ISRAEL JOURNAL OF MATHEMATICS **176** (2010), 29–41
DOI: 10.1007/s11856-010-0019-8

THE CRITICAL RANDOM GRAPH, WITH MARTINGALES

BY

Asaf Nachmias*

Department of Mathematics, UC Berkeley
Berkeley, CA 94720, USA
e-mail: asafnach@math.berkeley.edu

AND

Yuval Peres*

Microsoft Research, One Microsoft Way
Redmond, WA 98052-6399, USA
e-mail: peres@stat.berkeley.edu

ABSTRACT

We give a short proof that the largest component $\mathcal{C}_1$ of the random graph $G(n, 1/n)$ is of size approximately $n^{2/3}$. The proof gives explicit bounds for the probability that the ratio is very large or very small. In particular, the probability that $n^{-2/3}|\mathcal{C}_1|$ exceeds A is at most e^{-cA^3} for some $c > 0$.

1. Introduction

The random graph $G(n, p)$ is obtained from the complete graph on n vertices, by independently retaining each edge with probability p and deleting it with probability $1 - p$. Erdős and Rényi [8] introduced this model in 1960, and discovered that as c grows, $G(n, c/n)$ exhibits a **double jump**: the cardinality of the largest component $\mathcal{C}_1$ is of order $\log n$ for $c < 1$, of order $n^{2/3}$ for $c = 1$ and linear in n for $c > 1$. In fact, for the critical case $c = 1$ the argument in [8]

* U.C. Berkeley and Microsoft Research. Research of both authors supported in part by NSF grants #DMS-0244479 and #DMS-0104073.
Received February 3, 2006 and in revised form August 2, 2007

only established the lower bound on $\mathbf{P}(|\mathcal{C}_1| > An^{2/3})$ for some constant $A > 0$; the upper bound was proved much later in [4] and [12].

Short proofs of the results stated above for the noncritical cases $c < 1$ and $c > 1$ can be found in the books [2], [5] and [10]. However, we could not find a short and self-contained analysis of the case $c = 1$ in the literature. We prove the following two theorems:

Theorem 1 (see [16] and [17] for similar estimates): *Let $\mathcal{C}_1$ denote the largest component of $G(n, 1/n)$, and let $C(v)$ be the component that contains a vertex v. For any $n > 1000$ and $A > 8$ we have*

$$\mathbf{P}(|C(v)| > An^{2/3}) \leq 4n^{-1/3}e^{-\frac{A^2(A-4)}{32}},$$

and

$$\mathbf{P}(|\mathcal{C}_1| > An^{2/3}) \leq \frac{4}{A}e^{-\frac{A^2(A-4)}{32}}.$$

Theorem 2: *For any $0 < \delta < 1/10$ and $n > 200/\delta^{3/5}$, the random graph $G(n, 1/n)$ satisfies*

$$\mathbf{P}\big(|\mathcal{C}_1| < \lfloor \delta n^{2/3} \rfloor\big) \leq 15\delta^{3/5}.$$

While the estimates in these two theorems are not optimal, they are explicit, so the theorems say something about $G(n, 1/n)$ for $n = 10^9$ and not just as $n \to \infty$. The theorems can be extended to the "critical window" $p = 1/n + \lambda n^{-4/3}$, see Section 6. As noted above, Erdős and Rényi [8] proved a version of Theorem 2; their argument was based on counting tree components of $G(n, 1/n)$. However, to prove Theorem 1 by a similar counting argument requires consideration of subgraphs that are not trees. Indeed, with such considerations, Pittel [16] proves tail bounds on $n^{-2/3}|C_1|$ that are asymptotically more precise than Theorems 1 and 2. For a probabilistic approach to Theorem 1 which does not use martingales, see Scott and Sorkin [17].

The systematic study of the phase transition in $G(n, p)$ around the point $p \sim 1/n$ was initiated by Bollobás [4] in 1984 and an upper bound of order $n^{2/3}$ for the median (or any quantile) of $|\mathcal{C}_1|$ was first proved by Łuczak [12]. Łuczak, Pittel and Wierman [13] subsequently proved the following more precise result.

Theorem 3 (Łuczak, Pittel and Wierman 1994): *Let $p = 1/n + \lambda n^{-4/3}$ where $\lambda \in \mathbb{R}$ is fixed. Then for any integer $m > 0$, the sequence*

$$(n^{-2/3}|\mathcal{C}_1|, n^{-2/3}|\mathcal{C}_2|, \ldots, n^{-2/3}|\mathcal{C}_m|)$$

converges in distribution to a random vector with positive components.

The proofs in [12], [13] and [16] are quite involved, and use the detailed asymptotics from [19], [4] and [3] for the number of graphs on k vertices with $k+\ell$ edges. Aldous [1] gave a more conceptual proof of Theorem 3 using diffusion approximation, and identified the limiting distribution in terms of excursion lengths of reflected Brownian motion with variable drift. The argument in [1] is beautiful but not elementary, and it seems hard to extract from it explicit estimates for specific finite n. A powerful approach, that works in the more general setting of percolation on certain finite transitive graphs, was recently developed in [6]. This work is based on the lace expansion, and is quite difficult.

Our proofs of Theorems 1 and 2 use an exploration process introduced in [14] and [11], and the following classical theorem (see, e.g., [7, Section 4], or [18]).

Theorem 4 (Optional stopping theorem): *Let $\{X_t\}_{t\geq 0}$ be a martingale for the increasing σ-fields $\{\mathcal{F}_t\}$ and suppose that τ_1, τ_2 are stopping times with $0 \leq \tau_1 \leq \tau_2$. If the process $\{X_{t\wedge\tau_2}\}_{t\geq 0}$ is bounded, then $\mathbf{E}\,X_{\tau_1} = \mathbf{E}\,X_{\tau_2}$.*

Remark: If $\{X_t\}_{t\geq 0}$ is a submartingale (supermartingale), then under the same boundedness condition, we have $\mathbf{E}\,X_{\tau_1} \leq \mathbf{E}\,X_{\tau_2}$ (respectively, $\mathbf{E}\,X_{\tau_1} \geq \mathbf{E}\,X_{\tau_2}$).

The rest of the paper is organized as follows. In Section 2, we present the exploration process mentioned above. In Section 3, we present a very simple proof of the fact that in $G(n, 1/n)$ we have $\mathbf{P}(|\mathcal{C}_1| > An^{2/3}) \leq 6A^{-3/2}$. The proofs of Theorems 1 and 2 are then presented in Sections 4 and 5. The technical modifications required to handle the "critical window" $p = 1/n + \lambda n^{-4/3}$ are presented in Section 6.

2. The exploration process

For a vertex v, let $\mathcal{C}(v)$ denote the connected component that contains v. We recall an exploration process, developed independently by Martin-Löf [14] and Karp [11]. In this process, vertices will be either **active, explored** or **neutral**. At each time $t \in \{0, 1, \ldots, n\}$, the number of active vertices will be denoted Y_t and the number of explored vertices will be t. Fix an ordering of the vertices, with v first. At time $t = 0$, the vertex v is active and all other vertices are neutral, so $Y_0 = 1$. In step $t \in \{1, \ldots, n\}$, if $Y_{t-1} > 0$ let w_t be the first active vertex; if $Y_{t-1} = 0$, let w_t be the first neutral vertex. Denote by η_t the number of neutral neighbors of w_t in $G(n, 1/n)$, and change the status of these vertices to **active**. Then, set w_t itself **explored**.

Write $N_t = n - Y_t - t - \mathbf{1}_{\{Y_t=0\}}$. Given $Y_1, \ldots, Y_{t-1}$, the random variable η_t is distributed $\mathrm{Bin}(N_{t-1}, 1/n)$, and we have the recursion

$$Y_t = \begin{cases} Y_{t-1} + \eta_t - 1, & Y_{t-1} > 0 \\ \eta_t, & Y_{t-1} = 0. \end{cases} \tag{1}$$

At time $\tau = \min\{t \geq 1 : Y_t = 0\}$ the set of explored vertices is precisely $\mathcal{C}(v)$, so $|\mathcal{C}(v)| = \tau$.

To prove Theorem 1, we will couple $\{Y_t\}$ to a random walk with shifted binomial increments. We will need the following lemma concerning the **overshoots** of such walks.

Lemma 5: *Let $p \in (0,1)$ and $\{\xi_i\}_{i\geq 1}$ be i.i.d. random variables with* $\mathrm{Bin}(n,p)$ *distribution and let $S_t = 1 + \sum_{i=1}^{t}(\xi_i - 1)$. Fix an integer $H > 0$, and define*

$$\gamma = \min\{t \geq 1 \,:\, S_t \geq H \text{ or } S_t = 0\}\,.$$

Let $\Xi \subset \mathbb{N}$ be a set of positive integers. Given the event $\{S_\gamma \geq H,\, \gamma \in \Xi\}$, the conditional distribution of the overshoot $S_\gamma - H$ is stochastically dominated by the binomial distribution $\mathrm{Bin}(n,p)$.

Proof. First observe that if ξ has a $\mathrm{Bin}(n,p)$ distribution, then the conditional distribution of $\xi - r$, given $\xi \geq r$ is stochastically dominated by $\mathrm{Bin}(n,p)$. To see this, write ξ as a sum of n indicator random variables $\{I_j\}_{j=1}^n$ and let J be the minimal index such that $\sum_{j=1}^{J} I_j = r$. Given J, the conditional distribution of $\xi - r$ is $\mathrm{Bin}(n-J, p)$ which is certainly dominated by $\mathrm{Bin}(n,p)$.

For any $\ell \in \Xi$, conditioned on $\{\gamma = \ell\} \cap \{S_{\ell-1} = H - r\} \cap \{S_\gamma \geq H\}$, the overshoot $S_\gamma - H$ equals $\xi_\ell - r$ where ξ_ℓ has a $\mathrm{Bin}(n,p)$ distribution conditioned on $\xi_\ell \geq r$. The assertion of the lemma follows by averaging. ∎

Corollary 6: *Let X be distributed* $\mathrm{Bin}(n,p)$ *and let f be an increasing real function. With the notation of the previous lemma, we have*

$$\mathbf{E}\left[f(S_\gamma - H) \mid S_\gamma \geq H,\, \gamma \in \Xi\right] \leq \mathbf{E}\, f(X)\,.$$

3. An Easy Upper Bound

Fix a vertex v. To analyze the component of v in $G(n, 1/n)$, we use the notation established in the previous section. We can couple the sequence $\{\eta_t\}_{t\geq 1}$ constructed there, to a sequence $\{\xi_t\}_{t\geq 1}$ of i.i.d. $\mathrm{Bin}(n, 1/n)$ random variables, such that $\xi_t \geq \eta_t$ for all $t \leq n$. The random walk $\{S_t\}$ defined in Lemma 5 satisfies $S_t = S_{t-1} + \xi_t - 1$ for all $t \geq 1$ and $S_0 = 1$. Fix an integer $H > 0$ and define γ as in Lemma 5. Couple S_t and Y_t such that $S_t \geq Y_t$ for all $t \leq \gamma$. Since $\{S_t\}$ is a martingale, optional stopping gives $1 = \mathbf{E}\,[S_\gamma] \geq H\mathbf{P}(S_\gamma \geq H)$, whence

$$\mathbf{P}(S_\gamma \geq H) \leq 1/H\,. \tag{2}$$

Write $S_\gamma^2 = H^2 + 2H(S_\gamma - H) + (S_\gamma - H)^2$ and apply Corollary 6 with $f(x) = 2Hx + x^2$ to get for $H \geq 2$ that

$$\mathbf{E}\,[S_\gamma^2 \mid S_\gamma \geq H] \leq H^2 + 2H + 2 \leq H^2 + 3H\,. \tag{3}$$

Now $S_t^2 - (1 - 1/n)t$ is also a martingale. By optional stopping, (2) and (3),

$$1 + (1 - 1/n)\mathbf{E}\,\gamma = \mathbf{E}\,(S_\gamma^2) = \mathbf{P}(S_\gamma \geq H)\mathbf{E}\left[S_\gamma^2 \mid S_\gamma \geq H\right] \leq H + 3\,,$$

hence we have for $2 \leq H \leq n - 3$ that

$$\mathbf{E}\,\gamma \leq H + 3\,. \tag{4}$$

We conclude that for $2 \leq H \leq n - 3$

$$\mathbf{P}(\gamma \geq H^2) \leq (H + 3)/H^2 \leq 2/H\,.$$

Define $\gamma^* = \gamma \wedge H^2$, and so by the previous inequality and (2) we have

$$\mathbf{P}(S_{\gamma^*} > 0) \leq \mathbf{P}(S_\gamma \geq H) + \mathbf{P}(\gamma \geq H^2) \leq 3/H\,. \tag{5}$$

Let $T = H^2$ and note that if $|C(v)| > H^2$ we must have $S_{\gamma^*} > 0$ so by (5) we deduce $\mathbf{P}(|C(v)| > T) \leq 3/\sqrt{T}$ for all $9 \leq T \leq (n-3)^2$. Denote by N_T the number of vertices contained in components larger than T. Then

$$\mathbf{P}(|\mathcal{C}_1| > T) \leq \mathbf{P}(|N_T| > T) \leq \frac{\mathbf{E}\,N_T}{T} \leq \frac{n\mathbf{P}(|C(v)| > T)}{T}.$$

Putting $T = (\lfloor\sqrt{An^{2/3}}\rfloor)^2$ for any $A > 1$ yields

$$\mathbf{P}(|\mathcal{C}_1| > An^{2/3}) \leq \mathbf{P}(|\mathcal{C}_1| > T) \leq \frac{3n}{(\lfloor\sqrt{An^{2/3}}\rfloor)^3} \leq \frac{6}{A^{3/2}}\,,$$

as $(\lfloor\sqrt{An^{2/3}}\rfloor)^3 \geq (\sqrt{An^{2/3}}-1)^3 \geq nA^{3/2}(1-3A^{-1/2}n^{-1/3}) \geq \frac{A^{3/2}n}{2}$. ∎

4. Proof of Theorem 1

We proceed from (5). Define the process $\{Z_t\}$ by

$$Z_t = \sum_{j=1}^{t}(\eta_{\gamma^*+j}-1)\,. \tag{6}$$

The law of η_{γ^*+j} is stochastically dominated by a $\mathrm{Bin}(n-j,1/n)$ distribution, for $j \leq n$. Hence,

$$\mathbf{E}\,[e^{c(\eta_{\gamma^*+j}-1)} \mid \gamma^*] \leq e^{-c}[1+1/n(e^c-1)]^{n-j} \leq e^{(c+c^2)\frac{n-j}{n}-c} \leq e^{c^2-\frac{cj}{n}}\,,$$

as $e^c - 1 \leq c + c^2$ for any $c \in (0,1)$ and $1+x \leq e^x$ for $x \geq 0$. Since this bound is uniform in S_{γ^*} and γ^*, we have

$$\mathbf{E}\left[e^{cZ_t} \mid S_{\gamma^*}\right] \leq e^{tc^2-\frac{ct^2}{2n}}\,.$$

Write $\mathbf{P}_S$ for the conditional probability given S_{γ^*}. Then for any $c \in (0,1)$, we have

$$\mathbf{P}_S\big(Z_t \geq -S_{\gamma*}\big) \leq \mathbf{P}_S\big(e^{cZ_t} \geq e^{-cS_{\gamma^*}}\big) \leq e^{tc^2-\frac{ct^2}{2n}}e^{cS_{\gamma^*}}\,.$$

By (1), if $Y_{\gamma^*+j} > 0$ for all $0 \leq j \leq t-1$, then $Z_j = Y_{\gamma^*+j} - Y_{\gamma^*}$ for all $1 \leq j \leq t$. It follows that

$$\begin{aligned}\mathbf{P}\big(\forall j \leq t \quad Y_{\gamma^*+j} > 0 \mid S_{\gamma^*} > 0\big) &\leq \mathbf{E}\,[\mathbf{P}_S(Z_t \geq -S_{\gamma*}) \mid S_{\gamma*} > 0] \\ &\leq e^{tc^2-\frac{ct^2}{2n}}\mathbf{E}\,[e^{cS_{\gamma^*}} \mid S_{\gamma^*} > 0]\,.\end{aligned} \tag{7}$$

By Corollary 6 with $\Xi = \{1,\ldots,H^2\}$, we have that for $c \in (0,1)$,

$$\mathbf{E}\,[e^{cS_{\gamma^*}} \mid \gamma \leq H^2,\, S_\gamma > 0] \leq e^{cH+c+c^2}\,. \tag{8}$$

Since $\{S_{\gamma^*} > 0\} = \{\gamma > H^2\} \cup \{\gamma \leq H^2,\, S_\gamma > 0\}$ (a disjoint union), the conditional expectation $\mathbf{E}\,[e^{cS_{\gamma^*}} \mid S_{\gamma^*} > 0]$ is a weighted average of the conditional expectation in (8) and of $\mathbf{E}\,[e^{cS_{\gamma^*}} \mid \gamma > H^2] \leq e^{cH}$. Therefore $E[e^{cS_{\gamma^*}} \mid S_{\gamma^*} > 0] \leq e^{cH+c+c^2}$, whence by (7),

$$\mathbf{P}(\forall j \leq t \quad Y_{\gamma^*+j} > 0 \mid S_{\gamma^*} > 0) \leq e^{tc^2-\frac{ct^2}{2n}+cH+c+c^2}\,. \tag{9}$$

By our coupling, for any integer $T > H^2$, if $|C(v)| > T$ then we must have $S_{\gamma^*} > 0$ and $Y_{\gamma^*+j} > 0$ for all $j \in [0, T - H^2]$. Thus, by (5) and (9), we have

$$\begin{aligned} \mathbf{P}(|C(v)| > T) &\le \mathbf{P}(S_{\gamma^*} > 0)\mathbf{P}(\forall j \in [0, T-H^2] \quad Y_{\gamma^*+j} > 0 \mid S_{\gamma^*} > 0) \\ &\le \frac{3}{H} e^{(T-H^2)c^2 - \frac{c(T-H^2)^2}{2n} + cH + c + c^2} . \end{aligned} \tag{10}$$

Take $H = \lfloor n^{1/3} \rfloor$ and $T = \lfloor An^{2/3} \rfloor$ for some $A > 4$; substituting c which attains the minimum of the parabola in the exponent of the right hand side of (10) gives

$$\begin{aligned} \mathbf{P}(|C(v)| > An^{2/3}) &\le 4n^{-1/3} e^{-\frac{((T-H^2)^2/(2n) - H - 1)^2}{4(T-H^2+1)}} \\ &\le 4n^{-1/3} e^{-\frac{((A-1-n^{-2/3})^2/2 - 1 - n^{-1/3})^2}{4(A-1+2n^{-1/3}+n^{-2/3})}} \\ &\le 4n^{-1/3} e^{-\frac{(\frac{(A-2)^2}{2} - 2)^2}{4(A-1/2)}} , \end{aligned}$$

since $H^2 \ge n^{2/3}(1 - 2n^{-1/3})$ and $n > 1000$. As $[(A-2)^2/2 - 2]^2 = A^2(A/2-2)^2$ and $(A/2 - 2)/(A - 1/2) > 1/4$ for $A > 8$ we get

$$\mathbf{P}(|C(v)| > An^{2/3}) \le 4n^{-1/3} e^{-\frac{A^2(A-4)}{32}} .$$

Denote by N_T the number of vertices contained in components larger than T. Then

$$\mathbf{P}(|\mathcal{C}_1| > T) \le \mathbf{P}(|N_T| > T) \le \frac{\mathbf{E}\, N_T}{T} \le \frac{n\mathbf{P}(|C(v)| > T)}{T} ,$$

and we conclude that for all $A > 8$ and $n > 1000$,

$$\mathbf{P}(|\mathcal{C}_1| > An^{2/3}) \le \frac{4}{A} e^{-\frac{A^2(A-4)}{32}} . \qquad \blacksquare$$

5. Proof of Theorem 2

Let h, T_1 and T_2 be positive integers, to be specified later. The proof is divided into two stages. In the first, we ensure, with high probability, ascent of $\{Y_t\}$ to height h by time T_1. In the second stage we show that Y_t is likely to remain positive for T_2 steps.

Stage 1: Ascent to height h: Define

$$\tau_h = \min\{t \le T_1 \, : \, Y_t \ge h\}$$

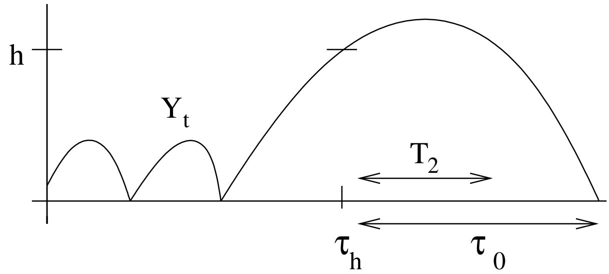

Figure 1. $\tau_0 \geq T_2$.

if this set is nonempty, and $\tau_h = T_1$ otherwise. If $Y_{t-1} > 0$, then $Y_t^2 - Y_{t-1}^2 = (\eta_t - 1)^2 + 2(\eta_t - 1)Y_{t-1}$. Recall that η_t is distributed as $\mathrm{Bin}(N_{t-1}, 1/n)$ conditioned on Y_{t-1}, and hence if we also require $Y_{t-1} \leq h$ then

$$\mathbf{E}\left[Y_t^2 - Y_{t-1}^2 \mid Y_{t-1}\right] \geq \frac{n-t-h}{n}(1-1/n) - 2\frac{t+h}{n}\,h\,.$$

Next, we require that $h < \sqrt{n}/4$ and $t \leq T_1 = \lceil \frac{n}{8h} \rceil$, whence

$$\mathbf{E}\left[Y_t^2 - Y_{t-1}^2 \,\middle|\, Y_{t-1}\right] \geq 1/2 \tag{11}$$

as long as $0 < Y_{t-1} \leq h$. Similarly, (11) holds if $Y_{t-1} = 0$. Thus $Y_{t\wedge\tau_h}^2 - (t\wedge\tau_h)/2$ is a submartingale. The proof of Lemma 5 implies that conditional on $Y_{\tau_h} \geq h$, the overshoot $Y_{\tau_h} - h$ is stochastically dominated by a $\mathrm{Bin}(n, 1/n)$ variable. So, apply Corollary 6 as in (3) with $f(x) = 2hx + x^2$ to get that $\mathbf{E}\,Y_{\tau_h}^2 \leq h^2 + 3h \leq 2h^2$ for $h \geq 3$. By optional stopping,

$$2h^2 \geq \mathbf{E}\,Y_{\tau_h}^2 \geq \frac{1}{2}\mathbf{E}\,\tau_h \geq \frac{T_1}{2}\mathbf{P}(\tau_h = T_1)\,,$$

so

$$\mathbf{P}(\tau_h = T_1) \leq \frac{4h^2}{T_1} \leq \frac{32h^3}{n}\,. \tag{12}$$

Stage 2: Keeping Y_t positive for T_2 steps: Define $\tau_0 = \min\{s : Y_{\tau_h+s} = 0\}$ if this set is nonempty, and $\tau_0 = T_2$ otherwise. Let $M_s = h - \min\{h, Y_{\tau_h+s}\}$. If $0 < M_{s-1} < h$, then

$$M_s^2 - M_{s-1}^2 \leq (\eta_{\tau_h+s} - 1)^2 + 2(1 - \eta_{\tau_h+s})M_{s-1}\,,$$

so provided $h < \sqrt{n}/4$ and $s \leq T_2 \leq n/(8h)$, and recalling that $\tau_h \leq T_1 = \lceil n/(8h) \rceil$ we have $\mathbf{E}\left[M_s^2 - M_{s-1}^2 \mid Y_{\tau_h+s-1}, \tau_h\right] \leq 2$. This also holds if $Y_{\tau_h+s-1} \geq$

h, so $\{M^2_{S\wedge\tau_0} - 2(s\wedge\tau_0)\}_{s=0}^{T_2}$ is a supermartingale. Given the event $\{Y_{\tau_h} \geq h\}$ write $\mathbf{P}_h$ for conditional probability and $\mathbf{E}_h$ for conditional expectation. Since $\{M^2_{s\wedge\tau_0} - 2(s\wedge\tau_0)\}_{s=0}^{T_2}$ is a supermartingale beginning at 0 under $\mathbf{E}_h$, optional stopping yields

$$\mathbf{E}_h M^2_{\tau_0\wedge T_2} \leq 2\mathbf{E}_h[\tau_0\wedge T_2] \leq 2T_2\,, \tag{13}$$

whence

$$\mathbf{P}_h(\tau_0 < T_2) \leq \mathbf{P}_h(M_{\tau_0\wedge T_2} \geq h) \leq \frac{\mathbf{E}_h M^2_{\tau_0\wedge T_2}}{h^2} \leq \frac{2T_2}{h^2}\,. \tag{14}$$

In conjunction with (12), this yields

$$\mathbf{P}(\tau_0 < T_2) \leq \mathbf{P}(\tau_h = T_1) + \mathbf{E}\,\mathbf{P}_h(\tau_0 < T_2) \leq \frac{32h^3}{n} + \frac{2T_2}{h^2}\,. \tag{15}$$

Let $T_2 = \lfloor \delta n^{2/3} \rfloor$ and choose h to approximately minimize the right-hand side of (15). This gives $h = \lfloor \frac{\delta^{1/5} n^{1/3}}{(24)^{1/5}} \rfloor$, which satisfies $T_2 \leq n/(8h)$ and makes the right-hand side of (15) less than $15\delta^{3/5}$. Since $|\mathcal{C}_1| < T_2$ implies $\tau_0 < T_2$, this concludes the proof. ∎

6. The Critical Window

As noted in the introduction, the proofs of Theorems 1 and 2 can be extended to the critical "window" $p = \frac{1+\lambda n^{-1/3}}{n}$ for some constant λ. For Theorem 2 this adaptation is straightforward, and we omit it. However, our proof of Theorem 1 used the fact that for $\lambda = 0$ (that is, $p = 1/n$) the exploration process is stochastically dominated by a mean zero random walk, so we include the necessary adaptation below.

Theorem 7: *Set $p = (1+\lambda n^{-1/3})/n$ for some $\lambda \in \mathbb{R}$ and consider $G(n,p)$. For $\lambda > 0$ and $A > 2\lambda + 3$ we have that for large enough n*

$$\mathbf{P}(|C(v)| \geq An^{2/3}) \leq \Big(\frac{4\lambda}{1-e^{-4\lambda}} + 16\Big) n^{-1/3} e^{-\frac{((A-1)^2/2-(A-1)\lambda-2)^2}{4A}}\,,$$

and

$$\mathbf{P}(|\mathcal{C}_1| \geq An^{2/3}) \leq \Big(\frac{4\lambda}{A(1-e^{-4\lambda})} + \frac{16}{A}\Big) e^{-\frac{((A-1)^2/2-(A-1)\lambda-2)^2}{4A}}\,.$$

For $\lambda < 0$ and $A > 3$ we have that for large enough n

$$\mathbf{P}(|C(v)| \geq An^{2/3}) \leq \Big(\frac{-2\lambda}{e^{-\lambda}-1} + \min(5, -\frac{1}{\lambda})\Big) n^{-1/3} e^{-\frac{((A-1)^2/2-(A-1)\lambda-2)^2}{4A}}\,,$$

and

$$\mathbf{P}(|\mathcal{C}_1| \geq An^{2/3}) \leq \Big(\frac{-2\lambda}{A(e^{-\lambda}-1)} + \min(5, -\frac{1}{\lambda})\Big) e^{-\frac{((A-1)^2/2-(A-1)\lambda-2)^2}{4A}} \,.$$

Proof. Assume $p = 1/n + \lambda n^{-4/3}$ and that n is large enough; again we bound the exploration process with a process $\{S_t\}$ defined by $S_t = S_{t-1} + \xi_t - 1$ where ξ_t are distributed as $\mathrm{Bin}(n,p)$ and $S_0 = 1$. The two cases of λ being positive or negative are dealt with separately; assume first $\lambda > 0$. Since $1 - e^{-a} \leq a - a^2/3$ for small enough $a > 0$, we have

$$\mathbf{E}\, e^{-a(\xi_t - 1)} = e^a[1 - p(1-e^{-a})]^n \geq e^a(1 - p(a - a^2/3))^n \,.$$

By Taylor expansion of $\log(1-x)$, for small a we have

$$\begin{aligned}\log \mathbf{E}\, e^{-a(\xi_t-1)} &\geq a + n\big(-p(a - a^2/3) + O(n^{-2})\big) \\ &= a - (1 + \lambda n^{-1/3})(a - a^2/3) + O(n^{-1}) \,,\end{aligned}$$

and so for $a = 4\lambda n^{-1/3}$ and n large, we have $\mathbf{E}\, e^{-a(\xi_t-1)} \geq 1$ hence $\{e^{-aS_t}\}$ is a submartingale. Take $H = \lceil n^{1/3} \rceil$, and define γ as in Lemma 5. Then by optional stopping we have

$$e^{-a} \leq 1 - \mathbf{P}(S_\gamma \geq H) + \mathbf{P}(S_\gamma \geq H) e^{-aH} \,,$$

and as $1 - e^{-a} \leq a$ for $a > 0$ we get

$$\mathbf{P}(S_\gamma \geq H) \leq \frac{4\lambda n^{-1/3}}{1 - e^{-4\lambda}} \,. \tag{16}$$

Also, observe that $S_t - \lambda n^{-1/3} t$ is a martingale, hence by optional stopping $1 + \lambda n^{-1/3}\mathbf{E}\,\gamma = \mathbf{P}(S_\gamma \geq H)\mathbf{E}\,[S_\gamma \mid S_\gamma \geq H]$ and so by Corollary 6 we get $\mathbf{E}\,\gamma \leq \frac{8n^{1/3}}{1-e^{-4\lambda}}$. For $\lambda > 1/4$, as $(1 - e^{-4\lambda})^{-1} \leq 2$, this gives that $\mathbf{E}\,\gamma \leq 16n^{1/3}$. It is immediate to check that $S_t^2 - \frac{1}{2}t$ is a submartingale as long as $t \leq \gamma$, hence by optional stopping $\frac{\mathbf{E}\,\gamma}{2} \leq \frac{4\lambda n^{-1/3}}{1-e^{-4\lambda}} \mathbf{E}\,[S_\gamma^2 | S_\gamma \geq H]$. Using Corollary 6 as in (3) and estimating $\frac{4x}{1-e^{-4x}} \leq 2$ for $x \in (0, 1/4]$ gives the same estimate for $\lambda \in (0, 1/4]$. Thus

$$\mathbf{E}\,\gamma \leq 16n^{1/3} \,, \tag{17}$$

for all $\lambda > 0$. Take again $\gamma^* = \gamma \wedge H^2$, and as in (5), by (16) and (17) we get

$$\mathbf{P}(S_{\gamma^*} > 0) \leq \Big(\frac{4\lambda}{1 - e^{-4\lambda}} + 16\Big) n^{-1/3} \,. \tag{18}$$

Define Z_t as in (6) and note that this time its increments can be stochastically dominated by variables distributed as $\mathrm{Bin}(n-j,p)-1$. Similar computations to the one made in the beginning of Section 4 give that for $c \in (0,1)$

$$\mathbf{E}\left[e^{cZ_t} \mid S_{\gamma^*}\right] \le e^{ct\lambda n^{-1/3} - \frac{ct^2}{2n} + c^2 t(1+\lambda n^{-1/3})},$$

and so as before we have

$$\begin{aligned}\mathbf{P}(\forall j \le t \quad Y_{\gamma^*+j} > 0 \mid S_{\gamma^*} > 0) &\le \mathbf{E}\left[\mathbf{P}_S(Z_t \ge -S_{\gamma*}) \mid S_{\gamma*} > 0\right] \\ &\le e^{ct\lambda n^{-1/3} - \frac{ct^2}{2n} + c^2 t(1+\lambda n^{-1/3})} \mathbf{E}\left[e^{cS_{\gamma^*}} \mid S_{\gamma^*} > 0\right] \\ &\le e^{ct\lambda n^{-1/3} - \frac{ct^2}{2n} + c^2 t(1+\lambda n^{-1/3}) + c(n^{1/3}+1) + 2(c+c^2)}\end{aligned}$$

where the last inequality is due to Corollary 6. Write $t = \lfloor Bn^{2/3} \rfloor$ for some constant B and take $c \in (0,1)$ which attains the minimum of the parabola in the exponent of the last display. This gives that for large enough n and fixed $B > 2\lambda + 2$ we have

$$\mathbf{P}(\forall j \le t \quad Y_{\gamma^*+j} > 0 \mid S_{\gamma^*} > 0) \le e^{-\frac{(B^2/2 - B\lambda - 2)^2}{4(B+1)}}.$$

Together with (18), as in the proof of Theorem 1, we conclude that for any $A > 2\lambda + 3$ we have

$$\mathbf{P}(|C(v)| \ge An^{2/3}) \le \Big(\frac{4\lambda}{1-e^{-4\lambda}} + 16\Big) n^{-1/3} e^{-\frac{((A-1)^2/2 - (A-1)\lambda - 2)^2}{4A}},$$

and as before this implies that

$$\mathbf{P}(|\mathcal{C}_1| \ge An^{2/3}) \le \Big(\frac{4\lambda}{A(1-e^{-4\lambda})} + \frac{16}{A}\Big) e^{-\frac{((A-1)^2/2 - (A-1)\lambda - 2)^2}{4A}}.$$

Assume now $p = 1/n + \lambda n^{-4/3}$ for some fixed $\lambda < 0$. For $a > 0$, as $1 + x \le e^x$ we have

$$\mathbf{E}\, e^{a(\xi_t - 1)} = e^{-a}[1 + p(e^a - 1)]^n \le e^{-a + np(e^a - 1)}.$$

By Taylor expansion of $e^x - 1$ we have

$$\log \mathbf{E}\, e^{a(\xi_t - 1)} \le -a + (1 + \lambda n^{-1/3})(a + a^2/2 + O(a^3)),$$

and so for $a = -\lambda n^{-1/3} > 0$ we have that $\mathbf{E}\, e^{a(\xi_t-1)} \le 1$ hence $\{e^{aS_t}\}$ is a supermartingale. With the same H and γ as before, optional stopping gives

$$e^a \ge 1 - \mathbf{P}(S_\gamma \ge H) + \mathbf{P}(S_\gamma \ge H) e^{an^{1/3}},$$

and as $e^x - 1 \le 2x$ for x small enough we get

$$\mathbf{P}(S_\gamma \ge H) \le \frac{-2\lambda n^{-1/3}}{e^{-\lambda} - 1} .$$

Also, as γ is bounded above by the hitting time of 0, Wald's Lemma (see [7]) implies that $\mathbf{E}\,\gamma \le -n^{1/3}/\lambda$. For $\lambda \in [-1/5, 0]$ it is straight forward to verify that $S^2_{t\wedge\gamma} - \frac{1}{2}(t \wedge \gamma)$ is a submartingale, hence as before we deduce by optional stopping that $\mathbf{E}\,\gamma \le 5n^{1/3}$ for such λ's. Thus we deduce that for all $\lambda < 0$,

$$\mathbf{E}\,\gamma \le \min(5, -1/\lambda)n^{1/3} .$$

The rest of the proof continues from (17), as in the case of $\lambda > 0$. ∎

Remark: Using similar methods, in [15], we analyze component sizes of bond percolation on random regular graphs.

Acknowledgments

We thank David Aldous, Christian Borgs and Jeff Steif for useful suggestions and comments. Part of this research was done while the first author was an intern at Microsoft Research. We are grateful to the referee for a remarkably careful reading of the manuscript and important corrections.

References

[1] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, The Annals of Probability **25** (1997), 812–854.

[2] N. and J. H. Spencer, *The Probabilistic Method*, 2nd edition, Wiley, New York, 2000.

[3] E. A. Bender, E. R. Canfield and B. D. McKay, *The asymptotic number of labeled connected graphs with a given number of vertices and edges*, Random Structures & Algorithms **1** (1990), 127–169.

[4] B. Bollobás, *The evolution of random graphs*, Transactions of the American Mathematical Society **286** (1984), 257–274.

[5] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[6] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade and J. Spencer, *Random subgraphs of finite graphs. I. The scaling window under the triangle condition*, Random Structures & Algorithms **27** (2005), 137–184.

[7] R. Durrett, *Probability: Theory and Examples*, 2nd edition, Duxbury Press, Belmont, California, 1996.

[8] P. Erdős and A. Rényi, *On the evolution of random graphs,* Magyar Tud. Akad. Mat. Kutató Int. Kőzl. **5** (1960), 17–61.

[9] S. Janson, D. E. Knuth, T. Łuczak and B. Pittel, *The birth of the giant component. With an introduction by the editors,* Random Structures & Algorithms **4** (1993), 231–358.

[10] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[11] R. M. Karp, *The transitive closure of a random digraph,* Random Structures & Algorithms **1** (1990), 73–93.

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

[13] T. Łuczak, B. Pittel and J. C. Wierman, *The structure of a random graph at the point of the phase transition,* Transactions of the American Mathematical Society **341** (1994), 721–748.

[14] A. Martin-Löf, *Symmetric sampling procedures, general epidemic processes and their threshold limit theorems*, Journal of Applied Probability **23** (1986), 265–282.

[15] A. Nachmias and Y. Peres, *Critical percolation on random regular graphs,* Random Structures & Algorithms, to appear. Available at http://www.arxiv.org/abs/0707.2839.

[16] B. Pittel, *On the largest component of the random graph at a nearcritical stage,* Journal of Combinatorial Theory. Series B **82** (2001), 237–269.

[17] A. D. Scott and G. B. Sorkin, *Solving sparse random instances of Max Cut and Max 2-CSP in linear expected time,* Combinatorics, Probability and Computing **15** (2006), 281–315.

[18] D. Williams, *Probability with martingales*, Cambridge University Press, Cambridge, 1991.

[19] E. M. Wright, *The number of connected sparsely edged graphs,* Journal of Graph Theory **1** (1977), 317–330.