AN ELEMENTARY APPROACH TO COMPONENT SIZES IN CRITICAL RANDOM GRAPHS

UMBERTO DE AMBROGGIO ![ORCID],* *University of Bath*

Abstract

In this article we introduce a simple tool to derive polynomial upper bounds for the probability of observing unusually large maximal components in some models of random graphs when considered at criticality. Specifically, we apply our method to a model of a random intersection graph, a random graph obtained through *p*-bond percolation on a general *d*-regular graph, and a model of an inhomogeneous random graph.

Keywords: Random walk; ballot theorem

2020 Mathematics Subject Classification: Primary 05C80
Secondary 60G50; 60C05

1. Introduction

The purpose of this paper is to introduce an elementary tool to obtain simple upper bounds for the probability of observing unusually large maximal components in some important models of random graphs at criticality.

Given any (undirected) graph $\mathbb{G} = (V, E)$ and vertices $v, u \in V$, we write $v \sim u$ if the edge $\{v, u\}$ is present in $\mathbb{G}$ and say that vertices u and v are neighbours. We write $v \leftrightarrow u$ if there exists a path of occupied edges connecting vertices v and u and we adopt the convention that $v \leftrightarrow v$ for every $v \in V$. We let $\mathcal{C}(v) = \{u \in V : v \leftrightarrow u\}$ be the component (or cluster) of vertex $v \in V$ and denote its size by $|\mathcal{C}(v)|$. Moreover, we define a largest component $\mathcal{C}_{\max}$ to be any cluster $\mathcal{C}(v)$ for which $|\mathcal{C}(v)|$ is maximal, so that $|\mathcal{C}_{\max}| = \max_{v \in V} |\mathcal{C}(v)|$.

The Erdős–Rényi random graph on $[n] = \{1, \ldots, n\}$, denoted by $\mathbb{G}(n, p)$, is the random graph obtained from the complete graph on n vertices by independently retaining each edge with probability $p \in [0, 1]$ and deleting it with probability $1 - p$.

One of the most surprising aspects of this model is that when p is of the form $p = p(n) = \gamma/n$. then the $\mathbb{G}(n, p)$ random graph undergoes a *phase transition* as γ passes 1. Specifically, if $\gamma < 1$ then $|\mathcal{C}_{\max}|$ is of order $\log(n)$, if $\gamma = 1$ then $|\mathcal{C}_{\max}|$ is of order $n^{2/3}$, and if $\gamma > 1$ then $|\mathcal{C}_{\max}|$ is of order n. See for instance the books [3], [13], or [15] for proofs of these statements and other interesting properties of this model. See also Krivelevich and Sudakov [19] for a simple proof of the phase transition in $\mathbb{G}(n, p)$.

In [9] De Ambroggio and Roberts introduced a ballot-type result (Lemma 1 below) to provide a new, purely probabilistic proof of the fact that in the $\mathbb{G}(n, p)$ model considered in the so-called *critical window*, i.e. when p is of the form $p = p(n) = n^{-1} + \lambda n^{-4/3}$, the probability

---

of observing a maximal cluster of size larger than $An^{2/3}$ tends to zero (as $n \to \infty$) exponentially fast in A. More precisely, they proved that, for large enough n,

$$\frac{c}{A^{3/2}} \exp\left(-\frac{A^3}{8} + \frac{\lambda A^2}{2} - \frac{\lambda^2 A}{2}\right) \leq \mathbb{P}\left(|\mathcal{C}_{\max}| > An^{2/3}\right) \leq \frac{c'}{A^{3/2}} \exp\left(-\frac{A^3}{8} + \frac{\lambda A^2}{2} - \frac{\lambda^2 A}{2}\right), \tag{1}$$

where $0 < c \leq c'$ are two finite constants, thus showing that in the near-critical $\mathbb{G}(n, p)$ model the number of vertices contained in the maximal component is unlikely to be much larger than $n^{2/3}$.

We remark that the correct asymptotic for $\mathbb{P}\left(|\mathcal{C}_{\max}| > An^{2/3}\right)$ in this critical model was obtained first by Pittel [26] (whose paper is partially based on an earlier article by Łuczak, Pittel, and Wierman [21]) and more recently by Roberts [27]. We also mention that Nachmias and Peres [23] used a general martingale argument to establish an exponential upper bound for the probability in (1), but their bound is not optimal.

The purpose of this work is to show that part of the argument used in [9] to prove the upper bound in (1) is quite general and can be used to obtain, in a surprisingly simple way, polynomial upper bounds for $\mathbb{P}(|\mathcal{C}_{\max}| > k)$ in different models of random graphs when considered at criticality. Specifically, we apply our method to three different models, namely a model of a random intersection graph, a random graph obtained through p-bond percolation on a general d-regular graph, and a model of an inhomogeneous random graph, and we show that $|\mathcal{C}_{\max}|$ is unlikely to be much larger than $n^{2/3}$ in these models. In this sense, these random graphs exhibit a similar critical behaviour.

2. Results

In order to better understand the statement of our main result (Theorem 1 below), we first need to recall the definition of an *exploration process*, which is an algorithmic procedure used to reveal the components of a given graph; see e.g. [9], [22], [23], [27], and references therein. As we will see in a moment, when the graph under investigation is *random* such an exploration process reduces the study of component sizes to the analysis of the trajectory of a random process, which looks like (but is not quite) a random walk.

Let $\mathbb{G} = ([n], E)$ be any (undirected) graph, and let V_n be a vertex selected uniformly at random from $[n]$. During the exploration of $\mathcal{C}(V_n)$, each vertex will be either *active*, *explored*, or *unseen*, and its status will change during the course of the exploration. At each step $t \in \{0\} \cup [n]$ of the algorithm, the number of explored vertices will be t, whereas the number of active vertices will be denoted by Y_t. At time $t = 0$, if V_n is an isolated vertex we stop the procedure; otherwise there exists some vertex $u \in [n] \setminus V_n$ with $\{V_n, u\} \in E$. In this case vertices V_n and u are declared active, whereas all other vertices are declared unseen (so that $Y_0 = 2$). At each step $t \in [n]$ of the algorithm we proceed as follows.

(a) If $Y_{t-1} > 0$, we let u_t be the active vertex with the smallest label. We reveal all unseen neighbours of u_t in $\mathbb{G}$ and change the status of these vertices to active. Then we set u_t itself to be explored.

(b) If $Y_{t-1} = 0$, we let u_t be the unseen vertex with the smallest label, and:

(b.1) if u_t is isolated, we halt the procedure;

(b.2) otherwise, there is at least one unseen vertex v such that $\{u_t, v\} \in E$, and we declare both u_t and v active; then we continue with step (a).

Letting η_t denote the number of unseen vertices in $\mathbb{G}$ which become active at step t of the exploration process, we see that:

(i) if $Y_{t-1} > 0$ then $Y_t = Y_{t-1} + \eta_t - 1$;

(ii) if $Y_{t-1} = 0$ then $Y_t = \eta_t$.

Remark 1. We remark that our description of an exploration process is slightly different from those provided in [9], [23], and [27], for example. Indeed, in our setting the algorithm is actually run only in the case where V_n is not an isolated vertex, and the exploration starts from two active vertices and not from one active vertex, as usually happens. Moreover, whenever $Y_{t-1} = 0$ and u_t is *not* isolated, we first reveal one of the neighbours of u_t before proceeding with the exploration. These small modifications will be particularly useful in one of our applications.

Now let $\mathbb{G} = ([n], E)$ be any (undirected) random graph, and use the above algorithm to reveal the components of $\mathbb{G}$. It is clear that the η_i are now random variables. Observe that, given any $k \in \mathbb{N} = \{1, 2, \dots\}$, if V_n is an isolated vertex then $|\mathcal{C}(V_n)| = 1$ and hence, in particular, we cannot have $|\mathcal{C}(V_n)| > k$ (as $k \geq 1$). On the other hand, if V_n is not isolated then $|\mathcal{C}(V_n)| > k$ implies that $Y_t = 2 + \sum_{i=1}^{t}(\eta_i - 1) > 0$ for all $t \in [k]$. Therefore we can write

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \mathbb{P}\left(2 + \sum_{i=1}^{t}(\eta_i - 1) > 0 \text{ for all } t \in [k]\right). \tag{2}$$

Note that the η_i are not independent and, moreover, they have different distributions (one of the reasons is that the number of unseen vertices in the graph decreases during the course of the exploration). Therefore Y_t does not define a random walk.

In order to bound the probability in (2) from above, the idea is to produce a sequence of independent and identically distributed (i.i.d.) random variables X_i, *bigger* than the η_i, that allow us to replace the probability on the right-hand side of (2) with the probability that a random walk (started at 2) stays positive up to time k.

In some random graphs this is an immediate consequence of the model construction, while in other instances one needs more care in order to produce these X_i.

Here is our main result.

Theorem 1. *Let $\mathbb{G} = ([n], E)$ be any (undirected) random graph. Suppose that there exists a sequence of i.i.d. random variables $(X_i)_{i \geq 1}$ taking values in $\mathbb{N}_0$, such that the distribution of X_1 may depend on n, satisfying:*

(i) $\mathbb{P}(X_1 = 3) \geq c$ *for all sufficiently large n, for some constant $c > 0$;*

(ii) $\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \mathbb{P}\left(2 + \sum_{i=1}^{t}(X_i - 1) > 0 \text{ for all } t \in [k]\right)$ *for every $k = k(n) \in \mathbb{N}$;*

(iii) *there exist $\delta, \rho = \rho(n) > 0$ and $\epsilon = \epsilon(n) \geq 0$ with $\epsilon_n \to 0$ (as $n \to \infty$) such that $\mathbb{E}(e^{rX_1}) \leq e^{r(1+\epsilon)+r^2\delta}$ for every $r \in (0, \rho)$ and all sufficiently large n.*

Suppose that $k = k(n) \in \mathbb{N}$ satisfies $\epsilon\sqrt{k} \leq c$ and $\rho\sqrt{k} \geq 1$ for all large enough n, for some finite constant $c > 0$. Then

$$\mathbb{P}(|\mathcal{C}_{\max}| > k) \leq \frac{C}{\mathbb{P}(X_1 = 3)} \frac{n}{k^{3/2}}$$

for all sufficiently large n, where $C > 0$ is a finite positive constant which depends solely on δ and c.

Remark 2. In all our applications the probability $\mathbb{P}(X_1 = 3)$ is bounded away from zero, so that if $k = k(n) = \lceil An^{2/3} \rceil$ satisfies the two assumptions in the statement of the theorem, then $\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3} \rceil)$ is indeed $\mathrm{O}(A^{-3/2})$.

Remark 3. We note that condition (iii) in Theorem 1 might be stated in different (possibly more general) terms, but we decided to state it in this way because of its simplicity to verify, as shown in our applications.

Our claim that the approach introduced in [9] is robust and that Proposition 1 leads to simple upper bounds for $\mathbb{P}(|\mathcal{C}_{\max}| > k)$ in several models of random graphs at criticality is justified in Sections 2.1, 2.2, and 2.3 below, where we use Proposition 1 to obtain polynomial upper bounds for the above probability in three particular models of random graphs.

We remark that our methodology does not lead to upper bounds for the probability of observing unusually *small* largest components; in this direction the martingale argument introduced by Nachmias and Peres in [23] seems to be more robust and adaptable to different models of random graphs.

2.1. Critical random intersection graph

Our first application of Theorem 1 involves a model of a random intersection graph; for an introduction to this class of models, we refer the reader to [11].

Here we are interested in the random graph described by Lageras and Lindholm [20]. Such a random graph, denoted by $\mathbb{G}(n, m, p)$, with a set of vertices $V = \{v_i : i \in [n]\}$ and a set of edges E, is constructed from a bipartite graph $\mathbb{B}(n, m, p)$ with two sets of vertices: $A = \{a_j : j \in [m]\}$, which we call the set of *auxiliary vertices*, and V (i.e. the vertex set of $\mathbb{G}(n, m, p)$). Edges in $\mathbb{B}(n, m, p)$ between vertices and auxiliary vertices are present independently with probability $p \in [0, 1]$. Two distinct vertices v_i and v_j are neighbours in $\mathbb{G}(n, m, p)$ (i.e. $\{v_i, v_j\} \in E$) if and only if there exists at least one $a_k \in A$ such that both edges $\{a_k, v_i\}$ and $\{a_k, v_j\}$ are present in the bipartite graph $\mathbb{B}(n, m, p)$.

We are interested in the case where $p = p(n) = \gamma/n^{(1+\alpha)/2}$ and $m = m(n) = \lfloor \beta n \rfloor$, where $\alpha, \beta, \gamma > 0$ are fixed parameters of the model.

Stark [28] has shown that the vertex degree distribution (i.e. the distribution of the degree of a vertex selected uniformly at random) is highly dependent on the value of α. However, as shown by Deijfen and Kets [10], the clustering is controllable only when $\alpha = 1$.

The component structure of the graph was studied for $\alpha \neq 1$, $\gamma > 0$, and $\beta = 1$ by Behrisch [2], whereas it was studied for $\alpha = 1$ and $\beta, \gamma > 0$ in [20]. Specifically, Lageras and Lindholm [20] proved that the $\mathbb{G}(n, m, p)$ model undergoes a phase transition as $\beta\gamma^2$ passes 1. Indeed, setting $\mu = \beta\gamma^2$, they proved that if $\mu < 1$ (sub-critical case), then with probability tending to one there is no component in $\mathbb{G}(n, m, p)$ with more than $\mathrm{O}(\log(n))$ vertices, while if $\mu > 1$ (super-critical case), then with probability tending to one there exists a unique giant component of size $n\delta$ where $\delta \in (0, 1)$, and the size of the second largest component is at most of order $\log(n)$.

By means of Theorem 1 we show that, in the *critical* case $\mu = 1$, it is unlikely for the largest component to contain more than $n^{2/3}$ vertices.

Proposition 1. *Let $\mathbb{G}(n, m, p)$ be the random intersection graph described above. Let $m = \lfloor \beta n \rfloor$, $p = \gamma/n$, and $\mu = \beta\gamma^2$. If $\mu = 1$ then, given any $A > 1$, when n is sufficiently large we have*

$$\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3} \rceil) \leq \frac{c_1}{A^{3/2}},$$

where c_1 is a finite constant which depends solely on γ and β.

2.2. Critical *p*-bond percolation on *d*-regular graph

In this section we consider a second application of Theorem 1. Here we analyse a random graph $\mathbb{G}_p$ obtained through *p*-bond percolation on a general *d*-regular graph.

In [22] Nachmias and Peres adapted the martingale method they developed in [23] to prove that, for any $d \geq 3$, when $p \leq (d-1)^{-1}$ then

$$\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3} \rceil) \leq \frac{8}{A^{3/2}}; \tag{3}$$

see [22, Proposition 1.2]. For a random regular graph $\mathbb{G}(n, d, p)$ they were also able to sharpen the upper bound in (3) and to prove a corresponding lower bound. (The $\mathbb{G}(n, d, p)$ random graph is obtained by the following two-step procedure: first we draw uniformly at random a graph from the set of all simple *d*-regular graphs on $[n]$, and then we retain each edge independently with probability *p* and delete it with probability $1-p$.) Specifically, in Theorem 2 of [22] it is shown that, when *p* is of the form

$$p = p(n, d) = \left(1 + \lambda n^{-1/3}\right)/(d-1) \quad (\lambda \in \mathbb{R}) \tag{4}$$

and $d \geq 3$ is *fixed*, then there are constants C_1, $C_2 \in (0, \infty)$ depending on λ and *d* such that, for every $A > 0$ and all *n*, $\mathbb{P}(|\mathcal{C}_{\max}| > An^{2/3}) \leq A^{-1}C_1 e^{-C_2 A^3}$. In [22] it is also shown that there exists a constant $C_3 \in (0, \infty)$ (also depending on λ and *d*) such that, for large enough *A* and all *n*, then $\mathbb{P}(|\mathcal{C}_{\max}| < \lceil A^{-1}n^{2/3} \rceil) \leq C_3 A^{-1/2}$, thus proving that the size of $|\mathcal{C}_{\max}|$ is indeed of order $n^{2/3}$ in this model when considered at criticality.

We remark that in [22] the parameter *d* is not allowed to depend on *n*. The problem of determining the size of $|\mathcal{C}_{\max}|$ in the critical $\mathbb{G}(n, d, p)$ model when $d = d(n)$ depends on *n* has been investigated by Joos and Perarnau [16], where the authors proved (among many other things) that for any $d \in \{3, \ldots, n-1\}$ and when *p* is of the form (4), then for all sufficiently large *n* and $A = A(\lambda)$ we have that $\mathbb{P}(|\mathcal{C}_{\max}| \notin [A^{-1}n^{2/3}, An^{2/3}]) \leq 20/\sqrt{A}$.

Our goal here is to show that, by means of Theorem 1, we can recover (up to a multiplicative constant) the bound in (3), in a very simple way.

Proposition 2. *Let $\mathbb{G}$ be a d-regular graph, $d > 3$, and let $\mathbb{G}_p$ denote the random graph obtained by bond percolation on $\mathbb{G}$ with probability p. If $p \leq 1/(d-1)$ then, given any $A > 1$, when n is sufficiently large we have*

$$\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3} \rceil) \leq \frac{c_2}{A^{3/2}}$$

for some finite positive constant c_2 which depends solely on d.

Remark 4. The requirement $d > 3$ is needed because the i.i.d. random variables $(X_i)_{i \geq 1}$ that dominate the η_i in the exploration process satisfy

$$\mathbb{P}(X_1 = 3) = \frac{(d-2)(d-3)}{6(d-1)^2} \left(1 - \frac{1}{d-1}\right)^{d-4};$$

see Section 3.3. Hence, if we want $\mathbb{P}(X_1 = 3) > 0$ (a condition required in Theorem 1), we do need $d > 3$.

2.3. Critical inhomogeneous random graph

In this section we discuss our final application of Theorem 1. In the random graph model that we investigate here, the *n* vertices are endowed with weights, and edges between a pair of vertices are placed independently with probabilities moderated by such weights.

Specifically, let $\mathbf{w} = (w_i)_{i \in [n]}$ be a sequence of positive real numbers, which we call the sequence of *vertex weights*; we think of w_i as the weight assigned to vertex $i \in [n]$. Define $l_n = \sum_{i \in [n]} w_i$, the sum of all weights.

We consider the so-called Norros–Reittu random graph [24] as described by Van der Hofstad [12]. This is an inhomogeneous random graph, that we denote by $NR_n(\mathbf{w})$, in which the probability that the edge $\{i, j\}$ is present in $NR_n(\mathbf{w})$ (for $1 \leq i < j \leq n$) is given by

$$p_{ij}^{NR} = \mathbb{P}(\{i, j\} \in E(NR_n(\mathbf{w}))) = 1 - \mathrm{e}^{-w_i w_j / l_n},$$

and edges are present independently.

Inhomogeneous random graphs were studied extensively by Bollobás, Janson, and Riordan [4]. As explained by Janson [14] and further noted by Van der Hofstad [12], the $NR_n(\mathbf{w})$ random graph is closely related to the models studied by Chung and Lu [5, 6, 7] and Norros and Reittu [24], so that the results proved for the $NR_n(\mathbf{w})$ random graph apply as well to these other models.

Other models of inhomogeneous random graphs have been studied more recently by Penrose [25] and by Kang, Pachon, and Rodriguez [18].

It is clear that the topology of the $NR_n(\mathbf{w})$ model depends on the choice of the sequence $\mathbf{w}$, which we now specify.

Let $F \colon (0, \infty) \mapsto [0, 1]$ be a distribution function, and define

$$[1 - F]^{-1}(u) = \inf\{s \colon [1 - F(s)] \leq u\}, \quad u \in (0, 1).$$

By convention, we set $[1 - F]^{-1}(1) = 0$. We construct the weights as in [12], namely we set

$$w_j = [1 - F]^{-1}(j/n), \quad j \in [n]. \tag{5}$$

In [4, Theorem 3.13] it was shown that in the $NR_n(\mathbf{w})$ random graph with vertex weights as in (5), the proportion of vertices having degree $k \geq 0$, denoted by N_k, converges in probability (as $n \to \infty$) to

$$p_k = \mathbb{E}\left(\mathrm{e}^{-W} \frac{W^k}{k!}\right),$$

where W is a random variable taking values in $(0, \infty)$ with distribution function F. The limiting sequence $(p_k)_{k \geq 0}$ is a so-called *mixed Poisson distribution with mixing distribution F*.

In order to describe the phase transition for the size of the largest component, define

$$\nu = \mathbb{E}(W^2) / \mathbb{E}(W).$$

As explained by Van der Hofstad [12] (see also [8]), this (positive) real number corresponds to the asymptotic mean of the offspring distribution in a branching process approximation of the exploration of $\mathcal{C}(V_n)$.

In [4, Theorem 3.1] it is shown that the graph undergoes a phase transition as ν passes 1. In particular, if $\nu > 1$, the largest component contains a positive proportion of the total number of vertices, whereas if $\nu \leq 1$ the largest component contains a vanishing proportion of vertices.

Van der Hofstad [12] provided a complete picture of the component structure in the critical $NR_n(\mathbf{w})$ model. Specifically, he proved that in the case where

$$\lim_{x \to \infty} x^{\tau - 1}(1 - F(x)) = c_F \tag{6}$$

for some constant $c_F > 0$ and some $3 < \tau < 4$, there is a constant $b > 0$ such that for all $A > 1$ and $n \geq 1$, the $NR_n(\mathbf{w})$ random graph satisfies

$$\mathbb{P}(|\mathcal{C}_{\max}| \notin [A^{-1}n^{(\tau-2)/(\tau-1)}, An^{(\tau-2)/(\tau-1)}]) \leq b/A. \tag{7}$$

On the other hand, when

$$1 - F(x) \leq c_F x^{-(\tau-1)} \quad (x \geq 0) \tag{8}$$

for some $c_F > 0$ and some $\tau > 4$, then there is a constant $b > 0$ such that, for all $n \geq 1$ and all $A > 1$, the $NR_n(\mathbf{w})$ random graph satisfies

$$\mathbb{P}(|\mathcal{C}_{\max}| \notin [A^{-1}n^{2/3}, An^{2/3}]) \leq b/A. \tag{9}$$

(In fact Van der Hofstad [12] proved a more general result, namely that the lower bounds (7) and (9) also remain valid after a *small perturbation* of the vertex weights; see [12, Theorems 1.1 and 1.2].)

For a heuristic explanation of the critical behaviour described by (7) and (9), we refer to [12, Section 1.3].

We also mention that De Ambroggio and Pachon [8] used the first part of the martingale argument introduced by Nachmias and Peres [23] to obtain simple upper bounds for the probability of observing unusually large maximal components in the (critical) $NR_n(\mathbf{w})$ random graph for both regimes $\tau \in (3, 4)$ and $\tau > 4$, even if in the former case (i.e. for $\tau \in (3, 4)$) the distribution function F is required to satisfy a stronger condition with respect to the one stated in (6).

Our goal here is to use Theorem 1 to provide a very simple proof of the fact that, in the critical $NR_n(\mathbf{w})$ model with vertex weights as in (5) and distribution function F satisfying (8), the largest component is unlikely to contain more than $n^{2/3}$ vertices. More precisely, we prove the following.

Proposition 3. *Consider the $NR_n(\mathbf{w})$ random graph with weights defined as in (5) above. Suppose that there exists a constant $c_F > 0$ and a $\tau > 4$ such that $1 - F(x) \leq c_F x^{-(\tau-1)}$ for all $x \geq 0$. Then, given any $A > 1$, when n is large enough we have that*

$$\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3} \rceil) \leq \frac{c_3}{A^{3/2}},$$

where c_3 is a finite positive constant which depends solely on c_F and τ.

3. Proofs

Here we are going to prove the results stated in Section 2. We start by proving Theorem 1 and subsequently we prove the remaining results, namely Propositions 1, 2, and 3.

The proof of Theorem 1 relies on the following *ballot-type* estimate, which is taken from [9]. For a general introduction to classical ballot theorems and their generalisations, see for instance [1], [17], and references therein.

Lemma 1. *Fix $n \in \mathbb{N}$ and let $(W_i)_{i \geq 1}$ be a sequence of i.i.d. valued random variables taking values in $\mathbb{Z}$. Let $r \in \mathbb{N}$, and suppose that $\mathbb{P}(W_1 = r) > 0$. Define $S_t = \sum_{i=1}^{t} W_i$ for $t \in \mathbb{N}_0$. Then, for any $j \geq 1$, we have*

$$\mathbb{P}(r + S_t > 0 \text{ for all } t \in [n], r + S_n = j) \leq \mathbb{P}(X_1 = r)^{-1} \frac{j}{n+1} \mathbb{P}(S_{n+1} = j).$$

3.1. Proof of Theorem 1

Let $k = k(n) \in \mathbb{N}$. By hypothesis, there is a sequence of i.i.d. random variables X_i taking values in $\mathbb{N}_0$ such that, setting $S_t = \sum_{i=1}^{t}(X_i - 1)$,

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \mathbb{P}(2 + S_t > 0 \text{ for all } t \in [k]).$$

Using Lemma 1 with $W_i = X_i - 1$ and $r = 2$, we obtain

$$\mathbb{P}(2 + S_t > 0 \text{ for all } t \in [k]) = \sum_{h=1}^{\infty} \mathbb{P}(2 + S_t > 0 \text{ for all } t \in [k], \, 2 + S_k = h)$$

$$\leq a \sum_{h=1}^{\infty} \frac{h}{k+1} \mathbb{P}(S_{k+1} = h), \tag{10}$$

where we set $a = 1/\mathbb{P}(X_1 = 3) \in (0, \infty)$. Now let m be a non-negative integer to be specified later. By splitting the series in (10) at $h = m$ we can write

$$a \sum_{h=1}^{\infty} \frac{h}{k+1} \mathbb{P}(S_{k+1} = h) \leq a \frac{m}{k+1} + \frac{a}{k+1} \sum_{h=m+1}^{\infty} h \mathbb{P}(S_{k+1} = h). \tag{11}$$

Now the series in (11) equals

$$\frac{a}{k+1} \sum_{h=m+1}^{\infty} h \mathbb{P}\left(\sum_{i=1}^{k+1} X_i = h + k + 1\right)$$

$$= \frac{a}{k+1} \sum_{z=m+k+2}^{\infty} z \mathbb{P}\left(\sum_{i=1}^{k+1} X_i = z\right) - a \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq m + k + 2\right). \tag{12}$$

To proceed, we observe the following: if X is a random variable taking values in $\mathbb{N}_0$, then for any $h \geq 1$, we have

$$\mathbb{E}(X \mathbb{1}_{\{X \geq h\}}) = \mathbb{E}\left(\sum_{i=1}^{\infty} \mathbb{1}_{\{i \leq X\}} \mathbb{1}_{\{X \geq h\}}\right) = h \mathbb{P}(X \geq h) + \sum_{i=h+1}^{\infty} \mathbb{P}(X \geq i).$$

Thus the series in (12) equals

$$a \frac{m+1}{k+1} \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq m + k + 2\right) + a \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq m + k + 2\right) + \frac{a}{k+1} \sum_{z=m+k+3}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq z\right).$$

Substituting the series in (12) with these three terms, we obtain

$$\frac{a}{k+1} \sum_{h=m+1}^{\infty} h \mathbb{P}(S_{k+1} = h) = a \frac{m}{k+1} \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq m + k + 2\right) + \frac{a}{k+1} \sum_{z=m+k+2}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq z\right). \tag{13}$$

Now observe that the series in (13) can be rewritten as follows:

$$\frac{a}{k+1} \sum_{z=m+k+2}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq z\right) = \frac{a}{k+1} \sum_{h=m+1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq h + k + 1\right).$$

Summarizing, so far we have shown that

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq a\frac{m}{k+1} + a\frac{m}{k+1}\mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq k + 1 + (m+1)\right)$$

$$+ \frac{a}{k+1}\sum_{h=m+1}^{\infty}\mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq k + 1 + h\right). \tag{14}$$

Using our assumption on $\mathbb{E}(e^{rX_1})$ and Markov's inequality, we have, for all $h \geq m+1$ and all $r \in (0, \rho)$,

$$\mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq k + 1 + h\right) \leq e^{-r(k+1)-rh}\mathbb{E}(e^{rX_1})^{k+1}$$

$$\leq \exp\{-r(k+1) - rh + r(1+\epsilon)(k+1) + \delta r^2(k+1)\}$$

$$= \exp\{-rh + r\epsilon(k+1) + \delta r^2(k+1)\}.$$

Now if k is such that $\rho\sqrt{k} \geq 1$ for all large enough n, then $r = 1/\sqrt{k+1} < \rho$, and hence using this specific value of r we obtain

$$\mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq k + 1 + h\right) \leq e^{-h/\sqrt{k+1}}\,e^{\epsilon\sqrt{k+1}+\delta}.$$

Also, if k satisfies $\epsilon\sqrt{k} \leq c$ for all sufficiently large n, we see that $\epsilon\sqrt{k+1} \leq 2c$ and hence we can bound the series in (14) from above as follows:

$$\frac{a}{k+1}\sum_{h=m+1}^{\infty}\mathbb{P}\left(\sum_{i=1}^{k+1} X_i \geq h + k + 1\right) \leq \frac{a\,e^{\delta+2c}}{k+1}\sum_{h=m+1}^{\infty} e^{-h/\sqrt{k+1}}. \tag{15}$$

Now observe that

$$\sum_{h=m+1}^{\infty} e^{-h/\sqrt{k+1}} = e^{-(m+1)/\sqrt{k+1}}\frac{1}{1 - e^{-1/\sqrt{k+1}}}.$$

Using the inequality $e^{-x} \leq 1 - x + x^2/2$ (which is valid for all $x \geq 0$), we see that

$$1 - e^{-1/\sqrt{k+1}} \geq \frac{1}{\sqrt{k+1}} - \frac{1}{2(k+1)} \geq \frac{1}{2\sqrt{k+1}},$$

and hence the expression on the right-hand side of (15) is at most

$$\frac{a\,e^{\delta+2c}}{k+1}e^{-(m+1)/\sqrt{k+1}}2\sqrt{k+1} = \frac{2a\,e^{\delta+2c}}{\sqrt{k+1}}e^{-(m+1)/\sqrt{k+1}}.$$

Thus we obtain

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq a\frac{m}{k+1} + a\,e^{\delta+2c}\frac{m}{k+1}e^{-(m+1)/\sqrt{k+1}} + \frac{2a\,e^{\delta+2c}}{\sqrt{k+1}}e^{-(m+1)/\sqrt{k+1}}.$$

Taking $m = \lfloor \sqrt{k+1} \rfloor$ we see that

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \frac{a}{\sqrt{k+1}} + 3a\,\mathrm{e}^{\delta+2c}\frac{\mathrm{e}^{-1}}{\sqrt{k+1}} = C\frac{a}{\sqrt{k+1}},$$

where we set $C = C(\delta, c) = 1 + 3\,\mathrm{e}^{\delta+2c-1}$. Finally, letting

$$N_k = \sum_{v \in [n]} \mathbb{1}_{\{|\mathcal{C}(v)| > k\}}$$

denote the number of vertices that are contained in components of size at least k, we obtain

$$\mathbb{P}(|\mathcal{C}_{\max}| > k) = \mathbb{P}(N_k > k) \leq \frac{1}{k}\mathbb{E}(N_k) = \frac{n}{k}\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \frac{C}{\mathbb{P}(X_1 = 3)}\frac{n}{k^{3/2}},$$

completing the proof of the theorem. □

3.2. Proof of Proposition 1

Let $\mathbb{H}(n, m, p)$ be a random (multi-)graph constructed from the bipartite graph $\mathbb{B}(n, m, p)$ by letting the number of edges between $v_i, v_j \in V$ equal the number of auxiliary vertices a_k that are adjacent to both v_i and v_j. (Recall that V is the vertex set of the random intersection graph $\mathbb{G}(n, m, p)$ under investigation.) Notice that $\mathbb{G}(n, m, p)$ can be obtained from $\mathbb{H}(n, m, p)$ by coalescing multiple edges between vertices into one single edge. Hence, thanks to this construction, we see that the degree distribution in $\mathbb{G}(n, m, p)$ is dominated by the degree distribution in $\mathbb{H}(n, m, p)$. Notice that the latter is a compound binomial distribution with moment generating function

$$\mathbb{E}(\mathrm{e}^{rX_1}) = \left\{ 1 - \frac{\gamma}{n} + \frac{\gamma}{n}\left(1 - \frac{\gamma}{n} + \frac{\gamma}{n}\,\mathrm{e}^r \right)^{n-1} \right\}^{\lfloor \beta n \rfloor}, \tag{16}$$

since (by construction) a vertex $v \in \mathbb{H}(n, m, p)$ is connected to a $\mathrm{Bin}(m, p)$ number of auxiliary vertices, and each one of them is connected to an independent $\mathrm{Bin}(n-1, p)$ number of vertices in $V \setminus \{v\}$.

Therefore, by revealing the components of $\mathbb{G}(n, m, p)$ using the exploration process described at the beginning of Section 2, we can write

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \mathbb{P}\left(2 + \sum_{i=1}^{t} (X_i - 1) > 0 \text{ for all } t \in [k] \right),$$

where the X_i are i.i.d. compound binomial random variables with moment generating function given in (16).

Using the probability generating function of X_1 (which coincides with (16) after substituting e^r with r), it is not difficult to show that (for large enough n) the probability $\mathbb{P}(X_1 = 3)$ is bounded from below by a positive constant which depends solely on γ and β.

Next, in order to apply Theorem 1, we simply need to prove an upper bound for $\mathbb{E}(\mathrm{e}^{rX_1})$. Recalling the expression of the moment generating function of X_1 given in (16), we obtain

$$\mathbb{E}(\mathrm{e}^{rX_1}) = \exp\left\{ \lfloor \beta n \rfloor \log\left(1 - \frac{\gamma}{n} + \frac{\gamma}{n}\left[1 + \frac{\gamma}{n}(\mathrm{e}^r - 1) \right]^{n-1} \right) \right\}. \tag{17}$$

Taking $r \in (0, 1)$, we have that $e^r - 1 \leq r + r^2$. Then, since $1 + x \leq e^x$ for all $x \in \mathbb{R}$, we see that (17) is at most

$$\exp\left\{\lfloor \beta n \rfloor \log\left(1 + \frac{\gamma}{n}(\exp\{\gamma(r + r^2)\} - 1)\right)\right\}.$$

Thus, taking $r < \min\{1, 1/2\gamma\}$ (so that $\gamma(r + r^2) < 1$), we see that

$$\exp\{\gamma(r + r^2)\} - 1 \leq \gamma(r + r^2) + \gamma^2(r + r^2)^2$$

and hence, using the fact that $\log(1 + x) \leq x$ for all $x > -1$, we can write

$$\exp\left\{\lfloor \beta n \rfloor \log\left(1 + \frac{\gamma}{n}(\exp\{\gamma(r + r^2)\} - 1)\right)\right\} \leq \exp\left\{\lfloor \beta n \rfloor \left(\frac{\gamma^2}{n}(r + r^2) + \frac{\gamma^3}{n}(r + r^2)^2\right)\right\}.$$

Recalling that $\beta\gamma^2 = 1$, we obtain

$$\exp\left\{\lfloor \beta n \rfloor \left(\frac{\gamma^2}{n}(r + r^2) + \frac{\gamma^3}{n}(r + r^2)^2\right)\right\} \leq \exp\{r + r^2(1 + 4\gamma)\}.$$

Therefore, for all $r \in (0, \min\{1, 1/2\gamma\})$, we have

$$\mathbb{E}(e^{rX_1}) \leq e^{r + r^2(1 + 4\gamma)}$$

and hence condition (iii) in Theorem 1 is satisfied for $\rho = \min\{1, 1/2\gamma\}$, $\delta = 1 + 4\gamma$, and $\epsilon = 0$. Hence, taking $k = \lceil An^{2/3}\rceil$ (which clearly satisfies the requirement $\epsilon\sqrt{k} \leq c$ for $c = 0$ as well as the condition $\rho\sqrt{k} \geq 1$) and applying Theorem 1, we obtain

$$\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3}\rceil) \leq \frac{c_1}{A^{3/2}}$$

for some finite positive constant c_1 which depends solely on γ and β. $\qquad\square$

3.3. Proof of Proposition 2

Since $\mathbb{G}$ is d-regular, we can use the exploration process described at the beginning of Section 2 to conclude that

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \mathbb{P}\left(2 + \sum_{i=1}^{t}(X_i - 1) > 0 \text{ for all } t \in [k]\right),$$

where the X_i are i.i.d. random variables with $X_i \sim \text{Bin}(d - 1, p)$, so that

$$\sum_{i=1}^{k+1} X_i \sim \text{Bin}((k + 1)(d - 1), p).$$

(Note that if we had started the exploration process with only one active vertex, now we would have $\eta_1 \sim \text{Bin}(d, p)$, and hence in particular it would be impossible to dominate η_1 with a $\text{Bin}(d - 1, p)$ random variable.) Using a monotonicity argument we can focus on the (critical) case $p = 1/(d - 1)$. Note that, since $d > 3$,

$$\mathbb{P}(X_1 = 3) = \frac{(d - 2)(d - 3)}{6(d - 1)^2}\left(1 - \frac{1}{d - 1}\right)^{d-4} > 0.$$

Next, for all $r \in (0, 1)$, using the inequality $1 + x \leq e^x$ (valid for all $x \in \mathbb{R}$) we see that

$$\mathbb{E}(e^{rX_1}) = \left(1 + \frac{1}{d-1}(e^r - 1)\right)^{d-1} \leq e^{e^r - 1} \leq e^{r+r^2}.$$

Hence condition (iii) of Theorem 1 is satisfied for ρ, $\delta = 1$ and $\epsilon = 0$. Thus, taking $k = \lceil An^{2/3} \rceil$ (which satisfies $\epsilon\sqrt{k} \leq c$ for $c = 0$, as well as $\rho\sqrt{k} \geq 1$), we arrive at

$$\mathbb{P}(|\mathcal{C}_{\max}| > k) \leq c_2 \frac{n}{k^{3/2}}$$

for some finite positive constant c_2 which depends solely on d. $\qquad\square$

3.4. Proof of Proposition 3

Before starting the actual proof, we need to recall the definition of *size-biased* distribution of a non-negative random variable and to introduce a few facts.

Definition 1. For a non-negative random variable X with $\mathbb{E}(X) > 0$, the size-biased distribution of X, denoted by X^*, is the random variable defined by

$$\mathbb{P}(X^* \leq x) = \frac{\mathbb{E}(X\mathbb{1}_{\{X \leq x\}})}{\mathbb{E}(X)}.$$

For proofs of the assertions that appear in the statement of the next result, see [8].

Lemma 2. *Suppose that* $1 - F(x) \leq c_F x^{-(\tau-1)}$ *for all* $x \geq 0$, *for some* $c_F > 0$ *and* $\tau > 4$. *Let* w_i *be as in* (5). *Then* $\max\{w_i : i \in [n]\} \leq (c_F n)^{1/(\tau-1)}$. *Moreover, defining*

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{w_i \leq x\}} \tag{18}$$

and letting W_n *being a random variable with distribution function* F_n *and size-biased distribution* W_n^*, *then* $\mathbb{E}((W_n^*)^2) \leq C_1$ *and* $|1 - \mathbb{E}(W_n^*)| \leq C_2 n^{-(\tau-3)/(\tau-1)}$ *for all large enough* n, *where* C_1 *and* C_2 *are two positive constants which depend on* c_F *and* τ.

As explained in Van der Hofstad [12] (see also [8]), the cluster exploration of V_n in the $NR_n(\mathbf{w})$ random graph can be dominated by the total progeny of a (marked mixed-Poisson) branching process. Specifically, following Van der Hofstad [12], we can write

$$\mathbb{P}(|\mathcal{C}(V_n)| > k) \leq \mathbb{P}\left(2 + \sum_{i=1}^{t}(X_i - 1) > 0 \text{ for all } t \in [k]\right),$$

where the X_i are independent mixed Poisson random variables with $X_i \sim \text{Poi}(w_{M_i})$ and the M_i are i.i.d. random variables, all distributed as a random variable M with distribution given by

$$\mathbb{P}(M = m) = \frac{w_m}{l_n}, \quad m \in [n].$$

As remarked in [12], a $\text{Poi}(w_M)$ random variable converges in distribution to a mixed Poisson random variable with random parameter W^*, where W^* is the size-biased distribution of W, the latter being a positive random variable with distribution function F. Therefore

$\mathbb{P}(X_1 = 3)$ converges to $\mathbb{P}(Z = 3)$, where $Z \sim \text{Poi}(W^*)$. It follows that $\mathbb{P}(X_1 = 3) \geq \mathbb{P}(Z = 3)/2$ for all large enough n, and hence we obtain

$$\mathbb{P}(X_1 = 3) \geq \frac{1}{2}\mathbb{E}(\mathbb{P}(Z = 3 | W^*)) = \mathbb{E}(e^{-W^*}(W^*)^3)/12 > 0.$$

Taking $r \in (0, 1)$ so that $e^r - 1 \leq r + r^2$, we obtain

$$\begin{aligned}
\mathbb{E}[e^{rX_1}] &= \sum_{h=0}^{\infty} e^{rh} \sum_{i \in [n]} \frac{w_i}{l_n} e^{-w_i} \frac{w_i^h}{h!} \\
&= \sum_{i \in [n]} \frac{w_i}{l_n} e^{w_i(e^r - 1)} \\
&\leq \sum_{i \in [n]} \frac{w_i}{l_n} e^{w_i(r + r^2)} \\
&= \exp\left\{ \log\left(\sum_{i \in [n]} \frac{w_i}{l_n} e^{w_i(e^r - 1)} \right) \right\}.
\end{aligned}$$

If moreover $r < (2 \max\{w_i : i \in [n]\})^{-1}$, then $w_i(r + r^2) < 2rw_i \leq 1$ for every $i \in [n]$ and hence, for $0 < r < \min\{1, 1/2 \max\{w_i : i \in [n]\}\}$, we can bound

$$\begin{aligned}
\log\left(\sum_{i \in [n]} \frac{w_i}{l_n} e^{w_i(e^r - 1)} \right) &\leq \log\left(\sum_{i \in [n]} \frac{w_i}{l_n}(1 + w_i(r + r^2) + w_i^2 4r^2) \right) \\
&= \log\left(1 + (r + r^2) \sum_{i \in [n]} \frac{w_i^2}{l_n} + 4r^2 \sum_{i \in [n]} \frac{w_i^3}{l_n} \right).
\end{aligned}$$

Note that if W_n is a random variable with distribution function F_n given in (18) and W_n^* is its size-biased distribution, then

$$\sum_{i \in [n]} \frac{w_i^2}{l_n} = \mathbb{E}(W_n^*) \quad \text{and} \quad \sum_{i \in [n]} \frac{w_i^3}{l_n} = \mathbb{E}((W_n^*)^2).$$

Therefore we arrive at

$$\log\left(\sum_{i \in [n]} \frac{w_i}{l_n} e^{w_i(e^r - 1)} \right) \leq \log\left(1 + (r + r^2)\mathbb{E}(W_n^*) + 4r^2 \mathbb{E}((W_n^*)^2) \right).$$

Thanks to Lemma 2 we know that $\mathbb{E}((W_n^*)^2) \leq C_1$ and $\mathbb{E}(W_n^*) \leq 1 + C_2 n^{-(\tau-3)/(\tau-1)}$ for all sufficiently large n (for some finite constants $C_1, C_2 > 0$ which depend on c_F and τ), we see that

$$\log\left(1 + (r + r^2)\mathbb{E}(W_n^*) + 4r^2 \mathbb{E}((W_n^*)^2) \right) \leq r\left(1 + C_2 n^{-(\tau-3)/(\tau-1)} \right) + r^2(1 + 5C_1)$$

for all large enough n. Summarizing, for all positive $r < \min\{1, 1/2 \max\{w_i : i \in [n]\}\}$ we have shown that

$$\mathbb{E}(e^{rX_1}) \leq \exp\left\{ r\left(1 + C_2 n^{-(\tau-3)/(\tau-1)} \right) + r^2(1 + 5C_1) \right\},$$

provided n is sufficiently large. Hence condition (iii) in Theorem 1 is satisfied for $\rho = \rho(n) = \min\{1, 1/2 \max\{w_i \colon i \in [n]\}\}$, $\delta = 1 + 5C_1$, and $\epsilon = \epsilon(n) = C_2 n^{-(\tau-3)/(\tau-1)}$. Note that $k = \lceil An^{2/3} \rceil$ satisfies

$$\epsilon \sqrt{k} = \mathrm{O}\big(A^{1/2} C_2 n^{-(\tau-4)/(3(\tau-1))}\big) \le 1$$

for all large enough n since A is fixed and $\tau > 4$, and $\rho \sqrt{k} \ge 1$ since

$$(2 \max\{w_i \colon i \in [n]\})^{-1} \sqrt{k} \ge A^{1/2} n^{(\tau-4)/(3(\tau-1))} \big(2 c_F^{1/(\tau-1)}\big)^{-1} \ge 1.$$

Hence we can apply Theorem 1 to conclude that

$$\mathbb{P}(|\mathcal{C}_{\max}| > \lceil An^{2/3} \rceil) \le \frac{c_3}{A^{3/2}},$$

for some finite positive constant c_3 that depends solely on c_F and τ. $\qquad\square$

Acknowledgement

The author would like to thank two anonymous referees and M. Roberts for useful suggestions that helped to improve the presentation of the paper. Moreover, the author thanks G. Perarnau and A. Pachon for interesting discussions about random intersection graphs on the occasion of the workshop *Graphs and Randomness in Turin* (January 2019).

Funding information

The author thanks the Royal Society for his PhD scholarship.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ADDARIO-BERRY, L. AND REED, B. A. (2008). Ballot theorems, old and new. In *Horizons of Combinatorics* (Bolyai Society Mathematical Studies **17**), pp. 9–35. Springer, Berlin and Heidelberg.

[2] BEHRISCH, M. (2007). Component evolution in random intersection graphs. *Electron. J. Combin.* **14**, R17.

[3] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd edn. Cambridge University Press, Cambridge.

[4] BOLLOBÁS, B., JANSON, S. AND RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31**, 3–122.

[5] CHUNG, F. AND LU, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Combinatorics* **6**, 125–145.

[6] CHUNG, F. AND LU, L. (2003). The average distance in a random graph with given expected degrees. *Internet Math.* **1**, 91–113.

[7] CHUNG, F. AND LU, L. (2006). The volume of the giant component of a random graph with given expected degrees. *SIAM J. Discrete Math.* **20**, 395–411.

[8] DE AMBROGGIO, U. AND PACHON, A. (2020). Simple upper bounds for the largest components in critical inhomogeneous random graphs. Available at arXiv:2012.09001.

[9] DE AMBROGGIO, U. AND ROBERTS, M. (2022). Unusually large components in near-critical Erdős–Rényi graphs via ballot theorems. *Combinatorics Prob. Comput.*, to appear. Available at doi: 10.1017/S0963548321000584.

[10] DEIJFEN, M. AND KETS, W. (2009). Random intersection graphs with tunable degree distribution and clustering. *Prob. Eng. Inf. Sci.* **23**, 661–674.

[11] FRIEZE, A. AND KAROŃSKI, M. (2015). *Introduction to Random Graphs*. Cambridge University Press, Cambridge.

[12] VAN DER HOFSTAD, R. (2013). Critical behaviour in inhomogeneous random graphs. *Random Structures Algorithms* **42**, 480–508.

[13] VAN DER HOFSTAD, R. (2016). *Random Graphs and Complex Networks*, Vol. 1 (Cambridge Series in Statistical and Probabilistic Mathematics **43**). Cambridge University Press, Cambridge.

[14] JANSON, S. (2009). Asymptotic equivalence and contiguity of some random graphs. *Random Structures Algorithms* **36**, 26–45.

[15] JANSON, S., ŁUCZAK, T. and RUCIŃSKI, A. (2011). *Random Graphs*. John Wiley, New York.

[16] JOOS, F. AND PERARNAU, G. (2018). Critical percolation on random regular graphs. *Proc. Amer. Math. Soc.* **146**, 3321–3332.

[17] KAGER, W. (2011). The hitting time theorem revisited. *Amer. Math. Monthly* **118**, 735–737.

[18] KANG, M., PACHON, A. AND RODRIGUEZ, P. M. (2018). Evolution of a modified binomial random graph by agglomeration. *J. Statist. Phys.* **170**, 509–535.

[19] KRIVELEVICH, M. AND SUDAKOV, B. (2012). The phase transition in random graphs: a simple proof. *Random Structures Algorithms* **43**, 131–138.

[20] LAGERAS, A. N. AND LINDHOLM, M. (2008). A note on the component structure in random intersection graphs with tunable clustering. *Electron. J. Combin.* **15**, N10.

[21] ŁUCZAK, T., PITTEL, B. AND WIERMAN, J. C. (1994). The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.* **341**, 721–748.

[22] NACHMIAS, A. AND PERES, Y. (2010). Critical percolation on random regular graphs. *Random Structures Algorithms* **36**, 111–148.

[23] NACHMIAS, A. AND PERES, Y. (2010). The critical random graph, with martingales. *Israel J. Math.* **176**, 29–41.

[24] NORROS, I. AND REITTU, H. (2006). On a conditionally Poissonian graph process. *Adv. Appl. Prob.* **38**, 59–75.

[25] PENROSE, M. D. (2018). Inhomogeneous random graphs, isolated vertices, and Poisson approximation. *J. Appl. Prob.* **55**, 112–136.

[26] PITTEL, B. (2001). On the largest component of the random graph at a nearcritical stage. *J. Combinatorial Theory B* **82**, 237–269.

[27] ROBERTS, M. I. (2017). The probability of unusually large components in the near-critical Erdős–Rényi graph. *Adv. Appl. Prob.* **50**, 245–271.

[28] STARK, D. (2004). The vertex degree distribution of random intersection graphs. *Random Structures Algorithms* **24**, 249–258.