Reprinted from JOURNAL OF COMBINATORIAL THEORY, Series B
All Rights Reserved by Academic Press, New York and London

Vol. 40, No. 2, April 1986
Printed in Belgium

Maximum Matchings in a Class of Random Graphs

A. M. FRIEZE

Department of Computer Science and Statistics, Queen Mary College, University of London, London E1 4NS, England

Communicated by the Editors

Received March 1984

Let $V_n = \{1, 2,..., n\}$ and $\mathscr{D}(n, m)$ be the set of digraphs with vertex set V_n in which each $v \in V_n$ has outdegree m. $\vec{D}(n, m)$ is chosen uniformly at random from $\mathscr{D}(n, m)$ and then $D(n, m)$ is obtained by ignoring the orientation of the edges of $\vec{D}(n, m)$. We show that

$$\lim_{\substack{n \to \infty \\ n \text{ even}}} \Pr(D(n, 1) \text{ has a perfect matching}) = 0,$$

$$\lim_{\substack{n \to \infty \\ n \text{ even}}} \Pr(D(n, 2) \text{ has a perfect matching}) = 1.$$

1. INTRODUCTION

This paper is concerned with the following three related models of a random graph $D_\delta(n, m)$, $\delta = 0, 1, 2$: let $V_n = \{1, 2,..., n\}$ and suppose that each $v \in V_n$ independently chooses m vertices $w_1, w_2,..., w_m$ and adds the arcs (v, w_i), $i = 1, 2,..., m$, to create a random digraph $\vec{D}_\delta(n, m)$ with nm arcs. When

$\delta = 2$: $w_1, w_2,..., w_m$ are distinct members of $V_n - \{v\}$ and each m-subset is equally likely to be chosen;

$\delta = 1$: $w_1, w_2,..., w_m$ are chosen independently and uniformly from $V_n - \{v\}$;

$\delta = 0$: $w_1, w_2,..., w_m$ are chosen independently and uniformly from V_n.

Thus $\vec{D}_\delta(n, m)$, $\delta = 0, 1$, may contain parallel arcs but $\vec{D}_2(n, m)$ cannot.

We obtain the graph $D_\delta(n, m)$ by ignoring orientation, removing loops, and allowing parallel edges to coalesce.

From now on we use the following convention: all probabilistic statements are based on the probability space of $\vec{D}_\delta(n, m)$ and all graph theoretic statements concern $D_\delta(n, m)$, unless specifically stated otherwise.

0095-8956/86 $3.00

The three models obviously have very similar properties, but we are mainly interested in the case $\delta = 2$. These graphs suffice as an approximate model of a sparse random graph with a lower bound of m on the vertex degrees. On the other hand, we find in one point of our proof that it is useful to have the little bit of extra independence available in the cases $\delta = 0$ or 1.

For a graph property Π we say that $D_\delta(n, m)$ *almost surely* (a.s.) has property Π if

$$\lim_{n \to \infty} \Pr(D_\delta(n, m) \text{ has } \Pi) = 1.$$

In Fenner and Frieze [2] we studied the connectivity of $D_2(n, m)$ and in [3] we showed that $D_2(n, 23)$ is a.s. Hamiltonian. An interesting open problem is that of determining m_0, the smallest m such that $D_2(n, m)$ is a.s. Hamiltonian. It is known that $m_0 \geqslant 3$ and that the value 23 can be reduced, but the exact value of m_0 is not known, although we strongly suspect $m_0 = 3$.

Shamir and Upfal [7] showed that $D_2(n, 6)$ a.s. has a perfect matching for n even. The main aim of this paper is to tighten this.

It is clear that if $D_\delta(n, m)$ a.s. has a perfect matching, then so do $D_{\delta+1}(n, m)$ for $\delta = 0, 1$. Thus we obtain a complete answer to when these graphs a.s. have a perfect matching by proving

THEOREM 1.1.

(a) $$\lim_{\substack{n \to \infty \\ n \text{ even}}} \Pr(D_2(n, 1) \textit{ has a perfect matching}) = 0,$$

(b) $$\lim_{\substack{n \to \infty \\ n \text{ even}}} \Pr(D_0(n, 2) \textit{ has a perfect matching}) = 1.$$ ∎

It is interesting to note that Walkup [9] obtained the same result in the bipartite analogues of these graphs.

Finally, note here that recently Grimmett [5], Grimmett and Pulleyblank [6] have studied $D_1(n, m)$ in relation to the vertex packing problem.

2. Proof of Theorem 1.1

(a) $D_2(n, 1)$ has no perfect matching if there exists a vertex having two neighbors of degree 1. A standard application of the Chebycheff inequality shows that there will a.s. be a large number of such vertices.

(b) In this case we let $D_n = D_0(n, 2)$ and introduce the following notation: For a graph G let $V(G)$, $E(G)$ denote the sets of vertices and edges of G, respectively. For $S \subseteq V(G)$ let $G[S] = (S, E_S)$ where $E_S = \{e \in E(G): e \subseteq S\}$, let $N_G(S) = \{w \notin S$: there exists $v \in S$ such that $\{v, w\} \in E(G)\}$ and let $N(S) = N_{D_n}(S)$. As usual a stable set S of G is a set of vertices S for which $E_S = \emptyset$.

Our main tool is a refinement of Tutte's Theorem [8] on the existence of a perfect matching. It is due independently to Gallai [4] and Edmonds [1]. We do not need the full theorem, only:

LEMMA 2.1. *If a graph G does not have a perfect matching then there exists $K \subseteq V(G)$, $|K| = k \geqslant 0$ such that if $H = G[V(G) - K]$ then*

H has at least $k+1$ components with an odd number of vertices; (2.1a)

no odd component of H, which is not an isolated vertex, is a tree. (2.1b) ∎

We call such a set K a bad set.

We proceed via a sequence of Lemmas to show that D_n a.s. has no bad sets, for n even.

LEMMA 2.2. *For positive integers k, l define the event $E_1(k, l)$ by*

$$E_1(k, l) = \text{there exists } K, L \subseteq V_n,\ K \cap L = \emptyset,$$
$$|K| = k,\ |L| = l \text{ such that } N(L) \subseteq K.$$

For $0 < \varepsilon < 1$ let $u(\varepsilon) = ((1-\varepsilon)/e^4(1+\varepsilon)^{1+\varepsilon}))^{1+\varepsilon}$ and suppose that $u = u(\varepsilon)$ satisfies $e^{4u}/u^u \leqslant 2^{1/6}$.

Then where $n_1 = \lfloor un \rfloor$ and $l_1 = \lceil (1+\varepsilon)k \rceil$ and

$$E_1(\varepsilon) = \bigcup_{k=1}^{n_1} \bigcup_{l=l_1}^{\lfloor n/2 \rfloor} E_1(k, l)$$

we have

$$\lim_{n \to \infty} \Pr(E_1(\varepsilon)) = 0.$$

Proof.

$$\begin{aligned}
\Pr(E_1(k, l)) &\leqslant \frac{n!}{k!\,l!\,(n-k-l)!} \left(\frac{k+l}{n}\right)^{2l} \left(1 - \frac{l}{n}\right)^{2(n-k-l)} \\
&\leqslant \frac{(ne)^{k+l}(k+l)^{2l} e^{-2l(n-k-l)/n}}{k^k l^l n^{2l}} \\
&\leqslant \frac{l^l e^{4k}}{k^k n^{l-k}}, \qquad 1 \leqslant k \leqslant n_1,\ l_1 \leqslant l \leqslant n/2.
\end{aligned}$$

If we put $u_l = (l/n)^l$ we find $u_l u_{l-1} \leqslant le/n$. Thus, by dividing the range of l at $\lfloor n/2e \rfloor$, say, we find

$$\sum_{l=l_1}^{\lfloor n/2 \rfloor} \left(\frac{l}{n}\right)^l \leqslant 2\left(\frac{l_1}{n}\right)^{l_1} + n2^{-n/2e}.$$

Thus

$$\begin{aligned}\sum_{k=1}^{n_1} \sum_{l=l_1}^{\lfloor n/2 \rfloor} \Pr(E_1(k, l)) &\leqslant \sum_{k=1}^{n_1} \left(\frac{ne^4}{k}\right)^k \left(2\left(\frac{l_1}{n}\right)^{l_1} + n2^{-n/2e}\right) \\ &\leqslant 2e \sum_{k=1}^{n_1} \left(\frac{(1+\varepsilon)^{1+\varepsilon} e^4 k^\varepsilon}{n^\varepsilon}\right)^k + n^2 \left(\frac{e^{4u}}{u^u 2^{1/2e}}\right)^n \\ &= o(1). \quad \blacksquare\end{aligned}$$

We now consider the case where D_n contains a bad set K, $|K| \leqslant u(\varepsilon)n$ for some ε, and $E_1(\varepsilon)$ does not occur. The next lemma proves the occurrence of a second event $E_2(\varepsilon)$ which we deal with in Lemma 2.4.

LEMMA 2.3. *Suppose D_n contains a bad set K, $1 \leqslant k = |K| \leqslant u(\varepsilon)n$, and no subset of K is bad. Let $H = G[V_n - K]$ have $s \geqslant k+1$ odd components $C_1, C_2, \dots, C_s$ with $n_1 = n_2 = \cdots = n_p = 1 < 3 \leqslant n_{p+1} \leqslant \cdots \leqslant n_s$ vertices, respectively.*

Assume that $E_1(\varepsilon)$ does not occur. Then there exists a partition K, P, Q, M of V_n with $p = |P|$, $q = |Q|$ satisfying

$N(M) \subseteq K$, $N(P) \subseteq K$, $N(Q) \subseteq K$; (2.2a)

P is a stable set; (2.2b)

Q contains at least q edges; (2.2c)

each vertex of K is adjacent to at least one member of $P \cup Q$; (2.2d)

$2 \leqslant k \leqslant u(\varepsilon)n$, $0 \leqslant p+q < (1+\varepsilon)k$, $p + \lfloor q/3 \rfloor \geqslant k$

and $q = 0$ implies $p \geqslant k+1$. (2.2e)

Proof. Define r by $n_{p+1} \leqslant \cdots \leqslant n_r < (1+\varepsilon)k \leqslant n_{r+1} \leqslant \cdots \leqslant n_s$. We show first that

$$n_1 + n_2 + \cdots + n_r < (1+\varepsilon)k \quad \text{and} \quad s \leqslant r+1. \tag{2.3}$$

Case 1. $s \geqslant r+1$.

If $n_{r+1} \leqslant n/2$ then $E_1(\varepsilon)$ occurs with $L = C_{r+1}$ and so $n_{r+1} > n/2$ and $s = r+1$. If $r = 1$ then (2.3) follows immediately. Otherwise, if

$n_1+\cdots+n_r \geqslant (1+\varepsilon)k$ then there exists $t \leqslant r$ such that $n_1+\cdots+n_{t-1} < (1+\varepsilon)k \leqslant n_1+\cdots+n_t \leqslant 2(1+\varepsilon)k$. But as $2(1+\varepsilon)\,u(\varepsilon)n < n/2$ we see that $E_1(\varepsilon)$ occurs with $L=\bigcup_{i=1}^{t} C_i$. Thus we have demonstrated (2.3) in this case.

Case 2. $s=r$.

We have only to show $n_1+\cdots+n_r < (1+\varepsilon)k$ and this is shown as above.

So now let $P=\bigcup_{i=1}^{p} C_i$, $Q=\bigcup_{i=p+1}^{r} C_i$, and $M=V_n-(K\cup P\cup Q)$. Now (2.2a), (2.2b) are immediate consequences of these definitions. Equation (2.2c) follows from (2.1b). To prove (2.2d) we use the minimality of K.

We show that $v\in K$ must be adjacent to vertices in at least two of $C_1,\ldots,C_s$. If v is adjacent to none of these components then $C_1,\ldots,C_s$ remain as odd components of $H_1=G[V_n-(K-\{v\})]$. If v is only adjacent to one of these components then at least $s-1$ of these remain as components of H_1.

It only remains to prove (2.2e). The bounds of k are part of the assumption and $p+q<(1+\varepsilon)k$ follows from (2.3). $p+\lfloor q/3\rfloor \geqslant k$ follows from $r\geqslant k$ and the fact that $|C_i|\geqslant 3$ for $i=p+1,\ldots,r$. To examine the case $q=0$ let n_0 be the number of vertices in even components of H. Suppose $p=k$, then $s=t+1$ and $2k+n_s+n_0=n$. But as n_0, n are even, this implies n_s is even, a contradiction. ∎

Let us refer to the existence of a partition satisfying (2.2) as the occurrence of $E_2(\varepsilon)$.

We can immediately show for any fixed integer k_0

$$\lim_{n\to\infty} \Pr(D_n \text{ has a bad set } K, \text{ with } 1\leqslant |K| \leqslant k_0)=0. \tag{2.4}$$

Let us take $\varepsilon=1/2k_0$ and assume $E_1(\varepsilon)$ does not hold. If there is a bad set K with $1\leqslant|K|\leqslant k_0$ then Lemma 2.3 implies that (2.2) holds for some $k\leqslant k_0$. But (2.2e) implies

$$q<3\varepsilon k/2 \tag{2.5}$$

which in this case implies $q<1$, or $q=0$. But then $p\geqslant k+1$ contradicts $p<(1+\varepsilon)k$.

In the proof of the following lemma we assume $k\geqslant k_0$ for some suitably large k_0 whose size need not be discussed until (2.10).

LEMMA 2.4.

$$\lim_{n\to\infty} \Pr(E_2(\varepsilon))=0 \qquad \textit{for small } \varepsilon.$$

Proof. For a given small ε let $E_2(k, p, q)$ refer to $E_2(\varepsilon)$ with given values for k, p, q. Then

$$\Pr(E_2(k, p, q)) \leqslant \frac{n!}{k!\, p!\, q!(n-p-q-k)!}\left(\frac{k}{n}\right)^{2p} \times\left(1-\frac{p+q}{n}\right)^{2(n-p-q-k)}\left(\frac{k+q}{n}\right)^{2q} \Pi_1 \Pi_2 \tag{2.6}$$

where

$$\Pi_1 = \Pr((2.2\text{c}) \text{ holds for a fixed } K, P, Q \mid (2.2\text{a}), (2.2\text{b}))$$

and

$$\Pi_2 = \Pr(2.2\text{d}) \text{ holds for a fixed } k, P, Q \mid (2.2\text{a}), (2.2\text{b}), (2.2\text{c})).$$

We can take K, P, Q as fixed in these definitions as we have taken the expectation over all possible set K, P, Q here.

In the construction of D_n we shall refer to each $v \in V_n$ *choosing* two neighbors at random.

Now

$$\Pi_1 \leqslant \sum_{t=q}^{2q} \binom{2q}{t}\left(\frac{q}{q+k}\right)^t\left(1-\frac{q}{q+k}\right)^{1q-t} \quad \text{for} \quad q>0.$$

To see this we have to consider the choice of neighbors for each $q \in Q$. One can see that for each $q \in Q$ and each choice of neighbor, the probability that the neighbor chosen is in Q (a *success*), given (2.2a), is $q/(q+k)$ regardless of any other choices made. Now if Q contains q or more edges then there must have been at least q successes. But, by the above remarks, the probability that there are at least q successes is given by the above binomial summation.

Now (2.5) implies that $q<k/2$ for $\varepsilon<\frac{1}{3}$ which implies

$$\Pi_1 \leqslant 2\binom{2q}{q}\left(\frac{q}{q+k}\right)^q\left(1-\frac{q}{q+k}\right)^q \leqslant\left(\frac{4kq}{(q+k)^2}\right)^q \quad \text{for} \quad q>0. \tag{2.7}$$

Now clearly

$$\Pi_2 \leqslant \Pr((2.2\text{d}) \mid (2.2\text{a}), (2.2\text{b}) \text{ and } Q \text{ is stable}). \tag{2.8}$$

It is now an immediate consequence of Lemma 2.5 (below) that

$$\Pi_2 \leqslant \Theta^k$$

where, for v any fixed vertex of K,

$$\Theta = \Pr(v \text{ is adjacent to at least one vertex of } P \cup Q)$$
$$= 1 - \left(1 - \frac{p+q}{n}\right)^2 \left(1 - \frac{1}{k}\right)^{2(p+q)}.$$

(To apply the lemma consider the vertex set K as a set of k boxes and a vertex v of K being chosen by a member of $P \cup Q$ as a ball falling into box v. Thus we take $m = 2(p+q)$, $a_i = 1$, and $Y_i =$ the number of choices made by i in $P \cup Q$ for $i \in K$.) Since $p+q \leqslant (1+\varepsilon)k \leqslant (1+\varepsilon)\, u(\varepsilon) n$ we have

$$\Theta \leqslant 1 - (1 - (1+\varepsilon)\, u(\varepsilon))^2 \left(1 - \frac{1}{k}\right)^{2(1+\varepsilon)k}.$$

Now taking $k \geqslant k_0$, see (2.4), we see that

$$\Pi_2 \leqslant \delta^k \qquad \text{for some } \delta < 1 \text{ when } k \geqslant k_0 \text{ and } \varepsilon \text{ is small.} \tag{2.9}$$

Using Stirling's formula and (2.7), (2.9) in (2.6) gives

$$\Pr(E_2(k, p, q)) \leqslant \left(\frac{k}{en}\right)^{p+q-k} \left(\frac{k}{p}\right)^p 4^q e^{2(p+q)(p+q+k)/n} \delta^k.$$

We can deduce from (2.2e) that (i) $p+q-k \geqslant 1$, (ii) $k/p \leqslant 1 + q/p$, and (iii) $q \leqslant 3\varepsilon k/2$ and so we can write, for $k \geqslant k_0$,

$$\Pr(E_2(k, p, q)) \leqslant \frac{k}{en} \cdot a^{\varepsilon k} e^{5k^2/n} \delta^k \qquad \text{for some } a > 0.$$

For $k \geqslant k_0$ let $S(k) = \{(p, q)\colon k, p, q \text{ satisfy (2.2e)}\}$ and note that $|S(k)| \leqslant 2k^2$ for small ε. Now choose ε small enough so that $a^\varepsilon e^{5u(\varepsilon)}\delta \leqslant \eta < 1$ and then

$$\sum_{k=k_0}^{u(\varepsilon)n} \sum_{(p,q) \in S(k)} \Pr(E_2(k, p, q)) \leqslant \frac{2}{en} \sum_{k=k_0}^{u(\varepsilon)n} k^2 \eta^k = O(1/n). \quad \blacksquare$$

By taking ε suitably small and k_0 suitably large, we can sum up what we have proved so far by: there is an absolute constant $u_0 > 0$ such that

$$\lim_{n \to \infty} \Pr(D_n \text{ contains a bad set } K, |K| \leqslant u_0 n) = 0.$$

This of course does depend on us proving

LEMMA 2.5. *Suppose r distinct balls are placed randomly in s boxes where $r \geqslant 0$ is a random variable. Let X_i, $i = 1, 2, \ldots, s$, be the number of balls placed in box i. Let Y_i, $i = 1, 2, \ldots, s$, be independent integer random variables in the range* 0, 1,..., L. *Let $a_1, a_2, \ldots, a_s \geqslant 0$ be given fixed integers and let E_i be the event $\{X_i + Y_i \geqslant a_i\}$. Then*

$$\Pr\left(\bigcap_{i=1}^{p} E_i\right) \leqslant \prod_{i=1}^{p} \Pr(E_i) \qquad \text{for} \quad p = 1, 2, \ldots, s. \tag{2.10}$$

Proof. Let A_i be the event $\{X_i \geqslant b_i\}$ for some integer b_i, $i = 1, 2, \ldots, s$. We first prove that

$$\Pr\left(\bigcap_{i=1}^{p} A_i\right) \leqslant \prod_{i=1}^{p} \Pr(A_i) \qquad \text{for} \quad i = 1, 2, \ldots, p. \tag{2.11}$$

We prove (2.11) by induction on p. Note that it is trivially true when $p = 1$. Let

$$\lambda_b = \Pr\left(\bigcap_{i=1}^{p} A_i \cap \{X_{p+1} = b\}\right)$$

and

$$\mu_b = \Pr(X_{p+1} = b) \qquad \text{for} \quad b = 0, 1, \ldots .$$

We note that

$$\lambda_b / \mu_b = \Pr\left(\bigcap_{i=1}^{p} A_i \,\middle|\, X_{p+1} = b\right) \qquad \text{provided } \mu_b > 0.$$

Hence, by considering re-directing one ball, we have

$$\lambda_b / \mu_b \leqslant \lambda_{b-1} / \mu_{b-1} \qquad \text{provided } \mu_b > 0.$$

Hence

$$\begin{aligned}
\Pr\left(\bigcap_{i=1}^{p} A_i \,\middle|\, X_{p+1} \geqslant l_{p+1}\right) &= \left(\sum_{l \geqslant l_{p+1}} \lambda_l\right) \Big/ \left(\sum_{l \geqslant l_{p+1}} \mu_l\right) \\
&\leqslant \left(\sum_{l \geqslant 0} \lambda_l\right) \Big/ \left(\sum_{l \geqslant 0} \mu_l\right) \\
&= \Pr\left(\bigcap_{i=1}^{p} A_i\right). \\
&\leqslant \prod_{l=1}^{p} \Pr(A_i)
\end{aligned}$$

which completes the inductive step.

Let now $\Omega = \{0, 1,..., L\}^n$ and $\mathbf{Y} = (Y_1, Y_2 \cdots Y_n)$. Then

$$\Pr\left(\bigcap_{i=1}^{p} E_i\right) = \sum_{\mathbf{y} \in \Omega} \Pr\left(\sum_{i=1}^{p} E_i \,\middle|\, \mathbf{Y} = \mathbf{y}\right) \Pr(\mathbf{Y} = \mathbf{y}).$$

But

$$\begin{aligned}\Pr\left(\bigcap_{i=1}^{p} E_i \,\middle|\, \mathbf{Y} = \mathbf{y}\right) &= \Pr(X_i \geqslant k_i - y_i,\ i = 1, 2,..., s)\\ &\leqslant \prod_{i=1}^{p} \Pr(X_i \geqslant k_i - y_i)\end{aligned}$$

by (2.11), assuming $\mathbf{y} = (y_1, y_2,..., y_n)$. Thus

$$\begin{aligned}\Pr\left(\bigcap_{i=1}^{p} E_i\right) &\leqslant \sum_{\mathbf{y} \in \Omega} \prod_{i=1}^{p} (\Pr(X_i \geqslant k_i - y_i) \Pr(Y_i = y_i))\\ &= \prod_{i=1}^{p} \left(\sum_{y_i=0}^{L} \Pr(X_i \geqslant k_i - y_i) \Pr(Y_i = y_i)\right)\\ &= \prod_{i=1}^{p} \Pr(E_i). \quad \blacksquare\end{aligned}$$

We complete the proof by showing that the existence of a large bad set a.s. implies the existence of a large stable set P with $|N(P)|$ of comparable size to $|P|$. We then show that this a.s. cannot happen.

LEMMA 2.6. *Let E_3 denote the following event*:

"*D_n contains at least* $(\log n)^3$ *sets $S \subseteq V_n$ satisfying*"

$$|S| \leqslant \log \log n; \tag{2.12a}$$

$$|E_S| \geqslant |S|. \tag{2.12b}$$

Then $\lim_{n \to \infty} \Pr(E_3) = 0$.

Proof. The expected number s_k of sets S with $|S| = k$ and $|E_S| \geqslant k$ satisfies

$$s_k \leqslant \binom{n}{k}\binom{2k}{k}\left(\frac{k}{n}\right)^k \leqslant (4e)^k.$$

Thus

$$\sum_{k=3}^{\lfloor \log\log \rfloor} s_k \leqslant 2(4e)^{\log\log n}.$$

The result follows from the Markov inequality. ∎

Let a pair of sets K, $P \subseteq V_n$ be *matched* if

(i) P is stable in D_n;

(ii) $N(P) = K$ and hence $K \cap P = \varnothing$;

(iii) $|P| \geqslant |K| - \delta(n)$

where $\delta(n) = \lceil n/\log\log n + \log n)^3 \rceil$.

LEMMA 2.7. *Suppose that D_n has no bad sets of size $u_0 n$ or less but D_n contains a bad set K, $|K| = k > u_0 n$. Suppose also that E_3 does not occur. Then there exists a matched pair K', P with $k - \delta(n) \leqslant |K'| \leqslant k$.*

Proof. We can assume that no proper subset of K is bad. Let $C_1, C_2, ..., C_s$, P be as defined in Lemma 2.3. We show first that $|P| \geqslant k - \delta(n)$. This is true as fewer than $n/\log\log n$ out of $C_{p+1}, ..., C_s$ can have size exceeding $\log\log n$ and (2.1b) implies that fewer than $(\log n)^3$ have size no more than $\log\log n$, assuming that E_3 does not occur. Let $K' = N(P) \subseteq K$. If $|K'| < |P|$ then K' is a bad set and the lemma follows. ∎

The final lemma that completes the proof of the theorem is

LEMMA 2.8. *Let $E_4(k)$ be the event*

"*There exists a matched pair K, P with $|K| = k$.*"

Let $k_0 = \lfloor u_0 n \rfloor - \delta(n)$. Then

$$\lim_{n\to\infty} \Pr\left(\bigcup_{k=k_0}^{n/2} E_4(k)\right) = 0.$$

Proof. Now

$$\Pr(E_4(k)) \leqslant \binom{n}{k} \sum_{p=k-\delta(n)}^{n-k} \binom{n-k}{n} \left(\frac{k}{n}\right)^{2p} \left(1 - \frac{p}{n}\right)^{2(n-k-p)} u(k, p)$$

where

$$u(k, p) = \Pr(N(P) = K \mid P \text{ is stable and } N(P) \subseteq K)$$

$$\leqslant \left(1 - \left(1 - \frac{1}{k+1}\right)^{2p} \left(1 - \frac{p}{n}\right)^2\right)^k \qquad \text{by Lemma 2.5.}$$

Using Stirling's inequalities and making simple approximations yield

$$\Pr(E_4(k)) \leqslant \sum_{p=k-\delta(n)}^{n-k} \left(\frac{k}{ne}\right)^{p-k} \left(\frac{k}{p}\right)^p e^{p^2/n} \left(1-\left(1-\frac{1}{k+1}\right)^{2p} \left(1-\frac{p}{n}\right)^2\right)^k. \tag{2.13}$$

Let u_p equal the summand in (2.13). For $p \geqslant 4k/3$ and $k \leqslant n/20$ we have

$$u_p \leqslant \left(\frac{1}{20e}\right)^{p/4} \left(\frac{3}{4}\right)^p e^p \leqslant (.97)^p. \tag{2.14}$$

For $p < 4k/3$ we have

$$u_p \geqslant v_p = \left(\frac{k}{ne}\right)^{p-k} \left(\frac{k}{p}\right)^p e^{p^2/n} \left(1-\left(1-\frac{1}{k+1}\right)^{8k/3} \left(1-\frac{4k}{3n}\right)^2\right)^k.$$

Now

$$v_{p+1}/v_p = (1+o(1))(k/e^2n)e^{2p/n} \tag{2.15}$$

and so (2.14) and (2.15) imply that if $k \leqslant n/20$

$$\Pr(E_4(k)) \leqslant e^{o(n)} v_k + n(.97)^k.$$

But

$$v_k = e^{k^2/n}(1+O(1/k) - e^{-8/3}\left(1-\frac{4k}{3n}\right)^2\Big)^k \leqslant (.99)^k$$

and it follows that

$$\Pr(E_4(k)) \leqslant \delta^k, \qquad k_0 \leqslant k \leqslant n/20 \tag{2.16}$$

where $\delta < 1$ is a constant.

We now wish to show that there exists a constant $\lambda_1 < \frac{1}{2}$ such that

$$\Pr\left(\bigcup_{k=k_1}^{n/2} E_4(k)\right) = o(1) \qquad \text{where} \quad k_1 = \lfloor \lambda_1 n \rfloor. \tag{2.17}$$

Let $\alpha(D_n)$ denote the size of the largest stable set in D_n. We prove the existence of $\lambda_2 < \frac{1}{2}$ such that

$$\Pr(\alpha(D_n) \geqslant \lfloor \lambda_2 n \rfloor = o(1). \tag{2.18}$$

Equation (2.17) then follows for any $\lambda_1 > \lambda_2$. Now for $k > n/20$

$$\Pr(\alpha(D_n) \geqslant k) \leqslant \binom{n}{k}\left(1-\frac{k}{n}\right)^{2k}$$

$$= e^{o(n)}((1-\lambda)^{3\lambda-1}\lambda^{-\lambda})^n \qquad \text{where} \quad k = \lambda n.$$

It follows that

$$\Pr(\alpha(D_n) > (\tfrac{1}{2}+\varepsilon)n) = o(1) \tag{2.19}$$

for any constant $\varepsilon > 0$.

A maximal stable set S is also a dominating set, i.e., each $w \in V_n - S$ is adjacent to at least one vertex of S. Hence, if $n/20 \leqslant k \leqslant 3n/4$, say,

$$\begin{aligned} &\Pr(\text{there exists a maximal stable set of size } k) \\ &\quad \leqslant \binom{n}{k}\left(1-\frac{k}{n}\right)^{2k} (1+O(1/n)-e^{-2k/(n-k)})^{n-k} \qquad \text{using Lemma 2.5} \\ &\quad = e^{o(n)} f(\lambda)^n \end{aligned}$$

where $f(\lambda) = (1-\lambda)^{3\lambda-1}\lambda^{-\lambda}(1-e^{-2\lambda/(1-\lambda)})^{1-\lambda}$. Now f is continuous for $\lambda < 1$ and $f(\frac{1}{2}) < 1$. Hence there exists $\varepsilon > 0$ such that

$$\Pr(\text{there exists a maximal stable set of size } k,\ (\tfrac{1}{2}-\varepsilon)n \leqslant k \leqslant (\tfrac{1}{2}+\varepsilon)n) = o(1).$$

This combined with (2.19) yields (2.18) and (2.17).

Next let a matched pair K, P be *maximal* if there does not exist $P' \supsetneqq P$ such that K, P' is a matched pair. We note first that

if D_n contains a matched pair K, P then D_n contains a maximal pair K, P' where $P' \supseteq P$. (2.20)

Furthermore

K, P is maximal implies that if $v \in R = V_n - K \cup P$ then there exists $w \in R$ such that $\{v, w\} \in E(D_n)$. (2.21)

(Otherwise K, $P \cup \{v\}$ is a matched pair.)

We show next that at least one of the three following events A, B, C occurs if $E_4(k)$ occurs with $k_1 \geqslant k > n/20$ and n is large.

A: "D_n contains a vertex of degree exceeding $\log n$";

B: $\bigcup_{k=k_0}^{n/20} E_4(k)$;

C: "D_n contains a maximal pair K, P, $n/20 < |K| \leqslant k_1$, and each vertex of K is adjacent to at least $|P| + \delta(n) - |K| + 2$ vertices of P."

Assume that $E_4(k)$ occurs with $k > n/20$ and neither A nor B occurs. Let $k' = \min\{k > n/20\colon E_4(k) \text{ occurs}\}$ and let K, P be a maximal pair with $|K| = k'$ (see (2.20)). Suppose now that there exists $v \in K$ such that

$$|W| \leqslant |P| + \delta(n) - |K| + 1 \qquad \text{where } W = \{w \in P\colon \{v, w\} \in E(D_n)\}.$$

Let $P' = P - W$ and $K' = N(P')$. If A does not occur then

$$\begin{aligned} |K'| &\geqslant |K| - |W| \log n \\ &\geqslant |K| - (\log n)^2 \\ &> n/20 - (\log n)^2. \end{aligned}$$

Furthermore

$$\begin{aligned} |P'| - |K'| &\geqslant (|P| - |W|) - (|K| - 1) \\ &\geqslant -\delta(n). \end{aligned}$$

Thus K', P' is a matched pair and so we have contradicted either the definition of k' or the fact that B does not occur.

If a vertex v has degree at least $\log n$ then at least $\log n - 2$ vertices have v as one of their choices. Thus

$$\Pr(A) \leqslant n \binom{n}{\lceil \log n \rceil - 2} \left(\frac{2}{n}\right)^{\lceil \log n \rceil - 2} = o(1).$$

Since $\Pr(B) = o(1)$, by (2.16), we have only to prove that $\Pr(C) = o(1)$. Thus let a maximal pair K, P be *extreme* if each vertex of K is adjacent to at least $|P| + \delta(n) - |K| + 2$ vertices of P. Now since there are at most $2(|K| + |P|)$ edges joining K and P we find that if K, P is extreme then

$$2(|K| + |P|) \geqslant (|P| + \delta(n) - |K| + 2)|K|.$$

Putting $t = |P| + \delta(n) - |K| \geqslant 0$ we obtain

$$t \leqslant (2 - 2\delta(n)/|K|)/(1 - 2/|K|).$$

For n large, this implies $0 \leqslant t \leqslant 2$, as t is integer. Now let

$E_5(k, t) =$ "D_n contains an extreme pair K, P with $|K| = k$ and $|P| = k - \delta(n) + t$."

We need only show that

$$\Pr\left(\bigcup_{k=k_0}^{k_1} \bigcup_{t=0}^{2} E_5(k, t) \right) = o(1).$$

Let us first consider the cases $t = 1, 2$. Then

$$\Pr(E_5(k, t)) \leqslant \frac{n!}{k!\, p!\, (n-k-p)!} \left(\frac{k}{p}\right)^{2p} \left(1 - \frac{p}{n}\right)^{2(n-k-p)} \pi_3(k, p)$$

where $p=k-\delta(n)+t$ and $\pi_r(k, p)=\Pr$(each vertex of a fixed k-set K is adjacent to at least r vertices of a fixed p-set P | each vertex of P makes both choices in K). Letting $k=\lambda n$ and applying Lemma 2.5 we find that for $t=1, 2$

$$\Pr(E_5(k, t)) \leqslant e^{o(n)} f(\lambda)^n$$

where

$$\begin{aligned} f(\lambda) &= \frac{(1-\lambda)^{2(1-2\lambda)}}{(1-2\lambda)^{1-2\lambda}} (1-e^{-2}(1+2(1-\lambda^2)+2(1-\lambda)^2))^{\lambda} \\ &\leqslant e^{\lambda^2}(1-3e^{-2})^{\lambda} \\ &< (.98)^{\lambda} \qquad \text{for} \quad 1/20 \leqslant \lambda \leqslant 1/2. \end{aligned}$$

It follows that

$$\Pr\left(\bigcup_{k=k_0}^{k_1} \bigcup_{t=1}^{2} E_5(k, t)\right) = o(1).$$

We are left with the events $E_5(k, 0)$ for $n/20 \leqslant k \leqslant k_1$. Now

$$\Pr(E_5(k, 0)) \leqslant \frac{n!}{k!p!(n-k-p)!} \left(\frac{k}{n}\right)^{2p} \left(1-\frac{p}{n}\right)^{2(n-k-p)} \pi_2(k, p)\, \mu(k, p)$$

where $p=k-\delta(n)$ and for fixed disjoint k-set K, p-set P, $R=V_n-K\cup P$,

> $\mu(k, p)=\Pr$(each $v\in R$ is adjacent to at least one other vertex of R | there are no $R-P$ edges).

Now, by Lemma 2.5,

$$\mu(k, p) \leqslant \sum_{s=0}^{n-k-p} \binom{n-k-p}{s} \left(\frac{k}{n-p}\right)^{2s} \left(1-\left(\frac{k}{n-p}\right)^2\right)^{n-k-p-s} \gamma^s. \tag{2.22}$$

In (2.22) we are summing over $s=|S|$ where S is the set of vertices of R which make both choices in K and

> $\gamma=\Pr$(some fixed vertex of S is chosen by at least one vertex in $R-S$ | the vertices in $R-S$ make at least one choice in R and no choices in P)

$$\begin{aligned} &= 1-\beta^{n-k-p-s} \\ &\leqslant 1-\beta^{n-k-p} \end{aligned}$$

where

$\beta = \Pr(u \in S$ is not chosen by $v \in R - S \mid v$ makes at least one choice in $R)$

$$= \frac{\left(\frac{n-k-p}{n-p}\right)^2\left(1-\left(\frac{1}{n-k-p}\right)\right)^2 + \frac{2k(n-k-p)}{(n-p)^2}\left(1-\frac{1}{n-k-p}\right)}{1-\left(\frac{k}{n-p}\right)^2}$$

$$= \left(1-\frac{1}{n-k-p}\right)\left(1-\frac{1}{n-p+k}\right).$$

Hence, after some manipulation,

$$\gamma \leqslant (1-e^{-2(1-k/n)})(1+O(\delta(n)/n)).$$

Putting $k=\lambda n$, $s=\mu n$ we see that for $k \leqslant k_1$,

$$\Pr(E_5(k,0)) \leqslant e^{o(n)} \sum_{s=0}^{n-2k+\delta(n)} \phi(\lambda,\mu)n$$

where

$$\phi(\lambda,\mu) = (1-x)^{\lambda}\left(\frac{\lambda^{2\mu}}{\mu^{\mu}}\right)\left(1+\frac{\mu}{1-2\lambda-\mu}\right)^{1-2\lambda-\mu}(1-e^{-2(1-\lambda)})^{\mu}$$

$$\text{and } x = e^{-2}((1-\lambda^2)+2(1-\lambda)^2).$$

We will thus be finished if we can show that there exists a constant $\eta < 1$ such that

$$\phi(\lambda,\mu) \leqslant \eta \qquad \text{for} \quad 1/20 \leqslant \lambda \leqslant \lambda_1 < 1/2,\ 0 \leqslant \mu \leqslant 1-2\lambda. \tag{2.23}$$

(We should really write $0 \leqslant \mu \leqslant 1-2\lambda+O(\delta(n)/n)$. But this will follow from (2.23), the continuity of ϕ, and the boundedness of the range for λ, μ.)

Differentiating ϕ with respect to μ shows that $\partial\phi/\partial\mu = 0$ if and only if

$$\mu = \mu(\lambda) = \frac{\lambda^2(1-2\lambda)(1-e^{-2(1-\lambda)})}{(1-2\lambda)+\lambda^2(1-e^{-2(1-\lambda)})}.$$

Now

$$\phi(\lambda,0) = (1-x)^{\lambda}$$

$$\phi(\lambda,\mu(\lambda)) = (1-x)^{\lambda}\left(1+\frac{\lambda^2(1-e^{-2(1-\lambda)})}{1-2\lambda}\right)^{(1-2\lambda)}$$

$$\phi(\lambda,(1-2\lambda)) = (1-x)^{\lambda}\left(\frac{\lambda^2(1-e^{-2(1-\lambda)})}{1-2\lambda}\right)^{(1-2\lambda)}$$

and so

$$\phi(\lambda, \mu) \leqslant \psi(\lambda) = \phi(\lambda, \mu(\lambda))$$
$$= (1-x)^{\lambda}\left(1 + \frac{\lambda^2(1-e^{-2(1-\lambda)})}{1-2\lambda}\right)^{(1-2\lambda)}$$

for $1/20 \leqslant \lambda \leqslant \lambda_1$. We have not been able to construct any simple analytical proofs that $\psi(\lambda) < 1$ in our range of interest. Instead we offer two simple computational proofs that the reader can check with the aid of a computer.

Computational Proof 1. It is easy to show, by crude estimations, that $\psi'(\lambda) < 10$ for $.05 \leqslant \lambda \leqslant .5$. Numerical computations yield $x_t = \psi(.05 + .0001t) < .996$ for $t = 0, 1, ..., 4500$. The mean value theorem then implies that $\psi(\lambda) < .997$ throughout our range.

Computational Proof 2. Suppose we have $.05 \leqslant \lambda_a < \lambda_b \leqslant \lambda_1$. Then simple monotonicity arguments show that within $[\lambda_a, \lambda_b]$

$$\psi(\lambda) \leqslant \phi(\lambda_a, \lambda_b; \lambda)$$
$$= (1 - e^{-2}(3 - 4\lambda_b + \lambda_b^2))^{\lambda}$$
$$\times (1 - 2\lambda_a + \lambda_a^2(1 - e^{-2(1-\lambda_a)}))/(1 - 2\lambda_b)^{1-2\lambda_b}.$$

Now $\Theta(\lambda_a, \lambda_b; \lambda) = \log \phi(\lambda_a \lambda_b; \lambda)$ is linear in λ. Hence

$$\max(\Theta(\lambda_a, \lambda_b; \lambda_a), \Theta(\lambda_a, \lambda_b; \lambda_b)) \leqslant -\varepsilon$$

implies $\psi(\lambda) \leqslant e^{-\varepsilon}$ for $\lambda_a \leqslant \lambda \leqslant \lambda_b$. We have only therefore to divide the interval $[.05, .5]$ into a sequence of intervals $[\mu_0, \mu_1]$, $[\mu_1, \mu_2], ..., [\mu_{p-1} \mu_p]$ where $\mu_0 = .05$ and $\mu_p = .5$ and check that both $\Theta(\mu_i, \mu_{i+1}; \mu_i)$ and $\Theta(\mu_i, \mu_{i+1}; \mu_{i+1})$ are strictly less than zero for $i = 0, 1, ..., p-1$. This works if we take $\mu_i = .05 + .002i$ for $i = 0, 1, ..., 225$.

Acknowledgments

I should particularly like to thank Geoffrey Grimmett for many useful conversations on this paper. I should also like to thank David Barton, Gerard Cornuéjols, and Linda Rass.

References

1. J. Edmonds, Paths, trees and flowers. *Canad. J. Math.* **17** (1965), 449–467.
2. T. I. Fenner and A. M. Frieze, On the connectivity of random m-orientable graphs and digraphs, *Combinatorica* **2** (1982), 347–359.

3. T. I. FENNER AND A. M. FRIEZE, On the existence of hamiltonian cycles in a class of random graphs, *Discrete Math.* **45** (1983), 301–305.
4. T. GALLAI, Über extreme Punkt- und Kantenmengen, *Ann. Univ. Sci. Budapest, Eötvös Sect. Math* **2** (1959), 133–138.
5. G. R. GRIMMETT, An exact threshold theorem for random graphs and the node-packing problem, *J. Combin. Theory Ser. B* **40** (1986).
6. G. R. GRIMMETT AND W. R. PULLEYBLANK, Random near-regular graphs and the node packing problem, to appear.
7. E. SHAMIR AND E. UPFAL, One-factor in random graphs based of vertex choice, *Discrete Math.* **41** (1982), 281–286.
8. W. TUTTE, The factorisation of linear graphs, *J. London Math. Soc.* **22** (1947), 107–111.
9. D. W. WALKUP, Matchings in random regular bipartite graphs, *Discrete Math.* **31** (1980), 59–64.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium