AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **89(2)** (2024), Pages 249–267

Regular decomposition of the edge set of a graph with applications

Béla Csaba*

Bolyai Institute
University of Szeged
Szeged, Hungary
bcsaba@math.u-szeged.hu

Abstract

We introduce a new method for decomposing the edge set of a graph, and use it to replace the regularity lemma of Szemerédi in some graph embedding problems. In particular, we prove a conditional triangle removal lemma, and also consider tree packing questions. An algorithmic version is also given.

1 Introduction

The Szemerédi regularity lemma is among the most powerful tools in graph theory. The lemma was published in 1978 [30], although a weaker version had already been used earlier by Szemerédi to prove the Erdős-Turán conjecture [14, 29]. Since then the lemma and its ramifications have found several applications in graph and hypergraph theory, number theory, algebra, geometry, and computer science.

We consider only simple graphs in this paper, i.e., no loops, no multiple edges. Given a graph $G = (V, E)$ and a number $\varepsilon \in (0, 1)$, the regularity lemma asserts that V can be partitioned into $k \leq N(\varepsilon)$ subsets $V_1 \cup \ldots \cup V_k$ and another set $V_0 = V - (\cup_i V_i)$ such that $G[V_i, V_j]$ is ε-regular (roughly speaking, this means approximate randomness; we will define this notion in the next section) for every $1 \leq i \neq j \leq k$, except at most εk^2 pairs of indices, $|V_0| \leq \varepsilon n$ and $|V_i| = (n - |V_0|)/k$ for every $1 \leq i \leq k$.

In Szemerédi's original proof of the regularity lemma, the threshold $N(\varepsilon)$ for the number of parts is a tower of twos of height $O(\varepsilon^{-5})$. As Gowers proved in [21], this tower-type bound is unavoidable in general. More precisely, there are graphs for which $N(\varepsilon)$ has to be at least a tower of twos of height $\Omega(\varepsilon^{-1/16})$. Conlon and Fox [5] further improved the lower bound to $\Omega(\varepsilon^{-1})$. This shows the major drawback

* This research was supported by project no. TKP2021-NVA-09, which has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.

ISSN: 2202-3518

of the regularity lemma: in order to get meaningful results, one has to work with enormously large graphs, thereby ruling out practical applications of the lemma. We remark that if the number of edges between vertex sets A and B is $o(|A| \cdot |B|)$, then the (A, B)-pair is ε-regular by definition, unless ε is very small compared to $|A|$ and $|B|$. Hence, the lemma is useful only when the density of the graph is essentially bounded from below by a positive constant.

In order to avoid these disadvantages, several alternatives for substituting the regularity lemma have been discovered, e.g., the weak regularity lemma of Frieze and Kannan [18], or the cylindrical regularity lemma by Eaton and Rödl [12, 13], see also in [1] by Alon, Duke, Leffmann, Rödl and Yuster. Gowers proved the existence of large quasirandom subgraphs in dense bipartite graphs [22], a somewhat similar one was given by the author [8]. There are also versions for dense subgraphs of random graphs or C_4-free graphs [7], for graphs with bounded VC-dimension [2, 26] or for graphs with low threshold-rank [19]. This list is long, but still far from being complete. However, none of the above results have the full strength of the original regularity lemma.

In this paper our goal is to present a method which in some cases may be used to replace the regularity lemma, even for graphs with vanishing densities. Besides, the number of vertices in the graph is allowed to be just "reasonably" large. In the decomposition of this paper one does not partition the vertex set of the graph as in the Szemerédi regularity lemma, rather the edge set. The basic building blocks of the decomposition are pseudorandom subgraphs, so called lower-regular or regular pairs. The most important difference between this method and the many versions of the regularity lemma is that the pseudorandom subgraphs in the decomposition of the present paper are edge-disjoint, but their vertex sets may intersect. These intersections make the use of the edge decomposition method more difficult, but it still can replace the regularity lemma in some cases. We will demonstrate this via some examples. Still, it seems that one cannot use this technique for such diverse problems as the regularity lemma.

Eaton and Rödl [12, 13] proved a so-called cylindrical regularity lemma (an algorithmic version was given in [1] by Alon, Duke, Leffmann, Rödl and Yuster) which for the special case of bipartite graphs gives a similar, though somewhat more restrictive, decomposition to the one presented in this paper. However, our bounds are stronger, due to the different approach, based on the results of Peng, Rödl and Ruciński [27][1]. This turns out to be important for graphs having vanishing densities. We will discuss the cylindrical regularity lemma together with the necessary notions in Section 3.1.

The outline of the paper is as follows. In Section 2 we review the necessary notation, notions, and results for our decompositions. Section 3 includes our main decomposition theorems. In Section 4 we prove a conditional triangle removal lemma for graphs having relatively few C_5's. Finally, in Section 5 we prove an algorithmic

[1]In an earlier version of this paper [9] we used the graph functional method of Komlós, which already gave stronger bounds, but it turned out that in general the results of [27] are slightly better. See the remark after Theorem 2.2.

version of our main theorem, and apply it for approximating an NP-complete problem in relatively dense graphs.

We made no attempts to optimize on the constants in the paper, and will not be concerned with floor signs and divisibility in the proofs. This makes the notation simpler, easier to follow. Throughout the paper $\log x$ will denote the *natural logarithm* of x, and $\exp(f) = e^f$ for any expression f.

2 Notation, definitions, main tools

Given a graph $G = (V, E)$ we use the notation $v(G) = |V|$ and $e(G) = |E|$. For disjoint subsets $X, Y \subset V$, we let $G[X, Y]$ denote the bipartite subgraph of G with parts X and Y that contains all the edges of G with one endpoint in X and the other endpoint in Y. For every vertex $v \in V$, the *neighborhood* of v is denoted by $N(v)$, and the *degree* of v is denoted by $\deg(v) = |N(v)|$. Given a set $S \subset V$, we let $N(v, S) = N(v) \cap S$ and $\deg(v, S) = |N(v, S)|$.

The *density* of G is defined to be $d_G = e(G) \cdot \binom{v(G)}{2}^{-1}$. The *bipartite density* of bipartite subgraphs of G with parts A and B is defined to be $d_G(A, B) = \frac{e(G[A,B])}{|A|\cdot|B|}$. Sometimes the subscript may be omitted when there is no confusion. Similarly, when a graph in question is bipartite, density will mean bipartite density.

Definition 2.1. Let $0 < \varepsilon, \delta < 1$ be real numbers. We say that a bipartite graph $H = (A, B; E)$ is an (ε, δ)-*lower-regular pair*, if

$$d_H(X, Y) \geq \delta$$

whenever $X \subseteq A$ and $Y \subseteq B$ such that $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$. We call H an (ε, δ)-*super-lower-regular pair*, if in addition every $v \in A$ has at least $\delta|B|$ neighbors and every $u \in B$ has at least $\delta|A|$ neighbors.

Let $\mathcal{B}$ denote the class of balanced bipartite graphs, that is, bipartite graphs having equal-sized parts, and for a positive integer m let $\mathcal{B}_m$ denote the class of balanced bipartite graphs having m vertices in both parts.

Peng, Rödl and Ruciński proved the theorem below in [27] which plays a crucial role in the first decomposition result of the present paper.

Theorem 2.2. *Let $0 < \varepsilon, d < 1$ be two numbers, and assume that $G \in \mathcal{B}_n$ is a graph with density d. Then G contains an $(\varepsilon, d/2)$-lower-regular pair $H \in \mathcal{B}_m$ with*

$$m \geq \frac{1}{2} d^{12/\varepsilon} n.$$

As it is also proved in [27], Theorem 2.2 is essentially tight; in the above lower bound for m the term $d^{12/\varepsilon}$ cannot be replaced by $d^{c/\varepsilon}$ if $c < 1/2000$. Let us remark that a slightly weaker bound of $m \geq d^{3(\log 1/\varepsilon)/\varepsilon} n$ was proved by Komlós [24, 25] using the graph functional method.

We also need another notion, already mentioned in the introduction.

Definition 2.3. Let $0 < \varepsilon, \delta < 1$ be real numbers. We say that a bipartite graph $H = (A, B; E)$ is an *ε-regular pair*, if

$$|d_H(A,B) - d_H(X,Y)| \leq \varepsilon$$

whenever $X \subseteq A$ and $Y \subseteq B$ such that $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$. We call H an *(ε, δ)-super-regular pair*, if in addition every $v \in A$ has at least $\delta|B|$ neighbors and every $u \in B$ has at least $\delta|A|$ neighbors.

Note the difference between ε-lower-regularity and ε-regularity: in the latter we have about the same density of every sufficiently large bipartite subgraph.

For our second decomposition theorem we need another result from [27].

Theorem 2.4. *Let $0 < \varepsilon, d < 1$. Then every bipartite graph $G \in \mathcal{B}_n$ with $d_G = d$ contains an ε-regular pair $H \in \mathcal{B}_m$ with density not smaller than $(1 - \varepsilon/3)d$ and*

$$m \geq d^{50/\varepsilon^2} \frac{n}{2}.$$

The lemma below shows that every lower-regular pair contains a slightly smaller super-lower-regular pair, and similarly, every regular pair contains a slightly smaller super-regular pair.

Lemma 2.5. *Let $0 < \varepsilon < 1/2$ and $F \in \mathcal{B}_m$.*

(i) Assume that $0 < \delta < 1$ and F is an (ε, δ)-lower-regular pair with density d_F. Then there exists $H \subset F$, $H \in \mathcal{B}_{m'}$ such that H is a $(2\varepsilon, \delta - \varepsilon)$-super-lower-regular with $m' \geq (1 - \varepsilon)m$; moreover, $e(H) \geq (1 - 3\varepsilon)e(F)$.

(ii) Let us assume now that F is an ε-regular pair with density $d_F \geq \varepsilon$. Then there exists $H \subset F$, $H \in \mathcal{B}_{m'}$ such that H is a $(2\varepsilon, d_F - \varepsilon)$-super-regular with $m' \geq (1 - \varepsilon)m$.

Proof: We first prove (i). Let A and B denote the vertex parts of F. Let $X \subset A$ and $Y \subset B$ denote the sets of those vertices which have degree smaller than δm in the opposite part. By definition of (ε, δ)-lower-regularity, we have $|X|, |Y| \leq \varepsilon m$. Without loss of generality we may assume that $|X| \geq |Y|$.

Denote the vertices of B by $\{v_1, \ldots, v_m\}$, and assume that $\deg(v_1) \leq \deg(v_2) \leq \cdots \leq \deg(v_m)$. Then we must have $\deg(v_{\varepsilon m}) \leq d_F m/(1 - \varepsilon)$, or otherwise B had at least $(1-\varepsilon)m$ such vertices which have degree larger than $d_F m/(1-\varepsilon)$, which is clearly a contradiction. Using the fact that $|Y| \leq |X| \leq \varepsilon m$, we have $Y \subseteq \{v_1, \ldots, v_{|X|}\}$ and $\deg(v_{|X|}) \leq d_F m/(1 - \varepsilon) < 2d_F m$.

We discard the vertices of X from A and the vertices $v_1, \ldots, v_{|X|}$ from B. Call the remaining subgraph H. It is easy to see that $H \in \mathcal{B}_{m'}$ is a $(2\varepsilon, \delta - \varepsilon)$-super-lower-regular pair with $m' \geq (1 - \varepsilon)m$.

Now we prove the lower bound for $e(H)$. Observe that $\delta \leq d_F$ by definition of lower-regularity. Hence, the total number of edges that were incident to vertices of $X \cup Y$ in F is less than

$$\delta m|X| + 2d_F m|X| \leq 3\varepsilon d_F m^2 = 3\varepsilon e(F),$$

implying that $e(H) \geq (1 - 3\varepsilon)e(F)$.

Since one can prove (ii) very similarly, we leave the details for the reader. $\square$

3 Decomposition theorems

We are ready to present the first decomposition theorem.

Theorem 3.1. *Let $G = (V, E)$ be a balanced bipartite graph on $2n$ vertices with density d_G, and let $0 < d < 1$ and $0 < \varepsilon < 1/2$ such that $n > \exp(12 \log(1/d)/\varepsilon)$. Then there exist natural numbers $m = m(\varepsilon, d)$ and $K = K(\varepsilon, d)$ such that the following holds: $E(G)$ can be written as the edge-disjoint union of bipartite graphs $H_1, \ldots, H_K \in \mathcal{B}$, and another,* exceptional *graph H_0, such that for every $i \geq 1$, H_i is an (ε, δ_i)-lower-regular pair with $m_i \geq m$ vertices in both parts, and $\delta_i \geq d/2$, while H_0 has density less than d. Furthermore,*

$$m \geq \frac{1}{2} d^{12/\varepsilon} n \quad \textit{and} \quad K \leq 8 \frac{d_G}{d} \cdot d^{-24/\varepsilon}.$$

Proof: For finding the decomposition of G we apply Theorem 2.2 repeatedly. In the first step, we check whether $d_G \geq d$. If not, we let $H_0 = G$, and stop. Otherwise, let H_1 be a balanced (ε, δ_1)-lower-regular pair which is provided by Theorem 2.2; here $\delta_1 = d_G/2$. Let $G_1 = (V, E_1)$ be the subgraph of G which we obtain by deleting the edges of H_1 from G; that is, $E_1 = E(G) - E(H_1)$.

In general, let us assume that we have found the first i lower-regular pairs $H_1, H_2, \ldots, H_i$. Let $G_i = (V, E_i)$, where $E_i = E(G_{i-1}) - E(H_i)$. We check whether $d_{G_i} \geq d$. If not, then we let $H_0 = G_i$, and stop. Otherwise we use Theorem 2.2 for finding H_{i+1}.

Next we consider the bounds for m and K. Note that the function $f(d) = d^{12/\varepsilon} n/2$ is monotone increasing. When finding the decomposition, we always apply Theorem 2.2 for graphs which have density at least d. Hence $m_i \geq f(d) = d^{12/\varepsilon} n/2$ for every $1 \leq i \leq K$. For the upper bound for K, notice that $e(H_i) \geq dm^2/2$ for every $i \geq 1$. Hence

$$K \leq 2 \frac{d_G n^2}{dm^2} \leq 8 \frac{d_G}{d} \cdot d^{-24/\varepsilon},$$

as desired. $\square$

For easier reference we formulate another version of the above decomposition theorem, which follows easily from the first one. In the theorem below we decompose the graph into super-lower-regular pairs, which we obtain from lower-regular pairs by applying Lemma 2.5.

Theorem 3.2. *Let $G = (V, E)$ be a balanced bipartite graph on $2n$ vertices with density d_G, and let $0 < d < 1$ and $0 < \varepsilon \le d/3$ such that $n > \exp(24 \log(1/d)/\varepsilon)$. Then there exist natural numbers $m = m(\varepsilon, d)$ and $K = K(\varepsilon, d)$ such that the following holds: $E(G)$ can be written as the edge-disjoint union of bipartite graphs $H_1, \ldots, H_K \in \mathcal{B}$, and another* exceptional *graph H_0, such that for every $i \ge 1$, H_i is an (ε, δ_i)-super-lower-regular with $m_i \ge m$ vertices in both parts, and $\delta_i \ge d/3$, while H_0 has density less than $2d$. Furthermore,*

$$m \ge \frac{1}{3} d^{24/\varepsilon} n \quad \text{and} \quad K \le 8 \frac{d_G}{d} \cdot d^{-48/\varepsilon}.$$

Proof: First we apply Theorem 3.1 to G with parameters $\varepsilon_0 = \varepsilon/2$ and d. We obtain the lower-regular pairs $H'_1, \ldots, H'_K$ and the exceptional subgraph H'_0, where H'_i is an $(\varepsilon_0, \delta'_i)$-lower-regular pair with m_i vertices in both parts for $i \ge 1$, and H'_0 has density at most d. Next we apply Lemma 2.5 for every H'_i $(i \ge 1)$ in order to turn it to the super-lower-regular H_i. We obtain an (ε, δ_i)-super-lower-regular pair, where, using the fact that that $\varepsilon_0 \le d/6$ and $\delta'_i \ge d/2$, we obtain $\delta_i \ge d/3$. We discarded at most $\varepsilon_0 m_i < m_i/6$ vertices from both parts of H_i; this explains the bound for m. Moreover, by Lemma 2.5 we have $e(H_i) \ge (1 - 3\varepsilon_0) e(H'_i)$. Since $\sum_i e(H'_i) \le e(G)$, the total number of edges lost, when turning the lower-regular pairs into super-lower-regular, is at most $3\varepsilon_0 e(G) = 3\varepsilon_0 d_G n^2 < dn^2$; here we used $d_G \le 1$ and $\varepsilon_0 \le d/6$. This implies that at the end H_0 has density less than $2d$. It is clear that the number K of pairs in the decomposition does not change. □

Instead of iteratively applying Theorem 2.2 as in the above theorems, one can also apply the same iterative scheme with Theorem 2.4 to obtain ε-regular pairs rather than ε-lower-regular.

Theorem 3.3. *Let $G = (V, E)$ be a balanced bipartite graph on $2n$ vertices with density d_G, and let $0 < d < 1$ and $0 < \varepsilon < 1/2$ such that $n > \exp(50 \log(1/d)/\varepsilon^2)$. Then there exist natural numbers $m = m(\varepsilon, d)$ and $K = K(\varepsilon, d)$ such that the following holds: $E(G)$ can be written as the edge-disjoint union of bipartite graphs $H_1, \ldots, H_K \in \mathcal{B}$, and another,* exceptional*, graph H_0, such that for every $i \ge 1$, H_i is an ε-regular pair with $m_i \ge m$ vertices in both parts and density $d_i \ge (1 - \varepsilon/3)d$, while H_0 has density less than d. Furthermore,*

$$m \ge \frac{1}{2} d^{50/\varepsilon^2} n \quad \text{and} \quad K \le 8 \frac{d_G}{d} \cdot d^{-100/\varepsilon^2}.$$

Proof: For finding the decomposition of G we apply Theorem 2.4 repeatedly, and similarly to the previous decompositions, we finally arrive at the ε-regular pairs $H_1, H_2, \ldots, H_K$, and the exceptional subgraph H_0, the latter having density at most d.

We obtain the bound for m from Theorem 2.4, with d plugged in for the density of the graph. For the upper bound for K, notice that $e(H_i) \ge dm^2/2$ for every $i \ge 1$. Hence

$$K \;\le\; 2 \frac{d_G n^2}{dm^2} \;\le\; 8 \frac{d_G}{d} \cdot d^{-100/\varepsilon^2},$$

which was desired. □

3.1 Comparison with the cylindrical regularity lemma

In this subsection we compare the bounds of the so called cylindrical regularity theorems for the case of bipartite graphs in [1, 12, 13] with the bounds in our decomposition theorems.

Let $k \geq 2$ be an integer, and $V_1, \dots, V_k$ be n-element sets. A *cylinder* of $V_1 \times V_2 \times \cdots \times V_k$ is of the form $W_1 \times W_2 \times \cdots \times W_k$, $W_i \subseteq V_i$ for $i = 1, 2, \dots, k$. Let G be a k-partite graph with parts $V_1, \dots, V_k$. Given an $\varepsilon > 0$ we say that the cylinder $W_1 \times \cdots \times W_k$ is *ε-regular* if the subgraph of G induced on the set $\cup_{1 \leq i \leq k} W_i$ is such that all $\binom{k}{2}$ of the pairs (W_i, W_j), $1 \leq i < j \leq k$, are ε-regular.

The cylindrical regularity lemma is as follows.

Lemma 3.4. *Let G be a k-partite graph with parts $V_1, \dots, V_k$ such that $|V_1| = \cdots = |V_k| = N$. Then for every $\varepsilon > 0$ there exists a partition $\mathcal{P}$ of $V_1 \times \cdots \times V_k$ into q cylinders such that*

- *$q \leq 4^h$, where $h \leq \binom{k}{2}/\varepsilon^5$, and*
- *all but εN^k of the k-tuples $(v_1, v_2, \dots, v_k)$, $v_i \in V_i$, $i = 1, 2, \dots, k$, are in ε-regular cylinders of $\mathcal{P}$.*

For bipartite graphs, that is, when $k = 2$, the upper bound for the number of cylinders is $4^{1/\varepsilon^5}$, which is much larger than the bounds in our decomposition theorems, whenever ε is small: the bounds for K in Theorems 3.1, 3.2, and 3.3 are $\exp(O(1/\varepsilon))$, $\exp O(1/\varepsilon)$, and $\exp(O(1/\varepsilon^2))$, respectively. This extends the applicability of regularity methods for graphs with vanishing densities. For example, let $G \in \mathcal{B}_n$ with density $d_G = 1/\sqrt{\log n}$. One can apply Theorem 3.1 to G with $d = d_G$ and $\varepsilon = d/2$, obtaining a non-trivial decomposition of G. On the other hand, the bound for the number q of cylinders in Lemma 3.4 is $4^{1/\varepsilon^5} = 4^{(\log n)^{5/2}} > n^2$, clearly larger than $e(G)$. Hence, for this graph the cylindrical regularity lemma may not give a useful decomposition.

Note that in the decomposition theorems of this paper we do not require that the Cartesian product of the vertex parts of G has to be partitioned by the Cartesian products of the vertex parts of the H_i graphs (here $i \geq 1$).

The cylindrical regularity lemma has an algorithmic version [1]. In Section 5 we present an algorithmic version of Theorem 3.3, and compare it with the algorithmic cylindrical regularity lemma.

3.2 Decomposition when G is not bipartite

Our decomposition theorems are formulated for balanced bipartite graphs. In this subsection, we present two methods which allow us to decompose arbitrary graphs.

The following well-known result is folklore; we omit the proof.

Lemma 3.5. *Let $G = (V, E)$ be an n-vertex graph with density d_G. Then there exists $X \subset V$, $|X| = \lfloor n/2 \rfloor$, such that $d_G(X, V - X) \geq d_G$.*

Suppose that G is not bipartite. Then by Lemma 3.5 its vertex set $V(G)$ has a bipartition $V(G) = X \cup Y$ such that $||X| - |Y|| \le 1$, and the density of $G[X,Y]$ is at least as large as that of G. Apply any of the decomposition theorems for $G[X,Y]$, delete the edges of the non-exceptional subgraphs from G, and repeat the partitioning of Lemma 3.5, until the density of what is left becomes sufficiently small.

Later we will discuss an algorithmic decomposition theorem, again for balanced bipartite graphs. The following method can be used to extend it for non-bipartite graphs. We need the following result; it is Theorem 16.1.2 from [3].

Theorem 3.6. *Let A be a $k \times k$ matrix of reals with every entry from the $[-1,1]$ interval. Then there exists a deterministic polynomial time algorithm that finds numbers $\kappa_1, \ldots, \kappa_k \in \{-1,1\}$ such that for every i $(1 \le i \le k)$ we have*

$$\left| \sum_{j=1}^{k} A_{ij}\kappa_j \right| \le \sqrt{2k\log(2k)}.$$

Proposition 3.7. *Let $\mu \in (0,1)$ be a number and $G = (V,E)$ be a graph on $n \ge 5$ vertices with density $d_G \ge 2\sqrt{\log n}/(\mu\sqrt{n})$. Then there exists a deterministic polynomial time algorithm which finds a bipartition $V = X \cup Y$ such that $||X| - n/2| \le \sqrt{n\log n}$ and the density of $G[X,Y]$ is at least $d_G(1-\mu)$.*

Proof: Let M be the $(n+1)\times(n+1)$ matrix which we obtain in the following way. Add an all 1 row as the $(n+1)$st row to the $n \times n$ adjacency matrix of G, and then add an all 0 column (containing $n+1$ elements). Clearly, M is an $(n+1)\times(n+1)$ matrix.

We use Theorem 3.6 with $k = n+1$ and $A = M$. Set $\sigma = \sqrt{(2n+2)\log(2n+2)}$. The desired bipartition is determined by the κ_i numbers: if $\kappa_i = 1$, then the ith vertex of G belongs to X; otherwise it belongs to Y.

The inequalities of Theorem 3.6 imply that the inner product of the last row of M and the vector $\kappa^T = (\kappa_1, \ldots, \kappa_{n+1})$ is at most σ; hence, $||X| - |Y|| \le \sigma$.

Let t denote $n/2 - |X|$. Then $|X| = n/2 - t$ and $|Y| = n/2 + t$, and $||X| - |Y|| = |2t| \le \sigma$. Hence we have

$$\frac{n}{2} - \frac{\sigma}{2} \ \le\ |X|,\ \ |Y| \ \le\ \frac{n}{2} + \frac{\sigma}{2}.$$

Let $v \in V$ be an arbitrary vertex of G. Using Theorem 3.6 and the definition of X and Y, we obtain $|\deg(v,X) - \deg(v,Y)| \le \sigma$. Very similarly to the above calculations we obtain the following upper and lower bounds:

$$\frac{\deg(v)}{2} - \frac{\sigma}{2} \ \le\ \deg(v,X), \quad \deg(v,Y) \ \le\ \frac{\deg(v)}{2} + \frac{\sigma}{2}.$$

Finally, we prove the lower bound for the density of $G[X,Y]$ as follows:

$$\frac{e(G[X,Y])}{|X|\cdot|Y|} \ \ge\ \frac{1}{2}\frac{\sum_{v\in V}\left(\frac{\deg(v)-\sigma}{2}\right)}{\frac{n^2}{4}} \ \ge\ \frac{1}{n^2}\sum_{v\in V}\deg(v) - \frac{\sigma}{n}.$$

It is easy to see that $(\sum_{v\in V}\deg(v))/n^2 \geq d_G(1-1/n)$. Hence

$$\frac{1}{n^2}\sum_{v\in V}\deg(v) - \frac{\sigma}{n} \;\geq\; d_G(1-\frac{1}{n}) - \frac{\sigma}{n} \;\geq\; d_G - 2\sqrt{\frac{\log n}{n}} \geq (1-\mu)d_G,$$

where in the last inequality we used the lower bound $\frac{2}{\mu}\sqrt{\frac{\log n}{n}}$ for d_G. □

Having the above theorem, one can iteratively apply it, as previously in the non-algorithmic method, until the vast majority of the edges of G are in edge-disjoint pseudorandom subgraphs.

4 A conditional triangle removal lemma

The celebrated result of Ruzsa and Szemerédi [28] states, roughly speaking, that if a graph of order n has $o(n^3)$ triangles, then it can be made triangle-free by deleting $o(n^2)$ edges from it.

Theorem 4.1. *Let G be a graph on n vertices. Then for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that if G has at most δn^3 triangles, then it can be made triangle-free by removing at most εn^2 edges.*

This very important result has far reaching implications in graph and hypergraph theory, number theory, etc.; see for example in [6]. The best bound for δ, due to Fox [17], is the reciprocal of a tower of twos of height $O(\log 1/\varepsilon)$.

Bollobás and Győri raised the following question in [4]: How many triangles can a graph G on n vertices have, if G has no cycle of length 5? They proved that the number of triangles in such a graph is at most $(5/4)n^{3/2} + o(n^{3/2})$. Since then several improvements were found on the constant multiplier of $n^{3/2}$ [15, 16].

In [7], Conlon, Fox, Sudakov and Zhao considered a conditional removal problem, which relates the above two questions. They proved that, given a graph of order n with $o(n^2)$ copies of C_4 and $o(n^{5/2})$ copies of C_5, it can be made $\{C_3, C_5\}$-free by deleting $o(n^{3/2})$ edges. Below we state and prove a somewhat similar theorem in which we have no condition on the number of C_4's in the graph, only for the number of C_5's. Since C_4's emerge when the number of edges is $\Omega(n^{3/2})$, the bound for the number of edges to be removed is much larger. Still, the δ below is just a single exponential function of a polynomial of $1/\varepsilon$, a huge gain compared to the δ in the known proofs of the (unconditional) triangle removal lemma.

Theorem 4.2. *Let $0 < \varepsilon \leq 1/4$ be a number, and set $n_0 = \exp(24\log(1/(3\varepsilon))/\varepsilon)$. Let G be a tripartite graph with vertex parts X, Y and Z such that $|X| = |Y| = |Z| = n > n_0$. Set $K_0 = 16\cdot\exp((48/\varepsilon)\log(1/(3\varepsilon)))$, $m = \exp(24\log(3\varepsilon)/\varepsilon)n/3$, and $\delta = \varepsilon^6\exp(48\log(3\varepsilon)/\varepsilon)/(9K_0^2)$. If the number of C_5's in G is at most δn^5, then G can be made triangle-free by deleting at most $8\varepsilon n^2$ edges.*

Proof: We begin with applying Theorem 3.2 with parameters ε and $d = 3\varepsilon$ for the bipartite subgraph $G[X, Y]$. We obtain the $(\varepsilon, \varepsilon)$-super-lower-regular pairs $H_1, H_2, \ldots, H_K$, each having parts on at least m vertices, and H_0 with less than $2dn^2 = 6\varepsilon n^2$ edges. Delete the edges of H_0 from G. Note that $K \le K_0$.

Call a $C_5 = x_1y_1x_2y_2z$ *good*, if $x_1, x_2 \in X$, $y_1, y_2 \in Y$, $z \in Z$, and x_1y_1, y_1x_2, x_2y_2 belong to the same super-lower-regular pair of the decomposition. Observe that if $x'y'z'$ is a triangle in $G - H_0$ such that $x' \in X, y' \in Y$ and $z' \in Z$, then the $x'y'$ edge must belong to some H_i $(i \ge 1)$.

By assumption, G has at most $\delta n^5 = \varepsilon^6 m^2 n^3 / K_0^2$ good C_5's, so there are at most εn vertices in Z which appear in at least $\delta n^5/(\varepsilon n) = \varepsilon^5 m^2 n^2 / K_0^2$ good C_5's. Call such vertices of Z *bad*; the rest of Z are the *good* vertices.

Delete every edge from G that joins a bad vertex in Z to any vertex in X; this way we achieve that there are no triangles with a bad vertex from Z. We deleted at most εn^2 edges in this step.

Let $z \in Z'$ be any good vertex. Assume that it has at least $\varepsilon n/K$ neighbors in both parts of some H_i. Denote the parts of H_i by $X_i \subset X$ and $Y_i \subset Y$, and let $m_i = |X_i| = |Y_i| \ge m$. Notice that if $x \in X_i$, then by the $(\varepsilon, \varepsilon)$-super-lower-regularity it has at least εm_i neighbors in Y, and similar holds for any $y \in Y_i$. Hence, if $x \in X_i, y \in Y_i$ and $x, y \in N(z)$, then, using ε-lower-regularity, there are at least $\varepsilon^3 m_i^2$ paths of length 3 that connects x and y in H_i. Using our assumption that z has at least $\varepsilon n/K$ neighbors in X_i and Y_i, there are at least $\frac{\varepsilon^5 m^2 n^2}{K^2}$ such good C_5's which contain z and three edges from H_i. This contradicts the fact that z is good, so we conclude that for every good vertex $z \in Z$ there is no H_i $(1 \le i \le K)$ for which z has many neighbors in both parts of it.

Next we repeat the following for every good $z \in Z$, for every $1 \le i \le K$: if $\deg(z, X_i) \le \deg(z, Y_i)$, then delete all edges that join z to X_i; otherwise delete all the edges that join z to Y_i. This way we delete at most $n \cdot K \cdot \varepsilon n/K = \varepsilon n^2$ edges from G. Since after these deletions there is no good $z \in Z$ which is contained in any triangle, we have removed every triangle. $\square$

We remark that in [7] it was proved that there exist n-vertex graphs with $o(n^{2.442})$ C_5's that cannot be made triangle-free by deleting $o(n^{3/2})$ edges. Note the large gap between the two bounds for conditional removal.

The proof method of Theorem 4.2 can easily be generalized to prove statements of the following type. Let $2 \le k \le \ell$ be integers. If an n-vertex, $(2k-1)$-partite graph G has $o(n^{2\ell+1})$ copies of $C_{2\ell+1}$'s, then it can be made C_{2k-1}-free by deleting $o(n^2)$ edges.

5 Algorithmic aspects

The algorithmic version of Szemerédi's regularity lemma [1] by Alon, Duke, Leffmann, Rödl and Yuster proved to be very useful in many problems in computer science, e.g., by providing good approximation algorithms for several NP-complete questions. The methods we considered earlier in the paper can be used to show the existence of a large super-regular pair, but they are not capable of finding one efficiently. In this section we present a deterministic polynomial time algorithm for finding a large ε-regular subgraph in a sufficiently dense graph. This algorithm can then be used for decomposing the edge set of a graph into large ε-regulars pairs, similarly to Theorem 3.1, albeit the result will be somewhat weaker. The algorithmic decomposition theorem is as follows.

Theorem 5.1. *Let G be a balanced bipartite graph on $2n$ vertices with density d_G, and let $0 < \delta$ and $0 < \varepsilon < 1/16$ be real numbers such that the following is satisfied: $e^{r/4}2n^{-1/4} < \varepsilon < \delta \leq 1$, where $r = (16^3/\varepsilon^{12})\log(16/\varepsilon^4)$. Then using a polynomial time deterministic algorithm we can decompose the edge set of G as follows: $E(G)$ can be written as the edge-disjoint union of ε-regular balanced bipartite graphs $F_1, \ldots, F_K$, and another balanced bipartite graph F_0, where $K = K(\varepsilon, \delta, d_G) \leq d_G e^{2r}/\delta$. For $i \geq 1$ each F_i has at least $m = m(\varepsilon) \geq e^{-r}n$ vertices and density at least δ, while F_0 has density less than δ.*

For proving the correctness of Theorem 5.1 we need to make preparations.

Let $M(n)$ denote the time needed to multiply two $n \times n$ matrices with $0, 1$ entries over the integers (so $M(n) = O(n^{2.376})$). For proving an algorithmic version of the regularity lemma, the authors of [1], among other lemmas, used the following:

Lemma 5.2. *Let H be a bipartite graph with equal parts $|A| = |B| = n$. Let $2n^{-1/4} < \varepsilon < 1/16$. Then there is a $O(M(n))$ time deterministic algorithm which verifies that H is ε-regular, or finds two subsets, $A_1 \subset A$, $B_1 \subset B$, $|A_1|, |B_1| \geq \varepsilon^4 n/16$, such that $|d(A, B) - d(A_1, B_1)| \geq \varepsilon^4$. The algorithm can be parallelized and implemented in NC^1.*

We call the sets A_1 and B_1 the *witnesses of ε^4-irregularity.* Note that the cardinalities of the witnesses for irregularity could be much smaller than εn. As proved in [1], this is unavoidable unless $P = NP$. We need a lemma for presenting the decomposition algorithm, but before that we consider a simple fact, which will be used in the proof of the lemma.

Fact 5.3. *Let F be a bipartite graph with vertex parts X and Y such that $|X| \geq 2$. Then X contains vertices x_1 and x_2 such that $d_F(X, Y) \geq d_F(X - x_1, Y)$ and $d_F(X, Y) \leq d_F(X - x_2, Y)$.*

Proof: We choose x_1 and x_2 such that

$$\deg(x_1) = \max\{\deg(x) : x \in X\} \quad \text{and} \quad \deg(x_2) = \min\{\deg(x) : x \in X\}.$$

By the above we have that

$$\deg(x_1) \geq \frac{\sum_{x \in X - x_1} \deg(x)}{|X| - 1} \quad \text{and} \quad \deg(x_2) \leq \frac{\sum_{x \in X - x_2} \deg(x)}{|X| - 1}.$$

Hence,

$$\begin{aligned} e_F(X,Y) &\geq \sum_{x\in X-x_1} \deg(x) + \frac{\sum_{x\in X-x_1} \deg(x)}{|X|-1} \\ &= \frac{|X|}{|X|-1} \sum_{x\in X-x_1} \deg(x) \\ &= \frac{|X|}{|X|-1} e_F(X-x_1,Y). \end{aligned}$$

This implies that

$$d_F(X-x_1,Y) = \frac{e_F(X-x_1,Y)}{(|X|-1)|Y|} \leq \frac{e_F(X,Y)}{|X||Y|} = d_F(X,Y).$$

Very similarly, we can prove $d_F(X-x_2,Y) \geq d_F(X,Y)$; the details are left for the reader. □

Lemma 5.4. *Let H be a bipartite graph with vertex parts A, B such that $|A| = |B| = n$, and let $0 < \eta < d \leq 1$. Assume that the density of H is d, and that there exists $A_1 \subset A, B_1 \subset B$ such that $|A_1|, |B_1| \geq \eta n$, and $|d - d(A_1, B_1)| \geq \eta$. Then in polynomial time we can find $A' \subset A, B' \subset B$ with $|A'| = |B'| \geq \eta n$ such that $d(A', B') \geq d + \eta^3$.*

Proof: We begin with the case $d(A_1, B_1) \leq d - \eta$. Applying Fact 5.3 repeatedly, if necessary, we may assume that $|A_1| = |B_1| \leq (1-\eta)n$. Define the sets $A_2 = A - A_1$ and $B_2 = B - B_1$. Below we show that one of the densities $d(A_1, B_2)$, $d(A_2, B_1)$ or $d(A_2, B_2)$ must be at least $d + \eta^3$.

Suppose not. Then

$$e(H) = dn^2 < (d-\eta)|A_1| \cdot |B_1| + (d+\eta^3)(n^2 - |A_1| \cdot |B_1|).$$

This implies that

$$\eta|A_1| \cdot |B_1| < \eta^3 n^2 - \eta^3 |A_1| \cdot |B_1|.$$

Using the fact that $|A_1|, |B_1| \geq \eta n$, we arrived at a contradiction. Hence one of the subgraphs $H[A_1, B_2]$, $H[A_2, B_1]$ or $H[A_2, B_2]$ must have density at least $d + \eta^3$. Denoting the parts of the most dense of these subgraphs by A' and B', we are done with proving this case.

Let us now assume that $d(A_1, B_1) \geq d + \eta$. If $|A_1| = |B_1|$, we are done; otherwise apply Fact 5.3 for discarding vertices from the larger set while not decreasing the density. In both cases we obtain the pair of sets A', B' such that $d(A', B') \geq d + \eta$ and $|A'| = |B'| \geq \eta n$.

It is easy to see that the above subsets A', B' can be found in $O(n^2)$ time: we need to compute densities of at most four subgraphs, and order vertices according to their degrees. This completes the proof. □

Using Lemma 5.4, we can easily formulate a polynomial time algorithm which finds a large ε-regular subgraph in a graph. Let $G = (A, B; E)$ be a bipartite graph with $|A| = |B| = n$ and density d, and assume that $0 < \varepsilon < d \leq 1$, such that $\exp(\frac{4 \cdot 16^3 \log 1/\varepsilon}{\varepsilon^{12}}) 2n^{-1/4} < \varepsilon < 1/16$.

1. Apply Lemma 5.2. If $G[A, B]$ is ε-regular, stop.
2. If not, by Lemma 5.4, substituting $\varepsilon^4/16$ for η, we can find a balanced subgraph $G[A', B']$ with $|A'| = |B'| \geq \varepsilon^4 |A|/16$ having density at least $d(G[A, B]) + \varepsilon^{12}/16^3$.
3. Let $A = A',\ B = B'$, and continue with Step 1.

The above algorithm stops in at most $16^3/\varepsilon^{12}$ steps, since if the density of a bipartite graph is 1, it must be ε-regular. With the above we have proved the following.

Proposition 5.5. *Let $G = (V, E)$ be a balanced bipartite graph on $2n$ vertices with density d_G and let $0 < \varepsilon < 1/16$ be a real number such that the following is satisfied: $e^{r/4} 2n^{-1/4} < \varepsilon < d_G$, where $r = (16^3/\varepsilon^{12}) \log(16/\varepsilon^4)$. Then in polynomial time we can find an ε-regular subgraph $H \in \mathcal{B}$ of G with density $\delta \geq d_G$ and $v(H) \geq e^{-r} n$.*

Proof: (of Theorem 5.1) The proof follows easily by repeatedly applying Proposition 5.5. The bounds for m and K, the density bounds for the F_i $(i \geq 0)$ bipartite subgraphs can be obtained similarly to the proofs of Theorem 3.1 or 3.3. We leave the details for the reader. □

Note that in Proposition 5.5 we guarantee only ε-regularity, not (ε, δ)-super-regularity for some δ. If one needs super-regularity, similarly to the proof of Theorem 3.2, Lemma 2.5 can be used, resulting only in a small decrease in the size of the pairs. Of course, this remark also applies to Theorem 5.1.

An algorithmic version of the cylindrical regularity lemma can be found in [1]. Similarly for the non-algorithmic version discussed in Section 3.1 they consider k-partite graphs with $k \geq 2$. For bipartite graphs they prove the bound $\exp(O(\varepsilon^{-17}))$ for the number of ε-regular subgraphs in the decomposition; hence our bound of $\exp(O(\varepsilon^{-12} \log 1/\varepsilon))$ in Theorem 5.1 is a substantial improvement.

5.1 Algorithmic applications for packing problems

The study of packing of graphs dates back more than a century, and recently it has received much attention; see for example [10, 20]. The generic packing question is as follows: Given a "large" host graph G and a "small" graph H, is it possible to cover the edge set of G edge-disjointly by copies of H? Equivalently, we sometimes say that the edge set of G is decomposed by edge-disjoint copies of H. In many cases the single small graph H is replaced by a family of graphs. Perhaps one of the most beautiful questions in the area is the still open tree packing conjecture of

Gyárfás [23] (often referred to as the Gyárfás–Lehel conjecture) from 1976 on the decomposition of the edge set of K_n into trees having every order between 1 and n. In general even an approximate form of these problems, when we may leave a small percentage of the edge set of the host graph uncovered, is very challenging. In the rest of this section we consider algorithmic versions of such questions. We remark that substituting Proposition 5.5 by Theorem 2.4 in the following proofs we could get non-algorithmic results with better bounds.

5.1.1 Packing edge-disjoint 3-paths

In [11], Dreier, Fuchs, Hartmann, Kuinke, Rossmanith, Tauer, and Wang study the edge-disjoint k-path packing problem: given a graph G on n vertices find ℓ edge-disjoint copies of a path of length k (hence having $k+1$ vertices) in it. It is easy to see that in case $k = 1$ and $\ell = n/2$ the question is: find a perfect matching in G, if there exists any. This, and the $k = 2$, case can be solved in polynomial time. However, for $k = 3$ path packing already becomes an NP-complete question, as proved in [11]. Hence, it makes sense to consider an approximate version of the problem.

Below we give the details of an algorithm which finds a set of edge-disjoint 3-paths, covering all but a small proportion of the edges. First we need a simple lemma which is at the heart of the algorithm.

Lemma 5.6. *Let $0 < \varepsilon < d \le 1$ be real numbers, and assume that $H = (A, B; E)$ is an ε-regular pair with density at least d. Then we can find a 3-path in H by a deterministic polynomial time algorithm.*

Proof: Using the lower bound for the density of H we can choose vertices $u \in A$ and $v \in B$ such that $\deg(u) \ge d|B|$ and $\deg(v) \ge d|A|$. Since $d > \varepsilon$, using ε-regularity we have the following lower bound for the number of edges between $N(u)$ and $N(v)$:

$$e_H(N(u), N(v)) \ge (d-\varepsilon)d|A| \cdot d|B| = (d-\varepsilon)d^2|A| \cdot |B| > 0.$$

That is, there is at least one edge connecting $N(u)$ and $N(v)$, which can be found in $O(n^2)$ time, implying the existence of a polynomial time algorithm for finding the 3-path, as desired. □

Theorem 5.7. *Let $0 < d \le 1$ and $0 < \varepsilon < 1/16$ be real numbers, and $G \in \mathcal{B}_n$ with density $d_G \ge d$, such that the following is satisfied: $e^{r/4}2n^{-1/4} < \varepsilon < d$, where $r = (16^3/\varepsilon^{12})\log(16/\varepsilon^4)$. Then in deterministic polynomial time we can find a set of edge-disjoint 3-paths which covers at least $e(G) - dn^2$ edges of G.*

Proof: The following algorithm finds the desired path packing in G.

1. Apply Proposition 5.5 for G in order to obtain an ε-regular pair $H \subset G$.
2. Find a 3-path P in H using Lemma 5.6.

3. Delete the edges of P from G. For simplicity, the remaining graph is also called G.
4. If the density of G is at least d, then continue with step 1, otherwise stop.

It is easy to see that the above algorithm finds edge-disjoint 3-paths. Moreover, when the algorithm terminates, it leaves at most dn^2 edges uncovered. □

5.1.2 Packing large edge-disjoint trees

Observe that the algorithm of the proof of Theorem 5.7 generalizes for packing by subgraph F if we can substitute Lemma 5.6 by another result that finds a copy of F in an ε-regular pair. This also holds for super-regular pairs, since by Lemma 2.5 a regular pair contains a super-regular pair, with only slightly worse parameters.

Next we show that one can find large trees in an (ε, δ)-super-regular pair, which immediately implies an analogue of Theorem 5.7 for those trees. Depending on ε and δ, these trees may have linear size and linear maximum degree.

Let T be a tree on t vertices and denote its root by r. Given any $x \in V(T)$, let $N^*(x)$ denote the set of its children, and $\deg^*_T(x) = |N^*(x)|$. We define the L_i *level sets* of T as follows: $L_1 = \{r\}$, and for $i \geq 1$ we have that $L_{i+1} = \cup_{x \in L_i} N^*(x)$. Hence, if a vertex y lies in L_j, then y is at distance $j-1$ from r. Let us denote the total number of levels by s. We remark that the only upper bound imposed on the maximum degree of T is $\max\{|L_i| : 1 \leq i \leq s\}$.

Lemma 5.8. *Let T be a rooted tree on t vertices. Assume that $H(A, B; E)$ is an (ε, δ)-super-regular pair with vertex parts A and B, such that $m = |A| = |B| \geq 2t$, and $\delta \geq 6\varepsilon$. We further assume that $|L_i| \leq \delta m/4$ for every $i \geq 1$. Then we can find a copy of T in H using a deterministic polynomial time algorithm.*

Proof: Our goal is to find an edge-preserving injection $\varphi : V(T) \longrightarrow A \cup B$. Denote the root of T by r. The algorithm we use constructs φ step-by-step, alternately embedding consecutive levels of T, starting at $L_1 = \{r\}$.

At any point in time we denote the uncovered subset of A by A_u, and similarly, the uncovered subset of B by B_u. We need to define two more subsets:

$$A' = \{v \in A_u : \deg_H(v, B_u) \geq (\delta - \varepsilon)|B_u|\}$$

and

$$B' = \{v \in B_u : \deg_H(v, A_u) \geq (\delta - \varepsilon)|A_u|\}.$$

In the beginning we have $A_u = A' = A$ and $B_u = B' = B$. Observe that A_u, B_u, A' and B' shrink dynamically as we embed more and more levels. It is easy to see that $|A_u| \geq |A| - t \geq m/2$, and similarly, $|B_u| \geq |B| - t \geq m/2$.

After succesfully embedding, say, level L_i into A', we update these sets as follows: $A_u = A_u - \varphi(L_i)$, $A' = A' - \varphi(L_i)$, and we also have to leave out those vertices of B' that have degree less than $(\delta - \varepsilon)|A_u|$ in the newly updated A_u; the set B_u remains the same. We do the updating analogously when a level is embedded into B'.

For the root r we greedily pick a vertex $v \in A$ and let $\varphi(r) = v$. Now assume inductively, that we have embedded levels $L_1, \dots, L_i$ such that the following holds: if i is odd, then $\varphi(L_i) \subset A'$, otherwise $\varphi(L_i) \subset B'$. Without loss of generality we assume that i is odd, and prove that one can embed L_{i+1} into B'. The other case can be dealt with similarly.

We need the following claim.

Claim 5.9. *At any point in time during the embedding we have $|A'| \geq |A_u| - \varepsilon m$ and $|B'| \geq |B_u| - \varepsilon m$.*

Proof: Since the proofs of the two inequalities are essentially the same, below we prove the first one. Note that $d_H \geq \delta$, where d_H is the density of H. Let $X = A_u - A'$. Assume on the contrary that $|X| \geq \varepsilon m$. Then by ε-regularity of H the following is satisfied:

$$\left| \frac{e_H(X, B_u)}{|X| \cdot |B_u|} - d_H \right| \leq \varepsilon.$$

In particular,

$$\frac{e_H(X, B_u)}{|X| \cdot |B_u|} \geq d_H - \varepsilon \geq \delta - \varepsilon.$$

Multiplying by $|B_u|$ we get that there exists a vertex $v \in X$ having at least $(\delta - \varepsilon)|B_u|$ neighbors in B_u. This implies that v belongs to A', contradicting the definition of X. □

Let $x \in L_i$ be an arbitrary vertex. Since $\varphi(x) = v \in A'$, we have that

$$\deg_H(v, B_u) \geq (\delta - \varepsilon)|B_u| \geq (\delta - \varepsilon)\frac{m}{2} \geq \frac{\delta m}{4} + \varepsilon m,$$

where the last inequality follows from the condition $\delta \geq 6\varepsilon$.

Using Claim 5.9 we have

$$\deg(v, B') \ \geq \ \deg(v, B_u) - \varepsilon m \ \geq \ \frac{\delta m}{4} \geq |L_{i+1}|.$$

Hence, for every $x \in L_i$, we can greedily choose a subset $S_x \subset N(\varphi(x), B')$ such that $|S_x| = \deg^*_T(x)$, and $S_x \cap S_{x'} = \emptyset$ for every $x' \neq x$, $x' \in L_i$. Therefore we may choose the $\varphi(y)$ images for $y \in N^*(x)$ greedily from S_x. This extension of φ clearly preserves the edges of T going between L_i and L_{i+1}; moreover, $\varphi(L_{i+1}) \subset B'$. Hence, level by level, we can embed T, as desired. □

Theorem 5.10. *Let $0 < d \leq 1$ and $0 < \varepsilon < 1/16$ be real numbers, and $G \in \mathcal{B}_n$ with density $d_G \geq d \geq 13\varepsilon$, such that the following is satisfied: $e^{r/4}2n^{-1/4} < \varepsilon < d$, where $r = (16^3/\varepsilon^{12})\log(16/\varepsilon^4)$. Set $m = e^{-r}n(1-\varepsilon)$. Assume that $T_1, T_2, \dots, T_l$ are rooted trees, each on at most $t \leq m/2$ vertices such that every level set of each tree has at most $(d-\varepsilon)m/4$ vertices. Denote the total number of edges in the trees by $e_T = \sum_i e(T_i)$. If $e(G) - e_T \geq dn^2$, then we can find edge-disjoint copies of $T_1, \dots, T_l$ in G by a deterministic polynomial time algorithm.*

Proof: The proof of the theorem is very similar to the proof of Theorem 5.7. We use the following algorithm.

1. Let $i = 1$.
2. Apply Proposition 5.5 to G in order to obtain a balanced ε-regular pair F on at least $2m' = 2m/(1-\varepsilon)$ vertices with density at least d.
3. Apply Lemma 2.5 to F in order to obtain a balanced $(2\varepsilon, d-\varepsilon)$-super-regular pair H on at least $2m$ vertices.
4. Apply Lemma 5.8 to embed T_i in H.
5. Delete the edges of T_i from G. For simplicity, the remaining graph is also called G.
6. If $i < l$, then let $i = i + 1$, and continue with Step 2. Otherwise, if $i = l$, stop.

For proving the correctness of the above algorithm, note that ε and d are fixed throughout. Let us assume that we have successfully found edge-disjoint copies of the first $i-1$ trees. The density of what is left from G must still be at least d. This density is sufficiently large for finding an ε-regular pair F with required density and order by Proposition 5.5 in Step 2. Then, in Step 3, we find a $(2\varepsilon, d-\varepsilon)$-super-regular pair $H \subset F$ by Lemma 2.5. Since $d \geq 13\varepsilon$, we get that $d - \varepsilon \geq 6 \cdot 2\varepsilon$. Hence, in Step 4, we can apply Lemma 5.8 with parameters 2ε and $d - \varepsilon$, and find a copy of T_i. If $i < l$, then we can repeat this procedure, since, by definition, the density of what is left from G is at least d. □

Acknowledgments

The author wishes to thank Péter Hajnal and András Pluhár for their valuable suggestions, and anonymous referees, whose remarks substantially improved the presentation of the paper.

References

[1] N. Alon, R. Duke, H. Leffmann, V. Rödl and R. Yuster, The algorithmic aspects of the regularity lemma, *J. Algorithms* **16** (1994), 80–109.

[2] N. Alon, E. Fischer and I. Newman, Efficient testing of bipartite graphs for forbidden induced subgraphs, *SIAM J. Comput.* **37** (2007), 959–976.

[3] N. Alon and J. Spencer, "The probabilistic method", 4th ed., Wiley Ser. Discrete Math. Optim., 2016.

[4] B. Bollobás and, E. Győri, Pentagons vs. triangles, *Discrete Math.* **308** (2008), 4332–4336.

[5] D. Conlon and J. Fox, Bounds for graph regularity and removal lemmas, *GAFA* **22** (2012), 1191–1256

[6] D. Conlon and J. Fox, Graph removal lemmas, *Surveys in combinatorics* (2013), 1–50, https://doi.org/10.1017/CBO9781139506748.002.

[7] D. Conlon, J. Fox, B. Sudakov and Y. Zhao, The regularity method for graphs with few 4-cycles, *J. London Math. Soc.*, https://doi.org/10.1112/jlms.12500.

[8] B. Csaba, A new graph decomposition method for bipartite graphs, *Proc. MATCOS* (2019), 11–14.

[9] B. Csaba, Regular decomposition of the edge set of graphs with applications, https://arxiv.org/abs/2109.12394.

[10] B. Csaba, D. Kühn, A. Lo, D. Osthus and A. Treglown, Proof of the 1-factorization and Hamilton decomposition conjectures, *Mem. Amer. Math. Soc.* vol. 244, no. 1154, (2016).

[11] J. Dreier, J. Fuchs, T. Hartmann, P. Kuinke, P. Rossmanith, B. Tauer and H. Wang, The complexity of packing edge-disjoint paths, https://arxiv.org/pdf/1910.00440.pdf.

[12] N. Eaton, Ramsey numbers for sparse graphs, *Discrete Math.* **185** (1998), 63–75.

[13] N. Eaton and V. Rödl, "A Regularity Lemma", Tech. Report, Math. Department, Emory University, 1992.

[14] P. Erdős and P. Turán, On some sequences of integers, *J. London Math. Soc.* **11** (1936), 261-264.

[15] B. Ergemlidze, E. Győri, A. Methuku and N. Salia, A note on the maximum number of triangles in a C_5-free graph, *J. Graph Theory* **90** (2019), 227–230.

[16] B. Ergemlidze and A. Methuku, Triangles in C_5-free graphs and hypergraphs of girth six, *J. Graph Theory* **99** (2022), 26–39.

[17] J. Fox, A new proof of the graph removal lemma, *Ann. Math.* (2) 174 (2011), 561–579.

[18] A. Frieze and R. Kannan, Quick approximations to matrices and applications, *Combinatorica* **19** (1999), 175–220.

[19] S. Gharan and L. Trevisan, A new regularity lemma and faster approximation algorithms for low threshold rank graphs, in: *Approximation, Randomization, and Combinatorial Optimization*, Lec. Notes in Comput. Sci., 8096, Springer, Heidelberg, 2013, pp. 303–316.

[20] S. Glock, F. Joos, J. Kim, D. Kühn and D. Osthus, Resolution of the Oberwolfach problem, *J. Europ. Math. Soc.* 23 (8) (2021), 2511–2547.

[21] W. T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, *GAFA* **7** (1997), 322–337.

[22] W. T. Gowers, Bipartite graphs of approximate rank one, (preprint).

[23] A. Gyárfás and J. Lehel, Packing trees of different order into K_n, *Coll. Math. Soc. J. Bolyai* **18** (1976), 463–469.

[24] J. Komlós, Using a graph functional for finding a large quasirandom subgraph in a dense graph, unpublished manuscript.

[25] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, In: *Combinatorics, Paul Erdős is Eighty*, Vol. II (Eds.: D. Miklós, V. T. Sós and T. Szőnyi), János Bolyai Math. Soc., Budapest (1996), 295–352.

[26] L. Lovász and B. Szegedy, Regularity partitions and the topology of graphons, An Irregular Mind, (Eds.: I. Bárány, J. Solymosi and G. Sági), *Bolyai Soc. Math. Studies* **21** (2010), 415–446.

[27] Y. Peng, V. Röd and A. Ruciński, Holes in graphs, *Electron. J. Combin.* **9** (2002), #R1.

[28] I. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, *Combinatorics* (Keszthely, 1976), 18 (1978), Vol. II. pp. 939–945, North-Holland, Amsterdam–New York.

[29] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arithmetica* **27** (1975), 199–245.

[30] E. Szemerédi, Regular partitions of graphs Colloques Internationaux C.N.R.S. No 260 — *Problémes Combinatoires et Théorie des Graphe*s, Orsay (1976), 399–401.

(Received 30 May 2023; revised 4 Jan 2024, 10 Apr 2024)