# Large independent sets in regular graphs of large girth

Joe Lauer* and Nicholas Wormald†
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON, Canada
lauer@math.mcgill.ca, nwormald@uwaterloo.ca

**Abstract**

Let $G$ be a $d$-regular graph with girth $g$, and let $\alpha$ be the independence number of $G$. We show that $\alpha(G) \geq \frac{1}{2}\left(1 - (d-1)^{-2/(d-2)} - \epsilon(g)\right) n$ where $\epsilon(g) \to 0$ as $g \to \infty$, and we compute explicit bounds on $\epsilon(g)$ for small $g$. For large $g$ this improves previous results for all $d \geq 7$. The method is by analysis of a simple greedy algorithm which was motivated by the differential equation method used to bound independent set sizes in random regular graphs. We use a "nibble"-type approach but require none of the sophistication of the usual nibble method arguments, relying only upon a difference equation for the expected values of certain random variables. The difference equation is approximated by a differential equation (though we do not use the "differential equation method").

## 1 Introduction

An *independent set*, $I$, of a graph, $G$, is a subset of the vertices of $G$ such that no two vertices of $I$ are joined by an edge in $G$. The *independence number* of $G$ is the cardinality of a maximum independent set, and is denoted by $\alpha(G)$. The *independence ratio* of $G$ is $\alpha(G)/n$, where $G$ has $n$ vertices. We are concerned with finding lower bounds for the independence ratio (and thus $\alpha(G)$) when $G$ has large girth and given maximum degree.

Hopkins and Staton [3] first gave results on the independence number of cubic graphs with large girth. They showed that for a cubic graph $G$, with girth $g$, $\alpha(G) \geq \frac{7}{18}(1 - \epsilon(g))$, where $\epsilon(g) \to 0$ as $g \to \infty$. Shearer [6] improved these bounds on large girth graphs of a given degree sequence, along with explicit results for small girth. For cubic graphs, Shearer's results imply that the constant $\frac{7}{18}$ in the above formula can be replaced by $\frac{125}{302}$, and he gave results for graphs of maximum degree $d$ in terms of $f(d)$ where $f(d)$ is defined iteratively.

It is known that, given such a bound for regular graphs of large girth, the same bound carries over to an asymptotic bound for random regular graphs. This is true because, for every $C > 0$, the probability that a random $d$-regular graph has more than $k$ cycles of length less than $C$ tends to 0 as $k \to \infty$. In [7] the second author analysed two greedy algorithms which give rise to large independent sets in random regular graphs. One of those gave weaker bounds than the other. Here, we establish the same result (asymptotically speaking) for regular graphs of large girth as was established in [7] by an algorithm resembling the weaker of these two, and establish a result for all regular graphs of large girth, that coincides (as

*Supported by NSERC via a USRA.
†Supported by the Canada Research Chairs Program and NSERC.

the girth goes to infinity) with the corresponding result for random graphs. We use an iterative approach, increasing the size of a prototype independent set $I$ at each iteration in a fashion that is often called greedy but is perhaps best described as purely random. After the final iteration, we modify $I$ slightly to ensure the set is independent. This gives a result for any given girth. With an arbitrarily large number, $k$, of iterations, the results hold for all graphs of sufficiently large girth, depending on $k$. If just one iteration were used, the method would be similar to the argument for dominating sets described by Alon and Spencer [1, Theorem 2.2].

The result for specific girth is presented in Section 2, along with a table of values. The result for large large girth is the following. First, define $\alpha(d, g)$ to be the supremum of those $\alpha$ such that for all $n > 0$, all $n$-vertex graphs $G$ of girth at least $g$ and maximum degree at most $d$ satisfy $\alpha(G) \geq \alpha n$. Then $\alpha(d, g)$ is monotonic nondecreasing in $g$, and is bounded above by 1, so has a limit as $g \to \infty$. Define this limit to be $\alpha(d)$. We ignore the trivial case $d = 2$. Also define

$$\beta(d) = \frac{1 - (d-1)^{-2/(d-2)}}{2}.$$

**Theorem 1** *For all $d \geq 3$, we have $\alpha(d) \geq \beta(d)$.*

It follows that for all $\epsilon > 0$, there exists $g$ such that every $n$-vertex graph $G$ with maximum degree $d$ and girth at least $g$ has $\alpha(G) \geq (\beta(d) - \epsilon)n$. The corresponding result for random regular graphs was established in [7] using a simple greedy algorithm. (There is also a stronger result in [7] that we conjecture also to hold for regular graphs with large girth, asymptotically speaking.) Theorem 1 is proved in Section 2.

McKay [4] gave the best known upper bounds $\gamma(d)$ on $\alpha(d)$. This was done by applying expectation arguments to the model of random regular graphs, and the graphs may be taken to have arbitrarily large girth. Shearer's result [6, Theorem 4] applied to regular graphs was a lower bound $\beta_0(d)$ on $\alpha(d)$. Our new result improves on this for regular graphs when $d \geq 7$, as shown in Lemma 1. Moreover, it has an additional appeal in that the formula is simple and nonrecursive. The constants $\beta_0$ and $\gamma$ are shown in Table 1, along with the value of $\beta$ from Theorem 1 (all to four significant figures).

| $d$ | $\beta_0(d)$ | $\beta(d)$ | $\gamma(d)$ |
|---|---|---|---|
| 3 | 0.4139 | 0.3750 | 0.4554 |
| 4 | 0.3510 | 0.3333 | 0.4163 |
| 5 | 0.3085 | 0.3016 | 0.3844 |
| 6 | 0.2771 | 0.2764 | 0.3580 |
| 7 | 0.2528 | 0.2558 | 0.3357 |
| 8 | 0.2332 | 0.2386 | 0.3165 |
| 9 | 0.2169 | 0.2240 | 0.2999 |
| 10 | 0.2032 | 0.2113 | 0.2852 |
| 20 | 0.1297 | 0.1395 | 0.1973 |
| 50 | 0.0682 | 0.0748 | 0.1108 |
| 100 | 0.0406 | 0.0447 | 0.0679 |

Table 1. Lower and upper bounds on the minimum independence ratio of all $d$-regular graphs with large girth.

The *greedy* algorithm randomly orders the vertices and then picks them in order to add to a growing independent set. Equivalently, it adds vertices to the independent set, starting with none and randomly choosing from the remaining allowed vertices. As a simple extension of our proof of Theorem 1 we also obtain the following in Section 2.

**Theorem 2** *The greedy algorithm applied to a $d$-regular graph of girth $g$ produces an independent set of expected size asymptotic to $\beta(d)n$ as $g \to \infty$.*

Further comparisons with past and future work are given in Section 3.

**Lemma 1** *$\beta(d) > \beta_0(d)$ for all $d \geq 7$.*

**Proof.** The two functions have the following recursive formulae, the first from [6, equation (21)] and the second can be verified by substitution:

$$\beta_0(d) = \frac{d(d-1)}{d^2+1}\beta_0(d-1) + \frac{1}{d^2+1} \tag{1.1}$$

$$\beta(d) = A\beta(d-1) + \frac{1}{2}(1-A),$$

where $A = (d-2)^{2/(d-3)}(d-1)^{-2/(d-2)}$. It is easy show using elementary arguments that $A > 1$ for all $d \geq 4$, and that $\beta(d) > \frac{1}{d-1}$ for all $d \geq 5$.

Then we have

$$A\Big(1 - \frac{d-2}{2}\Big) > \frac{d^2-2}{d^2+1} - \frac{d-2}{2},$$

which implies

$$\frac{1}{d-2}\Big(A - \frac{d(d-1)}{d^2+1}\Big) > \frac{1}{d^2+1} - \frac{1}{2}(1-A),$$

and hence,

$$\beta(d-1)\Big(A - \frac{d(d-1)}{d^2+1}\Big) > \frac{1}{d^2+1} - \frac{1}{2}(1-A)$$

for all $d \geq 6$. Thus

$$A\beta(d-1) + \frac{1}{2}(1-A) > \frac{d(d-1)}{d^2+1}\beta(d-1) + \frac{1}{d^2+1},$$

which implies the full result by induction beginning with $\beta(7) > \beta_0(7)$. ∎

We conclude this section with a description of the probabilistic algorithm and a sketch of its analysis that gives our main result.

Fix $k$ and let $G$ be a graph of girth at least $g \geq 2k+3$. We call the following the *p-greedy algorithm* for producing an independent set of vertices $I$. Initially set $I = \emptyset$. Choose probabilities $p_i$, $i = 1, 2, \ldots$. At the $i^{th}$ iteration, select vertices from $V(G) \setminus (I \cup N(I))$ independently at random with probability $p_i$ each. Add all these vertices to $I$. After the $k^{th}$ iteration, delete all vertices of $I$ that are adjacent to other vertices of $I$, and then stop. (In the last step, it is possible to retain some of these vertices, in order to leave a slightly larger independent set. This would produce an incremental improvement to the result for fixed girth $g$ but the same result for the limit as $g \to \infty$. We make no attempt to optimise our results for fixed $g$.)

Some notation will be useful for the rest of the paper. Let $U_0 = V(G)$, and let $U_i$ be the set of all vertices which are not in $I$ and not adjacent to a vertex in $I$ just after the

$i^{th}$ iteration. In particular, if $i = k$, this definition is made before the final adjustment of deleting adjacent vertices from $I$. The vertices in $U_i$ are "uncovered" vertices, in the sense that they have no neighbour in $I$ and so are still free to be chosen in the next iteration of the algorithm.

Our main result comes from probabilities independent of $i$, i.e. $p_i = p$ for all $i$. An outline of the argument, which is justified in the rest of the paper, is as follows. Let $r_i$ denote the probability that $u \in U_i$ for some vertex $u$ in any $d$-regular graph of sufficiently large girth. For small $i$, this probability is the same as if the vertex were in a $d$-regular infinite tree with the same algorithm applied. Similarly define $q_i$ for the probability that both of two neighbouring vertices are in $U_i$. We assume independence of events concerning the neighbours of $u$. This is intuitively valid when $u \in U_i$ because, by the girth condition, the only paths between such neighbours avoiding $u$ have length at least $2k+1$, and in $k$ steps of the algorithm, the "influence" of any random choice cannot spread to diameter bigger than $2k$. For $p$ extremely small the probability that $u \in U_i$ due to the selections in the $i$th round is

$$r_i \approx r_{i-1}(1-p)(1-dpq_{i-1}/r_{i-1})$$

where the first $p$ is for the choice of $u$ in round $i$ and the other terms are for the choice of one of its $d$ neighbours (for small $p$, two such events are unlikely). Of course, $q_{i-1}/r_{i-1}$ is the probability that a neighbour is in $U_{i-1}$, given that $u$ is. Similarly

$$q_i \approx q_{i-1}(1-p)^2(1-pq_{i-1}/r_{i-1})^{2d-2}.$$

Defining $w_i = r_i/q_i$ and dividing the two equations above,

$$w_i = w_{i-1}(1-p)(1-pw_{i-1})^{d-2} \approx w_{i-1} - p(w_{i-1} - (d-2)w_{i-1}^2).$$

Taking a limit as $p \to 0$ gives the differential equation

$$w' = -(w + (d-2)w^2), \qquad w(0) = 1. \tag{1.2}$$

Solving (2.5) analytically, we obtain $w(x) = ((d-1)e^x - (d-2))^{-1}$. By the same method, the first equation above leads to associated equation

$$r' = -r - drw, \qquad r(0) = 1 \tag{1.3}$$

and hence $r(x) = e^{-x}/((d-1) - (d-2)e^{-x})^{\frac{d}{d-2}}$. Since each round places vertices in $I$ with probability $p$, the size of the resulting independent set is, in the limit as $p \to 0$, $\int_0^\infty r(x)\mathrm{d}x = \beta(d)$.

## 2 The Main Result

We begin with a simple result on independence that enables us to compute probabilities of certain events during the course of the $p$-greedy algorithm running on a $d$-regular graph of large girth. This shows that the graph can be regarded locally as part of a $d$-regular infinite tree.

This lemma requires some notation to state. Let $k \geq 1$ and fix a $d$-regular graph $G$ of girth $g \geq 2k+3$, with $V = V(G)$. Given $u \in V(G)$, and $s \leq k+1$, let $T_1(u,s), \ldots, T_d(u,s)$ denote the $d$ components of $G - \{u\} - \{v : d(u,v) \geq s+1\}$, where $d(u,v)$ is the distance from $u$ to $v$. Due to the girth condition, these are indeed separate components, and $T_j(u,s)$ is a tree of height $j$ for each $j$.

Recall that events $H_1, \dots, H_r$ are *mutually independent* if the probability of any event $\bigwedge_{j=1}^{r} B_j$, with $B_j = H_j$ or $B_j = \overline{H_j}$ for each $j$, is equal to $\prod_{j=1}^{r} \mathbf{P}(B_j)$. We say that random subsets $A_1, \dots, A_r$ of $V$ are *mutually independent* if, for all subsets $W_1, \dots, W_r$ of $V$, the events $\{A_j = W_j\}$, $1 \le j \le r$, are mutually independent.

**Lemma 2** *Let $1 \le i \le k-1$ and let $uv \in E(G)$.*

*(a) $\mathbf{P}(u \in U_i)$ and $\mathbf{P}(u \in U_i \wedge v \in U_i)$ do not depend on the choice of $uv$.*

*(b) Conditional upon the event $u \in U_i$, the sets $U_i \cap T_j(u, 2)$ are mutually independent for $1 \le j \le d$.*

**Proof.** An alternative way to select the $U_i$ with the required distribution is as follows. Choose sets $\hat{S}_i$, $i = 1, \dots, k$ independently at random, each set $\hat{S}_i$ including each vertex with probability $p_i$ independently for each vertex. Set $U_0 = 0$. Then, for each $i \ge 1$, set $S_i = \hat{S}_i \cap U_{i-1}$ and $U_i = U_{i-1} \setminus (S_i \cup N(S_i))$.

The *$j$-neighbourhood* of $v$ is the set of vertices of distance at most $j$ from $v$. By induction on $i$, for $i \ge 1$ the events $v \in U_i$ and $v \in S_i$ are determined by the intersections of the sets $\hat{S}_j$, $j \le i$, with the $i$-neighbourhood of $v$, and the $(i-1)$-neighbourhood of $v$, respectively.

It follows that the events in (a) depend only on the restriction of the sets $\hat{S}_j$, $j = 1, \dots, k-1$, to the vertices of distance at most $k-1$ from $u$ and $v$. Since $g > 2k-1$, the subgraph of $G$ induced by these vertices is the same as for two adjacent vertices in an infinite $d$-ary tree. As this is the same graph for all $u$ and $v$, part (a) holds.

Part (b) is a bit more subtle. Fix $j$. Conditional on $u \in U_i$, induction on $i' \le i$ shows that the intersections of $S_{i'}$ and $U_{i'}$ with the vertices of $T_j(u, 2)$ depend only on the restrictions of the sets $\hat{S}_j$, $j = 1, \dots, i'$, to the vertices of $T_j(u, 2+i')$. Since $g > 2k+2$, these trees are pairwise disjoint, and (b) follows. ∎

Given a sequence $p_1, \dots, p_k$, set $r_0 = q_0 = 1$. Then iteratively for $1 \le i \le k$ define

$$r_i = r_{i-1}(1-p_i)(1-p_i q_{i-1}/r_{i-1})^d \tag{2.1}$$

and

$$q_i = q_{i-1}(1-p_i)^2(1-p_i q_{i-1}/r_{i-1})^{2d-2}, \tag{2.2}$$

and set

$$f(d, p_1, \dots, p_k) = \sum_{i=1}^{k} p_i (1-p_i)^d r_{i-1}, \tag{2.3}$$

where the value of $d$ we will find useful is the degree of the vertices in a regular graph.

**Theorem 3** *Let $k \ge 1$ and $G$ be a $d$-regular graph with girth $g \ge 2k+3$. Then*

$$\alpha(G) \ge n \max_{p_1, \dots, p_k} f(d, p_1, \dots, p_k).$$

Note that every choice of $p_i$ gives a lower bound on the independence ratio. Lower bounds on the function being maximised are given in Table 2 for some specific values of $k$ and $d$, obtained by setting $p_i = p$ for all $i$. Here we used $p$ ranging from 0.0264 and 0.0149 for $k = 100$, $d = 3$ and 10 respectively, and down to 0.0006 for $k = 10000$ and $d = 10$.

| $k$ | $d=3$ | $d=4$ | $d=5$ | $d=10$ |
|---|---|---|---|---|
| 100 | 0.3416 | 0.2962 | 0.2623 | 0.1692 |
| 200 | 0.3548 | 0.3105 | 0.2770 | 0.1836 |
| 500 | 0.3650 | 0.3219 | 0.2891 | 0.1965 |
| 1000 | 0.3693 | 0.3267 | 0.2943 | 0.2024 |
| 5000 | 0.3735 | 0.3316 | 0.2996 | 0.2088 |
| 10000 | 0.3742 | 0.3323 | 0.3005 | 0.2099 |

Table 2. Lower bound on independence ratio of all $d$-regular graphs with girth at least $2k+3$.

**Proof of Theorem 3** Let $u$ and $v$ be adjacent vertices in $G$. For each $i \geq 0$, define $R_i = \mathbf{P}(u \in U_i)$ and $Q_i = \mathbf{P}(u \in U_i \wedge v \in U_i)$. By Lemma 2, these probabilities are independent of the choice of $u$ and its neighbour $v$.

Let $i < k$. Conditional upon $u \in U_{i-1}$, the probability that any $v_j \in N(u)$ is in $U_{i-1}$ equals $Q_{i-1}/R_{i-1}$. So the probability that $v_j$ is not placed in $I$ in the $i^{th}$ round is $1 - p_i Q_{i-1}/R_{i-1}$. These events are mutually independent for each $v_j$, by Lemma 2 together with the independence of the choices of vertices in $U_{i-1}$. Hence

$$R_i = R_{i-1}(1-p_i)(1-p_i Q_{i-1}/R_{i-1})^d,$$

where the factor $1-p_i$ equals the probability that $u$ itself is not placed in $I$ in the $i^{th}$ round. Following from this independence, conditional upon $u, v \in U_{i-1}$, the probability that any given $v_j \in N(u) \setminus \{v\}$ is not placed in $I$ in the $i^{th}$ round is $1 - p_i Q_{i-1}/R_{i-1}$, and the same holds for $v$. Furthermore, these events are all independent of each other by Lemma 2(b). Thus

$$Q_i = Q_{i-1}(1-p_i)^2(1-p_i Q_{i-1}/R_{i-1})^{2d-2}.$$

Noting that $Q_0 = 1$, we now have by induction $Q_i = q_i$ for all $0 \leq i \leq k$, and hence from the equation above, $R_i = r_i$ for all $0 \leq i \leq k$.

We also need to account for the deletions of adjacent vertices of $I$ at the end. Note that two adjacent vertices can only arise in $I$ by being selected in the same iteration. The expected number of vertices selected in the $i^{th}$ iteration is $p_i R_{i-1} n$. Also, given a selected vertex, the probability that none of its neighbours is also selected is at least $(1-p_i)^d$, since a vertex has at most $d$ neighbours that can be selected. Such vertices will not be deleted from $I$ in the final deletion round. Therefore, the expected number of vertices added to $I$ at the $i^{th}$ stage that survive the final deletion round is at least $p_i(1-p_i)^d R_{i-1} n$, and so

$$\begin{aligned} \mathbf{E}|I| &\geq n \sum_{i=1}^{k} p_i (1-p_i)^d R_{i-1} \\ &= n \sum_{i=1}^{k} p_i (1-p_i)^d r_{i-1} \\ &= n f(d, p_1, \ldots, p_k). \end{aligned}$$

This holds for any sequence $p_1, \ldots, p_k$, and in particular, $\mathbf{E}|I| \geq n \max_{p_1,\ldots,p_k} f(d, p_1, \ldots, p_k)$. The result then follows since $\alpha(G) \geq \mathbf{E}|I|$ by the first moment principle. ∎

**Proof of Theorem 1** We first prove the conclusion for a $d$-regular graph $G$ with $n$ vertices and girth at least $2k+3$. Putting $w_i = q_i/r_i$ (which is actually a conditional probability) and dividing (2.2) by (2.1) gives (for $0 \le i < k$)

$$w_{i+1} = w_i(1-p_i)(1-p_i w_i)^{d-2}.$$

Setting $p_i = \epsilon$ for all $i$ gives

$$w_{i+1} - w_i = -\epsilon w_i - \epsilon(d-2){w_i}^2 + O(\epsilon^2). \tag{2.4}$$

Thus we are interested in $\hat{w}$ satisfying the differential equation (c.f. (1.2))

$$\hat{w}' = -(\hat{w} + (d-2)\hat{w}^2), \qquad \hat{w}(0) = 1. \tag{2.5}$$

This satisfies $\hat{w}(i\epsilon+\epsilon) - \hat{w}(i\epsilon) = -\epsilon\hat{w}(i\epsilon) - \epsilon(d-2)\hat{w}(i\epsilon)^2 + O(\epsilon^2)$. Hence by induction and using Taylor's theorem, $w_i = \hat{w}(i\epsilon) + O(i\epsilon^2)$ uniformly for all $i$ (that is, the constant implicit in $O()$ is absolute) provided $i\epsilon$ is restricted to an interval $[0, c]$ on which $\hat{w}$ is bounded. Thus, on such an interval, $w_i = \hat{w}(i\epsilon) + O(\epsilon)$, and we also have

$$\hat{w}(x) = \frac{e^{-x}}{(d-1)-(d-2)e^{-x}}.$$

By the same method, substituting $p_i = \epsilon$ in (2.1) gives an associated equation (c.f. (1.3))

$$\hat{r}' = -\hat{r} - d\hat{r}\hat{w}, \qquad \hat{r}(0) = 1.$$

Hence

$$\hat{r}(x) = \frac{e^{-x}}{((d-1)-(d-2)e^{-x})^{\frac{d}{d-2}}} \tag{2.6}$$

and we also obtain

$$r_i = \hat{r}(i\epsilon) + O(\epsilon). \tag{2.7}$$

Since the girth $g$ of $G$ is arbitrarily large, we may take $k = c/\epsilon$ for any fixed $c > 0$. Theorem 3 then implies that the independence number of $G$ satisfies

$$\begin{aligned}
\alpha(G) &\ge nf(\epsilon, \ldots, \epsilon) \\
&= n(1-\epsilon)^d \sum_{i=0}^{k-1} \epsilon r_i \\
&= O(\epsilon n) + n\sum_{i=0}^{k-1} \epsilon\hat{r}(i\epsilon) \quad \textit{by (2.7)} \\
&= O(\epsilon n) + n\int_0^c \hat{r}(x)\,\mathrm{d}x
\end{aligned}$$

since the summation is a Riemann sum for the integral with maximum rectangle width $\epsilon$. Letting $\epsilon \to 0$ and then $c \to \infty$, we get

$$\alpha(G) \ge n\int_0^\infty \hat{r}(x)\,\mathrm{d}x.$$

The conclusion of the theorem, for the graph $G$, now follows from

$$\int \hat{r}(x)\,\mathrm{d}x = -\frac{1}{2((d-1)-(d-2)e^{-x})^{\frac{2}{d-2}}}.$$

This proves the theorem for regular graphs.

If $G$ is not regular but has maximum degree $d$, then we can create a regular graph by taking many copies of $G$ and joining some pairs of vertices from two different copies to make the resulting graph, $H$, $d$-regular. By using sufficiently many copies, and a large-girth graph for the plan of connections between different copies, this can be done in such a way that $H$ has the same girth as $G$. Clearly the independence ratio of $H$ is no more than that of $G$ (just consider the copy of $G$ containing the most vertices of a maximum independent set of $H$). The theorem then follows for graphs with maximum degree at most $d$. ∎

**Proof of Theorem 2** Instead of choosing vertices in the first round of the algorithm $p$-greedy with probability $p$ each, first decide how many vertices go in the first round (by sampling $\mathbf{Bin}(n,p)$) and then choose this many distinct vertices sequentially at random. The greedy algorithm adds them to $I$ only if they are not neighbours of vertices already in $I$. The expected number of vertices excluded in this way is $O(p^2n)$. Then modify the following rounds of the $p$-greedy algorithm in the same fashion, running the greedy algorithm until the correct number of vertices of the set $U_i$ have been chosen. The result is an instance of the greedy algorithm, performed up until some random number $t$ of vertices have been chosen. The expected number of vertices retained in this instance of the greedy algorithm and not in the $p$-greedy algorithm, of vice versa, is clearly $O(pn)$. Now taking $k$ (and hence $g$) large enough, we may take $p$ arbitrarily close to 0 and $t > n(1-\epsilon)$ for any $\epsilon > 0$. It follows that the expected number of vertices treated differently by the two algorithms is an arbitrarily small fraction of $n$. The theorem follows. ∎

## 3 A wider perspective

As described in the introduction, our method resembles the use of the simple greedy algorithm of [7], in which vertices are added to the independent set one by one, to obtain results for random regular graphs. Here we add them group by group, but the analogy does not stop there: as for the random graph case, we analyse the algorithm by obtaining a differential equation that describes its behaviour. Our method also strongly resembles the nibble method of Rödl [5]. See also Alon and Spencer [1]. However we do not require sharp concentration inequalities as normally used in incremental or pseudorandom methods such as the nibble method, and also in the (justification of the) differential equation method of [7]. Instead, we make do just with expectation and direct applications of Markov's inequality. This is possible because we are able to obtain independence of many variables by taking sufficiently large girth. In this respect our argument has strong similarities with the upper bound on dominating set size given in [1, Theorem 2.2]. The method for the latter can be viewed as just one step in the sort of iterative procedure that we use here.

From (1.1), one easily gets that $\beta_0(d) \sim \ln d/d \sim \beta(d)$ as $d \to \infty$. On the other hand, the best known asymptotic upper bound known for this problem is asymptotically $2\ln d/d$ coming from random regular graphs [2, 4].

Shearer's result in [**?**] for triangle-free graphs gives a lower bound on the performance of the greedy algorithm on such graphs that is presumably not sharp. The results in the same paper on large girth, i.e. given by $\beta_0(d)$, are not obtained by bounding the greedy algorithm but by treating some vertices differently depending on the set of adjacent vertices having two neighbours "uncovered". A quite sophisticated algorithm (called "degree-greedy") was analysed for random regular graphs in [7], and we believe the same constants will arise from the corresponding algorithm applied to large girth regular graphs. These constants improve $\beta_0(d)$ quite significantly, for all $d \geq 3$. Work on this is presently in progress.

We believe that the approach of this paper is a convenient launching pad for many further investigations of properties of graphs with bounded degree and large girth.

**Acknowledgement** The authors are grateful to a referee for making some comments on a related approach, which led to a simplification in the proof of Lemma 2.

## References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New York 1992.

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

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

[4] B.D. McKay, Independent sets in regular graphs of high girth, Proc. Australia-Singapore Joint Conference on Information Processing and Combinatorial Mathematics (Singapore, 1986). *Ars Combinatoria* **23A** (1987), 179–185.

[5] V. Rödl, On a packing and covering problem, *Eur. J. Combinat.* **5** (1985), 69–78.

[6] J.B. Shearer, A note on the independence number of triangle-free graphs, II, *J. Combinatorial Theory (Series B)* **53** (1991), 300–307.

[7] N.C. Wormald, Differential equations for random processes and random graphs, *Annals of Applied Probability* **5** (1995), 1217–1235.