

# Random regular graphs of non-constant degree: connectivity and Hamiltonicity

Colin Cooper\*      Alan Frieze†      Bruce Reed‡

September 9, 2001

## Abstract

Let  $G_r$  denote a graph chosen uniformly at random from the set of  $r$ -regular graphs with vertex set  $\{1, 2, \dots, n\}$  where  $3 \leq r \leq c_0 n$  for some small constant  $c_0$ . We prove that with probability tending to 1 as  $n \rightarrow \infty$ ,  $G_r$  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,  $0 \leq r \leq n - 1$ , 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

---

\*Department of Mathematical and Computing Sciences, Goldsmiths College, University of London, New Cross, London SE14 6NW. Research supported by the STORM Research Centre, UNL. E-mail: [c.cooper@gold.ac.uk](mailto:c.cooper@gold.ac.uk).

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, U.S.A., Supported in part by NSF grant CCR-9818411. E-mail: [alan@random.math.cmu.edu](mailto:alan@random.math.cmu.edu).

‡Equipe Combinatoire, CNRS, Univ. de Paris VI, 4 Place Jussieu, Paris 75005, France. E-mail: [reed@ecp6.jussieu.fr](mailto:reed@ecp6.jussieu.fr).

and Wormald [17, 18] subsequently gave alternative approaches which are useful for  $r = o(n^{1/2})$  or  $r = \Omega(n)$ .

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  $\mathcal{G}_r$  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  $r \rightarrow \infty$  as  $n \rightarrow \infty$ , 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  $r(n) \geq \sqrt{n} \log n$ . Our paper complements [13] both in both in the range of  $r$  studied and in the techniques applied.

**Theorem 1** *Assume  $3 \leq r \leq c_0 n$  for some small positive absolute constant  $c_0$ . Then with probability tending to 1 as  $n \rightarrow \infty$ ,*

- (a)  $G_r$  is  $r$ -connected.
- (b)  $G_r$  is Hamiltonian.

The results of Theorem 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  $r \leq n^{.002}$  by Łuczak [15].

As [13] proves the case where  $r \geq n^{1/2} \log n$ , this implies  $G_r$  is  $r$ -connected and Hamiltonian **whp**<sup>1</sup> for all  $3 \leq r \leq n - 4$ .

## 2 Generating graphs with a fixed degree sequence.

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

Let  $\Omega$  be the set of all  $(2D)!/(D!2^D)$  partitions of  $W = [2D]$  into  $D$  2-element sets. An element of  $\Omega$  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  $i < n$ .

---

<sup>1</sup>A sequence of events  $\mathcal{E}_n$  is said to occur *with high probability (whp)* if  $\lim_{n \rightarrow \infty} \Pr(\mathcal{E}_n) = 1$ .

Let  $\Omega_{\mathbf{d}}$  be  $\Omega$  with the understanding that the underlying set  $W$  is partitioned into  $P_{\mathbf{d}}$ . The *degree sequence* of an element  $F$  of  $\Omega_{\mathbf{d}}$  is  $\mathbf{d}$ . We often write  $\Omega$  for  $\Omega_{\mathbf{d}}$  when the context is clear. Define  $\phi_{P_{\mathbf{d}}} : W \rightarrow [n]$  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  $\mathcal{G}_{\mathbf{d}}$  is the image under  $\gamma$  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  $\mathcal{G}_{\mathbf{d}}$ .

If  $d_i = r$ , ( $1 \leq i \leq n$ ) 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 below. This result allows us to prove many results directly via configurations and then condition the probability estimates for simple graphs.

**Lemma 1** Let  $\Delta = \max_{i \in [n]} d_i$ . Suppose that  $\Delta \leq n/1000$  and that  $\mathbf{d}$  satisfies  $\min_{i \in [n]} d_i \geq \Delta/4$ . Given  $a, b \in [n]$ , if  $G$  is sampled u.a.r. from  $\mathcal{G}_{\mathbf{d}}$ , then

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

**Proof** Let

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

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  $D \geq n\Delta/8$  it follows that

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

□

**Lemma 2** Suppose  $100 \leq r \leq n/1000$ . Let  $d_j = r$ ,  $1 \leq j \leq n$ . If  $F$  is chosen uniformly at random (u.a.r) from  $\Omega$  then for  $n$  sufficiently large,

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

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

**Algorithm** GENERATE

**begin**

$D := rn/2$

$F_0 := \emptyset$

Let  $\sigma = (x_1, x_2, \dots, x_{2D-1}, x_{2D})$  be an ordering of  $W$

**For**  $i = 1$  **to**  $D$  **do**

**begin**

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

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**

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

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

Now consider a bipartite graph  $H$  with vertex bipartition  $(\Omega_{i-1}, \Omega_i)$  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  $\Omega_{i-1}$  (induction) and then uniformly chooses an  $H$ -edge leaving  $F$ . This implies uniformity in  $\Omega_i$ .

Label the configuration points in set  $W_k$  of the partition, as  $\{(k-1)r + j : 1 \leq j \leq r\}$ . For the ordering  $\sigma$  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  $i = (j-1)n + k$ , ( $1 \leq k \leq n$ ,  $1 \leq j \leq r$ ), 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

$$\Pr(F_i \in \Omega_i^* \mid F_{i-1} \in \Omega_{i-1}^*) \geq \begin{cases} 1 & 2i \leq n \\ \left(1 - \frac{60\Delta_i}{(2i-1)n} - \frac{2\Delta_i^2 + 2\Delta_i}{i-1}\right) & n < 2i \leq rn. \end{cases} \quad (1)$$

If  $i \leq n/2$  then  $F_i$  induces a matching. If  $i > n/2$  and if at the  $i$ th 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 the edge  $\{a, b\}$  exists, in the corresponding simple random graph, is at most  $\frac{20\Delta_i}{n}$ , by Lemma 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  $i$ th 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 (1) for  $i \leq rn/2$ .  $\square$

**Remark 2** In Lemma 7 we need to run algorithm GENERATE starting with a configuration  $F_0$  on  $[2D']$  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.

### 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  $P_d$  associated with the configuration. After iteration  $i$ ,  $F_i$  is a u.a.r element of  $\Omega_i$ . We can, if we wish, complete a certain number  $I$  of iterations using CONSTRUCT and then switch to GENERATE. Instead of initializing the ordering  $\sigma$  used in algorithm GENERATE with  $W$  we initialize  $\sigma$  with  $R_I$ , the remaining unmatched points.

### 3 $r$ -Connectivity

We now prove Theorem 1(a). Since the result is already known for  $r$  constant, we can assume that  $10^6 \leq r \leq c_0 n$ , 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** *Let  $\mathcal{Q}_1 \subseteq \mathcal{G}_r$  be the event that for all vertices  $v, w \in V$  of  $G_r$ :*

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

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

**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  $\mathcal{F}(S) = \{G : G = G_r - v, N(v) = S\}$  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  $\mathcal{F}$  into sets  $\mathcal{F}(k) = \{G : |N(w) \cap S| = k\}$  where  $0 \leq k \leq r$  if  $w \notin S$  and  $0 \leq k \leq r - 1$  if  $w \in S$ .

For sets  $R, T \subseteq V - v$  let  $\mathcal{N}(R, T) = \mathcal{N}(R, T; S, w)$  be the set of graphs in  $\mathcal{F}$  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  $\mathcal{B}$  with left vertex set  $\mathcal{N}(R, T)$  and right vertex set  $\mathcal{N}(R + x, T - a)$ .

If  $G \in \mathcal{N}(R, T)$  and  $\{w, a\}, \{x, b\}$  are edges of  $G$  we make a *switching*  $G : (wa, xb) \rightarrow (wx, ab)$  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  $\mathcal{B}$ , and  $d_L(G)$  (resp.  $d_R(G')$ ) is the number of edges incident with  $G$  (resp.  $G'$ ) in  $\mathcal{B}$ .

Let  $\nu(a, x; G) = |N(a) \cap N(x)|$  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  $G : (wa, xb) \rightarrow (wx, ab)$  gives a simple  $G'$  we have

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

for  $G'$  is simple iff  $b \neq a$  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  $\nu(a, x; G') = \nu(a, x; G)$  as  $(\{a\} \cup N(a)) \cap N(x)$  is the same set in both graphs.

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

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

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

The graph  $\mathcal{B}$  consists of components within which  $\delta, \nu$  (and hence  $d_L, d_R$ ) are invariant. Consider a component with bipartition size  $(N_L, N_R)$ . We now prove that  $N_L \geq N_R$ . 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  $G'$ ,

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

and so

$$d_R(G') = |N(a)| - \nu - \delta \geq 2.$$

Thus

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

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

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

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

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

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  $\mathbf{Pr}(Z \geq k) \leq \mathbf{Pr}(X \geq k)$  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  $B(\rho, \rho/N)$ , is at most  $\exp(\rho^2/(N - \rho))$  (see [7] p57). Thus provided  $r = o(\sqrt{n})$ , using the following bound (2) on Binomial tails (see [1]),

$$\mathbf{Pr}(Y \geq \beta\mu) \leq \left(\frac{e}{\beta}\right)^{\beta\mu} \tag{2}$$

we see that

$$\Pr(X \geq \beta\mu) \leq 2 \left( \frac{e}{\beta} \right)^{\beta\mu}.$$

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

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

For  $\log^2 n \leq r \leq n$  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

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

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

**Lemma 4** *Let  $\mathcal{Q}_2$  be the event that no set of vertices  $U \subset V$  of  $G_r$ ,  $1 \leq |U| \leq n/70$ , induces more than  $r|U|/12$  edges. Then  $\Pr(\mathcal{Q}_2) = 1 - O(1/n^2)$ .*

**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  $u \leq 2\beta r$ ,

$$\binom{u}{2} \leq \beta ru.$$

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

In  $\Omega$  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

$$\begin{aligned} \Pr_{\Omega}(Y \geq \beta ru) &= \Pr(Y \geq (\beta(n - u)/u)\mathbf{E}Y) \\ &\leq \left( \frac{ue}{\beta(n - u)} \right)^{\beta ru} \quad \text{by (2)} \\ &\leq \left( \frac{34u}{n} \right)^{\beta ru}. \end{aligned}$$



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

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

provided  $r \leq c_0 n$ ,  $c_0$  sufficiently small.  $\square$

**Proof of Theorem 1(a).** Assume the events  $\mathcal{Q}_1, \mathcal{Q}_2$  described in Lemmas 3,4. If  $G_r$  is not  $r$ -connected then there is a separator  $X$  of size  $x \leq r-1$ . Let  $G_r - X = A + B$  and  $|A| = a \leq |B| = b$ .

**Case 1:**  $2 \leq a \leq r/2$ .

Let  $u, v \in A$  be arbitrary. If  $r = o(n)$  then as  $\mathcal{Q}_1$  occurs,

$$|N(u) \cup N(v)| \geq 2r - |N(u) \cap N(v)| \geq 2r - o(r) - 10 \quad (3)$$

However

$$|N(u) \cup N(v)| \leq |A \cup X| \leq a + r - 1 < 3r/2 \quad (4)$$

which contradicts (3).

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

**Case 2:**  $r/2 \leq a \leq n/80$ .

As  $|A \cup X| \leq a + r - 1$  and  $A \cup X$  contains at least  $ar/2$  edges we see that because  $\mathcal{Q}_2$  occurs

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

**Case 3:**  $n/80 \leq a \leq \lceil n/2 \rceil$ .

If configuration  $F$  is chosen randomly from  $\Omega$  then the existence of a separator of size  $x \leq r-1$ , where the smaller component has size  $a \geq n/80$ , 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 the probability of this event in  $\mathcal{G}_r$  is at most

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

for  $r \leq c_0 n$ ,  $c_0$  sufficiently small.  $\square$

## 4 Hamilton cycles

We prove Theorem 1(b) on the assumption that  $10^7 \leq r \leq c_0 n$ .

**Definition:** Let  $\mathcal{G}_r^*$  denote the subset of  $\mathcal{G}_r$  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 4 and Theorem 1(a) imply that

**Lemma 5**  $|\mathcal{G}_r^*| = (1 - o(1))|\mathcal{G}_r|$ .

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  $G$ , ( $R \in \text{Del}(G)$ ), if

**D1:**  $R$  avoids  $P$ .

**D2:** For all  $v \in V$ ,  $\frac{r}{4} \leq d_R(v) \leq \frac{r}{2}$ .

**Lemma 6** *Let  $G \in \mathcal{G}_r$  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*

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

**Proof**

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

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

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

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

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

□

The size of the set  $R$  of deleted edges is binomial  $B(rn/2, 1/3)$  and thus **whp**  $|R| = (1 + o(1))rn/6$ . For the purposes of Lemma 7 below, we condition on  $|R| \in [(.16)rn, (.17)rn]$ . We note that there exists some  $\delta > 10^{-7}$  such that

$$\Pr(|R| \notin [(.16)rn, (.17)rn]) \leq e^{-\delta rn}. \quad (5)$$

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

**A1:**  $H + S \in \mathcal{G}_r$ .

**A2:** No longest path of  $H$  is closed to a cycle by  $S$ .

Let

$$\begin{aligned} \mathcal{N} &= \{G \in \mathcal{G}_r^* : G \text{ is not Hamiltonian}\} \\ \mathcal{E} &= \{(G, R) : G \in \mathcal{N}, R \in \text{Del}(G)\} \\ \Psi &= \{H : H = G - R, (G, R) \in \mathcal{E}, |R| \in [(.16)rn, (.17)rn]\} \\ \mathcal{F} &= \{(G, S) : G \in \mathcal{G}_r, G - S \in \Psi, S \in \text{Add}(G - S)\}. \end{aligned} \quad (6)$$

**Remark 4** We note that  $\mathcal{E} \subseteq \mathcal{F}$ : Let  $(G, R) \in \mathcal{E}$  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  $(G, R) \in \mathcal{F}$ .

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

$$\Pr(S \in \text{Add}(H)) \leq e^{-\delta rn}.$$

## Proof

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

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

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  $\nu = |U \cup N(U)|$ . Thus  $u > \nu/3$ . The condition (D2) guarantees that  $U \cup N(U)$  induces at least  $ru/4 > r\nu/12$  edges in  $H$ . Thus (C1) implies  $\nu > n/70$  and  $u > n/210$ .

Let the degree sequence of  $R$  be  $\mathbf{d} = (d_1, \dots, d_n)$  and that of  $H$  be  $(r - d_1, \dots, r - d_n)$ . We choose a replacement set of edges  $S$  of size  $D = |R| = (d_1 + d_2 + \dots + d_n)/2$  uniformly among all edge sets with degree sequence  $\mathbf{d}$  such that  $H + S \in \mathcal{G}_r$ . 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  $\mathcal{S}(H)$ .

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  $\mathbf{d}$ , of  $R$ . We show that

$$\Pr(H + \gamma(F) \text{ is simple}) \geq n^{-2} e^{-4r^2}. \quad (7)$$

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

$$\frac{r/2 + r^2/2}{(.15)rn} \leq 4r/n.$$

Indeed, when CONSTRUCT starts there are  $2D \in [(.32)rn, (.34)rn]$  configuration points to be paired. At the last iteration of CONSTRUCT there are  $2D - rn/6 \geq (.15)rn$  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  $\leq r/2$  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  $(.15)rn$  choices for  $v_i$ .

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

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

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  $\{u_1, u_2, \dots, u_{rn/6}\}$ . For steps  $i = rn/12 + 1, \dots, D$  we run GENERATE with the minimum degree ordering  $\sigma$  of  $L - \{u_1, u_2, \dots, u_{rn/6}\}$  similar to the ordering described in the proof of Lemma 2. Observe that

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

The probability that the algorithm makes a Type B choice at step  $i$  is  $1 - \frac{1}{2^{i-1}}$ . 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  $\prod_{i=rn/12+1}^D (1 - \frac{1}{2^{i-1}}) \geq n^{-2}$ . The number of edges inserted by GENERATE is at most  $(.087)rn$  and  $(1 - \frac{25r}{n})^{(.087)rn} \geq e^{-3r^2}$  and so (7) 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  $Y = \{(a, b) : a \in END(x_0), b \in END(a)\}$ . Then  $S \in Add(H)$  implies  $\gamma(F) \cap Y = \emptyset$ . For otherwise the edge  $ab$  would close some longest path of  $H$  to a cycle.

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

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

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

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

Thus

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

and the lemma follows. □

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

Lemma 6, we see that

$$\begin{aligned}\Pr(\mathcal{E}) &= \sum_{G \in \mathcal{N}} \sum_{R \in \text{Del}(G)} \Pr((G, R)) \\ &\geq e^{-n} \Pr(\mathcal{N}).\end{aligned}$$

From the definitions (6), inequality (5) and Lemma 7 it follows that

$$\begin{aligned}\Pr(\mathcal{F}) &\leq \Pr(|R| \notin [(.16)rn, (.17)rn]) \\ &\quad + \sum_{H \in \Psi} \sum_{S \in \text{Add}(H)} \Pr((H + S, S) \mid G - R = H) \Pr(G - R = H) \\ &\leq \sum_{H \in \Psi} e^{-\delta rn} \Pr(G - R = H) + e^{-\delta rn} \\ &\leq 2e^{-\delta rn}.\end{aligned}$$

Now, by Remark 4,  $\mathcal{E} \subseteq \mathcal{F}$  and so  $\Pr(\mathcal{E}) \leq \Pr(\mathcal{F})$ , thus

$$\Pr(\mathcal{N}) \leq 2e^{n-\delta rn} = o(1)$$

and the theorem follows from Lemma 5. □

**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.

**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] N.Alon and J.Spencer, *The Probabilistic Method*, Wiley, 1992.
- [2] E.A.Bender and E.R.Canfield, *The asymptotic number of labelled graphs with given degree sequences*, Journal of Combinatorial Theory, Series A 24 (1978) 296-307.
- [3] B.Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European Journal on Combinatorics 1 (1980) 311-316.
- [4] B.Bollobás, *Random Graphs*, in Combinatorics (ed. H.N.V.Temperley), London Mathematical Society Lecture Note Series 52, Cambridge University Press, Cambridge (1981) 80-102.

- [5] B.Bollobás, *Almost all regular graphs are Hamiltonian*, European Journal on Combinatorics 4, (1983) 97-106.
- [6] B.Bollobás, *Random Graphs*, Academic Press, 1985.
- [7] W. Feller. *An Introduction to Probability Theory*. Volume I, Wiley (1960).
- [8] T.I.Fenner and A.M.Frieze, *Hamiltonian cycles in random regular graphs*, Journal of Combinatorial Theory, Series B 37 (1984) 103-112.
- [9] A.M.Frieze, *Random regular graphs of non-constant degree*, Technical Report, Department of Mathematical Sciences, Carnegie Mellon University (1988).
- [10] A.M.Frieze, *Finding Hamilton cycles in sparse random graphs*, Journal of Combinatorial Theory Series B 44 (1988) 230-250.
- [11] A.M.Frieze and T.Luczak, *On the independence and chromatic numbers of random regular graphs*, Journal of Combinatorial Theory Series B 54 (1992) 123-132.
- [12] S.Janson, T.Luczak and A.Ruciński, *Random Graphs*, John Wiley and Sons, 2000.
- [13] M.Krivelevich, B.Sudakov, V.Vu and N.C.Wormald, *Random regular graphs of high degree*, to appear in Random Structures and Algorithms.
- [14] L.Kučera, *Canonical labelling of regular graphs in linear expected time*, Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science (1987) 271-279.
- [15] T.Luczak, *Sparse random graphs with a given degree sequence*, in Random Graphs Vol. 2, (eds. A.M.Frieze and T.Luczak), Wiley, New York (1992) 165-182.
- [16] B.D.Mckay, *Asymptotics for symmetric 0-1 matrices with prescribed row sums*, Ars Combinatoria 19A (1985) 15-25.
- [17] B.D.Mckay and N.C.Wormald, *Asymptotic enumeration by degree sequence of graphs with degree  $o(n^{1/2})$* , Combinatorica 11 (1991) 369-382.
- [18] B.D.Mckay and N.C.Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, European Journal of Combinatorics 11 (1990) 565-580.
- [19] L.Pósa, *Hamilton circuits in random graphs*, Discrete Mathematics 14 (1976) 359-364.
- [20] R.W.Robinson and N.C.Wormald, *Almost all cubic graphs are Hamiltonian*, Random Structures and Algorithms 3 (1992) 117-126.
- [21] R.W.Robinson and N.C.Wormald, *Almost all regular graphs are Hamiltonian*, Random Structures and Algorithms 5 (1994) 363-374.

- [22] N.C.Wormald, *Models of random regular graphs*, Surveys in Combinatorics, London Mathematical Society Lecture Note Series 267, Cambridge University Press, Cambridge, 1999 (J.D.Lamb and D.A.Preece, Eds.), Proceedings of the 1999 British Combinatorial Conference, Cambridge University Press, 239-298.

---

This research was sponsored in part by National Science Foundation (NSF) grant no. CCR-0122581.