given an n-vertex graph G as in Theorem 4.0.6, with $n \geq n_0$, apply Proposition 1.2.4 to G to obtain a subgraph $G_d \subseteq G$ that admits an (ε, d)-upper regular partition, with at most M_0' clusters, and such that $|V(G) \setminus V(G_d)| \leq \varepsilon n$ and $\deg_{G_d}(x) \geq \deg_G(x) - (d+\varepsilon)n$, for all $x \in V(G_d)$. This implies that $\delta(G_d) \geq (1+\frac{\delta}{2})\frac{k}{2}$ and that $|\{v \in V(G_d) : \deg(v) \geq (1+\frac{\delta}{2})k\}| \geq \frac{\delta}{2}n$.

Hence, we can apply Proposition 4.0.5 to G_d to conclude. □

Chapter 5

On the Erdős-Sós conjecture

In this final chapter we give an affirmative answer to a partial approximated version of the Erdős-Sós conjecture in the case where the host graph is dense, i.e., the case where the number of edges grows quadratically on the number of vertices. Our result is just a simple consequence of Theorem 4.0.6.

Theorem 5.0.1 Let $\delta \in (0,1)$. There exist $n_0 = n_0(\delta) \in \mathbb{N}$ and $c = c(\delta) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph on n vertices which satisfies

$$\mathrm{d}(G) > k - 1 + \delta k,$$

with $n \geq k \geq \delta n$. If T is a tree with k edges and maximum degree at most ck, then T is a subgraph of G.

Proof. Given $\delta' := \frac{\delta^2}{2}$ consider k_0 and c from Theorem 4.0.6. Let G be an n-vertex graph with $\mathrm{d}(G) > k - 1 + \delta k$, where $n \geq k \geq \delta n$ and $k \geq k_0$. By Fact 1.1.1 we obtain a graph $H \subseteq G$ such that

$$\delta(H) \geq (1+\delta)\frac{k}{2} \quad \text{and} \quad \mathrm{d}(H) \geq (1+\delta)k. \tag{5.1}$$

We denote by n_0 the order of H. Let us define $L := \{v \in V(H) : \deg_H(v) \geq (1+\frac{\delta}{2})k\}$ and $S := V(H) \setminus L$. Note that $|L| \geq \frac{\delta^2}{2} n_0$, otherwise

$$\frac{\delta^2}{2} n_0^2 + (1+\frac{\delta}{2})k \cdot n_0 > |L| \cdot n_0 + |S| \cdot (1+\frac{\delta}{2})k \geq \sum_{v \in V(H)} \deg(v) \geq (1+\delta)k \cdot n_0,$$

which in turn implies that $\delta^2 n_0 > \delta k$, a contradiction to the fact that $k \geq \delta n \geq \delta n_0$. Therefore, applying Theorem 4.0.6 to H with parameter δ' we get the result. □

Conclusion

As we have previously said, part of this work can be seen as a generalization of the result of Komlós, Sárközy and Szemerédi [30], but part of its relevance also lies in the fact that it provides support for the $\frac{2}{3}$ conjecture and for the Erdős-Sós conjecture, by means of Theorems 3.2.8 and 5.0.1 respectively.

As we have seen in Section 3.1, Propositions 3.1.2 and 3.1.7 establish an explicit relation between the diameter of the host graph and the degree of the trees we are able to embed by means of our technique. In both propositions we embed the first levels of each piece through a path leading to a pair of clusters with space. In order to maintain control on the embedding we need that the number of vertices that go into these paths is not too large; that is why we impose a restriction on the degree of the trees: the greater the diameter, the longer might be the paths connecting good pairs and, therefore, the smaller has to be the bound on the degree of the trees. Proposition 3.1.14 shows that this dependence is not necessary, and that it can be replaced by a dependence on the minimum degree, thus improving the bound on the degree of the trees. The key idea here is that we do not need to go so far to find a pair of clusters with space to continue the embedding. Thus, we think that a possible improvement or extension of the present work would be lowering this distance, or finding a workable structure in the cases requiring longer distances, for this would enlarge the class of embeddable trees. The interesting thing about this is that it would have direct impact in all the Theorems of Section 3.2, raising the bound on the degree of the trees.

Another possible extension to this work would consist in solving the approximated dense case of the Erdős-Sós conjecture, i.e., to be able to remove the bound on the degree of the trees in Theorem 5.0.1. It is not difficult to prove that, when a graph has average degree at least k, the average degree of its line graph is at least $2k$ and, therefore, one can find an edge which is adjacent to at least $2k$ edges. In the approximated dense case this result could be applied to the reduced graph to find a pair of adjacent clusters (C, D) satisfying $\overline{\deg}(C) + \overline{\deg}(D) \geq (1 + \delta')2k$. We believe this special pair of clusters could be used to do the embedding in a similar way to Proposition 4.0.5, by considering two sets of seeds instead of one, as Piguet and Stein do in [38].

Bibliography

[1] Noga Alon, Michael Krivelevich, and Benny Sudakov. Embedding nearly-spanning bounded degree trees. *Combinatorica*, 27(6):629–644, 2007.

[2] Suman Balasubramanian and Edward Dobson. Constructing trees in graphs with no $K_{2,s}$. *Journal of Graph Theory*, 56(4):301–310, 2007.

[3] József Balogh, Béla Csaba, Martin Pei, and Wojciech Samotij. Large bounded degree trees in expanding graphs. *The Electronic Journal of Combinatorics*, 17(1):6, 2010.

[4] József Balogh, Béla Csaba, and Wojciech Samotij. Local resilience of almost spanning trees in random graphs. *Random Structures & Algorithms*, 38(1-2):121–139, 2011.

[5] Guido Besomi, Matías Pavez, and Maya Stein. Embedding trees of bounded degree in dense graphs. In preparation, 2018.

[6] Béla Bollobás. *Extremal graph theory*. Courier Corporation, 2004.

[7] Stephan Brandt and Edward Dobson. The Erdős-Sós conjecture for graphs of girth 5. *Discrete Mathematics*, 150(1-3):411–414, 1996.

[8] Béla Csaba, Judit Nagy-György, Ian Levitt, and Endre Szemerédi. Tight bounds for embedding bounded degree trees. In *Fete of combinatorics and computer science*, pages 95–137. Springer, 2010.

[9] Edward Dobson. Constructing trees in graphs whose complement has no $K_{2,s}$. *Combinatorics, Probability and Computing*, 11(4):343–347, 2002.

[10] Nancy Eaton and Gary Tiner. On the Erdős-Sós conjecture and graphs with large minimum degree. *Ars Combinatoria*, 95:373–382, 2010.

[11] Nancy Eaton and Gary Tiner. On the Erdős-Sós conjecture for graphs having no path with k+4 vertices. *Discrete Mathematics*, 313(16):1621–1629, 2013.

[12] Paul Erdős, Zoltán Füredi, Martin Loebl, and Vera T. Sós. Discrepancy of trees. *Studia Scientiarum Mathematicarum Hungarica*, 30(1):47–58, 1995.

[13] Paul Erdős and Tibor Gallai. On maximal paths and circuits of graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 10(3-4):337–356, 1959.

[14] Paul Erdős, Janos Pach, Richard Pollack, and Zsolt Tuza. Radius, diameter, and minimum degree. *Journal of Combinatorial Theory, Series B*, 47(1):73–79, 1989.

[15] Paul Erdős and Vera T. Sós. Mentioned in Paul Erdős, Extremal problems in graph theory, Theory of graphs and its applications. In *Proc. of the Symposium held in Smolenice in June*, pages 29–38, 1963.

[16] Asaf Ferber, Rajko Nenadov, and Ueli Peter. Universality of random graphs and rainbow embedding. *Random Structures & Algorithms*, 48(3):546–564, 2016.

[17] Asaf Ferber and Wojciech Samotij. Packing trees of unbounded degrees in random graphs. *arXiv preprint arXiv:1607.07342*, 2016.

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

[19] Agnieszka Goerlich and Andrzej Żak. On Erdős-Sós conjecture for trees of large size. *The Electronic Journal of Combinatorics*, 23(1):1–52, 2016.

[20] Frédéric Havet, Bruce Reed, Maya Stein, and David R. Wood. A variant of the Erdős-Sós conjecture. *arXiv preprint arXiv:1606.09343*, 2016.

[21] Penny E. Haxell. Tree embeddings. *Journal of Graph Theory*, 36(3):121–130, 2001.

[22] Dan Hefetz, Michael Krivelevich, and Tibor Szabó. Sharp threshold for the appearance of certain spanning trees in random graphs. *Random Structures & Algorithms*, 41(4):391–412, 2012.

[23] Jan Hladký, János Komlós, Diana Piguet, Miklós Simonovits, Maya Stein, and Endre Szemerédi. The approximate Loebl–Komlós–Sós Conjecture I: The sparse decomposition. *SIAM Journal on Discrete Mathematics*, 31(2):945–982, 2017.

[24] Jan Hladký, János Komlós, Diana Piguet, Miklós Simonovits, Maya Stein, and Endre Szemerédi. The approximate Loebl–Komlós–Sós Conjecture II: The rough structure of LKS graphs. *SIAM Journal on Discrete Mathematics*, 31(2):983–1016, 2017.

[25] Jan Hladký, János Komlós, Diana Piguet, Miklós Simonovits, Maya Stein, and Endre Szemerédi. The approximate Loebl–Komlós–Sós Conjecture III: The finer structure of LKS graphs. *SIAM Journal on Discrete Mathematics*, 31(2):1017–1071, 2017.

[26] Jan Hladký, János Komlós, Diana Piguet, Miklós Simonovits, Maya Stein, and Endre Szemerédi. The approximate Loebl–Komlós–Sós Conjecture IV: Embedding techniques and the proof of the main result. *SIAM Journal on Discrete Mathematics*, 31(2):1072–1148, 2017.

[27] Jan Hladký and Diana Piguet. Loebl–Komlós–Sós conjecture: dense case. *Journal of Combinatorial Theory, Series B*, 116:123–190, 2016.

[28] Daniel Johannsen, Michael Krivelevich, and Wojciech Samotij. Expanders are universal

for the class of all spanning trees. *Combinatorics, Probability and Computing*, 22(2):253–281, 2013.

[29] Jeff Kahn, Eyal Lubetzky, and Nicholas Wormald. The threshold for combs in random graphs. *Random Structures & Algorithms*, 48(4):794–802, 2016.

[30] János Komlós, Gábor N Sárközy, and Endre Szemerédi. Proof of a packing conjecture of Bollobás. *Combinatorics, Probability and Computing*, 4(3):241–255, 1995.

[31] János Komlós, Gábor N Sárkózy, and Endre Szemerédi. Spanning trees in dense graphs. *Combinatorics, Probability and Computing*, 10(5):397–416, 2001.

[32] Michael Krivelevich. Embedding spanning trees in random graphs. *SIAM Journal on Discrete Mathematics*, 24(4):1495–1500, 2010.

[33] Michael Krivelevich, Matthew Kwan, and Benny Sudakov. Bounded-degree spanning trees in randomly perturbed graphs. *SIAM Journal on Discrete Mathematics*, 31(1):155–171, 2017.

[34] Daniela Kühn and Deryk Osthus. Embedding large subgraphs into dense graphs. *arXiv preprint arXiv:0901.3541*, 2009.

[35] Wang Min, Li Guo-jun, and Liu Ai-de. A result of Erdős-Sós cojecture. *Ars Combinatoria*, 55:123–128, 2000.

[36] Richard Montgomery. Embedding bounded degree spanning trees in random graphs. *arXiv preprint arXiv:1405.6559*, 2014.

[37] Richard Montgomery. Sharp threshold for embedding combs and other spanning trees in random graphs. *arXiv preprint arXiv:1405.6560*, 2014.

[38] Diana Piguet and Maya Stein. An approximate version of the Loebl–Komlós–Sós conjecture. *Journal of Combinatorial Theory, Series B*, 102(1):102–125, 2012.

[39] Bruce Reed and Maya Stein. Personal communication, 2018.

[40] Jean-François Saclé and Mariusz Woźniak. The Erdős-Sós conjecture for graphs without C_4. *Journal of Combinatorial Theory, Series B*, 70(2):367–372, 1997.

[41] P.J. Slater, S.K. Teo, and H.P. Yap. Packing a tree with a graph of the same size. *Journal of Graph Theory*, 9(2):213–216, 1985.

[42] Sara Nadiv Soffer. The Komlós–Sós conjecture for graphs of girth 7. *Discrete Mathematics*, 214(1-3):279–283, 2000.

[43] Endre Szemerédi. Regular partitions of graphs. Technical report, Stanford University, Stanford, CA, USA, 1975.

[44] Gary Tiner. On the Erdős-Sós conjecture for graphs on $n = k + 3$ vertices. *Ars Combinatoria*, 95:143–150, 2010.

[45] Paul Turán. On an extremal problem in graph theory. Matematikai és Fizikai Lapok, 48: 436–452, 1941.

[46] Mariusz Wozniak. On the Erdős-Sós conjecture. *Journal of Graph Theory*, 21(2):229–234, 1996.

[47] Long-Tu Yuan and Xiao-Dong Zhang. On the Erdős-Sós conjecture for graphs on $n = k + 4$ vertices. *arXiv preprint arXiv:1403.5430*, 2014.

[48] Yi Zhao. Proof of the $(n/2 - n/2 - n/2)$ conjecture for large n. *The Electronic Journal of Combinatorics*, 18(1):27, 2011.

[49] Bing Zhou. A note on the Erdős-Sós conjecture. *Acta Mathematica Scientia*, 4(3):287–289, 1984.