INDEPENDENT SETS IN REGULAR GRAPHS OF HIGH GIRTH

B. D. McKay

Department of Computer Science, Australian National University

Abstract.

The *independence ratio* of a graph G is the value of $|A|/|V(G)|$, where A is an independent set of G of the largest possible size. For $r \geq 0$, $g \geq 3$, define $i(r,g)$ to be the infimum of the independence ratio over regular graphs with degree r and girth at least g. Bollobás was the first to show that $i(r,g)$ is uniformly bounded below $1/2$ for any fixed $r \geq 3$. In this paper, we sharpen his bounds.

AMS Classification Numbers: 05C35, 05C30, 05C99.

0. Introduction.

A set A of vertices of a graph G is *independent* if no two vertices of A are joined by an edge. The set A is a *maximum* independent set of G if there is no larger independent set. The *independence ratio* of G is $|A|/|V(G)|$ for a maximum independent set A.

A number of authors have explored the relationship between the girth, degree and independence ratio of regular graphs. (Almost) following Bollobás [5], let $i(r,g)$ be the infimum of the independence ratio of regular graphs of degree r and girth at least g.

By considering unions of complete graphs, we see easily that $i(r,3) = 1/(r+1)$ for $r \geq 1$. The only other exact values known for $r \geq 3$ are $i(3,4) = i(3,5) = 5/14$ (Staton [11]) and $i(4,4) = 4/13$ (Jones [8]).

The lower bound $i(r,4) \geq 2/(r+3)$ was proved by Fajtlowicz [6] and improved (for $r \geq 5$) to $i(r,4) \geq (r \log r - r + 1)/(r-1)^2$ by Shearer [10]. Hopkins and Staton [7] show that $i(r,6) \geq (2r-1)/(r^2+2r-1)$, which is better than Shearer's result only for $r \leq 6$. Hopkins and Staton also analyse cubic graphs in more detail and find that $i(3,6) \geq 19/52$, $i(3,8) \geq 20/53$ and $i(3,4k+2) \geq (7k-1)/18k$ for $k \geq 1$. The case of very high degree (with respect to $|V(G)|$ is considered in [1].

In this paper we consider $i(r,\infty) = \lim_{g\to\infty} i(r,g)$. Bollobás [5] used probabilistic methods to show that $i(r,\infty) < 1/2$ for $r \geq 3$. In fact he obtained the upper bounds on $i(r,\infty)$ that are here reproduced in Corollary 2.1. In this paper we will use similar methods to sharpen these bounds.

1. Preliminary Lemmas.

The probabilistic model we use was introduced by Bender and Canfield [2]. Consider a collection of disjoint sets $v_1, v_2, \ldots, v_n$, each of which has cardinality r. A ***pairing*** of order n and degree r consists of a collection of *edges* $\{x,x'\}$ such that

(i) For each edge $\{x,x'\}$, x, $x' \in \bigcup_{i=1}^n v_i$ and $\{x \neq x'\}$, and

(ii) each element of $\bigcup_{i=1}^n v_i$ is in exactly one edge.

Given a pairing P, we can obtain a multigraph $G(P)$. The vertices of P are $v_1, v_2, \ldots, v_n$, and two vertices v_i and v_j are joined by a number of edges equal to the number of edges $\{x,x'\}$ of P such that $x \in v_i$ and $x' \in v_j$.

The properties of pairings relevant to us are summarised in our first theorem.

Theorem 1.1.

(a) For $r \geq 0$, $n \geq 1$, each labelled r-regular simple graph of order n is derivable from exactly the same number of pairings.

(b) For fixed $r \geq 1$, $g \geq 3$, the proportion of pairings of order n and degree r which yield

simple graphs of girth at least g is asymptotically

$$\exp\left(-\sum_{k=1}^{g-1}\frac{(r-1)^k}{2k}\right)$$

as $n \to \infty$.

Proof. Part (a) is elementary. Part (b) was proved for $g = 3$ and fixed r by Bender and Canfield [2]. For arbitrary fixed r and g, part (b) was proved by Wormald [12] and Bollobás [4]. ∎

Part (b) of the theorem was established for $g = 3$ and $r = o(n^{1/3})$ by McKay [9]. The case where both g and r can be increasing functions of n will be the subject of a forthcoming paper of McKay and Wormald.

An immediate consequence of Theorem 1.1 is the following.

Corollary 1.1. *Fix* $r \geq 1$, $g \geq 3$. *Then any property which is true of almost all pairings of degree* r *(as* $n \to \infty$*) is also true of almost all labelled simple regular graphs of degree* r *and girth at least* g. ∎

Corollary 1.1 allows us to conduct our probabilistic analyses within the class of pairings, rather than in the more difficult class of regular graphs. When discussing a pairing P we will feel free to use terminology derived from $G(P)$. For example, an *independent set* of P is a subset of $\{v_1, v_2, \ldots, v_n\}$ which corresponds to an independent set of $G(P)$.

For any formal power series $f(x_1, x_2, \ldots, x_n)$ and vector $k_1, k_2, \ldots, k_n$, the coefficient of $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ in f will be denoted by $[x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}] f(x_1, x_2, \ldots, x_n)$. The following elementary lemma will be very useful for bounding the size of these coefficients.

Lemma 1.1. *Let* $f(x_1, x_2, \ldots, x_n)$ *be a formal power series with nonnegative coefficients and let* $k_1, k_2, \ldots, k_n$ *be integers. Then for any nonnegative real numbers* $\alpha_1, \alpha_2, \ldots, \alpha_n$, *we have*

$$[x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}] f(x_1, x_2, \ldots, x_n) \leq \alpha_1^{-k_1} \alpha_2^{-k_2} \cdots \alpha_n^{-k_n} f(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

Proof. The expression on the right is exactly

$$\sum_{i_1, i_2, \ldots, i_n} \alpha_1^{i_1 - k_1} \alpha_2^{i_2 - k_2} \cdots \alpha_n^{i_n - k_n} [x_1^{i_1} \cdots x_n^{i_n}] f(x_1, x_2, \ldots, x_n).$$

The claim thus follows from the nonnegativity conditions. ∎

The bound of Lemma 1.1 can often be made quite sharp by careful choice of $\alpha_1, \alpha_2, \ldots, \alpha_n$. If f is analytic then the nonnegative solution to

$$\frac{\partial}{\partial \alpha_i} f(\alpha_1, \alpha_2, \ldots, \alpha_n) = \frac{k_i}{\alpha_i} f(\alpha_1, \alpha_2, \ldots, \alpha_n), \quad (i = 1, 2, \ldots, n),$$

if any, gives the optimum choice. For many examples of practical importance the bound obtained is high by only $O(k_1^{1/2} k_2^{1/2} \cdots k_n^{1/2})$. See [3] for some theorems of this nature.

r	$\alpha_1(r)$	$\beta_2(r)$	$\alpha_2(r)$	$\alpha_3(r)$
3	0.45906	0.26413	0.45870	0.45537
4	0.42061	0.22364	0.41989	0.41635
5	0.38868	0.19593	0.38777	0.38443
6	0.36203	0.17546	0.36102	0.35799
7	0.33944	0.15957	0.33840	0.33567
8	0.32001	0.14678	0.31898	0.31652
9	0.30309	0.13623	0.30208	0.29987
10	0.28820	0.12734	0.28722	0.28521
20	0.19886	0.08032	0.19822	0.19732
50	0.11126	0.04198	0.11102	0.11079
100	0.06803	0.02506	0.06793	0.06787

Table 2.1.

2. The Main Results.

Theorem 2.1. [5] *Let $r \geq 3$, $0 < \alpha < 1/2$ be fixed. Then the expected number of independent sets of size αn in a random pairing of order n and degree r is $O(n^{-1/2})f_1(r,\alpha)^n$ as $n \to \infty$, where*

$$f_1(r,\alpha) = \frac{(1-\alpha)^{(1-\alpha)(r-1)}}{(1-2\alpha)^{r(1-2\alpha)/2}\alpha^\alpha}.$$

Proof. For each set A if size αn, count the number of pairings which have A as an independent set. Sum that over A then divide by the total number of pairings. ∎

For each $r \geq 3$ there is a unique $\alpha_1(r)$ in $(0,1/2)$ for which $f_1(r,\alpha_1(r)) = 1$. Furthermore, $f_1(r,\alpha) < 1$ for $\alpha \in (\alpha_1(r), 1/2)$. In conjunction with Theorem 1.1 we thus have the following bound.

Corollary 2.1. [5] *For each $r \geq 3$, $i(r,\infty) \leq \alpha_1(r)$.*

Proof. For any $\alpha > \alpha_1(r)$, almost no pairing has any independent set of size αn or more. By Corollary 1.1 this is also true of almost every r-regular graph of any fixed girth. ∎

Values of $\alpha_1(r)$ can be found in Table 2.1. As $r \to \infty$, we can show that

$$\alpha_1(r) = \frac{2}{r}(\log r - \log\log r + \log 2 + 1 + o(1)).$$

The question now arises as to the precision of Corollary 2.1, and it turns out that the bound is not sharp. The reason for this seems to be that regular graphs with independent sets of size αn for any $\alpha \in (0,1/2)$ tend to have exponentially many such sets. A simple count of the average number thus obscures the size beyond which most graphs have none. We can partly avoid this problem by restricting the count to independent sets satisfying additional properties, which properties must be possessed by any independent sets of greatest size.

As a first attempt we count *maximal* independent sets, which are those which cannot be increased in size by appending an extra vertex.

Theorem 2.2. *Let $r \geq 3$, $1/(r+1) < \alpha < 1/2$ be fixed. Then the average number of maximal independent sets of size αn in a random pairing of order n and degree r is $O(1)f_2(r,\alpha)^n$ as $n \to \infty$, where*

$$f_2(r,\alpha) = \frac{((1+\eta)^r - 1)^{1-\alpha}\alpha^{\alpha(r-1)}(1-2\alpha)^{r(1-2\alpha)/2}}{\eta^{r\alpha}(1-\alpha)^{1-\alpha}},$$

where η is the unique positive solution to

$$\frac{\eta(1+\eta)^{r-1}}{(1+\eta)^r - 1} = \frac{\alpha}{1-\alpha}.$$

Proof. We repeat the calculation outlined in the proof of Theorem 2.1. Let B be the complement of A in the pairing. The effect of the maximality condition is to require that each vertex of B has at least one neighbour in A. Thus the number of choices of the edges between A and B is exactly $(\alpha nr)![x^{r\alpha n}]((1+x)^r - 1)^{(1-\alpha)n}$. The value of this coefficient can be bounded by using Lemma 1.1, where $\eta = \eta_1$ is chosen as described after the proof of that lemma. ∎

For any value of $r \geq 3$ there are numbers $\beta_2(r)$, $\alpha_2(r)$ such that $1/(r+1) < \beta_2(r) < \alpha_2(r) < 1/2$, and

$$f_2(\alpha) \begin{cases} < 1, & \text{if } 1/(r+1) < \alpha < \beta_2(r), \\ \geq 1, & \text{if } \beta_2(r) \leq \alpha \leq \alpha_2(r), \\ < 1, & \text{if } \alpha_2(r) < \alpha < 1/2. \end{cases}$$

The same reasoning as given for Corollary 2.1 leads us to the following result.

Corollary 2.2. *For each $r \geq 3$, $i(r,\infty) \leq \alpha_2(r)$.* ∎

Typical values of $\beta_2(r)$ and $\alpha_2(r)$ are given in Table 2.1. The difference between $\alpha_1(r)$ and $\alpha_2(r)$ is disappointingly small; with much computation it can be shown that

$$\alpha_1(r) - \alpha_2(r) \sim \frac{4\log r}{e^2 r^2}, \text{ as } r \to \infty.$$

We can improve the bound still further by examining the structure of a maximum independent set more closely. Let V be the vertex set of a regular graph G and let A be a maximum independent set. Define B_1 to be the subset of B consisting of those vertices adjacent to exactly one vertex of A and let A_1 be the subset of A consisting of those vertices adjacent to at least one vertex of B_1. Further define $A_2 = A \setminus A_1$ and $C = V \setminus (A \cup B)$.

Lemma 2.1. *Suppose G contains no triangles or pentagons. Then*
(a) each vertex of A_1 is adjacent to exactly one vertex of B_1, and
(b) each vertex of C is adjacent to at least two vertices of A, including at least one of A_2.

Proof.
(a) Suppose that some vertex $x \in A_1$ is adjacent to two vertices $y, z \in B$. Then y and z are not adjacent (since G has no triangles) and so $A \cup \{y,z\} \setminus \{x\}$ is an independent set larger than A.
(b) Any vertex of C not adjacent to A at all could be used to immediately augment A. Suppose instead that some vertex x of C is adjacent only to A_1, say to the subset $X \subseteq A_1$. Let the set $Y \subseteq B$ be the neighbours of X in B. By (a), $|Y| = |X|$ and, since G contains no pentagons, $Y \cup \{x\}$ is independent. Thus $A \cup Y \cup \{x\} \setminus X$ is an independent set larger than A. ∎

Let us call any independent set satisfying conditions (a) and (b) of Lemma 2.1 a *strongly-maximal* independent set.

Theorem 2.3. *Let $r \geq 3$, $(2r-1)/(r^2+2r-1) < \alpha < 1/2$ be fixed. Define $\eta_1, \eta_2 > 0$ by*

$$\eta_1 \frac{\partial}{\partial \eta_1} p(\eta_1, \eta_2) = \frac{\tau(r-1)}{1-\alpha-\tau} p(\eta_1, \eta_2), \text{ and}$$

$$\eta_2 \frac{\partial}{\partial \eta_2} p(\eta_1, \eta_2) = \frac{r(\alpha-\tau)}{1-\alpha-\tau} p(\eta_1, \eta_2), \text{ where}$$

$$p(x, y) = (1+x+y)^r - (1+x)^r - ry.$$

Define $f_3(r, \alpha)$ to be the maximum over $\tau \in (0, \alpha)$ of

$$F(\tau) = \frac{(1-2\alpha)^{r(1-2\alpha)/2}(\alpha-\tau)^{(r-1)(\alpha-\tau)}(r-1)^{\tau(r-1)}(\tau/r)^{(r-2)\tau} p(\eta_1, \eta_2)^{1-\alpha-\tau}}{(1-\alpha-\tau)^{1-\alpha-\tau} \eta_1^{\tau(r-1)} \eta_2^{r(\alpha-\tau)}}.$$

Then the average number of strongly-maximal independent sets of size αn in a random pairing of order n and degree r is $O(n) f_3(r, \alpha)^n$ as $n \to \infty$.

Proof. Consider a strongly-maximal independent set A and define A_1, A_2, B and C as before. Let $|A_1| = |B| = \tau n$, the equal size of those sets following from Lemma 2.1(a).

For such a collection of sets A_1, A_2, B and C, we count the number of pairings for which these sets are valid. The number of choices of the edges between A, and B is $(\tau n)! r^{2\tau n}$ and so on. The only difficult count is that of the edges between A and C: these must satisfy Lemma 2.1(b). Assigning the variable x to those edges incident with A_1 and y to those incident with A_2, we see that the possible neighbours of A for a single vertex in C (recall that we are considering pairings, not graphs) is enumerated by $p(x, y)$. The required coefficient of $p(x, y)^{|C|}$ can be bounded using Lemma 1.1. To handle the unknown τ we simply maximize over it and multiply by $O(n)$. ∎

Rather than applying Theorem 2.3 immediately to bounding $i(r, \infty)$, we consider a simple improvement. Suppose that $X \subseteq B$ is an independent set. Let $Y \subseteq A$, be its neighbours in A. Then $A \cup X \setminus Y$ is an independent set of the same size as A. Since we can do the same for any other independent set in B, obtaining a different independent set of the size of A, we can find quite a lot of such large independent sets.

This leads us to a subproblem of independent interest. For any graph H, let $I(H)$ be the total number of independent sets, including the empty set.

Lemma 2.2. *Let H be a graph of order m and maximum degree at most $r-1$. Then*

$$I(H) \geq (1+r)^{\lfloor m/r \rfloor}.$$

Proof. Let $0 \leq q \leq \lfloor m/r \rfloor$, and consider the number of independent sets of size q. We can choose one vertex in m ways, a second in at least $m-r$ ways, a third in at least $m-2r$ ways, and so on. Allowing $q!$ for the number of ways each independent set can be obtained, and summing over q, we find that the total number of independent sets is at least

$$\sum_{q=0}^{\lfloor m/r \rfloor} \binom{m/r}{q} r^q \geq \sum_{q=0}^{\lfloor m/r \rfloor} \binom{\lfloor m/r \rfloor}{q} r^q = (1+r)^{\lfloor m/r \rfloor}. \quad ∎$$

The bound of Lemma 2.2 is actually achieved for unions of equal complete graphs but can undoubtedly be improved if additional conditions are imposed. The conditions that interest us most are girth conditions.

For any integers $r \geq 0$, $g \geq 3$ define

$$I(r, g) = \inf I(H)^{1/m},$$

where the infimum is over all graphs H of maximum degree at most r and girth at least g, and m is the order of H. This function does not appear to have been studied at all. From Lemma 2.2 we have $I(r, g) \geq (r+2)^{1/(r+1)}$ for any $g \geq 3$. For $r = 2$ we have the following.

Lemma 2.3. *Define the Fibonacci numbers $F_1, F_2, \ldots$ by $F_1 = F_2 = 1$, and $F_i = F_{i-2} + F_{i-1}$ for $i \geq 3$. Then for any $g \geq 3$, $I(2,g) = (F_{s-1} + F_{s+1})^{1/s}$ where $s = 2\lfloor g/2 \rfloor + 1$. In particular,*

$$\lim_{g \to \infty} I(2,g) = (1 + \sqrt{5})/2.$$

Proof. Any graph of maximum degree at most two consists only of paths and cycles. It is readily proved by induction that $I(P_n) = F_{n+2}$ and $I(C_n) = F_{n-1} + F_{n+1}$. The rest is easy. ∎

Problem. *What is $\lim_{g \to \infty} I(r,g)$ for $r \geq 3$?*

Theorem 2.4. *Let $r \geq 3$ be fixed. Define $\alpha_3(r)$ to be the greatest solution for $\alpha \in (0, 1/2)$ of $f_4(r, \alpha) = 1$, where $f_4(r, \alpha)$ is the maximum over $\tau \in (0, \alpha)$ of $F(\tau)/\lambda(r)^\tau$, where $F(\tau)$ is defined in Theorem 2.3 and*

$$\lambda(r) = \begin{cases} (\sqrt{5} + 1)/2, & \text{if } r = 3, \text{ and} \\ (1 + r)^{1/r}, & \text{if } r \geq 4. \end{cases}$$

Then $i(r, \infty) \geq \alpha_3(r)$.

Proof. Fix $\alpha \in (\alpha_3(r), 1/2)$. For each pairing P which contains at least one maximum independent set of size αn, let $\tau(P)$ be the greatest value of $|B|/n$ over all maximum independent sets of P. By Lemmas 2.2 and 2.3, and the switching operation described above, there are positive constants $\epsilon_3, \epsilon_4, \ldots \to 0$ such that P has at least $(\lambda(r) - \epsilon_g)^{\tau(P)n}$ maximum independent sets. Since there are at most n possible values for $|B|$, there is some $\tau'(P) \leq \tau(P)$ such that P has at least $(\lambda(r) - \epsilon_g)^{\tau(P)n}/n$ maximum independent sets with $|B| = \tau'(P)n$, and hence at least $(\lambda(r) - \epsilon_g)^{\tau'(P)n}/n$ such sets. Applying Theorem 2.3, we see that the probability of a random pairing of girth g having at least one maximum independent set of size αn is at most $O(n) \sum_{\tau'} (F(\tau')/(\lambda(r) - \epsilon_g)^{\tau'})^n = O(n^2) \max_{\tau'} (F(\tau')/(\lambda(r) - \epsilon_g)^{\tau'})^n$. The theorem now follows on letting $g \to \infty$. ∎

Values of $\alpha_3(r)$ can be found in Table 2.1. We have not been able to determine its asymptotic behaviour.

It is clear that Theorem 2.4 can be improved by analysing the structure of an independent set in more detail. For example, the decompositions used by Hopkins and Staton [6] to prove lower bounds could be used, although the analysis would be difficult.

It is interesting to consider the limits to analysis of this nature. A natural barrier is perhaps the typical size of a maximum independent set in a random pairing. This is a very difficult value to determine but we can establish lower bounds on it by analysing the behaviour of heuristic algorithms. In the case of random cubic pairings we can show that the independence ratio is almost surely at least $\sqrt{2} - 1 - o(1)$ and have experimental evidence that it is almost surely at least 0.439. Details will appear elsewhere.

References.

[1] M. O. Albertson and L. Chan, Independence and graph homomorphisms, to appear.

[2] E. A. Bender and E. R. Canfield, The asymptotic number of labelled graphs with given degree sequences, *J. Combinatorial Theory A*, **24** (1978) 296–307.

[3] E. A. Bender and L. B. Richmond, Central and local limit theorems applied to asymptotic enumeration II: multivariate generating functions, *J. Combinatorial Theory A*, **34** (1983) 255–265.

[4] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of regular graphs, *European J. Combinatorics*, **1** (1980) 311–314.

[5] B. Bollobás, The independence ratio of regular graphs, *Proc. Amer. Math. Soc.*, **83** (1981) 433–436.

[6] S. Fajtlowicz, On the size of independent sets in graphs, *Proc. 9th S-E Conf. Combinatorics, Graph Theory, and Computing*, (Utilitas Math. 1978), 269-274.

[7] G. Hopkins and W. Staton, Girth and independence ratio, *Canad. Math. Bull.*, **25** (1982) 179–186.

[8] K. F. Jones, Independence in graphs with maximum degree four, *J. Combinatorial Theory B*, **38** (1984) 254–269.

[9] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars Combinatoria*, **19A** (1985) 15–25.

[10] J. B. Shearer, A note on the independence ratio of triangle-free graphs, *Discrete Math.*, **46** (1983) 83–87.

[11] W. Staton, Some Ramsey-type numbers and the independence ratio, *Trans. Amer. Math. Soc.*, **256** (1979) 353–370.

[12] N. C. Wormald, Some problems in the enumeration of labelled graphs, Doctoral Thesis, Dept. of Math., Univ. of Newcastle (1978).