THE ROYAL SWEDISH ACADEMY OF SCIENCES
INSTITUT MITTAG-LEFFLER

Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

The diameter behavior in the random graph process

T. Luczak and T. G. Seierstad

REPORT No. 5, 2008/2009, spring

ISSN 1103-467X
ISRN IML-R- -5-08/09- -SE+spring

THE DIAMETER BEHAVIOR IN THE RANDOM GRAPH PROCESS

TOMASZ ŁUCZAK AND TARAL GULDAHL SEIERSTAD

Abstract. We give uniform bounds for the diameter in the random graph process $\{G(n,M)\}_M$ near the critical phase. In particular, we show that a.a.s. the maximum diameter of $G(n,M)$ is achieved in the critical period. It is also proved that while in the late subcritical phase a.a.s. the component of largest diameter is not the largest one, in the supercritical phase a.a.s. the diameter of the graph is determined by the diameter of the giant component.

1. Introduction

The diameter of a connected graph is the largest possible distance between its vertices. For a disconnected graph, we define its diameter as the maximum diameter of its components. The purpose of this paper is to study the changes of the diameter of $G(n,M)$ during the random graph process $\{G(n,M)\}_{M=0}^{\binom{n}{2}}$, especially near the point of the phase transition, when $M = n/2 + o(n)$.

Before we state our main theorem, let us recall first a few known results about the diameter of $G(n,M)$ and its counterpart $G(n,p)$. The threshold for the property that the diameter of the random graph is smaller than a given number k was found by Burtin [4, 5]. His result was extended to sparse random graphs by Bollobás [1] (see also [2]). The diameter of the largest component of the random graph with finite expected average degree was studied by Bollobás [2] and Chung and Li [6]. Recently, Fernholz and Ramachandran [7] showed a fairly general result, from which it follows that the diameter of the random graph in which the expected average degree is $c > 1$ is $(\alpha(c)+o(1)) \ln n / \ln c$, where $\alpha(c)$ is an explicitly given constant such that $\alpha(c) \to 3$ as $c \to 1$. They also proved that for such $c > 1$, i.e., in the late supercritical phase, the diameter of the random graph is determined by the diameter of the largest component. As for the diameter of the random graph in the critical range, when the expected average degree is $1 + o(1)$, it follows from a much more general theorem of Nachmias and Peres [15] that for every given function $p = p(n)$ (and $M = M(n)$) the diameter of $G(n,p)$ (respectively $G(n,M)$) is a.a.s. $O_p(n^{1/3})$.

Date: February 20, 2009.

1991 *Mathematics Subject Classification.* Primary: 05C80. Secondary: 05C12.

Key words and phrases. random graphs, random graph process, phase transition, critical behavior, diameter.

TŁ is partially supported by the Foundation for Polish Science, while TGS is supported by the Research Council of Norway; contract/grant number 170620/V30. This work was done when the authors visited the Institut Mittag-Leffler (Djursholm, Sweden).

The goal of this paper is to prove 'global' results on the random graph process $\{G(n,M)\}_{M=0}^{\binom{n}{2}}$ rather that 'local' results on the random graph $G(n,M)$. We show that the diameter of $G(n,M)$ is maximized for $M = n/2 + O(n^{2/3})$ and that in the whole supercritical phase, the diameter of the largest component of $G(n,M)$ is larger than the diameter of each of its other components. In order to state our result, let us define $\mathrm{diam_L}(G)$ as the diameter of the (lexicographically first) largest component of G, let $\mathrm{diam_s}(G)$ be the maximum diameter among all other components of G, and $\mathrm{diam}(G) = \max\{\mathrm{diam_L}(G), \mathrm{diam_s}(G)\}$. Then our main result can be stated as follows.

Theorem 1. *Let $\omega = \omega(n) \to \infty$. Then, a.a.s. the random graph process $\{G(n,M)\}_M$ is such that for every $M = n/2 + s$, where $\omega n^{2/3} \le s \le n/\omega$, we have*

$$0.4\frac{n}{s}\ln\frac{s^3}{n^2} \le \mathrm{diam_s}(G(n,M)) \le 0.6\frac{n}{s}\ln\frac{s^3}{n^2}\,, \tag{1}$$

while

$$0.7\frac{n}{s}\ln\frac{s^3}{n^2} \le \mathrm{diam_L}(G(n,M)) = \mathrm{diam}(G(n,M)) \le 100\frac{n}{s}\ln\frac{s^3}{n^2}\,. \tag{2}$$

Let us remark that in the subcritical phase, when $M = n/2 - s$, $s = s(n) = o(n)$ but $sn^{-2/3} \to \infty$, a.a.s. the lexicographically first largest component is a tree with $(0.5 + o(1))n^2 s^{-2} \ln s^3 n^{-2}$ vertices which has diameter $O_p\big(ns^{-1}\ln^{1/2} s^3 n^{-2}\big)$, while the diameter of the $G(n,M)$ is a.a.s. $(0.5 + o(1))ns^{-1}\ln s^3 n^{-2}$ (see Łuczak [13] and Theorem 7 below). It is also not very surprising that after the phase transition a.a.s. $\mathrm{diam_L}(G(n,M)) = \mathrm{diam}(G(n,M))$. Outside the largest component there is a substantial number of tree components with the diameter $(1 - o(1))\,\mathrm{diam_s}(G(n,M))$. Moreover, it is known (see Łuczak [12]) that in the random trees of large diameter the vertices are uniformly distributed along the longest path; i.e., if P is a longest path of length m in the random tree on $n \gg m^2$ vertices, then, say, roughly $(0.02+o(1))n$ vertices of T belong to subtrees which are rooted at P at distance at most $0.01m$ from one of its two ends. Hence, the probability that the giant component will 'catch' such a tree 'near its ends' is quite substantial. Consequently, the diameter of the largest component should be at least twice as large as the largest diameter of the small component. Note also that Fernholz and Ramachandran [7] results do not extend to the whole supercritical phase – near the critical period the diameter of $G(n,M)$ is much smaller than $\ln n/\ln(2M/n)$.

The paper is organized as follows. First we recall a few elementary facts on the diameter of trees in $G(n,M)$ and random rooted forests. Then, we show the 'local' result on the diameter of $G(n,M)$ in the supercritcal phase (Lemma 4) which later, in Section 4, is used to prove Theorem 1. We conclude the article with some comments on the strengthening of our result and the behavior of the diameter in the critical period, when $M = n/2 + O(n^{2/3})$. Let us also remark that our argument does not rely on branching process analysis, but is based on direct counting.

2. The diameter of random forests

In this section we study the maximum diameter of trees and rooted trees which appear in $G(n, M)$. We start with the following lemma, which is a refinement of Theorem 11(iii) from [13].

Lemma 2. *Let $M = n/2 - s$, where $s = s(n) = o(n)$ and $sn^{-2/3} \to \infty$. Then, with probability at least $1 - 21s^{-0.15}n^{0.1}$, the following holds.*

(i) $G(n, M)$ consists of trees and unicyclic components;
(ii) the largest component of $G(n, M)$ is smaller than $k_{\max} = \frac{n^2}{s^2} \ln \frac{s^3}{n^2}$;
(iii) $\operatorname{diam}(G(n, M)) \leq 0.55\frac{n}{s} \ln \frac{s^3}{n^2}$;
(iv) $G(n, M)$ contains at least $s^{0.15}n^{-0.1}$ trees of diameter at least $0.45\frac{n}{s} \ln \frac{s^3}{n^2}$;
(v) no unicyclic component of $G(n, M)$ contains a path longer than $0.46\frac{n}{s} \ln \frac{s^3}{n^2}$.

Proof. It is easy to check (e.g., see [10]) that the expected number of components which are either larger than $k_{\max}$ or contain more than one cycle is smaller than $s^{-0.3}n^{0.2}$, so the probability that such a component appears in $G(n, M)$ is bounded from above by $s^{-0.15}n^{0.1}$. Let $X(D; n, M)$ denote the number of trees in $G(n, M)$ with diameter at least D. One can repeat the calculations from [13], where the binomial random graph model $G(n, p)$ has been used, for the uniform model $G(n, M)$ showing that

$$\mathbb{E}X(D; n, M) = (16 + o(1))\frac{s^3}{n^2} \exp\left(-\frac{2s}{n}D\right). \tag{3}$$

Now set $D_\pm = (0.5 \pm 0.05)\frac{n}{s} \ln \frac{s^3}{n^2}$. Then,

$$\Pr(X(D_+; n, M) > 0) \leq \mathbb{E}X(D_+; n, M) \leq 17\frac{n^{0.1}}{s^{0.15}} \tag{4}$$

Furthermore, one can easily show that the expected number of unicyclic components with diameter larger than D_- is smaller than $s^{-0.15}n^{0.1}$. Hence, with probability at least $1 - 18s^{-0.15}n^{0.1}$, the diameter of $G(n, M)$ is at most D_+. Moreover, standard calculations (similar to those in [9] and [10]) show that the expected number of unicyclic components which contain a cycle larger than $\frac{n}{100s} \ln \frac{s^3}{n^2}$ is smaller than $s^{-0.15}n^{0.1}$. Since the length of the longest path contained in a unicyclic graph is bounded from above by the sum of its diameter and the length of its cycle, (v) follows.

Note that $\mathbb{E}X(D_-; n, M) > 3s^{0.15}n^{-0.1}$. Since, as we have already mentioned, the expected number of components larger than $k_{\max}$ is bounded from above by $s^{-0.3}n^{0.2}$, for the second factorial moment of $\mathbb{E}X(D_-; n, M)$ we have

$$\begin{aligned}&\left|\mathbb{E}X(D_-; n, M)(X(D_-; n, M) - 1) - (\mathbb{E}X(D_-; n, M))^2\right| \\ &\leq (\mathbb{E}X(D_-; n, M))^2 \max_{k \leq k_{\max}} \left|\frac{\mathbb{E}X(D_-; n-k, M-k+1)}{\mathbb{E}X(D_-; n, M)} - 1\right| \\ &\quad + s^{-0.3}n^{0.2}.\end{aligned}$$

For $k \le k_{\max}$, $D \le 3ns^{-1}\ln s^3 n^{-2}$, we have

$$\begin{aligned}
&\Big|\frac{\mathbb{E}X(D_-;n-k,M-k+1)}{\mathbb{E}X(D_-;n,M)}-1\Big|\\
&\le \Big|\frac{(s+(k+1)/2)^3}{s^3}\frac{n^2}{(n-k)^2}\exp\left(-\frac{(2s+k+1)D}{n-k}-\frac{2sn}{D}\right)-1\Big|\\
&= O\Big(\frac{k}{s}+\frac{kD}{n}\Big) = O\Big(\frac{n^2}{s^3}\ln^2\frac{s^3}{n^2}\Big) \le o\big((\mathbb{E}X(D_-;n,M))^{-2}\big)\,.
\end{aligned}$$

Consequently, $\mathbb{V}\mathrm{ar}X(D_-;n,M) = (1+o(1))\mathbb{E}X(D_-;n,M)$ and from Chebyshev's inequality we infer that

$$\Pr(X(D_-;n,M) \le s^{0.15}n^{-0.1}) \le s^{-0.15}n^{0.1}\,. \quad \square$$

Our next result concerns the height of the random rooted forest $F(n,m)$ chosen uniformly at random from the family of all mn^{n-m-1} forests on the vertex set $\{1,2,\ldots,n\}$ which consist of m trees and are such that each of the vertices $\{1,2,\ldots,m\}$ belongs to a different tree.

Lemma 3. *Let $m = s(n) = o(n)$ and $m^2/n \to \infty$. Then, with probability at least $1-3n^{0.1}m^{-0.2}$, the following holds.*

(i) the largest component of $F(n,m)$ is smaller than $\ell_{\max} = \frac{4n^2}{m^2}\ln\frac{m^2}{n}$;
(ii) the height of every tree in $F(n,m)$ is smaller than $1.1\frac{n}{m}\ln\frac{m^2}{n}$;
(iii) $F(n,m)$ contains at least $m^{0.2}n^{-0.1}$ trees of height at least $0.9\frac{n}{m}\ln\frac{m^2}{n}$.

Proof. Since the argument is very similar to that of Lemma 2, we just outline it. Simple calculation shows that the expected number of components with more than $\ell_{\max}$ vertices is bounded from above by $n^{0.1}m^{-0.2}$ (let us recall that, as wwas proved by Pavlov [16], the maximum size of the component in $F(n,M)$ is a.a.s. $(1/2+o(1))\ell_{\max}$). Using the asymptotic formula for the number of rooted trees with ℓ vertices and height $h \gg \sqrt{\ell}$ (see Łuczak [12]), one can show that the expectation of the number $X(H;n,m)$ of trees of height larger than H in $F(n,m)$ is given by

$$\mathbb{E}X(H;n,m) = (2+o(1))\frac{m^2}{n}\exp\big(-\frac{m}{n}H\big)\,,$$

provided $m^2/n \to \infty$ but $m/n \to 0$. Thus, the expected number of trees of height larger than $H_+ = 1.1\frac{n}{m}\ln\frac{m^2}{n}$ is bounded from above by $3m^{-0.2}n^{0.1}$, and so is the probability that such a tree appears in $F(n,m)$. Now let us set $H_- = 0.9\frac{n}{m}\ln\frac{m^2}{n}$. Then,

$$\mathbb{E}X(H_-;n,m) \ge 1.5m^{0.2}n^{-0.1}\,.$$

Moreover, arguing as in the proof of Lemma 2, we infer that

$$\begin{aligned}
\mathbb{V}\mathrm{ar}X(n,m) &\le 2(\mathbb{E}X(H_-;n,m))^2\frac{m\ell_{\max}H_-}{n^2}\\
&\le 8(\mathbb{E}X(H_-;n,m))^2\frac{n}{m^2}\ln^2\frac{m^2}{n}\,,
\end{aligned}$$

so the assertion follows from Chebyshev's inequality. $\square$

3. The diameter of the core of $G(n, M)$

The aim of this section is to show a local result on the diameter of $G(n, M)$ in the supercritical phase. Before we shall state it, let us introduce some notation related to the structure of the giant component of $G(n, M)$ (for details see [9], Chapter 5.4). The excess $\mathrm{exc}(G)$ of a connected (multi)graph G is the difference between the number of its edges and the number of its vertices. A component of a graph is complex if its excess is positive, i.e., it is neither an isolated tree nor a unicyclic component. By $\mathbf{L}$ we mean the (lexicographically first) largest component of $G(n, M)$, and $\mathbf{C}$ stands for its core, i.e., the maximal subgraph of $\mathbf{L}$ with the minimum degree at least two. By $\mathbf{K}$ we denote the kernel of $\mathbf{L}$, i.e., the multigraph of the minimum degree at least three obtained from the core by replacing each of its vertices of degree two by an edge joining its neighbours. Note that $\mathbf{L}$, $\mathbf{C}$, and $\mathbf{K}$ have the same excess. By $L = L(n, M)$, $\mathrm{cr} = \mathrm{cr}(n, M)$, $\mathrm{ker} = \mathrm{ker}(n, M)$ we denote the number of vertices in $\mathbf{L}$, $\mathbf{C}$, and $\mathbf{K}$ respectively, and $\kappa = \kappa(n, M) = \mathrm{exc}(\mathbf{L})$. A whisker is a path contained in the largest component which shares with its core just one of its ends. Then, the main result of this section can be stated as follows.

Lemma 4. *Let $M = n/2 + s$, where $s = s(n) = o(n)$ and $sn^{-2/3} \to \infty$. Then, with probability at least $1 - 5n^{0.02}s^{-0.03}$, the following holds.*

- *(i) $G(n, M)$ consists of one giant component and some isolated trees and unicyclic components;*
- *(ii) the largest component of $G(n, M)$ has at least $s^{0.03}n^{0.02}$ vertex disjoint whiskers, each of length at least $0.4\frac{n^2}{s^2}\ln\frac{s^3}{n^2}$, and no whisker longer than $0.6\frac{n^2}{s^2}\ln\frac{s^3}{n^2}$;*
- *(iii) $\mathrm{diam}(G(n, M)) \le 6.5\frac{n}{s}\ln\frac{s^3}{n^2}$;*
- *(iv) $G(n, M)$ contains at least $s^{0.15}n^{-0.1}$ isolated trees, each of diameter at least $0.45\frac{n}{s}\ln\frac{s^3}{n^2}$, and no trees of diameter larger than $0.55\frac{n}{s}\ln\frac{s^3}{n^2}$;*
- *(v) no unicyclic component of $G(n, M)$ contains a path longer than $0.46\frac{n}{s}\ln\frac{s^3}{n^2}$.*

Our proof of Lemma 4 relies on the following results on the behavior of $G(n, M)$ in the supercritical phase. Let us first start with a lemma on the structure of the giant component of $G(n, M)$.

Lemma 5. *Let $M = n/2 + s$, where $s = o(n)$ but $sn^{-2/3} \to \infty$. Then with probability at least $1 - n^{0.02}s^{-0.03}$ the following holds for the random graph $G(n, M)$.*

- *(i) $|L(n, M) - 4s| \le s\frac{n^{0.02}}{s^{0.03}}$;*
- *(ii) $|\,\mathrm{cr}(n, M) - \frac{8s^2}{n}| \le \frac{8s^2}{n}\frac{n^{0.02}}{s^{0.03}}$;*
- *(iii) $|\kappa(n, M) - \frac{16}{3}\frac{s^3}{n^2}| \le \frac{s^3}{n^2}\frac{n^{0.02}}{s^{0.03}}$;*
- *(iv) $|\,\mathrm{ker}(n, M) - \frac{32}{3}\frac{s^3}{n^2}| \le \frac{s^3}{n^2}\frac{n^{0.02}}{s^{0.03}}$;*
- *(v) for every $j \ge 4$ the number of vertices with degree at least j in the kernel $\mathbf{K}$ is bounded from above by $10s^j/n^{j-1}\frac{n^{0.02}}{s^{0.03}}$.*

Proof. The estimates for $L(n, M)$ and $\kappa(n, M)$ are proved in [10], Theorems 5 and 9 (see also [9], Theorems 5.12 and 5.13). The bounds for

$\mathrm{cr}(n, M)$ as well as (v) follow from the results on the structure of a random connected graph with given number of vertices and edges that can be found in Łuczak [11], Theorem 10. Finally, (iv) is an immediate consequence of (iii) and (v). □

Let $G^L(n, M)$ denote the graph obtained from $G(n, M)$ by deleting all vertices of the largest component $\mathbf{L}$. We shall use the following symmetry rule principle (see [10] or [9] Theorem 5.24) which states, roughly, that if $\mathbf{L}$ has L vertices and $e(\mathbf{L})$ edges, then the properties of $G^L(n, M)$ are basically the same as those of $G(n - v(\mathbf{L}), M - e(\mathbf{L}))$.

Lemma 6. *Let $M = n/2 + s$, where $s = o(n)$ but $sn^{-2/3} \to \infty$. Then, for n large enough,*

$$\Pr(G^L(n, M) \text{ has } \mathcal{A}) \le \max\big\{ \Pr(G(n', M') \text{ has } \mathcal{A}) \colon$$
$$|n' - (n - 4s)| \le \omega^{-0.9} s,\ |M' - (n'/2 - s)| \le \omega^{-0.9} s\big\} + 8n^{2/9} s^{-1/3}. \quad \square$$

Proof of Lemma 4. Lemma 2(i), Lemma 5(i) and Lemma 6 give (i), while Lemma 2 and Lemma 6 imply (iv) and (v). Note that $\mathbf{L}$ can be viewed as obtained from the core by adding to it a forest with vertices rooted at the core, where each such forest is equally likely. Hence, if we delete all edges of the core from the giant component, we obtain a random rooted forest $F(L(n, M), \mathrm{cr}(n, M))$. Consequently, (ii) follows from Lemma 3(ii) and (iii) and Lemma 5(i) and (ii). Thus, to complete the proof, we need to show (iii).

To this end we estimate the diameter of the core of the giant component. Note that given a degree sequence of $\mathbf{C}$, each graph with this degree sequence is equally likely to appear as $\mathbf{C}$. Since most of vertices of $\mathbf{C}$ are of degree two (Lemma 5(ii) and (iv)), one can first consider only vertices of degree at least three, build the kernel, and then randomly place vertices of degree two at the edges of the kernel. It turns out that, given a degree sequence, the kernel is a random multigraph, where the probability of the instance of each multigraph is roughly the same as in the configuration model. From Lemma 5(iv) and (v) it follows that most of the vertices of the kernel have degree three; thus the diameter of kernel is a.a.s. $(1 + o(1)) \log_3 \ker(n, M)$ (see Bollobás, Fernandez de la Vega [3], or Fernholz, Ramachandran [7]). However we need an explicit bound for the probability that the diameter of the kernel is not too large. Fortunately, one can mimic the argument from [3] to prove that with probability at least $1 - 1/\ker(n, M)$ the diameter of the kernel is bounded from above by $\ln \ker(n, M)$. Now we need to place $\mathrm{cr}(n, M) - \ker(n, M)$ vertices of degree two at the $\ker(n, M) + \kappa(n, M)$ edges of the kernel. Note that the probability that we place more than $r - \ln \ker(n, M) - 1$ vertices of degree two on some path of length $\ln \ker(n, M) + 2$ is the same as the probability that a random variable Z with hypergeometric distribution with parameters $\mathrm{cr}(n, M) + \kappa(n, M) - 1 = (8 + o(1))s^2/n$, $r - 1$, and $\ker(n, M) + \kappa(n, M) - 1 = (16 + o(1))s^3/n^2$ takes a value smaller than $\ln \ker(n, M) + 2 = (1 + o(1)) \ln s^3/n^2$. Now take $r_0 = 5ns^{-1} \ln s^3 n^{-2}$. Then the expectation of Z is $(10 + o(1)) \ln \ker(n, M)$ and by known estimates for the tail of the hypergeometric distribution (see [9] Thm.2.10),

$$\Pr(Z \le 0.101\mathbb{E}Z) \le \ker(n, M)^{-2.6}\ .$$

Now for each of the $\binom{\ker(n,M)}{2}$ pairs of vertices v, v'' of the kernel, consider all paths which start at a neighbour of v, go from v to v' using the shortest path in the kernel, and end at a neighbour of v'. Since our graph is basically cubic (see Lemma 5(v)), there are at most $5\binom{\ker(n,M)}{2}$ such paths, each of length at most $\ln \ker(n, M) + 2$. Hence, the probability that we put fewer than r_0 vertices of degree two on some of them is smaller than $10 \ker(n, M)^{-0.6} \leq 5s^{-0.18}n^{-0.12}$ and so is the probability that the core has diameter larger than $5.1ns^{-1} \ln s^3n^{-2}$. This fact, together with (ii), gives (iii). □

4. Proof of the main result

Proof of Theorem 1. Let $s = s(n)$ be any function such that $s = o(n)$ but $sn^{-2/3} \to \infty$. Furthermore, let $M' = n/2+s$ and $M'' = n/2+s+s/\ln s^3n^{-2}$. We show first that with probability at least $1 - n^{0.01}s^{-0.015}$ for all M, $M' \leq M \leq M''$, the estimates (1) and (2) hold.

We shall look at the process backwards, i.e., we shall delete edges from $G(n, M'')$ rather than add them to $G(n, M')$. Due to Lemma 4(iv), with probability at least $1 - 5n^{0.02}s^{-0.03}$, $G^L(n, M'')$ contains at least $m_0 = s^{0.15}n^{-0.1}$ trees of diameter larger than $0.44ns^{-1} \ln s^3n^{-2}$ but smaller than $0.56ns^{-1} \ln s^3n^{-2}$. The probability that in the process of removing $s/\ln s^3n^{-2}$ randomly chosen edges from $G(n, M'')$ we delete some edge of a given path shorter than $0.56ns^{-1} \ln s^3n^{-2}$ is smaller than $1 - e^{-1.11} < 0.7$, so the probability that all of the paths of length larger than $0.44ns^{-1} \ln s^3n^{-2}$ will be destroyed is bounded from above by $0.7^{m_0} \leq s^{-3}n^2$. This shows that with probability at least $1 - 6s^{-0.03}n^{0.02}$ for all M, $M' \leq M \leq M''$, we have $\mathrm{diam_s}(G(n, M)) \geq 0.44ns^{-1} \ln s^3n^{-2}$.

Now let ρ_1 be the probability that in the process of deleting edges from $G(n, M'')$, for some M, $M' \leq M \leq M''$, we create a component H of diameter larger than $0.56ns^{-1} \ln s^3n^{-2}$ outside the giant component. Note that the probability that we do not destroy a given path of length $0.56ns^{-1} \ln s^3n^{-2}$ by deleting edges, is larger than $e^{-1.13} > 0.3$, so with probability at least $0.3\rho_1$ we will see such a component of large diameter in $G^L(n, M')$. Hence, Lemma 4(iv) and (v) imply that $\rho_1 \leq 17s^{-0.03}n^{0.02}$, which results in a uniform upper bound for the diameter of $G^L(n, M)$, $M' \leq M \leq M''$.

In a similar way let ρ_2 denote the probability that in the process of removing edges from $G(n, M'')$, for some M, $M' \leq M \leq M''$, we create a pair of vertices at distance larger than $6.6ns^{-1} \ln s^3n^{-2}$. Then, again, with probability at least $e^{-14}\rho_2$ we will see such a pair in $G(n, M')$, so, by Lemma 4(iii), we have $\rho_2 \leq e^{14}s^{-0.03}n^{0.02}$.

Finally, we show the lower bound for the diameter of the largest component. Note that with probability at least $1 - 5s^{-0.03}n^{0.02}$ the number of whiskers of length at least $0.4ns^{-1} \ln s^3n^{-2}$ in $G(n, M'')$ is at least $t = s^{0.03}n^{-0.02}$. This fact implies that there are at least $\lfloor t/2 \rfloor > t/3$ pairs of vertices $\{v, w\}$ in $G(n, M'')$ such that v lies at distance $0.8ns^{-1} \ln s^3n^{-2}$ from w, and the shortest paths joining different pairs are edge disjoint. Thus, arguing as in the previous cases, we infer that with probability at least $1 - s^{-3}n^2$ the distance between at least one such pair, say $\{v_0, w_0\}$, will be the same also in $G(n, M')$, after we have removed $M'' - M'$ edges from $G(n, M'')$. Note also, that by Lemma 4(iv) and (v), with large probability no pair of vertices in $G^L(n, M)$ are at such a large distance from each other; hence both v_0 and w_0 must belong to the giant component in $G(n, M')$, and thus also in $G(n, M)$, for all $M' \le M \le M''$.

Once we have shown that with probability at least $1 - n^{0.01}s^{-0.015}$ the estimates (1) and (2) hold for all M, $n/2 + s \le M \le n/2 + s + s/\ln s^3n^{-2}$, it follows that they hold with probability at least $1 - n^{0.001}s^{-0.0015}$ for all M such that $n/2 + s \le M \le n/2 + 2s$. Indeed, it is enough to split the interval $[n/2 + s, n/2 + 2s]$ into $3\ln s^3n^{-2}$ smaller ones and add the probabilities that some estimate for the diameter fails in one of them. Finally, take $M_i = n/2 + s_i$, $i = 0, 1, \ldots, j$, where $s_i = 2^i\omega n^{1/3}$ and j is the smallest number such that $s_j \ge n/\omega$. Then, the probability that one of the estimates (1) and (2) fails for some M, $M_0 \le M \le M_j$ is bounded from above by the probability that it fails in some interval $M \in [M_i, M_{i+1}]$, $i = 0, 1, \ldots, j-1$, which in turn is smaller than

$$\sum_{i=0}^{j-1} 2^{-0.0015i}\omega^{-0.0015} = O\big(\omega^{-0.0015}\big) = o(1)\,.$$

This concludes the proof of Theorem 1. □

5. Final remarks

One can use Lemma 2 and argue as in the proof of Theorem 1, to show the following global result for the subcritical phase.

Theorem 7. *Let $\omega = \omega(n) \to \infty$. Then, a.a.s. the random graph process $\{G(n, M)\}_M$ is such that for every $M = n/2 - s$, where $\omega n^{2/3} \le s \le n/\omega$, we have*

$$0.4\frac{n}{s}\ln\frac{s^3}{n^2} \le \operatorname{diam}(G(n, M)) \le 0.6\frac{n}{s}\ln\frac{s^3}{n^2}\,. \tag{5}$$

The constants 0.4 and 0.6 in the estimates (1) and (5) are by no means best possible and are chosen just to simplify the proof, which would otherwise contain even more odd-looking functions of s^3/n^2; after some work one can replace both of them by $0.5 + o(1)$. The same applies to the constant 0.7 in (2), which can be replaced by $1 + o(1)$.

Observe that for the late subcritical case, when $M = n/2 - s$, $s = o(n)$ and $sn^{-2/3} \to \infty$, the diameter of $G(n, M)$ is much larger than the square of the size of the largest tree component. Moreover, for this range of M, $G(n, M)$ has a.a.s. a unique largest component, and hence $G(n, M)$ a.a.s. is such that the diameter of the largest component is smaller than the diameter of the whole graph. The global version of this result also holds, although its proof requires more technical work. Furthermore, one can easily show that a.a.s. in the random process the largest component determines the diameter of the graph in the whole supercritical phase, i.e., the following result holds.

Theorem 8. *Let $\omega = \omega(n) \to \infty$. Then, a.a.s. the random graph process $\{G(n, M)\}_M$ is such that for every $M \ge n/2 + \omega n^{2/3}$,*

$$1.1\operatorname{diam}_{\mathrm{s}}(G(n, M)) \le \operatorname{diam}_{\mathrm{L}}(G(n, M)) = \operatorname{diam}(G(n, M))\,. \quad \square$$

Thus, the critical period can be identified as the last time in the random graph process when the diameter of the graph is not determined by the diameter of the largest component.

Let us set

$$Z(n,M)=\frac{\mathrm{diam}_{\mathrm{L}}(G(n,M))}{\mathrm{diam}_{\mathrm{s}}(G(n,M))}\,.$$

Fernholz and Ramachandran [7] proved that for $M = cn/2$, where $c > 1$ is a constant, the random variable $Z(n,M)$ converges in probability to an explicitly given constant $\beta(c) > 2$. Theorem 1 suggests however that in the early supercritical period, $Z(n,M)$ converges in probability to a constant β_0, which does not depend on M as far as $M = n/2+s$, $s = o(n)$ but $sn^{-2/3} \to \infty$. Clearly, a natural candidate for the value of β_0 is $\lim_{c\to 1}\beta(c) = 3$.

Let us also comment briefly on the behavior of the maximum diameter which occurs in the random process,

$$D_{\max}(n)=\max\left\{\mathrm{diam}(G(n,M))\colon 0\le M\le\binom{n}{2}\right\}.$$

Then, as one can expect, the following holds.

Theorem 9. *For every constant C*

$$\liminf_{n\to\infty}\Pr(D_{\max}(n)\ge Cn^{1/3})>0\,.$$

On the other hand, for every $\omega = \omega(n) \to \infty$, a.a.s.

$$n^{1/3}/\omega \le D_{\max}(n)\le \omega n^{1/3}\,.$$

Proof. The first part of the assertion follows from the known fact (see Łuczak, Pittel, Wierman [14], or Janson *et al.* [8]) that with probability bounded away from zero, the random graph $G(n,n/2)$ contains a random tree on at least $n^{2/3}$ vertices which, with positive probability (see Szekeres [17]), has diameter larger than $Cn^{1/3}$.

Now let $\omega = \omega(n) \to \infty$. Consider the random graph $G(n,M_\omega)$, where $M_\omega = n/2 + (\omega n)^{1/3}$. Then, a.a.s.

$$D_{\max}\ge \mathrm{diam}(G(n,M_\omega))\ge n^{1/3}/\omega\,.$$

Note also that from Theorem 1 it follows that a.a.s. the random graph process is such that for all $M \ge M_\omega$ we have $\mathrm{diam}(G(n,M)) \le \omega n^{1/3}$. Moreover, clearly for all $M \le M_\omega$ we have

$$\mathrm{diam}(G(n,M))\le \mathrm{long}_{\mathrm{s}}(G(n,M_\omega))+\mathrm{diam}_{\mathrm{L}}(G(n,M_\omega))+\mathrm{cr}(G(n,M_\omega))\,,$$

where $\mathrm{long}_{\mathrm{s}}(G(n,M)$ denotes the length of the longest path outside the largest component. Since due to Lemmas 4 and 5, the right hand side of the above inequality is a.a.s. bounded from above by $\omega n^{1/3}$, the assertion follows. □

Finally, let us remark that Theorems 1, 7 and 9 imply that the diameter of $G(n,M)$ is maximized for $M = n/2+O(n^{2/3})$, which can serve as yet another feature characterizing the critical period in the random graph process.

References

[1] B. Bollobás, *The diameter of random graphs*, Trans. Amer. Math. Soc. **267**(1981), 41–52.

[2] B. Bollobás, *"Random graphs"*, Cambridge University Press, Cambridge, 2001, xviii+498.

[3] B. Bollobás, W. Fernandez de la Vega, *The diameter of random regular graphs*, Combinatorica **2** (1982), 125–134.

[4] Ju.D. Burtin, *Extremal metric characteristics of a random graph. I. Asymptotic estimates. (Russian. English summary)* Teor. Verojatnost. i Primenen. **19** (1974), 740–754.

[5] Ju.D. Burtin, *Extremal metric characteristics of a random graph. II. Limit distributions. (Russian. English summary)* Teor. Verojatnost. i Primenen. **20** (1975), 82–99.

[6] F. Chung, L. Lu, *The diameter of sparse random graphs*, Adv. in Appl. Math. **26** (2001), 257–279.

[7] D. Fernholz, V. Ramachandran, *The diameter of sparse random graphs*, Random Structures Algorithms **31** (2007), 482–516.

[8] S. Janson, D.E. Knuth, T. Łuczak, B. Pittel, *The birth of the giant component*, Random Structures Algorithms **4** (1993), 231–358.

[9] S. Janson, T. Łuczak, and A. Ruciński, *"Random graphs"*, Wiley-Interscience, New York, 2000, xii+333.

[10] T. Łuczak, *Component behavior near the critical point of the random graph process*, Random Structures Algorithms **1** (1990), 287–310

[11] T. Łuczak, *Cycles in a random graph near the critical point*, Random Structures Algorithms **2** (1991), 421–439.

[12] T. Łuczak, *The number of trees with large diameter*, J. Austral. Math. Soc. Ser. A **58** (1995), 298–311.

[13] T. Łuczak, *Random trees and random graphs*, Random Structures Algorithms **13** (1998), 485–500.

[14] T. Łuczak, B. Pittel, J.C. Wierman, *The structure of a random graph at the point of the phase transition*, Trans. Amer. Math. Soc. **341** (1994), 721–748.

[15] A. Nachmias, Y. Peres, *Critical random graphs: diameter and mixing time*, Ann. Probab. **36** (2008), 1267–1286.

[16] Yu.L. Pavlov, *Limit distributions of the height of a random forest (Russian)*, Teor. Veroyatnost. i Primenen. **28** (1983), 449–457.

[17] G. Szekeres, *Distribution of labelled trees by diameter*, Combinatorial mathematics, X (Adelaide, 1982), 392–397, Lecture Notes in Math., 1036, Springer, Berlin, 1983.

Department of Discrete Mathematics, Adam Mickiewicz University, ul. Umultowska 87, 61-614 Poznań, Poland

E-mail address: tomasz@amu.edu.pl

University of Oslo, Institute of Basic Medical Sciences, Department of Biostatistics, Postboks 1122, Blindern N-0317, Oslo, Norway

E-mail address: taralgs@medisin.uio.no