Independent Transversals in r-partite Graphs

Raphael Yuster

Department of Mathematics

Raymond and Beverly Sackler Faculty of Exact Sciences

Tel Aviv University, Tel Aviv, Israel

Abstract

Let $\mathcal{G}(r,n)$ denote the set of all r-partite graphs consisting of n vertices in each partite class. An *independent transversal* of $G \in \mathcal{G}(r,n)$ is an independent set consisting of exactly one vertex from each vertex class. Let $\Delta(r,n)$ be the maximal integer such that every $G \in \mathcal{G}(r,n)$ with maximal degree less than $\Delta(r,n)$ contains an independent transversal. Let $C_r = \lim_{n\to\infty} \frac{\Delta(r,n)}{n}$. We establish the following upper and lower bounds on C_r, provided $r > 2$:

$$\frac{2^{\lfloor \log r \rfloor -1}}{2^{\lfloor \log r \rfloor} - 1} \geq C_r \geq \max\{\frac{1}{2e}, \frac{1}{2^{\lceil \log(r/3) \rceil}}, \frac{1}{3 \cdot 2^{\lceil \log r \rceil -3}}\}.$$

For all $r > 3$, both upper and lower bounds improve upon previously known bounds of Bollobás, Erdös and Szemerédi. In particular, we obtain that $C_4 = 2/3$, and that $\lim_{r\to\infty} C_r \geq 1/(2e)$, where the last bound is a consequence of a lemma of Alon and Spencer. This solves two open problems of Bollobás, Erdös and Szemerédi.

1 Introduction

All graphs considered here are finite, undirected and simple. Let $\mathcal{G}(r,n)$ denote the set of all r-partite graphs consisting of n vertices in each partite vertex class. An *independent transversal* of $G \in \mathcal{G}(r,n)$ is an independent set consisting of exactly one vertex from each vertex class. Let $\Delta(G)$

($\delta(G)$) denote the maximum (minimum) degree of G. Let $\Delta(r,n)$ be the maximal integer such that every $G \in \mathcal{G}(r,n)$ with $\Delta(G) < \Delta(r,n)$ contains an independent transversal. Alternatively, let $\delta(r,n)$ be the minimal integer such that every $G \in \mathcal{G}(r,n)$ with $\delta(G) > \delta(r,n)$ contains an r-clique. Clearly, $\delta(r,n) + \Delta(r,n) = (r-1)n$. Let $C_r = \lim_{n\to\infty} \frac{\Delta(r,n)}{n}$. (Similarly, $c_r = \lim_{n\to\infty} \frac{\delta(r,n)}{n}$). The fact that these limits exist is simple (cf. also, [6] p. 318). Hence, $c_r + C_r = r - 1$. Trivially, $\Delta(2,n) = n$, and therefore, $C_2 = 1$. It was shown by Graver (cf. [4]) that $\delta(3,n) = n$ and therefore, $C_3 = c_3 = 1$. The proof, although elementary, is non-trivial. For $r \geq 4$, the exact value was not known. In fact, the best known results ([5], [6] p. 318, there in terms of c_r) were that for all $r \geq 4$:

$$\frac{1}{2} + \frac{1}{r-2} \geq C_r \geq \frac{2}{r}. \tag{1}$$

For $r = 4$ an example was constructed to obtain $C_4 \leq 8/9$. In this paper we improve both upper and lower bounds, for all $r > 3$. In fact, we have:

Theorem 1.1 *1.* $\Delta(4,n) \geq 2n/3$.

2. For all $r \geq 3$, $\Delta(2r,n) \geq \Delta(r,n)/2$.

3. For all $r \geq 3$,

$$C_r \geq \max\{\frac{1}{2^{\log\lceil(r/3)\rceil}}, \frac{1}{3\cdot 2^{\lceil\log r\rceil - 3}}\}.$$

For the upper bound, we have:

Theorem 1.2 *For every* $r \geq 2$, $\Delta(r,n) \leq n \cdot \frac{2^{\lfloor \log r\rfloor - 1}}{2^{\lfloor \log r\rfloor} - 1}$ *holds for infinitely many values of* n. *Consequently,* $C_r \leq \frac{2^{\lfloor \log r\rfloor - 1}}{2^{\lfloor \log r\rfloor} - 1}$.

Note that the bounds in Theorems 1.1 and 1.2 improve upon those of inequality (1) for *all* $r > 3$. In particular, our upper and lower bounds coincide for $r = 4$, and we therefore obtain that $C_4 = 2/3$. This solves a problem of Bollobás, Erdös and Szemerédi for the case $r = 4$. Note that the previously best known bound was $8/9 \geq C_4 \geq 1/2$. Even for some other values the improvement is significant. For example, we have $2/3 \geq C_6 \geq 1/2$ while the previous bound was $3/4 \geq C_6 \geq 1/3$.

It is obvious that $\Delta(r+1,n) \leq \Delta(r,n)$, since we may add a disconnected vertex class. Hence, C_r is a monotone decreasing function of r, and $\mu = \lim_{r\to\infty} C_r$ exists. It was conjectured by Bollobás, Erdös and Szemerédi [5] that $\mu = 0.5$. They also asked whether $\mu > 0$ holds (note that by equation (1) or Theorem 1.2 we have $\mu \leq 0.5$). Alon and Spencer have shown in Proposition 5.3 of Chapter 5 of [2], that any r-partite graph with maximum degree d, and with every vertex class having at least $2ed$ vertices (e being the natural logarithm), contains an independent transversal. This implies that for all $r \geq 2$, $\Delta(r,n) > n/(2e)$ and therefore we have:

Proposition 1.3 *For all $r \geq 2$, $C_r \geq 1/(2e)$. Consequently, $\mu \geq 1/(2e)$.* $\Box$

Note that the bound for C_r in Proposition 1.3 supersedes that of Theorem 1.1 only for $r \geq 13$. We can summarize the results of Theorems 1.1, 1.2 and Proposition 1.3 in the following corollary:

Corollary 1.4 *For all $r \geq 3$,*

$$\frac{2^{\lfloor \log r \rfloor - 1}}{2^{\lfloor \log r \rfloor} - 1} \geq C_r \geq \max\{\frac{1}{2e}, \frac{1}{2^{\lceil \log(r/3) \rceil}}, \frac{1}{3 \cdot 2^{\lceil \log r \rceil - 3}}\}.$$

The rest of this paper is organized as follows: In section 2 we prove Theorem 1.1. In section 3 we prove Theorem 1.2. Section 4 contains some concluding remarks and open problems.

2 The lower bound

In this section we prove Theorem 1.1. We begin with the following definitions. Let $G = (V,E) \in \mathcal{G}(r,n)$ have vertex classes $V_1, \ldots, V_r$. Let $E(V_i, V_j)$ denote the set of edges of G with one endpoint in V_i and the other in V_j. The *bipartite complement* $BC(i,j)$ $(1 \leq i < j \leq r)$ is the bipartite graph whose vertex classes are V_i and V_j and whose edge-set is

$$E(BC(i,j)) = \{(u,v) \mid u \in V_i,\ v \in V_j,\ (u,v) \notin E\}.$$

We say that (V_i, V_j) is a *sparse pair* if $BC(i,j)$ contains a perfect matching. If $\Delta(G) < n/2$, we clearly have that every pair (V_i, V_j) is sparse, since $BC(i,j)$ satisfies Hall's condition, and must contain a perfect matching. Somewhat less obvious is the following lemma:

Lemma 2.1 *Assume $\Delta(G) < 2n/3$. If (V_i, V_j) is a non-sparse pair, and k is the size of the maximum matching in $BC(i,j)$, then $k > 2n/3$ and*

$$|E(V_i, V_j)| > 8n^2/9 - 2nk/3 > 2n^2/9.$$

Proof Let (V_i, V_j) be a non-sparse pair. and let $k < n$ be the size of a maximum matching in $BC(i,j)$. For $X \subset V_i$, put $N(X) = \{u \mid \exists x \in X,\ (x,u) \in E(BC(i,j))\}$. Let X_0 be a subset such that $|N(X_0)| = |X_0| - (n-k)$. Such a set X_0 must exist according to Hall's condition (See, e.g. [7]). Clearly, $|X_0| < 2n/3$ since if $|X| \geq 2n/3$ then $|N(X)| = n$. Also, $|N(X_0)| > n/3$ since even a one-vertex set $X = \{x\}$ has $|N(X)| > n/3$, and $X_0 \neq \emptyset$. We therefore have $k > 2n/3$ and $|X_0| > 4n/3 - k$. Note that every vertex of X_0 is connected to every vertex of $V_j \setminus N(X_0)$ in G. Hence

$$|E(V_i, V_j)| \geq |X_0|(n - |N(X_0)|) = |X_0|(2n - k - |X_0|).$$

Since $4n/3 - k < |X_0| < 2n/3$ we have by elementary calculus

$$|E(V_i, V_j)| > 2n/3(4n/3 - k) = 8n^2/9 - 2nk/3 > 2n^2/9.$$

□

We are now ready to prove the first part of the theorem. Let $G \in \mathcal{G}(4,n)$ have $\Delta(G) < 2n/3$. We must show that G contains an independent transversal. Two cases are considered. Assume first that there are two disjoint pairs of vertex classes that are non-sparse. W.l.o.g. assume that the maximum matching in $BC(1,2)$ is $k_1 < n$, and that the maximum matching in $BC(3,4)$ is $k_2 < n$. Assume, for the sake of contradiction, that there is no independent transversal. Then there are at least $k_1 \cdot k_2$ edges in E with one endpoint in $V_1 \cup V_2$ and the other in $V_3 \cup V_4$. Furthermore, by lemma 2.1 we have $|E(V_1, V_2)| > 8n^2/9 - 2nk_1/3$, and also $|E(V_3, V_4)| > 8n^2/9 - 2nk_2/3$. Summing it all, we obtain:

$$|E| > k_1 \cdot k_2 + \frac{16}{9}n^2 - \frac{2}{3}n(k_1 + k_2).$$

Since $k_1, k_2 > 2n/3$ (by lemma 2.1), we have by elementary calculus that $|E| > 4n^2/9 + 16n^2/9 - 8n^2/9 = 4n^2/3$. However, as $\Delta(G) < 2n/3$ and G has $4n$ vertices, we must have $|E| < 4n^2/3$, a contradiction.

We may now assume that in any two disjoint pairs of vertex classes, at least one pair is sparse. We claim that V_1 must be a member of at least one sparse pair. If this were not the case, we would have, by lemma 2.1, that $|E(V_1, V_j)| > 2n^2/9$ for $j = 2, 3, 4$. This means that more than $2n^2/3$ edges are incident with V_1, but this contradicts our assumption that $\Delta(G) < 2n/3$. Similarly, each V_i is a member of at least one sparse pair. Consider the graph H whose vertex set is $\{1, 2, 3, 4\}$ and $(i, j) \in E_H$ iff (V_i, V_j) is a sparse pair. We have shown that the minimal degree of H is at least 1, and according to our assumption the complement of H does not contain a matching. It follows that H must contain a vertex of degree 3. We may therefore assume w.l.o.g. that (V_1, V_j) is a sparse pair for $j = 2, 3, 4$. Let $k_{23}(k_{24}, k_{34})$ be the size of the maximum matching in $BC(2, 3)$ $(BC(2, 4)$, $BC(3, 4))$. By lemma 2.1, we have,

$$n \geq k_{23}, k_{24}, k_{34} > 2n/3. \tag{2}$$

Assume for the sake of contradiction that there is no independent transversal. Hence, considering k_{23}, we have $|E| > nk_{23} + |E(V_1, V_4)| + |E(V_2, V_3)|$. Similar inequalities are obtained when considering k_{24} and k_{34}. We will derive a contradiction by showing that at least one of the following inequalities holds:

$$nk_{23} + |E(V_1, V_4)| + |E(V_2, V_3)| > |E|$$

$$nk_{24} + |E(V_1, V_3)| + |E(V_2, V_4)| > |E|$$

$$nk_{34} + |E(V_1, V_2)| + |E(V_3, V_4)| > |E|.$$

Summing these inequalities, it suffices to show that

$$n(k_{23} + k_{24} + k_{34}) > 2|E|.$$

Recalling that $|E| < 4n^2/3$, it suffices to show that

$$k_{23} + k_{24} + k_{34} > 8n/3. \tag{3}$$

Consequently, establishing (3) will lead to the desired contradiction.

If two out of the three terms on the l.h.s. of (3) equal n, then by (2) we have that (3) is established. If only one of k_{23}, k_{24}, k_{34} equals n, we proceed as follows. W.l.o.g. $k_{23} = n$. By lemma 2.1, we have that $|E(V_4, V_2)| > 8n^2/9 - 2nk_{24}/3$ and that $|E(V_4, V_3)| > 8n^2/9 - 2nk_{34}/3$. Since $\Delta(G) < 2n/3$ we have

$$2n^2/3 > \frac{16}{9}n^2 - \frac{2}{3}n(k_{24} + k_{34}).$$

This implies $k_{24} + k_{34} > 5n/3$ which establishes (3).

We may now assume that $k_{23}, k_{24}, k_{34} < n$. Let G' be the 3-partite induced subgraph of G on the vertex classes V_2, V_3, V_4. Clearly, $\Delta(G') < 2n/3$. Using the fact that $\Delta(3, n) = n$ (mentioned in the introduction), we can obtain at least $n/3$ vertex disjoint independent transversals of G'. By our assumption, none of these transversals can be extended to an independent transversal of G. This means that the degree of each vertex of V_1 is at least $n/3$. Hence,

$$e(V_1, V_2) + e(V_1, V_3) + e(V_1, V_4) \geq n^2/3. \tag{4}$$

Consider the edges adjacent to V_2. We know that $e(V_1, V_2) + e(V_2, V_3) + e(V_2, V_4) < 2n^2/3$. Thus by lemma 2.1

$$e(V_1, V_2) + \frac{16}{9}n^2 - \frac{2}{3}n(k_{23} + k_{24}) < \frac{2}{3}n^2.$$

Corresponding inequalities can be obtained for V_3 and V_4. Summing these three inequalities and using (4) we have

$$\frac{n^2}{3} + \frac{16}{3}n^2 - \frac{4}{3}n(k_{23} + k_{24} + k_{34}) < 2n^2$$

which implies that $k_{23} + k_{24} + k_{34} > 11n/4 > 8n/3$, and (3) is established.

We now prove the second part of Theorem 1.1. Let $r \geq 3$ and let $G \in \mathcal{G}(2r, n)$ have $\Delta(G) < \Delta(r, n)/2$. We must show that G contains an independent transversal. Since $\Delta(G) < \Delta(r, n)/2 \leq n/2$, we have that every pair of vertex classes of G is sparse. Let M_i for $i = 1, \ldots, r$ be a perfect matching in $BC(2i-1, 2i)$. Note that every member of M_i is of the form (a, b) where a, b are non-connected vertices of G, $a \in V_{2i-1}$, $b \in V_{2i}$. We construct a graph $G' \in \mathcal{G}(r, n)$ as follows. The vertex classes of G' are $M_1, \ldots, M_r$. Two vertices $e = (a, b) \in M_i$ and $f = (c, d) \in M_j$ where $i \neq j$ are connected iff at least one of $(a, c), (a, d), (b, c), (b, d)$ is an edge of G. Clearly, $\Delta(G') \leq 2\Delta(G) < \Delta(r, n)$. By the definition of $\Delta(r, n)$ we know that G' contains an independent transversal. Let $(a_1, b_1), \ldots, (a_r, b_r)$ be an independent transversal of G'. It is easy to see from the construction of G' that $a_1, b_1, a_2, b_2, \ldots, a_r, b_r$ is an independent transversal of G, as required.

The last part of Theorem 1.1 follows easily from the facts that $\Delta(3, n) = n$, $\Delta(4, n) \geq 2n/3$ (established in the first part of the theorem), $\Delta(2r, n) \geq \Delta(r, n)/2$ (established in the second part of the theorem) and $\Delta(r, n) \geq \Delta(r + 1, n)$. $\square$

3 The upper bound

In this section we prove Theorem 1.2. The following lemma supplies the desired construction which yields the upper bound.

Lemma 3.1 *For every two positive integers p and q there exists a graph $G_{p,q} \in \mathcal{G}(2^p, q(2^p - 1))$ with $\Delta(G_{p,q}) = q2^{p-1}$ and which does not contain an independent transversal.*

Proof We will construct $G_{p,q}$ by induction on p. In fact we will construct $G_{p,1}$ and for $q > 1$, $G_{p,q}$ is defined as follows. Replace every vertex of $G_{p,1}$ by q copies of it. Two vertices are connected in the new graph $G_{p,q}$ iff their origins were connected in $G_{p,1}$. All vertices that originate from the same vertex are independent and belong to the same vertex class in $G_{p,q}$. Clearly, $G_{p,q} \in$

$calG(2^p, q(2^p-1))$ and $\Delta(G_{p,q}) = q\Delta(G_{p,1}) = q2^{p-1}$, and $G_{p,q}$ does not contain an independent transversal since $G_{p,1}$ does not. For $p = 1$, $G_{1,1}$ is simply the graph consisting of a single edge. Note that, trivially, $G_{1,1}$ satisfies our requirements. For $p = 2$, the graph $G_{2,1}$ resembles the one constructed in [1]. Let the vertex classes of $G_{2,1}$ be (a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3) and (d_1, d_2, d_3). $G_{2,1}$ contains twelve edges in three vertex disjoint cycles of length four, and hence is 2-regular. These cycles are (a_1, b_1, a_2, b_2), (c_1, d_1, c_2, d_2) and (a_3, c_3, b_3, d_3). Clearly $G_{2,1} \in \mathcal{G}(4, 3)$, and $G_{2,1}$ does not contain an independent transversal, since the first two cycles can contribute at most one vertex to an independent transversal, and this means that one of a_3 or b_3 and one of c_3 or d_3 must belong to the independent transversal, but this is impossible due to the third cycle.

Assume, by induction, that we have constructed $G_{p-1,q}$. We now show how to construct $G_{p,1}$. We will use $G_{p-1,2}$ in order to define $G_{p,1}$. Note that $G_{p-1,2} \in \mathcal{G}(2^{p-1}, 2^p - 2)$ and $\Delta(G_{p-1,2}) = 2^{p-1}$, and it does not contain an independent transversal. Denote the vertex classes of $G_{p,1}$ by $V_1, \ldots, V_{2^p}$. Each vertex class is partitioned into two subsets, $V_i = U_i \cup W_i$ where $|U_i| = 2^{p-1}$ and $|W_i| = 2^{p-1} - 1$. For each $j = 1, \ldots, 2^{p-1}$, we join all the vertices of U_{2j-1} to all the vertices of U_{2j}. Notice that the degree of every vertex that belongs to a U_i is exactly 2^{p-1}. We now show how to connect the vertices of the W_i's among themselves. Put $X_j = W_{2j-1} \cup W_{2j}$ for $j = 1, \ldots, 2^{p-1}$. Now assume that the X_j's are the vertex classes of $G_{p-1,2}$. Notice that the degree of every vertex that belongs to a W_i is exactly 2^{p-1}. This completes the construction of $G_{p,1}$. Note that, indeed, $G_{p,1} \in \mathcal{G}(2^p, 2^p - 1)$ and $\Delta(G_{p,1}) = 2^{p-1}$. It remains to show that $G_{p,1}$ does not contain an independent transversal. If T were such a transversal, there could be at most one vertex in T from each of $U_{2j-1} \cup U_{2j}$ for $j = 1, \ldots, 2^{p-1}$. Hence T must contain at least one vertex from each X_j. This, however, is impossible since $G_{p-1,2}$ does not contain an independent transversal. $\square$

Proof of Theorem 1.2: Fix $r \geq 2$, and put $p = \lfloor \log r \rfloor$. Recall from the introduction that $\Delta(r+1, n) \leq \Delta(r, n)$. Thus, $\Delta(r, n) \leq \Delta(2^p, n)$. Now, for every n which is divisible by $2^p - 1$, the graph $G_{p,q}$ constructed in Lemma 3.1, where $q = n/(2^p - 1)$ shows that $\Delta(2^p, n) \leq q2^{p-1}$ Thus,

$$\Delta(r,n) \leq n \cdot \frac{2^{\lfloor \log r \rfloor - 1}}{2^{\lfloor \log r \rfloor} - 1}$$

holds for every n divisible by $2^{\lfloor \log r \rfloor} - 1$. Thus

$$C_r = \lim_{n\to\infty} \frac{\Delta(r,n)}{n} \leq \frac{2^{\lfloor \log r \rfloor - 1}}{2^{\lfloor \log r \rfloor} - 1}.$$

□

4 Concluding remarks and open problems

Our proof of Theorem 1.1 is algorithmic. That is, given a graph $G \in \mathcal{G}(r,n)$ with

$$\Delta(G) < n \max\{\frac{1}{2^{\lceil \log(r/3) \rceil}}, \frac{1}{3 \cdot 2^{\lceil \log r \rceil - 3}}\}$$

we can find an independent transversal in it in $O(n^3)$ time. In the case $r = 3$ we can greedily search all n^3 sets of three vertices, one from each vertex class, until we find an independent transversal, which must exist. In case $r > 3$, we need to apply, constantly many times, an algorithm which finds a maximum matching in a bipartite graph. This requires $O(n^{2.5})$ time, utilizing the best known algorithm for bipartite matchings. However, recall from the proof that we still use as a subroutine the result for $r = 3$, and hence the performance of the algorithm is still dominated by $O(n^3)$. The other ingredients in the algorithmic version of the proof of Theorem 1.1 require less time. This running time is better than the naive $O(n^r)$ algorithm that scans all possible transversals. As mentioned in the introduction, for $r \geq 13$, the bound obtained in Proposition 1.3 is better than that of Theorem 1.1. However, the proof of the Alon-Spencer lemma which yields Proposition 1.3 is non-constructive, as it uses the Lovász Local Lemma (cf. e.g. [2]). Therefore, from an algorithmic perspective, Theorem 1.1 does not become worthless for $r \geq 13$. For a sufficiently large r, however, it will become worthless, as Beck in [3] has shown that in some instances (including ours) the Local

Lemma can be made constructive. The price to pay, however, is a significant loss in the constants. The $1/(2e)$ constant in Proposition 1.3 is replaced by a much smaller one, if an algorithmic version is sought.

The most obvious open problem is that of finding C_r for $r \geq 5$. Even for $r = 5$ we currently only have that $2/3 \geq C_5 \geq 1/2$. A (slightly) less ambitious open problem is that of finding the exact value of $\mu = \lim_{r \to \infty} C_r$ or, at least, improving the current bounds. We currently have $1/2 \geq \mu \geq 1/(2e)$. As mentioned in the introduction, it is conjectured in [5] that $\mu = 1/2$.

References

[1] N. Alon, The strong chromatic number of a graph, Random Structures and Algorithms 3 (1992), 1-7.

[2] N. Alon and J. H. Spencer, The Probabilistic Method, (John Wiley and Sons Inc., New York, 1991).

[3] J. Beck, An algorithmic approach to the Lovász Local Lemma, Random structures and algorithms 2 (1991), 343-365.

[4] B. Bollobás, P. Erdös and E. G. Strauss, Complete subgraphs of chromatic graphs and hypergraphs, Utilitas Math. 6 (1974), 343-347.

[5] B. Bollobás, P. Erdös and E. Szemerédi, On complete subgraphs of r-chromatic graphs, Discrete Math. 13 (1975) 97-107.

[6] B. Bollobás, Extremal Graph Theory, (Academic press, New York 1978).

[7] O. Ore, Graphs and matching theorems, Duke Math. Journal 22 (1955), 625-639.