Combinatorics, Probability and Computing (2002) 11, 249–261. © 2002 Cambridge University Press
DOI: 10.1017/S0963548301005090 Printed in the United Kingdom

Random Regular Graphs of Non-Constant Degree: Connectivity and Hamiltonicity

COLIN COOPER,[1]† ALAN FRIEZE[2]‡ and BRUCE REED[3]

[1] Department of Mathematical and Computing Sciences, Goldsmiths College, University of London, New Cross, London SE14 6NW, UK
(e-mail: c.cooper@gold.ac.uk)

[2] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, USA
(e-mail: alan@random.math.cmu.edu)

[3] Equipe Combinatoire, CNRS, Université de Paris VI, 4 Place Jussieu, Paris 75005, France
(e-mail: reed@ecp6.jussieu.fr)

Received 24 April 2000; revised 10 September 2001

Let G_r denote a graph chosen uniformly at random from the set of r-regular graphs with vertex set $\{1, 2, \ldots, n\}$, where $3 \leqslant r \leqslant c_0 n$ for some small constant c_0. We prove that, with probability tending to 1 as $n \to \infty$, G_r is r-connected and Hamiltonian.

1. Introduction

The properties of random r-regular graphs have received much attention. For a comprehensive discussion of this topic, see the recent survey by Wormald [22] or Chapter 9 of the book *Random Graphs*, by Janson, Łuczak and Ruciński [12].

A major obstacle in the development of the subject has been a lack of suitable techniques for modelling simple random graphs over the entire range, $0 \leqslant r \leqslant n-1$, of possible values of r. The classical method for generating uniformly distributed simple r-regular graphs is by rejection sampling, using the configuration model of Bollobás [3]. The configuration model is a probabilistic interpretation of a counting formula of Bender and Canfield [2]. The method is most easily applied when r is constant or grows slowly with n, the number of vertices, as n tends to infinity. The formative paper [3] on this topic considered the case where $r = O((\log n)^{1/2})$. McKay [16] and McKay and Wormald [17, 18] subsequently gave alternative approaches which are useful for $r = o(n^{1/2})$ or $r = \Omega(n)$.

† Research supported by the STORM Research Centre, UNL.
‡ Supported in part by NSF grant CCR-9818411.

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

We use *edge switching techniques* extensively in this paper, and note that these techniques have been successfully applied in a number of places, *e.g.*, [16], [17, 18], [9], [14] and [13].

Let G_r denote a graph chosen uniformly at random from the set $\mathscr{G}_r$ of simple r-regular graphs with vertex set $V = \{1, 2, \dots, n\}$. We consider properties of simple r-regular graphs for the case where $r \to \infty$ as $n \to \infty$, but $r = o(n)$. The properties we study are vertex r-connectivity and Hamiltonicity. These properties are also studied in a recent paper by Krivelevich, Sudakov, Vu and Wormald [13], for the case where $r(n) \geqslant \sqrt{n} \log n$. Our paper complements [13] both in the range of r studied and in the techniques applied.

Theorem 1.1. *There is a small positive absolute constant c_0 with the following property. Assume $3 \leqslant r \leqslant c_0 n$. Then, with probability tending to 1 as $n \to \infty$,*

(a) *G_r is r-connected,*
(b) *G_r is Hamiltonian.*

The results of Theorem 1.1 are well known for r constant. Result (a) is from Bollobás [4] and (b) is from Robinson and Wormald [20, 21], Bollobás [5], Fenner and Frieze [8]. For $r = o(n^{1/2})$ such results could have been proved with the help of the models of [16] and [17]. In fact this was done, for Hamiltonicity, up to $r = o(n^{1/5})$, in an unpublished work by Frieze [9], and for r-connectivity, up to $r \leqslant n^{0.002}$, by Łuczak [15].

As [13] proves the case where $r \geqslant n^{1/2} \log n$, this implies G_r is r-connected and Hamiltonian w.h.p.[1] for all $3 \leqslant r \leqslant n - 4$.

2. Generating graphs with a fixed degree sequence

Let $\mathbf{d} = (d_1, d_2, \dots, d_n)$, and let $2D = (d_1 + d_2 + \cdots + d_n)$. Let $\mathscr{G}_\mathbf{d}$ be the set of simple graphs G with vertex set $V = [n]$, degree sequence $\mathbf{d}$, and D edges.

Let Ω be the set of all $(2D)!/(D!2^D)$ partitions of $W = [2D]$ into D 2-element sets. An element of Ω is a *configuration*. The constituent 2-element sets of a configuration F are referred to as the *edges* of F.

Let $W_1, W_2, \dots, W_n$ be the natural ordered partition $P_\mathbf{d}$ of $W = [2D]$ into sets of size $|W_i| = d_i$, and where $(\max W_i) + 1 = \min W_{i+1}$ for $i < n$.

Let $\Omega_\mathbf{d}$ be Ω with the understanding that the underlying set W is partitioned into P_d. The *degree sequence* of an element F of $\Omega_\mathbf{d}$ is $\mathbf{d}$. We often write Ω for $\Omega_\mathbf{d}$ when the context is clear. Define $\phi_{P_\mathbf{d}} : W \to [n]$ by $\phi(w) = i$ if $w \in W_i$. Let $\gamma(F)$ denote the multigraph with vertex set $[n]$ and edge multiset $E_F = \{\{\phi(x), \phi(y)\} : \{x, y\} \in F\}$.

Definition. Let $\Omega^*_\mathbf{d}$ denote those configurations F for which $\gamma(F)$ is simple relative to $P_\mathbf{d}$.

Remark 1. Note that each member of $\mathscr{G}_\mathbf{d}$ is the image under γ of precisely $\prod_{i=1}^n d_i!$ members of $\Omega^*_\mathbf{d}$. Thus, sampling F uniformly from $\Omega^*_\mathbf{d}$ induces the uniform measure on $\gamma(F)$ and is equivalent to sampling uniformly from $\mathscr{G}_\mathbf{d}$.

1 A sequence of events $\mathscr{E}_n$ is said to occur *with high probability* (w.h.p.) if $\lim_{n\to\infty} \mathbf{Pr}(\mathscr{E}_n) = 1$.

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

If $d_i = r$, $(1 \leqslant i \leqslant n)$ we will say the configuration F is r-regular. The probability $|\Omega^*|/|\Omega|$ that the underlying r-regular multigraph $\gamma(F)$ of such a configuration F is simple is $\exp(-\Theta(r^2))$. For $r = o(n^{1/2})$ this follows from [17, 18], and for larger values of r from Lemma 2.2 below. This result allows us to prove many results directly via configurations and then condition the probability estimates for simple graphs.

Lemma 2.1. *Let* $\Delta = \max_{i\in[n]} d_i$. *Suppose that* $\Delta \leqslant n/1000$ *and that* $\mathbf{d}$ *satisfies* $\min_{i\in[n]} d_i \geqslant \Delta/4$. *Given* $a, b \in [n]$, *if* G *is sampled u.a.r. from* $\mathscr{G}_{\mathbf{d}}$, *then*

$$\mathbf{Pr}(\{a,b\} \in E(G)) \leqslant \frac{20\Delta}{n}.$$

Proof. Let

$$\Omega_1 = \{G \in \mathscr{G}_{\mathbf{d}} : \{a,b\} \in E(G)\} \text{ and } \Omega_2 = \mathscr{G}_{\mathbf{d}} \setminus \Omega_1.$$

We consider the set X of pairs $(G_1, G_2) \in \Omega_1 \times \Omega_2$ such that G_2 is obtained from G_1 by deleting disjoint edges $\{a,b\}, \{x_1,y_1\}, \{x_2,y_2\}$ and replacing them by $\{a,x_1\}, \{y_1,y_2\}$, $\{b,x_2\}$. Given G_1, we can choose $\{x_1,y_1\}, \{x_2,y_2\}$ to be any ordered pair of disjoint edges which are not incident with a, b or their neighbours and such that $\{y_1,y_2\}$ is not an edge of G_1. Thus, each $G_1 \in \Omega_1$ is in at least $(D-(2\Delta^2+1))(D-(4\Delta^2+2))$ pairs. Each $G_2 \in \Omega_2$ is in at most $2D\Delta^2$ pairs. The factor of 2 arises because a suitable edge $\{y_1,y_2\}$ of G_2 has an orientation relative to the switching back to G_1. As $D \geqslant n\Delta/8$, it follows that

$$\frac{|\Omega_1|}{|\Omega_2|} \leqslant \frac{2D\Delta^2}{(D-(2\Delta^2+1))(D-(4\Delta^2+2))} \leqslant \frac{20\Delta}{n}. \qquad \square$$

Lemma 2.2. *Suppose* $100 \leqslant r \leqslant n/1000$. *Let* $d_j = r$, $1 \leqslant j \leqslant n$. *If* F *is chosen uniformly at random* (*u.a.r.*) *from* Ω *then, for sufficiently large* n,

$$\mathbf{Pr}(F \in \Omega^*) \geqslant e^{-2r^2}.$$

Proof. Consider the following algorithm from Frieze and Łuczak [11].

Algorithm GENERATE
begin
$D := rn/2$
$F_0 := \emptyset$
Let $\boldsymbol{\sigma} = (x_1, x_2, \ldots, x_{2D-1}, x_{2D})$ be an ordering of W
For $i = 1$ **to** D **do**
begin

$$F_i := \begin{cases} F_{i-1} \cup \{\{x_{2i-1}, x_{2i}\}\} & \left(\text{with probability } \frac{1}{2i-1}\right) \quad \mathbf{A} \\ F_{i-1} \cup \{\{x_{2i-1}, z_1\}, \{x_{2i}, z_2\}\} - \{z_1, z_2\} & \left(\text{with probability } \frac{2i-2}{2i-1}\right) \quad \mathbf{B} \end{cases}$$

Here $\{z_1, z_2\}$ is chosen u.a.r. from F_{i-1} and z_1 is chosen u.a.r. from $\{z_1, z_2\}$.
end
Output $F := F_D$
end

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

We first prove that GENERATE produces a u.a.r. member of Ω whatever the ordering $\boldsymbol{\sigma} = (x_1, x_2, \dots, x_{2D})$ of W. We then describe an ordering $\boldsymbol{\sigma}$ from which we can prove the lemma.

Let $W^{(i)} = (x_1, x_2, \dots, x_{2i})$ and let Ω_i be the set of configurations of $W^{(i)}$. We show inductively that F_i is a random member of Ω_i. This is clearly true for $i = 1$, and so assume that, for some $i \geqslant 2$, we have that F_{i-1} is chosen u.a.r. from Ω_{i-1}.

Now consider a bipartite graph H with vertex bipartition (Ω_{i-1}, Ω_i) and an edge (F, F') whenever $F' = F \cup \{x_{2i-1}, x_{2i}\}$ or $F' = (F \setminus \{a, b\}) \cup \{\{a, x_{2i-1}\}, \{b, x_{2i}\}\}$ for some $\{a, b\} \in F$. Each $F \in \Omega_{i-1}$ has degree $2i - 1$ in H and each $F' \in \Omega_i$ has degree 1. Our algorithm chooses F uniformly from Ω_{i-1} (induction) and then uniformly chooses an H-edge leaving F. This implies uniformity in Ω_i.

Label the configuration points in set W_k of the partition, as $\{(k-1)r + j : 1 \leqslant j \leqslant r\}$. For the ordering $\boldsymbol{\sigma}$ of W, we specify that x_i is always chosen as one of the remaining points for which $\phi(x_i)$ occurs as little as possible in the sequence $(\phi(x_1), \dots, \phi(x_{i-1}))$. To be specific, when $i = (j-1)n + k$, $(1 \leqslant k \leqslant n,\ 1 \leqslant j \leqslant r)$, define x_i to be the point in W_k with label $(k-1)r + j$.

Let $\Omega_i^* = \{F \in \Omega_i : \gamma(F) \text{ is simple}\}$. Let $\Delta_i = \lceil 2i/n \rceil$ denote the maximum degree in $\gamma(F_i)$. Let the edge $\{\phi(x_{2i-1}), \phi(x_{2i})\} = \{a, b\}$ and let $\{\phi(z_1), \phi(z_2)\} = \{c, d\}$. We will prove that

$$\mathbf{Pr}(F_i \in \Omega_i^* \mid F_{i-1} \in \Omega_{i-1}^*) \geqslant \begin{cases} 1, & 2i \leqslant n, \\ \left(1 - \frac{60\Delta_i}{(2i-1)n} - \frac{2\Delta_i^2 + 2\Delta_i}{i-1}\right), & n < 2i \leqslant rn. \end{cases} \tag{2.1}$$

If $i \leqslant n/2$ then F_i induces a matching. If $i > n/2$ and if, at the ith step of GENERATE, $\{a, b\}$ already exists in case A or is equal to $\{c, d\}$ in case B, then F_i will not be simple. The probability that the edge $\{a, b\}$ exists, in the corresponding simple random graph, is at most $\frac{20\Delta_i}{n}$, by Lemma 2.1. Thus the probability the edge exists (case A) or exists and is selected (case B) is at most

$$\frac{20\Delta_i}{n}\left(\frac{1}{2i-1} + \frac{2i-2}{2i-1}\frac{1}{i-1}\right) = \frac{60\Delta_i}{(2i-1)n}.$$

Assume now that the ith step is type B and $\{a, b\} \neq \{c, d\}$.

When $\{a, b\} \cap \{c, d\} \neq \emptyset$, a loop may be created. This happens with probability at most $2\Delta_i/(i-1)$.

When one of a, b is adjacent to c or d, a parallel edge may be created. This happens with probability at most $2\Delta_i^2/(i-1)$.

All cases have been covered and the result follows from iterating (2.1) for $i \leqslant rn/2$. □

Remark 2. In Lemma 4.3 we need to run algorithm GENERATE starting with a configuration F_0 on $[2D']$ and and restricting our random choice of $\{z_1, z_2\}$ to $F \setminus F_0$. The output is then F_0 plus a random configuration on $W = [2D' + 1, 2D' + 2D]$.

At this point we describe a simpler algorithm CONSTRUCT for obtaining a u.a.r. configuration.

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

Algorithm CONSTRUCT
begin
$F_0 := \emptyset$; $R_0 := W := [2D]$
For $i = 1$ **to** D **do**
begin
Choose $u_i \in R_{i-1}$ *arbitrarily*
Choose v_i *uniformly at random* from $R_{i-1} \setminus \{u_i\}$
$F_i := F_{i-1} \cup \{\{u_i, v_i\}\}$; $R_i := R_{i-1} \setminus \{u_i, v_i\}$
end
Output $F := F_D$.
end

Remark 3. Neither of the algorithms generating F_D use any information about the partition $P_{\mathbf{d}}$ associated with the configuration. After iteration i, F_i is a u.a.r. element of Ω_i. We can, if we wish, complete a certain number I of iterations using CONSTRUCT and then switch to GENERATE. Instead of initializing the ordering $\boldsymbol{\sigma}$ used in algorithm GENERATE with W, we initialize $\boldsymbol{\sigma}$ with R_I, the remaining unmatched points.

3. r-connectivity

We now prove Theorem 1.1(a). Since the result is already known for r constant, we can assume that $10^6 \leqslant r \leqslant c_0 n$, where c_0 is sufficiently small.

For a simple graph G with edge set E, the disjoint neighbour set, $N(S)$, of a set of vertices S is defined as $N(S) = \{w \notin S : \exists v \in S \text{ s.t. } \{v, w\} \in E\}$. When S is a singleton $\{v\}$ we use the notation $N(v)$.

Lemma 3.1. *Let $\mathscr{Q}_1 \subseteq \mathscr{G}_r$ be the event that, for all vertices $v, w \in V$ of G_r,*

(a) *if $r = o(n)$ then $|N(v) \cap N(w)| \leqslant 10 + o(r)$,*
(b) *if $\log^2 n \leqslant r \leqslant n$ then $|N(v) \cap N(w)| \leq r^2/n + 5\sqrt{r \log n}$.*

Then $\mathbf{Pr}(\overline{\mathscr{Q}_1}) = O(1/n^2)$.

Proof. Throughout this proof, we fix a vertex v and the set $S = N(v)$, of vertices which are the (disjoint) neighbours of v. Let w be a fixed vertex of $V - v$.

Let $\mathscr{F}(S) = \{G : G = G_r - v,\ N(v) = S\}$ be the set of graphs G with vertex set $V - v$ formed by deleting v from those r-regular graphs, G_r, for which $N(v) = S$. Thus $|S| = r$, and the vertices in S have degree $r - 1$ in G.

The vertex w partitions $\mathscr{F}$ into sets $\mathscr{F}(k) = \{G : |N(w) \cap S| = k\}$ where $0 \leqslant k \leqslant r$ if $w \notin S$ and $0 \leqslant k \leqslant r - 1$ if $w \in S$.

For sets $R, T \subseteq V - v$ let $\mathscr{N}(R, T) = \mathscr{N}(R, T;\ S, w)$ be the set of graphs in $\mathscr{F}$ with $N(w) \cap S = R$ and $N(w) - S = T$. If $|R| < |S - w|$, choose $x \in (S - w) \setminus R$ and $a \in T$. We consider a bipartite graph $\mathscr{B}$ with left vertex set $\mathscr{N}(R, T)$ and right vertex set $\mathscr{N}(R + x, T - a)$.

If $G \in \mathscr{N}(R, T)$ and $\{w, a\}, \{x, b\}$ are edges of G we make a *switching* $G : (wa, xb) \to$

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

(wx, ab) in which edges $\{w,a\},\{x,b\}$ are replaced by $\{w,x\},\{a,b\}$ provided the resulting graph G' is simple. These switchings define the edges of $\mathscr{B}$, and $d_L(G)$ (resp. $d_R(G')$) is the number of edges incident with G (resp. G') in $\mathscr{B}$.

Let $v(a,x;G) = |N(a) \cap N(x)|$ be the number of common neighbours of a and x in G. Let $\delta(a,x;G) = 1$ if $a \in N(x)$.

Considering the possibilities for b when the switching $G : (wa, xb) \to (wx, ab)$ gives a simple G', we have

$$d_L(G) = |N(x)| - v(a,x;G) - \delta(a,x;G)$$

because G' is simple if and only if $b \neq a$ and $b \notin N(a)$. Here $|N(x)| = r-1$ as $x \in S$. The switching leaves $\delta(a,x;G') = \delta(a,x;G)$ and $v(a,x;G') = v(a,x;G)$ as $(\{a\} \cup N(a)) \cap N(x)$ is the same set in both graphs.

Considering the switching $G' : (wx, ab) \to (wa, xb)$ giving G, we have

$$d_R(G') = |N(a)| - v(a,x;G') - \delta(a,x;G').$$

We note that $|N(a)| = r$ as $a \notin S$.

The graph $\mathscr{B}$ consists of components within which δ, v (and hence d_L, d_R) are invariant. Consider a component with bipartition size (N_L, N_R). We now prove that $N_L \geqslant N_R$. In any component with edges we have $d_R = d_L + 1$ so that $N_R = N_L d_L/(d_L+1)$. The case $(N_L, N_R) = (0,1)$ of isolated vertices in the right bipartition cannot occur. For, in G',

$$v(a,x;G') + \delta(a,x;G') \leqslant |N(x) - w| = r - 2$$

and so

$$d_R(G') = |N(a)| - v - \delta \geqslant 2.$$

Thus

$$|\mathscr{N}(R,T)| \geqslant |\mathscr{N}(R+x, T-a)|.$$

Given S and w, the size of $\mathscr{N}(R,T;\ S,w)$ is invariant for all $R, T, |R| = k$ by a simple symmetry argument.

Let $|\mathscr{N}(R,T;\ S,w)| = \eta(k)$. Thus $\eta(k)$ is a non-increasing function of k. Let $f(k) = |\mathscr{F}(k)|$ be the number of graphs in $\mathscr{F}$ with $|N(w) \cap S| = k$. If $w \notin S$ then for all $k \geqslant 0$, $f(k) = \binom{r}{k}\binom{n-2-r}{r-k}\eta(k)$. Similarly, if $w \in S$ then, for all $k \geqslant 0$, $f(k) = \binom{r-1}{k}\binom{n-1-r}{r-1-k}\eta(k)$.

Suppose G is chosen u.a.r. from $\mathscr{F}(S)$ and let $Z(G) = |R|$. Then $\mathbf{Pr}(Z=k) = f(k)/|\mathscr{F}|$. Writing $N = n-2, \rho = r - 1_{w \in S}$,

$$\mathbf{Pr}(Z=k) = \binom{\rho}{k}\binom{N-\rho}{\rho-k}\frac{\eta(k)}{|\mathscr{F}|}.$$

Let X be a hypergeometric random variable with $\mathbf{Pr}(X=k) = \binom{\rho}{k}\binom{N-\rho}{\rho-k}/\binom{N}{\rho}$. Then $\mathbf{Pr}(Z=k)/\mathbf{Pr}(X=k)$ decreases with k. It follows that $\mathbf{Pr}(Z \geqslant k) \leqslant \mathbf{Pr}(X \geqslant k)$ for any k.

The hypergeometric random variable X has mean $\mu = \rho^2/N$. The proportional error in bounding $\mathbf{Pr}(X=j)$ above by $\mathbf{Pr}(Y=j)$, where Y is the binomial random variable

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

$B(\rho, \rho/N)$, is at most $\exp(\rho^2/(N-\rho))$ (see [7, p. 57]). Thus, provided $r = o(\sqrt{n})$, using the following bound (3.1) on binomial tails (see [1]),

$$\mathbf{Pr}(Y \geqslant \beta\mu) \leqslant \left(\frac{e}{\beta}\right)^{\beta\mu}, \tag{3.1}$$

we see that

$$\mathbf{Pr}(X \geqslant \beta\mu) \leqslant 2\left(\frac{e}{\beta}\right)^{\beta\mu}.$$

If $r \leqslant \log^2 n$ let $k = \alpha\rho + 10$, $\alpha = 1/\log\log n$, then

$$\mathbf{Pr}(X \geqslant \alpha\rho + 10) \leqslant 2\left(\frac{e\rho^2}{(\alpha\rho+10)(n-2)}\right)^{\alpha\rho+10} = o(n^{-4}).$$

For $\log^2 n \leqslant r \leqslant n$ let $k = \rho^2/(n-2) + 4\sqrt{\rho \log n}$. We can apply Azuma's inequality to the 0, 1 sequence of observations of the sampling process of X, with $c_i = 1$ to infer that

$$\mathbf{Pr}(X \geqslant \rho^2/(n-2) + 4\sqrt{r\log n}) = o(n^{-4}).$$

Note that if $r \geqslant \log^2 n$ and $r = o(n)$ then the bound in (b) implies that in (a). □

Lemma 3.2. *Let $\mathscr{Q}_2$ be the event that no set of vertices $U \subset V$ of G_r, $1 \leqslant |U| \leqslant n/70$, induces more than $r|U|/12$ edges. Then* $\mathbf{Pr}(\mathscr{Q}_2) = 1 - O(1/n^2)$.

Proof. Let $\beta = 1/12$ and $\theta = 1/70$. Let $|U| = u$.

Note first that, in a simple r-regular graph, a set of size u induces at most $\binom{u}{2}$ edges and, provided $u \leqslant 2\beta r$,

$$\binom{u}{2} \leqslant \beta r u.$$

Let $\mathscr{E} = \{F \in \Omega^* : \text{no vertex set } U,\ 2\beta r \leqslant |U| \leqslant \theta n \text{ induces more than } \beta r|U| \text{ edges}\}$. It suffices to prove that $\mathbf{Pr}(\overline{\mathscr{E}}) = O(n^{-2})$.

In Ω the number of edges X falling inside a set U is dominated by a binomial random variable $Y \sim B(ur, u/(n-u))$, in which each configuration point of U independently selects a pairing on the assumption that all configuration points of U are available, and that ru configuration points of $V \setminus U$ are unavailable. Now, $\mathbf{E}Y = ru^2/(n-u)$ and

$$\begin{aligned}
\mathbf{Pr}_\Omega(Y \geqslant \beta r u) &= \mathbf{Pr}(Y \geqslant (\beta(n-u)/u)\mathbf{E}Y) \\
&\leqslant \left(\frac{ue}{\beta(n-u)}\right)^{\beta r u} \qquad \text{by (3.1)} \\
&\leqslant \left(\frac{34u}{n}\right)^{\beta r u}.
\end{aligned}$$

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

As $r \geqslant 10^6$, $\beta r/2 \gg 1$ and so, by Lemma 2.2,

$$\begin{aligned}\mathbf{Pr}\left(\overline{\mathscr{E}}\right) &\leqslant e^{2r^2}\sum_{u=2\beta r}^{\theta n}\binom{n}{u}\left(\frac{34u}{n}\right)^{\beta r u} \leqslant e^{2r^2}\sum_{u=2\beta r}^{\theta n}\left(\frac{ne}{u}\right)^{u}\left(\frac{34u}{n}\right)^{\beta r u}\\ &\leqslant e^{2r^2}\sum_{u=2\beta r}^{\theta n}\left(\frac{34u}{n}\right)^{\beta r u/2} \leqslant 2e^{2r^2}\left(\frac{68\beta r}{n}\right)^{\beta^2 r^2} \leqslant 2\exp\left\{2r^2-\beta^2 r^2\log\frac{n}{6r}\right\}\\ &= O(n^{-2}),\end{aligned}$$

provided $r \leqslant c_0 n$, c_0 sufficiently small. □

Proof of Theorem 1.1(a). Assume the events $\mathscr{Q}_1, \mathscr{Q}_2$ described in Lemmas 3.1 and 3.2. If G_r is not r-connected then there is a separator X of size $x \leqslant r-1$. Let $G_r - X = A + B$ and $|A| = a \leqslant |B| = b$.

Case 1: $2 \leqslant a \leqslant r/2$.
Let $u, v \in A$ be arbitrary. If $r = o(n)$ then, as $\mathscr{Q}_1$ occurs,

$$|N(u) \cup N(v)| \geqslant 2r - |N(u) \cap N(v)| \geqslant 2r - o(r) - 10. \tag{3.2}$$

However,

$$|N(u) \cup N(v)| \leqslant |A \cup X| \leqslant a + r - 1 < 3r/2, \tag{3.3}$$

which contradicts (3.2).

If $cn \leqslant r \leqslant n/4$ for some $c > 0$, we see that because $\mathscr{Q}_1$ occurs we have $|N(u) \cup N(v)| \geqslant (1 - o(1))7r/4$, which contradicts (3.3).

Case 2: $r/2 \leqslant a \leqslant n/80$.
As $|A \cup X| \leqslant a + r - 1$ and $A \cup X$ contains at least $ar/2$ edges, we see that because $\mathscr{Q}_2$ occurs

$$\frac{ar}{2} \leqslant \frac{r}{12}(a + r - 1) \text{ and so } a < r/5.$$

Case 3: $n/80 \leqslant a \leqslant \lceil n/2 \rceil$.
If configuration F is chosen randomly from Ω, then the existence of a separator of size $x \leqslant r-1$, where the smaller component has size $a \geqslant n/80$, has probability at most

$$\sum_{a=n/80}^{\lceil n/2\rceil}\sum_{x=0}^{r-1}\binom{n}{a}\binom{n-a}{x}\left(1-\frac{b}{n}\right)^{ar/2}.$$

Thus from Lemma 2.2 the probability of this event in $\mathscr{G}_r$ is at most

$$e^{2r^2}\sum_{a=n/80}^{\lceil n/2\rceil} 4^n e^{-a(n-(a+r))r/2n} \leqslant e^{-rn/500} = o(1)$$

for $r \leqslant c_0 n$, c_0 sufficiently small. □

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

4. Hamilton cycles

We prove Theorem 1.1(b) on the assumption that $10^7 \leqslant r \leqslant c_0 n$.

Definition. Let $\mathscr{G}_r^*$ denote the subset of $\mathscr{G}_r$ consisting of those graphs G with the following properties:

C1: all sets of vertices U of size at most $n/70$ induce at most $r|U|/12$ edges;
C2: the graph G is connected.

Lemma 3.2 and Theorem 1.1(a) imply the following.

Lemma 4.1. $|\mathscr{G}_r^*| = (1 - o(1))|\mathscr{G}_r|$. □

Given a subset R of the edges of G, let $d_R(v)$ be the number of edges of R which are incident with the vertex v of G.

Definition. Let P be some *fixed* longest path of G. A set of edges $R \subseteq E(G)$ is *deletable* from G, ($R \in \mathrm{Del}(G)$), if:

D1: R avoids P;
D2: For all $v \in V$, $\frac{r}{4} \leqslant d_R(v) \leqslant \frac{r}{2}$.

Lemma 4.2. *Let $G \in \mathscr{G}_r$, and let R be a random subset of the edges of G where each edge of G is placed into R independently with probability $1/3$. Then*

$$\mathbf{Pr}(R \text{ is deletable} \mid G) \geqslant e^{-n}.$$

Proof.

$$\mathbf{Pr}(\mathrm{D1} \mid G) = \left(\frac{2}{3}\right)^{|P|} \geqslant \left(\frac{2}{3}\right)^{n} \geqslant e^{-n/2}.$$

For (D2) we condition on (D1). We use the symmetric version of the Lovász Local Lemma (see, *e.g.*, Alon and Spencer [1]) to show that

$$\mathbf{Pr}(\mathrm{D2} \mid \mathrm{D1}) \geqslant e^{-n/2}.$$

Let A_v be the event $\{d_R(v) \notin [\frac{r}{4}, \frac{r}{2}]\}$; then $\mathbf{Pr}(A_v \mid \mathrm{D1}) \leqslant e^{-r/100}$ and the dependency graph has degree at most r. For large r we can apply the lemma to show that, conditional on D_1,

$$\mathbf{Pr}\left(\bigcap_{v \in V} \overline{A_v} \mid \mathrm{D1}\right) \geqslant (1 - 2e^{-r/100})^n \geqslant e^{-n/2}.$$ □

The size of the set R of deleted edges is binomial $B(rn/2, 1/3)$ and thus w.h.p. $|R| = (1 + o(1))rn/6$. For the purposes of Lemma 4.3 below, we condition on $|R| \in [(0.16)rn, (0.17)rn]$. We note that there exists some $\delta > 10^{-7}$ such that

$$\mathbf{Pr}\left(|R| \notin [(0.16)rn, (0.17)rn]\right) \leqslant e^{-\delta rn}. \tag{4.1}$$

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

Definition. A set of edges S is *addable* to a simple graph H, ($S \in \mathrm{Add}(H)$), if:

A1: $H + S \in \mathscr{G}_r$;

A2: no longest path of H is closed to a cycle by S.

Let

$$\begin{aligned} \mathscr{N} &= \{G \in \mathscr{G}_r^* : G \text{ is not Hamiltonian}\}, \\ \mathscr{E} &= \{(G,R) : G \in \mathscr{N}, R \in \mathrm{Del}(G)\}, \\ \Psi &= \{H : H = G - R,\ (G,R) \in \mathscr{E},\ |R| \in [(0.16)rn, (0.17)rn]\}, \\ \mathscr{F} &= \{(G,S) : G \in \mathscr{G}_r, G - S \in \Psi, S \in \mathrm{Add}(G - S)\}. \end{aligned} \tag{4.2}$$

Remark 4. We note that $\mathscr{E} \subseteq \mathscr{F}$: Let $(G,R) \in \mathscr{E}$ so that $G - R \in \Psi$, and let P be any longest path of G avoided by R. By (C2), G is connected, so P cannot be contained in any cycle, as this would imply either that G was Hamiltonian, or that P was not a longest path. Thus R is addable for $G - R$ and $(G,R) \in \mathscr{F}$.

Lemma 4.3. *Let $H \in \Psi$. Let $\mathscr{S}(H) = \{S : H + S \in \mathscr{G}_r\}$. Let S be chosen u.a.r. from $\mathscr{S}(H)$. There exists a constant $\delta > 10^{-7}$ such that*

$$\mathbf{Pr}(S \in \mathrm{Add}(H)) \leqslant e^{-\delta rn}.$$

Proof. Given y_0, let $P_{y_h} = y_0 y_1 \cdots y_h$ be a longest path starting at y_0 in H. A *Pósa rotation* $P_{y_h} \to P_{y_{i+1}}$ [19, 6] gives the path $P_{y_{i+1}} = y_0 y_1 \cdots y_i y_h y_{h-1} \cdots y_{i+1}$ formed from P_{y_h} by adding the edge $y_h y_i$ and erasing the edge $y_i y_{i+1}$.

Let $\mathrm{END}(a)$ be any set of endpoint vertices formed by Pósa rotations, with a fixed, of a longest path aPb in H. We prove that $|\mathrm{END}(a)| \geqslant n/210$.

The Pósa condition for the rotation endpoint set U of a longest path P requires that $|N(U)| < 2|U|$, where $N(U)$ is the disjoint neighbour set of U. Let $u = |U|$ and let $v = |U \cup N(U)|$. Thus $u > v/3$. The condition (D2) guarantees that $U \cup N(U)$ induces at least $ru/4 > rv/12$ edges in H. Thus (C1) implies $v > n/70$ and $u > n/210$.

Let the degree sequence of R be $\mathbf{d} = (d_1, \ldots, d_n)$ and that of H be $(r - d_1, \ldots, r - d_n)$. We choose a replacement set of edges S of size $D = |R| = (d_1 + d_2 + \cdots + d_n)/2$ uniformly among all edge sets with degree sequence $\mathbf{d}$ such that $H + S \in \mathscr{G}_r$. If we generate a random configuration F on $\mathbf{d}$, then, conditional on $H + \gamma(F)$ being simple, $\gamma(F) = S$ is a u.a.r. element of $\mathscr{S}(H)$.

The probability that H+$\gamma(F)$ is simple

We generate u.a.r. a configuration F from the set L, size $|L| = 2D$, of configuration points corresponding to the degree sequence $\boldsymbol{d}$, of R. We show that

$$\mathbf{Pr}(H + \gamma(F) \text{ is simple}) \geqslant n^{-2} e^{-4r^2}. \tag{4.3}$$

We generate the first $rn/12$ random pairings using CONSTRUCT and the rest of F using GENERATE (see Remarks 2 and 3). Our reason for this approach is as follows. The ordering $\boldsymbol{\sigma} = (x_1, x_2, \ldots, x_{2D})$ of L in GENERATE is deterministic. At step $i = 1$, the algorithm

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

GENERATE defaults to choice A. We cannot ignore the possibility that H already contains the edge $\{\phi(x_1), \phi(x_2)\}$. Similarly, if, at step $i+1$, GENERATE uses choice B, then, as the edges of H are fixed, we cannot argue that the existing edges of F_i avoid neighbours of $\phi(x_1), \phi(x_2)$ in H until $i \gg r^2$.

Assuming that the u_i are chosen randomly for each of the first $rn/12$ iterations, we claim that the probability that CONSTRUCT inserts a loop or parallel edge is at most

$$\frac{r/2 + r^2/2}{(0.15)rn} \leqslant 4r/n.$$

Indeed, when CONSTRUCT starts there are $2D \in [(0.32)rn, (0.34)rn]$ configuration points to be paired. At the last iteration of CONSTRUCT there are $2D - rn/6 \geqslant (0.15)rn$ points remaining. Each vertex occurs at most $r/2$ times in the sequence (by D2).

CONSTRUCT picks a point u_i and then a random point v_i. Given u_i, there are $\leqslant r/2$ choices which make a loop. In the worst case $d(u_i) = r - 1$ in $H + \gamma(F_{i-1})$ and each neighbour is missing $r/2$ points. This leads to at most $r/2 + r^2/2$ bad choices out of at least $(0.15)rn$ choices for v_i.

Let S_1 be the subgraph of S produced by CONSTRUCT. It follows that

$$\mathbf{Pr}(H + S_1 \text{ is simple}) \geqslant e^{-r^2}.$$

We now continue with GENERATE for the remaining $D - rn/12$ edges to be inserted. The subgraph H remains fixed, and GENERATE is initialized with configuration $F_{rn/12}$ of S_1 on $\{u_1, u_2, \ldots, u_{rn/6}\}$. For steps $i = rn/12 + 1, \ldots, D$ we run GENERATE with the minimum degree ordering $\boldsymbol{\sigma}$ of $L - \{u_1, u_2, \ldots, u_{rn/6}\}$ similar to the ordering described in the proof of Lemma 2.2. Observe that

$$\mathbf{Pr}(H + \gamma(F_i) \text{ is simple} \mid H + \gamma(F_{i-1}) \text{ is simple}) \geqslant \left(1 - \frac{1}{2i-1}\right)\left(1 - \frac{25r}{n}\right).$$

The probability that the algorithm makes a type B choice at step i is $1 - \frac{1}{2i-1}$. Given a type B choice, the probability that a loop or multiple edge is formed is at most $25r/n$, for reasons that we now explain. To create a loop we must choose $\phi(z_t) = \phi(x_{2i+t-2})$ for $t = 1$ or 2, and there are at most $2r$ choices of $\{z_1, z_2\}$ that will lead to this. To create a parallel edge, $\phi(z_t)$ must be a neighbour of $\phi(x_{2i+t-2})$, for $t = 1$ or 2 and there are at most $2r^2$ choices of $\{z_1, z_2\}$ that will lead to this. These choices are made randomly from a set of edges of F_i of size at least $rn/12$.

Now

$$\prod_{i=rn/12+1}^{D}\left(1 - \frac{1}{2i-1}\right) \geqslant n^{-2}.$$

The number of edges inserted by GENERATE is at most $(0.087)rn$ and $\left(1 - \frac{25r}{n}\right)^{(0.087)rn} \geqslant e^{-3r^2}$, and so (4.3) follows.

The probability that $\gamma(F)$ is addable for H

Let x_0 be an end vertex of longest path P in H. Now let $Y = \{(a, b) : a \in \mathrm{END}(x_0), b \in \mathrm{END}(a)\}$. Then $S \in \mathrm{Add}(H)$ implies $\gamma(F) \cap Y = \emptyset$. Otherwise the edge ab would close some longest path of H to a cycle.

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

We will use CONSTRUCT to generate a configuration F with the required degree sequence $(d_1, \dots, d_n)$.

Since $|\mathrm{END}(x_0)| \geqslant n/210$, the sum of the values d_v over vertices $v \in \mathrm{END}(x_0)$ is at least $\frac{r}{4}\frac{n}{210}$. Thus, we can choose u_j so that $\phi(u_j) \in \mathrm{END}(x_0)$ for each of the first $v = rn/1680$ steps. For $j \leqslant v$, writing a for $\phi(u_j)$, let Y_j be the set of remaining configuration points y such that $\phi(y) \in \mathrm{END}(a)$. Then $|Y_j| \geqslant \frac{r}{4}\frac{n}{210} - 2j$. As F contains at most $rn/2$ configuration points,

$$\begin{aligned}\mathbf{Pr}(\gamma(F) \cap Y = \emptyset) &\leqslant \prod_{j=1}^{v}\left(1 - \frac{|Y_j|}{rn/2}\right)\\ &\leqslant \exp\left(-\sum_{j=1}^{v}\left(\frac{1}{420} - \frac{4j}{rn}\right)\right)\\ &= e^{-\delta_1 rn}\end{aligned}$$

where $\delta_1 \approx 1/(1680 \times 840)$.

Thus

$$\mathbf{Pr}(S \in \mathrm{Add}(H)) \leqslant e^{-\delta_1 rn} \times n^2 e^{4r^2},$$

and the lemma follows. □

We can now complete the proof of Theorem 1.1(b). Suppose G is chosen u.a.r. from $\mathscr{G}_r^*$ and then R is chosen by selecting edges independently with probability $1/3$. From Lemma 4.2, we see that

$$\begin{aligned}\mathbf{Pr}(\mathscr{E}) &= \sum_{G \in \mathscr{N}} \sum_{R \in \mathrm{Del}(G)} \mathbf{Pr}((G,R))\\ &\geqslant e^{-n}\mathbf{Pr}(\mathscr{N}).\end{aligned}$$

From the definitions (4.2), inequality (4.1) and Lemma 4.3 it follows that

$$\begin{aligned}\mathbf{Pr}(\mathscr{F}) &\leqslant \mathbf{Pr}(|R| \notin [(0.16)rn, (0.17)rn])\\ &\quad + \sum_{H \in \Psi} \sum_{S \in \mathrm{Add}(H)} \mathbf{Pr}((H+S, S) \mid G - R = H)\mathbf{Pr}(G - R = H)\\ &\leqslant \sum_{H \in \Psi} e^{-\delta rn}\mathbf{Pr}(G - R = H) + e^{-\delta rn}\\ &\leqslant 2e^{-\delta rn}.\end{aligned}$$

Now, by Remark 4, $\mathscr{E} \subseteq \mathscr{F}$ and so $\mathbf{Pr}(\mathscr{E}) \leqslant \mathbf{Pr}(\mathscr{F})$; thus

$$\mathbf{Pr}(\mathscr{N}) \leqslant 2e^{n - \delta rn} = o(1),$$

and the theorem follows from Lemma 4.1. □

Remark 5. We note that by following Frieze [10] we can, at the expense of complicating the proof, prove the existence of a polynomial time algorithm for finding a Hamilton cycle.

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press

Acknowledgement

We wish to thank an anonymous referee for several very careful and thorough reviews, which contributed greatly to the clarity of exposition and accuracy of the paper.

References

[1] Alon, N. and Spencer, J. (1992) *The Probabilistic Method*, Wiley.
[2] Bender, E. A. and Canfield, E. R. (1978) The asymptotic number of labelled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** 296–307.
[3] Bollobás, B. (1980) A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combin.* **1** 311–316.
[4] Bollobás, B. (1981) Random graphs. In *Combinatorics* (H. N. V. Temperley, ed.), Vol. 52 of *LMS Lecture Note Series*, Cambridge University Press, Cambridge, 80–102.
[5] Bollobás, B. (1983) Almost all regular graphs are Hamiltonian. *Europ. J. Combin.* **4** 97–106.
[6] Bollobás, B. (1985) *Random Graphs*, Academic Press.
[7] Feller, W. (1960) *An Introduction to Probability Theory*, Vol. I, Wiley.
[8] Fenner, T. I. and Frieze, A. M. (1984) Hamiltonian cycles in random regular graphs. *J. Combin. Theory Ser. B* **37** 103–112.
[9] Frieze, A. M. (1988) Random regular graphs of non-constant degree. Technical Report, Department of Mathematical Sciences, Carnegie Mellon University.
[10] Frieze, A. M. (1988) Finding Hamilton cycles in sparse random graphs. *J. Combin. Theory Ser. B* **44** 230–250.
[11] Frieze, A. M. and Łuczak, T. (1992) On the independence and chromatic numbers of random regular graphs. *J. Combin. Theory Ser. B* **54** 123–132.
[12] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley.
[13] Krivelevich, M., Sudakov, B., Vu, V. and Wormald, N. C. Random regular graphs of high degree. To appear in *Random Struct. Alg.*
[14] Kučera, L. (1987) Canonical labelling of regular graphs in linear expected time. In *Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science*, 271–279.
[15] Łuczak, T. (1992) Sparse random graphs with a given degree sequence. In *Random Graphs*, Vol. 2 (A. M. Frieze and T. Łuczak, eds), Wiley, New York, 165–182.
[16] McKay, B. D. (1985) Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combinatoria* **19A** 15–25.
[17] McKay, B. D. and Wormald, N. C. (1991) Asymptotic enumeration by degree sequence of graphs with degree $o(n^{1/2})$. *Combinatorica* **11** 369–382.
[18] McKay, B. D. and Wormald, N. C. (1990) Asymptotic enumeration by degree sequence of graphs of high degree. *Europ. J. Combin.* **11** 565–580.
[19] Pósa, L. (1976) Hamilton circuits in random graphs. *Discrete Math.* **14** 359–364.
[20] Robinson, R. W. and Wormald, N. C. (1992) Almost all cubic graphs are Hamiltonian. *Random Struct. Alg.* **3** 117–126.
[21] Robinson, R. W. and Wormald, N. C. (1994) Almost all regular graphs are Hamiltonian. *Random Struct. Alg.* **5** 363–374.
[22] Wormald, N. C. (1999) Models of random regular graphs. In *Surveys in Combinatorics*, Proceedings of the 1999 British Combinatorial Conference (J. D. Lamb and D. A. Preece, eds), Vol. 267 of *LMS Lecture Note Series*, Cambridge University Press, Cambridge, 239–298.

https://doi.org/10.1017/S0963548301005090 Published online by Cambridge University Press