Combinatorics, Probability and Computing (2001) 10, 239–249. Printed in the United Kingdom
© 2001 Cambridge University Press

Random Finite Topologies and their Thresholds†

C. F. MACLEAN[1]§ and NEIL O'CONNELL[2]

[1] Ecole Normale Supérieure, 45 rue d'Ulm, 75005 Paris, France
(e-mail: maclean@clipper.ens.fr)

[2] BRIMS, Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS34 8QZ, England
(e-mail: noc@hplb.hpl.hp.com)

Received 16 March 1998; revised 12 October 2000

This paper is dedicated to Phillip Heyes in gratitude.

For each integer n, there is a natural family of probability distributions on the set of topologies on a set of n elements, parametrized by an integer variable, m. We will describe how these are constructed and analysed, and find threshold functions (for m in terms of n) for various topological properties; we focus attention on connectivity and the size of the largest component.

1. Introduction and summary

A choice of m random subsets of the set $[n]$ is a random function

$$f : [m] \mapsto \wp([n]),$$

where $\wp([n])$ denotes the power set of $[n]$. There is a one-to-one correspondence C between functions $f : [m] \mapsto \wp([n])$ and functions $g : [n] \mapsto \wp([m])$ given by

$$i \in f(j) \Leftrightarrow j \in g(i).$$

Throughout, f is a random function from $[m]$ to $\wp([n])$ and g is $C(f)$. We let $T_m(n)$ denote the topology generated by f (the smallest topology containing the sets $f(i)$, $i \leqslant m$).

It is well known that a topology T on a finite set S is equivalent to:

(1) an equivalence relation, $R(T)$ on S, and
(2) a partial order $P(T)$ on the equivalence classes of R.

† An earlier version of this article appeared as BRIMS Technical Report HPL-BRIMS-9718.
§ Based on work completed as a summer student at BRIMS, 1997.

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

A subset of S is open if, and only if, it is a collection of equivalence classes, closed under ascending P. For $T_m(n)$, it is easily seen that

$$i =_R j \Leftrightarrow g(i) = g(j);\ [i] <_P [j] \Leftrightarrow g(i) \subset g(j).$$

In this paper we study the asymptotic properties of the random topology $T_m(n)$. Before stating the main results, we will first establish some context. The first thing to point out is that our model for a random topology induces what seems to be a natural model for a random partial order: although there are many standard models for random partial orders, this one does not seem to have been studied before. For an overview of the literature on random partial orders, see the survey of Brightwell [3]. However, the *Boolean lattice* model, which has been studied extensively by Kreuter [8], Kohayakawa and Kreuter [5] and Kohayakawa, Kreuter and Osthus [6], is related to this one in much the same way as the random graph $G(n, p)$ is related to the random graph on n vertices with exactly $\binom{n}{2}p$ edges. The properties considered for the Boolean lattice model are a little obscure from a topological perspective; similarly, the properties we consider here are a little obscure from the perspective of the study of random partial orders.

A slightly more tenuous connection can be made by noting that $T_m(n)$ is induced by a random mapping $g : [n] \mapsto \wp([m])$. There is a large literature on random mappings: see for example [7]. However, the properties of g which will prove to be important in the study of random topologies are somewhat unusual and are not very well motivated in the context of random mappings. For example, the property of connectedness depends on the order structure of $\wp([m])$.

We will begin our discussion of $T_m(n)$ by noting that we can immediately determine a threshold for discreteness by appealing to a classical result, due to Rényi [9] (in answer to a question of Erdős), on the minimum cardinality of a collection of random subsets of $[m]$ that guarantees, with high probability, two distinct sets A and B with $A \subset B$: if $n^2 = \lambda(4/3)^m(1 + o(1))$ then

$$\lim_{n\to\infty} P(T_m(n) \text{ is discrete}) = e^{-\lambda}.$$

We can also obtain, by elementary considerations, a threshold for the property 'distinguishability', namely, if $n^2 = \lambda 2^{m+1}(1 + o(1))$ then

$$\lim_{n\to\infty} P(T_m(n) \text{ distinguishes points}) = e^{-\lambda}.$$

We define the function $\phi(m)$ to be equal to 2 if m is even and $\frac{3}{\sqrt{2}}$ otherwise.

The main results of this paper are as follows.

Theorem 1.1. *Let*

$$n = \frac{2^{m/2}}{\phi}\left(\frac{m}{2}\log 2 + \frac{1}{2}\log m + \lambda_n\right). \tag{1.1}$$

Then the random topology is almost surely connected if $\lambda_n \to \infty$.

Theorem 1.2. *If*

$$n = \frac{2^{m/2}}{\phi}\left(\frac{m}{2}\log 2 + \frac{1}{2}\log m - \lambda_n\right) \tag{1.2}$$

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

then the random topology is almost surely disconnected if $\lambda_n \to \infty$. *If* $\kappa\sqrt{\frac{2}{\pi}\frac{\log 2e^{\lambda_n}}{2\phi}} \to \chi$, *then in the limit,* $(C-1)$ *has a Poisson distribution with mean* χ, *where* C *is the number of components of the topology, and* κ *is* 1 *if* m *is even,* 2 *otherwise.*

Theorem 1.3. *If* $n = 2^{m/2-\frac{\lambda}{2}\sqrt{m}}$, *then the relative size of the largest component of the induced topology,* $X_m(n)$, *converges in probability to*

$$\int_{|x|\geqslant\lambda} e^{-x^2}/\sqrt{2\pi x}\,dx.$$

Proofs are given in Sections 2–4; in Section 3 we also compare these results with their analogues for random graphs (exploring these comparisons further seems to be an interesting topic for future research).

The threshold implicit in the statement of Theorems 1.1 and 1.2 essentially boils down to the question: 'Are there any elements i, with $|g(i)|$ equal to $m/2$ if m is even or $(m\pm1)/2$ if m is odd, which are isolated in the partial order?' This is why there is a parity effect. Theorem 1.3 is true because most i with $|g(i)|$ within $2\lambda\sqrt{m}$ of $m/2$ are isolated, and almost all other elements are in the giant component.

Finally, we introduce some conventions which will be in force throughout. Given two functions of n, $g(n)$ and $f(n)$, we will use $f \approx g$ to mean

$$\exists \text{ a constant, } c, \text{ such that } c \neq 0, \infty \text{ and } \lim_{n\to\infty}\frac{f(n)}{g(n)} = c.$$

When c is 1, we will use $\sim$ rather than $\approx$. We say that a property, Q $(= Q_n)$ of the topology $T_m(n)$, holds asymptotically almost surely (a.a.s.) if $P(Q_n) \to 1$. Throughout, an event E_n is said to happen 'almost surely' or 'almost certainly' if it happens a.a.s.

We begin by examining the component structure of the random partial order $P(T_m(n))$, in the range of n and m we are interested in.

2. Component structure of $P(T_m(n))$

A topology, T, is connected if and only if $P(T)$ is connected. We need to show that, given (1.1), a random $g : [n] \mapsto \wp([m])$ a.s. induces a connected partial order.

In the range of n and m that interests us, the component structure of the partial order $P(T_m(n))$ has the following property.

There exists a giant component, C, of the comparison graph, and all elements of $[n]$ are either points of C or are comparable with no other element of $[n]$. In the latter case we will say that the point is a singleton.

How will we show this? The number of subsets of $[m]$, comparable with $g(i)$ grows exponentially with $\big||g(i)| - m/2\big|$. If $g(i)$ is unusually large or small, i is likely to be comparable with many points in $P(T_m(n))$. We show that, almost surely, there exists such an i, and then use it as a 'hub', out of which we will 'grow' a component. Initially, we will show that this component contains any i such that $g(i)$ is large or small. But, if $|g(j)| \approx \frac{m}{2}$, and $g(j)$ is comparable with $g(k)$, then $g(k)$ is almost certainly either large or small. We will then have proved the following theorem.

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

Theorem 2.1. *Let E_n be the event that there exists a component, C, such that all elements of $[n]$ are either in C or are singletons. For all $\mu < \frac{5}{2}$,*

$$\lim_{m\to\infty} \inf_{n \geqslant \frac{2^{m/2}}{\mu}(\frac{m}{2}\log 2)} P(E_n) = 1.$$

We start with the following lemma.

Lemma 2.2. *There exist $p < \sqrt{2}$ and $\eta > \frac{3}{4}$ such that*

$$\lim_{m\to\infty} \inf_{n \geqslant p^m} P \bigcup_{i=1}^{n} \{|g(i)| > \eta m\} = 1.$$

Proof. Stirling's formula tells us $\binom{m}{\eta m} \sim q^m / f(\sqrt{m})$, where $q = \left(\frac{1}{\eta}\right)^{\eta}\left(\frac{1}{1-\eta}\right)^{1-\eta}$ and f is a polynomial. Since $q(3/4) > \sqrt{2}$, we may choose η such that $\eta > 3/4$ and $q(\eta) > \sqrt{2}$. Consider $r(\eta)$, the probability that $|g(i)| > \eta m$; we have

$$r(\eta) \geqslant \frac{\binom{m}{\eta m}}{2^m} \sim \left(\frac{q(\eta)}{2}\right)^m / (f(\sqrt{m})).$$

Provided that p is such that $pq(\eta) > 2$, we then have

$$\lim_{m\to\infty} \inf_{n : n > p^m} r(\eta) n = \infty.$$

So, for any real number, x, eventually the probability that no i satisfies $|g(i)| > \eta m$ is $\leqslant (1 - \frac{x}{n})^n$. But this last converges to e^{-x}. The lemma follows. □

From now on, we fix some $\eta > \frac{3}{4}$. Of course, with probability 1, there exists some i such that $|g(i)| < (1-\eta)m$. We now start 'growing' a component out from i.

Lemma 2.3. *Let E_n be the following event: there exist i, j, in different components of our topology, such that $|g(i)| \geqslant \eta m$ and $|g(j)| \geqslant \gamma m$. Then there exist $p < \sqrt{2}$ and a $\gamma < 3/4$ such that*

$$\lim_{m\to\infty} \sup_{n : n > p^m} P(E_n) = 0.$$

Proof. We show that, almost surely, for every such pair i, j there exists k such that $g(k)$ is contained in $g(i) \cap g(j)$.

Choose γ such that $\gamma < 3/4$, $\gamma + \eta > 3/2$. Suppose $|g(i)| \geqslant \eta m$ and $|g(j)| \geqslant \gamma m$. Set $\gamma + \eta - 1 = v$. It is important that $v > 1/2$. Since $|g(i) \cap g(j)| > vm$, the probability that $g(k)$ is contained in $g(i) \cap g(j))$ is greater than $2^{(v-1)m}$. For sufficiently large $n > p^m$ and $p < \sqrt{2}$, there exists $\epsilon > 0$ such that, eventually, $2^{(v-1)m} \geqslant n^{\epsilon - 1}$.

Thus, for sufficiently large n, the probability that $g(k)$ is contained in $g(i) \cap g(j)$, given that $|g(i)|$ is greater than ηm and $|g(j)|$ is greater than γm, is greater than $n^{\epsilon-1}$.

So, under these hypotheses, the probability that no $g(k)$ is contained in $g(i) \cap g(j)$ is less than $(1 - n^{\epsilon-1})^{n-2}$. But this last is $(1 + o(1))e^{-n^{\epsilon/2}}$. It follows that

$$P(E) \leqslant n_{(2)} e^{-n^{\epsilon/2}} \to 0.$$

The result follows. □

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

We now fix $\gamma < 3/4$ such that $\gamma + \eta > 3/2$. Therefore, almost surely, there exists one component of the topology containing any point whose image under g is small, and one containing any point whose image under g is large. We now need to show that these two components are, in fact, almost surely one and the same.

Lemma 2.4. *Let E_n be the following event. There exist i, j such that*

(i) $|g(i)| \geqslant \eta m$,
(ii) $|g(j)| \leqslant (1-\gamma)m$, *and*
(iii) $g(j) \subset g(i)$.

Then there exists $p < \sqrt{2}$ such that

$$\lim_{m\to\infty} \inf_{n>p^m} P(E_n) = 1.$$

Proof. Let I be minimal such that $|g(I)| \geqslant \eta m$. Given I and $g(I)$, the distribution of $g(j)$ $(j \neq I)$ is as follows.

If $j < I$, $g(j)$ has uniform distribution on the set of elements of $\wp([m])$ with size $< \eta m$. Otherwise, $g(j)$ has uniform distribution on the set $\wp([m])$. Therefore, given $(I, g(I))$,

(1) for all j, the probability of (i)–(iii) for I and j is greater than

$$\frac{\binom{\eta m}{(1-\gamma)m}}{2^m},$$

(2) the events (i)–(iii) for (I, j) are independent.

In particular, the probability of (i)–(iii) for no (I, j) is less than

$$\left(1 - \frac{\binom{\eta m}{(1-\gamma)m}}{2^m}\right)^{n-1}.$$

Note that there exists an x greater than $\sqrt{2}$, such that

$$\binom{\eta m}{(1-\gamma)m} > \binom{\frac{3}{4}m}{\frac{1}{4}m} \approx x^m/q(\sqrt{m}).$$

It suffices to take $p > 2/x$. The result follows. □

The lemmas above show that, in the limit as m goes to infinity, the minimum over all $n > p^m$ of the probability of $E_{n,m}$ goes to 1, where $E_{n,m}$ is the following event: all i satisfying $\left||g(i)| - \frac{m}{2}\right| > \left(\eta - \frac{1}{2}\right)m$ are contained in a single component, C, of the comparison graph of $P(T_m(n))$.

We now need to show the following.

Lemma 2.5. *Let $\mu < \frac{5}{2}$, and let E_n be the following event. All i satisfying $\left||g(i)| - \frac{m}{2}\right| \geqslant 1$ are contained in the same component, C, of the comparison graph. Then*

$$\lim_{m\to\infty} \inf_{n > \frac{2^{\frac{m}{2}}}{\mu}(\frac{m}{2}\log 2)} P(E_n) = 1.$$

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

Proof. Note that, if $\big||g(i)|-\frac{m}{2}\big| \geqslant 1$, then, for all $\epsilon > 0$, and for sufficiently large m,

$$\big|\{S, S \subset g(i), |S| < (1-\eta)m\} \cup \{S, S \supset g(i), |S| > \eta m\}\big| > (1-\epsilon)\big(2^{m/2+1} + 2^{m/2-1}\big).$$

Now, $(1-\epsilon)\big(2^{m/2+1} + 2^{m/2-1}\big) = (1-\epsilon)\frac{5}{2}2^{m/2}$. Therefore, for sufficiently large m, if $\tilde{p}$ is the probability that there exists an i such that $|g(i)-\frac{m}{2}| \geqslant 1$, and i is not contained in C, then

$$p < \big(1-(\mu+\delta)2^{-\frac{m}{2}}\big)^{n-1}$$

for some small $\delta > 0$ (recall that $\mu < \frac{5}{2}$). Since $np \to 0$ uniformly in n, the lemma follows. □

Essentially, what is going on is as follows. The probability that a point is not connected to C falls very fast with $\big||g(i)|-\frac{m}{2}\big|$. Near the threshold values of n and m for connectedness, therefore, we expect that, if there are any points not connected to C, they should have $|g(i)|$ as close as possible to $\frac{m}{2}$.

Proof of Theorem 2.1. For n even there is no more work to be done. For n odd, we have to discount the possibility that there exist i, j such that $g(i) = g(j) \cup \{k\}, |g(i)| = \frac{m-1}{2}$ and i, j are not in C.

First, consider the expected number of pairs i, j such that

$$g(i) = g(j) \cup \{k\},\ |g(i)| = \frac{m+1}{2}.$$

This is

$$\binom{n}{2}\frac{\binom{m}{\frac{m-1}{2}}}{2^m}\frac{m}{2^m} \ll q(m),$$

where q is some polynomial function. But, for $|g(i)| = \frac{m+1}{2}$, the probability that i is not connected to C is of the order of x^m for some $x < 1$.

The theorem follows. □

If we want to prove only that almost all points of $[n]$ are either contained in C or are singletons, then we can get by with much smaller values of n.

Theorem 2.6. *Let $E_n(i)$ be the event that i is either in C or is a singleton. There exists $p < \sqrt{2}$ such that, for each i,*

$$\lim_{m\to\infty}\inf_{n>p^m} P E_n(i) = 1.$$

Proof. Choose $\epsilon > 0$ such that $2-2\gamma-\epsilon > 1/2 > 2\gamma-1+\epsilon$. (We can do this, since γ is less than $\frac{3}{4}$.) Since

$$P\big(\big||g(i)| - m/2\big| < \epsilon m/2\big) \to 1,$$

it will be enough to show that

$$P\big(E_i\big||g(i)| - m/2\big| < \epsilon m/2\big) \to 1).$$

Suppose i is not a singleton, and let j be minimal such that $g(j)$ is comparable with $g(i)$;

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

$g(j)$ is then a randomly chosen element of

$$\{S : S \subset [m], S \text{ comparable with } g(i)\},$$

and, from the bounds we have on $|g(i)|$, the probability that $\left||g(j)| - m/2\right|$ is greater than $(\gamma - 1/2)m$ converges to 1.

The theorem follows. □

Proof of Theorem 1.1. By the above, to prove Theorem 1.1 it will be sufficient to show that the expected number of singletons tends to 0 for such n. The probability that $\left||g(i)| - \frac{m}{2}\right| < 1$ is $O(1)\frac{1}{\sqrt{m}}$. The probability, given that $\left||g(i)| - \frac{m}{2}\right| < 1$, that i is a singleton, is $(1 - \phi 2^{\frac{-m}{2}})^{n-1}$. Therefore, we seek to show that

$$\frac{n}{\sqrt{m}}\left(1 - \phi 2^{\frac{-m}{2}}\right)^{n-1} \to 0$$

for n as in the statement of Theorem 1.1. But this is immediate from (1.1). □

3. Proof of Theorem 1.2

If there is a positive ψ such that $n < \psi m$, then the statement is trivially true, since the topology in this case is almost surely discrete. We will therefore assume that this is not the case.

We introduce the notation '$g(i)$ is of size $\tilde{m}/2$' to mean that $|g(i)| = m/2$ for m even and $|g(i)| = (m \pm 1)/2$ for m odd. Now, consider the event B_i, given by

- $g(i)$ has size $\tilde{m}/2$,
- $g(i)$ is a singleton.

We will attempt to show first that, in the case $\lambda \to \infty$, $\Sigma_i B_i$ is almost certainly not zero. It will be straightforward to prove that the expected value of $\Sigma_i B_i$ is infinite. However, this will not be enough to establish the theorem, as the events B_i are not independent. Indeed, when $g(i)$ and $g(j)$ have small symmetric difference, the events B_i and B_j are highly positively correlated. To get around this difficulty, we make the following definition: X_i^δ is the event that

- $g(i)$ has size $\tilde{m}/2$,
- $g(i)$ is a singleton,
- there does not exist a j such that $g(j)$ has size $\tilde{m}/2$ and the symmetric difference $g(i) \triangle g(j)$ has size at most δm.

Define $N_\delta(m)$ to be the number of sets of size $\tilde{m}/2$ which have Hamming distance at most δm from some given set of size $\tilde{m}/2$. Choose δ such that, for the values of n, m under consideration, $\frac{N_\delta(m)n}{2^m} \leqslant n^{-\frac{1}{2}}$. Henceforth we will denote X_i^δ by X_i.

To show that the topology is almost certainly disconnected, it will be enough to show that $\Sigma_i X_i$ is almost certainly nonzero, or, alternatively (by the Chebyshev inequality), that its expected value tends towards infinity and that $E(X_i \cap X_j) = (1 + o(1))p(X_i)^2$.

Lemma 3.1. *Under the hypotheses of Theorem 1.2, the expected value of $\Sigma_i X_i$ tends towards infinity.*

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

Proof. Note that $\log n = m\log 2/2 + \log m + \log(\log 2/2\phi) + o(1)$, and hence

$$\begin{aligned} 2^{-m/2} &= 1/(\phi n)(m\log 2/2 + \log m/2 - \lambda_n) \\ &= 1/(\phi n)(\log n - \log m/2 - \log(\log 2/2\phi) - \lambda_n + o(1)). \end{aligned}$$

We have

$$\begin{aligned} P(X_i) &= \frac{\kappa\binom{m}{\lfloor m/2\rfloor}}{2^m}\left(1 - \phi 2^{-m/2} - \frac{N_\delta}{2^m}\right)^n, \\ P(X_i) &= (1+o(1))\kappa\sqrt{\frac{2}{\pi m}}\left(1 - n^{-1}\left(\log n - \log\frac{\log 2}{2\phi} - \frac{\log m}{2} - \lambda_n\right)\right)^n, \\ \log P(X_i) &= \log\sqrt{\frac{2}{\pi}} + \log\kappa\frac{\log 2}{2\phi} - \log n + \lambda_n + o(1), \\ P(X_i) &= (1+o(1))n^{-1}\left(\sqrt{\frac{2}{\pi}}\frac{\kappa\log 2}{2\phi}\right)e^{\lambda_n}, \\ E(\Sigma_i(X_i)) &= (1+o(1))\left(\sqrt{\frac{2}{\pi}}\frac{\kappa\log 2}{2\phi}\right)e^{\lambda_n}. \end{aligned}$$

The lemma follows. □

Lemma 3.2. *Under the hypotheses of Theorem 1.2,* $E(X_i \cap X_j) = (1+o(1))P(X_i)^2$.

Proof. We have

$$P(X_i \cap X_j) \geqslant \frac{\kappa\binom{m}{\lfloor m/2\rfloor}}{2^m}\frac{\kappa\binom{m}{\lfloor m/2\rfloor} - N_\delta}{2^m}\left(1 - 2\frac{\phi}{2^{m/2}} - 2\frac{N_\delta}{2^m}\right)^n.$$

By the same argument as given above, the right-hand side of this equation is $(1 + o(1))P(X_i)^2$. But we further have

$$P(X_i \cap X_j) \leqslant \frac{\kappa\binom{m}{\lfloor m/2\rfloor}}{2^m}\frac{\kappa\binom{m}{\lfloor m/2\rfloor} - N_\delta}{2^m}\left(1 - 2\frac{\phi}{2^{m/2}} + 2.2^{-\frac{1+\delta}{2}m}\right)^n$$

and, again, the argument given in the proof of Lemma 3.1 shows that the right-hand side is of order $(1+o(1))P(X_i)^2$.

The lemma and the first part of the theorem both follow. □

Now, consider the second part of the theorem. We will prove first that, in the limit, $\Sigma_i X_i$ and $\Sigma_i B_i$ have the same asymptotic distribution. We have shown (Lemma 2.5 and Theorem 2.1) that $\Sigma_i B_i = C - 1$ with probability 1. We will then prove that the factorial moments of X_i converge on χ^r, from which the theorem follows.

Lemma 3.3. *The expected value of* $\Sigma_i(X_i - B_i)$ *converges to* 0 *as* n, m *go to infinity for* n, m, *as in the second part of the statement of Theorem 1.2.*

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

Proof. Consider

$$\begin{aligned} P(B_i/X_i) &= \frac{\kappa\binom{m}{\lfloor m/2\rfloor}}{2^m}(1-\phi.2^{-m/2})^n\left(1-\left(1-\frac{N_\delta(m)}{2^m}\right)^n\right), \\ P(B_i/X_i) &= (1+o(1))n^{-1}\chi\frac{N_\delta(m)n}{2^m}, \\ nP(B_i/X_i) &\ll 1. \end{aligned}$$

The lemma follows. □

Lemma 3.4. $P(X_1 \cap X_2 \ldots \cap X_r) = (1+o(1))n^{-r}\chi^r$.

Proof.

$$P(X_1 \cap X_2 \ldots \cap X_r) \leqslant \left(\frac{\kappa\binom{m}{\lfloor m/2\rfloor}}{2^m}\right)^r\left(1-r\frac{\phi}{2^{m-1}}+\binom{r}{2}\frac{N_\delta(m)}{2^m}\right)^{(n-r)}.$$

But

$$\left(\frac{\kappa\binom{m}{\lfloor m/2\rfloor}}{2^m}\right)^r\left(1-r\frac{\phi}{2^{m-1}}+\binom{r}{2}\frac{N_\delta(m)}{2^m}\right)^{(n-r)} = (1+o(1))n^{-r}\chi^r,$$

by an argument similar to that given in the proof of Lemma 3.1. On the other hand, we have

$$P(X_1 \cap X_2 \ldots \cap X_r) \geqslant \left(\frac{\kappa\binom{m}{\lfloor m/2\rfloor}-rN_\delta(m)}{2^m}\right)^r\left(1-r\frac{\phi}{2^{m-1}}+rN_\delta(m)\right)^{(n-r)}.$$

Again, the right-hand side of this equation is of order $(1+o(1))n^{-r}\chi^r$. The lemma and the second part of the theorem follow. □

There are obvious analogies between this result and the results of Erdős and Rényi in [4]. An element of $G(n,p)$ almost surely becomes connected when $p \sim \frac{\log n}{n}$, and the graph almost surely becomes connected when the last isolated vertex disappears. There is a range from $p = \frac{1}{2n}(\log n + \log\log n)$ to $p = \frac{\log n}{n}$ where the random graph almost surely consists of a giant component plus a few ($< n^{1/2}$) isolated vertices. There are two notable differences, both due to the 'clumping' of edges in our model as compared with $G(n,p)$. Firstly, the corresponding edge probability, p, at which the last isolated vertex vanishes, is rather high: $\left(\frac{\log n}{n}\right)^{\log_{1/\sqrt{2}}(3/4)}$. Secondly, at or above the connectedness threshold in our model, the diameter of the (connected) graph is almost surely at most 6, and it seems very probable that this bound can be improved somewhat. A standard reference on random graphs is the book of Bollobás [1].

4. Proof of Theorem 1.3

We have already seen that, for $n > p^m$,

$$\lim_{n\to\infty} E(X_m(n) + X'_m(n)) = 1,$$

where $X'_m(n) = \frac{|R|}{n}$, and R is the set of singletons. As $X_m(n) + X'_m(n) \leqslant 1 + \frac{1}{n}$,

$$X_m(n) + X'_m(n) \to_{\mathscr{P}} 1.$$

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

Choose $a < \frac{\lambda}{2}$ and $b > \frac{\lambda}{2}$. We will show that $\lim_{n\to\infty} P(E) = 0$, where E is the event that *either*

(†) $|(|g(i)| - m/2)| \leqslant a\sqrt{m}$ and i is not a singleton, *or*
(‡) $|(|g(i)| - m/2)| \geqslant b\sqrt{m}$ and i is a singleton.

$$P(\dagger) = \Sigma_{|(l-m/2)|\leqslant a\sqrt{m}} c_l (1 - (1 - 2^{-l} - 2^{l-m} + 2^{-m})^{n-1}),$$

where $c_l = P(|g(i)| = l)$. Since, for $|(l - m/2)| < a\sqrt{m}$,

$$1 - \left(1 - 2^{-l} - 2^{l-m} + 2^{-m}\right)^{n-1} \leqslant 1 - \left(1 - 2^{1-m/2+a\sqrt{m}}\right)^{n-1},$$

we have

$$P(\dagger) \leqslant 1 - \left(1 - 2^{1-m/2+a\sqrt{m}}\right)^{n-1}.$$

Also,

$$\lim_{n\to\infty} n2^{1-m/2+a\sqrt{m}} = \lim_{n\to\infty} 2^{1+(a-\frac{\lambda}{2})\sqrt{m}} = 0,$$

and therefore

$$\lim_{n\to\infty} \left(1 - 2^{1-m/2+a\sqrt{m}}\right)^{n-1} = 1.$$

Therefore, $\lim_{n\to\infty} P(\dagger) = 0$. A similar argument shows that

$$P(\ddagger) \leqslant \left(1 - 2^{-m/2+b\sqrt{m}}\right)^{n-1} \to 0.$$

Our claim follows.

So, if C is our largest component, and $S = \{i \text{ s.t. } |(|g(i)| - m/2)| \geqslant \frac{\lambda}{2}\sqrt{m}\}$, then $\lim_{n\to\infty} E(|S \triangle C|) = 0$: since $|S \triangle C| \geqslant 0$ necessarily,

$$|S \triangle C| \to_{\mathscr{P}} 0.$$

The central limit theorem tells us that

$$\lim_{n\to\infty} P(g(i) \in S) = \int_{|x|\geqslant\lambda} \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}\, dx.$$

So, now, the weak law of large numbers tells us that

$$|S| \to_{\mathscr{P}} \int_{|x|\geqslant\lambda} \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}\, dx.$$

The theorem follows. □

Acknowledgements

The authors wish to thank: Eleni Katirtzoglou, who was instrumental in planting the seed for this research; Jeremy Gunawardena, director of Hewlett-Packard's Basic Research Institute in the Mathematical Sciences, for his support; all those at Hewlett-Packard Laboratories in Bristol who have helped us; and Hewlett-Packard, for their support.

Finally, we would also like to express our sincere gratitude to an anonymous referee for several very detailed and helpful reports; the presentation and content have been improved considerably as a result.

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press

References

[1] Bollobás, B. (1985) *Random Graphs*, Academic Press.
[2] Bollobás, B. (1986) *Combinatorics*, Cambridge University Press.
[3] Brightwell, G. R. (1993) Models of random partial orders. In *Surveys in Combinatorics 1993* (K. Walker, ed), Vol. 187 of *London Mathematical Society Lecture Notes*, Cambridge University Press.
[4] Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Mat Kutato Int. Kozl* **5** 17–60.
[5] Kohayakawa, Y. and Kreuter, B. The width of random subsets of Boolean lattices. To be found at: `http://www.ime.usp.br/~yoshi/pronex/publications/`.
[6] Kohayakawa, Y., Kreuter, B. and Osthus, D. (2000) The length of random subsets of Boolean lattices. *Random Struct. Alg.* **16** 177–194.
[7] Kolchin, V. F. (1986) *Random Mappings*, Optimization Software, New York. (Translation of Russian original.)
[8] Kreuter, B. (1998) Small sublattices in random subsets of Boolean lattices. In *Proceedings of the Eighth International Conference 'Random Structures and Algorithms'* (Poznan, 1997). *Random Struct. Alg.* **13** 383–407.
[9] Rényi, A. (1961) On random subsets of a finite set. *Mathematica* (*Cluj*) **3** 355–396.

https://doi.org/10.1017/S0963548301004680 Published online by Cambridge University Press