§1. Introduction

Let H be some fixed graph with r vertices and s edges. H is assumed to be strictly balanced i.e.

$$\frac{s}{r} > \frac{\mu(H')}{\nu(H')}$$

for all non–trivial subgraphs H′ of H, H′ ≠ H, where $\nu(H')$, $\mu(H')$ are the numbers of vertices, edges in H′ respectively. (From now on $H' \subset H$ will always mean such subgraphs).

Consider now the random graph $G_{n,m}$ chosen uniformly from $\mathcal{G}_{n,m}$ = {graphs with vertex set $[n] = \{1,2,...,n\}$ and m edges} and let X_H denote the number of distinct copies of H in $G_{n,m}$. Suppose now $m = \frac{1}{2}\omega n^{2-r/s}$ where $\omega = \omega(n)$. Erdős and Rényi [3] showed that

$$\Pr(X_H = 0) = 1 - o(1) \qquad \text{if } \omega \to 0$$
$$\Pr(X_H \neq 0) = 1 - o(1) \qquad \text{if } \omega \to \infty.$$

Here, as usual, we consider limits etc. as $n \to \infty$. Using $a(n) \sim b(n)$ to stand for $a(n) = (1 - o(1))\, b(n)$, we remark that

$$E(X_H) \sim \frac{\omega^s}{\alpha} = \lambda, \text{ say,}$$

where α denotes the number of automorphisms of H.

Erdős and Rényi's result has been refined in many ways. In particular, Bollobás [1] and Karonski and Rucinski [6] independently showed that if ω tends to a constant and k is a fixed non–negative integer then

(1.1) $$\Pr(X_H = k) \sim e^{-\lambda}\frac{\lambda^k}{k!}.$$

The aim of this paper is to show that the Poisson expression (1.1) is good for $\omega \to \infty$ reasonably fast. In particular we prove

THEOREM 1.1

Let H be strictly balanced and λ be as defined above. Then there exists a positive real constant $\theta = \theta(H)$ such that if $\omega = o(n^\theta)$ then

(1.2) $$\Pr(X_H = k) \sim e^{-\lambda}\frac{\lambda^k}{k!} \qquad 0 \lesssim k \lesssim (1+\epsilon_1)\lambda$$

where $\epsilon_1 = \dfrac{A_1(\log n)^{r/(2r-1)}}{\lambda^{(r-1)/(2r-1)}}$ for some constant $A_1 > 0$. □

(1.3) $$\Pr(X = k) >> e^{-\lambda}\frac{\lambda^k}{k!} \qquad (1+\epsilon_2)\lambda \lesssim k \lesssim \lambda\log n$$

where $\epsilon_2 = A_2\left(\frac{\log n}{\lambda^{1-2/r}}\right)^{r/2(r-1)}$ for some constant $A_2 > 0$, provided $\epsilon_2 \to 0$.

(The notation $a(n) >> b(n)$ is used for $a(n)/b(n) \to \infty$).

Remarks

1. We are not able to obtain the largest possible values for $\theta(H)$ although we hope to refine our analysis for particular graphs e.g. triangles.
2. Observe that $\epsilon_1\lambda >> \lambda^{1/2}$ and so (1.2) is valid into the tails of the Poisson distribution.

3. A somewhat stronger result for $k = 0$ and $G_{n,p}$ has been proved independently by Boppanna and Spencer [2] and Jansen, Luczak and Rucinski [4]. Jansen [5] has extended these result to estimate $\Pr(X_H \le k)$ for $k \le E(X_H)$.
4. See Rucinski [7] for a recent survey on the distribution of the number of copies of small subgraphs of random graphs.

§2. Proof of Theorem 1.1.

We will not specify $\theta(H)$ immediately but upper bounds for it will be derived along with the proof. We will use $A, A_1, A_2, \ldots$ to denote absolute constants whose values may or may not be explicitly stated.

We distinguish between <u>isolated</u> copies of H and <u>non–isolated</u> copies. Here a copy of H in $G_{n,m}$ is isolated if it shares no edge with any other copy of H.

Now let

$$\pi_{k,\ell} = \Pr(G_{n,m} \text{ contains exactly } k \text{ isolated and } \ell \text{ non–isolated copies of } H)$$

and

$$q_\ell = \sum_{k=0}^{\infty} \pi_{k,\ell} = \Pr(G_{n,m} \text{ contains exactly } \ell \text{ non–isolated copies of } H)$$

and

$$p_k = \sum_{\ell=0}^{k} \pi_{k-\ell,\ell} = \Pr(G_{n,m} \text{ contains exactly } k \text{ copies of } H).$$

The main work involved in the proof is to justify the following inequalities:

(2.1) $$n^{-A_3\ell^{2/r}} \le q_\ell \le n^{-A_4\ell^{1/r}} \qquad 0 \le \ell \le \lambda_0 = \lambda(\log n)^4$$

(2.2) $$\Pr(G_{n,m} \text{ contains at least } \lambda_0 \text{ isolated copies of } H) = o(e^{-\lambda_0})$$

and more importantly

(2.3) $$\frac{\pi_{k,\ell}}{\pi_{k-1,\ell}} = (1 + \epsilon_{k,\ell}) \frac{\lambda}{k} \qquad 0 \le k-1, \ell \le \lambda_0$$

where $|\epsilon_{k,\ell}| = o(\lambda_0^{-1})$.

We devote the remainder of this section to showing how our theorem follows from (2.1) – (2.3) and prove these inequalities later on.

Suppose now that $0 \le \ell \le \lambda_0$. It follows from (2.3) that

(2.4) $$\pi_{i,\ell} = (1 + o(1))\, \pi_{0,\ell} \frac{\lambda^i}{i!} \qquad 0 \le i \le \lambda_0$$

and so

$$q_\ell = (1 + o(1))\, \pi_{0,\ell} \sum_{i=0}^{\lambda_0} \frac{\lambda^i}{i!} + \sum_{i>\lambda_0} \pi_{i,\ell}$$

$$= (1 + o(1))\ \pi_{0,\ell}(e^{\lambda} - o(e^{-\lambda_0})) + o(e^{-\lambda_0})$$

on using (2.2). Hence

$$\pi_{0,\ell} = (1 + o(1))\ (q_\ell - o(e^{-\lambda_0}))e^{-\lambda}$$

and by (2.4)

$$\pi_{i,\ell} = (1 + o(1))\ q_\ell\, e^{-\lambda}\,\frac{\lambda^i}{i!} + o(\frac{\lambda^i}{i!}\, e^{-\lambda-\lambda_0}) \qquad 0 \le i \le \lambda_0$$

Thus

$$p_k = (1 + o(1)) \sum_{\ell=0}^{k} q_\ell\, e^{-\lambda}\,\frac{\lambda^{k-\ell}}{(k-\ell)!} + o(e^{-\lambda_0}\lambda_0) \qquad 0 \le k \le \lambda_0.$$

Now

$$p_k \ge q_k \ge n^{-A_3(\lambda_0)^{2/r}} >> e^{-\lambda_0}\lambda_0 \qquad \text{since } r \ge 3$$

and so

$$p_k \sim \sum_{\ell=0}^{k} q_\ell e^{-\lambda}\,\frac{\lambda^{k-\ell}}{(k-\ell)!} \qquad 0 \le k \le \lambda_0.$$

$$(2.5) \qquad = e^{-\lambda}\frac{\lambda^k}{k!}\Big(q_0 + \sum_{\ell=2}^{k}\frac{(k)_\ell}{\lambda^\ell}q_\ell\Big)$$

where $(k)_\ell = k(k-1)...(k-\ell+1)$.

To proceed from here we need $q_0 = 1 - o(1)$. To prove this we need a lemma on the edge density of intersecting copies of H. We need a general version of this to prove (2.1) and we prove this here. Let

$$\theta_1 = \min_{H' \subset H}\Big(\frac{2s - \mu(H')}{2r - \nu(H')}\Big) - \frac{s}{r} > 0.$$

Note that $\theta_1 > 0$ follows from the fact that H is strictly balanced. A collection $H_1, H_2, \ldots, H_k$ of copies of H in $G_{n,m}$ is said to be **linked** if for each i there is $j \neq i$ such that H_i, H_j share an edge.

Lemma 2.1

Let $H_1, H_2, \ldots, H_k$, $k \geq 2$ be a linked collection of copies of H. Let $K = \bigcup_{i=1}^{k} H_i$. Then

$$\mu(K) \geq (\theta_1 + \tfrac{s}{r})\nu(K).$$

Proof

Assume w.l.o.g. that $H_i \not\subseteq \bigcup_{j \neq i} H_j$ for $i = 1,2,\ldots,k$. We prove the result by induction on k. We discuss the base case and the inductive step in tandem. Let $K' = \bigcup_{i=1}^{k-1} H_i$. Then

$$(2.6) \qquad \frac{\mu(K)}{\nu(K)} = \frac{\mu(H_k) + \mu(K') - |E(H_k) \cap E(K')|}{\nu(H_k) + \nu(K') - |V(H_k) \cap V(K')|}.$$

Furthermore

$$uv \in E(H_k) \cap E(K') \rightarrow u,v \in V(H_k) \cap V(K')$$

and so if $H' = (V(H_k) \cap V(K'), E(H_k) \cap E(K'))$

then H' is a non–trivial proper subgraph of H and, by (2.6)

$$\frac{\mu(K)}{\nu(K)} = \frac{s + \mu(K') - \mu(H')}{r + \nu(K') - \nu(H')}.$$

Base Case: k = 2

Here $K' = H_2$ and $\mu(K)/\nu(K) \geq \theta_1 + \frac{s}{r}$ follows from the definition of θ_1.

Inductive Step

Write

$$\frac{\mu(K)}{\nu(K)} = \frac{2s - \mu(H') + (\mu(K') - s)}{2r - \nu(H') + (\nu(K') - r)}$$

and observe that

$$(\mu(K') - s) - (\theta_1 + \tfrac{s}{r})(\nu(K') - r)$$

$$= (\mu(K') - (\theta_1 + \tfrac{s}{r})\nu(K')) + r\theta_1 > 0$$

by induction. □

It is always more pleasant to do computation in the independent model $G_{n,p}$, $p = m/N$, $N = \binom{n}{2}$. We quote the following simple results (see Bollobas [ ], Section 2.1). Let $\mathcal{A}$ be any property of graphs. Then

$$\text{(2.7)} \qquad \Pr(G_{n,m} \in \mathcal{A}) \leq 3m^{1/2} \Pr(G_{n,p} \in \mathcal{A})$$

and if $\mathcal{A}$ is monotone then

$$\text{(2.8)} \qquad \text{a.e. } G_{n,p} \in \mathcal{A} \rightarrow \text{a.e. } G_{n,m} \in \mathcal{A}.$$

Lemma 2.2

If

$$\text{(2.9)} \qquad \theta < \theta_1 r^2/(s^2 + \theta_1 rs)$$

then $q_0 = 1 - o(1)$.

Proof

If $G_{n,m}$ has a pair of edge intersecting copies of H then it contains a set of $k \leq 2r-1$ vertices which span at least $\lceil k(\frac{s}{r} + \theta_1)\rceil$ edges. Now this property is monotone and

$$\Pr(G_{n,p} \text{ contains a pair of edge intersecting copies of } H)$$

$$\leq \sum_{k=r}^{2r-1} \binom{n}{k} 2^{2r^2} p^{k(\frac{s}{r}+\theta_1)} \leq \sum_{k=r}^{2r-1} n^k 2^{2r^2} \left(\frac{\omega}{n^{r/s}}\right)^{k(\frac{s}{r}+\theta_1)}$$

$$= 0(n^{r\{\theta(\frac{s}{r}+\theta_1)-\frac{r}{s}\theta_1\}})$$

$$= o(1).$$

Now use (2.8). □

Referring to (2.5), suppose first that $0 \leq k \leq \lambda$. then for θ sufficiently small

(2.10)
$$1-o(1) \leq q_0 + \sum_{\ell=2}^{k} \frac{(k)_\ell}{\lambda^\ell} q_\ell \leq q_0 + \sum_{\ell=2}^{k} q_\ell \leq 1$$

Now let $k = (1+\epsilon)\lambda$ where $0 \leq \epsilon \leq \epsilon_1 = A_1(\log n)^{r/(2r-1)}/\lambda^{(r-1)/(2r-1)}$. Then, using (2.1)

$$u_\ell = \frac{(k)_\ell}{\lambda^\ell} q_\ell \leq 2\left(\frac{k}{\lambda}\right)^\ell e^{-\ell^2/2k} n^{-A_4\ell^{1/r}}$$

$$\leq 2\exp\{\epsilon\ell - \frac{\ell^2}{2k} - A_4\ell^{1/r}\log n\}.$$

Case 1: $\ell \geq 3\,\epsilon\lambda$

$$u_\ell \leq 2n^{-A_4\ell^{1/r}}$$

Case 2: $\ell < 3\,\epsilon\lambda$

$$u_\ell \le 2\exp\{\ell^{1/r}(\epsilon\ell^{1-1/r} - A_4 \log n)\}$$

$$\le 2\exp\{\ell^{1/r}(3^{1-1/r}\,\epsilon^{2-1/r}\,\lambda^{1-1/r} - A_4 \log n)\}$$

$$\le 2\exp\{\ell^{1/r}\log n(3^{1-1/r}A_1^{2-1/r} - A_4)\}.$$

So if we make A_1 small enough so that $A_4 \ge 4A_1^2$ then we have

$$u_\ell \le 2n^{-A_1^2\ell^{1/r}}$$

which is also valid for Case 1.

Hence if $\lambda \le k \le (1+\epsilon_1)\lambda$ and θ is sufficiently small

$$1 - o(1) \le q_0 + \sum_{\ell=2}^{k} \frac{(k)_\ell}{\lambda^\ell} q_\ell \le 1 + 2\sum_{\ell=2}^{\infty} n^{-A_1^2\ell^{1/r}}$$

$$= 1 + o(1).$$

This together with (2.10) proves the first part of the theorem.

Suppose now that $k = (1+\epsilon)\lambda$ where $1 \ge \epsilon \ge \epsilon_2 = A_2(\frac{\log n}{\lambda^{1-2/r}})^{r/2(r-1)}$.

Then by (2.5)

$$p_k/(\frac{e^{-\lambda}\lambda^k}{k!}) \geq \frac{1}{2} \frac{k!}{\lambda^{k-\lfloor\lambda\rfloor}\lfloor\lambda\rfloor!} q_{k-\lfloor\lambda\rfloor}$$

$$\geq A\, (\frac{k}{e\lambda})^k e^\lambda n^{-A_3(\epsilon\lambda+1)^{2/r}}$$

$$\geq Ae^{\epsilon^2\lambda/3} n^{-2A_3(\epsilon\lambda)^{2/r}}$$

$$= A \exp\{\frac{\epsilon^2\lambda}{3}(1 - 2A_3\epsilon^{\frac{2}{r}-2}\lambda^{\frac{2}{r}-1} \log n)\}$$

$$\geq A \exp\{\frac{\epsilon^2\lambda}{3}(1 - 2A_3A_2^{\frac{2}{r}-2})\}.$$

Now $\epsilon^2\lambda \to \infty$ and we are free to choose A_2 so that $1 - 2A_3A_2^{\frac{2}{r}-2} = \frac{1}{2}$ and the result is proved for this case.

When $k \geq 2\lambda$ we use

$$\frac{(k+1)!}{\lambda^s(k+1-s)!} q_s \geq \frac{k!}{\lambda^s(k-s)!} q_s$$

to reduce to the previous case.

□

§3. Proof of (2.1) and (2.2)

The upper bound in (2.1) follows fairly easily from Lemma 2.2. Indeed suppose $G_{n,m}$ contains exactly ℓ non–isolated copies of H. Let K denote the graph induced by the union of these copies. If K has ρ vertices then, by Lemma 2.2, it has at least $\tau\rho$ edges where

$\tau = \theta_1 + \frac{s}{r}$. Note that

$$\ell^{1/r} \leq \rho \leq r\ell \leq r\lambda_0$$

where the lower bound on ρ is from $(\rho)_r \geq \ell$. Hence, on using (2.7),

$$\begin{aligned} q_\ell &\leq 3m^{1/2} \sum_{\rho=\ell^{1/r}}^{r\ell} \binom{n}{\rho} \begin{bmatrix} \binom{\rho}{2} \\ \tau\rho \end{bmatrix} p^{\tau\rho} \\ &\leq 3m^{1/2} \sum_{\rho=\ell^{1/r}}^{r\ell} \left(\frac{ne}{\rho}\right)^\rho \left(\frac{\rho^2 e}{2\tau\rho} p\right)^{\tau\rho} \end{aligned}$$

$$\leq 3m^{1/2} \sum_{\rho=\ell^{1/r}}^{r\ell} \left(\frac{A\rho^{\tau-1}\omega^\tau}{n^{\tau r/s-1}}\right)^\rho \tag{3.1}$$

$$\leq 3m^{1/2} \sum_{\rho=\ell^{1/r}}^{r\ell} \left(\frac{A'\omega^{s(\tau-1)+\tau}(\log n)^{4(\tau-1)}}{n^{r\,\theta_1/s}}\right)^\rho$$

and the upper bound in (2.1) follows provided

$$\theta(s(\tau-1) + \tau) < r\theta_1/s.$$

It is convenient to stop and prove a similar inequality which is needed later. Let $\lambda_1 = \lfloor \omega^{rs}(\log n)^{4r+1} \rfloor$. It follows from (3.1) that provided

$$\theta(rs(\tau - 1) + \tau) < r\theta_1/s \tag{3.2}$$

that

$$(3.3) \qquad \sum_{\ell=\lambda_1}^{2\lambda_1} q'_\ell = o(e^{-2\lambda_0})$$

where q'_ℓ is the probability that $G_{n,2m}$ contains precisely ℓ non–isolated copies. Furthermore, if $G_{n,2m}$ contains more than $2\lambda_1$ non–isolated copies of H then we can choose λ_1 of them. For each chosen copy of H that does not share an edge with another chosen copy we choose a further copy that does share an edge. In this way we build a linked collection of between λ_1 and $2\lambda_1$ copies. It then follows by the calculations above that

$$(3.4) \qquad \sum_{\ell=2\lambda_1+1}^{\infty} q'_\ell = o(e^{-2\lambda_0}), \text{ also.}$$

To prove the lower bound of (2.1) we consider the probability of the existence of a collection of disjoint complete subgraphs of specific sizes. Thus let $\sigma_t = \binom{t}{r}\frac{r!}{\alpha}$ for $t \geq r$ and observe that K_t contains σ_t distinct copies of H. For a given a define $\tau = \tau(a)$ by $\sigma_{\tau+1} > a \geq \sigma_\tau$. Next let $\ell_1 = \ell$ and $\ell_{i+1} = \ell_i - \sigma_{\tau(\ell_i)}$ and $T_i = \sum_{j=1}^{i} \tau(\ell_j)$ for $i = 1,2,\ldots,k$ where $\ell_k \geq \frac{(r+1)!}{\alpha} > \ell_{k+1}$.

Now let $\mathcal{E}$ denote the event that

(3.5a) $G_{n,m}$ contains complete subgraphs with vertex set $[T_1],[T_2]\backslash[T_1],\ldots,[T_k]\backslash[T_{k-1}]$

and

(3.5.b) ℓ_{k+1} copies of H containing the edge {1,2} but otherwise disjoint from all other copies. Let their vertices belong to $[T]\backslash[T_k]$ where $T - T_k = (r-2)\ell_{k+1}$

and

(3.5c) there are no other edges in [T] (this assumption simplifies the calculations but may be a bit drastic!)

and

(3.6) there are no other non–isolated copies of H is $G_{n,m}$.

Thus if $\mathscr{E}$ occurs then $G_{n,m}$ contains exactly ℓ non–isolated copies of H. We can write

$$Pr(\mathscr{E}) = \pi_1\pi_2$$

where

$$\pi_1 = Pr((3.5)) \text{ and } \pi_2 = Pr((3.6)|(3.5)).$$

But

$$\pi_1 = \begin{bmatrix} (N - \binom{T}{2}) \\ m - u \end{bmatrix} / \binom{N}{m} = (\tfrac{m}{N})^u \left(1 - 0(\tfrac{T^4}{N} + \tfrac{u^2}{m})\right)$$

where $u = \sum_{i=1}^{k} \begin{bmatrix} \tau(\ell_i) \\ 2 \end{bmatrix} + (s-1)\ell_{k+1}$. So

$$\pi_1 = \left(\frac{\omega}{n^{r/s}}\right)^u \left(1 - 0\left(\frac{T^4}{n} + \frac{u^2}{m} + \frac{u}{n}\right)\right)$$

(3.7) $$= \left(\frac{\omega}{n^{r/s}}\right)^u (1 - o(1)),$$

since we show later that

(3.8) $$\sum_{i=1}^{k} \tau(\ell_i)^x = 0(\ell^{x/r}) \qquad \text{for any fixed positive integer } x,$$

and we assume

(3.9) $$\theta < r(2 - \tfrac{r}{s})/4s.$$

We show next that $\pi_2 = 1 - o(1)$. Note that (3.6) given (3.5)) is monotone and so we can use the $G_{n,p}$ model to estimate π_2. Now by the FKG inequality

$$\pi_2 \geq \pi_2' \, \pi_2''$$

where

$$\pi_2' = \Pr\,(\text{there are no non–isolated copies of } H \text{ in } [n]\backslash[T])$$

and

$\pi_2'' \;=$ Pr(there are no extra copies of H which share an edge with those defined in (3.5)).

Now $\pi_2' = 1 - o(1)$ if (2.9) holds and

$$\pi_2'' \geq 1 - E \text{ (number of such copies of H)}$$

$$\geq 1 - \sum_{H' \subset H} \binom{n}{r-\nu(H')} 2^r p^{s-\mu(H')} \left(\sum_{i=1}^{k} \begin{bmatrix} \tau(\ell_i) \\ \nu(H') \end{bmatrix} \nu(H')! + 0(1) \right)$$

$$= 1 - 0\Big(\sum_{H' \subset H} n^{r-\nu(H')} \frac{\omega^{s-\mu(H')}}{n^{r - \frac{r}{s}\mu(H')}} \ell^{\nu(H')/r} \Big)$$

on using (3.8) to simplify the second summation

$$= 1 - o(1)$$

provided

$$\theta < \min_{H' \subset H} \frac{\nu(H') - \frac{r}{s}\,\mu(H')}{s - \mu(H') + \nu(H')\frac{s}{r}}. \tag{3.10}$$

The proof of (2.1) is completed once we have proved (3.8). For then (3.7) implies

$$\pi_1 \geq (\frac{\omega}{n^{r/s}})^{0(\ell^{2/r})}(1-o(1)).$$

Proof of (3.8)

When a is large we have, where $\tau = \tau(a)$,

$$\begin{aligned} a - \sigma_\tau &\leq \sigma_{\tau+1} - \sigma_\tau \\ &= r(\tau-1)_{r-1}\alpha^{-1} \\ &< r\tau^{r-1}. \end{aligned}$$

But

$$\begin{aligned} a \geq \sigma_\tau &\rightarrow \binom{\tau}{r} \leq a \\ &\rightarrow (\frac{\tau}{r})^r \leq a \end{aligned}$$

(3.11) $$\rightarrow \tau \leq ra^{1/r}$$

and so

$$a - \sigma_\tau \leq r^r\, a^{1-1/r}$$

which implies

$$\ell_i \le r^{ir}\ell^{(1-\frac{1}{r})^i} \qquad 1 \le i \le k \tag{3.12}$$

and

$$\tau(\ell_i) \le r^{(i+1)r}\,\ell^{(1-\frac{1}{r})^i/r}.$$

Now let $i_0 = \lceil r \log r \rceil$ and assume ℓ is large enough that $i_0 \le k$ ((3.8) is trivial for bounded ℓ). Then (3.12) implies

$$\ell_{i_0} \le A\ell^{1/r} \tag{3.13}$$

where $A = r^{i_0 r}$.

Now $\tau(\ell_1) \le r\ell^{1/r}$ and τ is monotone increasing and so

$$\sum_{i=1}^{i_0} \tau(\ell_i)^x \le i_0 r^x \ell^{x/r}. \tag{3.14}$$

On the other hand it is easy to see that

$$\sigma_\tau \ge \tau \quad \text{for } \tau \ge r+1$$

and thus

$$\ell \quad = (\ell_1 - \ell_2) + (\ell_2 - \ell_3) + \ldots + (\ell_k - \ell_{k+1}) + \ell_{k+1}$$

$$= \sigma_{\tau(\ell_1)} + \sigma_{\tau(\ell_1)} + \sigma_{\tau(\ell_2)} + \dots + \sigma_{\tau(\ell_k)} + \ell_{k+1}$$

$$\geq \tau(\ell_1) + \tau(\ell_2) + \dots + \tau(\ell_k)$$

and so replacing ℓ by ℓ_{i_0} above

$$\tau(\ell_{i_0+1}) + \dots + \tau(\ell_k) \leq \ell_{i_0+1}.$$

Hence

$$(3.15) \qquad \sum_{i=i_0+1}^{k} \tau(\ell_i)^x \leq \Big(\sum_{i=i_0+1}^{k} \tau(\ell_i)\Big)^x$$

$$\leq \ell_{i_0+1}^{\;x}$$

$$= 0(\ell^{x/r}) \quad \text{by (3.13).}$$

(3.8) follows from (3.14) and (3.16) and this completes the proof of (2.1).

We now turn to the proof of (2.2). For positive integer t

$$\Pr(\exists\ t \text{ isolated copies of } H \text{ in } G_{n,p}) \leq \frac{1}{t!}\binom{n}{r}^t\left(\frac{r!}{\alpha}\right)^t p^{ts}$$

$$\leq \left(\frac{e}{t} \cdot \frac{n^r}{r!} \cdot \frac{r!}{\alpha} \cdot p^s\right)^t$$

$$\le \left(\frac{3\omega^s}{t\alpha}\right)^t.$$

Now put $t = \lambda_0$ and apply (2.7).

The same argument gives

$$\Pr(G_{n,2m} \text{ contains at least } \lambda_1 \text{ isolated copies}) = o(e^{-2\lambda_0})$$

and so, using (3.3), (3.4), we find

(3.17) $$\text{Pt}(G_{n,2m} \text{ contains } 2\,\lambda_1 \text{ or more copies of } H) = o(e^{-2\lambda_0}).$$

§4. Proof of (2.3)

This section contains the main ideas of the proof of Theorem 1.1

Let $\mathscr{A}_{k\ell} = \{G \in \mathscr{G}_{n,m}\colon$ G has k isolated copies and ℓ non–isolated copies of H$\}$. Let $a_{k\ell} = |\mathscr{A}_{k,\ell}|$ so that (2.3) is actually concerned with the ratio $a_{k,\ell}/a_{k-1,\ell}$.

Now for $k > 0$, $\ell \ge 0$, let $BP_{k,\ell}$ denote the bipartite graph with vertex partition $\mathscr{A}_{k,\ell}$, $\mathscr{A}_{k-1,\ell}$ and edge set $\mathscr{E}_{k,\ell}$ where $G_1G_2 \in \mathscr{E}_{k,\ell}$, $G_1 \in \mathscr{A}_{k,\ell}$, $G_2 \in \mathscr{A}_{k-1,\ell}$ if the edge sets of G_1, G_2 are related by

$$E(G_2) = (E(G_1)\backslash\{e\}) \cup \{f\}$$

where e is an edge of some isolated copy of H in G_1 and f is some edge which does not create a new copy of H when added to G_1/e.

If $G \in \mathscr{A}_{k,\ell} \cup \mathscr{A}_{k-1,\ell}$ let $d(G)$ denote its degree in $BP_{k,\ell}$. Then

(4.1) $G \in \mathscr{A}_{k,\ell}$ implies

$$ks(N - m - \xi(G)) \le d(G) \le ks(N - m)$$

where $\xi(G)$ = the number of copies in G of a graph of the form $H - x$ for some edge $x \in E(H)$.

This is because we have ks choices for edge e in an isolated copy of H. Then of the $N - m$ possible edge replacements f there are at most $\xi(G-e)-1$ choices which create a new H when added. Finally observe that $\xi(G-e)-1 \le \xi(G)$.

Also

(4.2) $G \in \mathscr{A}_{k-1,\ell}$ implies

$$(m - s(k+\ell))(\xi(G) - 2\zeta(G)) \le d(G) \le m\xi(G)$$

where $\zeta(G)$ = the number of subgraphs of G of the form $(H_1 \cup H_2) - x$ where H_1, H_2 are copies of H which share x (so if e.g. H is a triangle then $(H_1 \cup H_2) - x$ must be a 4–cycle).

To see this we overestimate the number of choices of f by m and the number of choices of e by $\xi(G)$. To underestimate $d(G)$ we underestimate the number of choices of f by $m - s(k+\ell)$ since we do not wish to touch a copy of H. The number of choices for e, given f, is at least $\xi(G-f) - \zeta(G) \ge \xi(G) - 2\zeta(G)$ (crudely.)

The equation

$$\sum_{G \in \mathscr{A}_{k,\ell}} d(G) = \sum_{G \in \mathscr{A}_{k-1,\ell}} d(G)$$

and (4.1), (4.2) lead to

(4.3) $$\frac{(m-s(k-\ell))(\xi_{k-1,\ell} - 2\,\bar{\zeta}_{k-1,\ell})}{ks(N-m)} \le \frac{a_{k,\ell}}{a_{k-1,\ell}} \le \frac{m\bar{\xi}_{k-1,\ell}}{k\,s\,(N-m-\bar{\xi}_{k,\ell})}$$

where $\bar{\xi}_{k,\ell}$, $\bar{\zeta}_{k,\ell}$ denote the expectations of $\xi(G)$, $\zeta(G)$ over $\mathscr{A}_{k,\ell}$. It only remains now to estimate these quantities. For $G \in \mathscr{A}_{k,\ell}$ and $e \in E(\bar{G})$ ($\bar{G}$ = complement of G) let h_e denote the number of new copies of H created when e is added to G. Let $\mathscr{N}(G) = \{e \in E(\bar{G}) : h_e > 0\}$ and $\mathscr{N}(G) = |\mathscr{N}(G)|$. Let λ_1 be as in (3.2).

Lemma 4.3

Let $G = G_{n,m}$.

(a) $\Pr(\exists\, e \in E(\bar{G})\colon h_e \ge 2\lambda_1) = o(n^2 e^{-2\lambda_0})$.

(b) $\Pr(\eta(G) \ge n^{r/s}\lambda_1 \log n) = o(e^{-2\lambda_0})$.

Proof

Let $\mathscr{E}$ denote the event $\{G_{n,2m}$ has at least $2\lambda_1$ copies of $H\}$. Think of $G_{n,2m}$ as $G_{n,m}$ plus m random edges.

(a)

Let $\mathscr{E}_a = \{\exists\, e \in E(\bar{G})$ s.t. $h_e \ge 2\lambda_1\}$. Then

$$\begin{aligned}\Pr(\mathscr{E}) &\ge \Pr(\mathscr{E}\,|\,\mathscr{E}_a)\Pr(\mathscr{E}_a)\\ &\ge \tfrac{m}{N}\Pr(\mathscr{E}_a).\end{aligned}$$

Part (a) now follows from (3.17).

(b)

Let $\lambda_2 = n^{r/s}\lambda_1 \log n$ and $\mathcal{E}_b = \{\eta(G) \geq \lambda_2\}$. Then

$$\Pr(\mathcal{E}) \geq \Pr(\mathcal{E} \mid \mathcal{E}_b)\Pr(\mathcal{E}_b)$$

and (b) follows if we show that $\Pr(\mathcal{E} \mid \mathcal{E}_b) \geq \frac{1}{2}$. But to see this observe that the expected number of copies of H created by adding the second m edges is at least $\frac{m}{N}\eta(G_{n,m})$ and

$$\frac{m}{N}\lambda_2 \approx \omega\lambda_1 \log n$$

$$>> \lambda_1.$$

Note that we see now that the actual number added, given $\mathcal{E}_b$, majorizes a binomial with mean $>> \lambda_1$.

□

Let us now return to the consideration of (4.3). Suppose $\ell \leq \lambda_0$. It follows from (2.1) and (2.2) that there exists k_0 such that

$$\pi_{k_0,\ell} \geq n^{-A_3\ell^{2/r}}(2\lambda_0)^{-1}.$$

We prove that

$$\pi_{k,\ell} \geq \left(1 - \frac{1}{\lambda_0}\right)^{|k-k_0|} n^{-A_3\ell^{2/r}}(2\lambda_0)^{-1} \qquad 0 \leq k \leq \lambda_0. \tag{4.13}$$

This is true for $k = k_0$ and assume inductively that it is true for some $0 < k \le k_0$. $k > k_0$ will be dealt with subsequently and this is why we are assuming that $k_0 > 0$. We will be able to verify (2.3) as we proceed with the induction. We will estimate $\xi_{k,\ell}$, $\zeta_{k,\ell}$ by the same method and to do this we let Γ denote a generic graph of the form $H - x$ or $H_1 \cup H_2 - x$. Let Γ_0 denote some fixed copy of Γ with vertex set $\{1,2,...,t\}$, $t = \nu(\Gamma)$ and let $e_1,e_2,...,e_u$, $u = \mu(\Gamma)$ be an enumeration of its edges.

Let $\mathscr{A}^*_{k,\ell} = \{G \in \mathscr{A}_{k,\ell}$: for $i = 1,2,...,u$ we have either (i) $e_i \in E(G)$ and e_i does not lie in any copy of H or (ii) $e_i \notin E(G)$ and $e_i \notin \mathscr{N}(G)\}$.

Lemma 4.4

$$1 - \frac{1}{N} \ge \frac{|\mathscr{A}^*_{k,\ell}|}{|\mathscr{A}_{k,\ell}|} \ge 1 - \frac{2s\lambda_2}{N}).$$

Proof

By symmetry, we have

$$\frac{sk}{N} \le 1 - \frac{|\mathscr{A}^*_{k,\ell}|}{|\mathscr{A}_{k,\ell}|} \le E_{k,\ell}\left(\frac{(2s-1)(\eta(G) + s(k+\ell))}{N}\right)$$

where $E_{k,\ell}$ denotes expectation over G in $\mathscr{A}_{k,\ell}$ (4.13) and Lemma 4.3(b) imply that $E_{k,\ell}(\eta(G)) \le (1 + \frac{1}{2s})\lambda_2$ and the result follows.

□

So now let $\mathscr{A}^*_{k,\ell,i} = \{G \in \mathscr{A}^*_{k,\ell} : E(G) \cap \{e_1,...,e_u\} = \{e_1,...,e_i\}\}$ for $0 \le i \le u$ and consider the bipartite graph $BP^*_{k,\ell,i}$, $i \ge 0$, with bipartition $\mathscr{A}^*_{k,\ell,i}$, $\mathscr{A}^*_{k,\ell,i-1}$ and an edge G_1G_2 for $G_1 \in \mathscr{A}^*_{k,\ell,i}$, $G_2 \in \mathscr{A}^*_{k,\ell,i-1}$ if G_2 can be obtained from G_1 by deleting e_i and adding a new edge f. Using d to denote degree in $BP^*_{k,\ell,i}$ we have

(4.14) $G \in \mathscr{A}^*_{k,\ell,i}$ implies

$$N - m - \eta(G) \le d(G) \le N - m.$$

There are at most $N - m$ choices for f which gives the upper bound. On the other hand, if $f \notin E(G) \cup \eta(G)$ then $G - e_i + f \in \mathscr{A}^*_{k,\ell,i-1}$. To see this we first note that $G + f$ has the same $k + \ell$ copies of H as G. But then if $e_i \notin \mathscr{N}(G - e_i + f)$ we find that e_i belongs to a copy of H in $G + f$ and hence in G, which is disbarred by $G \in \mathscr{A}^*_{k,\ell}$

(4.15) $G \in \mathscr{A}^*_{k,\ell,i-1}$ implies

$$m - s(k+\ell) \le d(G) \le m.$$

There are at most m choices for f and if we choose to delete an f which is not in any copy of H then $G + e_i - f$ is in $\mathscr{A}^*_{k,\ell,i}$. The latter fact following from $e_i \notin \mathscr{N}(G)$.

Hence if $a^*_{k,\ell,i} = |\mathscr{A}^*_{k,\ell,i}|$ we have, analogously to (4.3),

$$\frac{m-s(k+\ell)}{N} \le \frac{a^*_{k,\ell,i}}{a^*_{k,\ell,i-1}} \le \frac{m}{N - m - \bar{\eta}_{k,\ell,i}}. \tag{4.16}$$

It follows from (4.13) and Lemma 4.4 that there exists i_0 such that

$$a^*_{k,\ell,i_0} \ge \frac{1}{6}\lambda_0^{-1} n^{-A_3\lambda_0^{2/r}} \binom{N}{m}.$$

Now (4.16) implies that $a^*_{k,\ell,i}/a^*_{k,\ell,i-1} \ge \frac{m}{2N}$ and so if $i > i_0$

$$a^*_{k,\ell,i} \geq \frac{1}{6}\lambda_0^{-1}\, n^{-A_3\lambda_0^{2/r}} \binom{N}{m}\left(\frac{m}{2N}\right)^{i-i_0}$$

and hence we see from Lemma 4.3(b) that $\bar{\eta}_{k,\ell,i} \leq 2\lambda_2$ for $i \geq i_0$. But this then implies that for $i > i_0$

$$(4.17) \qquad \left(1 - \frac{2s\lambda_0}{m}\right)\frac{m}{N} \leq \frac{a^*_{k,\ell,i}}{a^*_{k,\ell,i-1}} \leq \left(1 + \frac{3(m+\lambda_2)}{N}\right)\frac{m}{N}.$$

But if $i_0 \geq 1$ we see from (4.21) that $a^*_{k,\ell,i_0-1} \geq \frac{m}{2N} a^*_{k,\ell,i_0}$. This puts a bound of $2\lambda_2$ on $\bar{\eta}_{k,\ell,i_0-1}$ and proves (4.18) for $i = i_0$. Clearly we can repeat this argument a further $i_0 - 1$ times to show that (4.17) holds for $i \geq 1$.

It follows that

$$(4.18) \qquad \Pr(G \text{ contains } \Gamma_0 \mid G \in \mathcal{A}^*_{k,\ell}) = \left(\frac{m}{N}\right)^u (1 + \epsilon_{k,\ell,\Gamma})$$

where $|\epsilon_{k,\ell,\Gamma}| \leq A\omega^{rs}/n^{2-r/s}$.

Let us now deal with ξ. Let Λ_ξ denote the set of possible graphs of the form $H - x$. Then, from (4.18),

$$(4.19) \qquad E(\xi(G) \mid G \in \mathcal{A}^*_{k,\ell}) = \sum_{\Gamma\in\Lambda_\xi} \binom{n}{r}\frac{r!}{\alpha_\Gamma}\left(\frac{m}{N}\right)^{s-1}(1 + \epsilon_{k,\ell,\Gamma})$$

where α_Γ = the number of automorphisms of Γ.

To handle $E(\xi(G) \mid G \in \mathcal{A}_{k,\ell} - \mathcal{A}^*_{k,\ell})$ we note that for such G, (4.20) $$\xi(G) \le \sum_{e\in E(\bar{G})} h_e + s(k+\ell)$$

$$\le 2\lambda_1\eta(G) + n^r|\{e \in E(\bar{G}): h_e > 2\lambda_1\}| + s(k+\ell).$$

It follows now from Lemmas 4.3 and 4.4 that

(4.21) $$E(\xi(G)\,|\,G \in \mathscr{A}_{k,\ell} - \mathscr{A}^*_{k,\ell}) \le 3\lambda_1\lambda_2.$$

Lemma 4.4, (4.19) and (4.21) then imply that

$$\bar{\xi}_{k,\ell} = \omega^{s-1}n^{r/s} \sum_{\Gamma\in\Lambda_\xi} \frac{1}{\alpha_\Gamma}(1 + \epsilon_{k,\ell,\Gamma})$$

where $\epsilon_{k,\ell,\Gamma}$ now satisfies, $|\epsilon_{k,\ell,\Gamma}| \le A\omega^{3rs-s+1}/n^{2-r/s}$.

Before looking at ζ observe that

$$\sum_{\Gamma\in\Lambda_\xi} \frac{r!}{\alpha_\Gamma} = \frac{sr!}{\alpha},$$

since we obtain all copies of graphs of the form $H - x$ in K_r by taking all copies of H and deleting an edge. Thus we can write

(4.21) $$\bar{\xi}_{k,\ell} = \frac{s\omega^{s-1}}{\alpha} n^{r/s}(1 + \epsilon_{k,\ell})$$

where $|\epsilon_{k,\ell}| \le A\omega^{3rs-s+1}/n^{2-r/s}$.

Analogously to (4.19) we have

$$E(\zeta(G) \mid G \in \mathscr{A}^*_{k,\ell}) = \sum_{\Gamma \in \Lambda_\zeta} \binom{n}{\nu(\Gamma)} \frac{r!}{\alpha_\Gamma} \left(\frac{m}{N}\right)^{\mu(\Gamma)} (1 + \epsilon_{k,\ell,\Gamma}) \tag{4.22}$$

where Λ_ζ denotes the set of possible graphs of the form $H_1 \cup H_2 - x$.

Lemma 4.5

$\Gamma \in \Lambda_\zeta$ implies $\frac{r}{s}(\mu(\Gamma) + 1) - \nu(\Gamma) \geq 1 + \frac{r\theta_1}{s}$.

Proof

If $\Gamma = H_1 \cup H_2 - x$ let $H' = H_1 \cap H_2$. Then

$$\mu(\Gamma) = 2s - \mu(H') - 1$$

and

$$\nu(\Gamma) = 2r - \nu(H').$$

The result now follows from the definition of θ_1.

□

It follows from (4.22) and Lemma 4.5 that

$$E(\zeta(G) \mid G \in \mathscr{A}^*_{k,\ell}) \leq A\omega^{2s-1} n^{r/s(1-\theta_1)}. \tag{4.23}$$

For $G \in \mathscr{A}_{k,\ell} - \mathscr{A}^*_{k,\ell}$ we write, analogously to (4.20)

$$\zeta(G) \le \sum_{e \in E(\bar{G})} \binom{h_e}{2} + 2s \binom{\ell}{2}$$

$$\le 2\lambda_1^2 \eta(G) + n^{2r} |\{e \in E(\bar{G}) : h_e > 2\lambda_1\}| + s\ell^2.$$

It now follows from Lemmas 4.3 and 4.4 that

$$E(\zeta(G) \mid G \in \mathscr{A}_{k,\ell} - \mathscr{A}^*_{k,\ell}) \le 3\lambda_1^2 \lambda_2.$$

Combining this with (4.23) and $\theta_1 \le \frac{1}{r}$ and using Lemma 4.4 we obtain

(4.24) $$\bar{\zeta}_{k,\ell} \le A\omega^{2s-1} n^{r(1-\theta_1)/s}.$$

Remark: the above analysis, between here and (4.13) could equally well have been done with (4.13) replaced by $\pi_{k,\ell} \ge e^{-\lambda_0}$. This would lead to slightly larger "hidden" constants A.

Now (4.3) implies

(4.25) $$a_{k-1,\ell} \ge a_{k,\ell} \frac{ks\,(N-m-\bar{\xi}_{k,\ell})}{m\ \bar{\xi}_{k-1,\ell}}.$$

But clearly $\bar{\xi}_{k-1,\ell} \le n^r$ and so, using (4.13), $\pi_{k-1,\ell} \ge e^{-\lambda_0}$ and by the above remark (4.21) and (4.24) hold with k replaced by $k-1$. But using these estimates now in (4.3) gives

(4.26) $$\frac{a_{k,\ell}}{a_{k-1,\ell}} = \frac{\lambda}{k}(1 + \beta_{k,\ell})$$

where, $|\beta_{k,\ell}| = 0(\omega^s n^{-\theta_1 r/s} + \omega^{3rs-s+1} n^{rs-2}) = o(\lambda_0^{-1})$ provided

(4.27) $$\theta < \min\{\frac{r\theta_1}{2s^2}, \frac{2s-r}{s(3rs+1)}\}.$$

Note that (4.26) = (2.3) and that this completes the inductive step in the proof of (4.13) for $k \leq k_0$. For $k > k_0$ the only thing that changes is that we replace (4.23) by

$$a_{k+1,\ell} \geq \frac{(m-s(k+\ell))(\bar{\xi}_{k,\ell} - 2\bar{\zeta}_{k,\ell})}{ks(N-m)} a_{k,\ell}$$

which enables to use (4.21), (4.24) with k replaced by $k+1$. The rest is as before. This completes the proof of (2.3) and the theorem.

Remark: we have identified 5 upper bounds (2.9), (3.2), (3.9), (3.10) and (4.27). It turns out that (2.9) and (3.9) are implied by the others.

References

[1] B. Bollobás, "Threshold functions for small subgraphs", *Math. Proc. Camb. Phil. Soc.* 90 (1981) 197–206.

[2] R. Boppana and J. Spencer, "A useful correlation inequality", *J. Combinatorial Theory A* 50 (1989) 305–307.

[3] P. Erdös, an A. Rényi, "On the evolution of random graphs", *Publ. Math. Inst. Hung. Acad. Sci.* 5 (1960) 17–61.

[4] S. Janson, T. Łuczak and A. Rucinski, "An exponential bound for the probability of nonexistence of a specified subgraph in a random graph", in *Random Graphs* '87 (to appear).

[5] S. Janson, "Poisson approximation for large deviations", to appear.

[6] M. Karonski and A. Rucinski, "On the number of strictly balanced subgraphs of a random graph", *Graph Theory Lagów* 1981, *Lecture Notes in Mathematics* 1018 Springer–Verlag (1983) 79–83.

[7] A. Rucinski, "Small subgraphs of random graphs (a survey)", to appear.