1988

On the number of Hamilton cycles in a random graph

C Cooper
Carnegie Mellon University

Frieze

Follow this and additional works at: http://repository.cmu.edu/math
NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:
The copyright law of the United States (title 17, U.S. Code) governs the making
of photocopies or other reproductions of copyrighted material. Any copying of this
document without permission of its author may be prohibited by law.
ON THE NUMBER OF HAMILTON CYCLES
IN A RANDOM GRAPH

by

C. Cooper
Department of Computing, Mathematics and Statistics
Polytechnic of North London
Holloway Road
London N78DB
England

and

A. M. Frieze
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

Research Report No. 88-27

August 1988
On the number of hamilton cycles in a random graph

by

C. Cooper

Department of Computing, Mathematics and Statistics
Polytechnic of North London
Holloway Road
London N78DB
England

and

A.M. Frieze

Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213
USA

August 1988
Abstract

Let a random graph $G$ be constructed by adding random edges one by one, starting with $n$ isolated vertices. We show that with probability going to one as $n$ goes to infinity, when $G$ first has minimum degree two, it has at least $(\log n)^{(1-\epsilon)n}$ distinct hamilton cycles for any fixed $\epsilon > 0$. 
§1. Introduction

Let \( V_n = \{1, 2, \ldots, n\} \) and consider the random graph process (Bollobás [3]) \( G_0, G_1, \ldots, G_v \) where \( G_m = (V_n, E_m) \), \( E_0 = \emptyset \) and \( E_{m+1} \) is obtained from \( E_m \) by adding an edge \( e_{m+1} \) chosen randomly from \( [n]^{(2)} - E_m \). Now let

\[
m^* = \min\{m : \delta(G_m) \geq 2\}.
\]

Bollobás [2] (see also Ajtai, Komlos and Szemerédi [1]) showed that

\[
\lim_{n \to \infty} \Pr(G_{m^*} \text{ is hamiltonian}) = 1
\]

which was claimed but not proved by Komlós and Szemerédi [7] when they first established the exact threshold for the existence of hamilton cycles in a random graph.

Knowing that \( G_{m^*} \) usually has at least one hamilton cycle raises the question of how many distinct hamilton cycles does it usually contain. We prove

Theorem

If \( \varepsilon > 0 \) is fixed then

\[
\lim_{n \to \infty} \Pr(G_{m^*} \text{ has at least } (\log n)^{1-\varepsilon}n \text{ distinct hamilton cycles}) = 1.
\]

\( \Box \)

Thus at \( m^* \) the number of hamilton cycles jumps dramatically from 0 to at least \( (\log n)^{n-o(n)} \). On the other hand the expected number of hamilton cycles at this point is

\[
n!p^n = (\log n)^n e^{-n+o(n)}
\]

and so the theorem gives the right order of magnitude for the number of hamilton cycles in \( G_{m^*} \).
§2. Notation and preliminaries

We say that almost every (a.e.) graph process satisfies a certain property if this property holds with probability tending to 1 as \( n \) tends to \( \infty \). Let \( m_1 = \frac{1}{2} n(\log n + \log \log n - \log \log \log n) \) and \( m_2 = \frac{1}{2} n(\log n + \log \log n - \log \log \log n) \). It follows from Erdős and Rényi [4] that \( m_1 \leq m^* \leq m_2 \) in a.e. graph process.

In what follows our inequalities need only be true for large enough \( n \). It is always useful to bear in mind the relationship between \( G_m \) and \( G_p \), \( p = \frac{m}{\sqrt{\frac{m}{n}}} \), \( v = \binom{n}{2} \), the random graph in which each possible edge appears independently with probability \( p \). Let \( E_p \) denote the edge set of \( G_p \).

The properties we need are (see [2]): suppose \( \mathcal{A} \) is some property of graphs then

\[
\begin{align*}
(2.1a) \quad & \Pr(G_m \in \mathcal{A}) \leq 3\sqrt{n \log n} \Pr(G_p \in \mathcal{A}) \\
& m_1 \leq m \leq m_2 \\
(2.1b) \quad & \text{a.e. } G_p \in \mathcal{A} \text{ and } \mathcal{A} \text{ is monotone implies a.e. } G_m \in \mathcal{A}. \\
(2.1c) \quad & \text{a.e. } G_p \in \mathcal{A} \text{ implies there exists } m', m - \sqrt{n \log n} \leq m' \leq m \\
& \text{such that a.e. } G_{m'} \in \mathcal{A}. 
\end{align*}
\]

Now let \( \varepsilon > 0 \) be fixed and small from now on and \( V^+_n = V_n - V_{n-\varepsilon} \) where \( n_\varepsilon = \left\lceil (1-\varepsilon)n/2 \right\rceil \).
\[ L_m = \{ v \in V_n : d_m(v) \leq \log n/10 \} \]

where \( d_m(v) \) is the degree of \( v \) in \( G_m \) and

\[ L_m^+ = \{ v \in V_n : d_m^+(v) \leq \log n/10 \} \]

where \( d_m^+(v) \) is the number of neighbours of \( v \) in \( V_n^+ \).

For \( S \subseteq V_n \) let

\[ N_m(S) = \{ w \in V_n - S : \exists v \in S \text{ such that } vw \in E_m \} \]

and let \( N_p(S) \) be defined similarly.

For \( S, T \subseteq V_n, S \cap T = \emptyset \), \( e_m(S, T) = |\{vw \in E_m : v \in S, w \in T\}| \).

Let \( N_L = L_m \cup L_m^+ \cup (N_m (L_m \cup L_m^+) \cap V_n) \).

We now describe the basic properties of \( G_m, m_1 \leq m \leq m_2 \) which are needed for the paper.

**Lemma 2.1**

Almost every graph process is such that simultaneously for all \( m_1 \leq m \leq m_2 \), \( G_m \) satisfies

\[(2.2a) \quad \Delta(G_m) \leq 3 \log n. \quad \text{(maximum degree)}\]

\[(2.2b) \quad |L_m| \leq n^{2/5}, \quad |L_m^+| \leq n^{4/5}. \]

\[(2.2c) \quad \text{No pair of vertices } v, w \in L_m \text{ are within distance 4 of each other.} \]
(2.2d) No pair of vertices \( v, w \in V_n \) have 3 or more common neighbours

(2.2e) \( T \subseteq V, |T| < n \) implies that \( T \) contains at most \( 3|T| \) edges.

(2.2f) \( S \subseteq V_n, |S| < n \) implies \( (S) > |S| \).

(2.2g) \( S \subseteq V_n, |S| < n \) implies \( |\text{INJS} \cap V| > |S| \).

(2.2h) \( S, T \subseteq V_n, S \cap T = \emptyset, |S| = |T| = \left\lfloor \frac{n}{2} \right\rfloor \) implies \( e(S, T) \geq \frac{n \log n}{2(\log \log n)^6} \).

(2.2i) \( V_n \) contains at least \( \frac{4}{3} n \log n \) edges.

Proof (Outline: details of similar results can be found in [2])

Let \( p_j = n_{ij}/N, p_2 = m_2/N \).

Proof of (2.2a)

\[
\Pr(A(G) > 3 \log n) \leq n \sum_{k=3}^{\infty} (V^\cup U)^{11^n n^{11^n}} = o(1).
\]

Hence (2.1b) implies \( \Pr(A(G) > 3 \log n) = o(1) \) and then the result follows from \( A(G_{m1}) \leq A(G_{m2}) \).
Proof of (2.2b)

\[ E(|L_{p_1}|) = n \sum_{k \leq \frac{\log n}{10}} (n-1)p_1^k(1-p_1)^{n-1-k} \]

\[ = o(n^{0.34}). \]

Now use the Markov inequality and proceed as in the proof of (2.2a). The proof of the upper bound for \(|L_{m}^-|^+\) is similar.

Proof of (2.2c)

\[ \Pr((2.2c) \text{ fails in } G_{p_1}) \leq n^{5/4} \sum_{k \leq \frac{\log n}{10}} (n-1)p_1^k(1-p_1)^{n-1-k}^2 \]

\[ = o(1) \]

Let now \(m'\) be as in (2.1c). then

\[ \Pr((2.2c) \text{ fails from some } G_{m}, m' \leq m \leq m_2 \mid (2.2a) - (2.2c) \text{ holds in } G_{m'}) \]

\[ \leq \Pr(\exists e = uv \in E_{n_2} - E_{m'} \text{ such that dist}(u,L_{m'}), \text{ dist}(v,L_{m'}) \leq 3 \text{ in } G_{m'}) \]

\[ \leq n \log \log \log n \left( \frac{n}{\log n} \right)^3 \]

\[ = o(n^{2/5}(\log n)^3)^2/v) \]

\[ [v = \binom{n}{2}] \]

\[ = o(1). \]
Proof of (2.2d)

\[ \Pr(G_p \text{ has 2 vertices with 3 or more common neighbours}) \leq \binom{n}{2} \binom{n-2}{3} p^2 \]

\[ \leq (\log n)^6 / n. \]

We can now use (2.1b) to 'extend' this to \( G_{m_2} \). But if (2.2f) holds for \( G_{m_2} \), it must also hold for \( m \leq m_2 \).

Proof of (2.2e)

Fix \( m \) and \( p = \frac{m}{2} \). Then

\[ \Pr((2.2e) \text{ fails in } G_p) \leq \sum_{k=8}^{n/(\log n)^2} \binom{n}{k} \binom{k}{2} p^{3k+1} \]

\[ = o(n^{-16}). \]

Hence, by (2.1a),

\[ \Pr(\exists m, m_1 \leq m \leq m_2 \text{ such that (2.2e) fails in } G_m) = o(1). \]

Proof of (2.2f)

Now if (2.2e) holds then this on its own implies

\[ |N_m(S)| \geq \frac{\log n}{60} |S| \text{ for } S \subseteq V_n - L_m, |S| \leq \frac{n}{(\log n)^4}. \]

For larger \( S \), we drop the condition \( S \cap L_m = \phi \).
Suppose $S \subseteq V_n$, $|S| \leq \frac{n}{\log n}$. If $v \in V_n - S$ then $\Pr(v \in N_p(S)) = 1 - (1-p)^{|S|} \geq \frac{|S|}{2p}$. Hence

$$\Pr(\exists S \subseteq V_n: \frac{n}{(\log n)^4} \leq |S| \leq \frac{n}{\log n} \text{ and } |N_p(S)| \leq \frac{\log n}{60} |S|)$$

$$\leq \sum_{s = \frac{n}{(\log n)^4}}^{\frac{n}{\log n}} \binom{n}{s} \Pr(B(n-s, \frac{sp}{2}) \leq \frac{s \log n}{60})$$

$$\leq \sum_{s \geq \frac{n}{(\log n)^4}} \binom{ne\alpha}{s} e^{-\alpha ps}$$

for some constant $\alpha > 0$

$$= o(n^{-2}).$$

**Proof of (2.2g)**

Similar to that of (2.2f).

**Proof of (2.2h)**

Let $s = \left[\frac{n}{(\log \log n)^3}\right]$. Now $e_p(S, T)$ is distributed as the binomial random variable $B(s^2, p)$. But

$$\Pr(B(s^2, p) \leq \frac{1}{2} s^2 p) \leq e^{-\frac{1}{8} s^2 p}.$$ 

Hence
Pr((2.2h) fails in $G_p) \leq \left(\frac{n}{s}\right)^2 e^{-\frac{1}{8}s^2p}
= o(n^{-2})

and the result follows in the usual manner.

Proof of (2.2i)

The number of edges of $G_p$ which are contained in $V_n^+$ dominates $B(\frac{1}{8}n^2, p)$.

Now let $\mathcal{G} = \{G : (2.2) holds and \delta(G) \geq 2\}$.

§3. Proof of the theorem

We now describe a way of choosing a large set $\mathcal{G}$ of subgraphs of $G_m \in \mathcal{G}$, most of which are hamiltonian and such that if $C, C'$ are hamilton cycles of distinct $H, H' \in \mathcal{G}$ then $C \neq C'$.

Let $A_m = V_n - NL, B_m = V_n^+ - NL$ and for $v \in A_m$ let $W(v) = \{vw \in E_m : w \in B_m\}$.

Let $L_0 = \lfloor \log n/10 \rfloor$ and $r$ be a prime satisfying $(\log n)^2 \leq r \leq 2(\log n)^2$, let $k = \lfloor \log rL_0 \rfloor$ and $L = r^k$. We treat \{1,2,...,L\} as the points of the $k$-dimensional vector space over the field with $r$ elements, $\text{GF}_r$. This space has $K = r^{k-1}(r^{k-1})(r-1)$ lines. Let the point sets for these lines be the $r$-subsets $X_1, X_2, \ldots, X_K$ of $L$. The only property of these sets used is $|X_i \cap X_j| \leq 1$ for $i \neq j$.

For each $v \in A_m$, we choose a random $L$-subset $W'(v) \subseteq W(v)$ plus a random ordering $w_1, w_2, \ldots, w_L$ (of $W'(v)$). We then define $r$-subsets $W(v,k) \subseteq W'(v), k = 1,2,\ldots,K$ by letting $W(v,k) = \{w_{i_1}, w_{i_2}, \ldots, w_{i_r}\}$ when
Now let \( \phi = \{ f: A \rightarrow \{1,2,\ldots,K\} \} \). For each \( f \in \phi \) we will define a subgraph \( H_f \) of \( G_m \) as follows: delete from \( G_m \) all edges incident with \( A_m \) other than \( \bigcup_{v \in A_m} W(v,f(v)) \). Let now \( \# = \{ H_f: f \in \phi \} \). Observe

\[
|\phi| \geq K (n^4/5) = (\log n)^{(1-\epsilon-o(1))n}
\]

(3.1)

(3.2) If \( C_f, C_g \) are hamilton cycles of \( H_f, H_g \), \( f \neq g \) then \( C_f \neq C_g \).

For if \( f(v) \neq g(v) \) then \( C_f \) uses 2 edges of \( W(v,f(v)) \) and \( C_g \) can use at most one edge of \( W(v,f(v)) \).

Now let \( Z_m = |\{ f \in \phi: H_f \text{ is not hamiltonian} \}| \). We prove

\[
E(Z_m | G \in \phi_m) \leq |\phi|/n^3
\]

and so

\[
\Pr(Z_m \geq \frac{|\phi|}{n} | G \in \phi_m) = o(n^{-2}).
\]

Thus

\[
\Pr(G_m \text{ has fewer than } (1 - \frac{1}{n}) (\log n)^{(1-\epsilon-o(1))n} \text{ hamilton cycles } | G_m \in \phi_m) = o(n^{-2}).
\]

(3.4)
The theorem follows immediately from (3.4).

We must now show that most $H_f$ are hamiltonian.

Consider now a fixed $f \in \Phi$. To prove (3.3) we show

$$\Pr(H_f \text{ is not hamiltonian} | G \in \mathcal{G}_m) = O(n^{-3}).$$

First of all consider the distribution of the edges in the sets $W(v,f(v))$.

**Lemma 3.1**

Conditional on the sub-graph induced by $V_n - A_m$, the sets $W(v,f(v))$ are an independent collection of random $r$-subsets of $B_m$.

**Proof**

Consider a fixed $G_m$, $v \in A_m$ and $W(v) = N_m(v) \cap B_m$. (We cannot assume $G_m \in \mathcal{G}_m$ here.) Replacing $W(v)$ by another subset of $B_m$ of the same size does not change $A_m$ or $NL$. We use here the fact that $w \in B_m$ has at least $\log n / 10$ neighbours in $V_n^+$ and so changing the neighbours of $v \in A_m$ cannot place $w$ in $NL$. It follows that the sets $W(v)$ are independent random subsets and the lemma follows as the $W(v,f(v))$ are random subsets of these.

Let now $X \subseteq E_m$ and $H_f, X = H_f - X$. We say that $X$ is **deletable** if

(3.6a) \[ |X^+| = n \quad \text{where} \quad X^+ = \{ e \in X : e \subseteq V_n^+ \} . \]

(3.6b) \[ |X \cap W(v,f(v))| = 3 \quad \text{for} \quad v \in A_m. \]
(3.6c) \[ X \text{ is not incident with any vertex in} \]
\[ L_m = \{ v \in V : d(v) < \frac{2m2L + \log n}{\log n} \} \]

(3.6d) If \( v \in B_m \) and \( d^+(v) = \lfloor \log n/10 \rfloor + k \) then \( v \) is incident with at most \( k-1 \) edges in \( X \).

(3.6e) No \( v \in B_m \) is incident with \( \frac{2 \log n}{\log n} \) or more edges in \( X^* \).

(3.6f) \[ M^H_f = X(H_{f,x}) \] where \( X \) denotes the length of the longest path in the appropriate graph.

Observe that a calculation similar to that given for (2.2b) shows that \( \hat{L}_m \perp \perp \frac{2}{n} \) in a.e. \( G_m \). We now incorporate this condition into the definition of \( <S_m> \).

Our next lemma deals with the number of neighbours of subsets of \( A \).

For \( S \subseteq V \) and subgraph \( H \) of \( G_m \) let \( \mathcal{N}_{n}^m(S) = \{ w \in S : vw \in E(H) \text{ for some } v \in n \} \).

Lemma 3.2

The following hold with probability \( 1 - o(n^{-1}) \). Here let \( H = H_f \).

(i) \( S \subseteq A_m, \ 1 \leq |S| \leq \frac{n}{10} \) implies \( |N_n(S)| > 80 |s| \).

(ii) \( S \subseteq A, \ T \subseteq B, \ |S| = |T| = \frac{n}{\log n} \) implies that \( H \) contains at least \( n \log n \) edges joining \( S \) and \( T \).

(iii) \( T \subseteq B_m, \ |T| \leq \frac{n}{\log n} \) implies \( |N_n(T) \cap A_j| < 3r |T| \).
Proof

(i)

We first consider $|S| \leq n/3r$ and show $|N_H(S)| \geq r|S|/2$ with the required probability.

$$\Pr(\exists S: |S| \leq n/3r \text{ and } |N_H(S)| \leq r|S|/2) \leq \sum_{s=1}^{n/3r} \left( \frac{n}{s} \right)^{n-n/\epsilon} \left( \frac{r}{\epsilon} \right)^s \left( \frac{rs/2}{n-n/\epsilon} \right)^s$$

$$\leq \sum_{s=1}^{n/3r} \left( \frac{ne}{s} \left( \frac{2(n-n/\epsilon)}{rs} \right)^{r/2} \left( \frac{rs}{2(n-n/\epsilon)} \right)^r \right)^s$$

$$\leq \sum_{s=1}^{n/3r} \left( \frac{ne}{s} \left( \frac{ers}{2(n-n/\epsilon)} \right)^{r/2} \right)^s$$

$$= o(n^{-3}).$$

Suppose now $n/3r < |S| \leq n/600$. Let $S' \subseteq S$ be of size $[n/3r]$. Then

$$|N_H(S)| \geq |N_H(S')|$$

$$\geq r[n/3r]/2$$

$$\geq n/7$$

$$\geq 80 |S|.$$
(ii)

Consider the selection of the sets $W(v, f(v))$ for $v \in S$. This involves $rs$ ($s = |S|$) choices of elements in $B_m$ and each choice always has probability at least $\frac{s-r+1}{n/n_\epsilon}$ of being in $T$. Thus the number of choices, and hence edges in question, stochastically dominates the binomial $B(rs, \frac{s-r+1}{n/n_\epsilon})$. Hence

$$\Pr((iii) \text{ fails}) \leq \binom{n}{s}^2 \Pr(B(rs, \frac{s-r+1}{n/n_\epsilon}) \leq n \log\log n)$$

and the result follows from the Chernoff bound (see for example [3]) for the tails of the binomial since $E(B(rs, \frac{s-r+1}{n/n_\epsilon})) \approx \frac{2rs^2}{n(1+\epsilon)} \geq \frac{2n \log\log n}{1+\epsilon}$.

(iii)

Fix $T \subseteq B_m$, $\frac{n}{r \log n} \leq |T| = t \leq \frac{n}{6r}$ and $S \subseteq A_m$, $|S| = 3r|T|$. Now if $\hat{n} = |B_m|$ then

$$\Pr(W(v, f(v)) \cap T \neq \emptyset \text{ for all } v \in S) = \left(1 - \frac{\hat{n}-t}{\binom{n}{r}}\right)^{3rt} \leq (\frac{2rt}{n})^{3rt} \leq (\frac{2rt}{n})^{3rt}.$$

Hence
Pr((iii) fails) ≤ \( \sum_{t=n/(r \log n)}^{n/6r} \binom{n}{t} \left( \frac{r}{2n} \right)^{\frac{n}{2r}} \left( \frac{2r}{n} \right)^{3rt} \)

\[ ≤ \sum_{t=n/r \log n}^{n/6r} \binom{ne}{t} \left( \frac{r}{3} \right)^{3rt} \]

\[ = o(n^{-3}) \quad \Box \]

Let \( \xi_f \) be the event denoting the occurrence of the conditions in the above lemma.

**Lemma 3.3**

Suppose \( G \in \mathcal{G} \), \( f \in \Phi \), \( \xi_f \) occurs, \( X \) is deletable and \( H = H_{f,X} \). Then

(i) \( S \subseteq V_n \), \( |N_H(S)| < 2|S| \) implies

(a) \( |S| ≥ \frac{n}{600} \)

(b) \( |(S \cup N_H(S)) \cap (B_m)| ≥ \frac{n}{2} + \frac{en}{3} \).

(ii) \( H \) is connected.

**Proof**

Suppose \( S \subseteq V_n \). Let \( S_0 = S \cap L_m \), \( S_1 = S \cap (L_m^+ - L_m) \), \( S_2 = S \cap L_m \) and \( S_3 = S - (S_0 \cup S_1 \cup S_2) \).

Assume first that \( |S_3| ≤ \frac{n}{\log n} \) and \( |S_2| ≤ \frac{n}{600} \).

**Case 1:** \( |S_2| ≤ |S_1 \cup S_3| \).

(a) \( |S - S_2| < 2|NL| \).

Let \( S^* \) be the larger and \( \hat{S} \) the smaller of \( S_1, S_3 \). Then
\[ |N_H(S)| \geq |N_m(S_0)| + |N_m(S^*)| - \frac{2 \log n}{\log \log n} |S^*| - |S_2 \cup \hat{S}| \]

\[ - |N_m(S^*) \cap (S_0 \cup N_m(S_0))| \]

\[ \geq 2|S_0| + \left(\frac{\log n}{60} - \frac{2 \log n}{\log \log n}\right) |S^*| - 3|S^*| - |S^*| \]

\[ \geq 2|S|. \]

(after using (2.2c), (2.2f), (2.2g) and (3.6e) to obtain the second inequality).

(b) \[ |S - S_2| \geq 2|NL|. \]

\[ |N_H(S)| \geq |N_H(S_3)| - |NL \cup S_2| \]

\[ \geq (\frac{\log n}{60} - \frac{2 \log n}{\log \log n}) |S_3| - |NL| - |S_2| \]

\[ \geq 2|S|. \]

(using \( S_0 \cup S_1 \subseteq NL \) and \( |S_2| \leq |S_3| + |NL| \)).

Case 2: \[ |S_2| > |S_1 \cup S_3|. \]

\[ |N_H(S)| \geq 80|S_2| - 3|S_2| + 2|S_0| - |S_1 \cup S_3| \geq 2|S|. \]

Suppose now that \( |S_2| \leq \frac{n}{600} \) and \( \frac{n}{\log n} \leq |S_3| \leq \frac{n}{600} \). Choose \( S'_3 \subseteq S_3 \) of size \( \left\lfloor \frac{n}{\log n} \right\rfloor \) and let \( S' = (S - S_3) \cup S'_3 \). Then
\[ |N_H(S)| \geq |N_H(S')| - |S_3 - S'_3| \]
\[ \geq 2|S_0| + 22|S_2| + \frac{\log n}{200} (|S_1| + |S'_2|) - |S_3 - S'_3| \]
\[ \geq 2|S_0| + 22|S_2| + \frac{\log n}{200} |S_1| + \frac{n}{200} - \frac{\log n}{200} - |S_3| + \left\lfloor \frac{n}{\log n} \right\rfloor \]
\[ \geq 2|S|. \]

We have thus proved (i), part (a).

For part (b), we know, from part (a), that \(|S| \geq \frac{n}{600}\) and hence
\[ |S_2 \cup S_3| \geq \frac{n}{700}. \]

Assume first that \(|S_3| \geq \frac{n}{1400}\). Suppose \(|(S_3 \cup N_H(S_3)) \cap B_m| < \frac{1}{2} n + \frac{en}{3}\). Then there exists \(T \subseteq B_m\) of size at least \(\frac{en}{7}\) such that \(N_H(S_3) \cap T = \emptyset\). Now it follows from (2.2h) that \(G_m\) contains at least \(\frac{n \log n}{2(\log \log n)^6}\) edges joining \(S_3\) and \(T\). But \(X\) contains at most \(n\) edges joining \(S_3\) and \(T\) and so \(N_H(S_3) \cap T \neq \emptyset\) — contradiction.

Assume next that \(|S_2| \geq \frac{n}{1400}\). The proof here is similar to that above, but relying on Lemma 3.2(ii) in place of (2.2h), and the fact that \(X\) contains only 3 edges incident with each \(v \in A_m\).

(ii)

Suppose \(H\) is not connected and there exists \(S \subseteq V_n\), \(|S| \leq \frac{1}{2} n\) such that there are no \(S\) to \(V_n - S\) edges in \(H\). Now \(|(V_n - S) \cap (B_m)| \geq \frac{en}{3}\) and (i) implies \(|S| \geq \frac{n}{600}\). We obtain a contradiction using (2.2h) or Lemma 3.2(ii) as in (i)(b).

\(\square\)

Suppose now that \(H_f\) is not hamiltonian and \(X\) is deletable. Let \(P = (x_0, x_1, \ldots, x_\lambda)\) be a longest path of both \(H_f\) and \(H = H_f, X^*\). If
x_1x_\lambda \in E(H)$, $i \neq 0$, then the associated rotation with $x_0$ fixed and broken edge $x_1x_{i+1}$ yields a new longest path $\rho(P,x_0,x_1) = (x_0,x_1,\ldots,x_{\lambda-1},x_{i+1})$.

Let $\text{END}(P,x_0)$ denote the set of other endpoints of longest paths which are obtainable in $H$ from $P$ by a sequence of rotations, with $x_0$ fixed, and starting from $P$.

We will restrict our allowable rotations to those where the broken edge is an edge of the starting path $P$. We further restrict ourselves so that if $P'$ is obtained from $P$ by a sequence of rotations through paths $P = P_0,P_1,\ldots,P_k = P'$ then the paths $P_1,P_2,\ldots,P_k$ have distinct endpoints, other than $x_0$.

Suppose that the paths produced in the construction of $\text{END}(P,x_0)$ are $\mathcal{S} = \{P^0,P^1,P^2,\ldots\}$ where $P^0 = P$ and $P^{i+1}$ is obtained from some $P^j$, $j \leq i$, by a single rotation.

Let $\text{END} = \text{END}(P,x_0) \cup \{x_0\}$ and for each $x \in \text{END}$ let $P_x$ denote the first path (in the above ordering) with endpoint $x$ (so that $P_{x_0} = P$). For $x \neq x_0$ let $\text{END}(x) = \text{END}(P_x,x)$. Now a simple modification of the argument of Posa [6] shows that

$$|N_H(\text{END}(x))| < 2|\text{END}(x)|.$$ 

(Indeed, all we have to show is that if $v \in N_H(\text{END})$ with neighbours $w_1,w_2$ on $P$ then $\{w_1,w_2\} \cap \text{END} \neq \emptyset$. Suppose $w' \in \text{END}$ and $vw' \in E(H)$. Consider the neighbours $w'_1,w'_2$ of $v$ on $P_{w'}$. If $\{w_1',w_2'\} = \{w_1,w_2\}$ then some allowable rotation from $P_{w'}$ shows one of $w'_1,w'_2$ is in END. If say $w_1 \in \{w_1',w_2'\}$ then the sequence of rotations that created $P_{w'}$ deleted the edge $vw_1$ and so $w_1 \in \text{END}$.)
We deduce from Lemma 3.3 that

\[(3.7a)\quad |\text{END}(x)| > ^\gamma \quad \text{for} \quad x \in \text{END}\]

\[(3.7b)\quad |\text{END}| > ggy\]

\[(3.7c)\quad \text{Each} \quad P, \quad x \in \text{END}, \quad \text{contains at least} \quad \sum_{i=0}^{2} r_{i} \quad \text{edges with both endpoints in} \quad B_{m}.\]

To see (3.7c) let \(n_{i}, \quad i = 0, 1, 2\) denote the number of edges of \(P\) with \(i\) vertices in \(B_{m}\). Then

\[|\text{END}(x)| \geq n_{0} I (|V(P_{x}) \cap B| - |V(P_{x}) \cap (V_{n} \cup N_{L})|) - 1.\]

Since \(P_{x}\) is a longest path, it must contain \(N_{r_{i}}(\text{END}(x))\). But then Lemma 3.3 implies

\[|\text{END}(x) \cup N_{r_{i}}(\text{END}(x))| \geq n_{0} + \gamma - (|n_{0} - |n_{0} + \gamma| + o(n)) - 1\]

and (3.7c) follows.

Given (3.7) we consider two possibilities.

Case 1: there exists \(x \in \text{END}\) such that \(|\text{END}(x) \cap B_{m}| < \frac{1200}{n_{m}}\).

Case 2: \(|\text{END}(x) \cap B_{m}| < \frac{rr}{r_{r}}\) for all \(x \in \text{END}\).

Case 1 is easier to deal with and is considered first. Without loss of generality assume \(|\text{END} \cap B_{m}| > \text{STMTT}\) i.e. \(x = x_{n}\) suffices above. Observe that because \(H_{r}\) is connected,

\[(3.8)\quad x \in \text{END}, \quad y \in \text{END}(x) \implies xy \in E(H_{r}).\]

(We use the "colouring" argument of Fenner and Frieze [5] to show this is unlikely when a large number of \(x \in B\). Since \(A\) contains no edges in \(H_{r}\),
(3.8) does not help so much in Case 2 and we are in a similar situation to that encountered in the case of random bipartite graphs, Frieze [6]).

Suppose now that given $G_m \in \mathcal{G}_m$, we randomly pick $X \subseteq \mathcal{E}_m$ satisfying (3.6a), (3.6b). We consider two events:

$\mathcal{E}_1 = \mathcal{E}_f \cap \{G_m \in \mathcal{G}_m'; H_f$ is not hamiltonian, Case 1 occurs$\}$

$\mathcal{E}_2 = \mathcal{E}_1 \cap \{X$ is deletable$\}$.

We show

\begin{equation}
\Pr(\mathcal{E}_2 | \mathcal{E}_1) \geq \frac{1}{2}(1 - \frac{2}{r})^n (1 - \frac{20}{\log n})^n
\end{equation} \tag{3.9a}

\begin{equation}
\Pr(\mathcal{E}_2) \leq c_1^n \text{ for some constant } 0 \leq c_1 < 1.
\end{equation} \tag{3.9b}

We can then deduce

\begin{equation}
\Pr(\mathcal{E}_1) \leq (c_1 + o(1))^n.
\end{equation} \tag{3.10}

Proof of (3.9a)

Fix $G \in \mathcal{G}_m$ and the choices $W(v, f(v))$ for $v \in A_m$. Fix some longest path $P$ of $H_f$. Consider first the edges of $X$ that meet $A_m$. Each $W(v, f(v))$ contains at most 2 edges of $P$. This accounts for the term $(1 - \frac{2}{r})^n$. Now consider the remaining $n$ edges of $X$. Now to avoid $P$ and the edges incident with NL, $X$ must avoid at most $n + o(n)$ edges, given (2.2a), (2.2b). Using this and (2.2i) we obtain $(1 - \frac{20}{\log n})^n$ as a lower bound for the probability of avoiding these edges. Given that these edges are not selected, the probability that (3.6d) or (3.6e) fails is $o(1)$, which accounts
for the $\frac{1}{2}$.

**Proof of (3.9b)**

Consider fixed graphs $\hat{G}, \hat{H}$. We show

\begin{equation}
Pr(\xi_2 | G_m - X = \hat{G}, H_{f,X} = \hat{H}) \leq c_1^n
\end{equation}

and (3.9b) follows.

Observe that $G_m - X, H_{f,X}$ together determine $A_m$ by $v \in A_m$ iff $v \not\in n_e$ and it loses edges in $H_{f,X}$. $NL$ is then determined by $v \in NL$ iff $v \in A_m$ and $d^+(v) \leq \log n \over 10$ or $v \in V_{n_e}$ and $v$ is the neighbour of such a vertex.

If $Pr(\xi_2 | G_m - X = \hat{G}, H_{f,X} = \hat{H}) > 0$ then there exists $X$ such that $\xi_2$ occurs for $\hat{G} + X$, $\hat{H} + X$. Hence we may assume that (3.7) holds where $\text{END}$, $\text{END}(x), x \in \text{END}$ are determined by $\hat{H}$ only (and are independent of $X$). We may also assume Case 1 occurs in $\hat{H}$.

Furthermore the edges in $X$ are required to conform to (3.8). Thus let $\hat{\xi}_2$ denote the event $\{x \in \text{END}, y \in \text{END}(x) \implies xy \not\in X\}$. Then

\begin{equation}
Pr(\hat{\xi}_2 | G_m - X = \hat{G}, H_{f,X} = \hat{H}) \leq Pr(\xi_2 | G_m - X = \hat{G}, H_{f,X} = \hat{H}, (3.6c),(3.6d)).
\end{equation}

(For (3.12) use $Pr(A|BC) \geq Pr(AB|C)$ for events $A,B,C$).

Let us now consider the distribution of $X$ given $G_m - X, H_{f,X}$ and (3.6c), (3.6d). Let $X = X^+ \cup (\bigcup_{v \in \hat{A}_m} Y_v)$, where for $v \in \hat{A}_m$, $Y_v = \{vw \in X\}$. We claim that
(3.13a) \( X^+ \) is a random \( n \)-subset of \( B_m^{(2)} - E(G) \).

(3.13b) For \( v \in \Lambda_m \), \( Y_v \) is a random \( 3 \)-subset of \( \{vw \in E(G) : w \in B_m \} \) and these subsets are independent of each other.

(3.13a) follows from the fact that given (3.6c), (3.6d) holds for one \( X \), the addition (and subsequent deletion) of any \( n \)-subset of \( B_m^{(2)} - E(G) \) does not affect \( H_{f,X} \) and (3.6c), (3.6d) will still hold. (3.13b) follows from Lemma 3.1 and its proof.

Now for \( w \in \text{END} \cap B_m \) let \( \beta(w) = |\text{END}(w) \cap B_m | \). The following 2 subcases cover all possibilities:

Case 1a: \( |\{w : \beta(w) > \frac{n}{1200}\}| > \frac{n}{2400} \)
Case 1b: \( |\{w : \beta(w) < \frac{n}{1200}\}| > \frac{n}{2400} \).

It follows from (3.13a) that, where \( \nu^+ = \binom{n-n}{2} \) and \( \hat{m} \leq m \),

\[
\Pr(\hat{\xi}_2 | \text{Case 1a}) \leq \left[ \frac{\nu^+ - \hat{m} - 3n^2/(2(2400))^2}{n} \right] / \binom{n}{\nu^+ - \hat{m}} \leq \left( \frac{95999}{96000} \right)^n.
\]

It follows from (3.13b) that

\[
\Pr(\hat{\xi}_2 | \text{Case 1b}) \leq (1 - \frac{3}{2400})^{n/1200}.
\]

We have thus confirmed (3.9b).

Let us now consider Case 2. Let \( \xi_1 \) be as before, except that Case 2
replaces Case 1 and let $\xi_2$ now be defined with respect to the new $\xi_1$.

(3.9a) continues to hold. We prove

\[(3.9b') \quad \Pr(\xi_2 | G - X = \hat{G}, H_f, X = \hat{H}) \leq c_2^n \quad \text{for some constant} \quad 0 < c_2 < c_2(e) < 1\]

which combined with (3.9a) yields

\[(3.10') \quad \Pr(\xi_1) \leq (c_2 + o(1))^n.\]

From (3.10) and (3.10') and the fact that $\Pr(\xi_f | G \in \mathcal{G}_m) = 1 - o(n^{-3})$ we obtain (3.3) and the theorem.

We observe that (3.13) continues to hold. We can assume that $\hat{H}$ contains a longest path $P$ with endpoints $x_0, x_1$ and $\frac{n}{1200}$ vertices $\text{END} \subseteq A_m$ and for each $x \in \text{END}$ there is a set of $\frac{n}{600}$ paths $\phi_x$ with distinct endpoints (END(x)). These will have been constructed from a path $P_x$ by rotations as in the discussion prior to (3.7).

We now consider in more detail the construction of END($P, x_0$). Let $T = T(x_0)$ denote the tree with vertex set END($P, x_0$), rooted at $x_1$ and with an edge directed from $x$ to $y$ if $P_y$ is obtained by a single rotation from $P$. Let $\mathcal{J}$ be the set of possible trees that can be so constructed.

Consider the following condition:

\[\mathcal{A}: \text{there exists } T \in \mathcal{J} \text{ such that } T \text{ contains a subtree } T', \text{ rooted at } x_1, \text{ which has (i) } |V(T') \cap A_m| \geq \frac{n}{1200} \text{ and (ii) } |V(T') \cap B_m| \leq \frac{n}{4800}.\]

Suppose now that $\mathcal{A}$ holds. For each $v \in \text{END}' = V(T') \cap A_m$ let $\phi(v)$ denote the neighbour of $v$ on $P_v$. 

Lemma 3.4

If \( d \) holds then \( |\phi(\text{END'})| \geq \frac{n}{9600}. \)

Proof

We show first

\begin{equation}
(3.14) \quad y \in \phi(\text{END'}) - V(T') \implies |\phi^{-1}(y)| \leq 2.
\end{equation}

We do this by showing that if \( y = \phi(x) \) then \( xy \) is an edge of \( P \). This is clearly true if \( x = x_1 \). If \( x \neq x_1 \) then \( y \) is adjacent to \( x \) on \( P_x \). If \( xy \) is not an edge of \( P \) then \( y \) is an ancestor of \( x \) in \( T' \), a contradiction, as \( y \notin V(T') \).

Now (3.14) implies that

\begin{equation}
(3.15) \quad |\phi(\text{END'})| \geq \frac{1}{2}|\text{END'} - \phi^{-1}(\phi(\text{END'}) \cap V(T'))|.
\end{equation}

But since \( \phi^{-1}(\phi(\text{END'}) \cap V(T')) \subseteq \bigcap_{H_m} N_{B_n} \cap A_m \) we see from Lemma 3.3 and \( \delta(\text{ii}) \) that

\[ |\phi^{-1}(\phi(\text{END'}) \cap V(T'))| \leq \frac{n}{4800} \cdot 3r \]

and the lemma follows from this and (3.15).

It is important to note that any path obtained from \( P_x, x \in \text{END'} \) by a sequence of rotations with \( x \) fixed has \( \phi(x) \) as \( x \)'s neighbour.

Suppose now that \( \delta \) does not hold. We will obtain a contradiction. Let \( T \in \mathcal{F} \). Since \( |V(T) \cap A_m| \geq \frac{n}{1200} \) we must have \( |V(T) \cap B_m| > \frac{n}{4800} \). Then \( T \)...
contains a subtree $\tilde{T}$ with $|V(\tilde{T}) \cap B_m| = \frac{n}{\log n}$ and since $\ast \ast$ does not hold

$|V(\tilde{T}) \cap A \cap B|$. Let $S = V(\tilde{T}) \cap B_m$. It follows from (2.2h) that $|N^+(S) \cap A| = \frac{m}{\log m}$.

Now if $v \in S$, $w \in N^+(S) \cap B$ and $vw \in E(\tilde{T})$ then we can legitimately construct $p(P_x, x_0, w)$ unless the associated broken edge $ww' \in E(P)$. But this latter condition rules out at most $2|V(\tilde{T})|$ rotations: $2$ for each added edge of each $P_x, v \in V(\tilde{T})$. The same $w'$ can be produced at most twice in this way. Thus there exists $T \in J$ which contains a subtree which is obtained from $\tilde{T}$ by adding at least $1 - \frac{n}{1200} \log |N^+|$ leaves. Since $s \neq l$ does not occur, at least $\frac{n}{1200} \log |N^+| - \frac{n}{1200}$ of these new leaves are in $B_m$. But this means Case 1 holds, a contradiction.

Applying this argument for each $x \in \text{END}$ i.e. constructing a tree $T(x)$ of paths starting with $P_x$, we deduce, from Lemma 3.4 that the following is true:

**Lemma 3.5**

In $H$ there are $\frac{n}{9600}$ vertices $y_1^{1}, y_2^{2}, \ldots$ in END fl $A_m$ and a set of $\frac{n}{9600}$ vertices $z_1^{1}, z_2^{2}, \ldots$ in $B_m$ such that for each $i$ there are $\frac{n}{1200}$ longest paths with one endpoint $y_i^{1}$ adjacent to $y_i^{1}$ on each path and the other endpoints of each set of $\frac{n}{1200}$ paths are distinct members of $A$. D

Let $Y_i, i = 1, 2, \ldots, \frac{n}{9600}$ denote the set of other endpoints of the paths with one fixed endpoint $y_i^{1}$.

We can now confirm (3.9b). We must add random edges, as in (3.13), and show that with high probability these extra edges make the resulting graph hamiltonian or have a longer path than $H$.

We consider the edges in (3.13b) to be added randomly in 3 waves $X_1, X_2, X_3$. 
\(K_j U X^t\) where \(|X_j| = 1^1 = 1^1 = |A_m|\) and each \(v \in A_m\) is incident with one edge of each \(X_t^t\) \(t = 1, 2, 3\).

**Adding \(X_1\)**

For \(y \in Y = U_{i \in Y} y_i\) let \(\delta(y) = |\{i : y \in Y_i\}|\). Clearly \(|Y' I| \leq \infty\)

where \(Y' = \{y \in Y : \delta(y) \geq \frac{n}{8(1200)^2}\}\).

If \(y \in Y'\) then independently of other members of \(Y'\)

\[\text{Pr}(\text{for some } i, X_i \text{ contains an edge } yz, \text{ where } y \in Y_i) \geq \frac{1}{4(1200)^2}\]

Hence there exist constants \(0 < f_1, \theta_1 < 1\) such that

\[\text{Pr}(S_3) \geq 1 - \theta_1\]

where

\[S_3 = \{X_i \text{ contains } f_n \text{ edges of the form } z_y, y \in Y_i\}\].

Assume now that \(S^3\) occurs.

We now have \(f_n\) cycles \(C_1, C_2, \ldots\) say, plus an edge joining \(y_1\) to \(C_1\). Applying (3.7c) we see that each \(C_1\) contains a set of vertices \(K_1\), \(|K_1| > \frac{1}{8}z\) en, where \(v \in K_1\) implies \(v\) lies on an edge of \(C_1\) with both endpoints in \(B_1\).

**Adding \(X_2\)**

Now, independently, for each \(i\), \(\text{Pr}(X_i^* \text{ contains an edge } y_1u\) where
u \in K_1 \geq \epsilon$. By considering these cycles one by one, we see that there exist constants $0 < \xi_2 = \xi_2(\epsilon), \eta_2 = \eta_2(\epsilon) < 1$ such that

$$\Pr(\xi_4 | \xi_3) > 1 - \eta_2^n$$

where

$$\xi_4 = \{X_2 \text{ contains } \xi_2^n \text{ edges of the form } y_i u_i, u_i \in K_1$$

and the $B_m$ neighbours $v_1, v_2, \ldots$ of $u_1, u_2, \ldots$ on $C_1, C_2, \ldots$ are distinct).

Now each time $X_2$ contains an edge $y_i u_i, u_i \in K_1$, we can obtain a longest path of $\hat{H} + (X_1 \cup X_2)$ with one endpoint $y_i$ and the other endpoint in $B_m$ by using the edges $(C_1 \cup \{y_i u_i\}) - \{u_i, v_i\}$.

Assume that $\xi_4$ occurs.

Adding $X_3 \cup X^+$

We now have $\xi_3^n$ longest paths $Q_1, Q_2, \ldots$ of $\hat{H} + (X_1 \cup X_2)$, each with a distinct endpoint $v_i \in B_m$. We are now essentially in a Case 1 situation. Take each $Q_i$ and using $v_i$ as a fixed endpoint generate $\geq \frac{n}{600}$ longest paths by rotations. Now throw in $X_3 \cup X^+$. The probability that we fail to close one of these paths is exponentially small. (3.9b') follows and we are done.

References


