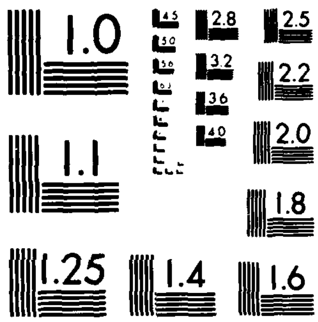
AD-A149 194 ON LARGE MATCHINGS AND CYCLES IN SPARSE RANDOM GRAPHS 1/1
(U) CARNEGIE-MELLON UNIV PITTSBURGH PA MANAGEMENT
SCIENCES RESEARCH GROUP A M FRIEZE JAN 84 MSRR-504
UNCLASSIFIED N00014-75-C-0621 F/G 12/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

(1)

ON LARGE MATCHINGS AND

CYCLES IN SPARSE RANDOM GRAPHS

by

A. M. Frieze*

January 1984

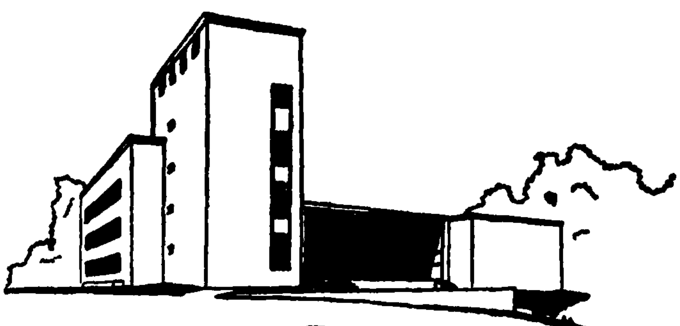
Contract N00014-75-C-0621

Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213

GRADUATE SCHOOL OF INDUSTRIAL ADMINISTRATION

WILLIAM LARIMER MELLON, FOUNDER



Management Science Research Report No. MSRR 504

ON LARGE MATCHINGS AND CYCLES IN SPARSE RANDOM GRAPHS

by

A. M. Frieze*

January 1984

Contract N00014-75-C-0621

* Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA 15213, U.S.A.
(On leave from Queen Mary College, London)

Management Sciences Research Group
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Abstract

Let $p = c/n$ where c is a large constant. We show that the random graph $G_{n,p}$ a.s. contains a matching of size $n(1 - (1+\varepsilon(c))e^{-c})/2$ and a cycle of size $n(1-(1+\varepsilon(c))ce^{-c})$ where $\varepsilon(c)$ is some function satisfying $\lim_{c\to\infty} \varepsilon(c) = 0$.

1. In this paper ~~we study~~ Studied the size of the largest matching and cycle in random graphs with edge probability c/n where c is a large constant. We continue the analysis of Bollobás [2], Bollobás, Fenner and Frieze [3] and confirm the conjecture in the final paragraph of the latter paper.

We shall let $G_{n,p}$ denote a random graph with vertex set $V_n = \{1,2,\ldots,n\}$ in which edges are chosen independently with probability p. We say that $G_{n,p}$ has a property Q <u>almost surely</u> (a.s.) if $\lim_{n\to\infty} \Pr(G_{n,p} \in Q) = 1$.

For $c > 0$ define $\alpha(c)$, $\beta(c)$ by

$$\text{(1.1)} \qquad \alpha(c) = \sup(\alpha \ge 0:\ G_{n,c/n} \text{ a.s. contains a matching of size at least } \alpha n/2)$$

and

$$\text{(1.2)} \qquad \beta(c) = \sup(\beta \ge 0:\ G_{n,c/n} \text{ a.s. contains a cycle of size at least } \beta n).$$

Our main result is an improved estimate of $\beta(c)$. However the same methods can be used to estimate $\alpha(c)$ and we shall do this first as the analysis is marginally simpler.

In what follows $p = c/n$ and $\varepsilon_1(c)$, $\varepsilon_2(c)$ are unspecified functions satisfying $\lim_{c\to\infty} \varepsilon_i(c) = 0$, $i=1,2$.

<u>Theorem 1.1</u>

$$\text{(1.3)} \qquad \alpha(c) = 1 - (1 + \varepsilon_1(c))e^{-c}$$

and this remains valid if $c\to\infty$.

As far as we know the only other paper dealing with this question is by Karp and Sipser [7] who prove some strong results about a simple heuristic for finding a large cardinality matching.

There has been more work done on estimating $\beta(c)$. Ajtai, Komlós and Szemerédi [1] and Fernandez de la Vega [6] showed that $\beta(c) \geq 1-c_0/c$. Bollobás made a significant step forward by showing that $G_{n,p}$ a.s. contains a large Hamiltonian subgraph and that $\beta(c) \geq 1-c^{24}e^{-c/2}$. By refining this analysis, Bollobás, Fenner and Frieze [3] showed that $\beta(c) \geq 1-c^6e^{-c}$. The main result of this paper is

Theorem 1.2

$$(1.4) \qquad \beta(c) = 1 - (1+\varepsilon_2(c))\, ce^{-c}$$

and this remains valid if $c \to \infty$.

Corollary 1.3

A random digraph with edge density c/n a.s. contains a directed cycle of size $n(1 - (1+\varepsilon_2(c))ce^{-c})$.

Notation

The following notation is used throughout. Let G be a graph. V(G), E(G) denote the sets of vertices and edges of G.

For $S \subseteq V(G)$ we let $G[S] = (S,E(S))$ where $E(S)=\{e \in E(G): e \subseteq S\}$.

$N_G(S) = \{w \in S:$ there exists $v \in S$ such that $\{v,w\} \in E(G)\}$.

For $v \in V(G)$ we write $N_G(v)$ for $N_G(\{v\})$ and $d_G(v)$ for the degree of v. $\mu(G)$ is the maximum cardinality of a matching of G.

$$BS(x,m) = \sum_{k=0}^{\lfloor x \rfloor} \binom{m}{k} p^k (1-p)^{m-k}$$

As the case $c > \log n$ is well known we shall assume for convenience that $ce \leq 3\log n$.

2. Lemma 2.1

Let $G = G_{n,p}$ and let vertex v be small if $d_G(v) \le c/10$ and large otherwise. Let SMALL, LARGE be the sets of small and large vertices respectively.

Let $W = W_1 \cup W_2$ where for k=1,2

$W_k = \{v : v$ is small and there exists a small w such that v and w are joined by a path of length $k\}$

Then for $c \ge 300$ G a.s. satisfies the following:

(2.1) $|\{v \in V_n: d_G(v) \le c/10 + 1\}| \le ne^{-2c/3}$;

(2.2) there does not exist $S \subseteq V_n$ with $|S| \ge ne^{-c}$ and $|\{e \in E(G): e \cap S \ne \emptyset\}| \ge 4c\,|S|$;

(2.3) $d_G(v) \le 4\log n$ for $v \in V_n$;

(2.4) $|W| \le c^2 e^{-4c/3} n$;

(2.5) $\emptyset \ne S \subseteq V_n$, $|S| \le n/14$ and $S \subseteq LARGE$ implies $|N_{G(S)}| \ge 6\,|S|$;

(2.6) $S \subseteq V_n$, $n/14 \le |S| \le n/2$ implies $|\{\{v,w\} \in E(G) : v \in S, w \in S\} \ge c\,|S|/10$;

Proof

To prove (2.1) note that for n large

$Exp(\ |\{v \in V_n : d_G(v) \le c/10 + 1\}\ ) = n\,BS(c/10 + 1, n-1) \le ne^{-.669c}$.

Now the variance of this set size can be shown to be $\le ne^{-2c/3}$.

Thus one can use either the Chebycheff or Markov inequality depending on whether or not c remains bounded as n tends to infinity.

Next note that the probability there exists a set S violating (2.2) is no more than

$$\sum_{s \ge ne^{-c}} \binom{n}{s}\binom{sn}{|4cs|} p^{|4cs|}$$

$$\le \sum_{s \ge ne^{-c}} \left(\frac{ne}{s}\right)^s \left(\frac{snep}{4cs}\right)^{4cs}$$

$$\le \sum_{s \ge ne^{-c}} \left(\frac{e^{5+1/c}}{256}\right)^{cs} = o(1).$$

To prove (2.3) we observe that

$$\mathrm{Exp}(|\{v \in V_n : d_G(v) > 4\log n\}|) = n \sum_{k > 4\log n} \binom{n-1}{k} p^k (1-p)^{n-k-1}$$

$$\le n \sum_{k > 4\log n} \left(\frac{ce}{k}\right)^k = o(1)$$

as $ce \le 3\log n$.

Next let $P_k = \{$paths of length k in G with small endpoints$\}$. Now clearly

(2.7) $\quad |W_k| \le 2\,|P_k| \quad$ for $k=1,2$.

Furthermore

(2.8) $\quad \mathrm{Exp}(|P_1|) = \binom{n}{2} p \lambda^2$

where $\lambda = BS(c/10 - 1, n-2) \le e^{-.669c}$

Now

$$\mathrm{Exp}(|P_1|^2) = \mathrm{Exp}(|P_1|) + \binom{n}{2}\binom{n-2}{2} p^2 \lambda_1 + 2(n-2)\binom{n}{2} p^2 \lambda_2$$

where

$$\lambda_1 = \Pr(\text{SMALL} \supseteq \{1,2,3,4\} \quad |E(G) \supseteq \{\{1,2\}, \{3,4\}\})$$
$$\leq \Pr(|N_G(1) \cap \{5,6,\ldots,n\}| \leq c/10 - 1)^4$$
$$\leq (\lambda(1-p)^{-2})^4$$

and

$$\lambda_2 = \Pr(\text{SMALL} \supseteq \{1,2,3\} \quad |E(G) \supseteq \{\{1,2\}, \{2,3\}\})$$
$$\leq (\lambda(1-p)^{-1})^3.$$

This gives

(2.9) $\quad \mathrm{Var}(|P_1|) \leq ce^{-4c/3}n \qquad$ for n large.

Similar calculations give

(2.10a) $\quad \mathrm{Exp}(|E_2|) = (1+o(1))n^3p^2\lambda^2/2$

and

(2.10b) $\quad \mathrm{Var}(|E_2|) \leq n^3p^2\lambda^2 \qquad$ for n large

(2.4) now follows from (2.7), (2.8), (2.9) and (2.10).

To prove (2.5) we first consider S for which $1 \leq s = |S| \leq n/35000e^4$. Let $T = S \cup N_G(S)$ and $t = |T|$. If (2.5) does not hold for S then $|T| \leq m_1 = \lceil n/5000e^4 \rceil$ and T contains at least $m_2 = \lceil ct/140 \rceil$ edges of G. The probability that such a T exists is no more than

$$\sum_{t=1}^{m_1} \binom{n}{t} \binom{\binom{t}{2}}{m_2} p^{m_2} \leq \sum_{t=1}^{m_1} \left(\frac{ne}{t}\right)^t \left(\frac{t^2ep}{2m_2}\right)^{m_2}$$

$$\leq \sum_{t=1}^{m_1} \left(\frac{ne}{t}\right)^t \left(\frac{70et}{n}\right)^{ct/140} \leq \sum_{t=1}^{m_1} \left(\frac{4900e^4t}{n}\right)^{ct/280} = o(1)$$

using $c \geq 300$.

For $|S| \geq m_3 = \lceil n/36000e^4 \rceil$ we can ignore the fact that the vertices of S are large. The probability that such an S exists violating (2.5) is no more than

$$\sum_{s = m_3}^{\lfloor n/14 \rfloor} \binom{n}{s}\binom{n}{6s}(1-p)^{s(n-7s)}$$

$$\leq \sum_{s = m_3}^{\lfloor n/14 \rfloor} \left(\frac{ne}{s}\right)^{s} \left(\frac{ne}{6s}\right)^{6s} e^{-cs/2}$$

$$\leq \sum_{s = m_3}^{\lfloor n/14 \rfloor} \left(6^8 \cdot 10^{21} \cdot e^{35} \cdot e^{-c/2}\right)^{s} = o(1)$$

which proves (2.5).

The probability that (2.6) does not hold is not more than

$$\sum_{s=\lceil n/14 \rceil}^{\lfloor n/2 \rfloor} \binom{n}{s} BS(cs/10,\ s(n-s))$$

$$\leq 2 \sum_{s=\lceil n/14 \rceil}^{\lfloor n/2 \rfloor} \left(\frac{ne}{s}\right)^{s} \left(\frac{10s(n-s)e}{cs}\right)^{cs/10} \left(\frac{c}{n}\right)^{cs/10} e^{-cs/3}$$

$$\leq 2 \sum_{s=\lceil n/14 \rceil}^{\lfloor n/2 \rfloor} \left(14e(10e)^{c/10} e^{-c/3}\right)^{s} = o(1).$$

∎

The proofs of our theorems rely on the removal of a certain set of vertices. We must show that this set is not too large. The following Lemma deals with part of this set.

<u>Lemma 2.2</u>

Let X_0 = SMALL and let the sequence of sets $X_1, X_2, \ldots, X_s$ be defined by

$$X_i = \{v \in V_n : |N_G(v) \cap \bigcup_{t=0}^{i-1} X_t| \geq 2\}$$

and let s be the smallest $i \geq 1$ such that $X_{i+1} = X_i$. Let $X = \bigcup_{i=1}^{s} X_i$, then

(2.11) $\quad |X| \leq 2e^4 c^4 e^{-4c/3} n \qquad$ a.s.

Proof

For $x \in X \cup X_0$ let $i(x)=\min\{i:x \in X_i\}$ and let $D(x)=(V(x), A(x))$ denote a digraph inductively constructed as follows: for $x \in X_0$, $D(x) = (\{x\},\emptyset)$ and for $x \in X_0$ let y_1, y_2 be 2 distinct neighbours of x satisfying $i(x) > i(y_1), i(y_2)$. Then

$$D(x) = (V(y_1) \cup V(y_2) \cup \{x\}, A(y_1) \cup A(y_2) \cup \{(x, y_1), (x, y_2)\})$$

Each $D(x)$ is acyclic, (weakly) connected and satisfies

(2.12) each $v \in V(x)$ has outdegree 0 or 2 and x is the unique vertex of indegree 0.

Let

k = the number of vertices of outdegree 2 = $|K(x)|$, where $K(x)=S(x)-X_0$.

and let

ℓ = the number of vertices of outdegree 0 = $|L(x)|$, where $L(x)=S(x) \cap X_0$.

It follows then that

(2.13a) $|A(x)| = 2k$

and we will show

(2.13b) $\ell \leq k+1$ and if $\ell=k+1$ then $D(x)$ is a binary tree rooted at x.

This is most easily proved by induction on k. A digraph satisfying (2.12) has at least one vertex y whose outneighbours z_1, z_2 both have outdegree zero. Removing arcs (y, z_1) and (y, z_2) and any vertex which becomes isolated we obtain a smaller digraph satisfying (2.12).

We obtain from the above that we can associate with each $x \in X$, a set $V(x)$ of vertices and a partition of $V(x)$ into $K(x)$, $L(x)$ satisfying

(2.14a) $x \neq x'$ implies $V(x) \neq V(x')$;

(2.14b) if $k = |K(x)|$, $\ell = |L(x)|$ then $2 \le \ell \le k+1$;

(2.14c) $L(x) \subseteq$ SMALL;

(2.14d) $G(x)=G[V(x)]$ is connected and has at least 2k edges;

(2.14e) if $\ell=k+1$ and $G(x)$ has 2k edges then $G(x)$ is a tree with leaves $L(x)$.

We estimate $|X_S - X_0|$ by counting sets of vertices satisfying (2.14). For a given k, ℓ, m let $\lambda_{k,\ell,m}$ be the expected number of sets K, L with $|K|=k$, $|L|=\ell$ satisfying (2.14) above, where $G[K \cup L]$ has m edges. Then

$$\lambda_{k,\ell,m} \le \binom{n}{k}\binom{n}{\ell}\binom{\binom{k+\ell}{2}}{m} p^m BS(c/10,\ n-k-\ell)^{\ell}$$

$$\le \left(\frac{ne}{k}\right)^k \left(\frac{ne}{\ell}\right)^{\ell} \left(\frac{(k+\ell)^2 e}{2m}\right)^m \left(\frac{c}{n}\right)^m e^{-2c\ell/3}\left(1 - \frac{c}{n}\right)^{-\ell(k+\ell)}$$

$$= \mu_{k,\ell,m}$$

Now if $c \le 2\log n$, $k, \ell \le n^{1/3}$ then $\mu_{k,\ell,m+1}/\mu_{k,\ell,m} \le n^{-1/4}$ for n large. Thus

$$\sum_{m=2k}^{\binom{k+\ell}{2}} \lambda_{k,\ell,m} \le (1+o(1))\ \mu_{k,\ell,2k}\ . \tag{2.15}$$

With the same bounds on c,k,ℓ and with n large and $\ell \le k+1$ we have

$$\mu_{k,\ell,2k} \le 21 n^{\ell-k}(e^4c^2k)^k \ell^{-\ell} e^{-2c\ell/3} \quad \text{which implies} \tag{2.16}$$

$$\sum_{\ell=2}^{k+1} \mu_{k,\ell,2k} \le 21(e^4c^2k/n)^k \sum_{\ell=2}^{k+1} (n/\ell e^{2c/3})^{\ell}$$

$$\le n(e^4c^2)^k e^{-2ck/3}$$

$$\le ne^{-ck/2} \qquad \text{as } c \ge 300.$$

It follows that $s \le \log n$ a.s., and we can assume $k \le \log n$. Now, using (2.16),

$$\sum_{k=2}^{\log n} \sum_{\ell=2}^{k} \mu_{k,\ell,2k} \le 21 \sum_{k=2}^{\log n} (e^4c^2)^k e^{-2ck/3}$$

$$\le 22(e^4c^2)^4 e^{-4c/3}$$

and so

(2.17) the number of sets K, L with $2 \le \ell \le k$ is a.s. less than $n^{1/2}e^{-4c/3}$.

We only need to consider the case $\ell=k+1$ from now on. But as $\mu_{k,k+1,m+1}/\mu_{k,k+1,m} \le 3ck/n$ we have

$$(2.18) \qquad \sum_{m\ge 2k} \mu_{k,k+1,m} \le (1+o(1))\mu_{k,k+1,2k}$$

So we are finally reduced to estimating

τ_k = the number of vertex induced binary trees with k leaves (k-b-trees) in which each leaf is small.

Let θ_k be the number of (vertex labelled) k-b-trees contained in a complete graph with 2k-1 vertices. (Clearly $\theta_k \le (2k-1)^{2k-3}$). Then

$$(2.19) \quad \mathrm{Exp}(\tau_k) = \binom{n}{2k-1}\theta_k p^{2k-2}(1-p)^{\binom{2k-1}{2}-2k+2} BS(c/10-1,n-2k+1)^k$$

$$\le n(e^2c^2e^{-2c/3})^k \qquad \text{for } n \text{ large.}$$

To estimate $\mathrm{Var}(\tau_k)$, let $\{T_1, T_2, \ldots, T_B\}$, $B = \binom{n}{2k-1}\theta_k$, be the set of k-b-trees contained in a complete graph with n vertices. Let A_i be the event that T_i is a vertex induced subgraph of G_p in which all leaves are small.

Next let $Y_p = \{(i,j): |V(T_i) \cup V(T_j)| = p\}$ for $p = 2k-1, \ldots, 4k-2$ and let $Z_{p,q} = \{(i,j) \in Y_p : |E(T_i) \cup E(T_j)| = q\}$. Then

$$\text{(2.20)} \qquad \mathrm{Exp}(\tau_k^2) = \mathrm{Exp}(\tau_k) + \Delta_1 + \Delta_2$$

where

$$\Delta_1 = \sum_{(i,j) \in Y_{4k-2}} \Pr(A_i \cap A_j)$$

and

$$\Delta_2 = \sum_{p=2k-1}^{4k-3} \sum_{(i,j) \in Y_p} \Pr(A_i \cap A_j)$$

Now

$$\Delta_1 \le \binom{n}{2k-1}^2 \left(\theta_k p^{2k-2} (1-p)^{\binom{2k-1}{2} - 2k + 2}\right)^2 \sigma$$

where

$$\sigma = BS(c/10-1,\ n-2k+1)^k \ BS(c/10-1,\ n-4k+2)^k$$

is an estimate of the probability that all leaves of 2 particular disjoint trees are small.

It follows that

$$\text{(2.21)} \qquad \Delta_1 \le \mathrm{Exp}(\tau_k)^2 (1-p)^{-2k^2}$$

Now for $p \le 4k-3$ we have

$$\sum_{(i,j)\varepsilon Y_p} \Pr(A_i \cap A_j) = \sum_{q=p-1}^{4k-4} \sum_{(i,j)\varepsilon Z_{p,q}} \Pr(A_i \cap A_j)$$

$$\le \sum_{q=p-1}^{4k-4} \binom{n}{p}\binom{\binom{p}{2}}{q}\binom{q}{2k-1}^2\left(\frac{c}{n}\right)^q e^{-2ck/3}(1-p)^{-8k^2}$$

(2.22) $\qquad \le ne^{-ck/2} \qquad\qquad$ for n large.

(2.19), (2.20), (2.21), (2.22) plus the Chebycheff inequality implies that τ_k is a.s. within a factor $(1+o(1))$ of the R.H.S. of (2.19). This together with (2.17) and (2.18) proves the result.

■

For a positive integer k, the <u>k-core</u> $V_k(G)$ is defined to be the largest set $S \subseteq V_n$ such that $\delta(G[S]) \ge k$. This is well defined, for if $\delta(G[S_i]) \ge k$ for $i=1,2$ then $\delta(G[S_1 \quad S_2]) \ge k$. We let G_k denote the subgraph of G induced by $V_k(G)$.

The k-core can be constructed using the following algorithm:

```
begin
    H:=G;
    while  δ(H) < k do
    begin
        Y: =  {v ε V(H) : d_H(v) < k};
        H: = H[V(H) - Y]
    end
end
```

On termination $H=G_k$. This is because one can easily show inductively that each iteration removes vertices that are not in $V_k(G)$ and as $\delta(H) \geq k$ we have $V(H) \subseteq V_k(G)$.

Clearly any matching of G is contained in G_1(= G minus isolated vertices) and any cycle of G is contained in G_2.

Now for k=1,2 let $A_k = A_k(G_{n,p}) = V_k(G_{n,p})-(W \cup X \cup Y_k)$ where W,X are as defined in Lemmas 2.1, 2.2 respectively and

$$Y_k = \{y \in V_n : d_{G_{n,p}}(y) = k \text{ and } N_{G_{n,p}}(y) \cap X \neq \emptyset\}.$$

Let $H_k = H_k(G_{n,p}) = G_{n,p}[A_k]$, then we have

Lemma 2.3

For k=1,2 let M be any matching of $G_{n,p}[A_k]$ which is not incident with any small vertex. Let $\hat{H}_k=H_k-M$, then (2.5) implies:

(2.23) $\emptyset \neq S \subseteq A_k$, $|S| \leq n/14$ implies $|N_{\hat{H}_k}(S)| \geq k|S|$.

Proof

Let $G=G_{n,p}$, $H=\hat{H}_k$ and for a given S let $S_1 = S \cap \text{SMALL}$ and $S_2 = S-S_1$. Now

(2.24) $|N_H(S)| \geq |N_H(S_1)| - |S_2| + |N_H(S_2)| - \min(|S_1|,|S_2|)$

We can write $\min(|S_1|, |S_2|)$ in place of $|S_1|$ as no vertex of S_2 is adjacent to more than one vertex of S_1, as $S_2 \cap X = \emptyset$.

Also, we claim

(2.25) $|N_H(S_1)| \geq k|S_1|$.

Note first that $v \in S_1$ implies $d_{G_k}(v) \geq k$ and no pair of vertices of S_1 are adjacent, since $S_1 \cap W_1 = \emptyset$. Note that no pair of vertices of S_1 have a common neighbour as $S_1 \cap W_2 = \emptyset$. Also $N_G(S_1) \cap (W \cup Y_k) = \emptyset$ as

$S_1 \cap W_1 = \emptyset$. Furthermore $v \in S_1$ implies $|N_G(v) \cap X| \leq 1$ as $S_1 \cap X = \emptyset$. Thus to prove (2.25) we need only show that if $v \in S_1$ and $d_G(v)=k$ then $N_{G(v)} \cap X = \emptyset$. But this follows from $S_1 \cap Y_k = \emptyset$.

We claim next that if (2.5) holds then

(2.26) $|N_H(S_2)| \geq 4|S_2|$

For then $|N_G(S_2)| \geq 6|S_2|$ and for each $v \in S_2$, $|N_G(v)| \leq |N_H(v)| + 2$. This is because v is incident with at most one edge of M and is adjacent to at most one vertex of W X Y_k. It is a simple matter to verify (2.23) from (2.24), (2.25) and (2.26) by considering $|S_1| \geq |S_2|$ and $|S_1| < |S_2|$ as separate cases.

■

3. Matchings

Let H_1 be the subgraph of G defined in Lemma 2.3. We are going to prove that H_1 a.s. has a perfect or near perfect matching. We first establish that H_1 is large.

Lemma 3.1

(3.1) $$|V(H_1)| = n(1 - (1+\varepsilon_1(c))e^{-c}) \qquad \text{a.s.}$$

where $\varepsilon_1(c) \to 0$ as $c \to \infty$.

Proof

$$|V(H_1)| \geq |V_1(G)| - |W| - |X| - |Y_1 - W| .$$

It is well known that

(3.2) $$|V_1(G)| = (1+o(1))n(1-e^{-c}) \qquad \text{a.s.}$$

where the o(1) term in (3.2) could for example be taken to be $\pm n^{-1/4}e^{-c/2}$, using the Chebycheff inequality.

Lemmas 2.1 and 2.2 give a.s. upper bounds on $|W|$, $|X|$ and (3.1) will follow from

(3.3) $$|Y_1 - W| \leq |X|$$

For $y \in Y_1$ there is, by definition, a unique $x(y) \in X$ such that y is adjacent to $x(y)$ in G. Now for distinct y_1, y_2 $\in Y_1 - W$ we have $x(y_1) \neq x(y_2)$ else $y_1 \in W_2$ and (3.4) follows.

We establish next the following condition that goes with a graph not having a (near) perfect matching.

Lemma 3.2

Suppose $\mu(H) < \lfloor |V(H)|/2 \rfloor$. Let $\mathfrak{M}$ be the set of maximum cardinality matchings of H. Let $U=\{u_1, u_2,\ldots,u_t\}$ be the set of vertices left isolated by some $M \varepsilon \mathfrak{M}$. For i=1,2,...,t there exists a set $U_i \subseteq U$ satisfying

(3.4a) $|N_H(U_i)| < |U_i|$;

(3.4b) $w \varepsilon U_i$ implies $e=\{u_i,w\} \notin E(H)$ and $\mu(H) < \mu(H+e)$.

Proof

Let $u_i \varepsilon U$ and let some $M_i \varepsilon \mathfrak{M}$ leave u_i isolated. Let $S_i \neq \emptyset$ be the set of vertices, different from u_i, left isolated by M_i. Let U_i' be the set of vertices reachable from S_i be an even length alternating path w.r.t. M_i. Let $U_i = S_i \bigcup U_i' \subseteq U$. Then (3.4b) holds otherwise M_i has an augmenting path.

If $u \varepsilon N_H(U_i)$ then $u \notin S_i$ and so there exists y_1 such that $\{u,y_1\} \varepsilon M_i$. We show that $y_1 \varepsilon U_i$ which will prove (3.4a). Now there exists $y_2 \varepsilon U_i$ such that $\{u,y_2\} \varepsilon E(H)$. Let P be an even length alternating path from some $s \varepsilon S_i$ terminating at y_2. If P contains $\{u,y_1\}$ we can truncate it to terminate with $\{u,y_1\}$, otherwise we can extend it using edges $\{y_2,x\}$ and $\{x,y_1\}$.

We are now ready for the

Proof of Theorem 1.1

We use a coloring argument that was introduced in Fenner and Frieze [5]. Suppose that after generating $G=G_{n,p}$ all its edges are colored blue, and then each edge of G is re-colored green with probability $p'=\log n/cn$ and left blue with probability $1-p'$. These recolourings are done independently of each other.

Let E^b, E^g denote the blue and green edges respectively and let $G^b = (V_n, E^b)$, $H_1 = H_1(G)$ and $H_1^b = H_1(G^b)$.

Remark 3.1

It is important to note that for a fixed value of E^b, E^g is a random subset of $\overline{E}^b$ where each $e \in \overline{E}^b$ is independently included in E^g with probability $p_1 = pp'/(1-p(1-p'))$ and excluded with probability $1-p_1$.

Consider next the following 2 events:

$\mathcal{G} \equiv G = G_{n,p}$ satisfies the conditions of Lemmas 2.1, 2.2 and $\mu(H_1) < \lfloor |V(H_1)| \rfloor /2$.

$\mathcal{E} \equiv$ (a) $\emptyset \neq S \subseteq A_1(G^b)$, $|S| \leq n/14$ implies $|N_{H_1^b}(S)| \geq |S|$;

(b) $\mu(H_1^b) < \lfloor |V(H_1^b)|/2 \rfloor$;

(c) there does not exist $e = \{v,w\} \in E^g$, $e \subseteq A_1(G^b)$ such that some maximum cardinality matching of H_1^b leaves both v and w isolated.

In consequence of what has already been proved, we need only prove

(3.5) $\lim_{n\to\infty} \Pr(\mathcal{G}) = 0.$

To prove (3.5) we shall prove

(3.6a) $\Pr(\mathcal{E} \mid \mathcal{G}) \geq (1 - o(1))(1-p')^{2n/3}$

(3.6b) $\Pr(\mathcal{E}) \leq (1-p_1)^{n^2/392}$

which together imply (3.5).

Proof of (3.6a)

Let $G_0 \in \mathcal{G}$ be fixed and let M_0 be any fixed maximum cardinality matching of H_1. We prove

(3.7) $\Pr(\mathcal{E} \mid G_{n,p} = G_0) \geq (1-p')^{2n/3} - 16(\log n)^4/c^2n.$

We can readily verify this once we have shown that

(3.8) $\mathcal{E} \cap \mathcal{G} \supseteq \mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3 \cap \mathcal{G}$

where

$\mathcal{E}_1 \equiv E^g$ is a matching of G_0;

$\mathcal{E}_2$ = no green edge meets any vertex of degree less than c/10+2 in G_0 or any vertex in $W \times Y_1$

$\mathcal{E}_3 = M_0 \cap E^g = \emptyset$

For $\mathcal{E}_1 \cap \mathcal{E}_2$ implies

(3.9) $A_1(G_0^b) = A_1(G_0)$

and then $\mathcal{E}_1$ implies (see Lemma 2.3) that (2.23) holds, which verifies $\mathcal{E}(a)$. $\mathcal{E}(b)$ follows directly from (3.9) and $G_0 \in \mathcal{G}$. $\mathcal{E}_3$ implies $\mu(H_1^b) = \mu(H_1)$ and $\mathcal{E}(c)$.

Now it follows from (2.3) that

(3.10) $\Pr(\bar{\mathcal{E}}_1) \leq 16(\log n)^4/c^2n.$

From Lemmas 2.1, 2.2 and (3.3) we find that the total number of edges of G_0 that are excluded by the conditions in $\mathcal{E}_2$, $\mathcal{E}_3$ is no more than

$$n((c/10 + 1)e^{-2c/3} + 4nce^{-ce})n + n/2 \leq 2n/3$$

Thus

$$\Pr(\bar{\mathcal{E}}_1 \cup \bar{\mathcal{E}}_2 \cup \bar{\mathcal{E}}_3) \leq 1-(1-p')^{2n/3} + 16(\log n)^4/c^2n$$

which proves (3.7).

Proof of (3.6b)

Now

(3.11) $\Pr(\mathcal{E}) = \sum_{\Gamma} \Pr(\mathcal{E} \mid G^b = \Gamma)\Pr(G^b = \Gamma)$

where Γ is an arbitrary graph with vertices V_n.

Now if $H_1(\Gamma)$ fails to satisfy $\mathcal{E}(a)$, $\mathcal{E}(b)$ then $\Pr(\mathcal{E}|G^b = \Gamma) = 0$. So let us assume that $\mathcal{E}(a)$, $\mathcal{E}(b)$ hold.

Now if U, $U_1,\ldots,U_t$ are as defined in Lemma 3.2 with $H=H_1$, then each set is of size at least $n/14$ and for $\mathcal{E}(c)$ to hold no green edge can join $u_i \in U$ to $w \in U_i$. But then in view of Remark 3.1 and $\mathcal{E}(a)$ we have

$$\Pr(\mathcal{E}(c)\ |G^b = \Gamma) \le (1-p_1)^{n^2/392}$$

which implies (3.6b).

We have thus shown that

$$\mu(G) \ge n(1 - (1+\varepsilon_1(c))e^{-c})/2 \qquad \text{a.s.}$$

On the other hand (3.2) implies

$$\mu(G) \le n(1+o(1))(1 - e^{-c})/2 \qquad \text{a.s.}$$

and Theorem 1.1 follows.

■

If we put $c=\log n + \omega$ where $\omega\to\infty$ then we have $\alpha(c)=1-(1+o(1))e^{-\omega}n^{-1}$ and then $G_{n,p}$ a.s. has a matching of size at least $(n - (1+o(1))e^{-\omega})/2$. This is Erdos and Rènyi's result [4], (what we have proved is that H_1 a.s. has a matching of size $\lfloor |V(H_1)|/2 \rfloor$ and one can see that when $c=\log n+\omega$, $H_1=G_{n,p}$ a.s.).

4. Cycles

Let H_2 be the subgraph of G defined in Lemma 2.3. We are going to prove that H_2 a.s. has a hamiltonian cycle. The proof is very similar to that of section 3 and as such we will only give the essential differences.

Lemma 4.1

(4.1) $\quad |V(H_2)| = n(1 - (1+\varepsilon_2(c))ce^{-c}) \qquad$ a.s.

where $\varepsilon_2(c) \to 0$ as $n \to \infty$

Proof

$$|V(H_2)| \geq |V_2(G)| - |W| - |X| - |Y_2 - W \cup X|$$

Now

$$|Y_2 - W \cup X| \leq |X|$$

follows by a similar argument to (3.3). Now let Z_0 be the set of vertices of degree 0 or 1 in G and let $Z_1, Z_2, \ldots$ be the sequence of sets removed in each iteration of the 2-core finding algorithm of section 2. Now, corresponding to (3.2), it is also well known that

$$Z_0 = (1-o(1))n(1-ce^{-c}) \qquad \text{a.s.}$$

We complete the proof of the lemma by showing that

$$Z_i \subseteq X \cup W_1 \cup Y_2 \qquad i=1,2,\ldots$$

Thus assume inductively that $Z_1, Z_2, \ldots Z_{i-1} \subseteq X \cup W_1 \cup Y_2$ for some $i \geq 1$ (true vacuously for i=1) and let $T = \bigcup_{t=0}^{i-1} Z_t$.

Then $y \in Z_i$ implies $d_G(y) \geq 2$ but $|N_G(y) - T| \leq 1$.

<u>Case 1</u>: $|N_G(y) \cap T| \geq 2$

By assumption $T \subseteq X \cup SMALL$ and so $y \in X$.

<u>Case 2</u>: $|N_G(y) \cap T| = 1$.

Then $d_G(y)=2$ implies $y \in X \cup W_1 \cup Y_2$.

■

<u>Lemma 4.2</u>

If c is large enough and G satisfies the conditions in Lemmas 2.1, 2.2 then H_2 is connected.

<u>Proof</u>

If $H=H_2$ is not connnected then there exists a *nonempty* $S \subseteq V(H)$ such that $N_H(S) = \emptyset$. We show that this is not possible for c large enough. (2.23) implies that $|S| \geq n/14$. (4.1) implies that, for c large, fewer than $2ce^{-c}n$ vertices are deleted from G in producing H. Then (2.2) implies that at most $8c^2e^{-c}n$ edges are lost in the construction. But then (2.6) implies that not all edges with one vertex in S have been deleted.

■

The analogue of Lemma 3.2 is

<u>Lemma 4.3</u>

Let H be a connected graph which is non-hamiltonian. Then

(a) (4.2) no edge of H joins the endpoints of any longest path of H.

(b) Let $U=\{u_1, u_2,\ldots,u_t\}$ be the set of vertices which are endpoints of longest paths of H. For i=1,2,...t there exists $U_i \subseteq U$ satisfying

(4.3a) $|N_H(U_i)| < 2\,|U_i|$;

(4.3b) $w \in U_i$ implies $\{u_i,w\} \notin E(H)$ and there is some longest path of H that joins u_i to w.

Proof

(4.2) is straightforward and (4.3) is from Posá [11] ∎

We can now give an outline of the

Proof of Theorem 1.2

We define E^b, E^g and G^b as in the proof of Theorem 1.1 and let $H_2^b = H_2(G^b)$. Let now

$\mathcal{G} \equiv G = G_{n,p}$ satisfies the conditions of Lemma's 2.1, 2.2 and H_2 is not hamiltonian, which implies that (4.2) holds with $H=H_2$.

We have only to show that (3.5) holds with this definition of $\mathcal{G}$. Let now

$\mathcal{E} \equiv$ (a) $\emptyset \neq S \subseteq A_2(G^b)$, $|S| \le n/14$ implies $|N_{H_2^b}(S)| \ge 2\,|S|$;

(b) there does not exist $e=\{v,w\} \in E^b \cup E^g$ such that v, w are the endpoints of some longest path of H_2^b.

We replace (3.6) by

(4.3a) $\Pr(\mathcal{E}|\mathcal{G}) \ge (1-o(1))(1-p')^{3n/2}$;

(4.3b) $\Pr(\mathcal{E}) \le (1-p_1)^{n^2/392}$.

This will prove the theorem.

To prove (4.3a) let $G_0 \in \mathcal{G}$ be fixed and let P_0 be some longest path of H_2. We define $\mathcal{E}_1$, $\mathcal{E}_2$ as before and define $\mathcal{E}_3 \equiv P_0 \cap E^g = \emptyset$.

Now $\mathcal{E}_1 \cap \mathcal{E}_2$ implies that $A_2(G_0^b) = A_2(G_0)$ and then (3.8) and (4.3a) will follow in the same way as (3.8) and (3.6a) previously.

To prove (4.3b) we use (3.11) and concentrate on the case where $H_2(\Gamma)$ satisfies $\mathcal{E}$(a). We note that for $\mathcal{E}$(b) to hold there is no $\{v,w\} \in E^g$, $v_i \in U$, $w \in U_i$ where $U, U_1, U_2, \ldots U_t$ are defined by (4.3) w.r.t. $H=H_2(\Gamma)$. (4.3b) follows from Remark 3.1 and $\mathcal{E}$(a) as before.

We note that if we put $c=\log n+\log\log n+\omega$ where $\omega\to\infty$ then we obtain the result of Komlós and Szemerédi [8]and Korsunov [9].

Finally note that our Corollary follows from the Percolation Theorem of McDiarmid [10].

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, 'The longest path in a random graph', Combinatorica 1 (1981), 1-12.

[2] B. Bollobás, 'Long paths in sparse random graphs', Combinatorica 2 (1982).

[3] B. Bollobás, T. I. Fenner and A. M. Frieze, 'Long cycles in sparse random graphs' to appear in the Proceedings of the 1983 Cambridge Conference on Combinatorics in honour of Paul Erdos.

[4] P. Erdos and A. Rényi, 'On the existence of a factor of degree one of a connected random graph', Acta Mathematica Academiae Scientiarum Hungaricae 17 (1966) 359-368.

[5] T. I. Fenner and A. M. Frieze. 'On the existence of hamiltonian cycles in a class of random graphs', Discrete mathematics 45 (1983).

[6] W. Fernandex De La Vega, 'Long paths in random graphs'.

[7] R. M. Karp and M. Sipser, 'Maximum matchings in sparse random graphs', 22nd IEEE Conference on the Foundations of Computer Science (1981) 364-375.

[8] J. Komlós and E. Szemerédi, 'Limit distribution for the existence of hamiltonian cycles in random graphs', Discrete Mathematics 43 (1982) 55-63.

[9] A. D. Korsunov, 'Solution of a problem of Erdos and Rényi on hamiltonian cycles in nonoriented graphs', Soviet Mathematics Doklaidy 17 (1976) 760-764.

[10] L. Pósa, 'Hamilton circuits in random graphs', Discrete Mathematics 14 (1976) 359-364.

END

FILMED

2-85

DTIC