ON PERFECT MATCHINGS IN RANDOM BIPARTITE GRAPHS WITH MINIMUM DEGREE AT LEAST TWO

by

Alan Frieze
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

Research Report No. 90-86

June 1990

ON PERFECT MATCHINGS IN RANDOM BIPARTITE GRAPHS WITH MINIMUM DEGREE AT LEAST TWO

Alan Frieze*
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213
U.S.A.

June 1990

*Supported in part by NSF grant CCR–8900112.

§1. INTRODUCTION

The threshold for the existence of a perfect matching in a random graph was established early on by Erdös and Rényi [ER]. Basically one needs enough random edges to ensure that the minimum degree is one with non–zero probability. Bollobás and Frieze [BF] considered the problem of the existence of perfect matchings in random graphs with minimum degree at least 1. Thus, for a positive integer k, let $\mathscr{G}^{(k)}_{n,m}$ denote the set of graphs which have vertex set $[n] = \{1,2,\ldots,n\}$, m edges and minimum degree at least k. Let $G^{(k)}_{n,m}$ be selected uniformly from $\mathscr{G}^{(k)}_{n,m}$. They found that the probability that $G^{(1)}_{n,m}$ has a perfect matching tends to the probability that it has no pair of vertices of degree 1 which have a common neighbour. Thus in this case about $(n \ln n)/4$ edges were needed to ensure a perfect matching with probability tending to 1. This is about half the number of edges required in the unrestricted case. Although we did not deal specifically with bipartite graphs, the changes required to deal with bipartite graphs would not be too difficult. More recently Bollobás , Fenner and Frieze [BFF] have improved the analysis and extended it to deal with Hamilton cycles. Broadly the probability that $G^{(k)}_{n,m}$ has the property

$\mathscr{A}_k$: $\lfloor k/2 \rfloor$ *edge disjoint Hamilton cycles plus a further edge disjoint perfect matching if* k *is odd*

tends to the probability that there is no set of $k + 1$ vertices of degree k which have a common neighbour. This means that about $(n \ln n)/(2(k + 1))$ edges are needed to ensure the property $\mathscr{A}_k$.

Bollobás, Cooper, Fenner and Frieze [BCFF] have considered the above property in relation to $G^{(k+1)}_{n,m}$. Here we find that by insisting on one higher minimum degree we can show that only a *linear number* (cn) of edges are needed to ensure the occurrence of $\mathscr{A}_k$. The results are however not sharp, in the sense that we only prove them for c sufficiently large.

In this paper we shall prove a sharp result for the simplest case: bipartite graphs and $k = 1$. (Even the simplest case seems to require a lot of work though.)

For disjoint sets V,W and positive integer m we let $\mathscr{B}(V,W;m) = \{$bipartite graphs with vertices V,W and m edges$\}$, $\mathscr{B}(V,W;\, m;\, \delta \geq 2\} = \{G \in \mathscr{B}(V_1,V_2;m) : \delta(G) \geq 2\}$.

Let now $c \geq 2$ be a fixed constant, $m = cn$ and B_n be sampled uniformly from $\mathscr{B}([n],[n];\, m;\, \delta \geq 2)$:.

THEOREM

$$\lim_{n\to\infty} \Pr(B_n \textit{ has a perfect matching}) = 1$$

□

If $c = 2$ then B_n is necessarily 2–regular and so *always* has a perfect matching. We assume therefore from now on that $c > 2$. Note that the theorem could be sharpened by consideration of $c = 2-o(1)$. We do not consider this possibility here.

The structure of the paper is as follows. In Section 2 we describe a useful "martingale" inequality. In Section 3 we describe the first of two models which enable us to analyze B_n. In Section 4 we describe the second of these two models. Finally in Section 5 we use the two models to prove the theorem.

GENERIC CONSTANTS

The calculations of this paper involve a large number of constants. For most of them the exact value is unimportant. We will often represent them by a generic positive constant A. When used in a formula it merely asserts the existence of some constant. It is thus legitimate to replace 2A by A etc.

§2. A MARTINGALE INEQUALITY

The use of martingale inequalities has become commonplace in the study of random graphs since their introduction by Shamir and Spencer [SS]. See Bollobás [B2] or McDiarmid [McD] for surveys.

Suppose that we have a probability space $\Omega = \Pi_{i=1}^{k}\Omega_i$ and a random variable $Z = Z(U_1,U_2,\ldots,U_k)$ defined on it. We will first assume that the measure on Ω is equal to the product of separate measures on the Ω_i so that $U_1,U_2,\ldots,U_k$ are independent.

For $U,V \in \Omega$ let $d(U,V) = |\{i : U_i \neq V_i\}|$. The following inequality has proved remarkably effective: suppose that the random variable Z is such that

$$d(U,V) = 1 \quad \text{implies} \quad |Z(U) - Z(V)| \leq a$$

then

$$\Pr(|Z - E(Z)| \geq t) \leq 2 \exp\{\frac{-2t^2}{ka^2}\}, \text{ for any } t > 0. \tag{2.1}$$

We will also need to apply (2.1) in the case $k = m$, $U_1,U_2,\ldots,U_m$ are the edges of $G \in \mathscr{B}([n_1],[n_2];m)$ for some n_1,n_2. Thus there is the dependence $U_i \neq U_j$ for $i \neq j$. Also in our applications Z will not depend on the ordering of $U_1,U_2,\ldots,U_m$ and so $d(U,V) = |\{i : U_i \notin \{V_1,V_2,\ldots,V_m\}\}|$. With all of these provisos (2.1) still holds.

§3. BALLS IN BOXES MODEL

We will use the following notation from [BFF]. For natural numbers s and t, let $[s]^t$ be the set of all s^t sequences to length t with the terms taken from the set $[s]$. Consider $[s]^t$ as a probability space in which any two sequences are equally likely. The space $[s]^t$ has the following intuitive interpretation which we shall use in the sequel. Put t distinguishable balls, say $b_1,b_2,\ldots,b_t$, into s boxes, with probability $1/s$ of putting a particular ball into any of the boxes. Every arrangement corresponds to a sequence of length t: if b_j goes into the i'th box then set $x_j = i$. Then the sequence $(x_1,x_2,\ldots,x_t)$ is a random element of the space $[s]^t$.

The *degree* of a number i in a sequence $X = (x_1,x_2,\ldots,x_t) \in [s]^t$, denoted by $d_X(i)$, is the number of times i occurs in the sequence: $d_X(i) = |\{j : x_j = i\}|$. Thus $d_X(i)$ is the number of balls in the i'th box. The *minimal degree* of X is $\delta(X) = \min\{d_X(i) : \in [s]\}$. Similarly the *maximal degree* of X is $\Delta(X) = \max(d_X(i) : i \in [s]\}$. We denote the number of boxes with k balls by $\nu_X(k) = |\{i : d_X(i) = k\}|$ for $k = 0,1,2....$

Let $[s]^t_2 = \{X \in [s]^t : \delta(X) \geq 2\}$ and consider this as a probability space of equiprobable elements. This space is much less pleasant that $[s]^t$ but we can make use of it. We now define a random bipartite *multigraph* in terms of this space. Thus let $X = (x_1,x_2,...,x_m)$ and $Y = (y_1,y_2,...,y_m)$ denote a pair of independently chosen random members of $[n]^m_2$ and let $B(X,Y)$ be the multigraph with vertex bipartition V,W and edges $E = x_1y_1,x_2y_2,...x_my_m$. The following lemma explains our interest in $B(X,Y)$.

Lemma 3.1 *Conditional on $B(X,Y)$ being a simple graph, it is equally likely to be any member of $\mathscr{B}_n$.*

Proof We simply have to observe that each member of $\mathscr{B}_n$ arises from exactly $m!$ distinct pairs X,Y. □

We will use fairly accurate estimates of the degree sequence of B_n. To obtain these we will use some ideas of [BFF]. We need to consider the following construction: given $X \in [s]^t$ we define a sequence $\rho(X)$ of minimal degree 2 in two steps. First we let

$$U(X) = \{i \in [s] : d_X(i) \geq 2\} = \{i_1,i_2,...,i_\sigma,\ \sigma = \sigma(X)\}$$

where $i_1 < i_2 < ... < i_\sigma$. Omit the terms of X not belonging to U and replace i_r by r to obtain the *reduced* sequence $\rho(X)$. By construction $\rho(X) \in [\sigma]^\tau_2$ for some $\tau = \tau(X)$. The following lemma is proved in [BFF]:

Lemma 3.2 *Let $Y_1, Y_2 \in [n]^m_2$. Then, where X is chosen randomly from $[s]^t$,*

$$\Pr(\rho(X) = Y_1) = \Pr(\rho(X) = Y_2) \qquad \square$$

(It is easy to see that Y_1, Y_2 arise from the same number of X's.)

The lemma is of course vacuous when $n > s$ or $m > t$ but we are interested in the case where $s > t > m$ and the probability of $\sigma = n$, $\tau = m$ is *not too small.*

We let $\omega = n^{\frac{1}{2}} \ln n$ and then define $\tilde{c}, M_b, N_b$ by

$$c = \frac{\tilde{c} - \tilde{c}e^{-\tilde{c}}}{1 - (1 + \tilde{c})e^{-\tilde{c}}}$$

$$N_b = \lceil (n - a\omega)(1 - (1 + \tilde{c})e^{-\tilde{c}})^{-1} \rceil$$

$$M_b = \lceil \tilde{c}N_b - \omega) \rceil$$

where a is large and positive.

To justify the implicit definition of $\tilde{c}$ we observe that the function

$$\varphi(x) = \frac{x - xe^{-x}}{1 - (1 + x)e^{-x}}$$

vanishes at $x = 0$, tends to infinity with x and is strictly monotonic increasing for $x > 0$. This last remark can be justified as follows:

$$\varphi'(x) = \frac{e^{2x} + 1 - (2 + x^2)e^x}{(e^x - 1 - x)^2}$$

and the numerator of this expression is

$$\sum_{k=1}^{\infty} \frac{2^k - 2 - k(k - 1)}{k!} x^k > 0.$$

We will prove

Lemma 3.3 *There exist* $\hat{N}_b = N_b + O(\omega)$, $\hat{M}_b = M_b + O(\omega)$ *such that if* X *is chosen randomly from* $[\hat{N}_b]^{\hat{M}_b}$ *then*

$$\Pr(\sigma(X) = n \text{ and } \tau(X) = m) \geq \exp\{-An^{\frac{1}{2}}(\ln n)^2\}.$$

Proof We first estimate the expected sizes of $\sigma(Y)$, $\tau(Y)$ when Y is chosen randomly from $[N_b]^{M_b}$. Now since

$$(1 - \frac{1}{N_b})^{M_b} = e^{-\tilde{c}} + (e^{-\tilde{c}} + o(1))\frac{\omega}{N_b},$$

we have

$$E(\nu_Y(0)) = N_b(1 - \frac{1}{N_b})^{M_b} = N_b e^{-\tilde{c}} + (e^{-\tilde{c}} + o(1))\omega.$$

$$E(\nu_Y(1)) = N_b \frac{M_b}{N_b}(1 - \frac{1}{N_b})^{M_b - 1} = N_b \tilde{c} e^{-\tilde{c}} + (\tilde{c} - 1 + o(1))e^{-\tilde{c}}\omega,$$

and then where

$$A_\sigma = A_\sigma(\tilde{c}) = a + \tilde{c}e^{-\tilde{c}}, \text{ and } A_\tau = A_\tau(\tilde{c}) = (1 + (\tilde{c} - 1)e^{-\tilde{c}}),$$

$$E(\sigma(Y)) = n - (A_\sigma + o(1))\omega, \; E(\tau(Y)) = m - (A_\tau + o(1))\omega.$$

Changing a single Y_i can change $\sigma(Y)$ by at most 1 and $\tau(Y)$ by at most 2. So applying (2.1) successively with $Z = \sigma(Y)$, $\alpha = 1$, $t = b\omega$ (b small and positive) and then with $Z = \tau(Y)$, $\alpha = 2$, $t = b\omega$ we deduce that with probability $1 - o(1)$

$$|\sigma(Y) - (n - A_\sigma\omega)| \leq 2b\omega$$

$$|\tau(Y) - (m - A_\tau\omega)| \leq 2b\omega.$$

Hence there exists $\tilde{n}$, $\tilde{m}$ such that

$$|n - A_\sigma \omega - \tilde{n}| \;,\; |m - A_\tau \omega - \tilde{m}| \leq 2b\omega$$

and

$$\Pr(\sigma(Y) = \tilde{n}, \tau(Y) = \tilde{m}) \geq \omega^{-2}. \tag{3.1}$$

Let $k = n - \tilde{n}$ and $\ell = m - \tilde{m}$ and observe that if a is sufficiently large and b is sufficiently small then $A_\sigma, A_\tau > 0$ and

$$\frac{A_\tau + 2b}{A_\sigma - 2b} \geq \frac{\ell}{k} \geq \frac{A_\tau - 2b}{A_\sigma + 2b}$$

$$= \frac{1 + ca + (\tilde{c} - 1)e^{-\tilde{c}} - 2b}{a + \tilde{c}e^{-\tilde{c}} + 2b}$$

$$> 2.$$

Now let $\hat{N}_b = N_b + k$ and $\hat{M}_b = M_b + \ell$ and consider constructing $X \in [\hat{N}_b]^{\hat{M}_b}$ as follows:

(i) first choose Y randomly from $[N_b]^{M_b}$,

(ii) independently, for $i = 1,2,\ldots,M_b$, and with probability $\frac{k}{\hat{N}_b}$ replace Y_i with a random integer in $[N_b + 1, \hat{N}_b]$.

If $\hat{Y}$ denotes the transformed value of Y then clearly $\hat{Y}$ is a random member of $[\hat{N}_b]^{M_b}$.

(iii) Independently, for $i = M_b + 1,\ldots,\hat{M}_b$ randomly choose an integer value $\tilde{Y}_i$ in $[\hat{N}_b]$. At this stage $X = (\hat{Y}_1,\ldots\hat{Y}_{M_b}, \tilde{Y}_{M_b+1},\ldots\tilde{Y}_{\hat{M}_b})$ is a random member of $[\hat{N}_b]^{\hat{M}_b}$.

To complete the proof of the lemma we observe that $\sigma(Y) = n$, $\tau(Y) = m$, if

(a) $\sigma(Y) = \tilde{n}$, $\tau(Y) = \tilde{m}$,

(b) $Y = \hat{Y}$,

(c) $\tilde{Y}_i \in [N_b + 1, \hat{N}_b]$ for $i = M_b + 1,\ldots,\hat{M}_b$ and each $j \in [N_b + 1, \hat{N}_b]$ occurs at least twice as a $\tilde{Y}_i$.

The events in (a), (b), (c) are independent and

$$\Pr((a)) \geq \omega^{-2} \quad \text{by (3.1)},$$

$$\Pr((b)) = \left(1 - \frac{k}{\hat{N}_b}\right)^{M_b} \geq \exp\left\{-M_b\left(\frac{k}{\hat{N}_b} + \frac{k^2}{\hat{N}_b^2}\right)\right\} \geq e^{-2\tilde{c}A_\sigma\omega},$$

$$\Pr((c)) \geq \left(\frac{1}{\hat{N}_b}\right)^{\ell}$$

(since there is at least one way for the $\tilde{Y}_i$ to produce (c).)

Multiplying these 3 lower estimates of probability gives the lemma.

□

Remark. The lower bound in the above lemma is small but it will suffice for this paper. It can be improved to n^{-A} using the methods of [BFF] and the next section.

We can now give reasonably good estimates for the degree sequence of $B(X,Y)$ (and hence B_n, as we will see.) Let $\tilde{\theta} = 1 - e^{-\tilde{c}}(1 + \tilde{c})$.

Lemma 3.4 Let X be chosen randomly from $[n]_2^m$. Then with probability $1 - o(1)$

$$|\nu_k(X) - \frac{\tilde{c}^k e^{-\tilde{c}}}{k!\,\tilde{\theta}} n| \leq n^{4/5} \tag{3.2}$$

for $k = 2,3,\ldots \lceil \log n \rceil$.

PROOF Let $\hat{X}$ be chosen randomly from $[\hat{N}_b]^{\hat{M}_b}$. We have

$$E(\nu_k(\hat{X})) = \hat{N}_b \begin{bmatrix}\hat{M}_b\\k\end{bmatrix} \left[\frac{1}{\hat{N}_b}\right]^k (1-\frac{1}{\hat{N}_b})^{\hat{M}_b-k}$$

$$= \frac{\hat{N}_b}{k!}\left[\frac{\hat{M}_b}{\hat{N}_b}\right]^k e^{-\hat{M}_b/\hat{N}_b} \; (1+0(\frac{k^2}{\hat{N}_b}))$$

$$= \frac{n\tilde{c}^k}{k!\,\tilde{\theta}} e^{-\tilde{c}} \; (1+0(\frac{k\omega}{n})).$$

But changing one $\hat{X}_i$ can only change $\nu_k(\hat{X})$ by at most 1 and so we deduce from (2.1) with $t = n^{\cdot 79}$ that

$$\Pr(|\nu_k(\hat{X}) - E(\nu_k(\hat{X}))| \geq n^{\cdot 79}) \leq e^{-2n^{1\cdot 58}/\hat{M}_b}$$

$$\leq e^{-n^{\cdot 57}}.$$

Let now $\mathscr{E}$ denote the event $\{\sigma(\hat{X}) = n, \tau(\hat{X}) = m\}$. Then from Lemma 3.3 and our expression for $E(\nu_k(\hat{X}))$ we have

$$\Pr(|\nu_k(\hat{X}) - \frac{\tilde{c}^k e^{-\tilde{c}}}{k!\,\tilde{\theta}} n \mid \geq n^{4/5} | \mathscr{E}) \leq \exp\{-n^{\cdot 57} + An^{1/2}(\log n)^2\}.$$

But $\nu_k(\hat{X}) = \nu_k(\rho(\hat{X}))$ for $k \geq 2$ and given $\mathscr{E}$, we have by Lemma 3.2, that $\rho(\hat{X})$ is a random member of $[n]_2^m$.

□

We should of course interpret $\nu_k(X)$ as the number of vertices of degree k in one half of the partition of B(X,Y).

The above lemma deals with vertices of low degree. Our next lemma will show that with high probability there are no vertices of large degree.

For the following lemma X is chosen randomly from $[n]_2^m$ and Z is chosen randomly from $[n]^m$. Let

$$D_X(\ell) = \sum_{i=1}^{\ell} d_X(i) \tag{3.3}$$

for $\ell \geq 1$ and let $D_Z(\ell)$ be defined analogously.

Lemma 3.5 Suppose $\ell \leq (\log n)^2$ and $\lambda \geq 0$. Then there exists a constant $\alpha = \alpha(c)$ such that

$$\Pr(D_X(\ell) \geq \lambda) \leq \alpha^{\ell} \Pr(D_Z(\ell) \geq \lambda).$$

Proof

$$\Pr(D_X(\ell) \geq \lambda) = \Pr(D_Z(\ell) \geq \lambda \mid d_Z(i) \geq 2, 1 \leq i \leq n)$$

$$= \frac{\Pr(D_Z(\ell) \geq \lambda,\ d_Z(i) \geq 2,\ 1 \leq i \leq \ell \mid d_Z(i) \geq 2,\ \ell < i \leq n)}{\Pr(d_Z(i) \geq 2,\ 1 \leq i \leq \ell \mid d_Z(i) \geq 2,\ \ell < i \leq n)}$$

$$\leq \frac{\Pr(D_Z(\ell) \geq \lambda,\ d_Z(i) \geq 2,\ 1 \leq i \leq \ell)}{\Pr(d_Z(i) \geq 2,\ 1 \leq i \leq \ell \mid d_Z(i) \geq 2,\ \ell < i \leq n)}$$

by a simple monotonicity argument.

$$\leq \frac{\Pr(D_Z(\ell) \geq \lambda)}{\prod_{i=1}^{\ell} \Pr(d_Z(i) \geq 2 \mid d_Z(j) \geq 2,\ i < j \leq n)}$$

$$\leq \frac{\Pr(D_Z(\ell) \geq \lambda)}{\prod_{i=1}^{\ell} \Pr(d_Z(i) = 2 \mid d_Z(j) \geq 2,\ i < j \leq n)}.$$

The lemma will follow immediately from

Claim:

$$\Pr(d_Z(i) = 2 \mid d_Z(j) \geq 2, i < j \leq n) \geq \frac{\Pr(d_Z(i) \leq 2)}{1 + \frac{2}{c-2} + \frac{2}{(c-2)^2}}.$$

This is because

$$\Pr(d_Z(i) \leq 2) = \sum_{t=0}^{2} \binom{m}{t}\left(\frac{1}{n}\right)^t\left(1 - \frac{1}{n}\right)^{m-t}$$

$$= (1 - o(1))e^{-c}\left(1 + c + \frac{c^2}{2}\right).$$

Proof of Claim

Let $\Omega_t = \{Z \in [n]^m : d_Z(i) = t,\ d_Z(j) \geq 2 \text{ for } i < j \leq n\}$, $t = 0,1,2$. Fix $t = 0$ or 1. We estimate the ratio $|\Omega_{t+1}|/|\Omega_t|$. For $Z \in \Omega_t$, $Z' \in \Omega_{t+1}$ we write $Z \sim Z'$ if $|\{k : z_k \neq z'_k\}| = 1$ (so that necessarily we can find some k for which $z_k \neq i$ and put $z'_k = i$ and leave other components of Z unchanged.)

For $Z \in \Omega_t$ let $D(Z) = |\{Z' \in \Omega_{t+1} : Z \sim Z'\}|$ and let $D_{min} = \min\{D(Z) : Z \in \Omega_t\}$. Now, for $Z' \in \Omega_{t+1}$ $|\{Z \in \Omega_{t+1} : Z \sim Z'\}| = (t + 1)(n - 1)$. Thus

$$\frac{|\Omega_{t+1}|}{|\Omega_t|} \geq \frac{D_{min}}{(t+1)(n-1)} .$$

But $D_{min} \geq (c-2)n$ and this yields

$$\frac{|\Omega_2|}{|\Omega_0| + |\Omega_1| + |\Omega_2|} \geq \frac{1}{1 + \frac{2}{c-2} + \frac{2}{(c-2)^2}}$$

and the claim follows.

□

COROLLARY 3.6 If X is chosen randomly from $[n]_2^m$ then

$$\Pr(\Delta(X) \geq \frac{2 \log n}{\log\log n}) = o(1).$$

PROOF

$$\Pr(\Delta(X) \geq \frac{2 \log n}{\log\log n}) \leq n \Pr(D_X(1) \geq \frac{2 \log n}{\log\log n})$$

$$\leq n\alpha \Pr(D_Z(1) \geq \frac{2 \log n}{\log\log n})$$

(using the notation of Lemma 3.5)

$$\leq n\alpha \binom{m}{2 \log n / \log\log n} (\frac{1}{n})^{2 \log n/\log\log n}$$

$$\le n\alpha\left(\frac{me\ \log\log n}{2n\ \log n}\right)^{2\log n/\log\log n}$$

$$= o(1).$$

□

Let $\mathscr{A}$ denote any graph property. We wish now to show the way that we will use the results of this section. Let $\hat{X}$, $\hat{Y}$ be sampled independently from $[\hat{N}_b]^{\hat{M}b}$ and let RIGHTSIZES denote the event $\{\sigma(\hat{X}) = \sigma(\hat{Y}) = n,\ \tau(\hat{X}) = \tau(\hat{Y}) = m\}$. Also let SIMPLE denote the event $\{B(\rho\hat{X}),\rho(\hat{Y}))$ is simple$\}$ (this would seem to require us to extend our definition of $B(X,Y)$ to cases where X,Y have different lengths. It does, and any consistent extension will do, because we will always condition on RIGHTSIZES.)

By Lemma's 3.1 and 3.2,

$$\Pr(B_n \in \mathscr{A}) = \Pr(B(\rho(\hat{X}),\rho(\hat{Y})) \in \mathscr{A} \mid \text{RIGHTSIZES,SIMPLE})$$

$$= \frac{\Pr(B(\rho(\hat{X}),\rho(\hat{Y})) \in \mathscr{A},\ \text{SIMPLE} \mid \text{RIGHTSIZES})}{\Pr(\text{SIMPLE} \mid \text{RIGHTSIZES})}$$

$$\le \frac{\Pr(B(\rho(\hat{X}),\rho(\hat{Y})) \in \mathscr{A} \mid \text{RIGHTSIZES})}{\Pr(\text{SIMPLE} \mid \text{RIGHTSIZES})}$$

$$= \frac{\Pr(B(X,Y) \in \mathscr{A})}{\Pr(\text{SIMPLE} \mid \text{RIGHTSIZES})}$$

$$\le A\ \Pr(B(X,Y) \in \mathscr{A}) \tag{3.4}$$

for some $A = A(c)$, by the following Lemma.

Lemma 3.7

$$\Pr(\text{SIMPLE} \mid \text{RIGHTSIZES}) \ge A^{-1}.$$

Proof (outline) Instead of just conditioning on RIGHTSIZES we can condition on the sequences $d_X(i)$, $d_Y(i)$, $i = 1,2,\ldots,n$ or equivalently on X,Y up to a permutation of their coordinates. By Lemmas 3.4, 3.5 we may assume (3.2) and
$\Delta(X), \Delta(Y) \le \ell_0 = \lfloor 2\log n/\log\log n \rfloor$. Under these conditions

$$
\begin{aligned}
E(\#\text{ loops in } B(X,Y)) &= m\Pr(x_1 = y_1) && \text{by symmetry}\\
&= m\Pr(x_1 = 1) && \text{by symmetry}\\
&= \frac{m}{n} && \text{by symmetry}\\
&= c.
\end{aligned}
$$

$$E(\#\text{ pairs of repeated edges in } B(X,Y)) = \binom{m}{2}\Pr(x_1{=}x_2, y_1{=}y_2)$$

by symmetry

$$= \binom{m}{2}\Pr(x_1{=}x_2)\Pr(y_1{=}y_2).$$

But

$$\Pr(x_1 = x_2) = \sum_{i=1}^{n} \Pr(x_1 = x_2 = i)$$

$$= \sum_{i=1}^{n} \frac{d_X(i)\,(d_X(i) - 1)}{m(m - 1)}$$

and

$$\sum_{i=1}^{n} d_X(i)^2 = \sum_{k=2}^{\ell_0} k^2 \nu_k(X)$$

$$= n \sum_{k=2}^{\infty} k^2 \frac{\tilde{c}^k e^{-\tilde{c}}}{k!\,\tilde{\theta}} + 0(n^{4/5}\ell_0^2).$$

Thus the expected number of loops and pairs of repeated edges tends to a constant $\lambda = \lambda(c)$, say. This suggests that number of these objects is asymptotically Poisson and that

$$\Pr(B(X,Y) \text{ is simple}) \to e^{-\lambda}.$$

It is straightforward to verify this rigorously (e.g. Bollobás [B1], Theorem I.20) and we leave this to the reader.

□

§4. 2–CORE MODEL

In this section we will study the random graph $B_{n_1,n_2,p}$. This is a random bipartite graph with bipartition $V_1 = [n_1]$, $V_2 = [n_2]$ (disjoint copies) and in which each of the n_1n_2 possible edges is independently included with probability p.

A subgraph H of $B_{n_1,n_2,p}$ is defined to have $\nu_i(H)$ vertices in V_i, $i = 1,2$ (or V_i–vertices) and $\mu(H)$ edges. The **2–core** $\kappa_2(G)$ of a graph G is the largest subgraph of G with minimum degree 2. Its use in this paper stems from the following known fact:

Lemma 4.1 Let $H = \kappa_2(B_{n_1,n_2,p})$. Let $V_i' = V(H) \cap V_i$. Then conditional on V_1', V_2', $m' = \mu(H)$, H is equally likely to be any member of $\mathscr{B}(V_1', V_2'; m'; \delta \geq 2)$.

□

(For each $H \in \mathscr{B}(V_1', V_2'; m; \delta \geq 2)$ and $m \geq m'$ there are the same number of graphs in $\mathscr{B}(V_1, V_2; m)$ which have H as their 2–core. Furthermore, for a given m, each graph in

$\mathscr{B}(V_1,V_2;\ m)$ is equally likely to occur as $B_{n_1,n_2,p}$.)

The aim of this section is to prove that given $c > 2$ there exists $d = d(c) > 1$, $C_1 > 0$ and integers N,N_1,N_2 such that

$$N_1 \approx N_2 \approx N = An, \tag{4.1}$$

$$p = \frac{d}{N}, \tag{4.2}$$

and where $H = \kappa_2(B_{N_1,N_2,p})$,

$$\Pr(\nu_i(H) = n,\ i = 1,2 \text{ and } \mu(H) = m) \geq n^{-C_1}. \tag{4.3}$$

It follows immediately from Lemma 4.1 and (4.3) that for any graph property $\mathscr{A}$:

$$\Pr(B_n \in \mathscr{A}) \leq n^{C_1}\Pr(\kappa_2(B_{N_1,N_2,p}) \in \mathscr{A}). \tag{4.4}$$

We will need to get accurate estimates of the likely size of the 2–core of $B_{n_1,n_2,p}$. These are given in Lemma 4.12. This will take quite a while. The reader who would like to get on with the proof of the main theorem could skip now to §5. While the size of the 2–core of $G_{n,p}$ is known accurately (Pittel [P]), the bipartite case has been somewhat neglected. On the other hand the methods of proof are well established and we will try to be brief.

We now examine the component structure of $B_{n_1,n_2,p}$ where

$$n_i = \alpha_i n\ ,\quad i = 1,2\ , \tag{4.5a}$$

$$p = \frac{d}{n}, \tag{4.5b}$$

(4.5c) $\alpha_i = \alpha_i(n) \to \bar{\alpha}_i$ as $n \to \infty$, $i = 1, 2$,

(4.5d) $\bar{\alpha}_1\bar{\alpha}_2 d^2 \notin I_\epsilon = [1-\epsilon, 1+\epsilon]$ where $\epsilon > 0$ is fixed and small.

We will show that the dichotomy $d < 1, d > 1$ for $G_{n,p}$ is represented by $\bar{\alpha}_1\bar{\alpha}_2 d^2 < 1$, $\bar{\alpha}_1\bar{\alpha}_2 d^2 > 1$ in the bipartite case.

So first of all let

$\tau_{i,k}$ = the number of V_i–vertices in tree components of size k , (in $B_{n_1,n_2,p}$).

$\bar{\tau}_{i,k}$ = the number of V_i–vertices in non–tree components of size k,

and

$$\sigma_{i,k} = \tau_{i,k} + \bar{\tau}_{i,k},$$

$i = 1,2$ and $1 \le k \le |V_i|$.

Let

$$h(\alpha_1,\alpha_2,d) = \frac{1}{d}\sum_{k=1}^{\infty}\sum_{\ell=1}^{k-1}\frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!}(\alpha_1 d e^{-\alpha_2 d})^{\ell}(\alpha_2 d e^{-\alpha_1 d})^{k-\ell}.$$

Let $\omega \to \infty$ slowly, and $T_i = \sum_{k=1}^{\omega \log n} \tau_{i,k}$, $i = 1,2$.

Lemma 4.2

$$E(T_i) = nh(\alpha_i,\alpha_{3-i},d) + 0((\omega \log n)^2)$$

for $i = 1,2$.

Proof Assume w.l.o.g. that $i = 1$.

$$E(\tau_{1,k}) = \sum_{\ell=1}^{k} \begin{bmatrix} n_1 \\ \ell \end{bmatrix} \begin{bmatrix} n_2 \\ k-\ell \end{bmatrix} \ell^{k-\ell}(k-\ell)^{\ell-1}(\tfrac{d}{n})^{k-1}(1-\tfrac{d}{n})^{n_2\ell+n_1(k-\ell)+0(k^2)}$$

$$= \frac{n}{d}\sum_{\ell=1}^{k} \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!}(\alpha_1 d e^{-\alpha_2 d})^{\ell}(\alpha_2 d e^{-\alpha_1 d})^{k-\ell}(1+0(\tfrac{k^2}{n})).$$

The lemma will follow from

$$\text{REM} \le n^{-A\omega} \tag{4.6}$$

where

$$\text{REM} = \sum_{k=\omega\log n}^{\infty} \sum_{\ell=1}^{k-1} \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!}(\alpha_1 d e^{-\alpha_2 d})^{\ell}(\alpha_2 d e^{-\alpha_1 d})^{k-\ell}$$

$$\le \sum_{k=\omega\log n}^{\infty} \sum_{\ell=1}^{k-1} (\tfrac{\ell}{k-\ell})^{k-2\ell}(\alpha_1 d e^{1-\alpha_2 d})^{\ell}(\alpha_2 d e^{1-\alpha_1 d})^{k-\ell}$$

$$= \sum_{k=\omega\log n}^{\infty} \sum_{\ell=1}^{k-1} (\alpha_1 d(\tfrac{k-\ell}{\ell})e^{1-\alpha_1 d(\frac{k-\ell}{\ell})})^{\ell}(\alpha_2 d(\tfrac{\ell}{k-\ell})e^{1-\alpha_2 d(\frac{\ell}{k-\ell})})^{k-\ell}. \tag{4.7}$$

Now

$$\gamma_1 = \gamma_1(k,\ell) = \alpha_1 d\,\frac{k-\ell}{\ell} \in I_\theta \longleftrightarrow \frac{\ell}{k} \in J_1(\theta) = \left[\frac{\alpha_1 d}{1+\theta+\alpha_1 d}, \frac{\alpha_1 d}{1-\theta+\alpha_1 d}\right]$$

$$\gamma_2 = \gamma_2(k,\ell) = \alpha_2 d\,\frac{\ell}{k-\ell} \in I_\theta \longleftrightarrow \frac{\ell}{k} \in J_2(\theta) = \left[\frac{1-\theta}{1-\theta+\alpha_2 d}, \frac{1+\theta}{1+\theta+\alpha_2 d}\right]$$

and

$$J_1(\theta) \cap J_2(\theta) \neq \emptyset \longleftrightarrow (1-\theta)^2 \leq \alpha_1\alpha_2 d^2 \leq (1+\theta)^2.$$

Hence, assuming,

$$\epsilon > 3\theta \quad \text{we have} \quad J_1(\theta) \cap J_2(\theta) = \emptyset.$$

Furthermore

$$\eta = \min\left\{\frac{\alpha_1 d}{1+\theta+\alpha_1 d}, \frac{1-\theta}{1-\theta+\alpha_2 d}, 1-\frac{\alpha_1 d}{1+\theta+\alpha_1 d}, 1-\frac{1+\theta}{1+\theta+\alpha_2 d}\right\}$$
$$> 0$$

and outside I_θ, $\gamma e^{1-\gamma} \leq (1-\theta)e^{\theta} \leq e^{-\theta^2/2}$. Thus

$$\sum_{\ell=1}^{k-1} (\gamma_1 e^{-\gamma_1})^{\ell}(\gamma_2 e^{-\gamma_2})^{k-\ell} \leq \sum_{\ell \in J_1(\theta) \cup J_2(\theta)} e^{-\eta\theta^2 k/2} + \sum_{\ell \notin J_1(\theta) \cup J_2(\theta)} e^{-\theta^2 k}$$

$$\leq A\zeta^k \tag{4.8}$$

for some constant $\zeta < 1$.

Putting this upper bound into (4.7) yields (4.6) and the lemma.

□

When we consider small non–tree components we find, as in the model $G_{n,p}$, that there are few vertices on them.

We will say as in Knuth, Motwhani and Pittel [KMP] that an event $\mathscr{E}_n$ occurs *quite surely* (q.s.) if $\Pr(\mathscr{E}_n) = 1 - o(n^{-K})$ for any positive constant K.

Lemma 4.3 Let T_i = the number of V_i–vertices in non–tree components of size at most $\omega \log n$, i = 1,2. Then

(a) $$E(\bar{T}_i) = 0(1) \qquad i = 1,2.$$

(b) $$\bar{T}_i \le \omega(\log n)^2 n^{\frac{1}{2}} \qquad \text{q.s.}$$

PROOF

(a)

$$E(\bar{T}_1) \le \sum_{k=4}^{\omega \log n} \sum_{\ell=2}^{k-2} \begin{bmatrix} n_1 \\ \ell \end{bmatrix} \begin{bmatrix} n_2 \\ k-\ell \end{bmatrix} \ell^{k-\ell}(k-\ell)^{\ell-1} \sum_{t=0}^{k\ell} \begin{bmatrix} k\ell \\ t+1 \end{bmatrix} \left(\frac{d}{n}\right)^{k+t} \left(1 - \frac{d}{n}\right)^{n_2\ell+n_1(k-\ell)+0(k^2)}$$

$$\le A \sum_{k=4}^{\infty} \sum_{\ell=2}^{k-2} k\ell \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!} (\alpha_1 d e^{\alpha_2 d})^{\ell} (\alpha_2 d e^{-\alpha_1 d})^{k-\ell}$$

$$\le A \sum_{k=4}^{\infty} k\zeta^k \qquad \zeta \text{ as in (4.8)}$$

$$< \infty.$$

(b)

We use (2.1) here. Adding or deleting an edge changes T_1 by at most $2\omega \log n$. Take $t = \frac{1}{2}\,\omega(\log n)^2\, n^{\frac{1}{2}}$ to obtain the required result.

□

Letting Δ denote maximum degree it is easy to show by the first moment method that

LEMMA 4.4

$$\Delta(B_{n_1,n_2,p}) \le \log n \qquad \text{q.s.}$$

□

Let now $S_i = T_i + \bar{T}_i$ for $i = 1,2$.

Lemma 4.5

$$|S_i - nh(\alpha_i,\alpha_{3-i},d)| = 0(\omega n^{\frac{1}{2}}(\log n)^3) \text{ q.s.}$$

for $i = 1,2$.

Proof Assume w.l.o.g. that $i = 1$. We want to apply (2.1) with $Z = C_1$ but vertices of large degree will cause problems. We circumvent this problem with a simple idea from Frieze and McDiarmid [FM]. For an integer $u \geq \log n$ let B_u be obtained from $B = B_{n_1,n_2,p}$ as follows: go through V_1 in order and if vertex v has degree $d > u$ delete $d-u$ edges incident with v; to be specific delete the edges incident with its $d-u$ highest neighbours. (Thus the degrees of V_1–vertices in B_u are bounded by u, but there is no such bound for V_2–vertices.)

Now,

$$\text{(4.9)} \qquad \begin{aligned} \Pr(B_u \neq B) &= \Pr(\Delta(B) > u) \\ &\leq \alpha_1 n \binom{n_2}{u} \left(\frac{d}{n}\right)^u \\ &\leq \alpha_1 n \left(\frac{\alpha_2 de}{u}\right)^u. \end{aligned}$$

Next let U_i, $i = 1,2,\ldots,n_1$ denote the set of edges incident with $i \in V_1$ and

$$Z = Z(U_1,U_2,\ldots U_{n_1}) = S_1(B_u).$$

Changing any U_i, changes Z by at most $u\omega\log n$ and so (2.1) implies

$$\text{(4.10)} \qquad \Pr(|Z - E(Z)| \geq t) \leq \exp\{-\frac{At^2}{nu^2\omega^2(\log n)^2}\}.$$

Let now $u = t^{2/3}/(n^{1/3}\omega^{2/3}(\log n)^{2/3})$ so that

$$\Pr(B_u \neq B) \leq e^{-u}$$

(assuming $t \geq \omega n^{1/2}(\log n)^3$) and

$$|E(Z) - E(S_1(B))| \leq ne^{-u}.$$

Hence, using Lemmas 4.2 and 4.3 and (4.10),

$$\Pr(|S_1 - nh(\alpha_1,\alpha_2,d)| \geq t + ne^{-u} + 0((\omega \log n)^2)) \leq e^{-u} + e^{-Au}.$$

Hence

(4.10) $$\Pr(|S_1 - nh(\alpha_1,\alpha_2,d)| \geq t) \leq \exp\{-\frac{At^{2/3}}{n^{1/3}\omega^{2/3}(\log n)^{2/3}}\}$$

provided $t \geq \omega n^{1/2}(\log n)^3$.

(The above inequality can be strengthened, but is good enough as it stands.)

The lemma follows easily from (4.10).

□

We now "eliminate" components of size between $\omega \log n$ and n/ω (and a little more).

Let $\hat{\tau}_{i,k}$ = the number of V_i–vertices in trees of size k which have at most one vertex with neighbours outside the tree. (Thus these trees can either span a component or span a subgraph attached to the rest of the graph by a unique vertex.)

Lemma 4.6

$$\sum_{i=1}^{2} \sum_{k=\omega\log n}^{n/\omega} \hat{\tau}_{i,k} = 0 \quad \text{q.s.}$$

Proof Fix k.

$$E(\hat{\tau}_{1,k}) \leq \sum_{\ell=1}^{k} \begin{bmatrix} n_1 \\ \ell \end{bmatrix} \begin{bmatrix} n_2 \\ k-\ell \end{bmatrix} \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!\,(k-\ell)!} k(\tfrac{d}{n})^{k-1} (1-\tfrac{d}{n})^{n_2\ell+n_1(k-\ell)-n_1-n_2}$$

$$\leq ne^{o(k)} \sum_{\ell=1}^{k} \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!} (\alpha_1 de^{-\alpha_2 d})^{\ell} (\alpha_2 de^{-\alpha_1 d})^{k-\ell}$$

$$\leq ne^{o(k)} \zeta^k$$

as in (4.8) of Lemma 4.2. The lemma now follows.

□

We now concentrate for a short while on the case $\bar{\alpha}_1\bar{\alpha}_2 d^2 < 1$. First of all there are usually few vertices on cycles.

Lemma 4.7 If $\bar{\alpha}_1\bar{\alpha}_2 d^2 < 1$ then

$$\Pr(B_{n_1,n_2,p} \text{ has } \geq \omega \text{ vertices on cycles}) = o(1).$$

Proof

$$E\text{ (number of vertices on cycles)} \leq \sum_{k=2}^{\infty} \begin{bmatrix} \alpha_1 n \\ k \end{bmatrix} \begin{bmatrix} \alpha_2 n \\ k \end{bmatrix} (k!)^2 (\tfrac{d}{n})^{2k}$$

$$\leq 2 \sum_{k=2}^{\infty} k(\bar{\alpha}_1\bar{\alpha}_2 d^2)^k$$

$$= 0(1).$$

The Markov inequality can then be used to complete the proof.

□

It will now be easy to prove that $B_{n_1,n_2,p}$ has no large components, with probability $1-o(1)$.

LEMMA 4.8 If $\bar{\alpha}_1\bar{\alpha}_2 d^2 < 1$ then

$$\Pr(B_{n_1,n_2,p} \text{ has a component of size } \geq \frac{n}{\omega}) = o(1).$$

PROOF Suppose $B = B_{n_1,n_2,p}$ has such a component C and T is a spanning tree of C. We can, by Lemma 4.4, assume that $\Delta(B) \leq \log n$. Thus T contains a subtree T' of size between $\omega \log n$ and $\omega(\log n)^2$ which is attached to the rest of T by a single vertex. But then we may assume, by Lemma 4.6, that there is an edge not in T which joints a vertex v of T' to a vertex of $T-T'$. But then v is in a cycle of B. We can remove T' and apply the above argument to $T-T'$, showing that B has at least $n/(\omega\log n)^2$ vertices on cycles. Now apply Lemma 4.7.

□

Thus we can now deduce that if $\bar{\alpha}_1\bar{\alpha}_2 d^2 < 1$ then

$$\text{(4.11)} \qquad E(\#V_1\text{–vertices in components of size} \leq \omega \log n) = \bar{\alpha}_1 n - o(n)$$

$$= nh(\bar{\alpha}_1,\bar{\alpha}_2,d) + o(n). \qquad \text{(Lemma 4.5)}$$

Thus

$$\text{(4.12)} \qquad h(\alpha_1,\alpha_2,d) = \alpha_1 \quad \text{if } \alpha_1\alpha_2 d^2 < 1.$$

(We have dropped the bars over the α's. (4.11) implies (4.12) is true for constant α_1,α_2 and of course, for fixed n, α_1,α_2 are constants.)

Assume from now on that $\alpha_1\alpha_2 d^2 > 1$. We first prove

Lemma 4.9 There exist (unique) β_1,β_2 such that

(a)

(i) $\beta_1\beta_2 d^2 < 1$

(ii) $\beta_1 d e^{-\beta_2 d} = \alpha_1 d e^{-\alpha_2 d}$

(iii) $\beta_2 d e^{-\beta_1 d} = \alpha_2 d e^{-\alpha_1 d}$

(b)

$$h(\alpha_i,\alpha_{3-i},d) = \beta_i \qquad i = 1,2.$$

Proof Let $x_i = \alpha_i d$ and $\lambda_i = \alpha_i d e^{-\alpha_{3-i} d}$ for $i = 1,2$. Then

$$x_i = \lambda_i e^{x_{3-i}} \quad \text{for } i = 1,2 \tag{4.13}$$

and so

$$f_i(x_i) = 0 \quad \text{for } i = 1,2 \tag{4.14}$$

where

$$f_i(x) = x - \lambda_i e^{\lambda_{3-i} e^x} \quad \text{for } i = 1,2. \tag{4.15}$$

Furthermore

$$f_i'(x) = 1 - (\lambda_i e^{\lambda_{3-i} e^x})(\lambda_{3-i} e^x) \qquad \text{for } i = 1,2$$

and so $f_i'' < 0$ and f_i is strictly concave for $i = 1,2$. We consider f_1.

$$f_1'(x_1) = 1 - x_1 x_2 = 1 - \alpha_1\alpha_2 d^2 < 0. \tag{4.16}$$

Thus $f_1(x) = 0$ has 2 distinct roots, x_1 and a second $y_1 < x_1$ satisfying

$$f_1'(y_1) > 0.$$

Let

$$y_2 = \lambda_2 e^{y_1} \tag{4.17}$$

so that

$$\lambda_1 e^{y_2} = \lambda_1 e^{\lambda_2 e^{y_1}} \tag{4.18}$$

$$= y_1.$$

Now put $\beta_i = y_i/d$ for $i = 1,2$. (4.17) and (4.18) are equivalent to (aiii), (aii) respectively. But then, as for (4.16), $f_1'(y_1) = 1 - y_1 y_2 > 0$ and so (ai) holds.

The simplest check for uniqueness is from (b) (which will not involve a circular argument.)

(b)

$$h(\alpha_1,\alpha_2,d) = \frac{1}{d} \sum_{k=1}^{\infty} \sum_{\ell=1}^{k} \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!} (\alpha_1 d e^{-\alpha_2 d})^{\ell} (\alpha_2 d e^{-\alpha_1 d})^{k-\ell}$$

$$= \frac{1}{d} \sum_{k=1}^{\infty} \sum_{\ell=1}^{k} \frac{\ell^{k-\ell}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!} (\beta_1 d e^{-\beta_2 d})^{\ell} (\beta_2 d e^{-\beta_1 d})^{k-\ell}$$

$$= \beta_1.$$

by (4.12).

□

We may therefore re–express Lemma 4.5 as

$$(4.19) \qquad |S_i - n\beta_i| = O(\omega n^{\frac{1}{2}}(\log n)^3) \quad \text{q.s.}$$

Not surprisingly, when $\bar{\alpha}_1\bar{\alpha}_2 d^2 > 1$ there is **q.s.** a unique giant component of size exceeding $\omega \log n$.

Lemma 4.10 If $\alpha_1\alpha_2 d^2 > 1$ then $B_{n_1,n_2,p}$ has a unique component GIANT of size exceeding $\omega \log n$. GIANT contains $n(\alpha_i - \beta_i) + O(\omega n^{\frac{1}{2}}(\log n)^3)$ V_i–vertices **q.s.**

Proof Let $p' = n^{-\frac{1}{2}}$. Let $B_1 = B_{n_1,n_2,p(1-p')}$ and B_2 be obtained from B_1 by joining up non–adjacent vertices with probability $p'' = pp'/(1-p+pp')$. It is easily checked that B_2 has the same distribution as $B_{n_1,n_2,p}$ i.e. an edge is included in B_2 with probability

$$p(1-p') + (1-p+pp')p'' = p.$$

Note that pp' is sufficiently small that the conclusion of Lemma 4.6 is valid for B_1. Suppose now that B_2 has r components of size exceeding n/ω (note that they comprise approximately $n(\alpha_1 - \beta_1)$ vertices **q.s.**.) Each such component of B_2 **q.s.** contains at least one component of B_2 of size at least $\sqrt{n}/\omega^2$ (deleting $\leq \omega\sqrt{n}$ edges from a component splits it into at most $\omega\sqrt{n}$ subcomponents.) But then, from Lemma 4.6, these B_1 components are of size at least n/ω **q.s.** But then the extra edges in $B_2 - B_1$ **q.s.** connect these $\leq \omega$ components together. Thus $r = 1$ **q.s.**.

□

We now estimate the number of edges in the small components. So let $m_\sigma = m_\sigma(n_1,n_2,p) =$ the number of edges of $B_{n_1,n_2,p}$ in trees of size at most $\omega \log n$.

Lemma 4.11

$$|m_\sigma - \beta_1\beta_2 n| = 0(n^{\frac{1}{2}}\omega(\log n)^2) \qquad \text{q.s.}$$

PROOF

$$E(m_\sigma) = \sum_{k=2}^{\omega \log n} (k-1) \sum_{\ell-1}^{k-1} \begin{bmatrix} n_1 \\ \ell \end{bmatrix} \begin{bmatrix} n_2 \\ k-\ell \end{bmatrix} \ell^{k-\ell-1}(k-1)^{\ell-1}\left(\frac{d}{n}\right)^{k-1}\left(1-\frac{d}{n}\right)^{n_2\ell+n_1k+0(k^2)}$$

$$= \frac{n}{d}\sum_{k=2}^{\infty}(k-1)\sum_{\ell=1}^{k}\frac{\ell^{k-\ell-1}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!}\left(\alpha_1 de^{-\alpha_2 d}\right)^{\ell}\left(\alpha_2 de^{-\alpha_1 d}\right)^{k-\ell} + 0(\omega^2(\log n)^2)$$

(4.20)
$$= \frac{n}{d}\sum_{k=2}^{\infty}(k-1)\sum_{\ell=1}^{k}\frac{\ell^{k-\ell-1}(k-\ell)^{\ell-1}}{\ell!(k-\ell)!}\left(\beta_1 de^{-\beta_2 d}\right)^{\ell}\left(\beta_2 de^{-\beta_1 d}\right)^{k-\ell} + 0(\omega^2(\log n)^2)$$

$$= \beta_1\beta_2 dn + 0(\omega^2(\log n)^2).$$

The last equation can be justified as follows: the R.H.S. of (4.20) is $E(m_\sigma(\beta_1 n,\beta_2 n,p)) = \beta_1\beta_2 dn - o(n)$, on using Lemmas 4.3, 4.4., 4.6 and 4.8. It then follows as in (4.11) and (4.12) that the summation in (4.20) is actually $\beta_1\beta_2 dn$.

The concentration result follows from (2.1). Indeed, adding or deleting an edge changes m_σ by at most $2\omega \log n$. Putting $t = n^{\frac{1}{2}}\, \omega(\log n)^2$ yields the lemma.

□

Having established the size of GIANT we will not study its 2–core, CORE say. Let MANTLE denote those vertices in GIANT but not CORE. We will estimate the size of MANTLE by using a simple and powerful idea of Pittel [P]. Observe that

(4.21) $v \in$ MANTLE $\longleftrightarrow$ v is jointed to the largest component of $B - v$ by a single edge.

Here $B = B_{n_1,n_2,p}$ and $B - v$ is obtained by deleting vertex v. Note that (4.20) only describes MANTLE **q.s.**

Furthermore

(4.22) $$v \in \text{CORE} \longleftrightarrow v \text{ is joined to the largest component of } B - v \text{ by at least 2 edges.}$$

Now let $n_{\text{CORE}i}$, $i = 1,2$, m_{CORE} denote the number of V_i–vertices, $i = 1,2$ and edges in the 2–core of B respectively. (4.21) and (4.22) will now be used to show

Lemma 4.12

(a) $|n_{\text{CORE}i} - n(\alpha_i - \beta_i - \beta_i d(\alpha_{3-i} - \beta_{3-i}))| \le n^{19/20}$ **q.s.**

(b) $|m_{\text{CORE}} - n(\alpha_1 - \beta_1)(\alpha_2 - \beta_2)| \le n^{19/20}$ **q.s.**

Proof Let $X = |\text{MANTLE} \cap V_1|$ and $k = \lfloor n^{1/10} \rfloor$. Then

(4.23) $$E((X)_k) = (n_1)_k \Pr([k] \subseteq \text{MANTLE} \cap V_1)$$

where $[k] \subseteq V_1$.

Now by considering the number of edges, t say, contained in $[k]$,

(4.24) $$\Pr([k] \subseteq \text{MANTLE}) \le \sum_{t=0}^{k-1} \binom{K}{t} \left(\frac{d}{n}\right)^t \Pr(\mathscr{E}_{k-t}) + \sum_{t=k}^{K} \binom{K}{t} \left(\frac{d}{n}\right)^t + \Pr(\bar{\mathscr{A}}),$$

where $K = \binom{k}{2}$ and

$$\mathscr{E}_{k-t} = \{\text{vertex } i,\ 1 \le i \le k-t, \text{ is joined by exactly one edge to the largest component of } B - [k]\}$$

and

$\mathscr{A} = \{B - [k]$ has a unique giant component in the sense of Lemma 4.10$\}$.

Now fix $1 \leq s \leq k$. Then

$$\Pr(\mathscr{E}_s) = E_\nu((\frac{\nu d}{n}(1 - \frac{d}{n})^{\nu-1})^s)$$

where ν denotes the number of V_2–vertices in the giant component of $B - [k]$,

$$= E_\nu((\frac{\nu d}{n} e^{-\frac{\nu d}{n}})^s)(1 + 0(\frac{s}{n})).$$

Now let $\nu_2 = n(\alpha_2 - \beta_2)$. Then

$$\Pr(\mathscr{E}_s) \leq (1 + o(1))E((\frac{\nu d}{n} e^{-\frac{\nu d}{n}})^s \mid |\nu - \nu_2| \leq n^{4/5}) + \Pr(|\nu - \nu_2| \geq n^{4/5})$$

$$\leq (1 + o(1))(\frac{\nu_2 d}{n} e^{-\frac{\nu_2 d}{n}})^s (1 + 0(\frac{s}{n^{1/5}})) + \exp\{-\frac{An^{1/5}}{\omega^{2/3}(\log n)^{2/3}}\}$$

by (4.10),

$$\leq (1 + o(1))\,\theta^s,$$

where $\theta = \frac{\nu_2 d}{n} e^{-\frac{\nu_2 d}{n}} = (\alpha_2 - \beta_2)de^{-(\alpha_2 - \beta_2)d} = \beta_1(\alpha_2 - \beta_2)d/\alpha_1$.

Hence from (4.23) and (4.24)

$$E((X)_k) \leq (1 + o(1))\,(n_1)_k\,\theta^k.$$

But for any $a > 0$

$$\Pr(X \geq a) \leq \frac{E((X)_k)}{(a)_k}.$$

Putting $a = n_1\theta + \frac{1}{3}n^{19/20}$ and observing that $(\theta n_1 + (1-\theta)k)/a \geq \theta(n_1 - t)/(a-t)$ for $t < k$ we find

$$\Pr(X \geq \beta_1(\alpha_2-\beta_2)dn + \tfrac{1}{3}n^{19/20}) \leq (1 + o(1))(1 - \frac{a - \theta n_1 - (1-\theta)k}{a})^k \tag{4.25}$$

$$\leq \exp\{-An^{1/20}\}.$$

By using a similar argument we obtain, where n'_{CORE1} is the number of V_1–vertices of CORE,

$$\Pr(n'_{CORE1} \geq (((\alpha_1 - \beta_1) - \beta_1 d(\alpha_2 - \beta_2))n + \tfrac{1}{3}n^{19/20}) \leq \exp\{-An^{1/20}\}. \tag{4.26}$$

Indeed the main step is to argue, in place of (4.24), that

$$\Pr([k] \subseteq \text{CORE}) \leq \sum_{t=0}^{\frac{1}{2}k} \binom{K}{t}\left(\frac{d}{n}\right)^t \Pr(\hat{\mathscr{E}}_{k-2t}) + \sum_{t=\frac{1}{2}k}^{K} \binom{K}{t}\left(\frac{d}{n}\right)^t + \Pr(\bar{\mathscr{A}})$$

where

$$\hat{\mathscr{E}}_s = \{\text{vertex } i,\ 1 \leq i \leq s, \text{ is joined by at least 2 edges to the largest component of } B - [k]\}.$$

A similar argument to that given for $\mathscr{E}_s$ yields

$$\Pr(\hat{\mathscr{E}}_s) \leq (1 + o(1))(1 - e^{-\frac{\nu_2 d}{n}} - \frac{\nu_2 d}{n} e^{-\frac{\nu_2 d}{n}})^s$$

and (4.26) can be easily obtained.

We can now finish the proof of the lemma quite easily.

(a)

(4.26) plus the bound on T_1 in Lemma 4.3(b) for the part of the 2–core outside of GIANT provides a probabilistic upper bound. On the other hand

$$n_1 - n_{CORE1} \leq |MANTLE \cap V_1| + S \quad \text{q.s.}$$

and the RHS of the above is bounded probabilistically by (4.25) and (4.19) and Lemma 4.3(b).

(b) Let $\tilde{m}$ ($= \alpha_1\alpha_2 dn + o(n^{1/2} \log n)$ q.s.) denote the number of edges of $B_{n_1,n_2,p}$. Then

$$\tilde{m} - |MANTLE| - m_\sigma - \bar{m}_\sigma \leq m_{CORE} \leq \tilde{m} - |MANTLE| - m_\sigma \quad \text{q.s.}$$

where $\bar{m}_\sigma$ is the number of edges in non–tree components of size at most $\omega \log n$.

Now $\bar{m}_\sigma \leq (T_1 + T_2)\Delta$ and so Lemmas 4.3 and 4.4 can be used to show that $\bar{m}_\sigma$ is "negligible" q.s.. Finally $|MANTLE| = |V(GIANT)| - (n'_{CORE1} + n'_{CORE2})$ and we can use Lemma 4.10 and (a) to bound $|MANTLE|$ probabilistically from above.

□

For each vertex v in the 2–core there is a unique maximal tree T_v containing v and those vertices $w \in MANTLE$ for which all paths joining w to the 2–core go through v.

Let $\mu_k = |\{v \in V_i : T_v \text{ has } k \text{ vertices}\}|$.

Lemma 4.13

(a) There exists $A_0 > 0$ such that

$$\Pr(\exists k \geq A_0 \log n : \mu_k > 0) \leq n^{-10}$$

(b)

$$\mu_k \leq n\hat{\zeta}^k + n^{1/5} \quad \text{q.s.}$$

where $\hat{\zeta} = \frac{1+\zeta}{2}$, ζ as in (4.8) and $1 \leq k \leq A_0 \log n$.

Proof

(a) Fix k and let Z = the number of vertices in trees T_V with k vertices. Then $Z \leq \hat{\tau}_{1,k} + \hat{\tau}_{2,k}$ (of Lemma 4.6.) Hence, from the proof of that lemma

$$\begin{aligned} \Pr(\mu_k > 0) &\leq \Pr(Z > 0) \\ &\leq E(Z) \\ &\leq 2\, ne^{0(k)}\, \zeta^k \end{aligned}$$

We can then choose A_0 such that

$$1 + A_0 \log \zeta \leq -12.$$

(b) Fix $k \leq A_1 \log n$ and let Y = the number of maximal trees of size k with at most one vertex with neighbours outside the tree. Clearly $\mu_k \leq Y$. Let $t = \lfloor n^{1/20} \rfloor$. Then

$$\begin{aligned} E((Y)_t) &\leq \sum_{1 \leq \ell_1, \ldots, \ell_t \leq k} \prod_{i=1}^{t} \left\{ \binom{n_1}{\ell_i} \binom{n_2}{k-\ell_i} \frac{\ell_i^{k-\ell_i-1} (k-\ell_i)^{\ell_i - 1}}{\ell_i! (k-\ell_i)!} \left(\frac{d}{n}\right)^{k-1} \left(1 - \frac{d}{n}\right)^{n_2 \ell_i + n_1(k-\ell_i) - n_1 - n_2 - 0(tk)} \right\} \\ &\leq k^t (ne^{o(k)} \zeta^k)^t \\ &= (ne^{o(k)} \zeta^k)^t. \end{aligned}$$

Hence

$$\Pr(Y \geq n\hat{\zeta}^k) \leq \frac{(ne^{o(k)}\zeta^k)^t}{(n\hat{\zeta}^k)_t}. \tag{4.27}$$

Suppose first that $k \leq k_0 = \lfloor -\frac{5}{6}\log n/\log \hat{\zeta} \rfloor$ so that $n\hat{\zeta}^k \geq n^{1/6}$. The result in this case follows easily from (4.27). For $k > k_0$ we use

$$\begin{aligned} \Pr(Y \geq n^{1/5}) &\leq \frac{(ne^{o(k)}\zeta^k)^t}{(n^{1/5})_t} \\ &\leq \frac{(n\hat{\zeta}^k)^t}{(n^{1/5}-t)^t} \\ &\leq (2n^{-1/30})^t. \end{aligned}$$

□

We now come to the second half of our proof of (4.1). In this part we consider a 3–dimensional array of random bipartite graphs defined as follows: from values N,d (to be defined later) we define a random bipartite graph $\Gamma_{\nu,\nu,\mu}$ with vertex partition $X = [\nu]$ and $Y = [\nu]$ where

$$\nu = \lceil N + N^{29/30} \rceil \text{ and } \mu = \lceil dN + N^{29/30} \rceil.$$

$$\nu' = \lceil N - N^{29/30} \rceil \text{ and } \mu' = \lceil dN - N^{29/30} \rceil.$$

For each $x \in X$, $y \in Y$ and $1 \leq i \leq \mu$ there exists an edge of colour i independently with probability $\frac{1}{N^2}$.

Once all the edges of $\Gamma_{\nu,\nu,\mu}$ are generated, parallel edges are coalesced to a single edge.

Let

$$J = \{(i,j,k) : \nu' \leq i,j \leq \nu,\ \mu' \leq k \leq \mu\}$$

and for $(i,j,k) \in J$ let $\Gamma_{i,j,k}$ be the subgraph of $\Gamma_{\nu_1,\nu_2,\mu}$ induced by the vertices of $[i] \subseteq X$, $[j] \subseteq Y$ and the edges of colour $t \in [k]$. Note that for a fixed i,j,k, $\Gamma_{i,j,k}$ is distributed as B_{i,j,p_k} where $p_k = 1-(1-\frac{1}{N^2})^k \approx \frac{k}{N^2}$, for $k = 0(N)$. Let the 2–core of $\Gamma_{i,j,k}$ have $\xi_1 = \xi_1(i,j,k)$ vertices from X, $\xi_2 = \xi_2(i,j,k)$ vertices from Y and $\xi_3 = \xi_3(i,j,k)$ edges. The aim is to show that for suitably chosen N,d

$$\exists(i,j,k) \in J : \Pr(\xi_1 = \xi_2 = n,\ \xi_3 = \lceil cn \rceil) \geq n^{-A} \tag{4.28}$$

which will imply (4.1).

We first have to show that ξ does not change by much for a unit change of i or j or k.

If $\sigma = (i,j,k) \in J$ and $t = 1,2,3$ then

$$\sigma_{+t} = (i + \delta_{it},\ j + \delta_{jt},\ k + \delta_{kt}) \qquad \text{(Kronecker delta.)}$$

Let

$$\zeta_{s,t} = \max\{\sigma \in J : \xi_s(\sigma_{+t}) - \xi_s(\sigma)\} \quad 1 \leq s,t \leq 3$$

and

$$\zeta = \max\{\zeta_{s,t} : 1 \leq s,t \leq 3\}.$$

Lemma 4.14 There exists $\rho_1 > 0$ such that

$$\Pr(\zeta > \rho_1 \log N) \leq N^{-10}.$$

Proof Consider first $\zeta_{1,1}$. The construction of $\Gamma_{i+1,j,k}$ from $\Gamma_{i,j,k}$ can be viewed as that of adding vertex $i+1$ and then independently adding edges $(i+1,y)$, $y \in [j] \subseteq Y$ with probability

$p \le \frac{k}{N^2}$.

For each such y let

$$\tau_y = \begin{cases} 0 \text{ if } y \notin \text{MANTLE } (\Gamma_{i',j',k'}) \\ \text{size of maximal tree of mantle containing } y, \text{ otherwise.} \end{cases}$$

Now conditional on the values of τ_y, $\zeta_{1,1}$ is dominated stochastically by

$$Z = \sum_{t=1}^{j} \eta_t$$

where

$$\eta_t = \begin{cases} \tau_t & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

and $\eta_1,\ldots,\eta_j$ are independent ($\eta_t \ge$ the number of vertices added to the 2–core if edge (i+1,t) exists in $\Gamma_{i+1,j,k}$.)

Now q.s. at most logN η_t's will be non–zero and so we can consider an alternative random variable

$$Z = Z_1 + Z_2 + \ldots + Z_{\log N}$$

where $Z_1.Z_2,\ldots,Z_{\log N}$ are independent identically distributed random variables with

$$\Pr(Z_1 = s) = \frac{\gamma_s}{j}$$

where $\gamma_s = |\{y : \tau_y = s\}|$.

Observe that $\Pr(Z_1 = 0)$ is bounded below by a constant. Now Lemma 4.13 implies that q.s.

$$\gamma_s \leq N\theta^s + N^{1/5}$$

for some $0 < \theta < 1$, and

$$\gamma_s = 0 \qquad s \geq A_0 \log N.$$

Hence we can assume

$$\Pr(Z_1 = s) \leq 2\theta^s + 2N^{-4/5} \quad 0 \leq s \leq A_0 \log N.$$

It follows that we can dominate Z_1 stochastically by $Z_1' + Z_1''$ where

$$\Pr(Z_1' = s) = \Pr(Z_1 = s) - 2N^{-4/5} \qquad 1 \leq s \leq A_0 \log N$$

$$\leq 2\theta^s$$

$$\Pr(Z_1' = 0) = 1 - \sum_{s=1}^{A_0 \log N} \Pr(Z_1' = s),$$

and

$$Z_1'' = \begin{cases} A_0 \log N & \text{with probability } \dfrac{(\log N)^2}{N^{4/5}} \\ 0 & \text{with probability } 1 - \dfrac{(\log N)^2}{N^{4/5}} \end{cases}.$$

$Z_1' + Z_1''$ dominates Z_1 since, for $s \geq 1$

$$\Pr(Z_1' + Z_1'' \geq s) = \frac{(\log N)^2}{N^{4/5}} + \left(1 - \frac{(\log N)^2}{N^{4/5}}\right) \Pr(Z_1' \geq s)$$

$$\geq \frac{(\log N)^2}{N^{4/5}} + \left(1 - \frac{(\log N)^2}{N^{4/5}}\right) \left(\Pr(Z_1 \geq s) - \frac{2A_0 \log N}{N^{4/5}}\right)$$

$$\geq \Pr(Z_1 \geq s) + \frac{1}{N^{4/5}}((\log N)^2 \Pr(Z_1 < s) - 2A_0 \log N)$$

$$\geq \Pr(Z_1 \geq s)$$

since $\Pr(Z_1 < s)$ is bounded away from zero by a constant.

Now

$$\Pr\left(\sum_{s=1}^{A_0 \log N} Z_s'' \geq 20A_0 \log N\right) \leq \begin{bmatrix} A_0 \log N \\ 20 \end{bmatrix} \left[\frac{(\log N)^2}{N^{4/5}}\right]^{20}$$

$$\leq N^{-(16-o(1))}.$$

$\sum_{s=1}^{A_0 \log N} Z_s'$ can be dealt with in a standard manner by applying the Markov inequality to its moment generating function. This deals with $\zeta_{1,1}$, $\zeta_{2,2}$ and (more or less) with $\zeta_{1,2}$ and $\zeta_{2,1}$.

The number of edges added to the 2–core by adding a vertex $i+1$ is at most the number of vertices added plus the degree of $i+1$, which is at most $\log N$ **q.s.** and this deals with $\zeta_{3,1}$, $\zeta_{3,2}$.

Finally, the number of vertices or edges added to the 2–core by adding an edge is at most twice the size of the largest tree in the mantle, plus 1.

□

We now consider a function $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$, $\Phi = (\phi_1, \phi_2, \phi_3)$ where

$$\phi_1(\alpha_1.\alpha_2, d) = \alpha_1 - \beta_1 - d\beta_1(\alpha_2 - \beta_2)$$
$$\phi_2(\alpha_1, \alpha_2, d) = \alpha_2 - \beta_2 - d\beta_2(\alpha_1 - \beta_1)$$
$$\phi_3(\alpha_1, \alpha_2, d) = d(\alpha_1 - \beta_1)(\alpha_2 - \beta_2)$$

where we will assume $\alpha_1\alpha_2 d^2 > 1$ and β_1, β_2 are as in Lemma 4.9. We know from Lemma 4.12 that $\Gamma_{\alpha_1 N, \alpha_2 N, dN}$ **q.s.** has a 2–core with (approximately) $N\phi_1 + N\phi_2$ vertices and $N\phi_3$

edges. We will choose N,d so that

$$N\phi_1(1,1,d) = N\phi_2(1,1,d) = n \; ; \quad N\phi_3(1,1,d) = cn \tag{4.29}$$

This guarantees that $\Gamma_{N,N,dN}$ has a 2–core with approximately n+n vertices and cn edges and we will show that a graph "close" to this has a 2–core of the required size with a sufficiently high probability.

We must verify first that (4.29) has a solution. Observe now that from their definition in Lemma 4.9, where $d > 1$

$$\beta_1(1,1,d) = \beta_2(1,1,d) = \frac{x}{d} \tag{4.30}$$

where x is the unique solution in (0,1) to

$$xe^{-x} = de^{-d}. \tag{4.31}$$

Hence

$$\begin{aligned} \phi_1(1,1,d) &= \phi_2(1,1,d) = (1-x)(1-\tfrac{x}{d}) \\ \phi_3(1,1,d) &= d(1-\tfrac{x}{d})^2 \end{aligned} \tag{4.32}$$

so (4.29) is equivalent to

$$\frac{d-x}{1-x} = c \geq 2, \qquad \text{where } x \text{ satisfies (4.31)} \tag{4.33}$$

$$n = N(1-x)(1-\tfrac{x}{d}),$$

So let $f(D) = \frac{D - x}{1 - x}$ where $xe^{-x} = De^{-D}$ and $0 < x < 1 < D$. Then the proof that (4.33) is solvable comes from

Lemma 4.15

(a) $f(D)$ is monotone increasing.

(b) $\lim_{D \to 1+} f(D) = 2$.

(c) $\lim_{D \to \infty} f(D) = \infty$.

Proof

(a)

$$f'(D) = \frac{1 - x + (D - 1)x'}{(1 - x)^2}$$

where

(4.34)
$$x' = \frac{dx}{dD} = \frac{x - Dx}{D - Dx}.$$

Thus

$$f'(D) = \frac{D(1 - x)^2 - x(D - 1)^2}{D(1 - x)^2}$$

$$= \frac{(D - x)(1 - Dx)}{D(1 - x)^3}.$$

Hence

(4.35) $$f'(D) > 0 \leftrightarrow Dx < 1$$

$$\leftrightarrow \frac{1}{D} e^{-\frac{1}{D}} > De^{-D}$$

$$\leftrightarrow g(D) < 1$$

where $g(D) = D^2 e^{\frac{1}{D} - D}$.

But $g(1) = 1$ and $g'(D) < 0$ for $D \neq 1$ and (a) follows.

(b)

Now f is clearly differentiable for $D > 1$ and so let $L = \lim_{D \to 1_+} f(D)$. But

(4.36) $$f(D) = 1 + \frac{D - 1}{1 - x}$$

and so

$$L = 1 + \lim_{D \to 1_+} \frac{1}{-x'} \qquad \text{by L'Hopital's rule.}$$

$$= 1 + \lim_{D \to 1_+} \frac{D}{x(f(D) - 1)} \qquad \text{by (4.34)}$$

$$= 1 + \lim_{D \to 1_+} \frac{1}{f(D) - 1}.$$

Now clearly this implies that $L \neq 1$ and so

$$L = 1 + \frac{1}{L - 1}$$

and $L = 0$ or 2. But (4.36) implies $L \geq 1$ and so $L = 2$.

(c)

$$x \to 0 \text{ as } D \to \infty.$$

□

Now let $\text{CBOX} = \{(\lambda_1,\lambda_2,\lambda_3) \in \mathbb{R}^3 : |\lambda_i - n| \le n^{24/25}, i = 1,2 \text{ and } |\lambda_3 - cn| \le n^{24/25}\}$ and $\text{BOX} = \text{CBOX} \cap \mathbb{Z}^3$.

For $(\gamma_1,\gamma_2,\gamma_3) \in \mathbb{R}^3$, $\gamma_1\gamma_2\gamma_3^2 > N^4$ we let

$$\hat{\Phi}(\gamma_1,\gamma_2,\gamma_3) = N\Phi\left(\frac{\gamma_1}{N}, \frac{\gamma_2}{N}, \frac{\gamma_3}{N}\right)$$

and $\text{DOM} = \hat{\Phi}^{-1}(\text{CBOX})$. Note that $(N,N,dN) \in \text{DOM}$.

Now a rather lengthy calculation shows that the Jacobian of Φ at $(1,1,d)$ is given by

$$\det \begin{bmatrix} \dfrac{(1-\frac{x}{d})(1-x+x^2d-x^2)}{1-x^2} & \dfrac{(1-\frac{x}{d})^2dx}{1-x^2} & (1-\frac{x}{d})^2\dfrac{x}{1-x} \\ \dfrac{(1-\frac{x}{d})^2dx}{1-x} & \dfrac{(1-\frac{x}{d})(1-x+x^2d-x^2)}{1-x^2} & (1-\frac{x}{d})^2\dfrac{x}{1-x} \\ \dfrac{d(1-\frac{x}{d})^2}{1-x} & \dfrac{d(1-\frac{x}{d})^2}{1-x} & (1-\frac{x}{d})^2\dfrac{1+x}{1-x} \end{bmatrix}$$

$$= \frac{d^2(1-\frac{x}{d})^6}{1-x^2}$$

$$\neq 0$$

since $0 < x < 1 < d$ and $dx < 1$ (see (4.35).) This implies that the Jacobian of Φ is non–zero in some neighbourhood of $(1,1,d)$. Hence there exists a fixed $\epsilon > 0$ such that the Jacobian of $\hat{\Phi}$ is non–zero in $\{(\gamma_1,\gamma_2,\gamma_3) : |\gamma_i - N| \le \epsilon N, i = 1,2 \text{ and } |\gamma_3 - dN| \le \epsilon N\}$. Thus for large N, $\hat{\Phi} : \text{DOM} \to \text{CBOX}$ is a bijection. Furthermore, simple continuity arguments imply that for some $\rho_2 > 0$

(4.37a) $$x,y \in \mathrm{CBOX} \to \|\Phi^{-1}(x) - \Phi^{-1}(y)\| \leq \rho_2\|x-y\|,$$

(4.37b) $$x,y \in \mathrm{DOM} \to \|\Phi(x) - \Phi(y)\| \leq \rho_2\|x-y\|.$$

Given the random graph $\Gamma_{\nu,\nu,\mu}$ we define a function $\Psi : \mathrm{BOX} \to \mathbb{Z}^3$ as follows: if $x \in \mathrm{BOX}$ let $y = (y_1,y_2,y_3) = \hat{\Phi}^{-1}(x)$. Then let $(i,j,k) = \lfloor y \rfloor = (\lfloor y_1 \rfloor, \lfloor y_2 \rfloor, \lfloor y_3 \rfloor)$. Finally, let

$$\Psi(x) = (\xi_1(i,j,k), \xi_2(i,j,k), \xi_3(i,j,k))$$

$$= (\psi_1(x), \psi_2(x), \psi_3(x)).$$

Observe in the above definition that, where $x_0 = (n,n,cn)$, $y_0 = \Phi^{-1}(x_0) = (N,N,dN)$, we have, by (4.37a),

$$\|y-y_0\| \leq \rho_2\|x-x_0\| \leq \rho_2\sqrt{3}\, n^{24/25}$$

and so $(i,j,k) \in J$.

Observe also that q.s. for all $x \in \mathrm{BOX}$

(4.38) $$\|\Psi(x) - x\| = \|\xi - \hat{\Phi}(y)\|$$

$$\leq \sqrt{3}\, N^{19/20} \qquad \text{by Lemma 4.12.}$$

We extend Ψ continuously to CBOX so that if x lies in the interior of a cube C of the integer lattice then the components of $\Psi(x)$ lie between the largest and smallest corresponding components at the vertices of C. The aim now is to show that q.s. there exists $\hat{x} \in \mathrm{CBOX}$ such that

(4.39) $$\Psi(\hat{x}) = \hat{y} = (n - 2\rho_1\rho_2 n,\ n - 2\rho_1\rho_2 \log n,\ cn - 13\rho_1\rho_2 \log n).$$

To do this we use a result in Non–linear Functional Analysis (Schwartz [S], Chapter III). For this result D refers to an open bounded subset of $\mathbb{R}^n$, $\bar{D}$ is its closure and ∂D is its boundary.

Theorem 4.16 To every continuous map $\phi : \bar{D} \to \mathbb{R}^n$ and every point $p \notin \phi(\partial D)$ there is an integer $\deg(p,\phi,D)$ with the properties:

(a) If ϕ_t is a family of continuous mappings depending continuously, in the uniform topology, on t, $0 \le t \le 1$ and such that $p \notin \phi_t(\partial D)$ for every t then

$$\deg(p,\phi_0,D) = \deg(p,\phi_1,D)$$

(b)

$$\deg(p,\phi,D) \neq 0 \text{ implies } p \in \phi(\bar{D})$$

(c)

$$\deg(p,I_n,D) = 1 \quad \text{for } p \in \bar{D}.$$

where I_n is the identity map on $\mathbb{R}^n$. □

To apply Theorem 4.16 we take $D = \mathrm{int}(\mathrm{CBOX})$ and $\phi = \Psi$ and then let $\phi_t = t\Psi + (1-t)I_3$ for $0 \le t \le 1$. (4.39) will follow once we verify that $\hat{y} \notin \phi_t(\partial D)$ for all t. Suppose then that $x \in \partial D$ and for example that $x_1 = n + n^{24/25}$. Then, where $\phi_t(x) = (\gamma_1,\gamma_2,\gamma_3)$ we have, by (4.38), that $\gamma_1 \ge x_1 - \sqrt{3}\, N^{19/20} > n$. Other cases are almost identical and (4.39) follows.

Now given $\hat{x}$ in (4.39) let $\tilde{x}$ be a vertex of the cube containing $\hat{x}$. It follows from (4.37) and Lemma 4.14 that **q.s.** if $(\tilde{i},\tilde{j},\tilde{k}) = \lfloor \Phi^{-1}(\tilde{x}) \rfloor$ then $\xi(\tilde{i},\tilde{j},\tilde{k}) \in K$ where $K = \{(a_1,a_2,b) \in \mathbb{Z}^3 : n - 3\rho_1\rho_2\log n \le a_i \le n - \rho_1\rho_2\log n,\ cn - 14\rho_1\rho_2\log n \le b \le cn - 12\rho_1\rho_2\log n\}$. Since $|\mathrm{BOX}| = O(n^{2\cdot 88})$ we deduce that there exists some $(i',j',k') \in J$ such that

$$\Pr(\mathcal{A}_3) \le n^{C_1} \sum_{k=A_1 \log n}^{A_2 \log n} \sum_{\ell=k-\mu}^{k-1} \binom{N_1}{k} \binom{N_2}{\ell} \binom{k\ell}{2k} \left(\frac{d}{N}\right)^{2k} \binom{N_1\ell}{\eta\ell} \left(\frac{d}{N}\right)^{\eta\ell}$$

$$\le \sum_{k=A_1 \log n}^{A_2 \log n} \sum_{\ell=k-\mu}^{k-1} e^{o(k)} \frac{N^{C_1+k+\ell} e^{k+\ell}}{k^k \ell^\ell} \left(\frac{k\ell e d}{2kN}\right)^{2k} \left(\frac{Ne\ell d}{\eta\ell N}\right)^{\eta\ell}$$

$$\le \sum_{k=A_1 \log n}^{A_2 \log n} \sum_{\ell=k-\mu}^{k-1} e^{o(k)} N^{C_1} \left(\frac{e^2 d}{2}\right)^{2k} \left(\frac{ed}{\eta}\right)^{\eta\ell}$$

$$= o(1)$$

for large η.

□

We now show that

$$\Pr(\mathrm{HALL}(A_1 \log n \le k \le A_2 \log n)) = o(1), \tag{5.9}$$

for all (large) constants $A_2 > 0$.)

We will use a similar but more careful estimate to that of (5.2) but we can now restrict our attention to K,L,k,ℓ s.t satisfying

(5.10a)	$\ell \ge k - \mu$	Corollary 5.2
(5.10b)	$s \le k+\ell+\mu$	Corollary 5.2
(5.10c)	L has at most ν large vertices	Lemma 5.3
(5.10d)	$t \le \eta\ell$	Lemma 5.4

So let HALL$'(k,\ell)$ denote HALL(k,ℓ) where (5.10) is incorporated as extra conditions. We need to show

$$\Pr(\mathrm{HALL}'(A_1\log n \le k \le A_2\log n)) = o(1). \tag{5.11}$$

Fix k within this range and ℓ satisfying (5.10a). Then

$$\Pr(\mathrm{HALL}'(k,\ell)) \le \binom{n}{k}\binom{n}{\ell} \sum_{s=2k}^{2k+\mu} \sum_{t=s}^{\eta\ell} \Pr(D_k = s, D_\ell = t, N([k]) \subseteq [\ell], (H4)).$$

We observe next that since $k = 0(\log n)$

$$\Pr(d_i = r \mid d_1,d_2,\ldots,d_{i-1}) \le \frac{\tilde{c}^r e^{-\tilde{c}}}{r!\tilde{\theta}} + 2n^{-1/5}, \qquad 1 \le i \le k,\ 2 \le r \le \log n, \tag{5.12a}$$

$$= (1 + 0(n^{-1/10}))\frac{\tilde{c}^r e^{-\tilde{c}}}{r!\tilde{\theta}}, \qquad 2 \le r \le \frac{\log n}{10\ \log\log n}. \tag{5.12b}$$

The first inequality follows from (3.2) and the fact that once the number of vertices with a given degree is fixed, the degrees of a particular set of vertices is obtained by sampling from the multi–set of degrees, without replacement. Let now

$$\Omega(s,1) = \{(s_1,s_2,\ldots,s_k) : 2 \le s_1,s_2,\ldots,s_k \le \frac{2\ \log n}{\log\log n},\ s_1 + s_2 + \ldots + s_k = s\}$$

and

$$\Omega(t,2) = \{(t_1,t_2,\ldots,t_\ell) \in \Omega(t,1) : |\{i : t_i \ge \frac{\log n}{10\ \log\log n}\}| \le \nu\}.$$

Then, where $K = [k] \subseteq V$, $L = [\ell] \subseteq W$,

$$(5.13) \qquad \Pr(d_K = s,\, d_L = t,\, N(K) \subseteq L,\, (H4)) \le$$

$$\sum_{x\in\Omega(s,1)} \Pr(d_X(i) = x_i,\, i \in K)\, \frac{(t)_s}{(cn)_s} \sum_{y\in\Omega(t,2)} \Pr(d_Y(j) = y_j,\, j \in L)\, \frac{\prod_{i=1}^{\ell} \binom{y_j}{2} t^{s-2\ell}}{\binom{t}{s}}.$$

EXPLANATION

Having fixed the degrees x,y for $i \in K$, $j \in L$ the probability that $N(K) \subseteq L$ is $(t)_s/(cn)_s$. Given that $N(K) \subseteq L$ the probability that each vertex of L has at least 2 neighbours in K is at most the final term.

Note next that (5.10a) and (5.10b) imply $|\Omega(s,1)| \le \binom{s-k-1}{k-1} < k^{\mu} = e^{o(k)}$ and that $x \in \Omega(s,1)$ implies there are at most μ x_i's $\neq 2$ and $\max\{x_i\} < \mu$. Hence, using (5.12b)

$$(5.14) \qquad \sum_{x\in\Omega(s,1)} \Pr(d_X(i) = x_i,\, i \in K) \le e^{o(k)}\, \frac{\tilde{c}^{2k} e^{-\tilde{c}k}}{2^k \tilde{\theta}^k}.$$

In the following estimate λ refers to the number of large vertices in L.

$$(5.15) \qquad \sum_{y\in\Omega(t,2)} \Pr(d_Y(j) = y_j,\, j \in L) \prod_{j=1}^{\ell} \binom{y_j}{2}$$

$$\le \sum_{\lambda=0}^{\nu} \binom{\ell}{\lambda} n^{-\lambda/11} \sum_{\tau=\tau_0}^{\tau_1} \sum_{y\in\Omega(t,2)} e^{o(\ell)}\, \frac{e^{-\tilde{c}(\ell-\lambda)}\tilde{c}^{\tau}}{\tilde{\theta}^{\ell-\lambda} \prod_{j=1}^{\ell-\lambda} y_j!} \prod_{j=1}^{\ell} \binom{y_j}{2}$$

(where $\tau_0 = t - \frac{2\lambda \log n}{\log\log n}$ and $\tau_1 = t - \frac{\lambda \log n}{10 \log\log n}$.)

$$\le e^{o(\ell)} \frac{e^{-\tilde{c}\ell}\tilde{c}^t}{(2\tilde{\theta})^\ell} \sum_{y\in\Omega(t,2)} \left(\prod_{j=1}^{\ell} (y_j-2)!\right)^{-1}$$

$$\le e^{o(\ell)} \frac{e^{-\tilde{c}\ell}\tilde{c}^t}{(2\tilde{\theta})^\ell} \frac{\ell^{t-2\ell}}{(t-2\ell)!}.$$

Combining this inequality with (5.14) in (5.13) we see that

$$\Pr(\mathrm{HALL}'(k,\ell)) \le \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} n \\ \ell \end{bmatrix} \sum_{s=2k}^{2k+\mu} \sum_{t=s}^{\eta\ell} e^{o(k)} \frac{\tilde{c}^{2k+t} e^{-\tilde{c}(k+\ell)}}{(2\tilde{\theta})^{k+\ell}} \frac{\ell^{t-2\ell} t^{s-2\ell}}{(t-2\ell)! \begin{bmatrix} t \\ s \end{bmatrix}} \frac{(t)_s}{(cn)_s}$$

Now in our range of interest $t^{s-2\ell} = e^{o(k)}$; $(t-2\ell)! = (\frac{t-2k}{e})^{t-2k} e^{o(k)}$; $\ell^{t-2\ell} \le k^{t-2k} e^{o(k)}$; $(t)_s / (\begin{bmatrix} t \\ s \end{bmatrix} (cn)_s) \le (\frac{2k}{cen})^{2k} e^{o(k)}$. Hence

$$\Pr(\mathrm{HALL}'(k,\ell)) \le \sum_{s=2k}^{2k+\mu} \sum_{t=s}^{\eta\ell} e^{o(k)} \frac{n^{k+\ell} e^{k+\ell}}{k^k \ell^\ell} \frac{\tilde{c}^{2k+t} e^{-2\tilde{c}k}}{(2\tilde{\theta})^{2k}} \left(\frac{ke}{t-2k}\right)^{t-2k} \left(\frac{2k}{cen}\right)^{2k}$$

$$\le \sum_{t=2k}^{2\eta k} e^{o(k)} \frac{\tilde{c}^{2k+t} e^{-2\tilde{c}k}}{(c\tilde{\theta})^{2k}} \left(\frac{ke}{t-2k}\right)^{t-2k}$$

$$= e^{o(k)} \sum_{t=2k}^{\eta k} \left(\frac{\tilde{c}^2 e^{-2\tilde{c}}}{(1-e^{-\tilde{c}})^2} \left(\frac{\tilde{c}e}{\frac{t}{k}-2}\right)^{\frac{t}{k}-2} \right)^k$$

(after using $c\tilde{\theta} = \tilde{c}(1-e^{-\tilde{c}})$,)

$$\le e^{o(k)} \sum_{t=2k}^{\eta k} \left(\frac{\tilde{c}^2 e^{-\tilde{c}}}{(1-e^{-\tilde{c}})^2}\right)^k$$

(after using $(\alpha/x)^x \le e^{\alpha/e}$.)

We obtain (5.11) and hence (5.9) once we observe that

$$\frac{x^2e^{-x}}{(1-e^{-x})^2} < 1 \quad \text{for } x > 0$$

$$(e^x((1-e^{-x})^2 - x^2e^{-x}) = \sum_{r=2}^{\infty} \frac{x^{2r}}{(2r)!}.)$$

Large k

We now finally consider $A_2\log n \le k \le \frac{n}{2}$. For this we use the 2–core model of §4 and obtain the inequality

$$\Pr(\mathrm{HALL}(k,\ell,s)) \le n^{C_1}\begin{bmatrix}N_1\\k\end{bmatrix}\begin{bmatrix}N_2\\\ell\end{bmatrix}g_{k,\ell,s}\left(\frac{d}{N}\right)^s\left(1-\frac{d}{N}\right)^{N_2k-s} \tag{5.16}$$

where $g_{k,\ell,s}$ = the number of bipartite graphs with vertex partition [k], [ℓ], s edges and minimum degree at least two.

The RHS of (5.16) is the expected number of subgraphs of $B_{N_1,N_2,p}$ with k vertices in V, ℓ vertices in W, s edges and minimum degree at least two, multiplied by n^{C_1}, since we are using (4.4).

Now

$$g_{k,\ell,s} \le \frac{1}{s!}h_{k,s}h_{\ell,s} \tag{5.17}$$

where $h_{k,s}$ = the number of ways of putting s distinguishable balls into k boxes with at least 2 balls in each box.

To see this let G be a bipartite graph with $k+\ell$ vertices and edges v_iw_i, $i = 1,2,\dots,s$. Any permutation ϕ of [s] yields a pair of allocations of balls into boxes: ball i into box $v_{\phi(i)}$, $i = 1,2,\dots,s$ and ball j into box $w_{\phi(j)}$, $j = 1,2,\dots,s$. Different orderings clearly yield different pairs of allocations. Note also that distinct graphs yield disjoints sets of pairs of allocations since such a pair completely determines the edges of the graph.

Now

$$\text{(5.18)} \qquad h_{k,s} = \sum_{d \in \Omega} s! \Big/ \prod_{i=2}^{s} (i!)^{d_i}$$

where

$$\Omega = \{(d_1,\dots,d_s) \geq 0\colon \sum_{i=2}^{s} d_i = k,\ \sum_{i=2}^{s} i\, d_i = s\}.$$

Now let $\rho = s/k$ and $\tilde{\rho}$ satisfy

$$\rho = \frac{\tilde{\rho} - e^{-\tilde{\rho}}\tilde{\rho}}{1 - e^{-\tilde{\rho}}(1+\tilde{\rho})}$$

and

$$\Omega_1 = \{d \in \Omega : \text{(i)} \quad d_i = 0 \text{ for } i \geq \log k,$$

$$\text{(ii)} \quad \left| d_i - \frac{\tilde{\rho}^i e^{-\tilde{\rho}}}{\tilde{\sigma}\, i!} \right| \leq i^{4/5}, \quad 2 \leq i < \log k\},$$

where $\tilde{\sigma} = 1 - e^{-\tilde{\rho}}(1 + \tilde{\rho})$.

It follows from Lemma 3.4 and (5.18) that, say,

$$\text{(5.19)} \qquad h_{k,s} \leq 2 \sum_{d \in \Omega_1} s! \Big/ \prod_{i=2}^{s} (i!)^{d_i}.$$

Now $|\Omega_1| \leq (2k^{4/5} + 1)^{\log k} = e^{o(k)}$ and so if

$$f(k,s) = \left(\left\lfloor \frac{s}{k} \right\rfloor !\right)^k \left\lceil \frac{s}{k} \right\rceil^{s \bmod k}$$

then (5.19) implies

$$h_{k,s} \le e^{o(k)}\, s!/f(k,s). \tag{5.20}$$

$\left(\prod_{i=2}^{s} (i!)^{d_i} \ge f(k,s)\right.$ follows from $(a+1)!(b-1)! > a!b!$ if $a \ge b$.)

To use (5.20) we need to know something about the behaviour of $f(k,s)$. What we need is summarised in the following lemma:

Lemma 5.5 Suppose $v = au + b$ where $a = \lfloor \frac{v}{u} \rfloor \ge 2$. then

(a) $\frac{f(u,v+1)}{f(u,v)} = a + 1$

(b) $\frac{f(u,v)}{f(u+1,v)} \ge 2$, assuming $a \le u + 1$.

Proof

(a) $f(u,v) = (a!)^u(a+1)^b$

and

$$f(u,v+1) = \begin{cases} (a!)^u(a+1)^{b+1} & 0 \le b \le u-2 \\ ((a+1)!)^{u+1} & b = u-1 \end{cases}$$

and (a) follows.

(b)

Suppose first that $\lfloor \frac{v}{u+1} \rfloor = a$ and so $v = a(u+1) + b - a$ where $b \ge a$.

Then

$$f(u+1,v) = (a!)^{u+1}(a+1)^{b-a}$$

and so

$$\frac{f(u,v)}{f(u+1,v)} = \frac{(a+1)^a}{a!} \ge \frac{9}{2}.$$

Suppose next that $\lfloor\frac{v}{u+1}\rfloor = a - 1$ and so $v = (a-1)(u+1) + b-a + u+1$ where $b-a + u+1 \geq 0$. Then

$$\frac{f(u,v)}{f(u+1,v)} = \frac{(a!)^u \ (a+1)^b}{((a-1)!)^{u+1} \ a^{b-a+u+1}}$$

$$= \frac{a^a}{a!}\left(\frac{a+1}{a}\right)^b$$

$$\geq 2.$$

Note finally that $\lfloor\frac{v}{u+1}\rfloor \leq a-2$ is ruled out by $a \leq u+1$.

□

It follows from(5.16), (5.17) and (5.20) that

(5.21) $$\Pr(\mathrm{HALL}(k,\ell,s)) \leq e^{o(k)}\, n^{C_1} r(k,\ell,s)$$

where

(5.22) $$r(k,\ell,s) = \begin{bmatrix} N_1 \\ k \end{bmatrix}\begin{bmatrix} N_2 \\ \ell \end{bmatrix}\left(\frac{d}{N}\right)^s\left(1 - \frac{d}{N}\right)^{N_2 k - s}\frac{s!}{f(k,s)f(\ell,s)}.$$

Now, by Lemma 5.5, for $\ell < k \leq \frac{1}{2}n < \frac{1}{2}N_2$

(5.23) $$\frac{r(k,\ell+1,s)}{r(k,\ell,s)} = \frac{N_2 - \ell}{\ell + 1}\,\frac{f(\ell,s)}{f(\ell+1,s)}$$

$$> 2$$

provided we can assume $\lfloor\frac{s}{\ell}\rfloor \leq \ell + 1$. But this is justified by $s < k \log n$ q.s., A_2 is large and

Lemma 5.6 Suppose $A_2 \log n \leq k \leq \frac{1}{2}n$. Let $\mathcal{A}_4$ denote the event that there exist K,L satisfying (H1), (H2) and $\ell \leq \ell_0 = e^{-4}d^{-2}k$. Then $\Pr(B_n \in \mathcal{A}_4) = o(1)$.

PROOF It follows from (4.1) – (4.4) that

$$\Pr(B_n \in \mathcal{A}_4) \leq n^{C_1} \sum_{\ell \leq \ell_0} \sum_{s \geq 2k} \binom{N_1}{k}\binom{N_2}{\ell}\binom{k\ell}{s}\left(\frac{d}{N}\right)^s$$

$$\leq n^{C_1} \sum_{\ell \leq \ell_0} \sum_{s \geq 2k} e^{o(k)} \frac{N^{k+\ell} e^{k+\ell}}{k^k \ell^\ell}\left(\frac{k\ell e d}{sN}\right)^s$$

$$\leq 2n^{C_1} \sum_{\ell \leq \ell_0} e^{o(k)} \frac{N^{k+\ell} e^{k+\ell}}{k^k \ell^\ell}\left(\frac{\ell e d}{2N}\right)^{2k}$$

$$\leq 2n^{C_1} \sum_{\ell \leq \ell_0} e^{o(k)} \left(\frac{e^4 d^2 \ell}{4k}\right)^k \left(\frac{\ell}{N}\right)^{k-\ell}$$

$$= o(1)$$

□

Thus,

(5.24) $$\Pr(\mathrm{HALL}(k \leq A_2 \log n \leq \tfrac{1}{2} n, \ell \leq e^{-4} d^{-2} k)) = o(1)$$

and we are justified in using (5.23) to obtain

(5.25) $$\sum_{\ell=\ell_0}^{k-1} r(k,\ell,s) \leq 2r(k,k,s).$$

Using Lemma 5.5 once more we have

$$\frac{r(k,k,s+1)}{r(k,k,s)} = \frac{d}{N}\left(1-\frac{d}{N}\right)^{-1}\frac{s+1}{(\lfloor\frac{s}{k}\rfloor + 1)^2}$$

(5.26) $$\leq (1 + o(1))\frac{dk^2}{sN}$$

$$\leq (1 + o(1))\frac{d}{4}.$$

Assume for the moment that $d < 4$. Then

(5.27) $$\sum_{s \geq 2k} r(k,k,s) \leq \frac{4 + o(1)}{4 - d}\, r(k,k,2k).$$

But

$$r(k,k,2k) \leq e^{o(k)} \frac{N^{2k}e^{2k}}{k^{2k}} \left(\frac{d}{N}\right)^{2k} e^{-dk} \frac{(2k)!}{2^{2k}}$$

$$= e^{o(k)}(d^2e^{-d})^k$$

(5.28) $$< \left(\frac{4}{7}\right)^k.$$

It follows from (5.21), (5.24), (5.25), (5.27) and (5.28) that if $d > 4$ and A_2 is sufficiently large then

$$\Pr(\mathrm{HALL}(A_2\log n \leq k \leq \tfrac{1}{2}n) = o(1)$$

and this completes the proof for $d < 4$.

Suppose now that $d \geq 4$. In view of (5.25) we need only consider $r(k,k,s)$. But

$$r(k,k,s) \le e^{o(k)} \frac{N^{2k}e^{2k}}{k^{2k}} (\frac{d}{N})^s e^{-dk} \sqrt{2\pi s} (\frac{s}{e})^s (\frac{ke}{s})^{2s}$$

(since Stirling's approximation implies $f(k,s) \ge (\frac{s}{ek})^s$.)

Writing $s = \alpha k$, $\alpha \ge 2$,

$$r(k,k,s) \le e^{o(k)} \sqrt{2\pi s} ((\frac{k}{N})^{\alpha-2} e^{2-d}(\frac{ed}{\alpha})^{\alpha})^k$$

$$\le e^{o(k)} \sqrt{2\pi s} (4e^{2-d} (\frac{ed}{2\alpha})^{\alpha})^k$$

(since $k < N/2$)

$$\le e^{o(k)} \sqrt{2\pi s} (4e^{2-d/2})^k$$

since $(\frac{ed}{2\alpha})^{\alpha} \le e^{d/2}$.

Now $4e^{2-d/2} < 1$ for $d > 4 + 2\log 4 = 6.7725887...$ and we have completed the proof for say, $d > 7$.

It remains only to consider $4 \le d \le 7$ and $s \le dk/2 < 3.5k$. If $s > dk/2$ then (5.26) allows us to reduce to this case.

Case 1: $s = (2 + \beta)k$ where $0 \le \beta < 1$.

$$f(k,s) = 2^k\, 3^{\beta k}$$

$$r(k,k,s) \le e^{o(k)} \sqrt{2\pi s} \frac{N^{2k}e^{2k}}{k^{2k}} (\frac{d}{N})^{(2+\beta)k} e^{-dk} (\frac{(2+\beta)k}{e})^{(2+\beta)k} \frac{1}{4^k 9^{\beta k}}$$

$$\le e^{o(k)} \sqrt{2\pi s} ((\frac{k}{N})^{\beta} \frac{d^{2+\beta} e^{-(d+\beta)} (2+\beta)^{2+\beta}}{4 \times 9^{\beta}})^k$$

$$\leq e^{o(k)} \sqrt{2\pi s}\, (d^2 e^{-d} (1 + \frac{\beta}{2})^2 (\frac{d(2+\beta)}{9e})^{\beta})^k$$

$$< e^{o(k)} \sqrt{2\pi s}\, (\frac{9}{4} d^2 e^{-d})^k$$

and we are done, since $d \geq 4$.

Case 2: $s = (3+\beta)k$ where $0 \leq \beta \leq \frac{1}{2}$.

$$f(k,s) = 6^k 4^{\beta k}$$

$$r(k,k,s) \leq e^{o(k)} \sqrt{2\pi s}\, \frac{N^{2k} e^{2k}}{k^{2k}} (\frac{d}{N})^{(3+\beta)k} e^{-dk} (\frac{(3+\beta)k}{e})^{(3+\beta)k} \frac{1}{36^k 16^{\beta k}}$$

$$= e^{o(k)} \sqrt{2\pi s}\, ((\frac{k}{N})^{1+\beta} \frac{d^{3+\beta} e^{-(d+1+\beta)} (3+\beta)^{3+\beta}}{36 \times 16^{\beta}})^k$$

$$\leq e^{o(k)} \sqrt{2\pi s}\, (\frac{d^3 e^{-(d+1)}}{72} (3+\beta)^{3+\beta} (\frac{d}{32e})^{\beta})^k$$

$$\leq e^{o(k)} \sqrt{2\pi s}\, (\frac{d^3 e^{-(d+1)}}{72} (3\cdot 5)^{3\cdot 5})$$

and we are done, since $d \geq 4$.

This completes the proof of our Theorem.

□

We claimed in the introduction that Bollobás and Frieze [BF] had shown that roughly $\frac{1}{4}$ n logn edges were needed in $G^{(1)}_{n,m}$ and that the proof could be extended to bipartite graphs. The proof as given there would not allow us to deal with $m = 0(n)$. So it behoves us to explain why with cn edges, we insist on minimum degree at least 2. If $c = 1$ then our bipartite graph is a perfect matching. However, for $c > 1$ we can use the method of §3 to show that if a graph G is sampled uniformly from $\mathscr{B}(V,W;\ \delta \geq 1)$ then it has a degree sequence as in Lemma 3.4 with a slightly different definition of $\tilde{c}$ and of course $k = 1,2,\ldots,\lceil \log n \rceil$. Of course Lemma 3.6 remains true. Under these conditions it is easy to show that with probability $1 - o(1)$ there will be 2 vertices of degree that share a common neighbour.

Acknowledgement: I thank Charlie Coffman and Mete Soner for their help, along the way.

REFERENCES

[B1] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[B2] B. Bollobás, "Martingales, isoperimetric inequalities and random graphs", in *Combinatorics*, A. Hajnal, L. Lovász and V.T. Sos Ed., *Colloq. Math. Sci. Janos Bolyai* 52, North Holland, 1988.

[BF] B. Bollobás and A.M. Frieze, "On matchings and Hamilton cycles in random graphs", *Annals of Discrete Mathematics* **28** (1985) 23–46.

[BFF] B. Bollobás, T.I. Fenner and A.M. Frieze, "Hamilton cycles in random graphs with minimal degree at least k", to appear in *Proceedings of Cambridge Combinatorics Conference in honour of Paul Erdős's 75th birthday.*

[BCFF] B. Bollobás, C. Cooper, T.I. Fenner and A.M. Frieze, "Hamilton cycles in sparse random graphs with minimal degree at least k", in preparation.

[EF] P. Erdős and A. Rényi, "On random matrices", *Publ. Math. Inst. Hungar. Acad. Sci.* **8**, (1964) 455–461.

[FMcD] A.M. Frieze and C.J.H. McDiarmid, "On random minimum length spanning trees", *Combinatorica* **9** (1989) 363–374.

[KMP] D.E. Knuth, R. Motwani and B. Pittel, "Stable husbands", *Random Structures and Algorithms* **1** (1990) 1–14.

[McD] C.J.H. McDiarmid, "On the method of bounded differences", *Surveys in Combinatorics 1989*, Invited papers at the Twelfth British Combinatorial Conference, Edited by J. Siemons, Cambridge University Press, 148–188.

[P] B. Pittel, "On trees census and the giant component in sparse random graphs", to appear.

[S] J.T. Schwartz, "Non–linear functional analysis", Lecture notes on a course given at the Courant Institute of Mathematical Sciences, New York, 1963–1964.

[SS] E. Shamir and J. Spencer, "Sharp concentration of the chromatic number of random graphs $G_{n,p}$", *Combinatorica* 7, (1987), 121–129.