

# Edge-disjoint induced subgraphs with given minimum degree

Raphael Yuster

Department of Mathematics  
University of Haifa  
Haifa 31905, Israel

`raphy@math.haifa.ac.il`

Submitted: Nov 9, 2012; Accepted: Feb 25, 2013; Published: Mar 8, 2013

Mathematics Subject Classifications: 05C70, 05C42

## Abstract

Let  $h$  be a given positive integer. For a graph with  $n$  vertices and  $m$  edges, what is the maximum number of pairwise edge-disjoint *induced* subgraphs, each having minimum degree at least  $h$ ? There are examples for which this number is  $O(m^2/n^2)$ . We prove that this bound is achievable for all graphs with polynomially many edges. For all  $\epsilon > 0$ , if  $m \geq n^{1+\epsilon}$ , then there are always  $\Omega(m^2/n^2)$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . Furthermore, any two subgraphs intersect in an independent set of size at most  $1 + O(n^3/m^2)$ , which is shown to be asymptotically optimal.

**Keywords:** minimum degree subgraph; edge packing; induced subgraph

## 1 Introduction

All graphs considered here are finite, undirected, and simple. For standard graph-theoretic terminology the reader is referred to [3]. Problems concerning edge-disjoint subgraphs that share some specified property are extensively studied in graph theory and many fundamental problems can be formulated in this way. For example, a proper edge coloring is just a set of pairwise edge-disjoint subgraphs, each having maximum degree at most one. A proper vertex coloring with  $k$  colors defines a set of  $\binom{k}{2}$  pairwise edge-disjoint induced bipartite subgraphs that cover all the edges. As another example, the problem of  $H$ -packing, where one seeks the maximum number of edge-disjoint subgraphs that are isomorphic to a given graph  $H$ , is a classical problem in graph theory and design theory. For some additional examples of results of this form see, e.g., [1, 4, 8, 9, 11, 12].

Observe that when we require the set of edge-disjoint subgraphs to be *induced* subgraphs, any two subgraphs can only intersect in an independent set, and it is natural to

add the additional requirement that any two such subgraphs have a small intersection. The main goal of this paper is to determine asymptotically the maximum number of edge-disjoint induced subgraphs that share the basic property of having a given minimum degree.

More formally, for a given integer  $h$ , let  $f_h(G)$  denote the maximum size of a set of pairwise edge-disjoint induced subgraphs of a graph  $G$ , each having minimum degree at least  $h$ . Trivially,  $f_1(G)$  equals the number of edges of  $G$ . However, it is not difficult to construct examples of graphs with  $n$  vertices and  $m$  edges for which already  $f_2(G) = O(m^2/n^2)$ . Our main result proves that this bound can always be attained for all graphs that have a polynomial number of edges, while, at the same time, keeping the intersection of any two subgraphs relatively small. We prove that for all fixed  $\alpha > 0$ , if  $m \geq n^{1+\alpha}$ , then  $f_h(G) = \Omega(m^2/n^2)$ . Furthermore, any two subgraphs in the given set intersect in only  $1 + O(n^3/m^2)$  vertices.

**Theorem 1.** *Let  $h$  be a positive integer and let  $\alpha$  be a positive real. There exists  $c = c(\alpha)$  and  $N = N(\alpha, h)$  such that any graph with  $n \geq N$  vertices and  $m \geq n^{1+\alpha}$  edges has at least  $cm^2/(h^2n^2)$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . Furthermore, any two such subgraphs intersect in an independent set of size at most  $1 + h^2n^3/(cm^2)$ .*

In the proof of Theorem 1 we make no attempt to optimize the constant  $c$ , although the proof shows that  $c$  is polynomial in  $\alpha$ . Also, observe that one can trivially replace “at least  $h$ ” with “precisely  $h$ ” in the statement of the theorem, by deleting vertices.

Recall that a graph is *balanced* if the average degree of each subgraph is not larger than the average degree of the graph. By carefully choosing constants and applying a minimality argument, the proof of Theorem 1 can be reduced to proving it for balanced graphs. A naive approach would then be to exploit the fact that balanced graphs have a relatively low chromatic number. Indeed, there is a vertex coloring of a balanced graph with  $O(m/n)$  colors. One therefore obtains  $O(m^2/n^2)$  pairwise edge-disjoint induced bipartite subgraphs. However, this approach suffers from two problems. It may be that most of these bipartite subgraphs are sparse, and do not contain subgraphs with the required minimum degree. Furthermore, the vertex intersection of any two of these subgraphs could be huge. Instead, our approach in proving Theorem 1 is different. We first obtain a coloring of a balanced graph with significantly more than  $m/n$  colors, but with more control on the number of edges incident with each color class. Using the properties of our coloring, it can be shown that when grouping many of the color classes together, one obtains an induced subgraph with the required minimum degree, with high probability. Combining this with an algebraic construction, it can be shown that many such pairwise edge-disjoint groups of color classes can be obtained.

It is not difficult to prove that the bound in Theorem 1 is tight up to a constant factor, and that the  $h^2$  term in the denominator is also sharp. As proved in Proposition 14, for any  $n, m$  and  $h \geq 2$ , there are graphs with  $n$  vertices and  $m$  edges that do not contain more than  $\binom{\lceil 2m/n \rceil}{2} / \binom{h}{2}$  pairwise edge-disjoint induced subgraphs each having minimum degree at least  $h$ . It is more difficult to prove that the intersection size  $1 + O(n^3/m^2)$

obtained in Theorem 1 is also asymptotically optimal. Observe first that if  $m > \Omega(n^{1.5})$  this intersection is a constant (in fact, 1). For  $\alpha < 1/2$ , if  $m = n^{1+\alpha}$ , then Theorem 1 says that the intersection is  $O(n^{1-2\alpha})$ . It is shown in Theorem 16 that  $n^{1-2\alpha}$  cannot be replaced with  $n^{1-2\alpha-\epsilon}$  for any  $\epsilon > 0$ .

The rest of this paper is organized as follows. In Section 2 we prove a lemma that establishes the desired properties of the aforementioned coloring of balanced graphs. Section 3 contains the proof of the main result, and is split into several subsections according to the different parts of the proof. Section 4 contains the proofs of Proposition 14 and Theorem 16. Concluding remarks are given in the final section.

## 2 Balanced graphs

**Definition 2.** A graph is *balanced* if its average degree is not smaller than the average degree of any of its subgraphs.

As the average degree of a graph with  $n$  vertices and  $m$  edges is  $2m/n$ , a balanced graph has the property that any subgraph with  $n'$  vertices has at most  $n'm/n$  edges. Some examples of balanced graphs are complete graphs, complete bipartite graphs, and trees (in fact, these examples are *strictly* balanced as the average degree of any proper subgraph is smaller than the average degree of the whole graph).

It is not difficult to see that a balanced graph is also  $d$ -degenerate for  $d = \lfloor 2m/n \rfloor$ . Indeed, as long as there is a vertex with degree at most  $d$ , delete it and continue in the same fashion. The process must end with the empty graph as otherwise there is a subgraph with minimum degree at least  $d + 1 > 2m/n$  violating the fact that the original graph is balanced. As a  $d$ -degenerate graph is  $(d + 1)$ -colorable, we have that any balanced graph can be properly colored with  $\lfloor 2m/n \rfloor + 1$  colors. Nevertheless, as mentioned in the introduction, in such a coloring we have little control over the densities of the  $O(m^2/n^2)$  induced bipartite subgraphs that correspond to pairs of color classes. Most of them may be too sparse and not contain subgraphs with the required minimum degree. Furthermore, color classes may be huge (in fact, the average size of a color class is  $\Theta(n^2/m)$ ) and thus, two subgraphs may have large intersection. Instead, we require a more “balanced” coloring of a balanced graph where we do have control over the density of edges between a color class and the other vertices. The main purpose of this section is to define and prove the existence of such a coloring. Before that, we need another definition and result.

**Definition 3.** A proper vertex coloring is *equitable* if the numbers of vertices in any two color classes differ by at most one.

In particular, in an equitable  $k$ -coloring of a graph with  $n$  vertices, every color class has either  $\lceil n/k \rceil$  vertices or  $\lfloor n/k \rfloor$  vertices. For convenience, we allow the case  $k \geq n$  in which case the coloring is trivial (every vertex gets a distinct color) and  $k - n$  additional “color classes” are empty sets. We require the following classical result of Hajnal and Szemerédi [5] on equitable colorings (see also [6]).

**Lemma 4.** [Hajnal and Szemerédi [5]] Any graph with maximum degree  $\Delta$  has an equitable coloring with  $\Delta + 1$  colors.  $\square$

We are now ready to prove the main lemma of this section.

**Lemma 5.** Let  $\delta$  and  $\alpha$  be positive reals. There exists  $N_2 = N_2(\delta, \alpha)$  such that the following holds for all  $n \geq N_2$ . Let  $G$  be a balanced graph with  $n$  vertices and  $m = n^{1+\beta}$  edges where  $\beta \geq \alpha$ . Then  $G$  has a proper coloring with the following properties:

- (i) The number of colors is at most  $\frac{9\delta}{\beta} n^{2\beta}$ ,
- (ii) the size of each color class is at most  $\max\{1, \frac{1}{\delta} n^{1-2\beta}\}$ ,
- (iii) any color class is incident with at most  $n$  edges.

**Proof.** The lemma is proved with  $N_2 = (4/\delta)^{2/\alpha}$ . We construct the following nested sequence  $\{X_i\}$  of subsets of vertices, starting with  $X_0 = V(G)$ . The subset  $X_1$  consists of all vertices of  $G$  with degree at least  $4n^{1.5\beta}$ . More generally, the subset  $X_i$  is constructed as follows. Let  $Y_i$  be the set of vertices of  $X_{i-1}$  having degree at least  $4n^{1.5\beta}$  in  $G[X_{i-1}]$  (recall that  $G[X]$  denotes the subgraph of  $G$  induced by  $X$ ). Let  $Z_i$  be the set of vertices of  $X_{i-1}$  having degree at least  $4n^{(i+2)\beta/2}$  in  $G$ . Set  $X_i = Y_i \cup Z_i$ .

We upper bound the size of  $X_i$ . Denote the number of edges of  $G[X_i]$  by  $m_i$ . We prove by induction on  $i$  that  $|X_i| \leq n^{1-i\beta/2}$  and  $m_i \leq n^{1-(i-2)\beta/2}$ . The base case  $i = 0$  trivially holds since  $|X_0| = n$  and  $m_0 = m = n^{1+\beta}$ .

Assume that we have proved the claim for  $i - 1$  and wish to prove it for  $i$ . Observe first that  $|Y_i| \leq 0.5n^{1-i\beta/2}$  as otherwise the sum of the degrees of the vertices of  $G[X_{i-1}]$  is more than

$$0.5n^{1-i\beta/2} \cdot 4n^{1.5\beta} = 2n^{1-(i-3)\beta/2} \geq 2m_{i-1}.$$

Observe next that  $|Z_i| \leq 0.5n^{1-i\beta/2}$  as otherwise the sum of the degrees of the vertices of  $G$  is more than

$$0.5n^{1-i\beta/2} \cdot 4n^{(i+2)\beta/2} = 2n^{1+\beta} = 2m.$$

It follows that

$$|X_i| = |Y_i \cup Z_i| \leq 0.5n^{1-i\beta/2} + 0.5n^{1-i\beta/2} = n^{1-i\beta/2}.$$

As  $G$  is balanced,  $m_i$  must satisfy:

$$m_i \leq |X_i| \frac{m}{n} = |X_i| n^\beta \leq n^{1-(i-2)\beta/2}.$$

The sequence  $\{X_i\}$  ends at the last  $t$  for which  $X_t \neq \emptyset$ . As we have shown that  $|X_i| \leq n^{1-i\beta/2}$ , we have that  $t \leq 2/\beta$ . For completeness, define  $X_{t+1} = \emptyset$ .

We now describe the required proper vertex coloring of  $G$ . We color separately each set of vertices  $X_{i-1} \setminus X_i$  for  $i = 1, \dots, t + 1$ . Set  $H_i = G[X_{i-1} \setminus X_i]$ . By the definition of  $Y_i$ , the maximum degree of  $H_i$  is less than  $4n^{1.5\beta}$ . Now, since  $n \geq N_2 = (4/\delta)^{2/\alpha}$  we have that  $n^{\alpha/2}\delta > 2$ . As  $\beta \geq \alpha$  we thus have

$$4n^{1.5\beta} \leq 2\delta n^{2\beta}.$$

Hence the maximum degree of  $H_i$  is less than  $2\delta n^{2\beta}$ . By Lemma 4, there is an equitable coloring of  $H_i$  with  $\lfloor 2\delta n^{2\beta} \rfloor + 1$  colors. The overall number of colors used by our coloring is at most

$$(t+1)(2\delta n^{2\beta} + 1) \leq \left(\frac{2}{\beta} + 1\right) (2\delta n^{2\beta} + 1) \leq \left(\frac{3}{\beta}\right) (3\delta n^{2\beta}) \leq \frac{9\delta}{\beta} n^{2\beta}.$$

Next, we prove that each color is incident with at most  $n$  edges. This is trivial for color classes that are singletons, so consider some color class  $C$  with  $|C| \geq 2$  which is defined in the equitable coloring of some  $H_i$ . As the number of color classes of  $H_i$  is  $\lfloor 2\delta n^{2\beta} \rfloor + 1$ , the equitability condition implies that

$$|C| \leq \frac{|X_{i-1} \setminus X_i|}{\lfloor 2\delta n^{2\beta} \rfloor + 1} + 1.$$

Recall, however that  $|X_{i-1}| \leq n^{1-(i-1)\beta/2}$  and therefore

$$|C| \leq \frac{n^{1-(i-1)\beta/2}}{\lfloor 2\delta n^{2\beta} \rfloor + 1} + 1 \leq \frac{1}{2\delta} n^{1-(i+3)\beta/2} + 1. \tag{1}$$

Since  $|C| \geq 2$  we must have  $\frac{1}{2\delta} n^{1-(i+3)\beta/2} \geq 1$ . By the definition of  $Z_i$ , each vertex of  $H_i$  has degree at most  $4n^{(i+2)\beta/2}$  in  $G$ . Hence, the number of edges of  $G$  incident with  $C$  is at most

$$\left(\frac{1}{2\delta} n^{1-(i+3)\beta/2} + 1\right) \cdot 4n^{(i+2)\beta/2} \leq \frac{4}{\delta} n^{1-\beta/2} \leq n$$

where in the last inequality we used the fact that  $n^{\beta/2} \geq n^{\alpha/2} \geq N_2^{\alpha/2} = 4/\delta$ . Finally, observe that the right hand side of (1) is maximized when  $i = 1$ , so we always have

$$|C| \leq \lfloor \frac{1}{2\delta} n^{1-2\beta} + 1 \rfloor \leq \max\{1, \frac{1}{\delta} n^{1-2\beta}\}.$$

□

### 3 Proof of the main result

Set the following constants:

$$\begin{aligned} c(\alpha) &= \frac{\alpha^6}{10^{18}}, \\ \delta(\alpha, h) &= \frac{\sqrt{c}}{h^2} = \frac{\alpha^3}{10^9 h^2}, \\ N_2(\alpha, h) &= (4/\delta)^{2/\alpha}, \\ N(\alpha, h) &= N_2^{1/\alpha}. \end{aligned}$$

The following lemma implies Theorem 1.

**Lemma 6.** Any balanced graph with  $n \geq N_2(\alpha, h)$  vertices and  $m \geq n^{1+\alpha}$  edges has  $cm^2/(h^2n^2)$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . Furthermore, any two such subgraphs intersect in at most  $1 + h^2n^3/(cm^2)$  vertices.

**Proof of Theorem 1 given Lemma 6.** We must prove that any graph  $G$  with  $n \geq N(\alpha, h)$  vertices and  $m \geq n^{1+\alpha}$  edges has  $cm^2/(h^2n^2)$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . Furthermore, any two subgraphs intersect in at most  $1 + h^2n^3/(cm^2)$  vertices.

Notice that  $G$  may not be balanced, so let  $G'$  be a subgraph of  $G$  with the least number of vertices  $n' < n$  and with  $m'$  edges for which  $m'/n' \geq m/n$ . Observe that the minimality of  $n'$  implies that  $G'$  must be balanced. We claim that  $n' \geq N_2$ . Indeed, as  $m/n \geq n^\alpha$  and since the average degree of any graph is less than its number of vertices, we must have  $n' > 2n^\alpha \geq 2N^\alpha \geq N_2$ . Thus, by Lemma 6,  $G'$  has a set  $L$  of  $c(m')^2/(hn')^2 \geq cm^2/(h^2n^2)$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . But induced subgraphs of  $G'$  are also induced subgraphs of  $G$ , so  $L$  is the required set also for  $G$ .

Now, again by Lemma 6, the intersection of any two subgraphs of  $L$  is of size at most  $1 + h^2(n')^3/(c(m')^2)$ . But

$$1 + \frac{h^2(n')^3}{c(m')^2} \leq 1 + \frac{h^2n}{c} \left(\frac{n'}{m'}\right)^2 \leq 1 + \frac{h^2n}{c} \left(\frac{n}{m}\right)^2 = 1 + \frac{h^2n^3}{cm^2}$$

so the intersection of any two elements of  $L$  is of at most the claimed size. □

**Proof of Lemma 6.** The proof is split into two cases, depending upon whether  $m = O(n^{3/2})$  or not. The following two lemmas accurately state the two cases.

**Lemma 7.** Any balanced graph with  $n \geq N_2(\alpha, h)$  vertices and  $m = n^{1+\beta} \geq n^{1+\alpha}$  edges where  $\delta \leq n^{1-2\beta}$  satisfies the conclusions of Lemma 6.

**Lemma 8.** Any balanced graph with  $n \geq N_2(\alpha, h)$  vertices and  $m = n^{1+\beta} \geq n^{1+\alpha}$  edges where  $\delta \geq n^{1-2\beta}$  satisfies the conclusions of Lemma 6.

The proof of Lemma 7 appears in the following two subsections. The proof of Lemma 8 is simpler and occupies Subsection 3.3. As these two lemmas cover all of the density range  $m \geq n^{1+\alpha}$ , they prove Lemma 6. □

### 3.1 Proof of Lemma 7

Recall that for any prime power  $p$ , there exists a projective plane of order  $p$ , denoted by  $PG(p)$  (see, e.g. [10]). In graph-theoretic terms, this means that for any prime power  $p$ , the complete graph  $K_{p^2+p+1}$  can be decomposed into  $p^2 + p + 1$  pairwise edge-disjoint cliques of order  $p + 1$  each. The vertices of  $K_{p^2+p+1}$  correspond to the points of  $PG(p)$  and the cliques in the partition correspond to the lines of  $PG(p)$ . For example, the Fano plane (the case  $p = 2$ ) corresponds to a decomposition of  $K_7$  into 7 pairwise edge-disjoint triangles.

A graph is  $r$ -partite if the set of vertices can be partitioned into  $r$  independent sets (the parts). For simplicity, we allow parts to be empty sets, so, in particular, a graph which is  $c$ -colorable is  $r$ -partite for all  $r \geq c$ . For an  $r$ -partite graph  $R$  and  $1 \leq t \leq r$ , an induced subgraph of  $R$  consisting of precisely  $t$  parts of the partition is called a *full  $t$ -partite subgraph* of  $R$ . Observe that an  $r$ -partite graph has precisely  $\binom{r}{t}$  full  $t$ -partite subgraphs (recall that we allow some parts to be empty sets, so under this assumption, any two empty parts are considered distinct).

Let  $R$  be an  $r$ -partite graph with  $r = p^2 + p + 1$ , where  $p$  is a prime power. A bijection  $\pi$  from the points of  $PG(p)$  to the parts of the partition induces a mapping between the lines of  $PG(p)$  and their corresponding full  $(p + 1)$ -partite subgraphs, one for each line. Equivalently,  $\pi$  defines a set  $L_\pi$  consisting of  $p^2 + p + 1$  full  $(p + 1)$ -partite subgraphs of  $R$ , where any two subgraphs in  $L_\pi$  are edge-disjoint (recall also that full partite subgraphs are, in particular, induced subgraphs). We call  $L_\pi$  a *projective decomposition* of  $R$ . There are precisely  $r!$  possible bijections  $\pi$  and hence there are precisely  $r!$  projective decompositions. The following lemma is immediate from the definition of a projective decomposition.

**Lemma 9.** *Let  $S_t$  be the set of all  $\binom{r}{t}$  full  $t$ -partite subgraphs of an  $r$ -partite graph  $R$ . Suppose that  $r = p^2 + p + 1$  where  $p$  is a prime power. Let  $L_\pi$  be a projective decomposition of  $R$  chosen uniformly at random among the  $r!$  projective decompositions. Then, each element of  $S_{p+1}$  has the same probability of being an element of  $L_\pi$ . In particular, a randomly chosen element of  $L_\pi$  corresponds to a random element of  $S_{p+1}$ .  $\square$*

Recall that our graph  $G$  is a balanced graph with  $m = n^{1+\beta} \geq n^{1+\alpha}$  edges and thus  $\beta \geq \alpha$ . As  $n \geq N_2$  and since  $N_2 = (4/\delta)^{2/\alpha}$  is precisely the constant  $N_2$  of Lemma 5, we may apply Lemma 5 to  $G$  and obtain a proper vertex coloring with the properties guaranteed by that lemma. By Lemma 5, this coloring has at most  $\frac{9\delta}{\beta}n^{2\beta}$  color classes.

Let  $p$  be the smallest prime power that is not smaller than  $3\sqrt{\delta/\beta}n^\beta$ . By Chebyshev's proof of Bertrand's postulate, there is a prime in  $[x, 2x]$  for any  $x \geq 1$ . In particular,

$$3\sqrt{\frac{\delta}{\beta}}n^\beta \leq p \leq 6\sqrt{\frac{\delta}{\beta}}n^\beta. \quad (2)$$

Let  $r = p^2 + p + 1$ . We have

$$\frac{9\delta}{\beta}n^{2\beta} \leq p^2 < r < 2p^2 < \frac{72\delta}{\beta}n^{2\beta}. \quad (3)$$

Thus, by artificially adding empty color classes to the coloring of  $G$  we may assume that the coloring has precisely  $r$  color classes. Denote the color classes by  $C_1, \dots, C_r$  (and recall that some of them may be empty sets). By properties (ii) and (iii) of Lemma 5, and since  $\delta \leq n^{1-2\beta}$ , we have:

$$|C_i| \leq \frac{1}{\delta}n^{1-2\beta}, \quad (4)$$

$$e(C_i, V(G) - C_i) \leq n. \quad (5)$$

In the last inequality,  $e(C_i, V(G) - C_i)$  denotes the number of edges connecting a vertex of  $C_i$  to a vertex outside  $C_i$ .

Recall that  $S_{p+1}$  denotes the set of all full  $(p+1)$ -partite subgraphs of  $G$ . The following two lemmas establish important properties of a randomly selected element of  $S_{p+1}$ .

**Lemma 10.** *The probability that a randomly selected element of  $S_{p+1}$  has at most  $\frac{4\sqrt{\beta}}{\sqrt{\delta}}n^{1-\beta}$  vertices is at least  $11/12$ .*

**Proof.** Consider a randomly selected element of  $S_{p+1}$ . In other words, we select precisely  $p+1$  color classes from the full set of  $p^2+p+1$  color classes. Each vertex has, therefore, a probability of  $\frac{p+1}{p^2+p+1}$  to belong to the randomly selected element. If  $X$  is the random variable corresponding to the number of vertices, then its expectation satisfies

$$\mathbb{E}[X] = n \frac{p+1}{p^2+p+1} \leq \frac{n}{p}.$$

By (2) we have

$$\mathbb{E}[X] \leq \frac{n}{p} \leq \frac{n}{3\sqrt{\frac{\delta}{\beta}}n^\beta} = \frac{\sqrt{\beta}}{3\sqrt{\delta}}n^{1-\beta}.$$

By Markov's inequality we have:

$$\Pr \left[ X > \frac{4\sqrt{\beta}}{\sqrt{\delta}}n^{1-\beta} \right] \leq \frac{1}{12}.$$

□

**Lemma 11.** *Either Lemma 3.2 holds, or else the probability that a randomly selected element of  $S_{p+1}$  has at least  $n^{1-\beta} \frac{\beta}{300\delta}$  edges is at least  $\frac{3}{10}$ .*

The proof of Lemma 11 is more involved and is deferred to the next subsection.

It follows from the last two lemmas that with probability at least  $3/10 - 1/12 > 1/5$ , a randomly selected element of  $S_{p+1}$  has at most  $\frac{4\sqrt{\beta}}{\sqrt{\delta}}n^{1-\beta}$  vertices and at least  $n^{1-\beta} \frac{\beta}{300\delta}$  edges. In particular the average degree of a randomly selected element of  $S_{p+1}$  is at least

$$\frac{(2\beta)/(300\delta)}{4\sqrt{\beta}/\sqrt{\delta}} = \frac{\sqrt{\beta}}{600\sqrt{\delta}} \geq 2h$$

with probability at least  $\frac{1}{5}$ . In the last inequality we used that fact that  $\delta < \alpha/(1200h)^2 \leq \beta/(1200h)^2$ . Observe that a graph with average degree at least  $2h$  has an induced subgraph with minimum degree at least  $h$ . Thus, a randomly selected element of  $S_{p+1}$  has an induced subgraph with minimum degree at least  $h$  with probability at least  $1/5$ . By Lemma 9, the expected number of elements of a randomly chosen projective decomposition that have subgraphs with minimum degree at least  $h$  is at least

$$\frac{r}{5} \geq \frac{9\delta}{5\beta}n^{2\beta} \geq \delta n^{2\beta} = \frac{\sqrt{cm^2}}{h^2n^2} > \frac{cm^2}{h^2n^2}.$$

In particular, there is a choice of  $\frac{cm^2}{h^2n^2}$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . Finally, the intersection of any two such subgraphs is a subset of at most one of the color classes  $C_i$ . By (4), the size of any such intersection is at most

$$\frac{1}{\delta}n^{1-2\beta} = \frac{h^2n^3}{\sqrt{cm^2}} < \frac{h^2n^3}{cm^2}. \quad \square$$

### 3.2 Proof of Lemma 11

As a randomly selected element of  $S_{p+1}$  contains precisely  $p+1$  vertex classes (corresponding to a hypergeometric distribution on the color classes), the event of any edge being in such an element is not independent of the corresponding event for any other edge. It will thus be more convenient to construct a random element of  $S_{p+1}$  as follows. Each color class is *independently* chosen with probability, say,  $1/(2p)$ . If more than  $p+1$  color classes are chosen (and this happens with small probability) we repeat the experiment. Otherwise, suppose only  $p' \leq p+1$  color classes are chosen. We randomly select  $p+1-p'$  additional color classes, uniformly at random. In this way we generate a truly random element of  $S_{p+1}$ . The following is a precise analysis of this process.

Assume that  $H$  is a full partite subgraph obtained by choosing each color class independently with probability  $1/(2p)$ . Let  $Y$  denote the random variable corresponding to the number of color classes chosen for  $H$ . Observe that  $Y$  has binomial distribution  $B(p^2+p+1, \frac{1}{2p})$ . Thus,  $\mathbb{E}[Y] = \frac{p^2+p+1}{2p}$  and hence, by Markov's inequality,

$$\Pr[Y > p+1] = \Pr[Y \geq p+2] \leq \frac{p^2+p+1}{2p} \cdot \frac{1}{p+2} < \frac{1}{2}. \quad (6)$$

Next, let  $X$  denote the random variable corresponding to the number of edges of  $H$ , and let  $X_e$  denote the indicator random variable which equals 1 if the edge  $e$  is chosen for  $H$  (namely, if the two color classes of the endpoints of  $e$  are both chosen). We observe that  $X = \sum_{e \in E(G)} X_e$  and  $\mathbb{E}[X_e] = 1/(4p^2)$ . We have,

$$\mathbb{E}[X] = \sum_{e \in E(G)} \mathbb{E}[X_e] = \frac{m}{4p^2}. \quad (7)$$

We would like to prove that  $X$  does not fall below its expected value by too much. To this end, we must estimate  $\mathbb{V}[X]$ , and hence the covariances  $Cov[X_e, X_f]$  for any pair of edges  $e, f$ . Observe that if the color classes of the endpoints of  $e$  are disjoint from the color classes of the endpoints of  $f$ , then  $X_e$  and  $X_f$  are independent. Following the notations in [2], write  $e \sim f$  if some endpoint of  $e$  is in the same color class as some endpoint of  $f$ . Let  $\Delta = \sum_{e \sim f} \Pr[X_e = 1 \wedge X_f = 1]$  (the sum is over ordered pairs of edges). We have, as in [2], that  $Cov[X_e, X_f] \leq \Pr[X_e = 1 \wedge X_f = 1]$  and thus

$$\mathbb{V}[X] \leq \mathbb{E}[X] + \Delta.$$

To estimate  $\Delta$ , we distinguish between two types of dependencies. If the two color classes of the endpoints of  $e$  are also the two color classes of the endpoints of  $f$  then, clearly  $\Pr[X_e = 1 \wedge X_f = 1] = \Pr[X_e = 1] = 1/(4p^2)$ . Assume that there are  $t_1$  ordered pairs with this *strong* type of dependence. If only one of the color classes of the endpoints of  $e$  is a color class of an endpoint of  $f$ , then clearly  $\Pr[X_e = 1 \wedge X_f = 1] = 1/(8p^3)$ , since the three endpoints of  $e$  and  $f$  must be chosen. Assume that there are  $t_2$  ordered pairs with this *weak* type of dependence. We therefore have

$$\Delta = \frac{t_1}{4p^2} + \frac{t_2}{8p^3}. \quad (8)$$

We first estimate  $t_1$ . Let  $B_{i,j}$  denote the bipartite subgraph induced by the color classes  $C_i$  and  $C_j$  and let  $m_{i,j}$  denote the number of edges of  $B_{i,j}$ . Observe that

$$t_1 \leq \sum_{1 \leq i < j \leq r} m_{i,j}^2. \quad (9)$$

We say that  $B_{i,j}$  is *dense* if  $m_{i,j} \geq \frac{2h}{\delta} n^{1-2\beta}$ . By (4),  $B_{i,j}$  has at most  $\frac{2}{\delta} n^{1-2\beta}$  vertices. Thus, if it is dense, it has average degree at least  $2h$ , and hence also a subgraph with minimum degree at least  $h$ . We may therefore assume that there are at most  $cm^2/(h^2n^2)$  dense  $B_{i,j}$ , as otherwise we have proved Lemma 7. Another point to notice is that even the dense  $B_{i,j}$  cannot be very dense. Indeed, as  $G$  is balanced,  $m_{i,j}$  cannot be more than  $m/n = n^\beta$  times the number of vertices of  $B_{i,j}$ . As the latter is at most  $\frac{2}{\delta} n^{1-2\beta}$ , we have that  $m_{i,j} \leq \frac{2}{\delta} n^{1-\beta}$ . We therefore have that the contribution of the dense  $B_{i,j}$  to (9) is at most

$$\sum_{\text{dense } B_{i,j}} m_{i,j}^2 \leq \frac{cm^2}{h^2n^2} \cdot \frac{4}{\delta^2} n^{2-2\beta} = \frac{4c}{h^2\delta^2} n^2.$$

By definition, there are  $\binom{r}{2}$  graphs  $B_{i,j}$  altogether, as there are  $r$  color classes. We therefore have that the contribution of the non-dense  $B_{i,j}$  to (9) is at most

$$\sum_{\text{non-dense } B_{i,j}} m_{i,j}^2 \leq \binom{r}{2} \frac{4h^2}{\delta^2} n^{2-4\beta} \leq \frac{20736h^2}{\beta^2} n^2$$

where in the last inequality we have used (3). It follows that

$$t_1 \leq n^2 \left( \frac{4c}{h^2\delta^2} + \frac{20736h^2}{\beta^2} \right) \leq \frac{20740h^2}{\beta^2} n^2 \quad (10)$$

where in the last inequality we have used the fact that  $\delta = \sqrt{c}/h^2$  implying that  $\frac{4c}{h^2\delta^2} \leq 4h^2 \leq 4h^2/\beta^2$ .

We next estimate  $t_2$ . If  $e$  is any edge, then  $e \sim f$  is a weakly dependent pair only if  $f$  is incident to one of the color classes of the endpoints of  $e$ . By (5) we therefore have that  $e$  belongs to at most  $2n$  weakly dependent pairs  $e \sim f$ . We therefore have  $t_2 \leq 2mn = 2n^{2+\beta}$ . Together with (8), (10) we have

$$\Delta = \frac{t_1}{4p^2} + \frac{t_2}{8p^3} \leq \frac{n^2 \left( \frac{20740h^2}{\beta^2} \right)}{4p^2} + \frac{2n^{2+\beta}}{8p^3}. \quad (11)$$

By Chebyshev's inequality:

$$\Pr \left[ X \leq \frac{\mathbb{E}[X]}{2} \right] \leq \frac{4\mathbb{V}[X]}{\mathbb{E}[X]^2} \leq \frac{4}{\mathbb{E}[X]} + \frac{4\Delta}{\mathbb{E}[X]^2}.$$

Trivially,  $\frac{4}{\mathbb{E}[X]} \leq \frac{1}{10}$ . By (7) and (11) we obtain:

$$\frac{4\Delta}{\mathbb{E}[X]^2} \leq \frac{16p^2n^2(\frac{20740h^2}{\beta^2})}{m^2} + \frac{16pn^{2+\beta}}{m^2} = \frac{16p^2(\frac{20740h^2}{\beta^2})}{n^{2\beta}} + \frac{16p}{n^\beta}.$$

Using (2) and the fact that  $\delta = \alpha^3/(10^9h^2)$  we get:

$$\frac{4\Delta}{\mathbb{E}[X]^2} \leq \frac{11946240\delta h^2}{\beta^3} + 96\sqrt{\frac{\delta}{\beta}} \leq \frac{1}{10}.$$

We therefore obtain that

$$\Pr \left[ X \leq \frac{\mathbb{E}[X]}{2} \right] \leq \frac{1}{10} + \frac{1}{10} = \frac{1}{5}.$$

By (7) and (2):

$$\mathbb{E}[X] = \frac{m}{4p^2} \geq \frac{\beta}{144\delta}n^{1-\beta}.$$

Therefore,

$$\Pr \left[ X \leq \frac{\beta}{300\delta}n^{1-\beta} \right] \leq \Pr \left[ X \leq \frac{\mathbb{E}[X]}{2} \right] \leq \frac{1}{5}.$$

Finally, recall that by (6), the probability that  $H$  has more than  $p + 1$  vertex classes is less than  $1/2$ . Hence, with probability at least  $1 - 1/2 - 1/5 = 3/10$ , a random element of  $S_{p+1}$  has at least  $\frac{\beta}{300\delta}n^{1-\beta}$  edges.  $\square$

### 3.3 Proof of Lemma 8

For any prime power  $p$ , there exist  $p - 1$  mutually orthogonal Latin squares of order  $p$  (see, e.g. [10]). In graph-theoretic terms, this means that for any prime power  $p$ , and for any  $2 \leq q \leq p + 1$ , the complete  $q$ -partite graph with  $p$  vertices in each vertex class, denoted by  $T(q, p)$ , can be decomposed into  $p^2$  pairwise edge-disjoint cliques of order  $q$  (thus, each such clique is a transversal, as it intersects each vertex class precisely once). Let  $Q(q, p)$  denote such a decomposition of  $T(q, p)$  into cliques. Note that  $Q(q, p)$  can be explicitly constructed since  $q - 2$  mutually orthogonal Latin squares are explicitly constructed from the addition and multiplication tables of a field with  $p$  elements.

Consider a graph  $R$  with  $r = pq$  vertices. A bijective mapping  $\pi$  from  $V(R)$  to the vertices of  $T(q, p)$  naturally defines a set  $L_\pi$  of  $p^2$  pairwise edge-disjoint induced subgraphs of  $R$ , each one with  $q$  vertices, where each subgraph in  $L_\pi$  corresponds to an element of  $Q(q, p)$ . Indeed,  $S \in Q(q, p)$  corresponds to the subgraph of  $R$  induced by the inverse

image of the vertex set of  $S$  in  $R$ . We call  $L_\pi$  a  $(p, q)$ -orthogonal decomposition of  $R$ . There are precisely  $r!$  possible bijections  $\pi$  (we think of the vertices of  $T(q, p)$  as labeled vertices) and hence there are precisely  $r!$  possible  $(p, q)$ -orthogonal decompositions of  $R$ . The following lemma is immediate from the definition of a  $(p, q)$ -orthogonal decomposition.

**Lemma 12.** *Let  $S_q$  be the set of all induced subgraphs with  $q$  vertices of a graph  $R$  with  $r$  vertices. Suppose that  $r = pq$  where  $p$  is a prime power and  $2 \leq q \leq p + 1$ . Let  $L_\pi$  be a  $(p, q)$ -orthogonal decomposition of  $R$  chosen uniformly at random among the  $r!$  orthogonal decompositions. Then, each element of  $S_q$  has the same probability of being an element of  $L_\pi$ . In particular, a randomly chosen element of  $L_\pi$  corresponds to a random element of  $S_q$ .  $\square$*

Consider now our graph  $G$  which is a balanced graph with  $m = n^{1+\beta}$  edges where  $\beta \geq \alpha$ . Let  $p$  be the smallest prime power that is not smaller than  $\sqrt{\delta}n^\beta$ . Thus,

$$\sqrt{\delta}n^\beta \leq p \leq 2\sqrt{\delta}n^\beta. \quad (12)$$

Let  $q$  be the smallest integer such that  $pq \geq n$ . Since we assume in Lemma 8 that  $\delta \geq n^{1-2\beta}$ , we have:

$$q \leq \frac{n}{p} + 1 \leq \frac{n^{1-\beta}}{\sqrt{\delta}} + 1 \leq \sqrt{\delta}n^\beta + 1 \leq p + 1.$$

Let  $r = pq$  and add  $r - n$  isolated vertices to  $G$  to obtain a graph  $R$  with  $r$  vertices. Recall that  $S_q$  denotes the set of all induced subgraphs of  $R$  with  $q$  vertices. The following lemma shows that a random element of  $S_q$  has relatively many edges, with some fixed probability.

**Lemma 13.** *The probability that a randomly selected element of  $S_q$  has at least  $2qh$  edges is at least  $1/4$ .*

**Proof.** The proof follows along the same lines as the proof of Lemma 11. As a randomly selected element of  $S_q$  contains precisely  $q$  vertices, the event of any edge being in such an element is not independent of the corresponding event for any other edge. It will thus be more convenient to construct a random element of  $S_q$  by choosing each vertex of  $R$  *independently* with probability  $1/(2p)$ , so that the expected number of chosen vertices is  $r/(2p) = q/2$ . Thus, the probability that more than  $q$  vertices are chosen is less than  $1/2$ . Hence, in order to prove the lemma, it suffices to prove that the probability of a random induced subgraph  $H$  obtained by choosing each vertex independently with probability  $1/(2p)$  to have less than  $2qh$  edges is at most  $1/4$ , as this will show that a randomly selected element of  $S_q$  has at least  $2qh$  edges with probability at least  $1 - 1/2 - 1/4 = 1/4$ .

Let  $X$  denote the random variable corresponding to the number of edges of  $H$ , and let  $X_e$  denote the indicator random variable which equals 1 if the edge  $e$  is chosen for  $H$ . As in Lemma 11 we have

$$\mathbb{E}[X] = \sum_{e \in E(G)} \mathbb{E}[X_e] = \frac{m}{4p^2}. \quad (13)$$

We define  $\Delta$  as in Subsection 3.2, where an ordered pair of edges  $e$  and  $f$  are dependent (denoted by  $e \sim f$ ) if  $e$  and  $f$  share a common endpoint. Thus,  $\Pr[X_e = 1 \wedge X_f = 1] = 1/(8p^3)$  since for this event to happen, we need the three endpoints of  $e$  and  $f$  to be chosen. As, trivially, there are less than  $2mn$  ordered pairs of dependent edges, we have

$$\Delta \leq \frac{2mn}{8p^3} = \frac{2n^{2+\beta}}{8p^3}. \quad (14)$$

By Chebyshev's inequality:

$$\Pr[X \leq \frac{\mathbb{E}[X]}{2}] \leq \frac{4\mathbb{V}[X]}{\mathbb{E}[X]^2} \leq \frac{4}{\mathbb{E}[X]} + \frac{4\Delta}{\mathbb{E}[X]^2}.$$

Trivially,  $\frac{4}{\mathbb{E}[X]} \leq \frac{1}{10}$ . By (13) and (14) we obtain:

$$\frac{4\Delta}{\mathbb{E}[X]^2} \leq \frac{16pn^{2+\beta}}{m^2} = \frac{16p}{n^\beta}.$$

Using (12) and the fact that  $\delta < 1/320^2$  we get:

$$\frac{4\Delta}{\mathbb{E}[X]^2} \leq 32\sqrt{\delta} \leq \frac{1}{10}.$$

We therefore obtain that

$$\Pr \left[ X \leq \frac{\mathbb{E}[X]}{2} \right] \leq \frac{1}{10} + \frac{1}{10} = \frac{1}{5}.$$

By (13) and (12):

$$\mathbb{E}[X] = \frac{m}{4p^2} \geq \frac{n^{1-\beta}}{16\delta}.$$

Therefore,

$$\Pr \left[ X \leq \frac{n^{1-\beta}}{32\delta} \right] \leq \Pr \left[ X \leq \frac{\mathbb{E}[X]}{2} \right] \leq \frac{1}{5}.$$

But now, observe that since  $\delta > 1/(128h)^2$  we have that

$$2qh \leq \left( \frac{n}{p} + 1 \right) 2h \leq \left( \frac{n^{1-\beta}}{\sqrt{\delta}} + 1 \right) 2h \leq \frac{4hn^{1-\beta}}{\sqrt{\delta}} \leq \frac{n^{1-\beta}}{32\delta}.$$

Thus,

$$\Pr[X \leq 2qh] \leq \Pr \left[ X \leq \frac{n^{1-\beta}}{32\delta} \right] \leq \frac{1}{5}. \quad \square$$

It follows from the last lemma that with probability at least  $1/4$ , a randomly selected element of  $S_q$  has average degree at least  $2h$  and hence a subgraph with minimum degree at least  $h$ . Therefore, by Lemma 12, the expected number of elements of a randomly chosen

$(p, q)$ -orthogonal decomposition that have subgraphs with minimum degree at least  $h$  is at least

$$\frac{p^2}{4} \geq \frac{\delta n^{2\beta}}{4} = \frac{\sqrt{cm^2}}{4h^2n^2} > \frac{cm^2}{h^2n^2}.$$

In particular, there is a choice of  $\frac{cm^2}{h^2n^2}$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . Finally, as any two elements of a  $(p, q)$ -orthogonal decomposition intersect in one vertex, Lemma 8 follows.  $\square$

## 4 Upper bound

We prove that Theorem 1 is asymptotically tight.

**Proposition 14.** *For positive integers  $n$ ,  $m \leq \binom{n}{2}$  and  $h \geq 2$ , there are graphs with  $n$  vertices and  $m$  edges that do not contain  $\binom{\lceil 2m/n \rceil}{2} / \binom{h}{2}$  pairwise edge-disjoint induced subgraphs each having minimum degree at least  $h$ .*

**Proof.** Let  $x$  be the smallest integer satisfying  $x(x-1)/2 + x(n-x) \geq m$ . By the minimality of  $x$  we have that  $\binom{x-1}{2} + (x-1)(n-x+1) \leq m$  and thus  $\binom{x}{2} + x(n-x) - m \leq n-x$ . Let  $X$  be a clique with  $x$  vertices and let  $Y$  be an independent with  $n-x$  vertices. Start by adding all  $x(n-x)$  possible edges between  $X$  and  $Y$ . This graph has  $\binom{x}{2} + x(n-x) \geq m$  edges, so remove  $\binom{x}{2} + x(n-x) - m \leq n-x$  edges between  $X$  and  $Y$  to obtain a graph  $G$  with  $n$  vertices and  $m$  edges. Suppose there is a set  $T$  of  $t$  pairwise edge-disjoint induced subgraphs, each having minimum degree at least  $h$ . As  $Y$  is an independent set, each element of  $T$  must contain at least  $h$  vertices of  $X$ , and hence also all the  $\binom{h}{2}$  edges between these  $h$  vertices of  $X$ . It follows that  $t \leq \binom{x}{2} / \binom{h}{2}$ . Finally, notice that  $x \leq \lceil 2m/n \rceil$  since

$$\frac{(2m/n)(2m/n-1)}{2} + \frac{2m}{n} \left( n - \frac{2m}{n} \right) \geq m.$$

It follows that

$$t \leq \frac{\binom{\lceil 2m/n \rceil}{2}}{\binom{h}{2}}. \quad \square$$

Our next goal is to prove that the intersection size in Theorem 1 is essentially tight. First, observe that if  $m = \Omega(n^{1.5})$ , then the intersection of any two subgraphs, being at most  $1 + h^2n^3/(cm^2)$ , is a constant. In fact, it is 1 already when  $m > hn^{3/2}/\sqrt{c}$ . For  $\alpha < 1/2$ , if  $m = n^{1+\alpha}$ , then Theorem 1 says that the intersection of any two subgraphs is  $O(n^{1-2\alpha})$ . Theorem 16 stated below proves that the term  $n^{1-2\alpha}$  cannot be replaced with  $n^{1-2\alpha-\epsilon}$  for any  $\epsilon > 0$ . Before proving it, we need the following simple lemma.

**Lemma 15.** *Let  $\mathcal{F}$  be a family of subsets of  $\{1, \dots, n\}$ . Assume that  $|\mathcal{F}| = t$  and that  $|X| \geq s$  for each  $X \in \mathcal{F}$ . Then there are two elements of  $\mathcal{F}$  that intersect in at least  $\frac{s}{t-1}(\frac{ts}{n} - 1)$  elements.*

**Proof.** Let  $d_i$  denote the number of subsets to which  $i$  belongs. We have  $d_1 + \dots + d_n \geq ts$ . For a pair  $X, Y \in \mathcal{F}$  let  $m_{X,Y}$  denote the size of their intersection. By Jensen's inequality we have:

$$\sum_{X,Y \in \mathcal{F}} m_{X,Y} = \sum_{i=1}^n \binom{d_i}{2} \geq n \cdot \frac{ts(ts-1)}{2}.$$

As the left hand side is a sum of  $\binom{t}{2}$  terms, we have that for some pair  $X, Y \in \mathcal{F}$ ,

$$m_{X,Y} \geq n \cdot \frac{ts(ts-1)}{t(t-1)} = \frac{s}{t-1} \left( \frac{ts}{n} - 1 \right). \quad \square$$

**Theorem 16.** *Let  $1/2 > \alpha > 0$ . For every  $\epsilon > 0$  there exists  $h = h(\epsilon)$  such that for all  $n$  sufficiently large there exist graphs with  $n$  vertices and at least  $n^{1+\alpha}$  edges such that in every set of  $n^{2\alpha-\epsilon}$  pairwise edge-disjoint induced subgraphs with minimum degree at least  $h$  there are at least two subgraphs that intersect in at least  $n^{1-2\alpha-5\epsilon}$  vertices.*

**Proof.** We use  $h = 8/\epsilon$ . Throughout the proof we assume, whenever necessary, that  $n$  is sufficiently large. Consider the random graph  $G(n, p)$  with  $p = 3n^{\alpha-1}$ . Let  $m$  denote the number of edges of the obtained sample  $G$  and observe that the expectation of  $m$  (which is binomially distributed) is  $\binom{n}{2}p$ . Thus, with probability at least  $1/2$ ,  $m \geq n^{1+\alpha}$  edges.

We next prove that with very high probability, for all positive integers  $k$  satisfying  $k \leq n^{1-\alpha-\epsilon}$ , every subgraph of  $G$  with  $k$  vertices has less than  $kh/2$  edges. Indeed, consider any such  $k$ , and let  $A_k$  be the event that there exists a vertex subset of  $k$  vertices with at least  $kh/2$  edges. Then,

$$\Pr[A_k] \leq \binom{n}{k} \binom{k^2}{kh/2} p^{kh/2} < \left( \frac{en}{k} \right)^k \left( \frac{2ek}{h} \right)^{kh/2} p^{kh/2} = \left( \frac{en}{k} \left( \frac{2ekp}{h} \right)^{h/2} \right)^k.$$

Since  $kp \leq 3n^{-\epsilon}$  for  $k \leq n^{1-\alpha-\epsilon}$ , one can add the inequality above for all  $k = 1, \dots, n^{1-\alpha-\epsilon}$  and still have that with small probability (say, smaller than  $1/4$ ), no event  $A_k$  occurs for  $k$  in that range.

Thus, with positive probability there exists a graph  $G$  with at least  $n^{1+\alpha}$  edges and with all subgraphs with at most  $n^{1-\alpha-\epsilon}$  vertices having less than  $kh/2$  edges. We therefore fix such a graph  $G$ .

Consider any set  $L$  with at least  $n^{2\alpha-\epsilon}$  pairwise edge-disjoint induced subgraphs with minimum degree at least  $h$ . By the properties of  $G$ , each element of  $L$  has at least  $n^{1-\alpha-\epsilon}$  vertices. By Lemma 15 used with  $s = n^{1-\alpha-\epsilon}$  and  $t = n^{2\alpha-\epsilon}$  we have that there are two elements of  $L$  that intersect in at least  $n^{1-2\alpha-5\epsilon}$  vertices.  $\square$

## 5 Concluding remarks

The proof of Theorem 1 is algorithmic and can be implemented in polynomial time. Given a graph with  $n$  vertices and  $m \geq n^{1+\alpha}$  edges, there is a polynomial time randomized algorithm that finds a set of edge-disjoint subgraphs with the cardinality and properties

stated in the theorem. Indeed, a projective plane of order  $p$  for a prime power  $p$  is elementary constructed from the addition and multiplication tables of a field with  $p$  elements. Similarly, a set of  $p - 1$  pairwise orthogonal Latin squares of order  $p$  are constructed from such tables. It is also straightforward to find a subgraph with minimum degree at least  $h$  in a graph with edge density at least  $2h$  in polynomial (in fact, linear) time. The major algorithmic component is therefore the implementation of Lemma 5. The coloring constructed in the lemma calls a procedure for finding an equitable coloring constantly many times. A recent result from [7] provides a polynomial time algorithm for this task.

Theorem 1 and Proposition 14 together prove that the term  $m^2/(h^2n^2)$  in Theorem 1 is tight up to an absolute constant, for all graphs with polynomially many edges (that is,  $m \geq n^{1+\epsilon}$  for any fixed  $\epsilon$ ). If  $m$  is very close to linear, the proof of Lemma 5 (specifically, the fact that there are  $(\log n)/(\log(m/n))$  iterations) introduces a logarithmic factor. It may be interesting to determine if this is essential. Similarly, the proof of Theorem 16 shows that the term  $n^3/m^2$  for the size of the intersection cannot be reduced by any polynomial factor. It may be of some interest to see if the term  $n^3/m^2$  cannot be improved even by logarithmic factors. We suspect that it cannot.

## Acknowledgment

I thank the referee for some very useful comments.

## References

- [1] N. Alon. Disjoint directed cycles. *Journal of Combinatorial Theory, Series B*, 68(2):167–178, 1996.
- [2] N. Alon and J.H. Spencer. *The Probabilistic Method*. Wiley-Interscience, 2000.
- [3] B. Bollobás. *Extremal Graph Theory*. Academic Press, 1978.
- [4] P. Erdős, J. Pach, and Pyber J. Isomorphic subgraphs in a graph. In *Combinatorics (Eger, 1987)*. Colloq. Math. Soc. János Bolyai, 52, North-Holland, Amsterdam, 1988.
- [5] A. Hajnal and E. Szemerédi. Proof of a conjecture of Erdős. In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 601–623. North-Holland, Amsterdam, 1970.
- [6] H.A. Kierstead and A.V. Kostochka. A short proof of the hajnal-szemerédi theorem on equitable colouring. *Combinatorics, Probability and Computing*, 17(2):265, 2008.
- [7] H.A. Kierstead, A.V. Kostochka, M. Mydlarz, and E. Szemerédi. A fast algorithm for equitable coloring. *Combinatorica*, 30(2):217–224, 2010.
- [8] C. Lee, P. Loh, and B. Sudakov. Self-similarity of graphs. submitted.
- [9] V. Rödl. On a packing and covering problem. *European Journal of Combinatorics*, 5(1):69–78, 1985.

- [10] J.H. Van Lint and R.M. Wilson. *A course in combinatorics*. Cambridge university press, 2001.
- [11] R.M. Wilson. Decomposition of complete graphs into subgraphs isomorphic to a given graph. *Congressus Numerantium*, XV:647–659, 1975.
- [12] R. Yuster. Combinatorial and computational aspects of graph packing and graph decomposition. *Computer Science Review*, 1(1):12–26, 2007.