Acta Math. Univ. Comenianae
Vol. LXXXVIII, 3 (2019), pp. 567–572

ON THE LARGEST COMPONENT OF THE CRITICAL RANDOM DIGRAPH

M. COULSON

Abstract. We consider the largest component of the random digraph $D(n,p)$ inside the critical window $p = n^{-1} + \lambda n^{-4/3}$. We show that the largest component $\mathcal{C}_1$ has size of order $n^{1/3}$ in this range. In particular we give explicit bounds on the probabilities that $|\mathcal{C}_1|n^{-1/3}$ is very large or very small that are analogous to those given by Nachmias and Peres for $G(n,p)$.

1. Introduction

Consider the random digraph model $D(n,p)$ where each of the $n(n-1)$ possibe edges is included with probability p independently of all others. This is analogous to the Erdős-Renyi random graph $G(n,p)$ in which each edge is again present with probabilty p independently of all others. Perhaps unsurprisingly the two models share a number of similar features. McDiarmid [**10**] showed that it is often possible to couple $G(n,p)$ and $D(n,p)$ to compare the probabilities of certain properties. In this paper we will study the size of the largest strongly connected component of $D(n,p)$ within the critical window. This is a setting in which one cannot apply the result of McDiarmid to deduce anything from the behavior of $G(n,p)$.

In the random graph $G(n,p)$ the component structure is well understood. In their seminal paper [**3**], Erdős and Rényi proved that for $p = c/n$ the largest component of $G(n,p)$ has size $O(\log(n))$ if $c < 1$, is of order $\Theta(n^{2/3})$ if $c = 1$ and has linear size when $c > 1$. This threshold behavior is known as the double jump. Zooming in further around the critical point, $p = 1/n$ and considering $p = (1+\varepsilon)/n$ such that $\varepsilon \to 0$ and $|\varepsilon|^3 n \to \infty$, Bollobás [**1**] proved the following theorem for $|\varepsilon| > (2\log(n))^{1/2} n^{-1/3}$ which was later extended to the whole range described above by Łuczak [**6**].

Theorem 1.1. *Let* $np = 1+\varepsilon$*, such that* $\varepsilon = \varepsilon(n) \to 0$ *but* $n|\varepsilon|^3 \to \infty$*, and* $k_0 = 2\varepsilon^{-2}\log(n|\varepsilon|^3)$*.*

i) *If* $n\varepsilon^3 \to -\infty$ *then a.a.s.* $G(n,p)$ *contains no component of size greater than* k_0*.*
ii) *If* $n\varepsilon^3 \to \infty$ *then a.a.s.* $G(n,p)$ *contains a unique component of size greater than* k_0*. This component has size* $2\varepsilon n(1+o(1))$*.*

Received May 21, 2019.
2010 *Mathematics Subject Classification.* Primary 05C80; Secondary 05C20, 60C05.

Within the critical window itself i.e. $p = n^{-1} + \lambda n^{-4/3}$ with λ fixed, the size of the largest component is not tightly concentrated as it is for larger p. Instead, there exists a random variable $X = X(\lambda)$ such that $|\mathcal{C}_1|n^{-2/3} \to X$ as $n \to \infty$ where $\mathcal{C}_1$ is the largest component of $G(n,p)$. Much is known about the distribution of X, for example Nachmias and Peres [**11**] proved the following (similar results can be found in [**13, 14**]).

Theorem 1.2. *Suppose* $0 < \delta < 1/10$, $A > 8$ *and* n *suitably large with respect to* A, δ. *Then if* $\mathcal{C}_1$ *is the largest component of* $G(n, 1/n)$, *we have*

i) $\mathbb{P}(|\mathcal{C}_1| < \lfloor \delta n^{2/3} \rfloor) \leq 15\delta^{3/5}$
ii) $\mathbb{P}(|\mathcal{C}_1| > An^{2/3}) \leq \frac{4}{A}\mathrm{e}^{-\frac{A^2(A-4)}{32}}$

Note we have only stated the theorem with $p = n^{-1}$ for clarity.

One finds that analogues of many of the above theorems for strongly connected components are true in $D(n,p)$. Also, the weak component structure of $D(n,p)$ is precisely the component structure of $G(n, 2p - p^2)$ and so is not very interesting. For $p = c/n$, Karp [**5**] and Łuckzak [**7**] independently showed that for $c < 1$ all strongly connected components are of size $O(1)$ and when $c > 1$ there is a unique complex component of linear order and every other component is of size $O(1)$ (a component is complex if it has more edges than vertices). The range $p = (1+\varepsilon)/n$ was studied by Łuczak and Seierstad [**8**] who proved the following result which can be viewed as a version of Theorem 1.1 for $D(n,p)$,

Theorem 1.3. *Let* $np = 1 + \varepsilon$, *such that* $\varepsilon = \varepsilon(n) \to 0$.

i) *If* $n\varepsilon^3 \to -\infty$ *then a.a.s. every component of* $D(n,p)$ *is an isolated vertex or a cycle of length* $O(1/|\varepsilon|)$.
ii) *If* $n\varepsilon^3 \to \infty$ *then a.a.s.* $D(n,p)$ *contains a unique complex component of size* $4\varepsilon^2 n(1 + o(1))$ *and every other component is an isolated vertex or cycle of length* $O(1/\varepsilon)$.

Our main result gives us information about the size of the largest component in the critical window, $p = n^{-1} + \lambda n^{-4/3}$ for constant $\lambda \in \mathbb{R}$. In particular we show that the size of the largest component of $D(n,p)$ is $\Theta(n^{1/3})$ and furthermore, we give bounds on the probabilities that it is very large or very small which resemble those of Nachmias and Peres for $G(n,p)$.

Theorem 1.4 (Lower Bound)**.** *Let* $\mathcal{C}_1$ *be the largest strongly connected component of* $D(n,p)$ *and let* $0 < \delta < 1/800$, $\lambda \in \mathbb{R}$ *and* $n \in \mathbb{N}$. *Then if* n *is large enough with respect to* δ, λ,

$$\mathbb{P}(|\mathcal{C}_1| < \delta n^{1/3}) \leq \begin{cases} 2^{\mathrm{e}^{2\lambda\delta^{1/2}}} \delta^{\frac{1}{2}\mathrm{e}^{\frac{\lambda\delta}{2}}} & \text{if } \lambda \geq 0 \\ 2^{\mathrm{e}^{\lambda\delta}} \delta^{\frac{1}{2}\mathrm{e}^{2\lambda\delta^{1/2}}} & \text{otherwise} \end{cases}$$

Note that while this theorem does not impose any restrictions upon the relationship between δ and λ, the bound obtained is only non-trivial provided that $|\lambda|$ is suitably small with respect to δ.

Theorem 1.5 (Upper Bound)**.** *There exists constants, $\zeta, \eta > 0$ such that for any $A > 0, \lambda \in \mathbb{R}$, if n is sufficiently large with respect to A, λ,*

$$\mathbb{P}(|\mathcal{C}_1| > An^{1/3}) \leq \zeta e^{-\eta A^{3/2} + \lambda^+ A}$$

Where $\lambda^+ = \max(\lambda, 0)$.

2. Sketch of proofs

We will present sketches of the proofs for the case $\lambda = 0$. When $\lambda \neq 0$ there are small adjustments to be made which are omitted for simplicity. See [**2**] for full details of the proofs of Theorems 1.4 and 1.5.

2.1. Enumeration of digraphs

A key ingredient in the proofs of Theorems 1.4 and 1.5 is a bound on the number of strongly connected digraphs with a given number of vertices and excess, here the *excess* of a strongly connected digraph D is k if D has k more edges than vertices.

Let $Y(m, k)$ be the number of labelled strongly connected digraphs with m vertices and excess k. A double counting of the number of ear decompositions of such digraphs allows us to conclude the following bound which is valid for any m and k.

Lemma 2.1.

$$Y(m, k) \leq \frac{(m + k)^k m^{2k} (m - 1)!}{k!} \tag{1}$$

When the excess is small we need a better bound than this. Pérez-Giminéz and Wormald [**12**] asmyptotically determined the value of $Y(m, k)$ for $k = \omega(1)$. An adaptation of their proof gives the following bound which is valid for any small enough k.

Lemma 2.2. *There exists a constant $C > 0$ such that for $1 \leq k \leq \sqrt{m}/3$ and $m \gg 1$ we have,*

$$Y(m, k) \leq C \frac{m! m^{3k-1}}{(2k - 1)!} \tag{2}$$

The proof of this is by showing an upper bound similar to [**12**, Theorem 1.1] holds for any $\lambda > 0$ and then setting $\lambda = 2k/m$ yields the result.

2.2. Lower bounds

The proof of Theorem 1.4 follows from an application of Janson's inequality in the following form [**4**, Theorem 2.18 (i)],

Theorem 2.3. *Let S be a set and $S_p \subseteq S$ chosen by including each element of S in S_p independently with probability p. Suppose that $\mathcal{S}$ is a family of subsets of S and for $A \in \mathcal{S}$, we define I_A to be the event $\{A \subseteq S_p\}$. Let X be the number of* *events I_A for $A \in \mathcal{S}$ which occur. Define $\mu = \mathbb{E}(X)$ and*

$$\Delta = \frac{1}{2} \sum_{A \neq B, A \cap B \neq \emptyset} \mathbb{E}(I_A I_B)$$

Then,

$$\mathbb{P}(X = 0) \leq \mathrm{e}^{-\mu + \Delta}$$

Let X be the random variable counting the number of directed cycles of length between $\delta n^{1/3}$ and $\delta^{1/2} n^{1/3}$ in $D(n, 1/n)$. X splits into the sum of a number of indicator random variables for the presence of long cycles in $D(n, 1/n)$. Furthermore, we can compute $\mathbb{E}(X) \geq -\log(\delta)/2$.

Directly evaluating Δ is difficult, however the terms $\mathbb{E}(I_A I_B)$ each represent the occurrence of a strongly connected digraph which is the union of two cycles. Thus we use the following auxiliary lemma to simplify the computations.

Lemma 2.4. *Each strongly connected digraph, D with excess k may be formed in at most 27^k ways as the union of a pair of directed cycles, C_1 and C_2.*

Applying this lemma yields the bound

$$\Delta \leq \frac{1}{2} \sum_{m=\delta n^{1/3}}^{2\delta^{1/2} n^{1/3}} \sum_{k=1}^{\infty} 27^k \mathbb{E}(Z(m, k)) \tag{3}$$

Where $Z(m, k)$ is the random variable counting the number of strongly connected digraphs with m vertices and excess k. One can easily compute these expectations, finding

$$\mathbb{E}(Z(m, k)) = \binom{n}{m} \frac{Y(m, k)}{n^{m+k}} \tag{4}$$

Replacing $Y(m, k)$ with the bound from equation (1) allows us to compute the expression in (3) and deduce it is at most $\log(2)$ for $\delta \in (0, 1/800)$ thus proving Theorem 1.4 when $\lambda = 0$.

2.3. Upper bounds

We prove the upper bound by computing the expected number of strongly connected components of size between $An^{1/3}$ and $n^{1/3} \log\log(n)$ which have excess at most $n^{1/6}$. It is easy to see that the probability of there being any larger components or components of excess at least $n^{1/6}$ is $o_n(1)$ by the result of Łuczak and Sierstad [**8**].

In order to count components, we first look at strongly connected subgraphs of $D \sim D(n, 1/n)$. For each subgraph H we run an out-exploration process starting from $V(H)$ until the process dies. We conclude that H is a strongly connected subgraph of D if there are no edges from the explored out-component of H which return to H.

The exploration process we consider was initially developed by Martin-Löf [**9**] and Karp [**5**]. During this process, vertices will be in one of three classes: *active*, *explored* or *unexplored*. At time $t \in \mathbb{N}$, we let X_t be the number of active vertices.

There will also be t explored vertices at time t. We will start from a set $\mathcal{X}_0$ of size x_0 and fix an ordering of the vertices, starting with $V(\mathcal{X}_0)$. At time $t = 0$ all vertices of $\mathcal{X}_0$ are active with all other vertices unexplored, so $X_0 = x_0$. For step $t \geq 1$, if $X_{t-1} > 0$ let w_t be the first active vertex. Otherwise, let w_t be the first unexplored vertex. Define η_t to be the number of unexplored out-neighbours of w_t in $D(n, 1/n)$. Change the class of each of these vertices to active and set w_t to explored. We set $N_t = n - X_t - t - \mathbb{1}(X_t = 0)$. Note that given the history of the process, η_t is distributed as a binomial random variable with parameters N_{t-1} and p. Furthermore, the following recurrence relation holds.

$$X_t = \begin{cases} X_{t-1} + \eta_t - 1 & \text{if } Y_{t-1} > 0 \\ \eta_t & \text{otherwise} \end{cases} \tag{5}$$

Let $\tau_1 = \min\{t \geq 1 : X_t = 0\}$. Note that this is a stopping time and at time τ_1 the set of explored vertices is precisely the out-component of $\mathcal{X}_0$. Let $\mathcal{X}_t$ be the set of active and explored vertices at time t. If $\mathcal{X}_0$ spans a strongly connected subdigraph D_0 of $D(n, 1/n)$, then D_0 is a strongly connected component if and only if there are no edges from $\mathcal{X}_{\tau_1} \setminus \mathcal{X}_0$ to X_0. So, firstly we show that the exploration process is very likely to last a significant length of time,

Lemma 2.5. *Let X_t be the exploration process defined above with starting set of vertices $\mathcal{X}_0$ of size m. Suppose $0 < c < \sqrt{2}$ is a fixed constant. Then,*

$$\mathbb{P}(\tau_1 < cm^{1/2}n^{1/2}) \leq 2e^{-\frac{(2-c^2)^2}{8c}m^{3/2}n^{-1/2} + O(m^2 n^{-1})}$$

The proof of this follows by lower bounding the process and using Doob's maximal inequality to deduce a Chernoff type bound. If we start the exploration process with the vertices of a strongly connected subdigraph H of $D(n, 1/n)$, then H is strongly connected component only if there are no edges returning to it from its explored out-component. Computing the probability no such edge is present yields the following lemma,

Lemma 2.6. *There exist $\beta, \gamma > 0$ such that if H is any strongly connected subgraph of $D(n, 1/n)$ with m vertices. Then the probability that H is a strongly connected component of $D(n, 1/n)$ is at most $\beta e^{-(1+\gamma)m^{3/2}n^{-1/2} + O(m^2 n^{-1})}$.*

Using Lemmas 2.5 and 2.6 we bound $\mathbb{E}(N(A))$ where $N(A)$ is the random variable which counts the number of strongly connected components of $D(n, 1/n)$ of size between $An^{1/3}$ and $n^{1/3}\log\log(n)$ and excess at most $n^{1/6}$. Let $X_i = |\mathcal{X}_i|$ and $Y_i = |E(\mathcal{X}_i \setminus \mathcal{X}_0, \mathcal{X}_0)|$. Then,

$$\mathbb{E}(N(A)) = \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \sum_{k=0}^{n^{1/6}} \binom{n}{m} p^{m+k} Y(m,k) \mathbb{P}(Y_{\tau_1} = 0 | X_0 = m) \tag{6}$$

Using Lemma 2.6 we bound $\mathbb{P}(Y_{\tau_1} = 0 | X_0 = m) \leq \beta e^{-(1+\gamma)m^{3/2}n^{-1/2} + O(m^2 n^{-1})}$. Furthermore, by Lemma 2.2,

$$\sum_{k=0}^{n^{1/6}} \frac{n^{-k} Y(m,k)}{(m-1)!} \leq 1 + C \sum_{k=1}^{n^{1/6}} \frac{n^{-k} m^{3k}}{(2k-1)!} \leq 2Cm^{3/2}n^{-1/2} \sinh(m^{3/2}n^{-1/2}). \tag{7}$$

Then, as $\sinh(x) \leq \mathrm{e}^x$ we get the following bound on $\mathbb{E}(N(A))$,

$$\mathbb{E}(N(A)) \leq \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \frac{1}{m}\Big(2Cm^{3/2}n^{-1/2}\mathrm{e}^{m^{3/2}n^{-1/2}}\Big) \times \Big(\beta \mathrm{e}^{-(1+\gamma)m^{3/2}n^{-1/2}+O(m^2n^{-1})}\Big), \tag{8}$$

which we can simplify and approximate with the following integral.

$$\mathbb{E}(N(A)) \leq \int_{m=An^{1/3}}^{n^{1/3}\log^2(n)+1} \frac{2\beta Cm^{1/2}}{n^{1/2}}\mathrm{e}^{-\frac{\gamma}{2}m^{3/2}n^{-1/2}}\mathrm{d}m. \tag{9}$$

Making the substitution $t = \gamma m^{3/2}n^{-1/2}/2$ in (9) reduces the integral to the integral of an exponential function which is easy to evaluate. The proof of Theorem 1.5 then follows by an application of Markov's inequality and a union bound.

References

1. Bollobás B., *The evolution of random graphs*, Trans. Amer. Math. Soc. **286** (1984), 257–284.
2. Coulson M., *The critical window in random digraphs*, `arXiv:1905.00624`.
3. Erdős P. and Rényi A., *On the evolution of random graphs*, Publ. Math. Inst. Hung. Acad. Sci. **5** (1960), 17–60.
4. Janson S., Łuczak T. and Ruciński A., *Random Graphs*, Wiley-Intersci. Ser. Discrete Math. Optim., Wiley-Interscience, New York, 2000.
5. Karp R. M., *The transitive closure of a random digraph*, Random Structures Algorithms **1** (1990), 73–93.
6. Łuczak T., *Component behavior near the critical point of the random graph process*, Random Structures Algorithms **1** (1990), 287–310.
7. Łuczak T., *The phase transition in the evolution of random digraphs*, J. Graph Theory **14** (1990), 217–223.
8. Łuczak T. and Seierstad T., *The critical behavior of random digraphs*, Random Structures Algorithms **35** (2009), 271–293.
9. Martin-Löf A., *Symmetric sampling procedures, general epidemic processes and their theshold limit theorems*, J. Appl. Probab. **23** (1986), 265–282.
10. McDiarmid C., *Clutter percolation and random graphs*, Math. Programming Stud. **13** (1980), 17–25.
11. Nachmias A. and Peres Y., *The critical random graph, with martingales*, Israel J. Math. **176** (2010), 29–41.
12. Pérez-Giménez X. and Wormald N., *Asymptotic enumeration of strongly connected digraphs by vertices and edges*, Random Structures Algorithms **43** (2013), 80–114.
13. Pittel B., *On the largest component of the random graph at a nearcritical stage*, J. Combin. Theory Ser. B **82** (2001), 237–269.
14. Scott A. D. and Sorkin G. B., *Solving sparse random instances of Max Cut and Max 2-CSP in linear expected time*, Combin. Probab. Comput. **15** (2006), 281–315.

M. Coulson, School of Mathematics, University of Birmingham, UK,
e-mail: `mjc685@bham.ac.uk`