Sets that are connected in two random graphs

Michael Molloy*

August 17, 2012

Abstract

We consider two random graphs G_1, G_2, both on the same vertex set. We ask whether there is a non-trivial set of vertices S, so that S induces a connected subgraph both in G_1 and in G_2. We determine the threshold for the appearance of such a subset, as well as the size of the largest such subset.

1 Introduction

The giant component dates back to the earliest days of random graph theory[5] and still remains as one of the most studied phenomena in the field. In this paper, we consider a very natural variation on the giant component which, somewhat surprisingly, does not appear to have been studied before: giant vertex sets that are connected in two different random graphs, simultaneously.

Consider two random graphs, G_1, G_2, each drawn from the model $G_{n,p=c/n}$ for some constant $c > 1$, and each on the same vertex set, $V = \{1, ..., n\}$. We say that $S \subseteq V$ is a *doubly connected set* of (G_1, G_2) if S induces a connected subgraph in G_1 and S induces a connected subgraph in G_2. We consider a doubly connected set to be *trivial* if it has size at most 2.

A simple argument shows that w.h.p.[1] a non-trivial doubly connected set S must lie in the intersection of the giant components of G_1 and G_2 (this follows, eg, from Lemma 3.1 below). It does not take long to realize that their intersection is w.h.p. not doubly connected. Indeed, it will contain a linear number of vertices which were, e.g. connected to the giant component of G_1 through a path of vertices that are not all in the giant component of G_2. At first thought, it is not clear whether there will be a non-trivial doubly connected set at all, even for very large constant c. In this paper, we show that there is. We determine the threshold for the appearance of a non-trivial doubly connected set, and we determine its size.

Define

$$c^* = \min_{\xi>0} \frac{\xi}{(1-e^{-\xi})^2} = 2.4554...,$$

*Dept of Computer Science, University of Toronto, molloy@cs.toronto.edu. Research supported by an NSERC Discovery Grant.

[1]We say that a property holds w.h.p. (with high probability) if the probability tends to 1 as n grows.

and define $\alpha = \alpha(c)$ to be the greatest solution to

$$\alpha = (1 - e^{-\alpha c})^2.$$

Substituting $\xi = \alpha c$, we obtain $c = \xi/(1 - e^{-\xi})^2$ and so $\alpha(c) > 0$ iff $c > c^*$.

Theorem 1.1 *(a) For any $c < c^*$, w.h.p. (G_1, G_2) does not have a doubly connected set of size greater than 2.*

(b) For any $c > c^*$, w.h.p. (G_1, G_2) has a doubly connected set, and the largest such set has size $\alpha(c)n + o(n)$.

At $c = c^*$, $\alpha = .5116...$ and α increases with c. So for $c > c^*$, w.h.p. the largest doubly connected set is on more than half the vertices. In comparison, the giant component of $G_{n,p=c^*/n}$ has size roughly $.8866n$ and squaring tells us that the intersection of the giant components of G_1, G_2 would have size roughly $.786n$.

Remark: A doubly connected set of size 2 is an edge that is selected for both graphs. The expected number of such edges is $c^2/2$, and a straightforward Method of Moments argument (see e.g. [8]) shows that the number is asymptotically distributed like a Poisson variable. Thus, for $c < c^*$, the probability that there is a doubly connected set of size 2 is $1 - e^{-c^2/2} + o(1)$.

Our proof extends to several copies of $G_{n,p}$ with varying edge-probabilities. Consider random graphs $G_1, ..., G_t$, on the same vertex set, where G_i is from $G_{n,p=c_i/n}$. If there is a solution to:

$$\alpha = \prod_{i=1}^{t} (1 - e^{-c_i \alpha}),$$

then w.h.p. there is a set S of size $\alpha n + o(n)$ such that S induces a connected subgraph in each of $G_1, ..., G_t$. If there is no such solution, then w.h.p. there is no such S on more than two vertices (in fact, on more than one vertex if $t \geq 3$). We omit the details of the proof adaptation.

The proof of our theorem will be reminiscent of studies of the k-core, the pure literal rule, and other similar problems [14, 11, 6, 7, 9]. There, one repeatedly removes vertices (literals, etc.) that have small degree, until a core remains. Here, we will repeatedly remove vertices that lie in small components of at least one of the two graphs. Analysis of the k-core process is enabled by the fact that, at each iteration, what remains is a random graph conditioned on certain degree sequence properties, which we know how to analyze. In our process, at each iteration, we find that what remains is, roughly, the giant component of a random graph on a smaller vertex set. This enables us to continue to analyze it.

The techniques developed in this paper are applied in [12], where we need to analyze a similar structure on the union of random bipartite graphs. In that setting, we have vertex sets $A_1, ..., A_k$ and a random bipartite graph on each pair (A_i, A_j). The *freezing threshold* for k-colourings of $G_{n,p}$ (see [12] for a definition) is determined via the threshold for the appearance of sets $S_i \subset A_i$ such that for each i, j, the bipartite subgraph induced by (S_i, S_j) is connected.

2 Some intuition

We begin with a short intuitive explanation for α and c^*.

Lemma 3.1, below, shows that any non-trivial doubly connected set of (G_1, G_2) must have linear size. So we can determine whether such a set exists by using an iterative stripping process. At each step, every vertex that is in a small (i.e. sublinear) component of either graph is removed from both graphs. We will show that w.h.p., at each stage each graph contains at most one component that is not small (i.e. the giant component). Lemma 3.1 implies that w.h.p. if this process removes all vertices then there is no doubly connected set; otherwise, the vertices that remain form the largest doubly connected set.

Let Θ_1, Θ_2 be the vertices that are removed, at any point during the procedure, because at the time of removal they lay in small components of G_1, G_2, respectively. Thus, the doubly connected component that we find is $S = V\backslash(\Theta_1 \cup \Theta_2)$. Note that S is the giant component of $G_1\backslash\Theta_2$ (the graph remaining after removing Θ_2 from G_1) and is also the giant component of $G_2\backslash\Theta_1$.

Now make a leap of faith and suppose that, somehow, G_1 and G_2 come up with Θ_1, Θ_2 *independently.*

Suppose $|\Theta_1| = (1-\rho)n+o(n)$ for some constant $\rho > 0$; by symmetry, it is reasonable to assume that we also have $|\Theta_2| = (1-\rho)n + o(n)$. S is the vertex-set of the giant component of $G_1\backslash\Theta_2$, and also of the giant component of $G_2\backslash\Theta_1$. Since Θ_2 is independent of G_1, we can treat $G_1\backslash\Theta_2$ as $G_{n',p=c/n}$ where $n' = n - |\Theta_2| = \rho n + o(n)$. So by Lemma 3.4 (below), $|S| = \alpha n + o(n)$ where

$$\alpha = \rho(1 - e^{-\alpha c}). \tag{1}$$

By the independence of Θ_1, Θ_2, our leap of faith can also lead us to assume that Θ_1 intersects $V\backslash\Theta_2$ in the same proportion that it intersects V. $\Theta_1\backslash\Theta_2$ is simply the set of vertices not appearing on the giant component of $G_1\backslash\Theta_2$, and so

$$\frac{|\Theta_1\backslash\Theta_2|}{|V\backslash\Theta_2|} = \frac{|\Theta_1|}{|V|} \qquad \to \qquad \frac{\rho-\alpha}{\rho} = 1-\rho \qquad \to \qquad \alpha = \rho^2. \tag{2}$$

(1) and (2) yield $\alpha = (1 - e^{-\alpha c})^2$ and hence the definition of c^*.

Of course, in our process, Θ_1, Θ_2 are *not* formed independently. However, our proof can be viewed as building a different pair of sets Θ_1', Θ_2' with $\Theta_i \subseteq \Theta_i'$, which yields the same doubly connected set; i.e. where $\Theta_1 \cup \Theta_2 = \Theta_1' \cup \Theta_2'$ and $S = V\backslash(\Theta_1' \cup \Theta_2')$ is the giant component of both $G_1\backslash\Theta_2'$ and $G_2\backslash\Theta_1'$. Furthermore, Θ_1', Θ_2' will be formed (essentially) independently, enabling an analysis similar to that above.

3 A stripping procedure

We start by proving that w.h.p. every doubly connected set with more than 2 vertices must have linear size.

Lemma 3.1 *For every $c > 0$, there exists a constant $\phi = \phi(c) > 0$ such that w.h.p. (G_1, G_2) does not have a doubly connected set of size greater than 2 and less than ϕn.*

Proof This is a very standard argument, using the principle that w.h.p. the subgraph induced by any small set of vertices must have very low edge-density.

Let X_a denote the number of doubly connected sets of size a. Each such set must contain a spanning tree in G_1 and in G_2. There are a^{a-2} spanning trees on a vertices, and each has exactly $a-1$ edges. So:

$$E(X_a) \le \binom{n}{a}(a^{a-2})^2\left(\frac{c}{n}\right)^{2a-2} < \frac{n^2}{a^4c^2}\left(\frac{en}{a}a^2\left(\frac{c}{n}\right)^2\right)^a < \frac{n^2}{a^4c^2}\left(\frac{ec^2a}{n}\right)^a.$$

From this, it is straightforward to obtain $\sum_{a=3}^{\phi n} E(X_a) = o(1)$ for $\phi < \frac{1}{ec^2}$. $\square$

Lemma 3.1 implies that any doubly connected set of size at least 3 must be contained in the intersection of the giant components of G_1 and G_2. Furthermore, the following procedure will w.h.p. find it.

At each iteration i, we will have a set of vertices $V_i \subset \{1, ..., n\}$. We define S_i, resp. T_i, to be the vertex set of the largest component of the subgraph of G_1, resp. G_2, induced by V_i.

STRIP:
Initialize $V_1 = \{1, ..., n\}$.
For $i = 1$ to ∞
    Expose the vertices of S_i.
    Expose the vertices of T_i.
    if $S_i = T_i = V_i$ then HALT SUCCEED.
    else
        $V_{i+1} := S_i \cap T_i$.
        if $V_{i+1} = \emptyset$ then HALT FAIL.

To clarify: at each iteration, we only expose the vertices in S_i, T_i; we do not expose any of the edges amongst those vertices. This will be important for our analysis (see e.g. Observation 4.1). Informally, suppose that an oracle sees all of G_1 and G_2, and at each iteration, only tells us which vertices are in S_i, T_i; any probabilistic analysis that we carry out is conditioned on the information that the oracle has given us thus far. Note that the information the oracle provides is enough for us to carry out the procedure.

More formally, $S_i \subset V_i$ is a random set of vertices, such that the probability that $S_i = S$ is the probability that S is the vertex set of the largest component of G_1 induced by V_i; after $S_i = S$ is selected we then condition all remaining random choices on the event that that vertex set is equal to S. (And similarly for T_i.)

We will prove that w.h.p.: If we halt Fail, then there is no doubly connected set of linear size. If we halt Succeed then $S_i = T_i$ is doubly connected, and is the maximum doubly connected set in (G_1, G_2).

This is the iterative procedure described in Section 2. If it halts Succeed, then Θ_1, Θ_2 are $\cup_{i\ge1} V_i \backslash S_i, \cup_{i\ge1} V_i \backslash T_i$.

To analyse STRIP, we define:

$$\rho_1 = 1,$$

for $i \ge 1$, γ_i is the greatest solution to

$$\gamma_i = \rho_i(1 - e^{-\gamma_i c}), \tag{3}$$

and

$$\rho_{i+1} = \gamma_i/\rho_i = 1 - e^{-\gamma_i c}.$$

Note: If $\rho_i \le 1/c$ then $\gamma_i = 0$, and if $\rho_i > 1/c$ then there is only one positive solution to (3).

Lemma 3.2 *The sequences $\gamma_1, \ldots$ and $\rho_1, \ldots$ are both strictly decreasing and*

(a) For any $c < c^*$, there exists I such that $\gamma_I = \rho_I = 0$.

(b) For any $c > c^*$, $\lim_{i\to\infty} \gamma_i = \alpha(c)$ and $\lim_{i\to\infty} \rho_i = \sqrt{\alpha(c)}$.

Proof It is easy to confirm that $\rho_2 < 1 = \rho_1$. Suppose $\rho_i < \rho_{i-1}$. Then it is straightforward to check that $\gamma_i < \gamma_{i-1}$ (in fact, the proof of Lemma 3.4 below implies this). Note that $\rho_{i+1} = 1 - e^{-\gamma_i c}$ and so $\gamma_i < \gamma_{i-1}$ implies $\rho_{i+1} < \rho_i$. So both sequences are decreasing. Since they are both positive, they have a limit. That limit must be a fixed point of the recursive equations. At that fixed point (ρ, γ), we have $\rho = \gamma/\rho$ and so $\gamma = \rho^2$. This yields:

$$\rho = 1 - e^{-c\rho^2} \tag{4}$$

For $c < c^*$ the only solution to (4) is $\rho = 0$, and so ρ_i, γ_i tend to 0. Therefore, there exists I such that $\rho_{I-1} \le 1/c$ and hence $\gamma_{I-1} = \rho_I = \gamma_I = 0$.

For $c \ge c^*$, the largest solution to (4) is $\rho = \sqrt{\alpha}$. A simple induction shows that $\rho_i \ge \sqrt{\alpha}$ for each i, and this yields part (b). Indeed, define $f(x) = 1 - e^{-yc}$ where $y = y(x)$ is the greatest solution to $y = x(1 - e^{-yc})$. Thus $\rho_{i+1} = f(\rho_i)$. Now $y(x)$ is non-decreasing and hence so is $f(x)$. Also note that $f(\sqrt{\alpha}) = \sqrt{\alpha}$. So if $\rho_i \ge \sqrt{\alpha}$ then $\rho_{i+1} = f(\rho_i) \ge f(\sqrt{\alpha}) = \sqrt{\alpha}$. Since $\rho_1 = 1 > \sqrt{\alpha}$, this completes the induction. $\square$

In the next section, we will prove:

Lemma 3.3 *For any constant i, and any constant $\phi > 0$, w.h.p.*

(a) $|V_i| = \rho_i^2 n + o(n)$;

(b) $|S_i|, |T_i| = \gamma_i n + o(n)$;

(c) every vertex in $V_i \backslash V_{i+1}$ lies in a component of size less than ϕn in either $G_1 \cap V_i$ or $G_2 \cap V_i$.

Lemmas 3.2, 3.3 show us that, by taking a sufficiently high constant number of iterations, S_i, T_i get quite close to what Theorem 1.1 would predict. Lemma 3.3(c) and Lemma 3.1 imply inductively that every vertex not in V_i does not lie in a non-trivial doubly connected set. Lemma 3.1 completes the proof of Theorem 1.1(a), by considering i high enough so that $\gamma_i < \phi$. The proof of Theorem 1.1(b) requires more work.

We close this section with some useful facts about the size of a giant component. We say $X \in Y \pm Z$ to mean that X is in the range $[Y - Z, Y + Z]$.

Lemma 3.4 *For $c\tau > 1$, w.h.p. the size of the largest component of $G_{n',p=c/n}$ where $n' = \tau n$ is in $\beta n \pm n^{3/5}$, where β is the positive solution to $\beta = \tau(1 - e^{-\beta c})$.*

Proof Note that $G_{n',p=c/n}$ is $G_{n',p=c\tau/n'}$. So the classical result by Erdős and Rényi[5] on the size of the giant component in $G_{n,p}$, along with the fact that the distribution of the size is asymptotically normal[16, 15, 3] implies that w.h.p. the size of the largest component is in $bn' \pm n^{3/5}$ where b is the positive solution to $b = 1 - e^{-c\tau b}$. The lemma follows by noting that $\beta = b\tau$. $\square$

We are interested in how β changes with τ near $\tau = \sqrt{\alpha}$.

Lemma 3.5 *For every $c > c^*$, at $\tau = \sqrt{\alpha}$ we have*

$$\frac{\partial}{\partial \tau}\beta < 2\sqrt{\alpha}.$$

Proof We use β' to denote $\frac{\partial}{\partial \tau}\beta$. So

$$\beta' = (1 - e^{-\beta c}) + \tau \frac{\partial}{\partial \tau}(1 - e^{-\beta c}) = (1 - e^{-\beta c}) + \tau c e^{-\beta c}\beta'.$$

Solving for β' yields:

$$\beta' = \frac{1 - e^{-\beta c}}{1 - \tau c e^{-\beta c}}.$$

Recall that c^* is a minimum of $\frac{\xi}{(1-e^{-\xi})^2}$. It is easy to check that $\frac{\xi}{(1-e^{-\xi})^2}$ is increasing above c^* and so, since $c > c^*$, the derivative with respect to ξ is postive. Differentiating and simplifying yields:

$$1 - e^{-\xi} > 2\xi e^{-\xi}.$$

Applying $\xi = \alpha c$, and $\sqrt{\alpha} = 1 - e^{-\alpha c}$, this yields $\sqrt{\alpha} > 2\alpha c e^{-\alpha c}$ so $\frac{1}{2} > \sqrt{\alpha} c e^{-\alpha c}$. Noting that at $\tau = \sqrt{\alpha}$ we have $\beta = \alpha$, yields that at $\tau = \sqrt{\alpha}$:

$$\frac{\partial}{\partial \tau}\beta = \frac{1 - e^{-\alpha c}}{1 - \sqrt{\alpha} c e^{-\alpha c}} < \frac{\sqrt{\alpha}}{1 - \frac{1}{2}} = 2\sqrt{\alpha}.$$

$\square$

We close this section with the Chernoff Bound[4]. We use the version presented in [13]. Here, $BIN(n,p)$ is the sum of n independent variables, each equal to 1 with probability p and 0 otherwise.

The Chernoff Bound *For any $0 \le t \le np$:*

$$\Pr\left(|BIN(n,p) - np| > t\right) < 2e^{-\frac{t^2}{3np}}.$$

4 Proof of Lemma 3.3

It is useful to consider what STRIP looks like from the perspective of G_1. At each iteration, G_1 removes all its small components, then is given a list T_i of vertices and removes all remaining vertices which are not in T_i. The following observation about T_i is crucial to our analysis:

Observation 4.1 *Given V_i and $t_i = |T_i|$, the set T_i is a uniformly random set of t_i vertices from V_i, and the choice of these vertices is independent of G_1.*

Proof At each step $j < i$ of STRIP, we expose the vertex sets S_j, T_j without exposing the edges. This implies that for any potential graph $H = G_2 \cap V_i$, every graph obtained by permuting the vertices of H is equally likely to be $G_2 \cap V_i$, and this holds even after conditioning on G_1 being equal to any particular graph. So consider any two vertex sets $T, T' \subset V_i$, both of size t_i, and any graph H such that if $G_2 \cap V_i = H$ then $|T_i| = t_i$. By symmetry, the number of ways to permute the vertices of H to obtain a new graph for which $T_i = T$ is equal to the number of ways to permute the vertices of H to obtain $T_i = T'$ (the number is $t_i!(|V_i| - t_i)!$). Summing over all possibilities for H yields that $\mathbf{Pr}(T_i = T) = \mathbf{Pr}(T_i = T')$, even after conditioning on G_1. This yields the observation. $\square$

This implies that if we are only viewing things from the perspective of G_1, then at each iteration, instead of exposing T_i and then deleting $V_i \backslash T_i$ from V_i, we can instead just expose $t_i = |T_i|$ and remove $\ell_i = |V_i| - t_i$ uniformly random vertices from V_i.

It will be convenient to keep track of an additional set $U_i \supseteq V_i$. Vertices are removed from U_i at the same proportional rate that G_2 causes vertices to be removed from V_i. G_1 does not cause any vertices to be removed from U_i. This set is useful because it extracts the effect that G_2 has on V_i and hence on S_i. Moreover, U_i is close to being independent of G_1; close enough to be useful. Very roughly speaking, in the limit as $i \to \infty$, $V - U_i$ can be thought of as Θ_2' from Section 2.

We specify U_i using the following modification of STRIP.

As before, at each iteration i, we will have a set of vertices $V_i \subset \{1, ..., n\}$, and we define S_i, resp. T_i, to be the vertex set of the largest component of the subgraph of G_1, resp. G_2, induced by V_i.

Again, we only expose the vertices of S_i, T_i. This time, in order to expose T_i, we will first expose $|T_i|$ and then choose a uniformly random set of $|T_i|$ vertices of V_i, making use of Observation 4.1. Equivalently, we expose $\ell_i = |V_i \backslash T_i|$ and obtain T_i by deleting ℓ_i uniformly random vertices from V_i. We choose those vertices in an unusual manner, to facilitate the specification of U_{i+1}.

STRIP1:
Initialize $V_1 = U_1 = \{1, ..., n\}$.
For $i = 1$ to ∞
  Expose the vertices of S_i.
  Expose $\ell_i = |V_i \backslash T_i|$.
  If $S_i = V_i$ and $\ell_i = 0$ then HALT SUCCEED
  Else
    Initialize $L_i = \emptyset$.
    Repeat ℓ_i times:
      Repeat until we choose a $u \in V_i$
        Choose a uniformly random vertex $u \in U_i$ (without replacement).
        Place u into L_i.
    $T_i := V_i \backslash L_i$.
    $V_{i+1} := S_i \backslash L_i$.   (Note that $S_i \backslash L_i = S_i \cap T_i$.)

$U_{i+1} := U_i \backslash L_i$.
If $V_{i+1} = \emptyset$ then HALT FAIL.

Note that, for each i, the ℓ_i vertices of L_i that are in V_i are uniformly random members of V_i. So by Observation 4.1, we can couple the choice of T_i in STRIP with the choice of $V_i \backslash L_i$ in STRIP1. Under this coupling, STRIP and STRIP1 produce the same sets S_i, T_i, V_i. More precisely:

We run STRIP and STRIP1 in parallel. At iteration i, we first expose the vertex sets $S_i = S, T_i = T$ for STRIP. Then we carry out iteration i of STRIP1 as follows: We set $\ell_i = |V_i \backslash T|$. Each time we select $u \in U_i$ we first determine whether $u \in V_i$. If $u \in V_i$ then we pick u to be a uniformly random member of T (without replacement). If $u \notin V_i$ then we pick u to be a uniformly random member of $U_i \backslash V_i$ (without replacement). Note that, by Observation 4.1, u is a uniformly random member of U_i (without replacement) and so this is a valid coupling; i.e. u is chosen with the correct distribution for STRIP1. Note also that this coupling ensures that we select $S_i = S, T_i = T$ for STRIP1, and so STRIP and STRIP1 produce the same sets.

Intuition: So long as S_i has linear size, S_i will be the largest component of $G_1 \cap U_i$. The reason is that all vertices in $U_i \backslash V_i$ were removed from V_i because they were not in S_j for some $j < i$; i.e. they were in small components of $G_1 \cap V_j$. Inductively, this means that they were in small components of $G_1 \cap U_j$ and hence are in small components of $G_1 \cap U_i$, as $U_i \subseteq U_j$. Therefore, all vertices in the giant component of $G_1 \cap U_i$ are in V_i, and so the largest component of $G_1 \cap U_i$ is also the largest component of $G_1 \cap V_i$.

This allows us to analyze $|S_i|$ by instead analyzing the largest component of $G_1 \cap U_i$, which is much easier. To do so, we determine the size of U_i:

Lemma 4.2 *For any constant i, w.h.p. $|U_i| = \rho_i n + o(n)$.*

Intuition: Suppose that we were to define another set U_i' which is analagous to U_i but from the perspective of G_2. Note that a simple induction shows $V_i = U_i \cap U_i'$: Indeed, every vertex in $V_i \backslash T_i$ is in L_i and hence is not in U_{i+1}; similarly every vertex in $V_i \backslash S_i$ is not in U_{i+1}'. Since $V_{i+1} = S_i \cap T_i$ and $V_{i+1} \subseteq U_i, U_i'$, we have $V_{i+1} = U_i \cap U_i'$. By symmetry, $|U_i'| = \rho_i n + o(n)$. Now U_i and U_i' are not independent, but they are very close - close enough that we have $|U_i \cap U_i'| = \rho_i^2 n + o(n)$. This is why $|V_i| = \rho_i^2 + o(n)$ (Lemma 3.3).

Proof of Lemmas 3.3, 4.2: We analyze STRIP1; by our coupling, this suffices to prove Lemma 4.2.

We will prove the lemmas by induction. More specifically, we prove there are two sequences of constants $\eta_1, \eta_2, ...$ and $\eta_1', \eta_2', ...$ such that w.h.p.

$$\begin{aligned} |S_i|, |T_i| &\in \gamma_i n \pm \eta_i n^{2/3} & (5)\\ |V_i| &\in \rho_i^2 n \pm \eta_i' n^{2/3} & (6)\\ |U_i| &\in \rho_i n \pm \eta_i' n^{2/3}. & (7) \end{aligned}$$

We will also prove:

$$\text{Every } v \in U_i \backslash V_i \text{ is in a component of size } o(n) \text{ in } G_1 \cap U_i \qquad (8)$$

We start with $\eta'_1 = 0$; η_1 will be implicitly defined below. Our base cases are that (6), (7) and (8) hold trivially for $i = 1$. Now we proceed by induction.

Suppose that V_i, S_i, ℓ_i are chosen. $S_i \subseteq V_i$ and $L_i \cap V_i$ is a set of ℓ_i uniformly random vertices of V_i, so $\mathbf{E}(|S_i \cap L_i|) = \frac{|S_i|\ell_i}{|V_i|}$. Since $V_{i+1} = S_i \backslash L_i$, we have

$$\mathbf{E}(|V_{i+1}|) = |S_i|(1 - \frac{\ell_i}{|V_i|}) = \frac{|T_i||S_i|}{|V_i|},$$

which is in the range $\frac{\gamma_i^2}{\rho_i^2} n \pm \frac{1}{2}\eta'_{i+1} n^{2/3}$, for sufficiently large η'_{i+1} so long as (5) and (6) hold for i.

For any $0 \le a \le |S_i|$, the probability that $|V_{i+1}| = |S_i| - a$ is

$$\binom{|S_i|}{a}\binom{|V_i| - |S_i|}{\ell_i - a} \Big/ \binom{|V_i|}{\ell_i}.$$

From this, it is straightforward to show that w.h.p. $|V_i|$ is within $\frac{1}{2}\eta'_i n^{2/3}$ of its mean. Recalling $\rho_{i+1} = \frac{\gamma_i}{\rho_i}$, this establishes that w.h.p. $|V_{i+1}| \in \rho_{i+1}^2 n \pm \eta'_{i+1} n^{2/3}$; i.e. this yields (6) for $i + 1$.

The induction step for U_{i+1} is similar. This time, we have $\mathbf{E}(|U_{i+1}|) = \frac{|T_i||U_i|}{|V_i|}$ which is in the range $\frac{\gamma_i}{\rho_i} n \pm \frac{1}{2}\eta'_{i+1} n^{2/3}$ for sufficiently large η'_{i+1}, so long as (5), (6) and (7) hold for i.

$|U_{i+1}|$ is determined by the number of uniformly random vertices of U_i that we select for L_i before selecting ℓ_i from V_i. The probability that this number is a is

$$\frac{\ell_i}{a}\binom{|V_i|}{\ell_i}\binom{|U_i| - |V_i|}{a - \ell_i} \Big/ \binom{|U_i|}{a}.$$

Explanation: consider the first a vertices selected for L_i. The event occurs iff (i) exactly ℓ_i of them are from V_i and (ii) one of the vertices from V_i is the ath vertex removed.

Again, from this it is straightforward to show that w.h.p. $|U_i|$ is within $\frac{1}{2}\eta'_i n^{2/3}$ of its mean. Recalling $\rho_{i+1} = \frac{\gamma_i}{\rho_i}$, this establishes that w.h.p. $|U_{i+1}| \in \rho_{i+1} n \pm \eta'_{i+1} n^{2/3}$; i.e. this yields (7) for $i + 1$.

Suppose that (8) holds for i. We will argue below that this implies that S_i is the largest component of $G_1 \cap U_i$, so long as that component has linear size; we will focus on the latter component. It would be convenient if $G_1 \cap U_i$ were distributed like $G_{n',p=c/n}$ where $n' = |U_i|$. But it is not, since there is some dependency between G_1 and $|U_i|$. So instead, we sandwich $G_1 \cap U_i$ between two graphs which really are from the $G_{n,p}$ model.

Suppose that (7) holds for i. We consider two sets of vertices U^-, U^+, which are defined to be uniformly random subsets of $\{1, ..., n\}$ of sizes $x^- = \lfloor \rho_i n - 2\eta'_i n^{2/3} \rfloor, x^+ = \lceil \rho_i n + 2\eta'_i n^{2/3} \rceil$. We couple these two sets with our process as follows:

Choose a sequence $\mathcal{A} = a_1, ..., a_{n-x^-}$ uniformly random vertices from $\{1, ..., n\}$, without replacement. Set U^- to be the set of x^- vertices that are not in $\mathcal{A}$. Set U^+ to be the set of x^+ vertices that are not amongst the first $n - x^+$ members of $\mathcal{A}$. Thus U^- is a uniformly random set of x^- vertices, U^+ is a uniformly random set of x^+ vertices, and $U^- \subset U^+$. During the first i iterations of STRIP1, each time we choose a uniform member of $U_j, j < i$ to place into L_j, we simply take the next member of $\mathcal{A}$; this is permissable unless the total number of such selections exceeds $n - x^-$, i.e. unless $|U_i| < x^-$. Note that this chooses the member of U_j with the correct (i.e. uniform) distribution.

If $x^- \le |U_i| \le x^+$, then our coupling succeeds and $U^- \subseteq U_i \subseteq U^+$; by our induction hypothesis, this is indeed the case w.h.p. and so we can assume that the coupling succeeds.

Because the vertices of U^- are selected uniformly at random, we can select them before exposing the edges of $G_1 \cap U^-$. Therefore, we see that $G_1 \cap U^-, G_1 \cap U^+$ are distributed as $G_{x^-,p=c/n}, G_{x^+,p=c/n}$.

Case 1: $\rho_i > 1/c$. Let S^-, S^+ be the giant components of $G_1 \cap U^-, G_1 \cap U^+$ respectively; our coupling implies $S^- \subseteq S^+$. We argue that w.h.p. $S^- \subseteq S_i \subseteq S^+$, which will easily yield (5) for S_i. First we analyze the sizes of S^-, S^+.

Lemma 3.4 yields w.h.p. $|S^-| \in \beta n \pm n^{3/5}$ where β is the positive solution to

$$\beta = \tau(1 - e^{-\beta c}), \qquad \text{with } \tau = x^-/n = \rho_i - 2\eta_i' n^{-1/3}.$$

Recalling that γ_i is the positive solution to $\gamma_i = \rho_i(1 - e^{-\gamma_i c})$, we see that for η_i sufficiently large (in terms of η_i' and the value of $\frac{\partial}{\partial \tau}\beta$ at $\tau = \rho_i$), we have $\beta \ge \gamma_i - \frac{1}{2}\eta_i n^{-1/3}$ and so w.h.p. $|S^-| \ge \gamma_i n - \eta_i n^{2/3}$. Similarly, we obtain that w.h.p. $|S^+| \le \gamma_i n + \eta_i n^{2/3}$.

Next we argue that w.h.p. $S^- \subseteq S_i \subseteq S^+$. By our coupling, S^- is contained in a component Q of $G_1 \cap U_i$. We showed above that w.h.p. S^- has linear size. So, assuming that (8) holds for i, all vertices of Q must lie in V_i, and hence Q is a component of $G_1 \cap V_i$. Since $V_i \subseteq U_i \subseteq U^+$, all vertices of every other component X of $G_1 \cap V_i$ lie in components of $G_1 \cap U^+$. If X does not lie in S^+ then w.h.p. X lies in a component of size $o(n)$ in $G_1 \cap U^+$. If X does lie in S^+ then X has size at most $|S^+| - |Q| \le |S^+| - |S^-| = o(n)$. Therefore w.h.p. every component of $G_1 \cap V_i$ other than Q has size $o(n)$ and so $S_i = Q$. This yields w.h.p. $S^- \subseteq S_i \subseteq S^+$, and hence (5) for S_i. The proof of (5) for T_i follows by symmetry.

We have also shown that every vertex in $U_i \backslash S_i$ lies in a component of size $o(n)$ in $G_1 \cap U_i$ and hence must lie in a component of size $o(n)$ in $G_1 \cap U_{i+1}$. Since $U_{i+1} \backslash V_{i+1} \subseteq U_i \backslash S_i$, this establishes (8) for $i + 1$.

Case 2: $\rho_i < 1/c$. In this case, $\gamma_i = 0$. W.h.p. all components of $G_{x^+,p=c/n}$ have size $O(\log n)$[5]. Since $V_i \subseteq U_i \subset U^+$, it follows that w.h.p. every component of $G_1 \cap V_i$ also has size $O(\log n)$ and so $|S_i| = O(\log n) < \eta_i n^{2/3}$ for any η_i (and large n). The same bound holds w.h.p. for $|T_i|$ by symmetry, thus yielding (5) for i.

Since $U_{i+1} \subseteq U_i \subseteq U^+$, every component of $G_1 \cap U_{i+1}$ also has size $O(\log n)$. This yields (8) for $i + 1$.

Case 3: $\rho_i = 1/c$. In this case, again we have $\gamma_i = 0$. Note that $p = \frac{1}{x^+} + O\left((x^+)^{-4/3}\right)$. It is well-known then that the largest component of $G_{x^+,p=c/n}$ w.h.p. has size $\Theta\left((x^+)^{2/3}\right) = \Theta(n^{2/3})$[2, 10, 8]. The rest follows as in Case 2, yielding (5) for i and (8) for $i + 1$.

To recap, we have shown that the bounds hold:

- (7) and (8) for i imply that w.h.p. (5) holds for i and (8) holds for $i + 1$;

- (5), (6) and (7) for i imply that w.h.p. (6) and (7) hold for $i+1$.

Recall that (6), (7) and (8) hold trivially for $i=1$. For any constant i, this induction requires that $O(i)$ events hold where each such event holds w.h.p. Since i is a constant, the union of those events holds w.h.p. This completes the induction and hence proves Lemma 3.3(a,b) and Lemma 4.2.

We close by noting that the arguments above imply Lemma 3.3(c). In particular, we showed that w.h.p. every vertex of $V_i \backslash S_i$ lies in a component of size $o(n)$ in $G_1 \cap U^+$ and hence lies in a component of size $o(n)$ in $G_1 \cap V_i$. By symmetry, the same argument shows that w.h.p. every vertex of $V_i \backslash T_i$ lies in a component of size $o(n)$ in $G_2 \cap V_i$. Since $V_{i+1} = S_i \cap T_i$, this yields Lemma 3.3(c). □

5 Proof of Theorem 1.1

Proof of Theorem 1.1(a): Suppose $c < c^*$. By Lemma 3.2 we can take I large enough that $\rho_I^2 = 0$; thus w.h.p. $|V_I| = o(n)$.

By Lemma 3.3(c), every vertex $v \in V_1 \backslash V_2$ lies in a component of either G_1 or G_2 which has size less than ϕn, where $\phi = \phi(c)$ comes from Lemma 3.1. Thus Lemma 3.1 implies that v cannot lie in any non-trivial doubly connected component. The same argument, applied inductively, shows that no vertex in $V_j \backslash V_{j+1}$ can lie in a non-trivial doubly connected component, for $j < I$. W.h.p. $|V_I| = o(n) < \phi n$ and so no vertex of V_I can lie in a non-trivial doubly connected component. This yields Theorem 1.1(a). □

To prove part (b), we will run STRIP1 until some large constant I so that $|V_I|, |S_I|$ are very close to their limits α, and $|U_I|$ is very close to its limit $\sqrt{\alpha}$. From this point on, Lemma 3.5 implies that each time we delete a vertex u from U_i, we expect to remove fewer than $2\sqrt{\alpha}$ vertices from the giant component, S_i. Note that the probability that u itself is in S_i is $\frac{|S_i|}{|U_i|} \approx \sqrt{\alpha}$. So each time we delete a vertex from S_i we expect to reduce the size of S_i by less than 2; i.e. the expected number of additional vertices to be removed from S_i is less than one. Because S_i, V_i are almost the same size, each time we delete a vertex from V_i, we expect to cause $|V_i \backslash S_i|$ to increase by less than one.

It follows that w.h.p. STRIP1 will halt very soon. We can view the vertices of $V_i \backslash T_i$ as a queue of vertices that must be removed from G_1 and the vertices of $V_i \backslash S_i$ as a queue of vertices that must be removed from G_2. Each time we process a vertex from one queue, it results in an expected increase in the other queue of less than one. So the total size of these queues has a negative drift and with high probability the queues empty quickly. Forthwith the details.

Inspired in part by the approach in [6], we consider a process that, after we find a doubly connected set, continues to remove vertices from U_i (S_i, T_i, V_i will remain unchanged). We will then consider an iteration that is greater than the time we expect STRIP1 to halt and show that we must indeed have found a doubly connected set before that iteration. Allowing ourselves to consider an iteration that is significantly larger than the actual stopping time of STRIP1 makes this task easier.

As before, at each iteration i, we will have a set of vertices $V_i \subset \{1, ..., n\}$, and we define S_i, resp. T_i, to be the vertex set of the largest component of the subgraph of G_1, resp. G_2, induced by

V_i.

KEEP-STRIPPING:
Initialize $V_1 = U_1 = \{1, ..., n\}$.
For $i = 1$ to ∞
    Expose the vertices of S_i.
    Expose $\ell_i = |V_i \backslash T_i|$.
    If $S_i = V_i$ and $\ell_i = 0$ then
        (*) Repeat until $U_i = \emptyset$:
            Pick a uniform vertex $u \in U_i$
            $L_i := \{u\}$
            $U_{i+1} := U_i \backslash L_i$
            $S_{i+1} := S_i; T_{i+1} := T_i; V_{i+1} := V_i$ and so these sets remain unchanged.
            $i := i + 1$.
    Else
        Initialize $L_i = \emptyset$.
        Repeat ℓ_i times:
            Repeat until we choose a $u \in V_i$
                Choose a uniformly random vertex $u \in U_i$ (without replacement).
                Place u into L_i.
        $T_i := V_i \backslash L_i$.
        $V_{i+1} := S_i \backslash L_i$.
        $U_{i+1} := U_i \backslash L_i$.
        If $V_{i+1} = \emptyset$ then HALT FAIL.

So once the procedure enters (*), we simply remove vertices from U_i one-at-a-time. S_i, T_i, V_i remain unchanged, and because we entered (*), we have $S_i = T_i = V_i$.

Lemma 5.1 *For any $c > c^*$ and $\delta > 0$, there exist constants $I = I(c, \delta)$ such that w.h.p.*

(a) $\alpha n < |V_I| < (\alpha + \delta)n$;

(b) STRIP halts after removing at most $\delta n + o(n)$ vertices from V_I.

Proof Note that to prove (b), it suffices to prove that w.h.p. KEEP-STRIPPING enters (*) within $I + \delta n + o(n)$ iterations. This is because (i) KEEP-STRIPPING is identical to STRIP1 up until the point that STRIP1 halts and KEEP-STRIPPING enters (*); and (ii) STRIP1 and STRIP can be coupled to produce the same sets S_i, T_i, V_i.

Recall from Lemma 3.4 that if $\tau c > 1$ then w.h.p. the size of the largest component in $G_{n', p=c/n}$ where $n' = \tau n$ is in $\beta n \pm n^{3/5}$ where $\beta = \beta(\tau)$ is the solution to $\beta = \tau(1 - e^{-\beta c})$. By Lemma 3.5 and the continuity of $\frac{\partial}{\partial \tau}\beta$, there exists $\xi, \zeta > 0$ such that:

$$\forall \tau \in [\sqrt{\alpha} - \xi, \sqrt{\alpha} + \xi], \text{ we have } \frac{\partial}{\partial \tau}\beta < (2 - \zeta)\sqrt{\alpha}. \tag{9}$$

We can assume that $\delta < \xi$ and is sufficiently small in terms of ζ. By Lemma 3.2, we can choose constant I so that $\sqrt{\alpha} \leq \rho_I < \sqrt{\alpha} + \frac{\delta}{2}$, $\alpha \leq \gamma_I < \alpha + \frac{\delta}{2}$, $\rho_{I-1} - \rho_I < \delta^2$ and $\gamma_{I-1} - \gamma_I < \delta^2$. By Lemmas 3.3 and 4.2, we have that w.h.p. $|U_I| = \rho_I n + o(n), |V_I| = \rho_I^2 n + o(n)$ and both $|S_I|, |T_I| = \gamma_I n + o(n)$.

We define U^* to be equal to U_i at the time where U_i has size exactly $(\rho_I - \delta)n$. We have to be careful what we mean by this, since all the vertices in L_i are removed from U_i at once. So:

Definition 5.2 *If, at any point during any iteration i, $|L_i| = \lfloor(\rho_I - \delta)n\rfloor - |U_i|$ then we set:*

- *L^* is the set of vertices that are in L_i at that point;*
- *$U^* = U_i \backslash L_i$.*
- *S^* is the set of vertices in the largest component of $G_1 \cap U^*$.*

We let J denote the iteration during which U^* is formed. By Lemma 3.3, w.h.p. $J \geq I$.

We will see that S^* has linear size. Note that this implies $S^* \subseteq S_I$. This is because, as argued in the proof of Lemmas 3.3 and 4.2, S_I is the largest component of $G_1 \cap U_I$, and every other component in $G_1 \cap U_I$ has size $o(n)$. Since $U^* \subset U_I$, every component of $G_1 \cap U^*$ that is not contained in S_I has size $o(n)$. Therefore if S^* has linear size then $S^* \subseteq S_I$.

Throughout all iterations of KEEP-STRIPPING, each vertex of L_i was selected uniformly from U_i. Therefore, U^* is a uniformly random set of $\lfloor(\rho_I - \delta)n\rfloor$ vertices from $U_1 = \{1, ..., n\}$. Thus, we can expose the subgraph of G_1 induced by U^* by first choosing the vertices of U^* and then choosing the edges; i.e. we can treat it as $G_{n'=\lfloor(\rho_I-\delta)n\rfloor, p=c/n}$.

We will prove that w.h.p. we enter line (*) before forming U^*; i.e. that w.h.p. we find a doubly connected set before removing $\delta n + o(n)$ vertices from U_I. So suppose that we form U^* before entering line (*); we will reach a contradiction by bounding $|S_{I-1} \backslash S^*|$ in two different ways.

Since $\alpha(c)$ is increasing with c, $\rho_I \geq \sqrt{\alpha(c)}$, and $c > c^*$, we have $c\rho_I > c^*\sqrt{\alpha(c^*)} = (2.4554...) \times (.5116..)^{1/2} > 1$. So by taking $\delta > 0$ sufficiently small, we have $c(\rho_I - \delta) > 1$, and so Lemma 3.4 yields that w.h.p. $|S^*| = gn + o(n)$ where $g = (\rho_I - \delta)(1 - e^{-gc})$.

Since $\delta < \xi$, and $\rho_{I-1} < \rho_I + \delta^2 < \sqrt{\alpha} + \frac{\delta}{2} + \delta^2$, we have $\sqrt{\alpha} \leq \rho_{I-1} \leq \sqrt{\alpha} + \xi$. So (9) yields:

$$g \geq \gamma_{I-1} - (2-\zeta)\sqrt{\alpha}(\rho_{I-1} - (\rho_I - \delta)) + o(1) > \gamma_{I-1} - (2-\zeta)\sqrt{\alpha}(\delta + \delta^2) + o(1) > \gamma_{I-1} - (2 - \frac{2}{3}\zeta)\sqrt{\alpha}\delta,$$

for δ sufficiently small in terms of ζ. W.h.p., $|S_{I-1}| = \gamma_{I-1}n + o(n)$. Therefore w.h.p.

$$|S^*| = gn + o(n) \geq |S_{I-1}| - (2 - \frac{\zeta}{2})\sqrt{\alpha}\delta n. \tag{10}$$

Consider the set

$$A = \left(\cup_{i=I}^{J-1} L_i \cap V_i\right) \cup (L^* \cap V_J).$$

Note that $A \subseteq S_{I-1} \backslash S^*$ since: (a) for each $i \geq I$, we have $V_i \subseteq V_I$, and so $A \subseteq V_I \subseteq S_{I-1}$; (b) $\left(L^* \cup_{i=I}^{J-1} L_i\right) \cap U^* = \emptyset$ so $A \cap S^* = \emptyset$.

Note that $|(\cup_{i=I}^{J-1} L_i) \cup L^*| = |U_I| - |U^*| = \delta n + o(n)$. The vertices of L_i are selected uniformly from U_i. Since $|V_i| \geq |V_I| - \delta n$, each choice is a member of V_i with probability at least

$$\frac{V_I - \delta n + o(n)}{U_I} = \frac{(\rho_I^2 - \delta)n + o(n)}{\rho_I n + o(n)} > \rho_I - \frac{2\delta}{\rho_I} > \sqrt{\alpha} - \frac{2\delta}{\sqrt{\alpha}},$$

since $\rho_I > \sqrt{\alpha}$. So the number of these vertices that are in V_i is dominated from below by the binomial variable $BIN(\delta n + o(n), \sqrt{\alpha} - \frac{2\delta}{\sqrt{\alpha}})$. The Chernoff Bound implies that w.h.p.

$$|A| \geq \delta(\sqrt{\alpha} - \frac{3\delta}{\sqrt{\alpha}})n. \tag{11}$$

Recall that STRIP1, KEEP-STRIPPING and U_i were based on viewing STRIP from the perspective of G_1. Now consider defining analogous procedures that view it from the perspective of G_2; i.e. we define $L_i', U_i', (L^*)', (U^*)', T^*$, and KEEP-STRIPPING$'$ by replacing G_1 by G_2 throughout the definitions of U_i, STRIP1 and KEEP-STRIPPING. We define J' to be the analogue of J.

We couple KEEP-STRIPPING and KEEP-STRIPPING$'$. Thus, they run in parallel and produce the same sets S_i, T_i, V_i. To be specific, note that we enter (*) during the same iteration in KEEP-STRIPPING and KEEP-STRIPPING$'$. Until that iteration, both these procedures are coupled with STRIP in the same way that STRIP1 is coupled with STRIP; thus they are coupled together. After entering (*), they are decoupled and U_i, U_i' evolve independently.

By symmetry, the same proof used for (10) yields that w.h.p.

$$|T^*| \geq |T_{I-1}| - (2 - \frac{\zeta}{2})\sqrt{\alpha}\delta n. \tag{12}$$

We define

$$A' = \left(\cup_{i=I}^{J'-1} L_i' \cap V_i\right) \cup ((L^*)' \cap V_{J'}).$$

By symmetry, the same proof used for (11) yields that w.h.p.

$$|A'| \geq \delta(\sqrt{\alpha} - \frac{3\delta}{\sqrt{\alpha}})n. \tag{13}$$

Note that if $J \geq J'$ and we do not enter (*) before iteration J then $A' \subseteq S_{I-1} \backslash S^*$. This is because: (a) $A' \subseteq V_I \subseteq S_{I-1}$ for the same reason that $A \subseteq V_I$ above. (b) For each iteration i before entering (*), $S_i = V_i \backslash L_i'$, just as $T_i = V_i \backslash L_i$. So every vertex in A' is in $V_i \backslash S_i$ for some $I \leq i \leq J' \leq J$ and hence is not in $S_J \supseteq S^*$

Similarly, if $J \leq J'$ and we do not enter (*) before iteration J' then $A, A' \subseteq T_{I-1} \backslash (S^*)'$.

Finally, we wish to bound the number of vertices in $A \cap A'$. If $J \geq J'$ and we do not enter (*) before iteration J, then we argued above that $A' \subseteq \cup_{i=I}^{J} V_i \backslash S_i$. It follows that every member of $A \cap A'$ is a member of $L_i \cap (V_i \backslash S_i)$ for some $I \leq i \leq J-1$ or $L^* \cap (V_J \backslash S_J)$.

Similarly, if $J \leq J'$ and we do not enter (*) before iteration J', then every member of $A \cup A'$ is a member of $L_i' \cap (V_i \backslash T_i)$ for some $I \leq i \leq J'-1$ or $(L^*)' \cap (V_{J'} \backslash T_{J'})$.

Recall that $V_i \backslash S_i \subseteq S_{I-1} \backslash S^*$ for $I \leq i \leq J$. So applying (10), we find that during iteration i, each of the $|L_i|$ (or $|L^*|$ if $i = J$) vertices chosen from U_i is in $V_i \backslash S_i$ with probability

$$\frac{|V_i \backslash S_i|}{|U_i|} < \frac{|S_{I-1} \backslash S_J|}{|U^*|} < \frac{(2 - \frac{\zeta}{2})\sqrt{\alpha}\delta n}{(\rho_I - \delta)n} < 3\delta.$$

So $\left|\left(\cup_{i=I}^{J-1} L_i \cap (V_i \backslash S_i)\right) \cup (L^* \cap (V_J \backslash S_J))\right|$ is dominated from above by the binomial variable $BIN(\delta n + o(n), 3\delta)$. So the Chernoff Bound yields that w.h.p.

$$\left|\left(\cup_{i=I}^{J-1} L_i \cap (V_i \backslash S_i)\right) \cup (L^* \cap (V_J \backslash S_J))\right| \leq 4\delta^2 n. \tag{14}$$

Similarly, w.h.p.

$$\left|\left(\cup_{i=I}^{J'-1} L'_i \cap (V_i \backslash T_i)\right) \cup ((L^*)' \cap (V_{J'} \backslash T_{J'}))\right| \leq 4\delta^2 n. \tag{15}$$

If $J \geq J'$ then $|A \cap A'|$ is bounded by (14), and if $J \leq J'$ then $|A \cap A'|$ is bounded by (15). Therefore, w.h.p.

$$|A \cap A'| \leq 4\delta^2 n. \tag{16}$$

If $J \geq J'$ and we do not enter (*) before iteration J, then $A \cup A' \subseteq S_{I-1} \backslash S^*$ and (11), (13), (16) imply that w.h.p.

$$|S^*| \leq |S_{I-1}| - 2 \times \delta(\sqrt{\alpha} - \frac{3\delta}{\sqrt{\alpha}})n + 4\delta^2 n < |S_{I-1}| - (2 - \frac{\zeta}{4})\sqrt{\alpha}\delta n,$$

for δ sufficiently small in terms of ζ, which contradicts (10).

Similarly, if $J \leq J'$ and we do not enter (*) before iteration J', then $A \cup A' \subseteq T_{I-1} \backslash T^*$ and (11), (13), (16) contradict (12).

Thus, w.h.p. KEEP-STRIPPING must enter (*) before $\max(J, J')$ iterations, and thus either before $\delta n + o(n)$ vertices are removed from U_I or before $\delta n + o(n)$ vertices are removed from U'_I. Either way, this is before $\delta n + o(n) < 2\delta n$ vertices are removed from V_I. This proves the lemma. □

The remaining part of our main theorem follows immediately:

Proof of Theorem 1.1(b): For any $\delta > 0$, Lemma 5.1 yields that STRIP will w.h.p. Halt Succeed and produce a set $S = T = V$ of size at most $(\alpha + \delta)n$ and at least $(\alpha - \delta)n - o(n)$. This will be a doubly connected set of (G_1, G_2).

Take $\delta < \phi = \phi(c)$ (from Lemma 3.1). The argument from the proof of Theorem 1.1(a) implies that no vertex outside of V_I is in a doubly connected set. Since $|V_I \backslash V| < \delta n + o(n) < \phi n$, every vertex in $V_j \backslash V_{j+1}, j \geq I$ is in a component of size less than ϕn in either $G_1 \cap V_j$ or $G_2 \cap V_j$. This, Lemma 3.3(c), and the same argument from the proof of Theorem 1.1(a) implies that no vertex outside of V is in a doubly connected set. Therefore V is the largest doubly connected set of (G_1, G_2).

Taking δ to be arbitrarily small yields Theorem 1.1(b). □

Acknowledgement

We would like to thank anonymous referees for helpful comments on an earlier draft of this paper.

References

[1] B. Bollobás, *Random Graphs.* 2nd edition. Cambridge University Press, 2001.

[2] B. Bollobás. *The evolution of random graphs.* Trans. Am. Math. Soc. **286** (1984), 257 - 274.

[3] B. Bollobás and O. Riordan, *Asymptotic normality of the size of the giant component via a random walk.* arXiv:1010.4595

[4] H. Chernoff. *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations.* Ann. Math. Statist. **23** 493 - 509 (1952).

[5] P. Erdős and A. Rényi. *On the evolution of random graphs.* Magayar Tud. Akad. Mat. Kutato Int. Kozl. **5** (1960), 17 - 61.

[6] D. Fernholz and V. Ramachandran. *Cores and Connectivity in Sparse Random Graphs.* The University of Texas at Austin, Department of Computer Sciences, technical report TR-04-13 (2004).

[7] S. Janson and M. Luczak. *A simple solution to the k-core problem.* Random Structures Algorithms **30** (2007) 50 - 62 (2007).

[8] S. Janson, T. Łuczak and A. Ruciński. Random Graphs. Wiley, New York (2000).

[9] J.H.Kim. *Poisson cloning model for random graphs.* arXiv:0805.4133v1

[10] T. Łuczak. *Component behaviour near the critical point of the random graph process.* Rand. Struc. & Alg. **1** (1990), 287 - 310.

[11] M. Molloy *Cores in random hypergraphs and boolean formulas.* Random Structures and Algorithms **27**, 124 - 135 (2005).

[12] M. Molloy *The freezing threshold for k-colourings of a random graph.* Proceedings of STOC 2012.

[13] M. Molloy and B. Reed, *Graph colouring and the probabilistic method*, Springer-Verlag, Berlin, 2001.

[14] B. Pittel, J. Spencer and N. Wormald, *Sudden emergence of a giant k-core in a random graph.* J.Comb.Th.(B) **67** (1996), 111 - 151.

[15] B. Pittel and N. Wormald, *Counting connected graphs inside-out.* J. Comb.Th. B **93** (2005), 127 - 172.

[16] V.E. Stepanov, *Some features of the structure of a random graph near a critical point.* (Russian) Teor. Veroyatnost. i Primenen. **32** (1987), 633657.