# How many random edges make a dense graph hamiltonian?

Tom Bohman* Alan Frieze† Ryan Martin‡

Department of Mathematical Sciences,
Carnegie Mellon University
Pittsburgh PA 15213.

April 9, 2002

**Abstract**

This paper investigates the number of random edges required to add to an arbitrary dense graph in order to make the resulting graph hamiltonian with high probability. Adding $\Theta(n)$ random edges is both necessary and sufficient to ensure this for all such dense graphs. If, however, the original graph contains no large independent set, then many fewer random edges are required. We prove a similar result for directed graphs.

## 1 Introduction

In the classical model of a random graph (Erdős and Rényi [3]) we add random edges to an empty graph, all at once or one at a time, and then ask for the probability that certain structures occur. At the present time, this model and its variants, have generated a vast number of research papers and at least two excellent books, Bollobás [1] and Janson, Łuczak and Ruciński [5]. It is also of interest to consider random graphs generated in other ways. For example there is a well established theory of considering random subgraphs of special graphs, such as the $n$-cube. In this paper we take a slightly different line. We start with a graph $H$ chosen *arbitrarily* from some class of graphs and then consider adding a random set of edges $R$. We then ask if the random graph $G = H + R$ has a certain property. This for example would model graphs which were basically deterministically produced, but for which there is some uncertainty about the complete structure. In any case, we feel that there is the opportunity here for asking interesting and natural questions.

As an example we consider the following scenario: Let $0 < d < 1$ be a fixed positive constant. We let $\mathcal{G}(n,d)$ denote the set of graphs with vertex set $[n]$

*Supported in part by NSF grant DMS-0100400
†Supported in part by NSF grant CCR-9818411.
‡Supported in part by NSF VIGRE grant DMS-9819950.

which have minimum degree $\delta \geq dn$. We choose $H$ arbitrarily from $\mathcal{G}(n,d)$ and add a random set of $m$ edges $R$ to create the random graph $G$. We prove two theorems about the number of edges needed to have $G$ Hamiltonian **whp**.[1] Since $d \geq 1/2$ implies that $H$ itself is Hamiltonian (Dirac's Theorem), this could be considered to be a probabilistic generalisation of this theorem to the case where $d < 1/2$. We henceforth assume $d < 1/2$.

**Theorem 1.** *Suppose $0 < d \leq 1/2$ is constant, $H \in \mathcal{G}(n,d)$ and let $\theta = \ln d^{-1} \geq .69$. Let $G = H + R$ where $|R| = m$ is chosen randomly from $\overline{E} = [n]^{(2)} \setminus E(H)$.*

**(a)** *If $m \geq (30\theta + 13)n$ then $G$ is Hamiltonian* **whp**.

**(b)** *For $d \leq 1/10$ there exist graphs $H \in \mathcal{G}(n,d)$ such that if $m < \theta n/3$ then* **whp** *$G$ is not Hamiltonian.*

So it seems that we have to add $\Theta(n)$ random edges in order to make $G$ Hamiltonian **whp**. Since a random member of $\mathcal{G}(n,d)$ is already likely to be Hamiltonian, this is a little disappointing. Why should we need so many edges in the worst-case? It turns out that this is due to the existence of a large independent set. Let $\alpha = \alpha(H)$ be the independence number of $H$.

**Theorem 2.** *Suppose $H \in \mathcal{G}(n,d)$ and $1 \leq \alpha < d^2 n/2$ and so $d > n^{-1/2}$ ($d$ need not be constant in this theorem). Let $G = H + R$ where $|R| = m$ is chosen randomly from $\overline{E}$. If*

$$\frac{md^3}{\ln d^{-1}} \to \infty \tag{1}$$

*then $G$ is Hamiltonian* **whp**.

Note that if $d$ is constant then Theorem 2 implies that $m \to \infty$ is sufficient.

We have also considered a similar problem in relation to adding random arcs to a dense digraph. For a digraph $D$ we denote its arc-set by $A(D)$. We denote its minimum out-degree by $\delta^+$ and its minimum in-degree by $\delta^-$ and then we let $\delta = \min\{\delta^+, \delta^-\}$. Let $0 < d < 1$ be a fixed positive constant. We let $\mathcal{D}(d)$ denote the set of digraphs with vertex set $[n]$ which have $\delta \geq dn$.

**Theorem 3.** *Suppose $0 < d < 1/2$ is constant and $H \in \mathcal{D}(d)$. Let $D = H + R$ where $|R| = m$ is chosen randomly from $\overline{E} = n^{(2)} \setminus E(H)$.*

**(a)** *If $m \geq (d^{-1}(15 + 6\theta) + 5d^{-2})n$ then $D$ is Hamiltonian* **whp**.

**(b)** *For $d \leq 1/10$ there exist digraphs $H \in \mathcal{D}(d)$ such that if $m < \theta n/3$ then* **whp** *$D$ is not Hamiltonian.*

If $\delta \geq n/2$ then $H$ itself is Hamiltonian, Ghouila-Houri [4].

Theorem 1 is proven in the next section, Theorem 2 is proven in Section 3 and Theorem 3 is proved in Section 4.

[1] A sequence of events $\mathcal{E}_n$ is said to occur "with high probability" (**whp**) if $\lim_{n\to\infty} \mathbf{Pr}(\mathcal{E}_n) = 1$

## 2 The worst-case for graphs

We will assume from now on that $m$ is exactly $\lceil 30\theta n\rceil + 13n$. We let $R = R_1 \cup R_2$ where $|R_1| = \lceil 30\theta n\rceil$. Then let $G_1 = H + R_1$.

We first show that

**Lemma 1.** *$G_1$ is connected* **whp**.

**Proof** Let $N = \binom{n}{2}$. If $u, v \in [n]$ then either they are at distance one or two in $H$ or

$$\mathbf{Pr}(dist_{G_1}(u,v) > 3) \le \left(1 - \frac{|R_1|}{N}\right)^{d^2n^2} \le e^{-60\theta d^2 n}.$$

Hence,

$$\mathbf{Pr}(diam(G) > 3) \le n^2 e^{-60\theta d^2 n} = o(1).$$

□

Given a longest path $P$ in a graph $\Gamma$ with end-vertices $x_0, y$ and an edge $\{y, v\}$ where $v$ is an internal vertex of $P$, we obtain a new longest path $P' = x_0..vy..w$ where $w$ is the neighbour of $v$ on $P$ between $v$ and $y$. We say that $P'$ is obtained from $P$ by a *rotation* with $x_0$ *fixed*.

Let $END_\Gamma(x_0, P)$ be the set of end-vertices of longest paths of $\Gamma$ which can be obtained from $P$ by a sequence of rotations keeping $x_0$ as a fixed end-vertex. Let $END_\Gamma(P) = \{x_0\} \cup END_\Gamma(x_0, P)$. Note that if $\Gamma$ is connected and non-Hamiltonian then there is no edge $\{x_0, y\}$ where $y \in END_\Gamma(x_0, P)$.

It follows from Posá [6] that

$$|N_\Gamma(END_\Gamma(P))| < 2|END_\Gamma(P)|, \tag{2}$$

where for a graph $\Gamma$ and a set $S \subseteq V(\Gamma)$

$$N_\Gamma(S) = \{w \notin S : \exists v \in S \text{ such that } vw \in E(\Gamma)\}.$$

**Lemma 2. Whp**

$$|N_{G_1}(S)| \ge 3|S| \tag{3}$$

*for all $S \subseteq [n]$, $|S| \le n/5$.*

**Proof** Now $|N_H(S)| \ge 3|S|$ for all $S \subseteq [n]$, $|S| \le dn/3$. So,

$$\begin{aligned}
\mathbf{Pr}(\exists |S| \le n/5 :\ |N_{G_1}(S)| < 3|S|) &\le \sum_{k=dn/3}^{n/5} \binom{n}{k}\binom{n}{3k}\left(1 - \frac{m}{N}\right)^{k(n-4k)} \\
&\le \sum_{k=dn/3}^{n/5} \left(\frac{n^4 e^4}{27k^4} e^{-12\theta}\right)^k \\
&= o(1).
\end{aligned}$$

□

It follows from Lemma 2 that for any longest path $P$ in a graph $\Gamma$ that contains $G_1$ as a subgraph we have $n/5 \le |END_\Gamma(P)| \le |P|$.

Now let $R_2$ be obtained from a random sequence $e_1, e_2, \dots$ of edges chosen from $\overline{E}$ *with replacement.*

Let $P_0$ be a longest path in $G_1$ of length $\lambda_0 \ge dn$. Now consider the following process: At a general stage we will have a path $P_i$ of length at least $\lambda_0 + i$. We will have used a set $S_i \subseteq R_2$ of size $Y_i$ to go from $P_{i-1}$ to $P_i$, for $i \ge 1$. Here $S_1 = \{e_1, e_2, \dots, e_{Y_1}\}$, $S_2 = \{e_{Y_1}, e_{Y_1+1}, \dots, e_{Y_1+Y_2}\}$ and so on. Let $Z_i = Y_1 + Y_2 + \cdots + Y_i$ and let $\Gamma_i = G_1 + \{e_1, e_2, \dots, e_{Z_i}\}$.

In order to see how the $Y_i$ are determined, let $P_i$ be a longest path in $\Gamma_i$ and let $END_{\Gamma_i}(P_i)$ be as defined above and note that by Lemma 2, we can assume that $|END_{\Gamma_i}(P_i)| \ge n/5$. We now add edges $e_{Z_i+1}, e_{Z_i+2}, \dots$ in turn until we find an edge $e_{Z_i+k} = \{a, b\}$ where $a \in END_{\Gamma_i}(P_i)$ and $b \in END_{\Gamma_i}(a, P_i)$. Since $\Gamma_i$ is connected the addition of $\{a, b\}$ will increase the length of the longest path or close a Hamilton cycle. We let $Y_{i+1} = k$ in this case. Finally, once we have formed a Hamilton cycle, at stage $r$ say, we let $Y_{r+1} = \cdots = Y_n = 0$.

Now the random variables $Y_1, Y_2, \dots, Y_n$ are independent random variables which are either geometric with probability of success at least $\frac{2}{25}$ or are zero valued. Thus

$$\mathbf{E}(Z_n) \le \frac{25n}{2}.$$

Since the variance of $Z_n$ is $O(n)$ it is easy to show by an application of Chebychev's inequality that $Z_n \le 13n$ **whp** and this completes the proof of (a).

**Remark 1.** *The calculations above go through quite happily for $\delta(H) \ge n^{3/4}$, say. For this degree bound the number of additional edges required in the worst-case is $\Omega(n \ln n)$. Since only $\frac{1}{2} n \ln n$ random edges are required for Hamiltonicity when we start with the empty graph, there is no point in considering smaller values of $d$, unless we can improve the constant factor.*

(b) Let $m = cn$ for some constant $c$ and let $H$ be the complete bipartite graph $K_{A,B}$ where $|A| = dn$ and $|B| = (1-d)n$. Let $I$ be the set of vertices of $B$ which are not incident with an edge in $R$. If $|I| > |A|$ then $G$ is not Hamiltonian. Instead of choosing $m$ random edges for $R$, we choose each possible edge independently with probability $p = \frac{2m}{(d^2+(1-d)^2)n^2}$. (We can use monotonicity, see for example Bollobás [1] II.1 to justify this simplification). Then

$$\mathbf{E}(|I|) = (1-d)n(1-p)^{(1-d)n-1} \sim (1-d)\exp\left\{-\frac{2(1-d)m}{(d^2+(1-d)^2)n}\right\} n.$$

It follows from the Chebychev inequality that $|I|$ is concentrated around its mean

and so $G$ will be non-Hamiltonian **whp** if $c$ satisfies

$$c < \frac{1}{2(1-d)}(d^2 + (1-d)^2)\ln(d^{-1}-1).$$

This verifies (b). □

## 3 Graphs with small independence number

**Proof of Theorem 2**
We will first show that we can decompose $H$ into a few large cycles.

**Lemma 3.** *Suppose that $G$ has minimum degree $dn$ where $d \le 1/2$ and that $\alpha(G) < d^2n/2$. Let $k_0 = \lfloor \frac{2}{d} \rfloor$. Then the vertices of $G$ can be partitioned into $\le k_0$ vertex disjoint cycles.*

**Proof** Let $C_1$ be the largest cycle in $H$. $|C_1| \ge dn+1$ and we now show that the graph $H \setminus C_1$ has minimum degree $\ge dn - \alpha$.

To see this, let $C_1 = v_1, \ldots, v_c, v_{c+1} = v_1$. Let $w \in V(H \setminus C_1)$. Because $C_1$ is maximum sized, no such $w$ is adjacent to both $v_i$ and $v_{i+1}$. Also, if $w \sim v_i$ and $w \sim v_j$ with $i < j$ and $v_{i-1} \sim v_{j-1}$, then

$$w, v_j, \ldots, v_c, v_1, \ldots, v_{i-1}, v_{j-1}, \ldots, v_i, w$$

is a larger cycle. So the predecessors of $N(w)$ in $C_1$ must form an independent set and $|N(w) \cap C_1| \le \alpha$. Similar arguments are to be found in [2].

We can repeat the above argument to create disjoint cycles $C_1, \ldots, C_r$ where $|C_1| \ge |C_2| \ge \cdots \ge |C_r|$ and $C_j$ is a maximum sized cycle in the graph $H_{j-1} = H \setminus (C_1 \cup \cdots \cup C_{j-1})$ for $j = 1, 2, \ldots, r$. Now $H_k$ has minimum degree at least $dn - k\alpha$ and at most $n - dn - 1 - (dn - \alpha + 1) - \cdots - (dn - (k-1)\alpha + 1) = n - k(dn + 1 - (k-1)\alpha/2)$ vertices. Since $d^2n > 2\alpha$, $H_{k_0}$, if it existed, would have minimum degree at least 2 and a negative number of vertices. □

Let $C_1, \ldots, C_r$ be the cycles given by Lemma 3.

In order to simplify the analysis, we assume the edges of $R$ are chosen from $\overline{E}$ by including each $e \in \overline{E}$ independently with probability $\frac{m}{|\overline{E}|}$. Because Hamiltonicity is a monotone property, showing that $G$ is Hamiltonian **whp** in this model implies the theorem. We get a further simplification in the analysis if we choose these random edges in rounds: set $R = R_1 \cup R_2 \cup \cdots \cup R_r$ where each edge set $R_i$ is independently chosen by including $e \in \overline{E}$ with probability $p$, where $1 - (1-p)^r = \frac{m}{|\overline{E}|}$. Each $R_i$ will be used to either extend a path or close a cycle and will only be used for one such attempt. In this way each such attempt is independent of the previous. To this end let $G_t = H \cup \bigcup_{s=1}^{t} R_t$ for $t = 0, 1, \ldots, r$. Thus $G_0 = H$ and $G_r = G$.

Let $e = \{x, y\}$ be an edge of $C_r$ and let $Q$ be the path $C_r - e$. In each phase of our procedure, we have a current path $P$ with endpoints $x, y$ together with a collection of vertex disjoint cycles $A_1, A_2, \ldots, A_s$ which cover $V$. Initially $P = Q$, $s = r - 1$ and $A_i = C_i$, $i = 1, 2, \ldots, r - 1$.

Suppose a path $P$ and collection of edge disjoint cycles have been constructed in $G_{t-1}$ (initially $t = 1$). Consider the set $Z = END_{G_{t-1}}(x, P)$ created from rotations with $x$ as a fixed endpoint, as in Section 2. We identify the following possibilities:

**Case 1**: There exists $z_1 \in Z$, $z_2 \notin P$ such that $f = (z_1, z_2)$ is an edge of $H$.
Let $Q$ be the corresponding path with endpoints $x, z_1$ which goes through $V(P)$. Now suppose that $z_2 \in C_i$ and let $f' = (z_2, z_3)$ be an edge of $C_i$ incident with $z_2$. Now replace $P$ by the path $Q, f, Q'$ where $Q' = C_i - f$. This construction reduces the number of cycles by one.

**Case 2**: $|V(P)| \le n/2$ and $z \in Z$ implies that $N_{G_{t-1}}(z) \subseteq V(P)$.
It follows from (2) that $|Z| \ge dn/3$. Now add the next set $R_t$ of random edges. Since $|V(P)| \le n/2$, the probability that no edge in $R_t$ joins $z_1 \in Z$ to $z_2 \in V \setminus V(P)$ is at most $(1-p)^{(dn/3)(n/2)}$. If there is no such edge, we fail, otherwise we can use $(z_1, z_2)$ to proceed as in Case 1. We also replace $t$ by $t + 1$.

**Case 3**: $|V(P)| > n/2$ and $z \in Z$ implies that $N_{G_{t-1}}(z) \subseteq V(P)$.
Now we close $P$ to a cycle. For each $z \in Z$ let $A_z = END_{G_{t-1}}(z, Q_z)$ where $Q_z$ is as defined in Case 1. Each $A_z$ is of size at least $dn/3$. Add in the next set $R_t$ of random edges. The probability that $R_t$ contains no edge of the form $(z, z')$ where $z \in Z$ and $z' \in A_z$ is at most $(1-p)^{d^2n^2/10}$. If there is no such edge, we fail. Otherwise, we have constructed a cycle $C$ through the set $V(P)$ in the graph $G_t$. If $C$ is Hamiltonian we stop. Otherwise, we choose a remaining cycle $C'$, distinct from $C$ and replace $P$ by $C' - e$ where $e$ is any edge of $C'$. Now $|V(P)| < n/2$ and we can proceed to Case 1 or Case 2.

After at most $r$ executions of each of the above three cases, we either fail or produce a Hamilton cycle. The probability of failure is bounded by

$$\begin{aligned} k_0((1-p)^{(dn/3)(n/2)} + (1-p)^{d^2n^2/10}) &\le 2d^{-1}\left(\left(1 - \frac{m}{|\overline{E}|}\right)^{\frac{dn^2}{6r}} + \left(1 - \frac{m}{|\overline{E}|}\right)^{\frac{d^2n^2}{10r}}\right) \\ &\le 4d^{-1}e^{-md^3/10} \\ &= o(1) \end{aligned}$$

provided (1) holds. □

An observation: We do not actually need the condition that $\alpha(H) \le d^2n/2$ to complete this proof. The weaker condition that $d^2n/2$ bounds the independence number of the neighborhood of each vertex is enough.

## 4 Directed graphs

For a digraph $D = ([n], A)$ we let $B_D$ be the bipartite graph $([1, n], [n + 1, 2n], E)$ such that $B_D$ contains an edge $(x, y)$ for every arc $(x, y-n) \in A$. Perfect matchings of $B_D$ correspond to *cycle covers* of $D$ i.e. sets of vertex disjoint directed cycles which contain all vertices of $D$.

We divide our arcs $R$ into two subsets: $R = R_1 \cup R_2$, where each $R_i$ is independently randomly chosen from $[n]^2 \setminus A(H)$. Here

$$|R_i| = \rho_i n \text{ where } \rho_1 = d^{-1}(15 + 6\theta) \text{ and } \rho_2 = 5d^{-2}.$$

**Lemma 4. Whp** *$H_1 = H + R_1$ has a cycle cover $\Sigma_1$.*

**Proof** We apply Hall's theorem to show that $B_{H_1}$ has a perfect matching **whp**. If $B_{H_1}$ does not have a perfect matching then there exists a *witness* $K \subseteq [1, n], |K| \le n/2($ or $L \subseteq [n + 1, 2n], |L| \le n/2)$ such that its neighbour set $N(K)$ in $B_{H_1}$ satisfies $|N(K)| \le |K| - 1$ (resp. $|N(L)| \le |L| - 1$). Clearly any such witness must be of size at least $dn$.

Since having a perfect matching is a monotone increasing property, we can assume that the arcs of $R_1$ are chosen independently with $p_1 = \frac{\rho_1}{n}$.

The probability that $B_{H_1}$ does not contain a perfect matching is therefore bounded by

$$2\sum_{k=dn}^{n/2} \binom{n}{k}\binom{n}{k-1}(1-p_1)^{k(n-k)} \le 2\sum_{k=dn}^{n/2}\left(\frac{n^2e^2}{k^2}\cdot e^{-\rho_1/2}\right)^k = o(1).$$

□

We also need to know that there are many arcs joining large sets. For disjoint $S, T \subseteq [n]$ let $N^+(S) = \{t \notin S : \exists s \in S \text{ such that } (s, t) \text{ is an arc of } H_1\}$. Define $N^-(S)$ similarly.

**Lemma 5. Whp**, *for all disjoint $S, T \subseteq [n]$ with $|S|, |T| \ge dn/2$, $|N^-(S) \cap T|, |N^+(S) \cap T| \ge |T|/2$.*

**Proof** Let $\mathcal{E}$ denote the event $\{\exists$ disjoint $S, T \subseteq [n] : |S|, |T| \ge dn/2$ and $|N^+(S) \cap T| < |T|/2\}$. Now fix $S, T$ with $|S| = s, |T| = t$. If $|N^+(S) \cap T| < |T|/2$ then there exists $T' \subseteq T$ of size $|T|/2$ such that there are no arcs from $S$ to $T'$ in $H_1$. The probability of this is at most

$$2^t(1-p_1)^{ts/2} \le \left(2e^{-\rho_1 s/(2n)}\right)^t.$$

Thus

$$\begin{aligned}
\mathbf{Pr}(\mathcal{E}) &\leq \sum_{s=dn/2}^{(1-d)n}\sum_{t=dn/2}^{n-s}\binom{n}{s}\binom{n}{t}\left(2e^{-\rho_1 s/(2n)}\right)^t\\
&\leq \sum_{s=dn/2}^{(1-d)n}\sum_{t=dn/2}^{n-s}\left(\frac{ne}{s}\right)^s\left(\frac{2ne^{1-\rho_1 s/(2n)}}{t}\right)^t\\
&\leq \sum_{s=dn/2}^{(1-d)n}\sum_{t=dn/2}^{n-s}\left(\frac{ne}{s}\right)^s e^{-\rho_1 st/(3n)}\\
&\leq n^2 e^{-\rho_1 dn^2/24}\\
&= o(1).
\end{aligned}$$

The proof for $N^-(S)\cap T$ is identical. □

**Corollary 4.** *$H_1$ is strongly connected* **whp**.

**Proof** If $H_1$ is not strongly connected then there exists $S\subseteq[n], |S|\leq n/2$ such that either $N^+(S)=\emptyset$ or $N^-(S)=\emptyset$. But this would contradict Lemma 5 with $T=\overline{S}$. □

Assume that $H_1$ is strongly connected and satifies the condition of Lemma 5.

We now describe a procedure for converting the cycle cover $\Sigma_1$ to a Hamilton cycle. We start with an arbitrary cycle $C$ for which there is an arc $(y,z)$ with $y\in C, z\in C'\neq C$. Such an arc must exist because $H_1$ is strongly connected. Let $(y,x)$ be the arc of $C$ leaving $y$ and $(y',z)$ be the arc of $C'$ entering $z$. Now delete arcs $(y,x),(y',z)$ from $\Sigma_1$ and add the arc $(y,z)$. This yields a path $P$ from $x$ to $y'$ plus a collection of disjoint cycles which covers $[n]$. Call this a *near cycle cover*.

Given an NCC we first try to perform an *out path extension*. We can do this if there is an arc $e$ joining the terminal endpoint of the path $P$ to a vertex $v$ of one of the cycles, $C'$ say. By adding the arc $e$ and deleting the arc of $C'$ entering $v$ we create an NCC with one fewer cycle. Note that this construction is the same as that of the previous paragraph, except that we do not invoke strong connectivity. We also try to perform an analogous *in path extension* by checking if there is an arc $(w,s)$ where $s$ is the start vertex of $P$ and $w\notin P$. If such an arc exists, we may extend the path $P$ by adding a path section at its beginning.

We continue with these path extensions until the NCC $\Sigma_2$ that we have no longer admits one. Let $Q=(u_0,u_1,\ldots,u_k)$ be the path of $\Sigma_2$ and define the successor function $\sigma$ by $\sigma(u_i)=u_{i+1}$ for $i<k$. Now $k\geq dn$ since there are no out path extensions available. All of $u_k$'s out neighbours are in $Q$. Let

$$S=\{u_{i-1}:\ i\leq k-dn/2, (u_k,u_i)\text{ is an arc of }H_1\},\ T=\{u_{k-dn/2},\ldots,u_k\}.$$

Clearly $|S|\geq dn/2$. Let $T'=N^+(S)\cap T$ so that $|T'|\geq dn/2$ by Lemma 5. For $v\in T'$ choose $\lambda(v)\in S$ such that $(\lambda(v),v),(u_k,\sigma\lambda(v))$ are both arcs of $H_1$. For

each such $v \in T'$ consider the path

$$Q_v = Q + (u_k, \sigma\lambda(v)) - (\lambda(v), \sigma\lambda(v)) + (\lambda(v), v) - (\sigma^{-1}(v), v).$$

Note that $Q_v$ has the same vertex set as $Q$ and has endpoints $u_0, \sigma^{-1}(v)$ — see Figure 1.

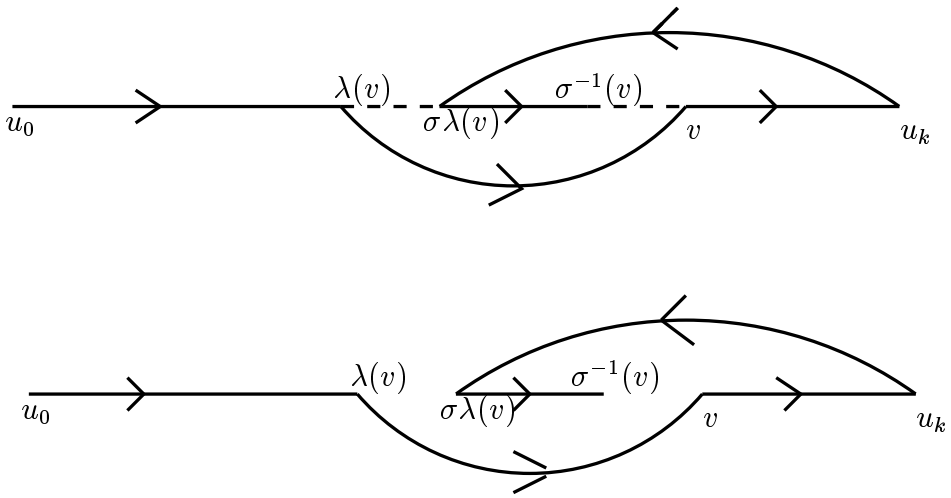


Figure 1

For each $v \in T'$ we see if there is an out path extension available for $Q_v$. Suppose no such out extension exists. By an analogous procedure to the creation of $Q_v$, we can, for each $v \in T'$, construct a set $\mathcal{Q}_v$ of $\geq dn/2$ paths each with a distinct start vertex and all with the same end vertex $\sigma^{-1}(v)$, and all covering the vertices of $Q$, the start vertices are distinct within $\mathcal{Q}_v$ that is. (There are no in path extensions available into $u_0$ and we just look at the arcs that enter $u_0$).

If there is an in path extension available for a $v \in T, Q' \in \mathcal{Q}_v$ then we carry it out.

Now suppose that we fail in all of these attempts at path extension. We generate a sequence of random arcs $e_1, e_2, \ldots$ , part of $R_2$. Each $e_i$ is chosen uniformly from the arcs not in $H_1$, with replacement. We continue until we find an arc which closes a path in some $\mathcal{Q}_v$, $v \in T'$ to a cycle $C^*$ say. Note that each $e_i$ has probability at least $d^2/4$ of achieving this.

Now note that the sequence, starting with a cycle cover, replacing two cycles by a path, doing path extensions, using random arcs to close a path to $C^*$, produces a new cycle cover with one less cycle. Thus eventually a Hamilton cycle is produced.

The number of random edges required can be bounded by the sum $Z = Z_1 + Z_2 + \cdots + Z_n$ where the $Z_i$ are independent geometric random variables with probability of success $d^2/4$. Thus $\mathbf{E}(Z) = 4d^{-2}n$ and **whp** $Z < 5d^{-2}n$. We could

use the Chebycheff inequality for example to prove the latter claim. Thus if we add $5d^{-2}n$ random edges to $H_1$ then we will create a Hamilton cycle **whp**. This completes the proof of part (a) of Theorem 3.

For part (b) we can start with $K_{A,B}$ of Theorem 1(a) and then replace each edge by an arc in both directions to create $H = \vec{K}_{A,B}$. Once again we let $I$ be the set of vertices of $B$ which are not incident with an arc in $R$. If $|I| > |A|$ then $D$ is not Hamiltonian. Instead of choosing $m = cn$ random arcs for $R$, we choose each possible arc independently with probability $p = \frac{m}{(d^2+(1-d)^2)n^2}$. Then

$$\mathbf{E}(|I|) = (1-d)n(1-p)^{2(1-d)n-2} \sim (1-d)\exp\left\{-\frac{2(1-d)m}{(d^2+(1-d)^2)n}\right\}n.$$

It follows from the Chebychev inequality that $|I|$ is concentrated around its mean and so $G$ will be non-Hamiltonian **whp** if $c$ satisfies

$$c < \frac{1}{2(1-d)}(d^2+(1-d)^2)\ln(d^{-1}-1).$$

This verifies (b).

**Acknowledgement** We thank the referees for a careful reading which has revealed several small errors and led to the simplified proof of Theorem 1(a).

## References

[1] B. Bollobás, *Random Graphs*, Academic Press, 1985.

[2] V. Chvátal and P. Erdős, A note on Hamiltonian circuits, *Discrete Math* 2 (1972) 111-113.

[3] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 5 (1960) 17-61. *Discrete Mathematics* 45 (1983) 301-305.

[4] A. Ghouila-Houri, Une condition suffisante d'existance d'un circuit hamiltonien, *Comptes Rendus Academie Scientifique Paris* 25 (1960) 495-497.

[5] S. Janson, T. Łuczak, and A. Ruciński, Random Graphs, Wiley - Interscience Series, New York, 2000.

[6] L. Pósa, Hamiltonian circuits in random graphs, *Discrete Mathematics* 14 (1976) 359-364.