

On the Independence and Chromatic Numbers of Random Regular Graphs

A. M. FRIEZE

*Department of Mathematics, Carnegie Mellon University,
Pittsburgh, Pennsylvania 15213*

AND

T. ŁUCZAK

*Institute of Mathematics, Adam Mickiewicz University,
Poznan, Poland*

Communicated by the Editors

Received May 24, 1988

Let G_r denote a random r -regular graph with vertex set $\{1, 2, \dots, n\}$ and $\alpha(G_r)$ and $\chi(G_r)$ denote respectively its independence and chromatic numbers. We show that with probability going to 1 as $n \rightarrow \infty$ respectively

$$\left| \alpha(G_r) - \frac{2n}{r} (\log r - \log \log r + 1 - \log 2) \right| \leq \frac{\epsilon n}{r}$$

and

$$\left| \chi(G_r) - \frac{r}{2 \log r} - \frac{8r \log \log r}{(\log r)^2} \right| \leq \frac{8r \log \log r}{(\log r)^2}$$

provided $r = o(n^\theta)$, $0 < \theta < 1/3$, $0 < \epsilon < 1$, are constants, and $r \geq r_\epsilon$, where r_ϵ depends on ϵ only. © 1992 Academic Press, Inc.

This paper is concerned with the independence and chromatic numbers of random regular graphs. Thus let $\text{REG}(n, r)$ denote the set of r -regular graphs with vertex set $[n] = \{1, 2, \dots, n\}$. Let G_r denote a random graph sampled uniformly from $\text{REG}(n, r)$. We use α and χ for independence and chromatic numbers, respectively.

In random graph theory, these have been studied by, inter alia, Matula [13], Grimmett and McDiarmid [9], Bollobás and Erdős [6], Shamir and Spencer [14], Bollobás [5], Frieze [8], and Łuczak [12]. The aim of this paper is to extend the results of [8, 12] to G_r and prove

THEOREM. (a) *Let $0 < \varepsilon < 1$ be fixed. There exists a constant r_ε such that if $r \geq r_\varepsilon$, $r = o(n^\theta)$, $\theta < 1/3$ constant, then*

$$\left| \alpha(G_r) - \frac{2n}{r} (\log r - \log \log r + 1 - \log 2) \right| \leq \frac{\varepsilon n}{r}$$

with probability going to 1 as $n \rightarrow \infty$.

(b) *Moreover, for some constant r_0 and $r_0 < r = o(n^\theta)$, we have*

$$\frac{r}{2 \log r} \leq \chi(G_r) \leq \frac{r}{2 \log r} \left(1 + \frac{32 \log \log r}{\log r} \right)$$

with probability going to 1 as $n \rightarrow \infty$.

(All logarithms are natural.)

Proof of the Theorem. We will first proceed under the assumption that r is constant. The extension to r growing will then be straightforward. We shall use the model of Bender and Canfield [2] and Bollobás [4] to study G_r . Specifically, we will adopt the configuration terminology of [4]. We let $W = [rn]$ and $W_i = \{(i-1)r + 1, \dots, ir\}$, $i = 1, 2, \dots, n$, be a partition of W into n sets of size r . For $w \in W$ we define $\psi(w) = \lceil w/r \rceil$ so that $w \in W_{\psi(w)}$ holds.

A configuration on a set Z , $|Z| = 2k$, is a partition of Z into k pairs. Φ_Z denotes the set of configurations on Z . If $Z \subseteq W$ and $F \in \Phi_Z$ we let $\mu(F)$ be the multigraph with vertex set $[n]$ and k edges $\{\psi(x)\psi(y) : \{x, y\} \in F\}$.

We consider $\Phi = \Phi_W$ as a probability space in which each $F \in \Phi$ is equally likely. Let Q be a property of the graphs in $\text{REG}(n, r)$ and let Q^* be a property of the configurations in Φ . Suppose these properties are such that for $G_r \in \text{REG}(n, r)$ and $F \in \mu^{-1}(G_r)$, G_r has Q if and only if F has Q^* . All we shall need from [2, 4] is

$$\Pr(G_r \in Q) \leq e^{r^2} \Pr(F \in Q^*). \tag{0}$$

In the analysis we only claim that inequalities hold for r and n sufficiently large and ε sufficiently small.

It is well known that the binomial random variable $B(n, p)$ is sharply concentrated around its expected value np , if this is large.

We will rather loosely refer to the following as the ‘‘Chernoff bounds’’:

$$\Pr(B(n, p) \leq (1 - \beta) np) \leq e^{-\beta^2 np/2}$$

$$\Pr(B(n, p) \geq (1 + \beta) np) \leq e^{-\beta^2 np/3}$$

for $0 \leq \beta \leq 1$.

We will concentrate first on the independence number. The most difficult task is to bound $\alpha(F) = \alpha(\mu(F))$ from below, with high probability. To do this we will generate a random F in a somewhat complicated way. Our purpose is to use the result of [8] "halfway" through the construction.

For a multigraph H and vertex v of H we let $d(v, H)$ denote the degree of v in H , where loops count twice.

Step 1. Let $r_1 = r - \lceil r^{1/2} \log r \rceil$ and $m_1 = \lfloor r_1 n / 2 \rfloor$, and let $X = (x_1, x_2, \dots, x_{2m_1})$ be a random member of $[n]^{2m_1}$; i.e., $x_1, x_2, \dots, x_{2m_1}$ are chosen independently at random from $[n]$. Let now G_1 denote the multigraph with vertex set $[n]$ and edge set $\{x_{2i-1}x_{2i} : 1 \leq i \leq m_1\}$.

Step 2. The next step is to delete edges from G_1 so that each vertex has degree at most r and to construct most of F .

Let $d_j = d(j, G_1)$ and Y_1, Y_2, \dots, Y_n be a partition of $Y = [2m_1]$ with $|Y_j| = d_j$ for $1 \leq j \leq n$. In fact let $Y_1 = [d_1]$ and $Y_i = [\sum_{t=1}^i d_t] \setminus [\sum_{t=1}^{i-1} d_t]$ for $i \geq 2$. We construct a configuration F_1 on $Y = \bigcup_{j=1}^n Y_j$.

begin

$F_1 := \emptyset$; $Y'_j := Y_j$ for $1 \leq j \leq n$;

for $t := 1$ to m_1 do

begin

randomly choose $p_{2t-1} \in Y'_{x_{2t-1}}$; $Y'_{x_{2t-1}} := Y'_{x_{2t-1}} - \{p_{2t-1}\}$;

randomly choose $p_{2t} \in Y'_{x_{2t}}$; $Y'_{x_{2t}} := Y'_{x_{2t}} - \{p_{2t}\}$;

$F_1 := F_1 \cup \{\{p_{2t-1}, p_{2t}\}\}$

end

end

CLAIM 1. F_1 is a random configuration on Y .

Proof. $p_1, p_2, \dots, p_{2m_1}$ is a random permutation of Y since interchanging p_k, p_{k+1} yields the same distribution of permutations. Each partition arises from $2^{m_1} m_1!$ permutations. ■

For $s \in Y$ let its rank $\rho(s) = s - \sum_{i=1}^{j-1} d_i$, where $s \in Y_j$, so that Y_j has elements of rank $1, 2, \dots, d_j$. Let

$$F'_1 = \{\{p, q\} \in F_1 : \max\{\rho(p), \rho(q)\} \leq r\}.$$

$$F_2 = \{\{\sigma(p), \sigma(q)\} : \{p, q\} \in F'_1\},$$

where if $p \in Y_j$, $\sigma(p) = (j-1)r + \rho(p)$, and $Z = \bigcup_{e \in F_2} e$.

CLAIM 2. F_2 is a random configuration on Z .

Proof. F_1 is a random configuration on Y implies F'_1 is a random configuration on $\bigcup_{e \in F'_1} e$, since any such configuration has the same number

of extensions to an F_1 . Since F_2 is obtained by a fixed relabelling of $\bigcup_{e \in F_1} e$, the result follows. ■

We let G_2 be the multigraph $\mu(F_2)$.

Step 3. We now enlarge F_2 so that it “covers” the whole of W .

Suppose $Z \neq W$, $x_1, x_2 \in W - Z$, and $Z' = Z \cup \{x_1, x_2\}$. We define a function $f: \Phi_{Z'} \rightarrow \Phi_Z$ as follows: let $F' \in \Phi_{Z'}$.

(a) If $\{x_1, x_2\} \in F'$ then

$$f(F') = F' - \{\{x_1, x_2\}\},$$

otherwise

(b) suppose $\{x_1, z_1\}, \{x_2, z_2\} \in F'$, ($z_1 \neq z_2$), then

$$f(F') = (F' \cup \{\{z_1, z_2\}\}) - \{\{x_1, z_1\}, \{x_2, z_2\}\}.$$

CLAIM 3. If $F \in \Phi_Z$ then $|f^{-1}(F)| = |Z| + 1$.

Proof. If $F' \in f^{-1}(F)$ then either

(a) $F' = F \cup \{\{x_1, x_2\}\}$ or

(b) $F' = F \cup \{\{x_1, z_1\}, \{x_2, z_2\}\} - \{\{z_1, z_2\}\}$ for some $\{z_1, z_2\} \in F$. ■

It follows from Claim 2 and Claim 3 that the following algorithm generates a random configuration F'_2 on Z' :

ADD(F_2, x_1, x_2):

begin

With probability $(|Z| + 1)^{-1}$ let $F'_2 = F_2 \cup \{\{x_1, x_2\}\}$ else randomly choose

$\{z_1, z_2\} \in F_2$ (randomly ordered z_1, z_2) and then let

$$F'_2 = (F_2 \cup \{\{x_1, z_1\}, \{x_2, z_2\}\}) - \{\{z_1, z_2\}\}$$

Output F'_2

end

Hence if $W - Z = \{x_1, x_2, \dots, x_{2s}\}$ the following algorithm constructs a random configuration F on W :

FINISH

begin

$$F := F_2,$$

for $i = 1$ to s do $F := \text{ADD}(F, x_{2i-1}, x_{2i})$

end

We will now show that with high probability

- (i) G_1 , and hence G_2 , has an independent set of the required size.
- (ii) Algorithm FINISH does not disturb this set too much.

To prove (i) we observe that if $G_{n,m}$ denotes the standard random graph with vertex set $[n]$ and m edges then G_{n,m_1} can be generated by adding a random number of extra random edges to the graph obtained by deleting loops and coalescing multiplied edges in G_1 .

Now it was shown in [8] that if $r_1 \geq r_\epsilon$ then

$$Pr \left(\alpha(G_{n,m_1}) \leq \frac{2n}{r_1} (\log r_1 - \log \log r_1 + 1 - \log 2 - \epsilon) \right) \leq \exp \left\{ -\frac{An}{r_1(\log r_1)^2} \right\}$$

for some "constant" $A = A(\epsilon)$.

It follows from this and the fact that $r_1 = r(1 - O(\log r/r^{1/2}))$ that

$$Pr \left(\alpha(G_1) \leq \alpha_\epsilon = \frac{2n}{r} (\log r - \log \log r + 1 - \log 2 - \epsilon) \right) \leq \exp \left\{ -\frac{Bn}{r(\log r)^2} \right\}, \tag{1}$$

where $B = B(\epsilon)$.

We now show that the transition from F_2 to F does not create too many edges contained in a given large independent set of G_1 and G_2 .

Now let $d'_j = d(j, G_2)$ for $j \in [n]$ and $S_0 = \{j: d'_j \leq r - 3r^{1/2} \log r\}$. Our next task is to prove

$$Pr \left(|S_0| \geq \frac{n}{r^{1+\theta}} \right) \leq e^{-Cn/r^{3+2\theta}} \tag{2}$$

for some constant $C > 0$.

Now if $k \in S_0$ then either

- (a) $k \in S_1 = \{j: d_j \leq r - 2r^{1/2} \log r\}$ or
- (b) $k \in S_2 = \{j: d_j - d'_j \geq r^{1/2} \log r\}$.

Now

$$Pr(1 \in S_1) = Pr \left(B \left(2m_1, \frac{1}{n} \right) \leq r_1 - r^{1/2} \log r \right)$$

(where $B(\cdot, \cdot)$ denotes a binomial random variable)

$$\leq e^{-(\log r)^{2/3}}$$

from the Chernoff bound for the tails of the binomial. Thus

$$E(|S_1|) \leq ne^{-(\log r)^2/3}.$$

Now the events $i \in S_1, j \in S_1$ for $i \neq j$ are not independent. But on the other hand changing any x_i can only change $|S_1|$ by at most one and so, by the Hoeffding–Azuma inequality [1]

$$Pr(|S_1| \geq ne^{-(\log r)^2/3} + u) \leq \exp \left\{ -\frac{2u^2}{rn} \right\}. \tag{3}$$

This inequality implies that if $\xi = \xi(X)$ is a random variable such that

$$|\xi(X) - \xi(X')| \leq d$$

whenever X and X' differ only in one component, then

$$Pr(\xi - E(\xi) \geq u) \leq \exp \left\{ -\frac{2u^2}{2m_1 d^2} \right\}$$

(see, for example, Bollobás [7] or McDiarmid [10]).

To handle S_2 we define $\delta_{ij}, i \in [m_1], j \in [n]$ by

$$\delta_{i,j} = \begin{cases} 1 & \text{if (a) } j \in \{x_{2i-1}, x_{2i}\}, \\ & \text{(b) } \max\{d_{x_{2i-1}}, d_{x_{2i}}\} > r \\ 0 & \text{otherwise.} \end{cases}$$

Then $j \in S_2$ implies $\sum_{i=1}^{m_1} \delta_{i,j} \geq r^{1/2} \log r$. So let $S'_2 = \{j : \sum_{i=1}^{m_1} \delta_{ij} \geq r^{1/2} \log r\} \supseteq S_2$ and observe that S'_2 depends only on X , unlike S_2 which depends on p_1, \dots, p_{2m_1} as well. Let now $\delta = \delta_{1,1}$. Then

$$\begin{aligned} Pr(\delta = 1) &\leq 2Pr(x_1 = 1 \text{ and } d_1 > r) + 2Pr(x_2 = 1 \text{ and } d_1 > r) \\ &= \frac{2}{n} (Pr(d_1 > r \mid x_1 = 1) + Pr(d_1 > r \mid x_1 = 2)) \\ &\leq \frac{4}{n} Pr(d_1 > r \mid x_1 = 1) \\ &= \frac{4}{n} Pr \left(B \left(2m_1 - 1, \frac{1}{n} \right) \geq r \right) \\ &\leq \frac{4}{n} e^{-(\log r)^2/4} \quad (\text{using the Chernoff bound}). \end{aligned}$$

Hence

$$E\left(\sum_{i=1}^{m_1} \delta_{i,j}\right) \leq 2re^{-(\log r)^{2/4}} \quad j \in [n]$$

and so

$$Pr\left(\sum_{i=1}^{m_1} \delta_{i,j} \geq r^{1/2} \log r\right) \leq 2(r^{1/2}/\log r) e^{-(\log r)^{2/4}} \quad j \in [n]$$

and so $E(|S'_2|) \leq 2n(r^{1/2}/\log r) e^{-(\log r)^{2/4}}$.

Now changing any x_i can change at most 4 $\delta_{i,j}$'s and hence $|S'_2|$ by at most 4 and so by the Hoeffding–Azuma inequality

$$Pr(|S'_2| \geq 2n(r^{1/2}/\log r) e^{-(\log r)^{2/4}} + u) \leq \exp\left\{-\frac{u^2}{8rn}\right\}. \quad (4)$$

Inequality (2) follows from (3) and (4) with $u = n/3r^{1+\theta}$. So let us now assume that $\alpha(G_2) > \alpha_\epsilon$ (see (1)) and $|S_0| < n/r^{1+\theta}$. We consider the effect of FINISH.

Let T be an independent set of vertices of G_2 of size $\lceil \alpha_\epsilon \rceil$. Assume that $W - Z = \{x_1, x_2, \dots, x_{2s}\}$ where $\psi(x_j) \in T$ iff $j \in \{1, 3, 5, 7, \dots, 2\tau - 1\}$. We must estimate the number γ of bad edges which (i) are in $\mu(F)$ and (ii) are contained in T .

Note that an execution of the statement σ_i :

$$F := \text{ADD}(F, x_{2i-1}, x_{2i})$$

can only contribute to γ if $i \leq \tau$ and that

$$\begin{aligned} \tau &\leq r |S_0| + 3r^{1/2} \log r |T| \\ &\leq r \frac{n}{r^{1+\theta}} + 3r^{1/2} \log r \lceil \alpha_\epsilon \rceil \\ &\leq \frac{2n}{r^\theta}. \end{aligned}$$

But σ_i creates a bad edge only if $\{x_{2i-1}, x_{2i}\}$ is not added to F and the pair $\{z_1, z_2\} \in F$ satisfies $\{\psi(z_1), \psi(z_2)\} \cap T \neq \emptyset$. Hence

$$\begin{aligned} Pr(\sigma_i \text{ creates a bad edge}) &\leq |\{z_1, z_2\} \in F : \{\psi(z_1), \psi(z_2)\} \cap T \neq \emptyset\}| / |F_2| \\ &\leq \frac{\lceil \alpha_\epsilon \rceil r}{|F_2|} \leq 4 \frac{\log r}{r} \end{aligned}$$

regardless of the outcome of the execution of FINISH to this point.

Hence γ is majorised by $B(\lceil 2n/r^{\theta} \rceil, 4 \log r/r)$ and so

$$\Pr \left(\gamma \geq \frac{16n}{r^{1+\theta}} \right) \leq \exp \left\{ -\frac{8n}{3r^{1+\theta}} \right\} \quad (5)$$

using the Chernoff bound.

Note that $\alpha(\mu(F)) \geq \alpha(G_2) - 2\gamma$.

Thus (1), (2), and (5) (and a surreptitious doubling of ε) imply that

$$\Pr(\alpha(F) \leq \alpha_\varepsilon) \leq e^{-Dn/r^{3+2\theta}} \quad (6)$$

for some $D = D(\varepsilon)$.

To bound $\alpha(F)$ from above is straightforward.

Let now $l = \lceil \alpha_{-\varepsilon} \rceil$ and Y be the random variable which counts the number of independent sets of $\mu(F)$ of size l . Then

$$\begin{aligned} P(\alpha(F) \geq l) &\leq E(Y) \\ &= \binom{n}{l} \prod_{i=1}^{l-1} \left(1 - \frac{rl-i}{rn-2i+1} \right) \\ &\leq \binom{n}{l} \prod_{i=1}^{l-1} \left(1 - \frac{rl-i}{rn} \right) \\ &\leq 2 \binom{n}{l} \exp \left\{ -\frac{rl^2}{2n} \right\} \\ &\leq 2 \left(\frac{ne}{l} \exp \left\{ -\frac{rl}{2n} \right\} \right)^l \\ &\leq 2e^{-l/2}. \end{aligned}$$

Hence, the first part of the theorem for constant r follows from (0), (6), and (7).

Now to the second part of the theorem.

The lower bound is immediate from the first part of the theorem, since $\chi(G_r) \geq n/\alpha(G_r)$. For the upper bound we use the fact that the main result of Łuczak [12] implies that for $r \geq r_0$ (= some sufficiently large constant)

$$\Pr \left(\chi(G_1) \geq k_0 = \left\lceil \frac{r}{2 \log r} \left(1 + \frac{30 \log \log r}{\log r} \right) \right\rceil \right) = o(1),$$

G_1 as in Step 1.

Step 2 can only decrease the chromatic number and (6) shows that

if G_1 has a k_0 -colouring and we use it for $\mu(F)$ then with probability $1 - o(1)$ $\mu(F)$ has at most

$$\frac{16n}{r^{1+\theta}} \cdot \frac{r}{2 \log r} \left(1 + \frac{30 \log \log r}{\log r} \right) \leq \frac{10n}{r^\theta}$$

edges with both ends of the same colour. The result (for constant r) will follow if we show that with probability $1 - o(1)$, all subgraphs of $\mu(F)$ with at most $s_0 = 20n/r^\theta$ vertices can be (re-)coloured with at most $l = r \log \log r / (\log r)^2$ colours. We prove this by showing that any subgraph H induced by $s \leq s_0$ vertices satisfies $\delta(H) < l$ and this is in turn implied by each such H having less than $ls/2$ edges. This latter statement is easy to prove.

$Pr(\exists s \leq s_0$ vertices of $\mu(F)$ containing $\geq ls$ edges)

$$\begin{aligned} &\leq \sum_{s=l}^{s_0} \binom{n}{s} \binom{\binom{s}{2}}{ls/2} \left(\frac{r^2}{rn - rs} \right)^{ls/2} \\ &\leq \sum_{s=l}^{s_0} \left(\frac{ne}{s} \right)^s \left(\frac{s^2 e}{ls} \right)^{ls/2} \left(\frac{r}{n-s} \right)^{ls/2} \\ &\leq \sum_{s=l}^{s_0} \left(\left(\frac{s}{n} \right)^{(l-2)/l} \frac{3r}{l} \right)^{ls/2} = o(1) \end{aligned}$$

and the whole theorem has been proved, for constant r .

Let us now consider the case $r \rightarrow \infty$ but $r = o(n^\theta)$. The above proof shows that $\mu(F)$ for a random $F \in \Omega_W$ has its independence and chromatic numbers in the right range. We have to show that this implies the same for G_r . We rely on the work of McKay and Wormald [11] for this. They give a procedure DEG which takes as input a random $F \in \Phi_W$ and tries to construct an r -regular simple graph by eliminating loops and multiple edges. The elimination of a loop or multiple edge involves the addition of at most 4 new edges. The procedure succeeds with probability $1 - o(1)$ and it produces each member of $\text{REG}(n, r)$ with the same probability. Also, it is easy to see that with probability $1 - o(1)$ F has $O(r^2)$ loops and multiple edges. Thus we need only show that adding $O(r^2)$ edges to a typical F does not change α or χ by much. But now $r^2 = o(n/r)$ for $r = o(n^\theta)$, and so part (a) requires no work. For part (b) we need to be convinced that the $O(r^2) = o(n^{2\theta})$ added edges are sufficiently random so that they usually induce the union of 4 forests which can be 8-coloured. This can be done fairly straightforwardly but requires a fair amount of detail from [11] which is inappropriate here.

ACKNOWLEDGMENT

We thank Colin McDiarmid for pointing out some subtle errors.

REFERENCES

1. K. AZUMA, Weighted sums of certain dependent random variables, *Tôhoku Math. J.* **19** (1967), 357–367.
2. E. A. BENDER AND E. R. CANFIELD, The asymptotic number of labelled graphs with given degree sequences, *J. Combin. Theory Ser. A* **24** (1978), 296–307.
3. B. BOLLOBÁS, “Random Graphs,” Academic Press, Orlando, FL, 1985.
4. B. BOLLOBÁS, A probabilistic formula for the number of labelled regular graphs, *European J. Combin.* **1** (1980), 311–316.
5. B. BOLLOBÁS, The chromatic number of random graphs, *Combinatorica* **8** (1988), 49–55.
6. B. BOLLOBÁS AND P. ERDŐS, Cliques in random graphs, *Math. Proc. Cambridge Philos. Soc.* **80** (1976), 419–427.
7. B. BOLLOBÁS, Martingales, isoperimetric inequalities and random graphs, in “Combinatorics” (A. Hajnal, L. Lovász, and V. T. Sos, Eds.), Colloq. Math. Sci. Janos Bolyai, Vol. 52, North-Holland, Amsterdam, 1988.
8. A. M. FRIEZE, On the independence number of a random graphs, *Discrete Math.*, in press.
9. G. R. GRIMMETT AND C. J. H. MCDIARMID, On colouring random graphs, *Math. Proc. Cambridge Philos. Soc.* **77** (1975), 313–324.
10. C. J. H. MCDIARMID, On the method of bounded differences, in “Surveys in Combinatorics” (J. Siemons, Ed.), pp. 148–188, London Mathematical Society Lecture Notes, Vol. 141, Cambridge Univ. Press, London/New York, 1989.
11. B. D. MCKAY AND N. C. WORMALD, Uniform generation of random regular graphs of moderate degree, to appear.
12. T. ŁUCZAK, The chromatic number of random graphs, *Combinatorica*, in press.
13. D. MATULA, On the complete subgraphs of a random graphs, in “Combinatory Mathematics and Its Applications,” pp. 356–369, Chapel Hill, 1970.
14. E. SHAMIR AND J. SPENCER, Sharp concentration of the chromatic number on random graphs $G_{n,p}$, *Combinatorica* **7** (1987), 121–129.