Almost H-factors in dense graphs

Noga Alon *

and Raphael Yuster

Department of Mathematics

Raymond and Beverly Sackler Faculty of Exact Sciences

Tel Aviv University, Tel Aviv, Israel

Abstract

The following asymptotic result is proved. For every fixed graph H with h vertices, any graph G with n vertices and with minimum degree $d \geq \frac{\chi(H)-1}{\chi(H)}n$ contains $(1-o(1))n/h$ vertex disjoint copies of H.

*Research supported in part by a United States Israel BSF Grant and by a Bergmann Memorial Grant

1 Introduction

All graphs considered here are finite, undirected and simple (i.e., have no loops and no parallel edges). If G is a graph on n vertices and H is a graph on h vertices, we say that G has an *H-factor* if it contains n/h vertex disjoint copies of H. Thus, for example, a K_2-factor is simply a perfect matching, whereas a C_4-factor is a spanning subgraph of G every connected component of which is a cycle of length 4.

Let H be a graph on h vertices, let G be a graph on n vertices, and suppose h divides n. There are several known results that show that in this case, if the minimum degree $d = d(G)$ of G is sufficiently large, then G contains an H-factor. Indeed, by Tutte's 1-factor Theorem (see, e.g. [1]) if $d \geq n/2$ then G has a K_2-factor. Similarly, if H is a path of length $h-1$ then, by Dirac's Theorem on Hamilton cycles (cf. [1]), $d \geq n/2$ suffices again for the existence of an H-factor. Corrádi and Hajnal [2] proved that for $H = K_3$, $d = 2n/3$ suffices and Hajnal and Szemerédi [4] proved that for $H = K_k$, $d = \frac{k-1}{k}n$ guarantees an H-factor. All these results are easily seen to be best possible.

A recent conjecture of Erdös and Faudree [3] asserts that any graph with $n = 4m$ vertices and with minimum degree $2m = n/2$ has a C_4-factor. At the moment we are unable to prove or disprove this conjecture, but we can prove that any such graph contains an *almost C_4-factor*, i.e., $m - o(m)$ vertex disjoint copies of C_4. In fact, we can prove a much more general result, that shows that for any fixed graph H, any graph on n vertices with a sufficiently large minimum degree contains a subgraph on $n - o(n)$ vertices which has an H-factor. The exact statement of the result is the following.

Theorem 1.1 *For every $\epsilon > 0$ and for every integer h, there exists an $n_0 = n_0(\epsilon, h)$ such that for every graph H with h vertices and for every $n > n_0$, any graph G with n vertices and with minimum degree $d \geq \frac{\chi(H)-1}{\chi(H)}n$ contains at least $(1-\epsilon)n/h$ vertex disjoint copies of H.*

The proof is based on the Uniformity Lemma of Szemerédi [5] together with some additional ideas, and is presented in the next two sections. The final section contains some concluding remarks and open problems.

2 Almost H-factors in graphs with a totally ϵ-regular partition

We start with a few definitions, most of which follow [5]. If $G = (V, E)$ is a graph, and A, B are two disjoint subsets of V, let $e(A, B) = e_G(A, B)$ denote the number of edges of G with an endpoint in A and an endpoint in B. If A and B are non-empty, define the *density of edges* between A and B by $d(A, B) = \frac{e(A,B)}{|A||B|}$. For $\epsilon > 0$, the pair (A, B) is called *ϵ-regular* if for every $X \subset A$ and $Y \subset B$ satisfying $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$, the inequality

$$|d(A, B) - d(X, Y)| < \epsilon$$

holds.

An *equitable partition* of a set V is a partition of V into pairwise disjoint classes $C_0, C_1, \ldots, C_k$, in which all the classes C_i for $1 \leq i \leq k$ have the same cardinality. The class C_0 is called the *exceptional class* and may be empty. An equitable partition of the set of vertices V of G into the classes $C_0, C_1 \ldots, C_k$, with C_0 being the exceptional class, is called *ϵ-regular* if $|C_0| \leq \epsilon|V|$, and all but at most ϵk^2 of the pairs (C_i, C_j) for $1 \leq i < j \leq k$ are ϵ-regular. The above partition is called *totally ϵ-regular* if C_0 is empty and all pairs (C_i, C_j), where $1 \leq i < j \leq k$, are ϵ-regular.

The following lemma is proved in [5].

Lemma 2.1 (The Uniformity Lemma [5]) *For every $\epsilon > 0$ and every positive integer t there is an integer $T = T(\epsilon, t)$ such that every graph with n vertices has an ϵ-regular partition into $k + 1$ classes, where $t \leq k \leq T$.* $\square$

When applying the Uniformity Lemma to derive Theorem 1.1 we have to prove the existence of almost H-factors in graphs with a totally ϵ-regular partition. When H is a complete multipartite graph, this is established in the following two lemmas.

Lemma 2.2 *Let $C_1, \ldots, C_k$ be a totally ϵ-regular partition of the set of vertices of a graph G, and suppose that $|C_i| = m$ for all i and that $d(C_i, C_j) \geq \delta$ for all $1 \leq i < j \leq k$. If $k \geq 2$ and*

$$(k - 1)\epsilon + \frac{h - 1}{m} < (\frac{\delta}{2})^{hk}$$

then G contains a complete k-partite graph with h vertices in each of its color classes $A_1, \ldots A_k$, where $A_i \subset C_i$ for $1 \leq i \leq k$.

Proof We prove that for every p, $1 \leq p \leq k$, and for every q, $0 \leq q \leq h$, there are (possibly empty) subsets $A_i \subset B_i \subset C_i$, $(1 \leq i \leq k)$, with the following properties.

(i) $|A_i| = h$ for all $i < p$, $|A_p| = q$ and $|A_i| = 0$ for all $i > p$.

(ii) $|B_i| \geq (\frac{\delta}{2})^{(i-1)h} m$ for all $1 \leq i \leq p$ and $|B_i| \geq (\frac{\delta}{2})^{(p-1)h+q} m$ for all $p < i \leq k$

(iii) For all $1 \leq i < j \leq k$, every vertex $u \in A_i$ is adjacent in G to every vertex $v \in B_j$.

The assertion of the lemma follows from the above statement for $p = k$ and $q = h$, since for these values of the parameters the sets A_i are the color classes of a complete multipartite subgraph of G with h vertices in each color class.

The subsets A_i and B_i are constructed by induction on $(p-1)h+q$. For $p = 1$ and $q = 0$ simply take $A_i = \emptyset$ and $B_i = C_i$ for all i. Given the sets A_i, B_i satisfying (i), (ii) and (iii) for p and q we show how to modify them for the next value of $(p-1)h+q$. If $q = h$ and $p < k$ we can replace p by $p+1$ and q by 0 with no change in the sets A_i, B_i. Thus we may assume that q is strictly smaller than h. Consider the set $D_p = B_p \setminus A_p$. Observe that by assumption the cardinality of each B_j, for $p < j \leq k$ is bigger than ϵm. For each such j, let D_p^j denote the set of all vertices in D_p that have less than $(\delta - \epsilon)|B_j|$ neighbors in B_j. We claim that $|D_p^j| < \epsilon m$ for each j. This is because otherwise the two sets $X = D_p^j$ and $Y = B_j$ would contradict the ϵ-regularity of the pair (C_p, C_j), since $d(D_p^j, B_j) < \delta - \epsilon$, whereas $d(C_p, C_j) \geq \delta$, by assumption. Therefore, the cardinality of the set $D_p \setminus (D_p^{p+1} \cup \ldots \cup D_p^k)$ is at least

$$|B_p| - |A_p| - (k-p)\epsilon m \geq (\frac{\delta}{2})^{(p-1)h} m - q - (k-1)\epsilon m > 0,$$

where the last inequality follows from the assumption in the lemma. We can now choose arbitrarily a vertex v in $D_p \setminus (D_p^{p+1} \cup \ldots \cup D_p^k)$, add it to A_p, and replace each B_j for $p < j \leq k$ by the set of neighbors of v in B_j. Since $\delta - \epsilon > \delta/2$ this will not decrease the cardinality of each B_j by more than a factor of $\delta/2$ and it is easily seen that the new sets A_i, B_i defined in this manner satisfy the conditions (i), (ii) and (iii) with $p' = p$ and $q' = q + 1$. This completes the proof of the lemma. $\square$

Corollary 2.3 *Let $C_1, \ldots, C_k$ be a totally γ^2-regular partition of the set of vertices of a graph G, and suppose that $|C_i| = c$ for all i and that $d(C_i, C_j) \geq \delta + \gamma$ for all $1 \leq i < j \leq k$. If $k \geq 2$ and*

$$(k-1)\gamma + \frac{h-1}{\gamma c} < (\frac{\delta}{2})^{hk}$$

then G contains at least $(1-\gamma)c/h$ vertex disjoint complete k-partite graphs with h vertices in each color class, so that each of these graphs has one color class in each C_i.

Proof Let F be a maximal family of vertex disjoint complete k-partite subgraphs of G, each having h vertices in each color class, and each having a color class in each C_i. We have to prove that the cardinality of F is at least $(1-\gamma)c/h$. Suppose this is false, and let G^* be the induced subgraph of G obtained by deleting from G all the vertices of the members of F. Let C_i^* be the set of all vertices of G^*contained in C_i. Clearly, $|C_i^*| \geq \gamma c$, and one can easily check that the sets C_i^* form a totally γ-regular partition of the set of vertices of G^*. Moreover $d(C_i^*, C_j^*) \geq \delta$ for all $1 \leq i < j \leq k$. By Lemma 2.2 (with $m = \gamma c$ and $\epsilon = \gamma$) G^* contains a complete k-partite graph with h vertices in each color class that can be added to F, contradicting its maximality. This completes the proof. $\square$

3 The proof of the main result

In order to deduce Theorem 1.1 from Lemma 2.1 and Corollary 2.3 we need some additional preparation. In particular, we need the theorem of Hajnal and Szemerédi mentioned in the introduction, which is the following.

Lemma 3.1 (Hajnal and Szemerédi [5]) *If k divides n then any graph with n vertices and with a minimum degree $d \geq \frac{k-1}{k}n$ has n/k vertex disjoint copies of K_k.* $\square$

Corollary 3.2 *Let $G = (V, E)$ be a graph with n vertices in which the degrees of all the vertices but at most βn are at least $(1-\beta)\frac{k-1}{k}n$. Then G contains a set of at least $\frac{n}{k} - k(\beta n + 1)$ vertex disjoint copies of K_k.*

Proof Let V' be the set of all vertices of G whose degrees in G are less than $(1-\beta)\frac{k-1}{k}n$. Let G' be the graph obtained from G by joining each vertex of V' to any other vertex of G. (Thus in G' the degree of each vertex in V' is $n-1$). Let G" be the graph obtained from G' by adding to it a complete graph on a set V" of at least $(k-1)\beta n$ and at most $(k-1)(\beta n + 1)$ new vertices and by joining each of them to every vertex of G'. The exact cardinality of V" is chosen so that the total number of vertices of G" will be divisible by k. In G" the degree of every vertex in $V' \cup V$" is $m-1$, where $m = n + |V"|$ is the number of vertices of G". The degree of each other vertex is at least

$(1-\beta)\frac{k-1}{k}n+|V"| \geq \frac{k-1}{k}m$. Therefore, by Lemma 3.1, G" has a set of m/k vertex disjoint copies of K_k. At most $|V'|+|V"| \leq k(\beta n+1)$ of these contain vertices of $V' \cup V$" and all the others are in fact subgraphs of G. Therefore, G contains a set of at least $m/k-|V'|-|V"| \geq n/k-k(\beta n+1)$ vertex disjoint copies of K_k. $\square$

Proof of Theorem 1.1 Given an integer h and a real positive $\epsilon < 1$, choose a real $\delta = \delta(\epsilon, h) > 0$ satisfying

$$\delta < \frac{\epsilon}{33h^2}. \tag{1}$$

Let $\gamma = \gamma(\epsilon, h)$ satisfy

$$\gamma < \frac{1}{2(h-1)}(\frac{\delta}{2})^{h^2}. \tag{2}$$

Put $t = \lceil 1/\delta \rceil$ and let $T(\cdot,\cdot)$ be the function appearing in the Uniformity Lemma (Lemma 2.1). We prove the theorem with

$$n_0 = n_0(\epsilon, h) = \frac{T(\gamma^2, t) \cdot 2h}{(1-\gamma^2)\gamma(\delta/2)^{h^2}}. \tag{3}$$

Let H be a graph with h vertices and let $k = \chi(H)$ denote its chromatic number. Clearly $k \leq h$. Suppose $n > n_0$ and let G be a graph with n vertices in which all degrees are at least $\frac{k-1}{k}n$. We must show that G contains a set of at least $(1-\epsilon)n/h$ vertex disjoint copies of H. Let K be the complete k-partite graph with h vertices in each color class. It is easy to check that K has an H-factor, i.e., it contains k vertex disjoint copies of H. Therefore, it suffices to prove that G contains a set of at least $(1-\epsilon)\frac{n}{kh}$ vertex disjoint copies of K. We next prove this assertion by the Uniformity Lemma, Corollary 2.3 and Corollary 3.2.

By the Uniformity Lemma G has a γ^2-regular partition into $q+1$ vertex disjoint classes $C_0, \ldots, C_q$, where C_0 is the exceptional class and $t \leq q \leq T(\gamma^2, t)$.

Let L be the graph on the vertices $1, 2, \ldots, q$ in which ij is an edge for $1 \leq i < j \leq q$ iff (C_i, C_j) is a γ^2-regular pair and the density of edges in this pair satisfies $d(C_i, C_j) \geq \delta + \gamma$. A vertex i of L is called *good* if there are at most γq other vertices j of L such that the pair (C_i, C_j) is not γ^2-regular. Obviously, all vertices of L but at most $2\gamma q$ are good.

Claim: The degree of any good vertex of L is at least

$$q(\frac{k-1}{k} - \gamma^2 - \frac{1}{q} - 2\gamma - \delta) \geq q\frac{k-1}{k}(1-10\delta).$$

Proof Let $c = \frac{n-|C_0|}{q} \leq n/q$ denote the number of vertices in each of the sets C_j, $1 \leq j \leq q$. For each fixed i, $1 \leq i \leq q$, the sum of the degrees in G of the vertices in C_i is at least $\frac{k-1}{k}nc$, by the hypotheses. On the other hand, if the degree of i in L is d, and i is a good vertex, then the sum of the degrees in G of the vertices in C_i can be bounded by the sum of five summands, as described below.

- The contribution to the sum of the edges between C_i and the exceptional class C_0 does not exceed $|C_0|c \leq \gamma^2 nc$.
- The contribution of edges joining two vertices of C_i does not exceed c^2.
- The contribution of edges between C_i and classes C_j for which the pair (C_i, C_j) is not γ^2-regular is at most c^2 times the number of such indices j and is thus at most $\gamma q c^2$. (Here we used the fact that i is a good vertex of L.)
- The contribution of edges between C_i and classes C_j for which $d(C_i, C_j) < \delta + \gamma$ does not exceed $q(\delta + \gamma)c^2$.
- The contribution of edges between C_i and classes C_j for which (C_i, C_j) is γ^2-regular and $d(C_i, C_j) \geq \delta + \gamma$ is at most dc^2 (since each such j is a neighbor of i in L).

Therefore
$$\frac{k-1}{k} nc \leq \gamma^2 nc + c^2 + \gamma q c^2 + q(\delta + \gamma)c^2 + dc^2.$$
Since $c \leq n/q$ this imples that
$$\frac{k-1}{k} n \leq n(\gamma^2 + \frac{1}{q} + \gamma + (\delta + \gamma) + \frac{d}{q}),$$
and thus
$$d \geq q(\frac{k-1}{k} - \gamma^2 - \frac{1}{q} - 2\gamma - \delta).$$
Since $q \geq t \geq 1/\delta$, we have $1/q \leq \delta$. By (2) $\gamma^2 < \gamma < \delta (< 1)$ and since $k \geq 2$ we conclude that
$$d \geq q(\frac{k-1}{k} - \gamma^2 - \frac{1}{q} - 2\gamma - \delta) \geq q(\frac{k-1}{k} - 5\delta) \geq q\frac{k-1}{k}(1 - 10\delta).$$
This completes the proof of the claim.

Returning to the proof of the theorem, recall that all the vertices of L but at most $2\gamma q < 10\delta q$ are good. Therefore, by the last claim and by Corollary 3.2 (with $\beta = 10\delta$ and $n = q$), L contains a set of at least
$$\frac{q}{k} - k(10\delta q + 1) \geq \frac{q}{k}(1 - 11\delta k^2) \tag{4}$$

vertex disjoint copies of K_k. (Here we used the fact that since $q \geq 1/\delta$ we have $10\delta q + 1 \leq 11\delta q$.)

Consider a copy of K_k in L, and let $i_1, i_2, \ldots, i_k$ be its vertices. Let G' be the induced subgraph of G on $C_{i_1} \cup \ldots \cup C_{i_k}$. The partition of the set of vertices of G' into the classes $C_{i_1}, \ldots, C_{i_k}$ is a totally γ^2-regular partition, by the definition of L. Moreover, by this definition, $d(C_{i_j}, C_{i_s}) \geq \delta + \gamma$ for all $1 \leq j < s \leq k$. In addition, $k \geq 2$ and the number of vertices c in each C_{i_j} satisfies:

$$c \geq \frac{n(1-\gamma^2)}{q} \geq \frac{n(1-\gamma^2)}{T(\gamma^2, t)} \geq \frac{2h}{\gamma(\delta/2)^{h^2}}, \tag{5}$$

where the last inequality follows from (3).

By (2) and (5) we have:

$$(k-1)\gamma + \frac{h-1}{\gamma c} < (h-1)\gamma + \frac{h}{\gamma c} \leq \frac{1}{2}(\delta/2)^{h^2} + \frac{1}{2}(\delta/2)^{h^2} = (\delta/2)^{h^2} \leq (\delta/2)^{hk}.$$

Therefore, by Corollary 2.3, G' contains a set of at least $(1-\gamma)c/h \geq (1-\gamma)\frac{n(1-\gamma^2)}{qh}$ vertex disjoint copies of K.

Since this holds for every copy of K_k in L, this and (4) implies that G contains a set of at least

$$(1-\gamma)\frac{n(1-\gamma^2)}{qh}\frac{q}{k}(1-11\delta k^2) = \frac{n}{kh}(1-\gamma)(1-\gamma^2)(1-11\delta k^2) \tag{6}$$

vertex disjoint copies of K.

However, as $\gamma^2 < \gamma < 11\delta k^2$ we conclude, by (1), that

$$(1-\gamma)(1-\gamma^2)(1-11\delta k^2) \geq (1-11\delta k^2)^3 \geq 1 - 33\delta k^2 \geq 1 - 33\delta h^2 \geq 1 - \epsilon.$$

Thus, by (6), G contains a set of at least $\frac{n}{kh}(1-\epsilon)$ vertex disjoint copies of K. Since each copy of K contains k vertex disjoint copies of H this completes the proof of the theorem. $\Box$

4 Concluding remarks and open problems

1. Theorem 1.1 is essentially best possible in the sense that the quantity $\frac{\chi(H)-1}{\chi(H)}$ appearing there cannot be replaced by any smaller constant. This is easily seen by letting G be a complete k-partite graph with non-equal color classes where H is any complete k-partite graph with equal color classes.

2. Some error term is needed in the statement of Theorem 1.1, even if h divides n, i.e., the statement of the theorem becomes false if we omit the ϵ even if we assume that h divides n. To see this, let G be the graph obtained from two vertex disjoint complete graphs on $n/2+1$ vertices each by identifying two vertices of the first with two vertices of the second. Then in G all the degrees are at least $n/2$. Let H be a 3-connected bipartite graph on $h = 2l$ vertices (e.g., the complete biparite graph $K_{l,l}$, where $l \geq 3$), and suppose that $n = (4s + 2)l$, for some integer s. Clearly, every copy of H in G must be contained completely in one of the two complete graphs consisting G. However, by the assumptions $\frac{n}{2} \equiv l (mod\ h)$ and hence h does not divide $n/2 - 1, n/2$ or $n/2 + 1$, implying that G does not have an H-factor.

   A similar (though slightly more complicated) argument shows that the error term is needed even if h divides n and H is a properly chosen tree with h vertices. In particular, one can show that if H is the complete full ternary tree of depth 3 with $h = 1 + 3 + 9 + 27 = 40$ vertices and G is obtained from two complete graphs on $n/2+1$ vertices each as above, then, if $n = (2s + 1)40$, G does not have an H-factor. We omit the detailed proof of this fact.

   Another example showing that some error term is needed in Theorem 1.1 is the following; let H be the complete bipartite graph $K_{l,l}$, where $l \geq 3$ is odd, and let G be the graph obtained from the complete bipartite graph with color classes of sizes $l(2s+1)+1$ and $l(2s+1)-1$ by adding a perfect matching on the vertices of the larger color class. Here, again, the number of vertices of H, which is $h = 2l$ divides the number of vertices of G, which is $n = (2s+1)2l$, and the minimum degree in G is $n/2$. It is, however, easy to check, that G does not have an H-factor. This example can be obviously extended to show that some error term is needed in Theorem 1.1 for certain graphs H of any desired chromatic number.

3. By the above remark, some error term is needed in any strengthening of Theorem 1.1. The following strengthening seems true.

   Conjecture 4.1 *For every integer h there exists a constant $c(h)$ such that for every graph H with h vertices, any graph G with n vertices and with minimum degree $d \geq \frac{\chi(H)-1}{\chi(H)}n$ contains at least $n/h - c(h)$ vertex disjoint copies of H.*

4. By the Hajnal-Szemerédi result stated in Section 3, the error term in Theorem 1.1 is not

needed in case H is a complete graph and h divides n. As mentioned in the introduction this is also trivially the case if H is a path. It would be interesting to find additional nontrivial graphs H for which no error term is needed. A possible interesting example is the case $H = C_4$, as conjectured by Erdös and Faudree [3].

References

[1] J. A. Bondy and U.S. R. Murty, *Graph Theory with Applications*, Macmillan Press, London, 1976.

[2] K. Corrádi and A. Hajnal, *On the maximal number of independent circuits in a graph*, Acta Math. Acad. Sci. Hungar. 14 (1963), 423-439.

[3] P. Erdös, *Some recent combinatorial problems* , preprint, November 1990.

[4] A. Hajnal and E. Szemerédi, *Proof of a conjecture of Erdös*, in: *Combinatorial Theory and its Applications*, Vol. II (P. Erdös, A. Renyi and V. T. Sós eds.), Colloq. Math. Soc. J. Bolyai 4, North Holland, Amsterdam 1970, 601-623.

[5] E. Szemerédi, *Regular partitions of graphs*, in: *Proc. Colloque Inter. CNRS* (J. -C. Bermond, J. -C. Fournier, M. Las Vergnas and D. Sotteau eds.), 1978, 399-401.