INSTITUTE OF MATHEMATICS

The approximate Loebl-Komlós-Sós Conjecture II: The rough structure of LKS graphs

Jan Hladký

János Komlós

Diana Piguet

Miklós Simonovits

Maya Stein

Endre Szemerédi

The approximate Loebl–Komlós–Sós Conjecture II: The rough structure of LKS graphs

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

Abstract

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

In the first paper of the series, we gave a decomposition of the graph G into several parts of different characteristics; this decomposition might be viewed as an analogue of a regular partition for sparse graphs. In the present paper, we find a combinatorial structure inside this decomposition. In the last two papers, we refine the structure and use it for embedding the tree T.

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

Contents

1 Introduction

This is the second of a series of four papers [HKP$^+$a, HKP$^+$b, HKP$^+$c, HKP$^+$d] in which we provide an approximate solution of the Loebl–Komlós–Sós Conjecture. The conjecture reads as follows.

Conjecture 1.1 (Loebl–Komlós–Sós Conjecture 1995 [EFLS95])**.** *Suppose that G is an n-vertex graph with at least $n/2$ vertices of degree more than $k - 2$. Then G contains each tree of order k.*

We discuss the history and state of the art in detail in the first paper [HKP$^+$a] of our series. The main result, which will be proved in [HKP$^+$d], is the approximate solution of the Loebl–Komlós–Sós Conjecture.

Theorem 1.2 (Main result [HKP$^+$d])**.** *For every $\alpha > 0$ there exists a number k_0 such that for any $k > k_0$ we have the following. Each n-vertex graph G with at least $(\frac{1}{2} + \alpha)n$ vertices of degree at least $(1 + \alpha)k$ contains each tree T of order k.*

In the first paper [HKP$^+$a] we exposed the decomposition techniques we use on the host graph. In particular, we saw in [HKP$^+$a, Lemma 3.14] that any graph satisfying the assumptions of Theorem 1.2 may be decomposed into a set of huge degree vertices, regular pairs, an expanding subgraph, and another set with certain expansion properties, which we call the avoiding set. We call this a *sparse decomposition* of a graph. We will recall the necessary notions from [HKP$^+$a] in Section 3.

Many embedding problems for dense host graphs are attacked using the following three-step approach: *(a)* the regularity lemma is applied to the host graph, *(b)* a suitable combinatorial structure is found in the cluster graph, and *(c)* the target graph is embedded into the combinatorial structure using properties of regular pairs. If we consider the sparse decomposition as a sparse counterpart to *(a)* then the main result of the present paper, Lemma 5.4, should be regarded as a counterpart to *(b)*. More precisely, for each graph satisfying the assertions of Theorem 1.2 that is given together with its sparse decomposition, Lemma 5.4 gives a combinatorial structure whose building blocks are the elements of the sparse decomposition. As in tree embedding problems in the dense setting (e.g. in [AKS95, PS12]), the core of this combinatorial structure is a well-connected matching consisting of regular pairs. We call such matchings *regularized*.

With the structure given by Lemma 5.4, one can convince oneself that the tree T from Theorem 1.2 can be embedded into the host graph, and indeed we provide such motivation in Section 5.1. However, the rigorous argument is far from trivial. One needs to refine the structure found here, which is done in [HKP$^+$c]. For this reason, we call the output of Lemma 5.4 *the rough structure*. In the last paper [HKP$^+$d] of our series we will develop embedding techniques for trees, and finally prove Theorem 1.2.

2 Notation and preliminaries

2.1 General notation

The set $\{1, 2, \ldots, n\}$ of the first n positive integers is denoted by $[n]$. We frequently employ indexing by many indices. We write superscript indices in parentheses (such as $a^{(3)}$), as opposed to notation of powers (such as a^3). We use sometimes subscript to refer to parameters appearing in a fact/lemma/theorem. For example $\alpha_{\mathrm{T}1.2}$ refers to the parameter α from Theorem 1.2. We omit rounding symbols when this does not affect the correctness of the arguments.

Table 2.1: Specific notation used in the series.

lower case Greek letters	small positive constants ($\ll 1$)
	ϕ reserved for embedding; $\phi : V(T) \rightarrow V(G)$
upper case Greek letters	large positive constants ($\gg 1$)
one-letter bold	sets of clusters
bold (e.g., $\mathbf{trees}(k), \mathbf{LKS}(n, k, \eta)$)	classes of graphs
blackboard bold (e.g., $\mathbb{H}, \mathbb{E}, \mathbb{S}_{\eta,k}(G), \mathbb{XA}$)	distinguished vertex sets except for
	$\mathbb{N}$ which denotes the set $\{1, 2, \ldots\}$
script (e.g., $\mathcal{A}, \mathcal{D}, \mathcal{N}$)	families (of vertex sets, "dense spots", and regular pairs)
$\nabla(=\text{nabla})$	sparse decomposition (see Definition 3.5)

Table 2.1 shows the system of notation we use in the series.

We write $V(G)$ and $E(G)$ for the vertex set and edge set of a graph G, respectively. Further, $v(G) = |V(G)|$ is the order of G, and $e(G) = |E(G)|$ is its number of edges. If $X, Y \subseteq V(G)$ are two, not necessarily disjoint, sets of vertices we write $e(X)$ for the number of edges induced by X, and $e(X, Y)$ for the number of ordered pairs $(x, y) \in X \times Y$ such that $xy \in E(G)$. In particular, note that $2e(X) = e(X, X)$.

For a graph G, a vertex $v \in V(G)$ and a set $U \subseteq V(G)$, we write $\deg(v)$ and $\deg(v, U)$ for the degree of v, and for the number of neighbours of v in U, respectively. We write $\text{mindeg}(G)$ for the minimum degree of G, $\text{mindeg}(U) := \min\{\deg(u) : u \in U\}$, and $\text{mindeg}(V_1, V_2) = \min\{\deg(u, V_2) : u \in V_1\}$ for two sets $V_1, V_2 \subseteq V(G)$. Similar notation is used for the maximum degree, denoted by $\text{maxdeg}(G)$. The neighbourhood of a vertex v is denoted by $\text{N}(v)$. We set $\text{N}(U) := \bigcup_{u \in U} \text{N}(u)$. The symbol $-$ is used for two graph operations: if $U \subseteq V(G)$ is a vertex set then $G - U$ is the subgraph of G induced by the set $V(G) \setminus U$. If $H \subseteq G$ is a subgraph of G then the graph $G - H$ is defined on the vertex set $V(G)$ and corresponds to deletion of edges of H from G. Any graph with zero edges is called *empty*. A family $\mathcal{A}$ of pairwise disjoint subsets of $V(G)$ is an *ℓ-ensemble* in G if $|A| \geqslant \ell$ for each $A \in \mathcal{A}$.

Finally, $\mathbf{trees}(k)$ denotes the class of all trees of order k.

2.2   Regular pairs

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

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

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

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

Fact 2.1. *Suppose that (U, W) is an ε-regular pair of density d. Let $U' \subseteq W, W' \subseteq W$ be sets of vertices with $|U'| \geqslant \alpha|U|$, $|W'| \geqslant \alpha|W|$, where $\alpha > \varepsilon$. Then the pair (U', W') is a $2\varepsilon/\alpha$-regular pair of density at least $d - \varepsilon$.*

The regularity lemma [Sze78] has proved to be a powerful tool for attacking graph embedding problems; see [KO09] for a survey.

Lemma 2.2 (Regularity lemma)**.** *For all $\varepsilon > 0$ and $\ell \in \mathbb{N}$ there exist $n_0, M \in \mathbb{N}$ such that for every $n \geqslant n_0$ the following holds. Let G be an n-vertex graph whose vertex set is pre-partitioned into sets $V_1, \ldots, V_{\ell'}$, $\ell' \leqslant \ell$. Then there exists a partition $U_0, U_1, \ldots, U_p$ of $V(G)$, $\ell < p < M$, with the following properties.*

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

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

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

We shall use Lemma 2.2 for auxiliary purposes only as it is helpful only in the setting of dense graphs (i.e., graphs which have ℓ vertices and $\Omega(\ell^2)$ edges).

2.3 LKS graphs

It will be convenient to restrict our attention to a class of graphs which is in a way minimal for Theorem 1.2. Write $\mathbf{LKS}(n, k, \alpha)$ for the class of all n-vertex graphs with at least $(\frac{1}{2} + \alpha)n$ vertices of degrees at least $(1 + \alpha)k$. With this notation Conjecture 1.1 states that every graph in $\mathbf{LKS}(n, k, 0)$ contains every tree from $\mathbf{trees}(k + 1)$.

Given a graph G, denote by $\mathbb{S}_{\eta,k}(G)$ the set of those vertices of G that have degree less than $(1 + \eta)k$ and by $\mathbb{L}_{\eta,k}(G)$ the set of those vertices of G that have degree at least $(1 + \eta)k$. When proving Theorem 1.2, we may of course restrict our attention to LKS-minimal graphs, that is, to graphs that are edge-minimal with respect to the membership to $\mathbf{LKS}(n, k, \alpha)$. It is easy to show that in each such graph the set $\mathbb{S}_{\eta,k}(G)$ is independent, all the neighbours of every vertex $v \in V(G)$ with $\deg(v) > \lceil (1 + \eta)k \rceil$ have degree exactly $\lceil (1 + \eta)k \rceil$, and $|\mathbb{L}_{\eta,k}(G)| \leqslant \lceil (1/2 + \eta)n \rceil + 1$. It turns out that our main decomposition result [HKP$^+$a, Lemma 3.14] outputs a graph with slightly weaker properties than being LKS-minimal. Let us therefore introduce this graph class.

Definition 2.3. *Suppose that $n, k \in \mathbb{N}$, and $\eta > 0$. Let $\mathbf{LKSsmall}(n, k, \eta)$ be the class of those graphs $G \in \mathbf{LKS}(n, k, \eta)$ for which we have the following three properties:*

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

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

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

3 Decomposing sparse graphs

In [HKP$^+$a] we introduced the notion of sparse decomposition, and proved that every graph can be (almost perfectly) decomposed. We define the sparse decomposition after introducing its basic building blocks: dense spots and avoiding sets. For motivation and more details we refer the reader to [HKP$^+$a, Section 3.2], of which this section is a condensed version.

We start by defining dense spots. These are bipartite graphs having positive density, and will (among other things) serve as a basis for regularization.

Definition 3.1 ((m, γ)-**dense spot**, (m, γ)-**nowhere-dense**). *Suppose that $m \in \mathbb{N}$ and $\gamma > 0$. An (m, γ)-dense spot in a graph G is a non-empty bipartite subgraph $D = (U, W; F)$ of G with $\mathrm{d}(D) > \gamma$ and $\mathrm{mindeg}(D) > m$. We call a graph G (m, γ)-nowhere-dense if it does not contain any (m, γ)-dense spot.*

When the parameters m and γ are not relevant, we refer to D simply as a dense spot.

Note that dense spots do not have any specified orientation. That is, we view $(U, W; F)$ and $(W, U; F)$ as the same object.

Definition 3.2 ((m, γ)-**dense cover**). *Suppose that $m \in \mathbb{N}$ and $\gamma > 0$. An (m, γ)-dense cover of a given graph G is a family $\mathcal{D}$ of edge-disjoint (m, γ)-dense spots such that $E(G) = \bigcup_{D \in \mathcal{D}} E(D)$.*

We now define the avoiding set. Informally, a set $\mathbb{E}$ of vertices is avoiding if for each set U of size up to Λk (where $\Lambda \gg 1$ is a large constant) and each vertex $v \in \mathbb{E}$ there is a dense spot containing v and almost disjoint from U. Favourable properties of avoiding sets for embedding trees are shown in [HKP$^+$a, Section 3.5].

Definition 3.3 (($\Lambda, \varepsilon, \gamma, k$)-**avoiding set**). *Suppose that $k \in \mathbb{N}$, $\varepsilon, \gamma > 0$ and $\Lambda > 0$. Suppose that G is a graph and $\mathcal{D}$ is a family of dense spots in G. A set $\mathbb{E} \subseteq \bigcup_{D \in \mathcal{D}} V(D)$ is ($\Lambda, \varepsilon, \gamma, k$)-avoiding with respect to $\mathcal{D}$ if for every $U \subseteq V(G)$ with $|U| \leqslant \Lambda k$ the following holds for all but at most εk vertices $v \in \mathbb{E}$. There is a dense spot $D \in \mathcal{D}$ with $|U \cap V(D)| \leqslant \gamma^2 k$ that contains v.*

We can now introduce an auxiliary notion of bounded decomposition on which we can build the key concept of sparse decomposition (see below). The main result in [HKP$^+$a] tells us that every graph has an almost perfect sparse decomposition. This sparse decomposition (and the bounded decomposition included in it) will provide us with control on the behaviour of the different edge and vertex sets involved, and thus be helpful for the embedding of the tree.

Definition 3.4 (($k, \Lambda, \gamma, \varepsilon, \nu, \rho$)-**bounded decomposition**). *Suppose that $k \in \mathbb{N}$ and $\varepsilon, \gamma, \nu, \rho > 0$ and $\Lambda > 0$. Let $\mathcal{V} = \{V_1, V_2, \ldots, V_s\}$ be a partition of the vertex set of a graph G. We say that $(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ is a ($k, \Lambda, \gamma, \varepsilon, \nu, \rho$)-bounded decomposition of G with respect to $\mathcal{V}$ if the following properties are satisfied:*

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

2. *$\mathbf{V}$ is a family of disjoint subsets of $V(G)$.*

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

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

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

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

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

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

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

The members of $\mathbf{V}$ are called *clusters*. Define the *cluster graph* $\mathbf{G}_{\mathrm{reg}}$ as the graph on the vertex set $\mathbf{V}$ that has an edge $C_1 C_2$ for each pair (C_1, C_2) which has density at least γ^2 in the graph G_{reg}. Further, we define the graph $G_{\mathcal{D}}$ as the union (both edge-wise, and vertex-wise) of all dense spots $\mathcal{D}$.

We now enhance the structure of bounded decomposition by adding one new feature: vertices of very large degree.

Definition 3.5 $((k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-**sparse decomposition**)**.** *Suppose that $k \in \mathbb{N}$ and $\varepsilon, \gamma, \nu, \rho > 0$ and $\Lambda, \Omega^*, \Omega^{**} > 2$. Let $\mathcal{V} = \{V_1, V_2, \ldots, V_s\}$ be a partition of the vertex set of a graph G. We say that $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ is a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-sparse decomposition of G with respect to $V_1, V_2, \ldots, V_s$ if the following hold.*

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

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

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

Definition 3.6 (**captured edges**)**.** *In the situation of Definition 3.5, we refer to the edges in $E(G_{\mathrm{reg}}) \cup E(G_{\mathrm{exp}}) \cup E_G(\mathbb{H}, V(G)) \cup E_G(\mathbb{E}, \mathbb{E} \cup \bigcup \mathbf{V})$ as* captured *by the sparse decomposition. We write G_∇ for the subgraph of G on the vertex set $V(G)$ which consists of the captured edges. Likewise, the captured edges of a bounded decomposition $(\mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ of a graph G are those in $E(G_{\mathrm{reg}}) \cup E(G_{\mathrm{exp}}) \cup E_{G_{\mathcal{D}}}(\mathbb{E}, \mathbb{E} \cup \bigcup \mathbf{V})$.*

It will be useful to have the following shorthand notation at hand.

Definition 3.7 ($\mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$ *and* $\bar{\mathcal{G}}(n, k, \Omega, \rho, \nu)$)**.** *Suppose that $k, n \in \mathbb{N}$ and $\nu, \rho, \tau > 0$ and $\Omega > 0$. We define $\mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$ to be the class of all tuples $(G, \mathcal{D}, H, \mathcal{A})$ with the following properties:*

(i) *G is a graph of order n with $\mathrm{maxdeg}(G) \leqslant \Omega k$,*

(ii) *H is a bipartite subgraph of G with colour classes A_H and B_H and with $e(H) \geqslant \tau k n$,*

(iii) *$\mathcal{D}$ is a $(\rho k, \rho)$-dense cover of G,*

(iv) *$\mathcal{A}$ is a (νk)-ensemble in G, and $A_H \subseteq \bigcup \mathcal{A}$,*

(v) *$A \cap U, A \cap W \in \{\emptyset, A\}$ for each $A \in \mathcal{A}$ and for each $D = (U, W; F) \in \mathcal{D}$.*

Those G, $\mathcal{D}$ and $\mathcal{A}$ satisfying all conditions except for (ii) *and the last part of* (iv) *will make up the triples $(G, \mathcal{D}, \mathcal{A})$ of the class $\bar{\mathcal{G}}(n, k, \Omega, \rho, \nu)$.*

5

4 Augmenting a matching

In previous papers [AKS95, Zha11, PS12, Coo09, HP15] concerning the LKS Conjecture in the dense setting the crucial turn was to find a matching in the cluster graph of the host graph possessing certain properties. We will prove a similar "structural result" in Section 5. In the present section, we prove the main tool for Section 5, namely Lemma 4.8. All statements preceding Lemma 4.8 are only preparatory. The only exception is (the easy) Lemma 4.4 which is recycled later, in [HKP+c].

4.1 Regularized matchings

We prove our first auxiliary lemma on our way towards Lemma 4.8.

Lemma 4.1. *For every $\Omega \in \mathbb{N}$ and $\varepsilon, \rho, \tau > 0$ there is a number $\alpha > 0$ such that for every $\nu \in (0, 1)$ there exists a number $k_0 \in \mathbb{N}$ such that for each $k > k_0$ the following holds.*

For every $(G, \mathcal{D}, H, \mathcal{A}) \in \mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$ there are $(U, W; F) \in \mathcal{D}$, $A \in \mathcal{A}$ and $X, Y \subseteq V(G)$ such that

1) $|X| = |Y| \geqslant \alpha \nu k$,

2) $X \subseteq A \cap U \cap A_H$ and $Y \subseteq W \cap B_H$, where A_H and B_H are the colour classes of H, and

3) (X, Y) is an ε-regular pair in G of density $\mathrm{d}(X, Y) \geqslant \frac{\tau\rho}{4\Omega}$.

Proof. Let Ω, ε, ρ and τ be given. Applying Lemma 2.2 to $\varepsilon_{\mathrm{L2.2}} := \min\{\varepsilon, \frac{\rho^2}{8\Omega}\}$ and $\ell_{\mathrm{L2.2}} := 2$, we obtain numbers n_0 and M. We set

$$\alpha := \frac{\tau\rho}{\Omega^2 M}, \tag{4.1}$$

and given $\nu \in (0, 1)$, we set

$$k_0 := \frac{2n_0}{\alpha \nu M}.$$

Now suppose we are given $k > k_0$ and $(G, \mathcal{D}, H, \mathcal{A}) \in \mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$.

Property (i) of Definition 3.7 gives that $e(G) \leqslant \Omega kn/2$, and Property (ii) says that $e(H) \geqslant \tau kn$. So $e(H)/e(G) \geqslant 2\tau/\Omega$. Averaging over all dense spots $\mathcal{D}$ in the dense cover of G we find a dense spot $D = (U, W; F) \in \mathcal{D}$ such that

$$e_D(A_H, B_H) = |F \cap E(H)| \geqslant \frac{e(H)}{e(G)}|F| \geqslant \frac{2\tau|F|}{\Omega}. \tag{4.2}$$

Without loss of generality, we assume that

$$e_D(U \cap A_H, W \cap B_H) \geqslant \frac{1}{2} \cdot e_D(A_H, B_H) \geqslant e_D(U \cap B_H, W \cap A_H), \tag{4.3}$$

as otherwise one can just interchange the roles of U and W. Then,

$$e_G(U \cap A_H, W \cap B_H) \overset{(4.3)}{\geqslant} \frac{1}{2} \cdot e_D(A_H, B_H) \overset{(4.2)}{\geqslant} \frac{\tau}{\Omega} \cdot |F|. \tag{4.4}$$

Let $\mathcal{A}' \subseteq \mathcal{A}$ denote the family of those $A \in \mathcal{A}$ with $0 < e_G(A \cap U \cap A_H, W \cap B_H) < \frac{\tau}{\Omega} \cdot |F| \cdot \frac{|A|}{|U|}$. Note that for each $A \in \mathcal{A}'$ we have $A \subseteq U$ by Definition 3.7 (v). Therefore,

$$e_G\left(\bigcup \mathcal{A}' \cap U \cap A_H, W \cap B_H\right) < \frac{\tau}{\Omega} \cdot |F| \cdot \frac{|\mathcal{A}'|}{|U|} \leqslant \frac{\tau}{\Omega} \cdot |F| \overset{(4.4)}{\leqslant} e_G(U \cap A_H, W \cap B_H).$$

As $\mathcal{A}$ covers A_H, G has an edge xy with $x \in U \cap A_H \cap A$ for some $A \in \mathcal{A} \setminus \mathcal{A}'$ and $y \in W \cap B_H$. Set $X' := A \cap U \cap A_H = A \cap A_H$ and $Y' := W \cap B_H$. Then directly from the definition of $\mathcal{A}'$ and since D is a $(\rho k, \rho)$-dense spot, we obtain that

$$\mathrm{d}_G(X', Y') = \frac{e_G(X', Y')}{|X'||Y'|} \geqslant \frac{\frac{\tau}{\Omega} \cdot |F| \cdot \frac{|A|}{|U|}}{|A||W|} > \frac{\tau \rho}{\Omega}. \tag{4.5}$$

Also, since $(U, W; F) \in \mathcal{D}$, we have

$$|F| \geqslant \rho k |U| . \tag{4.6}$$

This enables us to bound the size of X' as follows.

$$
\begin{aligned}
|X'| &\geqslant \frac{e_G(X', Y')}{\mathrm{maxdeg}(G)} \\
\text{(as } A \notin \mathcal{A}' \text{ and by D3.7(i))} \quad &\geqslant \frac{\frac{\tau}{\Omega} \cdot \frac{|F|}{|U|} \cdot |A|}{\Omega k} \\
\text{(by (4.6))} \quad &\geqslant \frac{\tau \cdot \rho k \cdot |A|}{\Omega^2 k} \\
&\geqslant \frac{\tau \rho \nu k}{\Omega^2} \\
&\overset{(4.1)}{=} \alpha \nu k M .
\end{aligned}
\tag{4.7}
$$

Similarly,

$$|Y'| \geqslant \alpha \nu k M . \tag{4.8}$$

Applying Lemma 2.2 to $G[X', Y']$ with prepartition $\{X', Y'\}$ we obtain a collection of sets $\mathcal{C} = \{C_i\}_{i=0}^p$, with $p < M$. By (4.7), and (4.8), we have that $|C_i| \geqslant \alpha \nu k$ for every $i \in [p]$. It is easy to deduce from (4.5) that there is at least one $\varepsilon_{\mathrm{L2.2}}$-regular (and thus ε-regular) pair (X, Y), $X, Y \in \mathcal{C} \setminus \{C_0\}$, $X \subseteq X'$, $Y \subseteq Y'$ with $\mathrm{d}(X, Y) \geqslant \frac{\tau \rho}{4\Omega}$. Indeed, it suffices to count the number of edges incident with C_0, lying in $\varepsilon_{\mathrm{L2.2}}$-irregular pairs or belonging to too sparse pairs. Their number is strictly smaller than

$$(\varepsilon_{\mathrm{L2.2}} + \varepsilon_{\mathrm{L2.2}} + \frac{\rho^2}{4\Omega})|X'||Y'| \leqslant \frac{\rho^2}{2\Omega}|X'||Y'| \overset{(4.5)}{\leqslant} e(X', Y'),$$

and thus not all edges between X' and Y'. This finishes the proof of Lemma 4.1. $\qquad\square$

Instead of just one pair (X, Y), as it is given by Lemma 4.1, we shall later need several disjoint pairs for embedding larger trees. For this purpose we introduce the following definition, generalizing the notion of a matching in the cluster graph in the traditional regularity setting.

Definition 4.2 ((ε, d, ℓ)-**regularized matching**)**.** *Suppose that $\ell \in \mathbb{N}$ and $d, \varepsilon > 0$. A collection $\mathcal{N}$ of ordered pairs (A, B) with $A, B \subseteq V(H)$ is called an (ε, d, ℓ)-regularized matching of a graph H if*

(i) $|A| = |B| \geqslant \ell$ for each $(A, B) \in \mathcal{N}$,

(ii) (A, B) *induces in H an ε-regular pair of density at least d, for each $(A, B) \in \mathcal{N}$, and*

(iii) all involved sets A and B are pairwise disjoint.

Sometimes, when the parameters do not matter (as for instance in Definition 4.5 below) we simply call it a regularized matching.

For a regularized matching $\mathcal{N}$, we shall write $\mathcal{V}_1(\mathcal{N}) := \{A \, : \, (A, B) \in \mathcal{N}\}$, $\mathcal{V}_2(\mathcal{N}) := \{B \, : \, (A, B) \in \mathcal{N}\}$ and $\mathcal{V}(\mathcal{N}) := \mathcal{V}_1(\mathcal{N}) \cup \mathcal{V}_2(\mathcal{N})$. Furthermore, we set $V_1(\mathcal{N}) := \bigcup \mathcal{V}_1(\mathcal{N})$, $V_2(\mathcal{N}) := \bigcup \mathcal{V}_2(\mathcal{N})$ and $V(\mathcal{N}) := V_1(\mathcal{N}) \cup V_2(\mathcal{N}) = \bigcup \mathcal{V}(\mathcal{N})$. As these definitions suggest, the orientations of the pairs $(A, B) \in \mathcal{N}$ are important. The sets A and B are called $\mathcal{N}$-*vertices* and the pair (A, B) is an $\mathcal{N}$-*edge*.

We say that a regularized matching $\mathcal{N}$ *absorbs* a regularized matching $\mathcal{M}$ if for every $(S, T) \in \mathcal{M}$ there exists $(X, Y) \in \mathcal{N}$ such that $S \subseteq X$ and $T \subseteq Y$. In the same way, we say that a family of dense spots $\mathcal{D}$ *absorbs* a regularized matching $\mathcal{M}$ if for every $(S, T) \in \mathcal{M}$ there exists $(U, W; F) \in \mathcal{D}$ such that $S \subseteq U$ and $T \subseteq W$.

We later need the following easy bound on the size of the elements of $\mathcal{V}(\mathcal{M})$.

Fact 4.3. *Suppose that $\mathcal{M}$ is an (ε, d, ℓ)-regularized matching in a graph H. Then $|C| \leqslant \frac{\mathrm{maxdeg}(H)}{d}$ for each $C \in \mathcal{V}(\mathcal{M})$.*

Proof. Let for example $(C, D) \in \mathcal{M}$. The maximum degree of H is at least as large as the average degree of the vertices in D, which is at least $d|C|$. $\qquad\qquad\square$

The second step towards Lemma 4.8 is Lemma 4.4. Whereas Lemma 4.1 gives one dense regular pair, in the same setting Lemma 4.4 provides us with a dense regularized matching.

Lemma 4.4. *For every $\Omega \in \mathbb{N}$ and $\rho, \varepsilon, \tau \in (0, 1)$ there exists $\alpha > 0$ such that for every $\nu \in (0, 1)$ there is a number $k_0 \in \mathbb{N}$ such that the following holds for every $k > k_0$.*

For each $(G, \mathcal{D}, H, \mathcal{A}) \in \mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$ there exists an $(\varepsilon, \frac{\tau\rho}{8\Omega}, \alpha\nu k)$-regularized matching $\mathcal{M}$ of G such that

(P1) *for each $(X, Y) \in \mathcal{M}$ there are $A \in \mathcal{A}$, and $D = (U, W; F) \in \mathcal{D}$ such that $X \subseteq U \cap A \cap A_H$ and $Y \subseteq W \cap B_H$, and*

(P2) $|V(\mathcal{M})| \geqslant \frac{\tau}{2\Omega} n$.

Proof. Let $\alpha := \alpha_{L4.1} > 0$ be given by Lemma 4.1 for the input parameters $\Omega_{L4.1} := \Omega$, $\varepsilon_{L4.1} := \varepsilon$, $\tau_{L4.1} := \tau/2$ and $\rho_{L4.1} := \rho$. For $\nu_{L4.1} := \nu$, Lemma 4.1 yields a number $k_0 \in \mathbb{N}$.

Now let $(G, \mathcal{D}, H, \mathcal{A}) \in \mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$. Let $\mathcal{M}$ be an inclusion-maximal $(\varepsilon, \frac{\tau\rho}{8\Omega}, \alpha\nu k)$-regularized matching with property **(P1)**. We claim that

$$e_G(A_H \setminus V_1(\mathcal{M}), B_H \setminus V_2(\mathcal{M})) < \frac{\tau}{2} kn. \tag{4.9}$$

Indeed, suppose otherwise. Then the bipartite subgraph H' of G induced by the sets $A_H \setminus V_1(\mathcal{M}) = A_H \setminus V(\mathcal{M})$ and $B_H \setminus V_2(\mathcal{M}) = B_H \setminus V(\mathcal{M})$ satisfies Property (ii) of Definition 3.7, with $\tau_{D3.7} := \tau/2$. So, we have that $(G, \mathcal{D}, H', \mathcal{A}) \in \mathcal{G}(n, k, \Omega, \rho, \nu, \tau/2)$.

Thus Lemma 4.1 for $(G, \mathcal{D}, H', \mathcal{A})$ yields a dense spot $D = (U, W; F) \in \mathcal{D}$ and a set $A \in \mathcal{A}$, together with two sets $X \subseteq U \cap A \cap (A_H \setminus V(\mathcal{M}))$, $Y \subseteq W \cap (B_H \setminus V(\mathcal{M}))$ such that $|X| = |Y| > \alpha_{\text{L4.1}} \nu k = \alpha \nu k$, and such that (X, Y) is $\varepsilon_{\text{L4.1}}$-regular and has density at least

$$\frac{\tau_{\text{L4.1}} \rho_{\text{L4.1}}}{4 \Omega_{\text{L4.1}}} = \frac{\tau \rho}{8\Omega}.$$

This contradicts the maximality of $\mathcal{M}$, proving (4.9).

In order to see (**P2**), it suffices to observe that by (4.9) and by Property (ii) of Definition 3.7, the set $V(\mathcal{M})$ is incident with at least $\tau k n - \frac{\tau}{2} k n = \frac{\tau}{2} k n$ edges. By Definition 3.7 (i), it follows that $|V(\mathcal{M})| \geqslant \frac{\tau}{2} k n \cdot \frac{1}{\Omega k} \geqslant \frac{\tau}{2\Omega} n$, as desired.   □

4.2   Augmenting paths for matchings

We now prove the main lemma of Section 4, namely Lemma 4.8. We will use an augmenting path technique for our regularized matchings, similar to the augmenting paths commonly used for traditional matching theorems. For this, we need the following definitions.

Definition 4.5 (Alternating path, augmenting path). *Suppose that $n, s \in \mathbb{N}$ and $\delta > 0$. Given an n-vertex graph G, and a regularized matching $\mathcal{M}$, we call a sequence $\mathcal{S} = (Y_0, \mathcal{A}_1, Y_1, \mathcal{A}_2, Y_2, \ldots, \mathcal{A}_h, Y_h)$ (where $h \geqslant 0$ is arbitrary) a (δ, s)-alternating path for $\mathcal{M}$ from Y_0 if for all $i \in [h]$ we have*

   (i) $\mathcal{A}_i \subseteq \mathcal{V}_1(\mathcal{M})$ and the sets $\mathcal{A}_i$ are pairwise disjoint,

   (ii) $Y_0 \subseteq V(G) \setminus V(\mathcal{M})$ and $Y_i = \bigcup_{(A,B) \in \mathcal{M}, A \in \mathcal{A}_i} B$,

   (iii) $|Y_{i-1}| \geqslant \delta n$, and

   (iv) $e(A, Y_{i-1}) \geqslant s \cdot |A|$, for each $A \in \mathcal{A}_i$.

If in addition there is a set $\mathcal{C}$ of disjoint subsets of $V(G) \setminus (Y_0 \cup V(\mathcal{M}))$ such that

   (v) $e(\bigcup \mathcal{C}, Y_h) \geqslant t \cdot n$,

then we say that $\mathcal{S}' = (Y_0, \mathcal{A}_1, Y_1, \mathcal{A}_2, Y_2, \ldots, \mathcal{A}_h, Y_h, \mathcal{C})$ is a (δ, s, t)-augmenting path for $\mathcal{M}$ from Y_0 to $\mathcal{C}$.

   The number h is called the length *of $\mathcal{S}$ (or of $\mathcal{S}'$).*

Next, we show that a regularized matching either has an augmenting path or admits a partition into two parts so that only few edges cross these parts in a certain way.

Lemma 4.6. *Given an n-vertex graph G with $\mathrm{maxdeg}(G) \leqslant \Omega k$, a number $\tau \in (0, 1)$, a regularized matching $\mathcal{M}$, a set $Y_0 \subseteq V(G) \setminus V(\mathcal{M})$, and a set $\mathcal{C}$ of disjoint subsets of $V(G) \setminus (V(\mathcal{M}) \cup Y_0)$, one of the following holds:*

(M1) *There is a regularized matching $\mathcal{M}'' \subseteq \mathcal{M}$ with*

$$e\left(\bigcup \mathcal{C} \cup V_1(\mathcal{M} \setminus \mathcal{M}''), Y_0 \cup V_2(\mathcal{M}'')\right) < \tau n k,$$

(M2) *$\mathcal{M}$ has a $(\frac{\tau}{2\Omega}, \frac{\tau^2}{8\Omega} k, \frac{\tau^2}{16\Omega} k)$-augmenting path of length at most $2\Omega/\tau$ from Y_0 to $\mathcal{C}$.*

9

Proof. If $|Y_0| \leqslant \frac{\tau}{2\Omega} n$ then **(M1)** is satisfied for $\mathcal{M}'' := \emptyset$. Let us therefore assume otherwise.

Choose a $(\frac{\tau}{2\Omega}, \frac{\tau^2}{8\Omega} k)$-alternating path $\mathcal{S} = (Y_0, \mathcal{A}_1, Y_1, \mathcal{A}_2, Y_2, \ldots, \mathcal{A}_h, Y_h)$ for $\mathcal{M}$ with $|\bigcup_{\ell=1}^{h} \mathcal{A}_\ell|$ maximal.

Now, let $\ell^* \in \{0, 1, \ldots, h\}$ be maximal with $|Y_{\ell^*}| \geqslant \frac{\tau}{2\Omega} n$. Then $\ell^* \in \{h, h-1\}$. Moreover, as $|Y_\ell| \geqslant \frac{\tau}{2\Omega} n$ for all $\ell \leqslant \ell^*$, we have that $(\ell^* + 1) \cdot \frac{\tau}{2\Omega} n \leqslant |\dot{\bigcup}_{\ell \leqslant \ell^*} Y_\ell| \leqslant n$ and thus

$$\ell^* + 1 \leqslant \frac{2\Omega}{\tau}. \tag{4.10}$$

Let $\mathcal{M}'' \subseteq \mathcal{M}$ consist of all $\mathcal{M}$-edges $(A, B) \in \mathcal{M}$ with $A \in \bigcup_{\ell \in [h]} \mathcal{A}_\ell$. Then, by the choice of $\mathcal{S}$,

$$e\left(V_1(\mathcal{M} \setminus \mathcal{M}''), \bigcup_{\ell=0}^{\ell^*} Y_\ell\right) = \sum_{\ell=0}^{\ell^*} e\left(V_1(\mathcal{M} \setminus \mathcal{M}''), Y_\ell\right)$$

$$< (\ell^* + 1) \cdot \frac{\tau^2}{8\Omega} k \cdot |V_1(\mathcal{M} \setminus \mathcal{M}'')| \overset{(4.10)}{\leqslant} \frac{\tau}{4} kn. \tag{4.11}$$

Furthermore, if $\ell^* = h - 1$ (that is, if $|Y_h| < \frac{\tau}{2\Omega} n$) then

$$e\left(V_1(\mathcal{M} \setminus \mathcal{M}'') \cup \bigcup \mathcal{C}, Y_h\right) < \frac{\tau}{2\Omega} n \cdot \max\deg(G) \leqslant \frac{\tau}{2\Omega} \Omega k n = \frac{\tau}{2} kn. \tag{4.12}$$

So, regardless of whether $h = \ell^*$ or $h = \ell^* + 1$, we get from (4.11) and (4.12) that

$$e\left(V_1(\mathcal{M} \setminus \mathcal{M}'') \cup \bigcup \mathcal{C}, Y_0 \cup V_2(\mathcal{M}'')\right) < \frac{3}{4} \tau kn + e\left(\bigcup \mathcal{C}, \bigcup_{\ell=0}^{\ell^*} Y_\ell\right).$$

Thus, if $e(\bigcup \mathcal{C}, \bigcup_{\ell=0}^{\ell^*} Y_\ell) \leqslant \frac{\tau}{4} kn$, we see that **(M1)** is satisfied for $\mathcal{M}''$. So, assume otherwise. Then, by (4.10), there is an index $j \in \{0, 1, \ldots, \ell^*\}$ for which

$$e\left(\bigcup \mathcal{C}, Y_j\right) > \frac{\tau^2}{16\Omega} kn,$$

and thus, $(Y_0, \mathcal{A}_1, Y_1, \mathcal{A}_2, Y_2, \ldots, \mathcal{A}_j, Y_j, \mathcal{C})$ is a $(\frac{\tau}{2\Omega}, \frac{\tau^2}{8\Omega} k, \frac{\tau^2}{16\Omega} k)$-augmenting path for $\mathcal{M}$. This shows **(M2)**. $\qquad \square$

The aim of this section is to find a regularized matching covering as many vertices from the graph as possible. This is done by iteratively improving a matching. Below, Lemma 4.7 provides with such an iterative step: given a regularized matching $\mathcal{M}$ we either find (II) a better regularized matching $\mathcal{M}'$, or there is (I) a natural barrier to finding such a matching. This barrier is a separation of the previous regularized matching into two blocks ($\mathcal{M}''$ and $\mathcal{M} \setminus \mathcal{M}''$) such that very few edges "cross" this separation. The absence of such a separation guarantees the existence of an augmenting path for $\mathcal{M}$, which can be used to find a better regularized matching. This matching $\mathcal{M}'$ has **(C1)** to improve $\mathcal{M}$ substantially and **(C2)** respect the structure of the graph and of $\mathcal{M}$.

Lemma 4.7. *For every $\Omega \in \mathbb{N}$ and $\tau \in (0, \frac{1}{2\Omega})$ there is a number $\tau' \in (0, \tau)$ such that for every $\rho \in (0, 1)$ there is a number $\alpha \in (0, \tau'/2)$ such that for every $\varepsilon \in (0, \alpha)$ there is a number $\pi > 0$ such that for every $\gamma > 0$ there is $k_0 \in \mathbb{N}$ such that the following holds for every $k > k_0$ and every $h \in (\gamma k, k/2)$.*

Let G be a graph of order n with $\mathrm{maxdeg}(G) \leqslant \Omega k$, with an (ε^3, ρ, h)-regularized matching $\mathcal{M}$ and with a $(\rho k, \rho)$-dense cover $\mathcal{D}$ that absorbs $\mathcal{M}$. Let $Y \subseteq V(G) \setminus V(\mathcal{M})$, and let $\mathcal{C}$ be an h-ensemble in G with $\mathcal{C} \cap (V(\mathcal{M}) \cup Y) = \emptyset$. Assume that $U \cap C \in \{\emptyset, C\}$ for each $D = (U, W; F) \in \mathcal{D}$ and each $C \in \mathcal{C} \cup \mathcal{V}_1(\mathcal{M})$.

Then one of the following holds.

(I) There is a regularized matching $\mathcal{M}'' \subseteq \mathcal{M}$ such that

$$e\left(\bigcup \mathcal{C} \cup V_1(\mathcal{M} \setminus \mathcal{M}''), Y \cup V_2(\mathcal{M}'') \right) < \tau n k.$$

(II) There is an $(\varepsilon, \alpha, \pi h)$-regularized matching $\mathcal{M}'$ such that

(C1) $|V(\mathcal{M}) \setminus V(\mathcal{M}')| \leqslant \varepsilon n$, and $|V(\mathcal{M}')| \geqslant |V(\mathcal{M})| + \frac{\tau'}{2} n$, and

(C2) for each $(T, Q) \in \mathcal{M}'$ there are sets $C_1 \in \mathcal{V}_1(\mathcal{M}) \cup \mathcal{C}$, $C_2 \in \mathcal{V}_2(\mathcal{M}) \cup \{Y\}$ and a dense spot $D = (U, W; F) \in \mathcal{D}$ such that $T \subseteq C_1 \cap U$ and $Q \subseteq C_2 \cap W$.

Proof. We divide the proof into five steps.

Step 1: Setting up the parameters. Suppose that Ω and τ are given. For $\ell = 0, 1, \ldots, \lceil 2\Omega/\tau \rceil$, we define the auxiliary parameters

$$\tau^{(\ell)} := \left(\frac{\tau^2}{32\Omega} \right)^{\lceil \frac{2\Omega}{\tau} \rceil - \ell + 2}, \tag{4.13}$$

and set

$$\tau' := \frac{\tau^{(0)}}{2\Omega}.$$

Given ρ, we define

$$\alpha := \frac{\tau' \rho}{16\Omega}.$$

Then, given ε, for $\ell = 0, 1, \ldots, \lceil 2\Omega/\tau \rceil$, we define the further auxiliary parameters

$$\mu^{(\ell)} := \alpha_{\mathrm{L4.4}}\left(\Omega, \rho, \varepsilon^3, \tau^{(\ell)} \right)$$

which are given by Lemma 4.4 for input parameters $\Omega_{\mathrm{L4.4}} := \Omega$, $\rho_{\mathrm{L4.4}} := \rho$, $\varepsilon_{\mathrm{L4.4}} := \varepsilon^3$, and $\tau_{\mathrm{L4.4}} := \tau^{(\ell)}$. Set

$$\pi := \frac{\varepsilon}{2} \cdot \min \left\{ \mu^{(\ell)} \ : \ \ell = 0, \ldots, \lceil 2\Omega/\tau \rceil \right\},$$

Given the next input parameter γ, Lemma 4.4 for parameters as above and the final input $\nu_{\mathrm{L4.4}} := \gamma$ yields $k_{0_{\mathrm{L4.4}}} =: k_0^{(\ell)}$, set

$$k_0 := \max \left\{ k_0^{(\ell)} \ : \ \ell = 0, \ldots, \lceil 2\Omega/\tau \rceil \right\}.$$

Step 2: Finding an augmenting path. We apply Lemma 4.6 to G, τ, $\mathcal{M}$, Y and $\mathcal{C}$. Since (**M1**) corresponds to (I), let us assume that the outcome of the lemma is (**M2**). Then there is a $(\frac{\tau}{2\Omega}, \frac{\tau^2}{8\Omega} k, \frac{\tau^2}{16\Omega} k)$-augmenting path $\mathcal{S}' = (Y_0, \mathcal{A}_1, Y_1, \mathcal{A}_2, Y_2, \ldots, \mathcal{A}_{j^*}, Y_{j^*}, \mathcal{C})$ for $\mathcal{M}$ starting from $Y_0 := Y$ such that $j^* \leqslant 2\Omega/\tau$.

Our aim is now to show that (II) holds.

Step 3: Creating parallel matchings. Inductively, for $\ell = j^*, j^* - 1, \ldots, 0$ we shall define auxiliary bipartite induced subgraphs $H^{(\ell)} \subseteq G$ with colour classes $P^{(\ell)}$ and Y_ℓ that satisfy

(a) $e(H^{(\ell)}) \geqslant \tau^{(\ell)} kn$,

and $(\varepsilon^3, 2\alpha, \mu^{(\ell)} h)$-regularized matchings $\mathcal{M}^{(\ell)}$ that satisfy

(b) $V_1(\mathcal{M}^{(\ell)}) \subseteq P^{(\ell)}$,

(c) for each $(A', B') \in \mathcal{M}^{(\ell)}$ there are a dense spot $(U, W; F) \in \mathcal{D}$ and a set $A \in \mathcal{V}_1(\mathcal{M})$ (or a set $A \in \mathcal{C}$ if $\ell = j^*$) such that $A' \subseteq U \cap A$ and $B' \subseteq W \cap Y_\ell$,

(d) $|V(\mathcal{M}^{(\ell)})| \geqslant \frac{\tau^{(\ell)}}{2\Omega} n$, and

(e) $|B \cap V_2(\mathcal{M}^{(\ell)})| = |A \cap P^{(\ell-1)}|$ for each edge $(A, B) \in \mathcal{M}$, if $\ell > 0$.

We take $H^{(j^*)}$ as the induced bipartite subgraph of G with colour classes $P^{(j^*)} := \bigcup \mathcal{C}$ and Y_{j^*}. Definition 4.5 (v) together with (4.13) ensures (a) for $\ell = j^*$. Now, for $\ell \leqslant j^*$, suppose $H^{(\ell)}$ is defined already. Further, if $\ell < j^*$ suppose also that $\mathcal{M}^{(\ell+1)}$ is defined already. We shall define $\mathcal{M}^{(\ell)}$, and, if $\ell > 0$, we shall also define $H^{(\ell-1)}$.

Observe that $(G, \mathcal{D}, H^{(\ell)}, \mathcal{A}_\ell) \in \mathcal{G}(n, k, \Omega, \rho, \frac{h}{k}, \tau^{(\ell)})$, because of (a) and the assumptions of the lemma. So, applying Lemma 4.4 to $(G, \mathcal{D}, H^{(\ell)}, \mathcal{A}_\ell)$ and noting that $\frac{\tau^{(\ell)} \rho}{8\Omega} \geqslant 2\alpha$ we obtain an $(\varepsilon^3, 2\alpha, \mu^{(\ell)} h)$-regularized matching $\mathcal{M}^{(\ell)}$ that satisfies conditions (b)–(d).

If $\ell > 0$, we define $H^{(\ell-1)}$ as follows. For each $(A, B) \in \mathcal{M}$ take a set $\tilde{A} \subseteq A$ of cardinality $|\tilde{A}| = |B \cap V(\mathcal{M}^{(\ell)})|$ so that

$$e(\tilde{A}, Y_{\ell-1}) \geqslant \frac{\tau^2}{8\Omega} k \cdot |\tilde{A}| . \tag{4.14}$$

This is possible by Definition 4.5 (iv): just choose those vertices from A for $\tilde{A}$ that send most edges to $Y_{\ell-1}$. Let $P^{(\ell-1)}$ be the union of all the sets $\tilde{A}$. Then, (e) is satisfied. Furthermore,

$$|P^{(\ell-1)}| = |V_2(\mathcal{M}^{(\ell)})| \overset{(d)}{\geqslant} \frac{\tau^{(\ell)}}{4\Omega} n.$$

So, by (4.14),

$$e(P^{(\ell-1)}, Y_{\ell-1}) \geqslant \frac{\tau^2}{8\Omega} k \cdot |P^{(\ell-1)}| \geqslant \frac{\tau^2 \cdot \tau^{(\ell)}}{32\Omega^2} kn \overset{(4.13)}{=} \tau^{(\ell-1)} kn . \tag{4.15}$$

We let $H^{(\ell-1)}$ be the bipartite subgraph of G induced by the colour classes $P^{(\ell-1)}$ and $Y_{\ell-1}$. Then (4.15) establishes (a) for $H^{(\ell-1)}$. This finishes step ℓ.[1]

Step 4: Harmonising the matchings. Our regularized matchings $\mathcal{M}^{(0)}, \ldots, \mathcal{M}^{(j^*)}$ will be a good base for constructing the regularized matching $\mathcal{M}'$ we are after. However, we do not know anything about $|B \cap V_2(\mathcal{M}^{(\ell)})| - |A \cap V_1(\mathcal{M}^{(\ell-1)})|$ for the $\mathcal{M}$-edges $(A, B) \in \mathcal{M}$. But this term will be crucial in determining how much of $V(\mathcal{M})$ gets lost when we replace some of its $\mathcal{M}$-edges with $\bigcup \mathcal{M}^{(\ell)}$-edges. For this reason, we refine $\mathcal{M}^{(\ell)}$ in a way that its $\mathcal{M}^{(\ell)}$-edges become almost equal-sized.

Formally, we shall inductively construct regularized matchings $\mathcal{N}^{(0)}, \ldots, \mathcal{N}^{(j^*)}$ such that for $\ell = 0, \ldots, j^*$ we have

---

[1] Recall that the matching $\mathcal{M}^{(\ell-1)}$ is only to be defined in step $\ell - 1$.

(A) $\mathcal{N}^{(\ell)}$ is an $(\varepsilon, \alpha, \pi h)$-regularized matching,

(B) $\mathcal{M}^{(\ell)}$ absorbs $\mathcal{N}^{(\ell)}$,

(C) if $\ell > 0$ and $(A, B) \in \mathcal{M}$ with $A \in \mathcal{A}_\ell$ then $|A \cap V(\mathcal{N}^{(\ell-1)})| \geqslant |B \cap V(\mathcal{N}^{(\ell)})|$, and

(D) $|V_2(\mathcal{N}^{(\ell)})| \geqslant |V_1(\mathcal{N}^{(\ell-1)})| - \frac{\varepsilon}{2} \cdot |V_2(\mathcal{M}^{(\ell)})|$ if $\ell > 0$ and $|V_2(\mathcal{N}^{(0)})| \geqslant \frac{\tau^{(0)}}{2\Omega} n = \tau' n$.

Set $\mathcal{N}^{(0)} := \mathcal{M}^{(0)}$. Clearly (B) holds for $\ell = 0$, (A) is easy to check, and (C) is void. Finally, Property (D) holds because of (d). Suppose now $\ell > 0$ and that we already constructed matchings $\mathcal{N}^{(0)}, \ldots, \mathcal{N}^{(\ell-1)}$ satisfying Conditions (A)–(D).
Observe that for any $(A, B) \in \mathcal{M}$ we have that

$$|B \cap V_2(\mathcal{M}^{(\ell)})| \overset{\text{(b),(e)}}{\geqslant} |A \cap V_1(\mathcal{M}^{(\ell-1)})| \geqslant |A \cap V_1(\mathcal{N}^{(\ell-1)})|, \qquad (4.16)$$

where the last inequality holds because of (B) for $\ell - 1$.

So, we can choose a subset $X^{(\ell)} \subseteq V_2(\mathcal{M}^{(\ell)})$ such that $|B \cap X^{(\ell)}| = |A \cap V(\mathcal{N}^{(\ell-1)})|$ for each $(A, B) \in \mathcal{M}$. Now, for each $(S, T) \in \mathcal{M}^{(\ell)}$ write $\widehat{T} := T \cap X^{(\ell)}$, and choose a subset $\widehat{S}$ of S of size $|\widehat{T}|$. Set

$$\mathcal{N}^{(\ell)} := \left\{ (\widehat{S}, \widehat{T}) \, : \, (S, T) \in \mathcal{M}^{(\ell)}, |\widehat{T}| \geqslant \frac{\varepsilon}{2} \cdot |T| \right\}. \qquad (4.17)$$

Then (B) and (C) hold for ℓ.

For (A), note that Fact 2.1 implies that $\mathcal{N}^{(\ell)}$ is an $(\varepsilon, 2\alpha - \varepsilon^3, \frac{\varepsilon}{2} \mu^{(\ell)} h)$-regularized matching. In order to see (D), it suffices to observe that

$$|V_2(\mathcal{N}^{(\ell)})| = \sum_{(\widehat{S}, \widehat{T}) \in \mathcal{N}^{(\ell)}} |\widehat{T}| \geqslant |X^{(\ell)}| - \sum_{(S,T) \in \mathcal{M}^{(\ell)}} \frac{\varepsilon}{2} \cdot |T|$$

$$\geqslant \sum_{(A,B) \in \mathcal{M}} |A \cap V_1(\mathcal{N}^{(\ell-1)})| - \frac{\varepsilon}{2} \cdot |V_2(\mathcal{M}^{(\ell)})| = |V_1(\mathcal{N}^{(\ell-1)})| - \frac{\varepsilon}{2} \cdot |V_2(\mathcal{M}^{(\ell)})|.$$

Step 5: The final matching. Suppose that $(A, B) \in \mathcal{M}$ with $A \in \mathcal{A}_\ell$ for some $\ell \in \{1, 2, \ldots, j^*\}$. Then, set $A' := A \setminus V_1(\mathcal{N}^{(\ell-1)})$. Also, choose a set $B' \subseteq B \setminus V_2(\mathcal{N}^{(\ell)})$ of cardinality $|A'|$. This is possible by (C). By (4.17) we deduce that

$$|B \setminus V_2(\mathcal{N}^{(\ell)})| - |B'| \leqslant \frac{\varepsilon}{2} |B| . \qquad (4.18)$$

We consider the set $\mathcal{L} \subseteq \mathcal{M}$ consisting of all $\mathcal{M}$-edges $(A, B) \in \mathcal{M}$ with $|A'| > \frac{\varepsilon}{2} \cdot |A|$.
Set

$$\mathcal{K} := \{(A', B') : (A, B) \in \mathcal{L}\}.$$

By the assumption of the lemma, for every $(A', B') \in \mathcal{K}$ there are an edge $(A, B) \in \mathcal{M}$ and a dense spot $D = (U, W; F) \in \mathcal{D}$ such that

$$A' \subseteq A \subseteq U \text{ and } B' \subseteq B \subseteq W. \qquad (4.19)$$

Since $\mathcal{M}$ is (ε^3, ρ, h)-regularized , Fact 2.1 implies that $\mathcal{K}$ is a $(\varepsilon, \rho - \varepsilon^3, \frac{\varepsilon}{2}h)$-regularized matching. Set

$$\mathcal{M}' := \mathcal{K} \cup \bigcup_{\ell=0}^{j^*} \mathcal{N}^{(\ell)}.$$

It is easy to check that $\mathcal{M}'$ is an $(\varepsilon, \alpha, \pi h)$-regularized matching. Using (4.19) together with (B) and (c), we see that **(C2)** holds for $\mathcal{M}'$.

In order to see **(C1)**, we calculate

$$|V(\mathcal{M}) \setminus V(\mathcal{M}')| = \sum_{(A,B) \in \mathcal{M}} \left(|A \setminus V_1(\cup_{\ell=0}^{j^*} \mathcal{N}^{(\ell)} \cup \mathcal{K})| + |B \setminus V_2(\cup_{\ell=0}^{j^*} \mathcal{N}^{(\ell)} \cup \mathcal{K})| \right)$$

$$\overset{(4.18)}{\leqslant} \sum_{(A,B) \in \mathcal{M} \setminus \mathcal{L}} \left(|A' \cup B'| + \frac{\varepsilon}{2}|B| \right)$$

$$+ \sum_{(A,B) \in \mathcal{L}} \left(|A \setminus V_1(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell-1)} \cup \mathcal{K}))| + |B \setminus V_2(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell)} \cup \mathcal{K})| \right).$$

In the second sum, consider an arbitrary term corresponding to (A, B). By the definition of $\mathcal{K}$, the term $|A \setminus V_1(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell-1)} \cup \mathcal{K})|$ is zero. To treat the term $|B \setminus V_2(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell)} \cup \mathcal{K})|$, we recall that $|A| = |B|$ and $|A'| = |B'|$ (in the definition of $\mathcal{K}$). This gives that $|B \setminus V_2(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell)} \cup \mathcal{K})| = |A \cap V_1(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell-1)})| - |B \cap V_2(\cup_{\ell=1}^{j^*} \mathcal{N}^{(\ell)})|$. This leads to

$$|V(\mathcal{M}) \setminus V(\mathcal{M}')| \leqslant \sum_{(A,B) \in \mathcal{M} \setminus \mathcal{L}} \left(|A' \cup B'| + \frac{\varepsilon}{2}|B| \right)$$

$$+ \sum_{(A,B) \in \mathcal{L}} \sum_{\ell=1}^{j^*} \left(|A \cap V_1(\mathcal{N}^{(\ell-1)})| - |B \cap V_2(\mathcal{N}^{(\ell)})| \right)$$

$$\leqslant \sum_{(A,B) \in \mathcal{M} \setminus \mathcal{L}} \left(\frac{\varepsilon}{2}|A| + \varepsilon|B| \right) + \sum_{\ell=1}^{j^*} \left(|V_1(\mathcal{N}^{(\ell-1)})| - |V_2(\mathcal{N}^{(\ell)})| \right)$$

$$\overset{(D)}{\leqslant} \frac{3\varepsilon}{4}n + \sum_{\ell=1}^{j^*} \frac{\varepsilon}{2} \cdot |V_2(\mathcal{M}^{(\ell)})| \leqslant \varepsilon n .$$

Using the fact that $V_2(\mathcal{N}^{(0)}) \subseteq V(\mathcal{M}') \setminus V(\mathcal{M})$ the last calculation also implies that

$$|V(\mathcal{M}')| - |V(\mathcal{M})| \geqslant |V_2(\mathcal{N}^{(0)})| - |V(\mathcal{M}) \setminus V(\mathcal{M}')| \overset{(D)}{\geqslant} \tau'n - \varepsilon n > \frac{\tau'}{2}n ,$$

since $\varepsilon < \alpha \leqslant \tau'/2$ by assumption. $\qquad\square$

Iterating Lemma 4.7 we prove the main result of the section.

Lemma 4.8. *For every $\Omega \in \mathbb{N}$ and $\rho \in (0, 1/\Omega)$ there exists a number $\beta > 0$ such that for every $\varepsilon \in (0, \beta)$, there are $\varepsilon', \pi > 0$ such that for each $\gamma > 0$ there exists $k_0 \in \mathbb{N}$ such that the following holds for every $k > k_0$ and $c \in (\gamma k, k/2)$.*

Let G be a graph of order n, with $\mathrm{maxdeg}(G) \leqslant \Omega k$. Let $\mathcal{D}$ be a $(\rho k, \rho)$-dense cover of G, and let $\mathcal{M}$ be an (ε', ρ, c)-regularized matching that is absorbed by $\mathcal{D}$. Let $\mathcal{C}$ be a c-ensemble in G with $\mathcal{C} \cap (V(\mathcal{M})) = \emptyset$. Let $Y \subseteq V(G) \setminus (V(\mathcal{M}) \cup \bigcup \mathcal{C})$. Assume that for each $(U, W; F) \in \mathcal{D}$, and for each $C \in \mathcal{V}_1(\mathcal{M}) \cup \mathcal{C}$ we have that

$$U \cap C \in \{\emptyset, C\}. \tag{4.20}$$

Then there exists an $(\varepsilon, \beta, \pi c)$-regularized matching $\mathcal{M}'$ such that

(i) $|V(\mathcal{M}) \setminus V(\mathcal{M}')| \leqslant \varepsilon n$,

(ii) for each $(T, Q) \in \mathcal{M}'$ there are sets $C_1 \in \mathcal{V}_1(\mathcal{M}) \cup \mathcal{C}$, $C_2 \in \mathcal{V}_2(\mathcal{M}) \cup \{Y\}$ and a dense spot $D = (U, W; F) \in \mathcal{D}$ such that $T \subseteq C_1 \cap U$ and $Q \subseteq C_2 \cap W$, and

(iii) $\mathcal{M}'$ can be partitioned into $\mathcal{M}_1$ and $\mathcal{M}_2$ so that

$$e\left(\left(\bigcup \mathcal{C} \cup V_1(\mathcal{M})\right) \setminus V_1(\mathcal{M}_1), \ (Y \cup V_2(\mathcal{M})) \setminus V_2(\mathcal{M}_2)\right) < \rho k n.$$

Proof. Let Ω and ρ be given. Let $\tau' := \tau'_{L4.7}$ be the output given by Lemma 4.7 for input parameters $\Omega_{L4.7} := \Omega$ and $\tau_{L4.7} := \rho/2$.

Set $\rho^{(0)} := \rho$, set $L := \lceil 2/\tau' \rceil + 1$, and for $\ell \in [L]$, inductively define $\rho^{(\ell)}$ to be the output $\alpha_{L4.7}$ given by Lemma 4.7 for the further input parameter $\rho_{L4.7} := \rho^{(\ell-1)}$ (keeping $\Omega_{L4.7} = \Omega$ and $\tau_{L4.7} = \rho/2$ fixed). Then $\rho^{(\ell+1)} \leqslant \rho^{(\ell)}$ for all ℓ. Set $\beta := \rho^{(L)}$.

Given $\varepsilon < \beta$ we set $\varepsilon^{(\ell)} := (\varepsilon/2)^{3^{L-\ell}}$ for $\ell \in [L] \cup \{0\}$, and set $\varepsilon' := \varepsilon^{(0)}$. Clearly,

$$\sum_{\ell=0}^{L} \varepsilon^{(\ell)} \leqslant \varepsilon. \tag{4.21}$$

Now, for $\ell + 1 \in [L]$, let $\pi^{(\ell)} := \pi_{L4.7}$ be given by Lemma 4.7 for input parameters $\Omega_{L4.7} := \Omega$, $\tau_{L4.7} := \rho/2$, $\rho_{L4.7} := \rho^{(\ell)}$ and $\varepsilon_{L4.7} := \varepsilon^{(\ell+1)}$. For $\ell \in [L] \cup \{0\}$, set $\Pi^{(\ell)} := \frac{\rho}{2\Omega} \prod_{j=0}^{\ell-1} \pi^{(j)}$. Let $\pi := \Pi^{(L)}$.

Given γ, let k_0 be the maximum of the lower bounds $k_{0_{L4.7}}$ given by Lemma 4.7 for input parameters $\Omega_{L4.7} := \Omega$, $\tau_{L4.7} := \rho/2$, $\rho_{L4.7} := \rho^{(\ell-1)}$, $\varepsilon_{L4.7} := \varepsilon^{(\ell)}$, $\gamma_{L4.7} := \gamma \Pi^{(\ell)}$, for $\ell \in [L]$.

Suppose now we are given G, $\mathcal{D}$, $\mathcal{C}$, Y and $\mathcal{M}$. Suppose further that $c > \gamma k > \gamma k_0$. Let $\ell \in \{0, 1, \ldots, L\}$ be maximal subject to the condition that there is a matching $\mathcal{M}^{(\ell)}$ with the following properties:

(a) $\mathcal{M}^{(\ell)}$ is an $(\varepsilon^{(\ell)}, \rho^{(\ell)}, \Pi^{(\ell)}c)$-regularized matching,

(b) $|V(\mathcal{M}^{(\ell)})| \geqslant \ell \cdot \frac{\tau'}{2} n$,

(c) $|V(\mathcal{M}) \setminus V(\mathcal{M}^{(\ell)})| \leqslant \sum_{i=0}^{\ell} \varepsilon^{(i)} n$, and

(d) for each $(T, Q) \in \mathcal{M}^{(\ell)}$ there are sets $C_1 \in \mathcal{V}_1(\mathcal{M}) \cup \mathcal{C}$, $C_2 \in \mathcal{V}_2(\mathcal{M}) \cup \{Y\}$ and a dense spot $D = (U, W; F) \in \mathcal{D}$ such that $T \subseteq C_1 \cap U$ and $Q \subseteq C_2 \cap W$.

15

Observe that such a number ℓ exists, as for $\ell = 0$ we may take $\mathcal{M}^{(0)} = \mathcal{M}$. Also note that $\ell \leqslant 2/\tau' < L$ because of (b).

We now apply Lemma 4.7 with input parameters $\Omega_{\text{L4.7}} := \Omega$, $\tau_{\text{L4.7}} := \rho/2$, $\rho_{\text{L4.7}} := \rho^{(\ell)}$, $\varepsilon_{\text{L4.7}} := \varepsilon^{(\ell+1)} < \beta \leqslant \rho^{(\ell+1)} = \alpha_{\text{L4.7}}$, $\gamma_{\text{L4.7}} := \gamma\Pi^{(\ell)}$ to the graph G with the $(\rho^{(\ell)}k, \rho^{(\ell)})$-dense cover $\mathcal{D}$, the $(\varepsilon^{(\ell)}, \rho^{(\ell)}, \Pi^{(\ell)}c)$-regularized matching $\mathcal{M}^{(\ell)}$, the set

$$\tilde{Y} := (Y \cup V_2(\mathcal{M})) \setminus V_2(\mathcal{M}^{(\ell)}),$$

and the $(\Pi^{(\ell)}c)$-ensemble

$$\tilde{\mathcal{C}} := \left\{ C \setminus V(\mathcal{M}^{(\ell)}) \ : \ C \in \mathcal{V}_1(\mathcal{M}) \cup \mathcal{C}, \ |C \setminus V_1(\mathcal{M}^{(\ell)})| \geqslant \Pi^{(\ell)}c \right\}.$$

Lemma 4.7 yields a regularized matching which either corresponds to $\mathcal{M}''_{\text{L4.7}}$ as in Assertion (I) or to $\mathcal{M}'_{\text{L4.7}}$ as in Assertion (II). Note that in the latter case, the matching $\mathcal{M}'_{\text{L4.7}}$ actually constitutes an $(\varepsilon^{(\ell+1)}, \rho^{(\ell+1)}, \Pi^{(\ell+1)}c)$-regularized matching $\mathcal{M}^{(\ell+1)}$ fulfilling all the above properties for $\ell+1 \leqslant L$. In fact, (b) and (c) hold for $\mathcal{M}^{(\ell+1)}$ because of (**C1**), and it is not difficult to deduce (d) from (**C2**) and from (d) for ℓ. But this contradicts the choice of ℓ. We conclude that we obtained a regularized matching $\mathcal{M}''_{\text{L4.7}} \subseteq \mathcal{M}^{(\ell)}$ as in Assertion (I) of Lemma 4.7.

Thus, in other words, $\mathcal{M}^{(\ell)}$ can be partitioned into $\mathcal{M}_1$ and $\mathcal{M}_2$ so that

$$e\left(\bigcup \tilde{\mathcal{C}} \cup V_1(\mathcal{M}_2) \ , \ \tilde{Y} \cup V_2(\mathcal{M}_1) \right) \ < \ \tau_{\text{L4.7}}kn \ = \ \rho kn/2. \tag{4.22}$$

Set $\mathcal{M}' := \mathcal{M}^{(\ell)}$. Then $\mathcal{M}'$ is $(\varepsilon, \beta, \pi c)$-regularized by (a). Note that Assertion (i) of the lemma holds by (4.21) and by (c). Assertion (ii) holds because of (d).

Since

$$(Y \cup V_2(\mathcal{M})) \setminus V_2(\mathcal{M}_2) \ \subseteq \ \tilde{Y} \cup V_2(\mathcal{M}_1),$$

and because of (4.22) we know that in order to prove Assertion (iii) it suffices to show that

$$\begin{aligned} X \ &:= \ \left(\left(\bigcup \mathcal{C} \cup V_1(\mathcal{M}) \right) \setminus V_1(\mathcal{M}_1) \right) \setminus \left(\bigcup \tilde{\mathcal{C}} \cup V_1(\mathcal{M}_2) \right) \\ &= \ \left(\bigcup \mathcal{C} \cup V_1(\mathcal{M}) \right) \setminus \left(\bigcup \tilde{\mathcal{C}} \cup V_1(\mathcal{M}^{(\ell)}) \right) \end{aligned}$$

sends at most $\rho kn/2$ edges to the rest of the graph. For this, it would be enough to see that $|X| \leqslant \frac{\rho}{2\Omega}n$, since by assumption, G has maximum degree Ωk.

To this end, note that by assumption, $|\mathcal{V}_1(\mathcal{M}) \cup \mathcal{C}| \leqslant \frac{n}{c}$. Further, the definition of $\tilde{\mathcal{C}}$ implies that for each $A \in \mathcal{C} \cup \mathcal{V}_1(\mathcal{M})$ we have that $|A \setminus \left(\bigcup \tilde{\mathcal{C}} \cup V_1(\mathcal{M}^{(\ell)}) \right)| \leqslant \Pi^{(\ell)}c$. Combining these two observations, we obtain that

$$|X| < \Pi^{(\ell)}n \leqslant \frac{\rho}{2\Omega}n \,,$$

as desired.

$\square$

5 Rough structure of LKS graphs

In this section we give a structural result for graphs $G \in \textbf{LKSsmall}(n, k, \eta)$, stated in Lemma 5.4. Similar structural results were essential also for proving Conjecture 1.1 in the dense setting in [AKS95,

PS12]. There, a certain matching structure was proved to exist in the cluster graph of the host graph. This matching structure then allowed us to embed a given tree into the host graph. We motivate the structure asserted by Lemma 5.4 in more detail in Section 5.1.

Naturally, in our possibly sparse setting the sparse decomposition ∇ of G will enter the picture (instead of just the cluster graph of G. For more on sparse decomposition, see [HKP$^+$a]). There is an important subtlety though: we may need to "re-regularize" the cluster graph $\mathbf{G}_{\mathrm{reg}}$ of ∇. In this case, we have to find another regularization of parts of G, partially based on G_{reg}. Lemma 4.8 is the main tool to this end. The re-regularization is captured by the regularized matchings $\mathcal{M}_A$ and $\mathcal{M}_B$.

Let us note that this step is one of the biggest differences between our approach and the announced solution of the Erdős–Sós Conjecture by Ajtai, Komlós, Simonovits and Szemerédi. In other words, the nature of the graphs arising in the Erdős–Sós Conjecture allows a less careful approach with respect to regularization, still yielding a structure suitable for embedding trees. We discuss the necessity of this step in further detail in Section 5.2. The main result of this paper Lemma 5.4, is given in Section 5.3.

5.1   Motivation for and intuition behind Lemma 5.4

Recall that [HKP$^+$a, Lemma 3.14] asserts that each graph $G = G_{\mathrm{T1.2}}$ satisfying the conditions of Theorem 1.2 has a sparse decomposition which captures almost all its edges. With this pre-processing at hand, we want Lemma 5.4 to provide specific structural properties of G under which we could make the embedding of the tree $T = T_{\mathrm{T1.2}}$ work. The complexity of these assertions (which span more than half a page) stems from the complicated nature of the sparse decomposition, and from the delicate features of the embedding techniques (worked out in [HKP$^+$d, Section 6]). In this section we try to explain and motivate the key assertions of Lemma 5.4. The reader may skip the section at his or her convenience. The only bit from this section that is needed for the main result is Definition 5.3.

At this stage, let us introduce informally the notion of fine partition which we use to cut up the tree T. Let $\tau \ll 1$. We find a constant number of *cut vertices* of T so that the components (which we refer to as *shrubs*) in the remainder of T are of order at most τk. The cut vertices will decompose into two sets W_A and W_B so that distance from any vertex of W_A to any vertex in W_B is odd. It can be shown that we can do the cutting so that each shrub either neighbours only one cut vertex from $W_A \cup W_B$, or it neighbours two, in which case both these cut vertices are in W_A. Thus, the set of all shrubs can be decomposed as $\mathcal{S}_A \dot{\cup} \mathcal{S}_B$ depending on the cut vertices that surround individual shrubs. The last property of the fine partition we shall use is that

$$\sum_{t \in \mathcal{S}_A} v(t) \geqslant \sum_{t \in \mathcal{S}_B} v(t) \, . \tag{5.1}$$

The quadruple $(W_A, W_B, \mathcal{S}_A, \mathcal{S}_B)$ is then called a (τk)-*fine partition* of T. The full definition which includes several additional properties is given in [HKP$^+$d, Section 3].

As said earlier, Lemma 5.4 is an extensive generalization of previous structural results on the LKS Conjecture in the dense setting. So, as a starting point for our motivation let us explain the structural result Piguet and Stein [PS12] use to prove the dense approximate version of the LKS Conjecture.

Theorem 5.1 ([PS12])**.** *For every $\eta > 0$ and $q > 0$ there exists n_0 such that for every $n > n_0$ and $k > qn$ we have the following. For every graph $G \in \mathbf{LKS}(n, k, \eta)$ we have $\mathbf{trees}(k) \subseteq G$.*

Here, of course, the structure we work with is encoded in the cluster graph (in the sense of the original regularity lemma) $\mathbf{G}_{\mathrm{reg}}$ of the graph $G_{\mathrm{T5.1}}$. Note that $\mathbf{G}_{\mathrm{reg}} \in \mathbf{LKS}(N, K, \eta/2)$, where N is the number of clusters and $K = k \cdot \frac{N}{n}$. The main structural result of Piguet and Stein then reads as follows.

Informal Lemma 5.2 ([PS12, Lemma 8], simplified)**.** *Suppose that* $\mathbf{G}_{\mathrm{reg}} \in \mathbf{LKS}(N, K, \alpha)$ *and let us write* $\mathbf{L} = \mathbb{L}_{K,\alpha}(\mathbf{G}_{\mathrm{reg}})$*. Then we have one of the following.*

(H1) *There exists a matching* $M \subseteq \mathbf{G}_{\mathrm{reg}}$ *and an edge* $A_1 A_2$ *so that* $\deg_{\mathbf{G}_{\mathrm{reg}}}(A_i, \mathbf{L} \cup V(M)) \geqslant K$, *for* $i = 1, 2$.

(H2) *There exists a matching* $M \subseteq \mathbf{G}_{\mathrm{reg}}$ *and an edge* AB *with* $\deg_{\mathbf{G}_{\mathrm{reg}}}(A, \mathbf{L} \cup V(M)) \geqslant K$, *and* $\deg_{\mathbf{G}_{\mathrm{reg}}}(B, \mathbf{L} \cup V(M)) \geqslant \frac{K}{2}$*. Further,*

$$\text{for every } e \in M, \quad |\mathrm{N}_{\mathbf{G}_{\mathrm{reg}}}(A) \cap e| \leqslant 1 . \tag{5.2}$$

Piguet and Stein use structures **(H1)** and **(H2)** to embed any given tree $T \in \mathbf{trees}(k)$ into G using the regularity method. A comprehensive description of the embedding procedure is given in Sections 3.6 and 3.7 in [PS12]. The embedding itself is quite technical but it follows a relatively pedestrian strategy which we present next. The regularity method tells us that a regular pair can be filled-up by an arbitrary family of shrubs, provided that the colour classes of these shrubs (viewed as one bipartite graph) do not overfill the end-clusters of that regular pair. The degree conditions in Informal Lemma 5.2 suggest that we will to utilize the clusters of M and of $\mathbf{L}$. More precisely, some of the shrubs will be accommodated in the edges of the matching M. Suppose next that we would like to proceed with embedding some shrubs using a cluster $X \in \mathbf{L}$. This can be done as follows. Using the high-degree property of X we can find a cluster Y adjacent to X that is not filled up completely by the image of T. We then use the pair XY to accommodate further shrubs. We keep embedding T by mapping W_A to A_1 (in **(H1)**) or to A (in **(H2)**), W_B to A_2 or to B, and the shrubs pendent from these cut-vertices either into the regular edges of M, or to edges incident to clusters $\mathbf{L}$ as described above. Thus, the degree conditions in Informal Lemma 5.2 guarantee that we can accommodate shrubs of total order up to k from A_1, A_2, and A. The degree bound for B guarantees that we can embed shrubs of total order up to $k/2$ from B, recall that this is sufficient thanks to (5.1). The fact that $A_1 A_2$ or AB forms an edge allows us to make connections between images of W_A and W_B.

So far, we have not explained the role of condition (5.2). We include a relatively detailed explanation in Figure 5.1. However, this condition is independent of the rest of the argument, and it may be sufficient for the reader to take granted that (5.2) is crucial for the embedding to work.

We now try to give an analogue to Informal Lemma 5.2 in the sparse setting when the structure of G is encoded in the sparse decomposition rather than in the cluster graph. Recall that in the dense setting sets suitable for embedding shrubs were clusters of a regular matching (that is, M), and clusters of large degree (that is, $\mathbf{L}$). In the sparse setting, in addition to using a suitable matching of regular pairs $\mathcal{M}$ and large degree vertices $\mathbb{L}_{k,\eta}(G)$ we can make use of two further sets for embedding shrubs: $V(G_{\mathrm{exp}})$ (as explained in [HKP$^+$a, Section 3.6]) and the set $\mathbb{E}$ (as explained in [HKP$^+$a, Section 3.5]). Thus, the counterpart of clusters A_1, A_2 and A from Informal Lemma 5.2 is the set $\mathbb{XA}$ of vertices, which have degree at least k into the set $\mathbb{L}_{k,\alpha}(G) \cup V(\mathcal{M}) \cup V(G_{\mathrm{exp}}) \cup \mathbb{E}$.[2]

---

[2]The rather different looking formal definition of $\mathbb{XA}$ is given in (5.4). Below, we give an explanation for this difference.
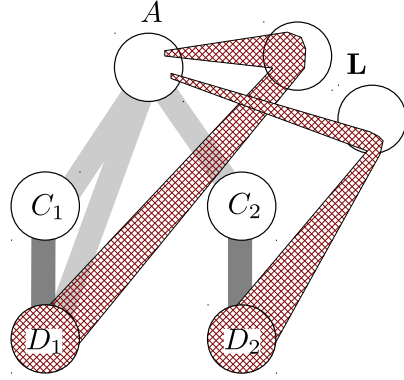
Partial image of T shown hatched.

The B-part of the embedding not shown.

Figure 5.1: The reason for requiring (5.2) in the setting of **(H2)**. Consider two edges $C_1D_1, C_2D_2 \in M$ such that only C_2D_2 satisfies (5.2). At some point during the embedding of T, we may need to use the high-degree property of clusters in **L**. When doing so we cannot guarantee that we will fill the edges of M in an efficient way. That is, we may end up filling D_1 and D_2 completely and leaving C_1 and C_2 untouched. If this happens, both regular pairs C_1D_1 and C_2D_2 are useless for embedding further shrubs. The used space in C_2D_2 equals the degree from A to C_2D_2. That is, we do not expect to embed anything more in the edge C_2D_2. The condition $\deg_{\mathbf{G}_{\mathrm{reg}}}(A, \mathbf{L} \cup V(M)) \geqslant K$ ensures we find free space somewhere else in the cluster graph to complete our embedding. Clearly, the pair C_1D_1 does not have this favourable feature: the number of vertices used by the embedding is only half the degree of A to C_1D_1. In this case, the condition $\deg_{\mathbf{G}_{\mathrm{reg}}}(A, \mathbf{L} \cup V(M)) \geqslant K$ is not strong enough.
We do not need a counterpart of (5.2) for $\mathrm{N}_{\mathbf{G}_{\mathrm{reg}}}(B)$. The reason is that we can schedule our embedding process in such a way that when we use the high-degree property of **L** we have already exhausted the degree from B to M.

Likewise, the counterpart of cluster B in Informal Lemma 5.2 are vertices of $\mathbb{XB}$, which have degree at least $k/2$ into $\mathbb{L}_{k,\alpha}(G) \cup V(\mathcal{M}) \cup V(G_{\mathrm{exp}}) \cup \mathbb{E}$.[3] We see that a sparse counterpart to **(H1)** would be two disjoint well-connected sets $\mathbb{XA}_1, \mathbb{XA}_2 \subseteq \mathbb{XA}$. In Lemma 5.4 we achieve this in one of two possible ways. One way is finding a large regularized matching $\mathcal{M}_{\mathrm{good}}$ inside $\mathbb{XA}$; one can then take $\mathbb{XA}_1 = V_1(\mathcal{M}_{\mathrm{good}})$ and $\mathbb{XA}_2 = V_2(\mathcal{M}_{\mathrm{good}})$. This corresponds to **(K2)** in Lemma 5.4(h). Next, suppose that $\mathbb{XA}$ induces sufficiently many edges. Then we take $\mathbb{XA}_1$ and $\mathbb{XA}_2$ to be a bipartition of $\mathbb{XA}$ corresponding to a maximum cut. Hence, the sets $\mathbb{XA}_1$ and $\mathbb{XA}_2$ are again well-connected. This corresponds to the case $e(\mathbb{XA}) \geqslant \eta k n / 12$ in **(K1)** in Lemma 5.4(h). Similarly, if the sets $\mathbb{XA}$ and $\mathbb{XB}$ are well-connected, we are on a good track to getting a sparse version of **(H2)**.

It remains to translate condition (5.2). The right counterpart to this condition is

$$\text{for every } XY \in \mathcal{M}, \quad \mathrm{N}_{G_\nabla}(\mathbb{XA}) \cap X = \emptyset \quad \text{or} \quad \mathrm{N}_{G_\nabla}(\mathbb{XA}) \cap Y = \emptyset \,. \tag{5.3}$$

It so happens that the actual statement of Lemma 5.4 deviates quite substantially from the informal account given above. So, let us now state an informal version of Lemma 5.4. After that, we explain how it relates to the description above. Also, we mark the correspondence between this informal version and the actual lemma by using the same numbering. In particular, assertions (d), (f), (g) in Lemma 5.4 are needed for reasons that cannot be explained in this high-level overview and are not reflected in the informal version. Further, statement of (c') of our informal lemma carries only half of the information compared to the full version in Lemma 5.4.

---

[3]The formal definition of $\mathbb{XB}$ is given again in (5.4).

Let us now give the actual definitions of the sets $\mathbb{XA}$, $\mathbb{XB}$. Later, we explain how these definitions imply the features described above.

Definition 5.3. *Suppose that* $k \in \mathbb{N}$, $\gamma, \eta, \varepsilon, \varepsilon', \nu, \rho > 0$ *and* $\Lambda, \Omega^*, \Omega^{**} > 0$. *Suppose that* G *is a graph with a* $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon, \nu, \rho)$-*sparse decomposition*

$$\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$$

with respect to $\mathbb{L}_{\eta,k}(G)$ *and* $\mathbb{S}_{\eta,k}(G)$. *Suppose further that* $\mathcal{M}_A, \mathcal{M}_B$ *are regularized matchings in* G. *We then define the triple* $(\mathbb{XA}, \mathbb{XB}, \mathbb{XC}) = (\mathbb{XA}, \mathbb{XB}, \mathbb{XC})(\eta, \nabla, \mathcal{M}_A, \mathcal{M}_B)$ *by setting*

$$
\begin{aligned}
\mathbb{XA} &:= \mathbb{L}_{\eta,k}(G) \setminus V(\mathcal{M}_B) \,, \\
\mathbb{XB} &:= \left\{ v \in V(\mathcal{M}_B) \cap \mathbb{L}_{\eta,k}(G) \ : \ \widehat{\deg}(v) < (1 + \eta)\frac{k}{2} \right\} \,, \\
\mathbb{XC} &:= \mathbb{L}_{\eta,k}(G) \setminus (\mathbb{XA} \cup \mathbb{XB}) \,,
\end{aligned}
\tag{5.4}
$$

where $\widehat{\deg}(v)$ *on the second line is defined by*

$$\widehat{\deg}(v) := \deg_G \big(v, \mathbb{S}_{\eta,k}(G) \setminus (V(G_{\mathrm{exp}}) \cup \mathbb{E} \cup V(\mathcal{M}_A \cup \mathcal{M}_B)) \big) \,. \tag{5.5}$$

It is enough to restrict ourselves for the proof to the class $\mathbf{LKSsmall}(n, k, \eta) \subseteq \mathbf{LKS}(n, k, \eta)$. We intentionally leave out (or simplify) almost all numerical parameters in this informal statement.

Informal version of Lemma 5.4. *Suppose* $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\mathrm{reg}}, G_{\mathrm{exp}}, \mathbb{E})$ *is a sparse decomposition of a graph* $G \in \mathbf{LKSsmall}(n, k, \eta)$. *We write* $S^0 := \mathbb{S}_{\eta,k}(G) \setminus (V(G_{\mathrm{exp}}) \cup \mathbb{E})$. *Then there exist regularized matchings* $\mathcal{M}_A$ *and* $\mathcal{M}_B$, *such that for the sets* $\mathbb{XA}$ *and* $\mathbb{XB}$ *defined as in Definition 5.3 we have*

(a) $V(\mathcal{M}_A) \cap V(\mathcal{M}_B) = \emptyset$,

(b) $V_1(\mathcal{M}_B) \subseteq S^0$,

(c′) *for each* $X \in \mathcal{V}(M_A) \cup \mathcal{V}(\mathcal{M}_B)$ *we have that* $X \subseteq \mathbb{S}_{\eta,k}(G)$ *or* $X \subseteq \mathbb{L}_{\eta,k}(G)$,

(e) $e\big(\mathbb{XA}, S^0 \setminus V(\mathcal{M}_A)\big) = 0$,

(h) *if* $\mathbb{XA}$ *induces almost no edges and does not contain a substantial regularized matching*[4] *then there is a substantial amount of edges between* $\mathbb{XA}$ *and* $\mathbb{XB}$.

The regularized matching $\mathcal{M}_A \cup \mathcal{M}_B$ from the lemma plays the role of $\mathcal{M}$ in the motivation above. It remains to justify the dissimilarities between the statement of the lemma and the text above. The first discrepancy is that the definitions of the sets $\mathbb{XA}$ and $\mathbb{XB}$ in (5.4) are quite different from the ones in the motivation above. The other discrepancy is a seeming absence of (5.3) in the statement. As for the first issue, consider an arbitrary vertex $v \in \mathbb{XA}$. Property (e) tells us that v sends no edges to $S^0 \setminus V(\mathcal{M}_A) \supseteq S^0 \setminus (V(\mathcal{M}_A) \cup V(\mathcal{M}_B))$. As $v \in \mathbb{L}_{\eta,k}(G)$, we have that $\deg(v, \mathbb{L}_{\eta,k}(G) \cup V(\mathcal{M}_A) \cup V(\mathcal{M}_B) \cup V(G_{\mathrm{exp}}) \cup \mathbb{E}) \geqslant (1 + \eta)k$, as needed. Next, consider a vertex

---

[4]The exact quantification of "almost no edges" and "substantial regularized matching" does not in guarantee the former property to imply the latter. See also Remark 5.5

$v \in \mathbb{X}\mathbb{B}$. The fact that $v \in \mathbb{L}_{\eta,k}(G)$ together with the definition of $\widehat{\deg}$ immediately gives that $\deg(v, \mathbb{L}_{\eta,k}(G) \cup V(\mathcal{M}_A) \cup V(\mathcal{M}_B) \cup V(G_{\exp}) \cup \mathbb{E}) > (1 + \eta)\frac{k}{2}$, again as needed.

Let us now turn to deriving (5.3). This property is required only for the counterpart of (**H2**). So, we can assume that we do not have the counterpart of (**H1**), that is, the set $\mathbb{X}\mathbb{A}$ induces (almost) no edges. Let us now consider an arbitrary regular pair (X, Y) in $\mathcal{M}_A \cup \mathcal{M}_B$. First assume that $(X, Y) \in \mathcal{M}_B$. Then (b) tells us that $X \subseteq S^0$. We then have $N(\mathbb{X}\mathbb{A}) \cap X = \emptyset$ by Property (e), as needed for (5.3). Next, assume that $(X, Y) \in \mathcal{M}_A$. Then Definition 2.3(ii) (together with (c') of our informal lemma) tells us that at least one of X and Y is contained in $\mathbb{L}_{\eta,k}(G)$. Say this is X. We then have $X \subseteq \mathbb{X}\mathbb{A}$. But the absence of edges inside $\mathbb{X}\mathbb{A}$ tells us that $e(X, \mathbb{X}\mathbb{A}) = 0$, again as needed for (5.3).

5.1.1 Rough versus fine structure

In the dense case [PS12], we can proceed with embedding T using the regularity method immediately after having established a statement like Informal Lemma 5.2. That is, we can zigzag-embed consecutive cut vertices $W_A \cup W_B$ of T in AB, or $A_1 A_2$. When we arrive to a shrub $t \in \mathcal{S}_A \cup \mathcal{S}_B$ stemming from cut vertex $u \in W_A \cup W_B$ embedded to a cluster D (that is, $D = A$, $D = B$, $D = A_1$, or $D = A_2$) we find using (**H1**) or (**H2**) an edge $XY \in E(\mathbf{G}_{\mathrm{reg}})$ such that $DX \in E(\mathbf{G}_{\mathrm{reg}})$ and the pair (X, Y) has not been filled-up. That is, *(i)* using that $DX \in E(\mathbf{G}_{\mathrm{reg}})$ we traverse from D to XY, *(ii)* we embed t in (X, Y), and *(iii)* if t is an internal shrub, we again use that $XD \in E(\mathbf{G}_{\mathrm{reg}})$ to traverse back[5] to D and continue embedding cut vertices in AB or $A_1 A_2$.

In the current setting of the sparse decomposition, the structure given by Lemma 5.4 would allow us to carry out counterparts to *(i)* and *(ii)* (even though there is a number of technical obstacles). That is, we would be able to embed consecutive cut vertices, to traverse to locations suitable for shrubs and to embed these shrubs. However, carrying a counterpart to *(iii)* is a major problem. The symmetry-based argument from the dense case "if DX is an edge then XD is an edge and thus we can traverse in both directions" does not work when working with $\mathbb{E}$ or $G_{\exp}$. So, the purpose of [HKP+c] is to clean the rough structure in such a way that it will allow a counterpart to *(iii)*.

5.2 The role of Lemma 4.8 in the proof of Lemma 5.4

In this section, we explain the role of Lemma 4.8 in our proof of Lemma 5.4. That is, we want to explain why it is not possible in general to find regular matchings $\mathcal{M}_A$ and $\mathcal{M}_B$ from the informal version of Lemma 5.4 inside the cluster graph $\mathbf{G}_{\mathrm{reg}}$. Because of this we will have to find a suitable "re-regularization" which turns out to be provided by Lemma 4.8.

Recall the motivation from Section 5.1. We wish to find two sets $\mathbb{X}\mathbb{A}$ and $\mathbb{X}\mathbb{B}$ (or two sets within $\mathbb{X}\mathbb{A}$) which are suitable for embedding the cut vertices W_A and W_B of a (τk)-fine partition $(W_A, W_B, \mathcal{S}_A, \mathcal{S}_B)$ of T. In this sketch we just focus on finding $\mathbb{X}\mathbb{A}$; the ideas behind finding a suitable set $\mathbb{X}\mathbb{B}$ are similar. To accommodate all the shrubs from $\mathcal{S}_A$ — which might contain up to $k - 2$ vertices in total — we need $\mathbb{X}\mathbb{A}$ to have total degree at least k into the sets $\mathbb{L}_{\eta,k}(G)$, $V(G_{\exp})$, $\mathbb{E}$, together with vertices of any fixed matching $\mathcal{M}$ consisting of regular pairs. This motivates us to look for a regularized matching $\mathcal{M}$ which covers as much as possible of the set

---

[5]As said at the beginning of Section 5.1, if t is internal then both of its neighboring cut vertices are in W_A. In particular, the distance between these two cut vertices is even. That means that to traverse back to D, we really use the pair $XD \in E(\mathbf{G}_{\mathrm{reg}})$ rather than YD (for which we have no guarantee that it is an edge in $\mathbf{G}_{\mathrm{reg}}$).

$S^0 := \mathbb{S}_{\eta,k}(G) \setminus (V(G_{\text{exp}}) \cup \mathbb{E})$, which consists of the vertices that are not utilizable otherwise. As a next step one would prove that there is a set $\mathbb{XA}$ with

$$\text{mindeg}\left(\mathbb{XA}, V(G) \setminus (S^0 \setminus V(\mathcal{M}))\right) \gtrsim k \ .$$

(By $\gtrsim k$ we mean larger than k by a suitable small additional approximation factor.)

In the dense setting [PS12], where the structure of G is determined by $\mathbf{G}_{\text{reg}}$, and where $S^0 = \mathbb{S}_{\eta,k}(G)$, such a matching $\mathcal{M}$ can be found inside $\mathbf{G}_{\text{reg}}$ using the Gallai–Edmonds Matching Theorem. But here, just working with $\mathbf{G}_{\text{reg}}$ is not enough for finding a suitable regularized matching as the following example shows.
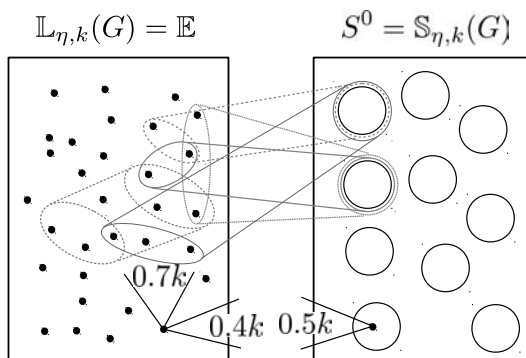
Figure 5.2: An example of a graph $G \in \mathbf{LKS}(n, k, \eta := \frac{1}{10})$ in which $\mathbf{G}_{\text{reg}}$ is empty, yet there is no candidate set for $\mathbb{XA}$ of vertices which have degrees at least k outside the set S^0. Sample dense spots are shown in grey.

Figure 5.2 shows a graph G with $\mathbb{L}_{\eta,k}(G) = \mathbb{E}$. Let us describe the construction of such a graph G. We partition the vertex sets into to-be sets $\mathbb{S}_{\eta,k}(G)$ and $\mathbb{L}_{\eta,k}(G)$. We further gather vertices of $\mathbb{S}_{\eta,k}(G)$ into clusters. We now insert edges into G. All the edges inserted will be in the form of dense spots. These dense spots either both parts in $\mathbb{L}_{\eta,k}(G)$, or one part $\mathbb{L}_{\eta,k}(G)$ and the other in $\mathbb{S}_{\eta,k}(G)$. We do this so that each inserted dense spot in the $\mathbb{S}_{\eta,k}(G)$-part respects the cluster structure, while it behaves in a random-like way in the $\mathbb{L}_{\eta,k}(G)$-part. Further, we require that each $\mathbb{L}_{\eta,k}(G)$-vertex sends $0.7k$ edges to $\mathbb{L}_{\eta,k}(G)$ and $0.4k$ edges to $\mathbb{S}_{\eta,k}(G)$, and each $\mathbb{S}_{\eta,k}(G)$-vertex receives $0.5k$ edges from $\mathbb{L}_{\eta,k}(G)$. Clearly, such a construction is possible.

The point of the construction is that $\mathbb{E} = \mathbb{L}_{\eta,k}(G)$, and that $S^0 = \mathbb{S}_{\eta,k}(G)$ form clusters which do not induce any dense regular pairs. No vertex has degree $\gtrsim k$ outside S^0, and the cluster graph $\mathbf{G}_{\text{reg}}$ contains no matching.

However, in this situation we can still find a large regularized matching $\mathcal{M}$ between $\mathbb{L}_{\eta,k}(G)$ and $\mathbb{S}_{\eta,k}(G)$, by regularizing the crossing dense spots $\mathcal{D}$ (which we can assume to be the original dense spots inserted in our construction). In general, obtaining a regularized matching is of course more complicated. Given the above example, one may ask whether there is any role of the graph $\mathbf{G}_{\text{reg}}$ at all. The answer is that for constructing $\mathcal{M}$, we can either use directly the edges of $\mathbf{G}_{\text{reg}}$, or, if we do not have these edges the information about their lack helps us to find $\mathcal{M}$ elsewhere.

5.3 Finding the structure

We can now state the main result of this paper.

Lemma 5.4. *For every $\eta \in (0,1)$, $\Omega > 0, \gamma \in (0, \eta/3)$ there is $\beta > 0$ so that for every $\varepsilon \in (0, \frac{\gamma^2 \eta}{12})$ there exist $\varepsilon', \pi > 0$ such that for every $\nu > 0$ there exists $k_0 \in \mathbb{N}$ such that for every Ω^* with $\Omega^* < \Omega$, every Ω^{**} with $\Omega^{**} > \max\{2, \Omega^*\}$ and every k with $k > k_0$ the following holds.*

Suppose $\nabla = (\mathbb{H}, \mathbf{V}, \mathcal{D}, G_{\text{reg}}, G_{\exp}, \mathbb{E})$ is a $(k, \Omega^{**}, \Omega^*, \Lambda, \gamma, \varepsilon', \nu, \rho)$-sparse decomposition of a graph $G \in \mathbf{LKSsmall}(n, k, \eta)$ with respect to $S := \mathbb{S}_{\eta,k}(G)$ and $L := \mathbb{L}_{\eta,k}(G)$ which captures all but at most $\eta k n /6$ edges of G. Let $\mathfrak{c}$ be the size of the clusters $\mathbf{V}$.[6] Write

$$S^0 := S \setminus (V(G_{\exp}) \cup \mathbb{E}) . \tag{5.6}$$

Then $G_{\mathcal{D}}$ contains two $(\varepsilon, \beta, \pi\mathfrak{c})$-regularized matchings $\mathcal{M}_A$ and $\mathcal{M}_B$ such that for the triple $(\mathbb{XA}, \mathbb{XB}, \mathbb{XC}) := (\mathbb{XA}, \mathbb{XB}, \mathbb{XC})(\eta, \nabla, \mathcal{M}_A, \mathcal{M}_B)$ we have

(a) $V(\mathcal{M}_A) \cap V(\mathcal{M}_B) = \emptyset$,

(b) $V_1(\mathcal{M}_B) \subseteq S^0$,

(c) for each $(X_1, X_2) \in \mathcal{M}_A \cup \mathcal{M}_B$, there is a dense spot $(D_1, D_2; E_D) \in \mathcal{D}$ with $X_1 \subseteq D_1, X_2 \subseteq D_2$, and furthermore, either $X_1 \subseteq S$ or $X_1 \subseteq L$, and $X_2 \subseteq S$ or $X_2 \subseteq L$,

(d) for each $X_1 \in \mathcal{V}_1(\mathcal{M}_A \cup \mathcal{M}_B)$ there exists a cluster $C_1 \in \mathbf{V}$ such that $X_1 \subseteq C_1$, and for each $X_2 \in \mathcal{V}_2(\mathcal{M}_A \cup \mathcal{M}_B)$ we have $X_2 \subseteq L \cap \mathbb{E}$ or there exists $C_2 \in \mathbf{V}$ such that $X_2 \subseteq C_2$,

(e) $e_{G_\nabla}\big(\mathbb{XA}, S^0 \setminus V(\mathcal{M}_A)\big) \leqslant \gamma k n$,

(f) $e_{G_{\text{reg}}}(V(G) \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)) \leqslant \varepsilon \Omega^* k n$,

(g) for the regularized matching $\mathcal{N}_{\mathbb{E}} := \{(X_1, X_2) \in \mathcal{M}_A \cup \mathcal{M}_B : (X_1 \cup X_2) \cap \mathbb{E} \neq \emptyset\}$ we have $e_{G_{\text{reg}}}\big(V(G) \setminus V(\mathcal{M}_A \cup \mathcal{M}_B), V(\mathcal{N}_{\mathbb{E}})\big) \leqslant \varepsilon \Omega^* k n$,

(h) for $\mathcal{M}_{\text{good}} := \{(X_1, X_2) \in \mathcal{M}_A : X_1 \cup X_2 \subseteq \mathbb{XA}\}$ we have that each $\mathcal{M}_{\text{good}}$-edge is an edge of $\mathbf{G}_{\text{reg}}$, and at least one of the following conditions holds

    (K1) $2e_G(\mathbb{XA}) + e_G(\mathbb{XA}, \mathbb{XB}) \geqslant \eta k n /3$,

    (K2) $|V(\mathcal{M}_{\text{good}})| \geqslant \eta n /3$.

Remark 5.5. *As explained in Section 5.1, property* (h) *is the most important part of Lemma 5.4. Note that the assertion* **(K2)** *implies a quantitatively weaker version of* **(K1)***. Indeed, consider $(X_1, X_2) \in \mathcal{M}_A$. An average vertex $v \in X_1$ sends at least $\beta \cdot \pi\mathfrak{c} \geqslant \beta \cdot \pi\nu k$ edges to X_2. Thus, if $|V(\mathcal{M}_{\text{good}})| \geqslant \eta n /3$ then $\mathcal{M}_{\text{good}}$ induces at least $(\eta n /6) \cdot \beta \cdot \pi\nu k = \Theta(kn)$ edges in $\mathbb{XA}$. Such a bound, however, would be insufficient for our purposes as later $\eta \gg \pi, \nu$. So, the deficit in the number of edges asserted in* **(K2)** *(compared to the $e_G(\mathbb{XA}) \geqslant \eta k n /12$ part of* **(K1)***) is compensated by the fact that these edges are already regularized.*

For the proof we need the well-known Gallai–Edmonds Matching Theorem, which we state next. A graph H is called *factor-critical* if $H - v$ has a perfect matching for each $v \in V(H)$.

Theorem 5.6 (Gallai–Edmonds matching theorem (see for instance [LP86]))**.** *Let H be a graph. Then there exist a set $Q \subseteq V(H)$ and a matching M of size $|Q|$ in H such that*

---

[6]The number $\mathfrak{c}$ is irrelevant when $\mathbf{V} = \emptyset$. In particular, note that in that case we necessarily have $\mathcal{M}_A = \mathcal{M}_B = \emptyset$ for the regularized matchings given by the lemma.

1) every component of $H - Q$ is factor-critical, and

2) M matches every vertex in Q to a different component of $H - Q$.

The set Q in Theorem 5.6 is often referred to as a *separator*.

Proof of Lemma 5.4. The idea of the proof is to first obtain some information about the structure of the graph $\mathbf{G}_{\mathrm{reg}}$ with the help of Theorem 5.6. Then the structure given by Theorem 5.6 is refined by Lemma 4.8 to yield the assertions of the lemma.

Let us begin with setting the parameters. Let $\beta := \beta_{\mathrm{L4.8}}$ be given by Lemma 4.8 for input parameters $\Omega_{\mathrm{L4.8}} := \Omega$, $\rho_{\mathrm{L4.8}} := \gamma^2$, and let ε' and π be given by Lemma 4.8 for further input parameter $\varepsilon_{\mathrm{L4.8}} := \varepsilon$. Last, let k_0 be given by Lemma 4.8 with the above parameters and $\gamma_{\mathrm{L4.8}} := \nu$.

Without loss of generality we assume that $\varepsilon' \leqslant \varepsilon$ and $\beta < \gamma^2$. We write $\mathbf{S} := \{C \in \mathbf{V} : C \subseteq S\}$ and $\mathbf{L} := \{C \in \mathbf{V} : C \subseteq L\}$. Further, let $\mathbf{S}^0 := \{C \in \mathbf{S} : C \subseteq S^0\}$.

Let $\mathbf{Q}$ be a separator and let N_0 be a matching given by Theorem 5.6 applied to the graph $\mathbf{G}_{\mathrm{reg}}$. We will presume that the pair $(\mathbf{Q}, N_0)$ is chosen among all the possible choices so that the number of vertices of $\mathbf{S}^0$ that are isolated in $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$ and are not covered by N_0 is minimized. Let $\mathbf{S}^{\mathrm{I}}$ denote the set of vertices in $\mathbf{S}^0$ that are isolated in $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$. Recall that the components of $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$ are factor critical.

Define $\mathbf{S}^{\mathrm{R}} \subseteq V(\mathbf{G}_{\mathrm{reg}})$ as a minimal set such that

- $\mathbf{S}^{\mathrm{I}} \setminus V(N_0) \subseteq \mathbf{S}^{\mathrm{R}}$, and

- if $C \in \mathbf{S}$ and there is an edge $DZ \in E(\mathbf{G}_{\mathrm{reg}})$ with $Z \in \mathbf{S}^{\mathrm{R}}$, $D \in \mathbf{Q}$, $CD \in N_0$ then $C \in \mathbf{S}^{\mathrm{R}}$.

Then each vertex from $\mathbf{S}^{\mathrm{R}}$ is reachable from $\mathbf{S}^{\mathrm{I}} \setminus V(N_0)$ by a path in $\mathbf{G}_{\mathrm{reg}}$ that alternates between $\mathbf{S}^{\mathrm{R}}$ and $\mathbf{Q}$, and has every second edge in N_0. Also note that for all $CD \in N_0$ with $C \in \mathbf{Q}$ and $D \in \mathbf{S}^0 \setminus \mathbf{S}^{\mathrm{R}}$ we have

$$\deg_{\mathbf{G}_{\mathrm{reg}}}(C, \mathbf{S}^{\mathrm{R}}) = 0 . \tag{5.7}$$

Let us prove another property of $\mathbf{S}^{\mathrm{R}}$.

Claim 5.5.1. $\mathbf{S}^{\mathrm{R}} \subseteq \mathbf{S}^{\mathrm{I}} \subseteq \mathbf{S}^{\mathrm{R}} \cup V(N_0)$. In particular, $\mathbf{S}^{\mathrm{R}} \subseteq \mathbf{S}^0$.

Proof of Claim 5.5.1. By the definition of $\mathbf{S}^{\mathrm{R}}$, we only need to show that $\mathbf{S}^{\mathrm{R}} \subseteq \mathbf{S}^{\mathrm{I}}$. So suppose there is a vertex $C \in \mathbf{S}^{\mathrm{R}} \setminus \mathbf{S}^{\mathrm{I}}$. By the definition of $\mathbf{S}^{\mathrm{R}}$ there is a non-trivial path R going from $\mathbf{S}^{\mathrm{I}} \setminus V(N_0)$ to C that alternates between $\mathbf{S}^{\mathrm{R}}$ and $\mathbf{Q}$, and has every second edge in N_0. Then, the matching $N_0' := N_0 \triangle E(R)$ covers more vertices of $\mathbf{S}^{\mathrm{I}}$ than N_0 does. Further, it is straightforward to check that the separator $\mathbf{Q}$ together with the matching N_0' satisfies the assertions of Theorem 5.6. This is a contradiction, as desired. □

Using a very similar alternating path argument we see the following.

Claim 5.5.2. If $CD \in N_0$ with $C \in \mathbf{Q}$ and $D \notin \mathbf{S}^{\mathrm{I}}$ then $\deg_{\mathbf{G}_{\mathrm{reg}}}(C, \mathbf{S}^{\mathrm{R}}) = 0$.

Using the factor-criticality of the components of $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$ we extend N_0 to a matching N_1 as follows. For each component K of $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$ which meets $V(N_0)$, we add a perfect matching of $K - V(N_0)$. Furthermore, for each non-singleton component K of $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$ which does not meet $V(N_0)$, we add a matching which meets all but exactly one vertex of $\mathbf{L} \cap V(K)$. This is possible

as by the definition of the class **LKSsmall**(n, k, η) we have that $\mathbf{G}_{\mathrm{reg}} - \mathbf{L}$ is independent, and so $\mathbf{L} \cap V(K) \neq \emptyset$. This choice of N_1 guarantees that

$$e_{\mathbf{G}_{\mathrm{reg}}}(\mathbf{V} \setminus V(N_1)) = 0 . \tag{5.8}$$

We set

$$M := \{ C_1 C_2 \in N_0 \; : \; C_1 \in \mathbf{S}^{\mathrm{R}}, C_2 \in \mathbf{Q} \} .$$

We have that

$$e_{\mathbf{G}_{\mathrm{reg}}}(\mathbf{V} \setminus V(N_1), V(M) \cap \mathbf{S}^{\mathrm{R}}) = 0 . \tag{5.9}$$

As $\mathbf{S}$ is an independent set in $\mathbf{G}_{\mathrm{reg}}$, we have that

$$\mathbf{Q}_M := V(M) \cap \mathbf{Q} \subseteq \mathbf{L} . \tag{5.10}$$

The matching M in $\mathbf{G}_{\mathrm{reg}}$ corresponds to an $(\varepsilon', \gamma^2, \mathfrak{c})$-regularized matching $\mathcal{M}$ in the underlying graph G_{reg}, with $V_2(\mathcal{M}) \subseteq \bigcup \mathbf{Q}$ (recall that regularized matchings have orientations on their edges). Likewise, we define $\mathcal{N}_1$ as the $(\varepsilon', \gamma^2, \mathfrak{c})$-regularized matching corresponding to N_1. The $\mathcal{N}_1$-edges are oriented so that $V_1(\mathcal{N}_1) \cap \bigcup \mathbf{Q} = \emptyset$; this condition does not specify orientations of all the $\mathcal{N}_1$-edges and we orient the remaining ones in an arbitrary fashion. We write $S^{\mathrm{R}} := \bigcup \mathbf{S}^{\mathrm{R}}$.

Claim 5.5.3. $e_{G_\nabla}(L \setminus (\mathbb{E} \cup V(\mathcal{M})), S^{\mathrm{R}}) = 0.$

Proof of Claim 5.5.3. We start by showing that for every cluster $C \in \mathbf{L} \setminus V(M)$ we have

$$\deg_{\mathbf{G}_{\mathrm{reg}}}(C, \mathbf{S}^{\mathrm{R}}) = 0 . \tag{5.11}$$

First, if $C \notin \mathbf{Q}$, then (5.11) is true since $\mathbf{S}^{\mathrm{R}} \subseteq \mathbf{S}^{\mathrm{I}}$ by Claim 5.5.1. So suppose that $C \in \mathbf{Q}$, and let $D \in V(\mathbf{G}_{\mathrm{reg}})$ be such that $DC \in N_0$. Now if $D \notin \mathbf{S}^{\mathrm{I}}$ then (5.11) follows from Claim 5.5.2. On the other hand, suppose $D \in \mathbf{S}^{\mathrm{I}} \subseteq \mathbf{S}^0$. As $C \notin V(M)$, we know that $D \notin \mathbf{S}^{\mathrm{R}}$, and thus, (5.11) follows from (5.7).

Now, by (5.11), G_{reg} has no edges between $L \setminus (\mathbb{E} \cup V(\mathcal{M}))$ and S^{R}. Also, no such edges can be in G_{exp} or incident with $\mathbb{E}$, since $\mathbf{S}^{\mathrm{R}} \subseteq \mathbf{S}^0$ by Claim 5.5.1. Finally, since $G \in$ **LKSsmall**(n, k, η) and $\Omega^{**} > 2 > (1 + \eta)$, there are no edges between $\mathbb{H}$ and S. This proves the claim. $\qquad \square$

We prepare ourselves for an application of Lemma 4.8. The numerical parameters of the lemma are $\Omega_{\mathrm{L4.8}}, \rho_{\mathrm{L4.8}}, \varepsilon_{\mathrm{L4.8}}$ and $\gamma_{\mathrm{L4.8}}$ as above. The input objects for the lemma are the graph $G_{\mathcal{D}}$ of order $n' \leqslant n$, the collection of $(\gamma k, \gamma)$-dense spots $\mathcal{D}$, the matching $\mathcal{M}$, the (νk)-ensemble $\mathcal{C}_{\mathrm{L4.8}} := \mathbf{S}^{\mathrm{R}} \setminus V(N_1)$, and the set $Y_{\mathrm{L4.8}} := L \cap \mathbb{E}$. Note that Definition 3.4, item 6, implies that $\mathcal{D}$ absorbs $\mathcal{M}$. Further, (4.20) is satisfied by Definition 3.4, item 7.

The output of Lemma 4.8 is an $(\varepsilon, \beta, \pi \mathfrak{c})$-regularized matching $\mathcal{M}'$ with the following properties.

(I) $|V(\mathcal{M}) \setminus V(\mathcal{M}')| < \varepsilon n' \leqslant \varepsilon n.$

(II) For each $(X_1, X_2) \in \mathcal{M}'$ there are sets $C \in \mathbf{S}^{\mathrm{R}}$ and $(D_1, D_2; E_D) \in \mathcal{D}$ such that $X_1 \subseteq C \cap D_1$ and either $X_2 \subseteq L \cap \mathbb{E} \cap D_2$ or there exists $C' \in \mathbf{Q}_M$ such that $X_2 \subseteq L \cap \mathbb{E} \cap C'$.

(Indeed, to see this we use that $\mathcal{V}_1(\mathcal{M}) \subseteq \mathbf{S}^{\mathrm{R}}$ and that $V_2(\mathcal{M}) \subseteq \bigcup \mathbf{Q}_M$ by the definition of $\mathcal{M}$.)
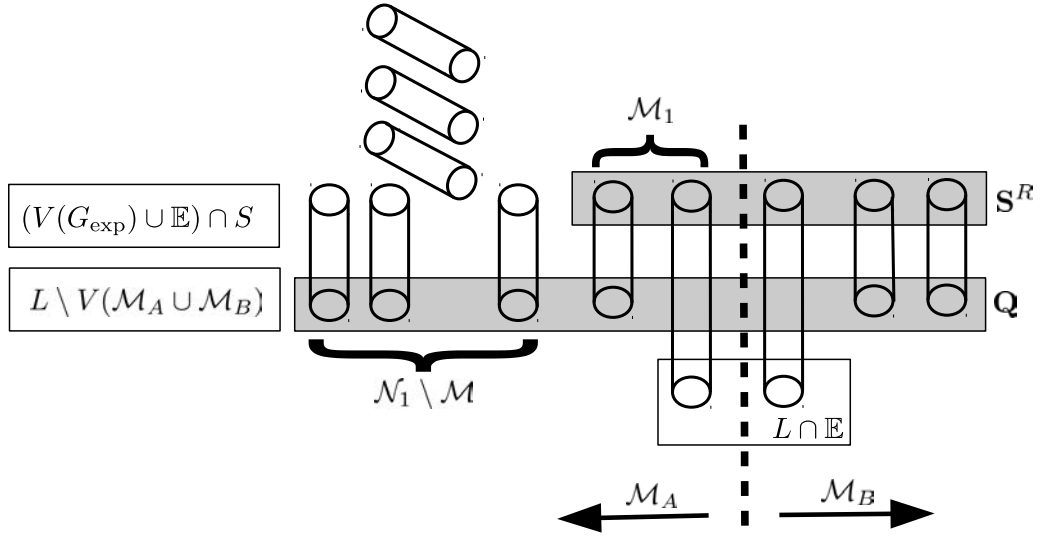
Figure 5.3: The situation in G after applying Lemma 4.8. The dotted line illustrates the separation as in (III).

(III) There is a partition of $\mathcal{M}'$ into $\mathcal{M}_1$ and $\mathcal{M}_B$ such that

$$e_{G_{\mathcal{D}}}\left(\left(\left(S^{\mathrm{R}} \setminus V(\mathcal{N}_1)\right) \cup V_1(\mathcal{M})\right) \setminus V_1(\mathcal{M}_1), \left((L \cap \mathbb{E}) \cup V_2(\mathcal{M})\right) \setminus V_2(\mathcal{M}_B)\right) < \gamma k n' .$$

We claim that also

(IV) $V(\mathcal{M}') \cap V(\mathcal{N}_1 \setminus \mathcal{M}) = \emptyset$.

Indeed, let $(X_1, X_2) \in \mathcal{M}'$ be arbitrary. Then by (II) there is $C \in \mathbf{S}^{\mathrm{R}}$ such that $X_1 \subseteq C$. By Claim 5.5.1, C is a singleton component of $\mathbf{G}_{\mathrm{reg}} - \mathbf{Q}$. In particular, if C is covered by N_1 then $C \in V(M)$. It follows that $X_1 \cap V(\mathcal{N}_1 \setminus \mathcal{M}) = \emptyset$. In a similar spirit, the easy fact that $(Y \cup \bigcup \mathbf{Q}_M) \cap V(\mathcal{N}_1 \setminus \mathcal{M}) = \emptyset$ together with (II) gives $X_2 \cap V(\mathcal{N}_1 \setminus \mathcal{M}) = \emptyset$. This establishes (IV).

Observe that (II) implies that $V_1(\mathcal{M}') \subseteq S^{\mathrm{R}}$, and so, by Claim 5.5.1 we know that

$$V_1(\mathcal{M}_B) \subseteq S^{\mathrm{R}} \subseteq \bigcup \mathbf{S}^{\mathrm{I}} \subseteq S^0. \tag{5.12}$$

Set

$$\mathcal{M}_A := (\mathcal{N}_1 \setminus \mathcal{M}) \cup \mathcal{M}_1 . \tag{5.13}$$

Then $\mathcal{M}_A$ is an $(\varepsilon, \beta, \pi\mathfrak{c})$-regularized matching. Note that from now on, the sets $\mathbb{XA}, \mathbb{XB}$ and $\mathbb{XC}$ are defined. The situation is illustrated in Figure 5.3. By (IV), we have $V(\mathcal{M}_A) \cap V(\mathcal{M}_B) = \emptyset$, as required for Lemma 5.4(a). Lemma 5.4(b) follows from (5.12). The claim below asserts that the next two properties are satisfied as well.

Claim 5.5.4. Lemma 5.4(c) and Lemma 5.4(d) are satisfied.

Proof of Claim 5.5.4. Consider an arbitrary pair $(X_1, X_2) \in \mathcal{M}_A \cup \mathcal{M}_B$. Either we have that $(X_1, X_2) \in \mathcal{N}_1$ or $(X_1, X_2) \in \mathcal{M}'$. In the former case, $X_1 X_2$ is an edge in $\mathbf{G}_{\mathrm{reg}}$. Then the properties for (X_1, X_2) asserted in Lemma 5.4(c) and Lemma 5.4(d) follow from the fact that the cluster graph is prepartitioned with respect to S and L, and from Definition 3.4(6).

In the case $(X_1, X_2) \in \mathcal{M}'$, the asserted properties are given by (II). □

We now turn to Lemma 5.4(e). First we prove some auxiliary statements.

Claim 5.5.5. We have $\mathbf{S}^0 \setminus V(N_1 \setminus M) \subseteq \mathbf{S}^{\mathrm{R}}$.

Proof of Claim 5.5.5. Let $C \in \mathbf{S}^0 \setminus V(N_1 \setminus M)$. Note that if $C \notin \mathbf{S}^{\mathrm{I}}$, then $C \in V(N_1)$. On the other hand, if $C \in \mathbf{S}^{\mathrm{I}}$, then we use Claim 5.5.1 to see that $C \in \mathbf{S}^{\mathrm{R}} \cup V(N_1)$. We deduce that in either case $C \in \mathbf{S}^{\mathrm{R}} \cup V(N_1)$. The choice of C implies that $C \in \mathbf{S}^{\mathrm{R}} \cup V(M)$. Now, if $C \in V(M)$, then $C \in \mathbf{S}^{\mathrm{R}}$ by (5.10) and by the definition of M. Thus $C \in \mathbf{S}^{\mathrm{R}}$ as desired. $\qquad\square$

It will be convenient to work with a set $\bar{S}^0 \subseteq S^0$, $\bar{S}^0 := (S \cap \bigcup \mathbf{V}) \setminus V(G_{\exp}) = \bigcup \mathbf{S}^0$. The next two easy claims assert absence of edges of certain types incident to S^0 and $\bar{S}^0$.

Claim 5.5.6. The vertices in $S^0 \setminus \bar{S}^0$ are isolated in G_∇.

Proof of Claim 5.5.6. Indeed, let us check Definition 3.6. Clearly, $S^0 \setminus \bar{S}^0$ is disjoint from $V(G_{\mathrm{reg}})$ and $V(G_{\exp})$. Further, $S^0 \setminus \bar{S}^0$ sends no edges to $\mathbb{H}$ by Definition 2.3(ii). Lastly, the set $S^0 \setminus \bar{S}^0$ is disjoint from the "avoiding edges" spanned by the vertex sets $\mathbb{E}$ and $\mathbb{E} \cup \bigcup \mathbf{V}$. $\qquad\square$

Claim 5.5.7. We have $G_\nabla[L \cap \mathbb{E}, \bar{S}^0] = G_\mathcal{D}[L \cap \mathbb{E}, \bar{S}^0]$.

Proof of Claim 5.5.7. The $\supseteq$-inclusion of the edge-sets is clear.

Next, recall that Definition 3.6 tells us that each edge in G_∇ between $\mathbb{E}$ and $\bigcup \mathbf{V}$ is either in $G_{\exp}$ or in $G_\mathcal{D}$. As $\bar{S}^0 \cap V(G_{\exp}) = \emptyset$, the $\subseteq$-inclusion follows. $\qquad\square$

By Claim 5.5.5, we have

$$\bar{S}^0 \setminus V(\mathcal{M}_A) \subseteq \left(\bigcup \mathbf{S}^0 \setminus V(N_1 \setminus M) \right) \setminus V(\mathcal{M}_A) \subseteq S^{\mathrm{R}} \setminus V(\mathcal{M}_A). \tag{5.14}$$

As every edge incident to $S^0 \setminus \bar{S}^0$ is uncaptured, we see that

$$\begin{aligned}
E_{G_\nabla}\big(\mathbb{XA} \cap \mathbb{E}, S^0 \setminus V(\mathcal{M}_A)\big) &= E_{G_\nabla}\big(\mathbb{XA} \cap \mathbb{E}, \bar{S}^0 \setminus V(\mathcal{M}_A)\big) \\
{\scriptstyle(\mathbb{XA} \cap \mathbb{E} = (L \cap \mathbb{E}) \setminus V(\mathcal{M}_B),\ \mathrm{C5.5.7})} \quad &= E_{G_\mathcal{D}}\big((L \cap \mathbb{E}) \setminus V(\mathcal{M}_B), \bar{S}^0 \setminus V(\mathcal{M}_A)\big) \\
{\scriptstyle(\text{by } (5.14))} \quad &\subseteq E_{G_\mathcal{D}}\big((L \cap \mathbb{E}) \setminus V(\mathcal{M}_B),\ S^{\mathrm{R}} \setminus V(\mathcal{M}_A) \big). \tag{5.15}
\end{aligned}$$

Claim 5.5.8. We have

$$E_{G_{\mathrm{reg}}}\big(\mathbb{XA} \cap \textstyle\bigcup \mathbf{V}, S^0 \setminus V(\mathcal{M}_A)\big) \subseteq E_{G_\mathcal{D}}\left(((L \cap \mathbb{E}) \cup V_2(\mathcal{M})) \setminus V_2(\mathcal{M}_B), S^{\mathrm{R}} \setminus V(\mathcal{M}_A) \right).$$

Before proving Claim 5.5.8, let us see that it implies Lemma 5.4(e). As $G \in \mathbf{LKSsmall}(n, k, \eta)$, there are no edges between $\mathbb{H}$ and S. That means that any captured edge from $\mathbb{XA}$ to $S^0 \setminus V(\mathcal{M}_A)$ must start in $\mathbb{E}$ or in $\bigcup \mathbf{V}$, and must be contained in $G_\mathcal{D}$. Thus Lemma 5.4(e) follows by plugging (III) into (5.15) and into Claim 5.5.8.

Let us now turn to proving Claim 5.5.8.

Proof of Claim 5.5.8. First, observe that by the definition of $\mathbb{XA}$ and by the definition of $\mathcal{M}$ (and M) we have

$$\mathbb{XA} \cap \textstyle\bigcup \mathbf{V} \subseteq (V_2(\mathcal{M}) \setminus V_2(\mathcal{M}_B)) \cup (L \setminus (\mathbb{E} \cup V(\mathcal{M}))) . \tag{5.16}$$

Further, by applying (5.14) and Claim 5.5.3 we get

$$E_{G_{\mathrm{reg}}}\big(L \setminus (\mathbb{E} \cup V(\mathcal{M})), \bar{S}^0 \setminus V(\mathcal{M}_A)\big) = \emptyset . \tag{5.17}$$

27

Therefore, we obtain

$$E_{G_{\mathrm{reg}}}\big(\mathbb{X}\mathbb{A}\cap\bigcup\mathbf{V},S^0\setminus V(\mathcal{M}_A)\big)\overset{\mathrm{C5.5.6}}{=}E_{G_{\mathrm{reg}}}\big(\mathbb{X}\mathbb{A}\cap\bigcup\mathbf{V},\bar{S}^0\setminus V(\mathcal{M}_A)\big)$$

$$\text{(by (5.16))}\quad\subseteq E_{G_{\mathrm{reg}}}\big(V_2(\mathcal{M})\setminus V_2(\mathcal{M}_B),\bar{S}^0\setminus V(\mathcal{M}_A)\big)$$

$$\cup E_{G_{\mathrm{reg}}}\big(L\setminus(\mathbb{E}\cup V(\mathcal{M})),\bar{S}^0\setminus V(\mathcal{M}_A)\big)$$

$$\text{(by (5.14), (5.17))}\quad\subseteq E_{G_{\mathrm{reg}}}\big(V_2(\mathcal{M})\setminus V_2(\mathcal{M}_B),S^{\mathrm{R}}\setminus V(\mathcal{M}_A)\big)\,,$$

as needed. $\qquad\square$

In order to prove (f) we first observe that

$$V(\mathcal{N}_1)\setminus V(\mathcal{M}_A\cup\mathcal{M}_B)\overset{(5.13)}{=}V(\mathcal{N}_1)\setminus V\big((\mathcal{N}_1\setminus\mathcal{M})\cup\mathcal{M}_1\cup\mathcal{M}_B\big)$$

$$=(V(\mathcal{N}_1)\cap V(\mathcal{M}))\setminus V(\mathcal{M}_B\cup\mathcal{M}_1)$$

$$\overset{(\mathrm{III})}{=}(V(\mathcal{N}_1)\cap V(\mathcal{M}))\setminus V(\mathcal{M}')=V(\mathcal{M})\setminus V(\mathcal{M}')\,.\qquad(5.18)$$

Now, we have

$$e_{G_{\mathrm{reg}}}(V(G)\setminus V(\mathcal{M}_A\cup\mathcal{M}_B))\leqslant e_{G_{\mathrm{reg}}}(V(G)\setminus V(\mathcal{N}_1))+\sum_{v\in V(\mathcal{N}_1)\setminus V(\mathcal{M}_A\cup\mathcal{M}_B)}\deg_{G_\nabla}(v)$$

$$\text{(by (5.8) and (5.18))}\quad\leqslant\sum_{v\in V(\mathcal{M})\setminus V(\mathcal{M}')}\deg_{G_\nabla}(v)\leqslant|V(\mathcal{M})\setminus V(\mathcal{M}')|\Omega^* k$$

$$\text{(by (I))}\quad<\varepsilon\Omega^* kn\,,$$

which shows (f).

Let us turn to proving (g). First, recall that we have $V(\mathcal{N}_\mathbb{E})\subseteq V(\mathcal{M}')\cup V(\mathcal{N}_1)$ (cf. 5.13). Since $V(\mathcal{N}_1)\cap\mathbb{E}=\emptyset$ we actually have

$$V(\mathcal{N}_\mathbb{E})=V(\mathcal{N}_\mathbb{E})\cap V(\mathcal{M}')\,.\qquad(5.19)$$

Using (5.19) and (II) we get

$$e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{N}_1),V(\mathcal{N}_\mathbb{E})\big)\leqslant e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{N}_1),V(\mathcal{M}')\cap S^{\mathrm{R}}\big)$$

$$\text{(by (5.9))}\quad\leqslant e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{N}_1),(V(\mathcal{M}')\setminus V(\mathcal{M}))\cap S^{\mathrm{R}}\big)$$

$$\text{(by (IV))}\quad\leqslant e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{N}_1),(V(\mathcal{M}')\setminus V(\mathcal{N}_1))\cap S^{\mathrm{R}}\big)$$

$$\leqslant 2e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{N}_1)\big)\overset{(5.8)}{=}0\,.\qquad(5.20)$$

We have

$$e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{M}_A\cup\mathcal{M}_B),V(\mathcal{N}_\mathbb{E})\big)\leqslant e_{G_{\mathrm{reg}}}\big(V(G)\setminus V(\mathcal{N}_1),V(\mathcal{N}_\mathbb{E})\big)$$

$$+e_{G_{\mathrm{reg}}}\big(V(\mathcal{N}_1)\setminus V(\mathcal{M}_A\cup\mathcal{M}_B),V(G)\big)$$

$$\text{(by (5.20))}\quad\leqslant 0+|V(\mathcal{N}_1)\setminus V(\mathcal{M}_A\cup\mathcal{M}_B)|\Omega^* k$$

$$\text{(by (5.18), (I))}\quad\leqslant\varepsilon\Omega^* kn\,,$$

as needed.

We have thus shown Lemma 5.4(a)–(g). It only remains to prove Lemma 5.4(h), which we will do in the remainder of this section.

We first collect several properties of $\mathbb{XA}$ and $\mathbb{XC}$. The definitions of $\mathbb{XC}$ and S^0 give

$$|\mathbb{XC}|(1+\eta)\frac{k}{2} \leqslant e_G\big(\mathbb{XC}, S^0 \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)\big) \leqslant |S^0 \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)|(1+\eta)k \,. \tag{5.21}$$

Each vertex of $\mathbb{XC}$ has degree at least $(1+\eta)\frac{k}{2}$ into S, and so,

$$e_G(S, \mathbb{XC}) \geqslant |\mathbb{XC}| \left\lceil (1+\eta)\frac{k}{2} \right\rceil \,. \tag{5.22}$$

has neighbours in S. Also, for each vertex $v \in \mathbb{XC}$, Definition 2.3(ii) gives that

$$\deg_G(v) = \lceil (1+\eta)k \rceil \tag{5.23}$$

Consequently (using the elementary inequality $\lceil a \rceil - \lceil \frac{a}{2} \rceil \leqslant \frac{a}{2}$),

$$e_G\big(\mathbb{XA}, \mathbb{XC}\big) \overset{(5.23)}{\leqslant} |\mathbb{XC}|\lceil (1+\eta)k \rceil - e_G(S, \mathbb{XC})$$

$$\overset{(5.22)}{\leqslant} |\mathbb{XC}|(1+\eta)\frac{k}{2} \tag{5.24}$$

$$\overset{(5.21)}{\leqslant} |S^0 \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)|(1+\eta)k \,. \tag{5.25}$$

Let $\mathcal{M}_{\text{good}}$ be defined as in Lemma 5.4(h), that is, $\mathcal{M}_{\text{good}} := \{(X_1, X_2) \in \mathcal{M}_A : X_1 \cup X_2 \subseteq \mathbb{XA}\}$. Note that (5.12) implies that $X_1 \subseteq S$ for every $(X_1, X_2) \in \mathcal{M}_B$. Thus by the definition of $\mathbb{XA}$,

$$\text{if } (X_1, X_2) \in \mathcal{M}_A \cup \mathcal{M}_B \text{ with } X_1 \cup X_2 \subseteq L \text{ then } (X_1, X_2) \in \mathcal{M}_{\text{good}}. \tag{5.26}$$

We will now prove the first part of Lemma 5.4(h), that is, we show that each $\mathcal{M}_{\text{good}}$-edge is an edge of $\mathbf{G}_{\text{reg}}$. Indeed, by (II), we have that $V_1(\mathcal{M}_1) \subseteq S$, so as $\mathbb{XA} \cap S = \emptyset$, it follows that $\mathcal{M}_1 \cap \mathcal{M}_{\text{good}} = \emptyset$. Thus $\mathcal{M}_{\text{good}} \subseteq \mathcal{N}_1$. As $\mathcal{N}_1$ corresponds to a matching in $\mathbf{G}_{\text{reg}}$, all is as desired.

Finally, let us assume that neither **(K1)** nor **(K2)** is fulfilled. After five preliminary observations (Claim 5.5.9–Claim 5.5.13), we will derive a contradiction from this assumption.

Claim 5.5.9. We have $|S \cap V(\mathcal{M}_A)| \leqslant |\mathbb{XA} \cap V(\mathcal{M}_A)|$.

Proof of Claim 5.5.9. To see this, recall that each $\mathcal{M}_A$-vertex $X \in \mathcal{V}(\mathcal{M}_A)$ is either contained in S, or in L. Further, if $X \subseteq S$ then its partner in $\mathcal{M}_A$ must be in L, as S is independent. Now, the claim follows after noticing that $L \cap V(\mathcal{M}_A) = \mathbb{XA} \cap V(\mathcal{M}_A)$. □

Claim 5.5.10. We have $|S \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)| + 2\eta n < |\mathbb{XA} \setminus V(\mathcal{M}_A)| + \eta n/3$.

Proof of Claim 5.5.10. As $G \in \mathbf{LKS}(n, k, \eta)$, we have $|S| + 2\eta n \leqslant |L|$. Therefore,

$$|S \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)| + 2\eta n \leqslant |L \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)| + \sum_{\substack{(X_1, X_2) \in \mathcal{M}_A \cup \mathcal{M}_B \\ X_1 \cup X_2 \subseteq L}} |X_1 \cup X_2|$$

$$\overset{(5.26)}{=} |\mathbb{XA} \setminus V(\mathcal{M}_A)| + |V(\mathcal{M}_{\text{good}})|$$

$$\overset{\neg \mathbf{(K2)}}{<} |\mathbb{XA} \setminus V(\mathcal{M}_A)| + \eta n/3 \,.$$

□

29

Claim 5.5.11. We have $e_{G_\nabla}\left(\mathbb{X}\mathbb{A} \cap (\mathbb{E} \cup V(\mathcal{M})), S^{\mathrm{R}} \setminus V(\mathcal{M}_A)\right) < \eta k n / 2$.

Proof of Claim 5.5.11. As

$$\mathbb{X}\mathbb{A} \cap (\mathbb{E} \cup V(\mathcal{M})) \subseteq ((L \cap \mathbb{E}) \cup V_2(\mathcal{M})) \setminus V_2(\mathcal{M}_B) \quad \text{and}$$
$$S^{\mathrm{R}} \setminus V(\mathcal{M}_A) \subseteq \left((S^{\mathrm{R}} \setminus V(\mathcal{N}_1)) \cup V_1(\mathcal{M})\right) \setminus V_1(\mathcal{M}_1) \,,$$

we get from (III) that

$$e_{G_\mathcal{D}}\left(\mathbb{X}\mathbb{A} \cap (\mathbb{E} \cup V(\mathcal{M})), S^{\mathrm{R}} \setminus V(\mathcal{M}_A)\right) \leqslant \gamma k n \,. \tag{5.27}$$

Observe now that both sets $\mathbb{X}\mathbb{A} \cap (\mathbb{E} \cup V(\mathcal{M}))$ and $S^{\mathrm{R}} \setminus V(\mathcal{M}_A)$ avoid $\mathbb{H}$. Further, no edges between them belong to $G_{\exp}$, because Claim 5.5.1 implies that $S^{\mathrm{R}} \setminus V(\mathcal{M}_A) \subseteq S^0 \subseteq V(G) \setminus V(G_{\exp})$. Therefore, we can pass from $G_\mathcal{D}$ to G_∇ in (5.27) to get

$$e_{G_\nabla}\left(\mathbb{X}\mathbb{A} \cap (\mathbb{E} \cup V(\mathcal{M})), S^{\mathrm{R}} \setminus V(\mathcal{M}_A)\right) \leqslant \gamma k n < \eta k n / 2 \,.$$

□

Claim 5.5.12. We have $S \setminus (S^{\mathrm{R}} \cup V(\mathcal{M}_A)) \subseteq S \setminus (\bar{S}^0 \cup V(\mathcal{M}_A \cup \mathcal{M}_B))$.

Proof of Claim 5.5.12. The claim follows directly from the following two inclusions.

$$S^{\mathrm{R}} \cup V(\mathcal{M}_A) \supseteq S \cap V(\mathcal{M}_A \cup \mathcal{M}_B) \,, \text{ and} \tag{5.28}$$
$$S^{\mathrm{R}} \cup V(\mathcal{M}_A) \supseteq \bar{S}^0 \,. \tag{5.29}$$

Now, (5.28) is trivial, as by (II) we have that $S^{\mathrm{R}} \supseteq S \cap V(\mathcal{M}_B)$. To see (5.29), it suffices by (5.13) to prove that $V(N_1 \setminus M) \cup \mathbf{S}^{\mathrm{R}} \supseteq \mathbf{S}^0$. This is however the assertion of Claim 5.5.5. □

Next, we bound $e_{G_\nabla}(\mathbb{X}\mathbb{A}, S)$.

Claim 5.5.13. We have

$$e_{G_\nabla}(\mathbb{X}\mathbb{A}, S) \leqslant |S \cap V(\mathcal{M}_A)|(1 + \eta)k + |S \setminus (S^0 \cup V(\mathcal{M}_A \cup \mathcal{M}_B))|(1 + \eta)k + \frac{1}{2}\eta k n \,.$$

Proof of Claim 5.5.13. We have

$$e_{G_\nabla}(\mathbb{X}\mathbb{A}, S) = e_{G_\nabla}(\mathbb{X}\mathbb{A}, S \cap V(\mathcal{M}_A)) + e_{G_\nabla}(\mathbb{X}\mathbb{A}, S \setminus (S^{\mathrm{R}} \cup V(\mathcal{M}_A)))$$
$$+ e_{G_\nabla}(\mathbb{X}\mathbb{A} \setminus (\mathbb{E} \cup V(\mathcal{M})), S^{\mathrm{R}} \setminus V(\mathcal{M}_A)) + e_{G_\nabla}(\mathbb{X}\mathbb{A} \cap (\mathbb{E} \cup V(\mathcal{M})), S^{\mathrm{R}} \setminus V(\mathcal{M}_A)) \,.$$

To bound the first term we use that each vertex in $S \cap V(\mathcal{M}_A)$ has degree at most $(1 + \eta)k$, and thus obtain $e_{G_\nabla}(\mathbb{X}\mathbb{A}, S \cap V(\mathcal{M}_A)) \leqslant |S \cap V(\mathcal{M}_A)|(1 + \eta)k$. To bound the second term, we again use a bound on degree of vertices of $S \setminus ((S^{\mathrm{R}} \cup V(\mathcal{M}_A)) \cup (S^0 \setminus \bar{S}^0))$, together with Claim 5.5.12. The third term is zero by Claim 5.5.3. The fourth term can be bounded by Claim 5.5.11. □

A relatively short double counting below will lead to the final contradiction. The idea behind this computation is given in Figure 5.4.
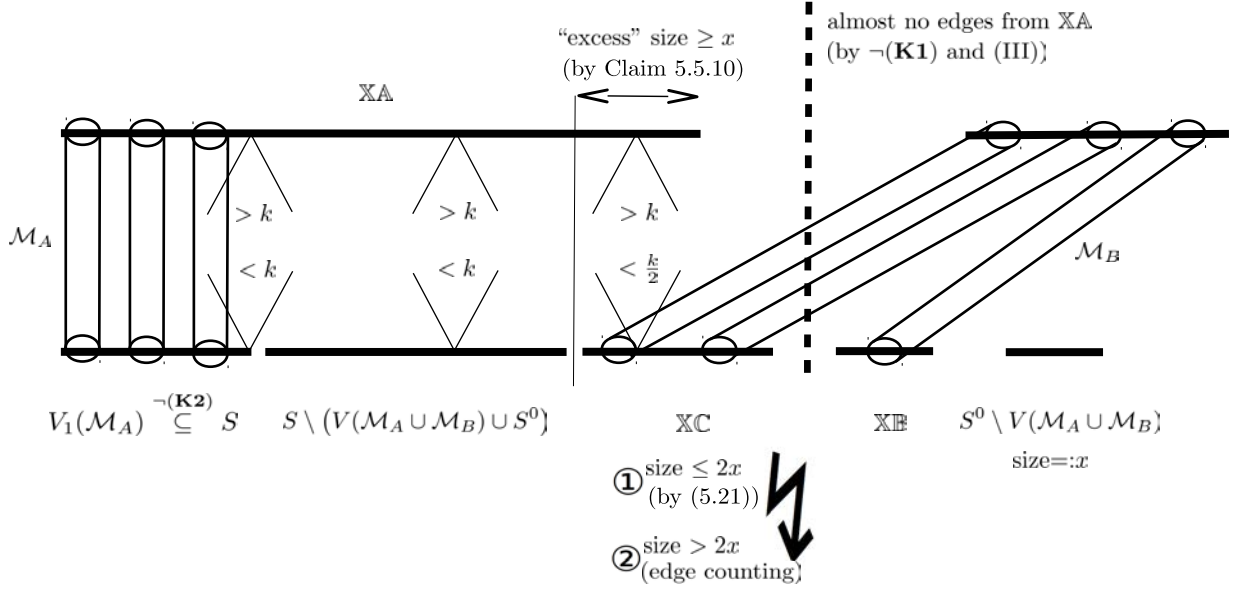
Figure 5.4: A simplified computation showing that ¬(**K1**), ¬(**K2**) leads to a contradiction. Denoting by x the size of $S^0 \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)$ we get ① $|\mathbb{XC}| \leqslant 2x$. On the other hand, each vertex of $\mathbb{XA}$ emanates $\gtrsim k$ edges which are absorbed by the sets $V_1(\mathcal{M}_A)$, $S \setminus (V(\mathcal{M}_A \cup \mathcal{M}_B) \cup S^0)$, and $\mathbb{XC}$. The vertices of $V_1(\mathcal{M}_A)$ and $S \setminus (V(\mathcal{M}_A \cup \mathcal{M}_B) \cup S^0)$ can absorb $\lesssim k$ edges. The vertices of $\mathbb{XC}$ receive $\lesssim \frac{k}{2}$ edges of $\mathbb{XA}$ by (5.24). This leads to ② $|\mathbb{XC}| > 2x$, doubling the size of the "excess" vertices of $\mathbb{XA}$.

$$
\begin{aligned}
|\mathbb{XA}|(1+\eta)k &\leqslant \sum_{v \in \mathbb{XA}} \deg_G(v) \leqslant \sum_{v \in \mathbb{XA}} \deg_{G_\nabla}(v) + 2\big(e(G) - e(G_\nabla)\big) \\
&\leqslant 2e_{G_\nabla}(\mathbb{XA}) + e_{G_\nabla}(\mathbb{XA}, \mathbb{XB}) + e_{G_\nabla}\big(\mathbb{XA}, \mathbb{XC}\big) + e_{G_\nabla}\big(\mathbb{XA}, S\big) + \frac{\eta k n}{3} \\
\text{\scriptsize (by ¬(\textbf{K1}), (5.25), C5.5.13)} \quad &\leqslant \frac{7}{6}\eta k n + \big|S^0 \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)\big|(1+\eta)k \\
&\quad + |S \cap V(\mathcal{M}_A)|(1+\eta)k \\
&\quad + |S \setminus (S^0 \cup V(\mathcal{M}_A \cup \mathcal{M}_B))|(1+\eta)k \\
\text{\scriptsize (by C5.5.9)} \quad &\leqslant \frac{7}{6}\eta k n + |S \setminus V(\mathcal{M}_A \cup \mathcal{M}_B)|(1+\eta)k \\
&\quad + |\mathbb{XA} \cap V(\mathcal{M}_A)|(1+\eta)k \\
\text{\scriptsize (by C5.5.10)} \quad &\leqslant \frac{7}{6}\eta k n + \big(|\mathbb{XA} \setminus V(\mathcal{M}_A)| - \frac{5}{3}\eta n\big)(1+\eta)k \\
&\quad + |\mathbb{XA} \cap V(\mathcal{M}_A)|(1+\eta)k \\
&< |\mathbb{XA}|(1+\eta)k - \frac{1}{2}\eta k n \, ,
\end{aligned}
\tag{5.30}
$$

a contradiction. This completes the proof of Lemma 5.4. $\qquad\square$

6 Acknowledgements

Symbol index

$[n]$, 1
$\mathrm{d}(U, W)$, 2
deg, 2
maxdeg, 2
mindeg, 2
$\mathcal{M}$-edge, 8
$E(G)$, 2
$e(G)$, 2
ℓ-ensemble, 2
$e(X)$, 2
$e(X, Y)$, 2
$\bar{\mathcal{G}}(n, k, \Omega, \rho, \nu)$, 5
$\mathcal{G}(n, k, \Omega, \rho, \nu, \tau)$, 5
$G_{\mathcal{D}}$, 5
$\mathbf{G}_{\mathrm{reg}}$, 5
G_{∇}, 5
$\mathbb{L}_{\eta,k}(G)$, 3
$\mathbf{LKS}(n, k, \eta)$, 3
$\mathbf{LKSsmall}(n, k, \eta)$, 3
$\mathrm{N}(v)$, 2
$\mathbb{S}_{\eta,k}(G)$, 3
$V_1(\mathcal{M})$, $V_2(\mathcal{M})$, $V(\mathcal{M})$, 8
$\mathcal{V}_1(\mathcal{M})$, $\mathcal{V}_2(\mathcal{M})$, $\mathcal{V}(\mathcal{M})$, 8
$\mathcal{M}$-vertex, 8
$V(G)$, 2
$v(G)$, 2
$\mathbb{X}\mathbb{A}(\eta, \nabla, \mathcal{M}_A, \mathcal{M}_B)$, 20
$\mathbb{X}\mathbb{B}(\eta, \nabla, \mathcal{M}_A, \mathcal{M}_B)$, 20
$\mathbb{X}\mathbb{C}(\eta, \nabla, \mathcal{M}_A, \mathcal{M}_B)$, 20
$\mathbf{trees}(k)$, 2

General index

References

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

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

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

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

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

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

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

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

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

[LP86]   L. Lovász and M. D. Plummer. *Matching theory*, volume 121 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1986. Annals of Discrete Mathematics, 29.

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

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

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