GROWTH OF COMPONENTS IN RANDOM GRAPHS

SVANTE JANSON

Abstract. The creation and growth of components of a given complexity in a random graph process are studied. In particular, the expected number and total size of all such components is found. It follows that the largest ℓ-component during the process is $O_p(n^{2/3})$ for any given ℓ. The results also yield a new proof of the asymptotic behaviour of Wright's coefficients.

1. Introduction

We consider the growth of size and complexity of the components of a random graph process $\{G(n,m)\}_{0\le m\le\binom{n}{2}}$ or $\{G(n,t)\}_{0\le t\le 1}$. Recall that both processes describe a randomly growing graph on n vertices, where edges are added one by one, and where each new edge is chosen uniformly at random among all remaining possibilities. The only difference between the two processes is that in $\{G(n,m)\}$, the edges are added at the fixed times 1,2,..., so at time m we have the random graph $G(n,m)$ with m edges, while in $\{G(n,t)\}$ the edges are added at random times in such a way that at time $t=p$, we have the random graph $G(n,p)$. ($\{G(n,t)\}$ may be constructed by letting each edge e in the complete graph K_n appear at a random time T_e, with T_e independent and uniformly distributed on $(0,1)$, and letting $G(n,t)$ contain the edges that have appeared before t.)

In this paper, we will study random variables that depend on the order the edges appear in the process, but not on the time scale. All such variables will thus have the same distribution for both processes (not only asymptotically, but also for each finite n). Hence all results below are valid for both processes.

We define the *complexity* of a connected graph to be its number of edges minus its number of vertices. A component of a graph of complexity ℓ is called an ℓ-component. Here $\ell\ge -1$; a (-1)-component is a tree, a 0-component is unicyclic, and ℓ-components with $\ell\ge 1$ are known as complex components.

In the beginning of the random graph process, there are no edges at all, and thus n components of order 1 and complexity -1; at the end, we have the complete graph, with a single component of complexity $\binom{n}{2}-n$. Several authors have studied what happens in between, see e.g. [4, 9, 6, 7, 11, 8]. We will here add some results obtained by studying, as in [6], the ways ℓ-components are created.

Each time a new edge is added, there are two possibilities:

Date: January 26, 2000; revised June 15, 2000.

This is a preprint of an article published in Random Structure & Algoritms **17** (2000), no. 3-4, 343-356.

(i) The edge joins two vertices already in the same component. The number of components and their orders remain the same, but if the edge was added to an ℓ-component, it becomes an $(\ell+1)$-component. We call this a transition $\ell \to \ell+1$.

(ii) The edge joins two vertices in different components. If these have complexities ℓ_1 and ℓ_2, they will merge to an $(\ell_1+\ell_2+1)$-component. We call this a transition $\ell_1 \oplus \ell_2 \to \ell_1+\ell_2+1$.

We say that an ℓ-component is *created* by a transition $\ell-1 \to \ell$ or $\ell_1 \oplus \ell_2 \to \ell$ with $\ell_1, \ell_2 \geq 0$; in contrast, we say (for $\ell \geq 0$) that an ℓ-component *grows* when it swallows a tree by a transition $\ell \oplus -1 \to \ell$.

We can regard the ℓ-components that appear in the random graph process $\{G(n,m)\}$ (or, similarly, $\{G(n,t)\}$) in two ways:

The static view: Let $\mathcal{C}_\ell(m)$ denote the collection of all ℓ-components in $G(n,m)$, and consider the family $\mathcal{C}_\ell^* = \bigcup_m \mathcal{C}_\ell(m)$ of every ℓ-component that appears at some stage of the process, ignoring when it appears. We sometimes, for emphasis, call the elements of $\mathcal{C}_\ell^*$ *static ℓ-components*.

The dynamic view: For $\ell \geq 0$, we can regard a component as "the same" even after it has grown by merging with a tree. In other words, we identify any two elements of $\mathcal{C}_\ell^*$ such that one is contained in the other, and regard the ℓ-components in the random graph process as functions of time; these time-dependent subgraphs of K_n are called *dynamic ℓ-components*. Each dynamic component thus exists during some time interval, in which it is created, grows, and disappears. In particular, a dynamic component does not have a fixed size.

The maximal elements of $\mathcal{C}_\ell^*$ (with respect to inclusion) are called *maximal ℓ-components*. Note that for $\ell \geq 0$, the maximal ℓ-components are in one-to-one correspondence with the dynamic ℓ-components; they are the final forms of the dynamic components before disappearing. For any ℓ, the maximal ℓ-components are disjoint, their vertex sets form a partition of the set $\mathcal{V}_\ell = \bigcup\{V(C) : C \in \mathcal{C}_\ell^*\}$ of every vertex that belongs to an ℓ-component at some stage of the random graph process, and every static ℓ-component is a subgraph of some maximal ℓ-component.

We further define $V_\ell = |\mathcal{V}_\ell|$, the number of vertices that at some time belong to an ℓ-component, and $V_\ell^{\max} = \max\{|V(C)| : C \in \mathcal{C}_\ell^*\}$, the order of the largest ℓ-component that ever appears. Note that, trivially, $V_\ell^{\max} \leq V_\ell$, and that for any m or t, each ℓ-component in $G(n,m)$ or $G(n,t)$ has at most $V_\ell^{\max}$ vertices, while the union of all ℓ-components has at most V_ℓ vertices.

Let $\alpha(\ell;k)$ be the expected number of times that a new edge is added to an ℓ-component of order k, with both ends of the edge in the component. Similarly let $\beta(\ell_1,\ell_2;k_1,k_2)$ be the expected number of ordered pairs (H_1,H_2) such that H_1 and H_2 are two distinct components at some stage of the process, which are joined by the next edge, and such that H_i has k_i vertices and complexity ℓ_i. (Note that, by considering ordered pairs, every edge joining two components is counted twice.)

Our results are based on an evaluation of these expected numbers in Lemma 1 below (Section 2). In the following sections we apply Lemma 1 to, respectively, complex components, unicyclic components, tree components, and complex components again. For example, we show that for any fixed $\ell \geq -1$, the largest ℓ-component that appears during the evolution is of order $O_p(n^{2/3})$. (See e.g. [8] for the definition of O_p and o_p.) In particular, this holds for tree components and unicyclic components, which extends the well-known result that for any given $m = m(n)$ or $p = p(n)$, the maximal order of such components in $G(n, m)$ or $G(n, p)$ is $O_p(n^{2/3})$.

Remark 1. In this paper we study only the expectations of various random variables. In principle, it is possible to compute variances and higher moments too; see [6], where this is done for the special case $\ell = 0$. Such calculations might perhaps be used to show that some of the random variables studied below are concentrated about their means (as we conjecture), but we have not attempted this.

Remark 2. Although we will not do it in this paper, it is possible to study also when the transitions occur in the random graph process by keeping track of the time t in the estimates above, instead of immediately integrating over all t. Again, see [6], where this is done for $\ell = 0$.

Remark 3. It might be possible to make a similar study of the disappearance of components of a given complexity, but so far we have not succeeded to deduce anything useful.

2. The expected number of transitions

Let $C(k, k+\ell)$ denote the number of connected graphs with $k+\ell$ edges on k labelled vertices. Then $C(k, k-1) = k^{k-2}$ (Cayley's formula for the number of labelled trees); more generally, by Wright [13], see also [7, 12], for any fixed $\ell \geq -1$, there exists a constant ρ_ℓ such that

$$C(k, k+\ell) = \rho_\ell k^{k-1/2+3\ell/2}\bigl(1 + O(k^{-1/2})\bigr). \tag{1}$$

In particular, $\rho_{-1} = 1$, $\rho_0 = \sqrt{\pi/8}$ and $\rho_1 = 5/24$.

We can now state our basic result. We use $n^{\underline{k}}$ to denote the falling factorial $n(n-1)\cdots(n-k+1)$.

Lemma 1. *For any $\ell \geq -1$ and $k \geq 1$,*

$$\alpha(\ell; k) = n^{\underline{k}} \frac{(k+\ell)!}{k!} C(k, k+\ell) \left(\binom{k}{2} - k - \ell \right) \frac{(nk - k^2/2 - 3k/2 - \ell - 1)!}{(nk - k^2/2 - k/2)!}, \tag{2}$$

and for any k_1, k_2, ℓ_1, ℓ_2, with $k = k_1 + k_2$ and $\ell = \ell_1 + \ell_2$,

$$\begin{aligned} &\beta(\ell_1, \ell_2; k_1, k_2) \\ &= n^{\underline{k}} (k+\ell)! \frac{k_1 C(k_1, k_1+\ell_1)}{k_1!} \frac{k_2 C(k_2, k_2+\ell_2)}{k_2!} \frac{(nk - k^2/2 - 3k/2 - \ell - 1)!}{(nk - k^2/2 - k/2)!}. \end{aligned} \tag{3}$$

Moreover, for fixed ℓ, and all k and n with $1 \le k \le n$,

$$\alpha(\ell;k) = \tfrac{1}{2}\rho_\ell k^{(3\ell+1)/2} n^{-\ell-1} e^{-k^3/24n^2}\Bigl(1 + O\Bigl(\frac{k}{n} + \frac{k^4}{n^3} + k^{-1/2}\Bigr)\Bigr), \tag{4}$$

and if further $k_1, k_2 \ge 1$, $\ell_1, \ell_2 \ge -1$, $k_1 + k_2 = k$ and $\ell_1 + \ell_2 = \ell$,

$$\beta(\ell_1,\ell_2;k_1,k_2) = \tfrac{1}{\sqrt{2\pi}}\rho_{\ell_1}\rho_{\ell_2} k_1^{3\ell_1/2} k_2^{3\ell_2/2} k^{-1/2} n^{-\ell-1} e^{-k^3/24n^2} \cdot \Bigl(1 + O\Bigl(\frac{k}{n} + \frac{k^4}{n^3} + k_1^{-1/2} + k_2^{-1/2}\Bigr)\Bigr), \tag{5}$$

and also

$$\beta(\ell_1,-1;k_1,k_2) = \rho_{\ell_1} k_1^{3\ell_1/2} \frac{k_2^{k_2-1}e^{-k_2}}{k_2!} k^{-1/2} n^{-\ell_1} e^{-k^3/24n^2} \cdot \Bigl(1 + O\Bigl(\frac{k}{n} + \frac{k^4}{n^3} + k_1^{-1/2}\Bigr)\Bigr). \tag{6}$$

Proof. We do the calculations for $G(n,t)$, which is more convenient. Similar calculations can be performed for $G(n,m)$, leading to sums of terms including several binomial coefficients. By the equivalence discussed above, these sums must equal the values in (2) and (3), although it seems difficult to show this directly. (Perhaps, the Gosper–Zeilberger algorithm [5, Section 5.8] can be used, but we have not tried it.) Asymptotic results are easily derived using these sums too.

For a transition $\ell \to \ell+1$ when a new edge is added to an ℓ-component of order k, there are $\binom{n}{k}C(k,k+\ell)$ possible ℓ-components and $\binom{k}{2} - k - \ell$ possible edges to add; moreover, the probability that a given one of these possible ℓ-components actually is a component of $G(n,t)$ is $t^{k+\ell}(1-t)^{(n-k)k+\binom{k}{2}-k-\ell}$, and the conditional probability that a given edge, not in this component, is added in the time interval $(t, t+dt)$, given that it was not added earlier, is $dt/(1-t)$. Hence, we obtain, integrating over all times,

$$\alpha(\ell;k) = \binom{n}{k} C(k,k+\ell)\left(\binom{k}{2} - k - \ell\right)\int_0^1 t^{k+\ell}(1-t)^{(n-k)k+\binom{k}{2}-k-\ell-1}\,dt.$$

We obtain (2) by evaluating the beta integral.

Similarly, with $k = k_1 + k_2$ and $\ell = \ell_1 + \ell_2$, since there are k_1k_2 ways to join two given components of orders k_1 and k_2,

$$\beta(\ell_1,\ell_2;k_1,k_2) = \binom{n}{k}\binom{k}{k_1} C(k_1,k_1+\ell_1)C(k_2,k_2+\ell_2)k_1k_2 \cdot \int_0^1 t^{k+\ell}(1-t)^{(n-k)k+\binom{k}{2}-k-\ell-1}\,dt,$$

which yields (3).

For the asymptotical results we observe that, for fixed ℓ and $k \le n$,

$$k(n-k/2) \ge nk - k^2/2 - k/2 \ge nk - k^2/2 - 3k/2 - \ell - 1 = k\bigl(n - k/2 + O(1)\bigr) = k(n-k/2)\Bigl(1 + O\Bigl(\frac{1}{n}\Bigr)\Bigr)$$

and thus

$$\begin{aligned}\frac{(nk-k^2/2-3k/2-\ell-1)!}{(nk-k^2/2-k/2)!} &= \frac{1}{(nk-k^2/2-k/2)^{\underline{k+\ell+1}}}\\ &= k^{-k-\ell-1}(n-k/2)^{-k-\ell-1}\Big(1+O\Big(\frac{k}{n}\Big)\Big). \qquad (7)\end{aligned}$$

Moreover, we claim that for $1 \le k \le n$,

$$\frac{n^{\underline{k}}}{(n-k/2)^k} = e^{-k^3/24n^2}\Big(1+O\Big(\frac{k}{n}+\frac{k^4}{n^3}\Big)\Big). \qquad (8)$$

Indeed, if $k \le n^{3/4}$, then, by Taylor expansions,

$$\begin{aligned}\log n^{\underline{k}} &= k\log n + \sum_{i=1}^{k-1}\log(1-i/n) = k\log n + \sum_{i=1}^{k-1}\Big(-\frac{i}{n}-\frac{i^2}{2n^2}\Big)+O\Big(\frac{k^4}{n^3}\Big)\\ &= k\log n - \frac{k^2}{2n} - \frac{k^3}{6n^2} + O\Big(\frac{k}{n}+\frac{k^4}{n^3}\Big)\end{aligned}$$

and

$$\log(n-k/2)^k = k\log n + k\log(1-k/2n) = k\log n - \frac{k^2}{2n} - \frac{k^3}{8n^2} + O\Big(\frac{k^4}{n^3}\Big),$$

which imply (8).

On the other hand, if $n^{3/4} \le k \le n$, then

$$\begin{aligned}0 \le \frac{n^{\underline{k}}}{(n-k/2)^k} &\le 2\prod_{i=1}^{k-1}\frac{n-i}{n-k/2} = 2\prod_{i=1}^{k/2}\frac{(n-i)(n-(k-i))}{(n-k/2)^2}\\ &= 2\prod_{i=1}^{k/2}\Big(1-\frac{(i-k/2)^2}{(n-k/2)^2}\Big) \le 2\exp\left(-\sum_{i=1}^{k/2}\frac{(i-k/2)^2}{(n-k/2)^2}\right)\\ &= 2\exp\left(-\frac{(k/2)^3}{3(n-k/2)^2}+O\Big(\frac{k^2}{n^2}\Big)\right) \le \exp\Big(-\frac{k^3}{24n^2}+O(1)\Big),\end{aligned}$$

and again (8) follows (since now $k^4/n^3 \ge 1$).

Combining (2), (7), (8) and (1), we obtain (4). Similarly, from (3), (7), (8), (1) and Stirling's formula we obtain (5) and (6) (for the latter using Cayley's formula for $C(k_2, k_2-1)$ instead of (1)). $\square$

When using Lemma 1, we will frequently find use for the following simple asymptotic formulas.

Lemma 2.

(i) *If $a > -1$, then*

$$\sum_{1}^{n} k^a e^{-k^3/24n^2} \sim 2^{a+1}3^{(a-2)/3}\Gamma\Big(\frac{a+1}{3}\Big)n^{2(a+1)/3}$$

as $n \to \infty$.

(ii) *If $a, b > -1$, then*

$$\sum_{\substack{k_1+k_2=k\\k_1,k_2\geq 1}} k_1^a k_2^b \sim \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)} k^{a+b+1}$$

as $k \to \infty$.

Proof. For (i) it is easily checked that

$$\sum_1^n k^a e^{-k^3/24n^2} \sim \int_0^n x^a e^{-x^3/24n^2}\,dx = n^{2(a+1)/3} \int_0^{n^{1/3}} y^a e^{-y^3/24}\,dy$$
$$\sim n^{2(a+1)/3} \int_0^\infty y^{a+1} e^{-y^3/24}\,\frac{dy}{y},$$

where the final integral equals, by the substitution $y = (24z)^{1/3}$,

$$\int_0^\infty (24z)^{(a+1)/3} e^{-z}\,\frac{dz}{3z} = (24)^{(a+1)/3} 3^{-1}\Gamma((a+1)/3).$$

For (ii), similarly,

$$\sum_{\substack{k_1+k_2=k\\k_1,k_2\geq 1}} k_1^a k_2^b = \sum_{k_1=1}^{k-1} k_1^a (k-k_1)^b \sim \int_0^k x^a (k-x)^b\,dx$$
$$= k^{a+b+1} \int_0^1 x^a (1-x)^b\,dx = \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)} k^{a+b+1}. \qquad \square$$

3. Creation of complex components

It was observed in [6] that a complex component must be created first as a 1-component, i.e. through a transition $0 \to 1$ or $0 \oplus 0 \to 1$, and the expected numbers of such transitions were calculated (by the same method as here). As a warm-up, we rederive this result using Lemma 1.

Indeed, (4) implies by four applications of Lemma 2(i), recalling $\rho_0 = \sqrt{\pi/8}$, that the expected numbers of transitions $0 \to 1$ is

$$\begin{aligned}\sum_{k=1}^n \alpha(0;k) &= \tfrac12 \rho_0 n^{-1} \sum_{k=1}^n \Big(k^{1/2} + O\Big(\frac{k^{3/2}}{n} + \frac{k^{9/2}}{n^3} + 1\Big)\Big) e^{-k^3/24n^2}\\ &= \tfrac12 \rho_0 2^{3/2} 3^{-1/2} \Gamma(1/2)\big(1+o(1)\big) + O(n^{5/3-2}) + O(n^{11/3-4})\\ &\qquad + O(n^{2/3-1})\\ &= \frac{\pi}{2\sqrt3} + o(1).\end{aligned}$$

Similarly, the expected numbers of transitions $0 \oplus 0 \to 1$ is, with the factor $1/2$ since β counts ordered pairs, by (5) and several applications of Lemma 2(i),(ii)

(we leave the details to the reader)

$$\tfrac12 \sum_{k_1,k_2=1}^{n} \beta(0,0;k_1,k_2) \sim \frac12\Big(\frac{1}{2\pi}\Big)^{1/2}\frac{\pi}{8}\sum_{k=2}^{n} k^{1/2}n^{-1}e^{-k^3/24n^2} \sim \frac{\pi}{8\sqrt3}.$$

Adding these numbers together, we obtain that the expected number of creations of complex components converges to $5\pi/8\sqrt3 \approx 1.134$.

Remark 4. By studying higher moments, it was in [6] further deduced that the probability that the random graph process never has more than one complex component converges to a limit strictly between 0 and 1; this limit was shown to equal $5\pi/18$ in [7].

4. Unicyclic components

Unicyclic components are created by the transition $-1 \to 0$ and grow further by $0 \oplus -1 \to 0$.

Unlike the creation of 1-components studied in the preceding section, the number of transitions $-1 \to 0$ is large; it follows from Lemma 1 (and Lemma 2 for the error terms), that its expectation is

$$\sum_{k=1}^{n} \alpha(-1;k) \sim \tfrac12 \sum_{k=1}^{n} k^{-1}e^{-k^3/24n^2} \sim \tfrac12 \int_1^n \frac{1}{x}e^{-x^3/24n^2}\,dx.$$

To estimate this integral, we write it as

$$\int_1^{n^{2/3}} \left(\frac1x + O\Big(\frac{x^2}{n^2}\Big)\right)dx + \int_1^{n^{1/3}} e^{-y^3/24}\,\frac{dy}{y} = \log n^{2/3} + O(1) + O(1),$$

and thus the expected number of creations is $\sim \frac13 \log n$. Hence, on the average, a random graph process has about $\frac13\log n$ different dynamic unicyclic components during the evolution (and thus about $\frac13\log n$ maximal unicyclic components in the static view). Equivalently, there are at the end about $\frac13\log n$ cycles that at some time have belonged to unicyclic components. (We believe that these random numbers are concentrated about their mean, i.e. that they are $\big(\frac13 + o_p(1)\big)\log n$, but we have not attempted to verify this.)

It can easily be verified by calculations similar to those in [8, Section 5.3] that the average number of unicyclic components in $G(n,n/2)$ or $G(n,1/n)$ is $\sim \frac16\log n$; moreover the actual random number is $\big(\frac16 + o_p(1)\big)\log n$, so it is concentrated about its mean (for example by using Corollary 4 below to show that unicyclic components of size larger than $n^{2/3}/\log\log n$ are too few to matter, and then estimating the variance of the number of smaller unicyclic components). Hence, roughly 1/2 of all (dynamic) unicyclic components that appear during the evolution exist at $m = n/2$ (or at $p = 1/n$). (I'm grateful to a referee for pointing out an error in a previous version of the present paper.)

Remark 5. The number of cycles in complex components in $G(n,n/2)$ (or $G(n,1/n)$) is $O_p(1)$ (see [8, Theorem 5.19]) and thus the total number of cycles is $\big(\frac16+o_p(1)\big)\log n$ too. On the other hand, it is also easily checked that the average number of cycles is $\sim \frac14\log n$. In other words, the number of cycles in $G(n,n/2)$ or $G(n,1/n)$ is concentrated at $\frac{1}{6}\log n$, which is only (asymptotically) 2/3 of the average. (The reason is of course that there is a small probability of having a component with very many cycles.) This is a striking example of the well-known fact that the average can be misleading.

The number of transitions $0 \oplus -1 \to 0$ when the unicyclic components grow is much larger; its expectation is, by Lemmas 1 and 2 as above (now using (6)) and the well-known formula $\sum_1^\infty k^{k-1}e^{-k}/k! = 1$,

$$\begin{aligned}\sum_{k_1,k_2} \beta(0,-1;k_1,k_2) &\sim \sqrt{\frac{\pi}{8}} \sum_{k=1}^{n} \sum_{k_2=1}^{k-1} \frac{k_2^{k_2-1}e^{-k_2}}{k_2!} k^{-1/2} e^{-k^3/24n^2} \\ &\sim \sqrt{\frac{\pi}{8}} \sum_{k=1}^{n} k^{-1/2} e^{-k^3/24n^2} \\ &\sim 2^{-1}3^{-5/6}\pi^{1/2}\Gamma(1/6)n^{1/3}. \end{aligned} \tag{9}$$

Summarizing, we have:

Theorem 1. *The expected number of dynamic* 0*-components is* $\sim \frac{1}{3}\log n$.

The expected number of static 0*-components is* $\sim 2^{-1}3^{-5/6}\pi^{1/2}\Gamma(1/6)n^{1/3} \approx 1.975\, n^{1/3}$. □

Next, we study $V_0 = |\mathcal{V}_0|$, the number of vertices that ever belong to unicyclic components. The expected number of vertices added to $\mathcal{V}_0$ by creations of unicyclic components (transitions $0 \to 1$) is, by Lemmas 1 and 2,

$$\sum_{k=1}^{n} k\alpha(-1;k) \sim \tfrac{1}{2} \sum_{k=1}^{n} e^{-k^3/24n^2} \sim 3^{-2/3}\Gamma(\tfrac{1}{3})n^{2/3},$$

and the expected number of vertices added to already existing unicyclic components (by transitions $0 \oplus -1 \to 0$) is

$$\begin{aligned}\sum_{k_1,k_2} k_2\beta(0,-1;k_1,k_2) &\sim \frac{1}{\sqrt{2\pi}}\sqrt{\frac{\pi}{8}} \sum_{k=1}^{n} \sum_{k_2=1}^{k-1} k_2^{-1/2} k^{-1/2} e^{-k^3/24n^2} \\ &\sim \tfrac{1}{4} \sum_{k=1}^{n} 2e^{-k^3/24n^2} \sim 3^{-2/3}\Gamma(1/3)n^{2/3}. \end{aligned}$$

Both types of transitions thus asymptotically contributes the same number of vertices, and by summing we obtain:

Theorem 2.

$$\mathbb{E}\, V_0 \sim 2 \cdot 3^{-2/3}\Gamma(1/3)n^{2/3} \approx 2.576\, n^{2/3}.$$ □

Since V_0 majorizes both $V_0^{\max}$ and the total order of the unicyclic components at any given stage of the process, we obtain immediate corollaries.

Corollary 3. *For any* $C > 2 \cdot 3^{-2/3}\Gamma(1/3)$ *and large enough* n,

$$\mathbb{E}\, V_0^{\max} \le Cn^{2/3}.$$

Thus the largest unicyclic component during the evolution has order $O_p(n^{2/3})$. $\square$

Corollary 4. *For any given $m = m(n)$ or $p = p(n)$, let $\widetilde{V}_0$ be the total size of the unicyclic components in $G(n,m)$ or $G(n,p)$. Then, for any $C > 2 \cdot 3^{-2/3}\Gamma(1/3)$ and large enough n,*

$$\mathbb{E}\,\widetilde{V}_0 \leq Cn^{2/3}. \qquad \square$$

For $m = n/2$ or $p = 1/n$,

$$\mathbb{E}\,\widetilde{V}_0 \sim \frac{6^{1/3}\Gamma(1/3)}{12} n^{2/3} = 2^{-5/3}3^{-2/3}\Gamma(1/3)n^{2/3}$$

[4, Theorem V.23], that is, by Theorem 2,

$$\mathbb{E}\,\widetilde{V}_0 \sim 2^{-8/3}\,\mathbb{E}\,V_0.$$

Thus, on the average, $2^{-8/3} \approx 0.157$ of the vertices that at some stage of the random graph process belongs to a unicyclic component, do so when $m = n/2$. (The random variables V_0 and $\widetilde{V}_0$ are not concentrated in the above sense, which follows e.g. from the results in [1] or [10, Theorem 9]. In view of the latter results, it seems highly likely that $n^{-2/3}V_0$ has an asymptotic distribution, but unlikely that it has a simple form.)

5. Tree components

We cannot estimate the size of the maximal tree components by the method of the previous section, since each vertex belongs to a tree at the beginning of the process and thus $V_{-1} = |\mathcal{V}_{-1}| = n$. Instead, let W_{-1} be the number of unordered pairs of vertices $\{x, y\}$ such that x and y belong to the same tree component at some stage of the process; equivalently, W_{-1} is the number of pairs $\{x, y\}$ such that x and y belong to the same maximal tree component. Hence, if the maximal tree components are $T_1,\dots,T_N$,

$$W_{-1} = \sum_i \binom{|T_i|}{2}. \tag{10}$$

Theorem 5.

$$\mathbb{E}\,W_{-1} \sim 2^{-1}3^{-5/6}\pi^{1/2}\Gamma(1/6)n^{4/3} \approx 1.975\,n^{4/3}.$$

Proof. The number of pairs $\{x, y\}$ in the same tree is 0 at the beginning of the process, and increases by k_1k_2 each time two tree components of orders k_1 and

k_2 merge. Consequently, again with the symmetry factor 1/2,

$$\begin{aligned}\mathbb{E}\, W_{-1} &= \tfrac{1}{2}\sum_{k_1,k_2} k_1k_2\beta(-1,-1;k_1,k_2)\\ &\sim \frac{1}{2}\sum_{k_1,k_2}\frac{1}{\sqrt{2\pi}}k_1^{-1/2}k_2^{-1/2}k^{-1/2}ne^{-k^3/24n^2}\\ &\sim \frac{\pi}{2\sqrt{2\pi}}n\sum_{k=2}^{n}k^{-1/2}e^{-k^3/24n^2}\\ &\sim \Big(\frac{\pi}{8}\Big)^{1/2}2^{1/2}3^{-5/6}\Gamma(1/6)n^{1+1/3}. \qquad \square\end{aligned}$$

By (10), $\big(V_{-1}^{\max}\big)^2 \le 2W_{-1}+n$. Consequently, we can now estimate the size of the largest tree component.

Corollary 6. *At least for large enough n,*

$$\mathbb{E}\big(V_{-1}^{\max}\big)^2 \le 4n^{4/3}$$

and thus $\mathbb{E}\, V_{-1}^{\max} \le 2n^{2/3}$. In particular, $V_{-1}^{\max} = O_p(n^{2/3})$. $\square$

Remark 6. We can similarly consider, for any $j \ge 2$,

$$W_{-1}^{(j)} = \sum_i \binom{|T_i|}{j},$$

the number of j-tuples of vertices that at some stage belong to the same tree component, cf. (10). The argument above extends to

$$\mathbb{E}\, V_{-1}^{(j)} \sim C_{-1}^{(j)}n^{2j/3}$$

for every $j \ge 2$ and some constants

$$C_{-1}^{(j)} = 2^{j-3}3^{j/3-3/2}\pi^{-1/2}\frac{\Gamma(j/3-1/2)}{\Gamma(j-1)}\sum_{i=1}^{j-1}\frac{\Gamma(i-1/2)\Gamma(j-i-1/2)}{i!\,(j-i)!}.$$

In particular, this implies that all moments of $n^{-2/3}V_{-1}^{\max}$ are bounded. We omit the details.

Moreover, similar results hold for ℓ-components for any $\ell \ge 0$ too (and any $j \ge 1$). Again, we leave the details to the reader; note that $j = 1$ gives V_ℓ treated in the next section.

6. Higher complexities

In this final section, we study creations of components of higher complexity. Let, for $\ell \ge 1$, U'_ℓ be the number of creations by transitions $\ell - 1 \to \ell$, and U''_ℓ the number of creations by transitions $\ell_1 \oplus \ell_2 \to \ell$ for some $\ell_1, \ell_2 \ge 0$ (i.e. mergers of two cyclic components). Furthermore, let $U_\ell = U'_\ell + U''_\ell$ be the total number of creations of ℓ-components, i.e. the number of dynamic ℓ-components in the random graph process. We begin by computing the expectations.

Theorem 7. *Let $\ell \geq 1$. Then, as $n \to \infty$,*

$$\mathbb{E}\, U'_\ell \to u'_\ell = 2^{3\ell/2-1} 3^{\ell/2-1} \rho_{\ell-1} \Gamma(\ell/2),$$

$$\mathbb{E}\, U''_\ell \to u''_\ell = 2^{3\ell/2-1/2} 3^{\ell/2-1} \pi^{-1/2} \frac{\Gamma(\ell/2)}{\Gamma(3\ell/2+1/2)} \cdot \tfrac{1}{2} \sum_{m=0}^{\ell-1} \rho_m \rho_{\ell-m-1} \Gamma(3m/2+1)\Gamma(3(\ell-m-1)/2+1)$$

and thus $\mathbb{E}\, U_\ell \to u'_\ell + u''_\ell$.

Proof. We argue as for the special case $\ell = 1$ in Section 3. First, by Lemmas 1 and 2,

$$\mathbb{E}\, U'_\ell = \sum_{k=1}^{n} \alpha(\ell-1; k) \sim \tfrac{1}{2} \rho_{\ell-1} n^{-\ell} \sum_{k=1}^{n} k^{3\ell/2-1} e^{-k^3/24n^2} \sim 2^{3\ell/2-1} 3^{\ell/2-1} \rho_{\ell-1} \Gamma(\ell/2).$$

Secondly,

$$\mathbb{E}\, U''_\ell = \tfrac{1}{2} \sum_{m=0}^{\ell-1} \sum_{k_1,k_2} \beta(m, \ell-m-1; k_1, k_2),$$

which by a similar computation yields the second formula. □

While the limits in Theorem 7 can be evaluated for any given ℓ, they are not very illuminating as they stand, but we can connect them to well-known properties of the evolution of random graphs.

Note first that, as a consequence of [7, Theorem 16], each U_ℓ converges in distribution to some random variable Y_ℓ, whose support is $\{1, 2, \dots\}$ for $\ell = 1$ and $\{0, 1, 2, \dots\}$ for $\ell \geq 2$. Furthermore, usually $U_\ell = 1$ (although thus both $U_\ell \geq 2$ and, for $\ell \geq 2$, $U_\ell = 0$ are possible with positive limit probabilities). Indeed, see Remark 4, with probability approching $5\pi/18$, $U_\ell = 1$ for all $\ell = 1, \dots, \binom{n}{2} - n$ simultaneously. Moreover, the tendency for U_ℓ to equal 1 becomes more pronounced as ℓ increases, and we have the following result.

Lemma 3. *For every $\varepsilon > 0$ there is an L such that, for any n, the probability that $U_\ell = 1$ for all $\ell = L, L+1, \dots, \binom{n}{2} - n$ is at least $1 - \varepsilon$.*

Sketch of proof. This is an easy consequence of [7, Theorem 16]. Alternatively (and simpler), it can be proved as follows; see [8, Theorem V.10] for more details (in a slightly different case). First, we can find (by methods close to those used in this paper) A such that the probability that a new complex component is created after $n/2 + An^{2/3}$ edges have been added is at most $\varepsilon/3$. Then we find B such that the probability that there exist two complex components in $G(n, \lfloor n/2 + An^{2/3} \rfloor)$ that are not joined by at least one of the following $Bn^{2/3}$ edges is at most $\varepsilon/3$. Finally there is an L such that the probability that there is a component in $G(n, \lfloor n/2 + (A+B)n^{2/3} \rfloor)$ with complexity at least L is at most $\varepsilon/3$. If none of these exceptional events occurs, then $U_\ell = 1$ for $\ell \geq L$. □

We next observe that U_ℓ is close to 1 in mean too, using an argument similar to dominated convergence. Since each complex component must begin as one or several ℓ-components, we have the majorization $U_\ell \le U_1$ for every $\ell \ge 1$, and thus, by the Cauchy–Schwarz inequality and the fact that $\mathbb{E}\, U_1^2 \le C$ for some constant C [6],

$$|\,\mathbb{E}\, U_\ell - 1| \le \bigl(\mathbb{E}(U_1+1)^2\,\mathbb{P}(U_\ell \ne 1)\bigr)^{1/2} \le C'\,\mathbb{P}(U_\ell \ne 1)^{1/2}.$$

By Lemma 3, this implies that $\mathbb{E}\, U_\ell$ is close to 1 for large ℓ (for all n), and thus, by Theorem 7,

$$u'_\ell + u''_\ell \to 1 \qquad \text{as } \ell \to \infty. \tag{11}$$

Indeed, (11) can also be derived from the well-known asymptotics of Wright's constants ([2, 3]; see also [7, Section 8])

$$2^{5\ell/2-1/2}3^{-\ell}\pi^{-1/2}\frac{\Gamma(3\ell/2)}{\Gamma(\ell)}\rho_\ell \to \frac{1}{2\pi} \qquad \text{as } \ell \to \infty. \tag{12}$$

Simple calculations using Stirling's formula show that (12) is equivalent to $u'_\ell \to 1$ as $\ell \to \infty$, and that it implies $u''_\ell \to 0$ as $\ell \to \infty$. (More precisely, $u''_\ell = O(\ell^{-1})$; the main contribution for large ℓ coming from the terms with $m = 0$ and $m = \ell - 1$, which together yield $u'_\ell/(3\ell-1)$ for $\ell \ge 2$.)

Conversely, our result (11) yields $u'_\ell \le C$ for some constant C. The calculations just mentioned now show first that the left hand side of (12) is bounded and then that $u''_\ell = O(\ell^{-1})$. Consequently (11) yields $u'_\ell \to 1$, and thus (by the same calculations again) we obtain a new proof of (12).

Remark 7. The random graph process eventually has a unique complex component, and by Lemma 3 this usually happens before the complexity gets very large. This complex component increases its complexity by receiving new edges added to it and by merging with unicyclic components. The facts that $u'_\ell \to 1$ and $u''_\ell \sim u'_\ell/(3\ell-1) \sim 1/3\ell$ as $\ell \to \infty$ show that usually the first case occurs, although sometimes a unicyclic component is swallowed.

We may also compute the expected number of transitions $\ell \oplus -1 \to \ell$ when the ℓ-components grow; this is, using Lemmas 1 and 2 as in (9),

$$\sum_{k_1,k_2} \beta(\ell,-1;k_1,k_2) \sim \rho_\ell 2^{3\ell/2+1/2}3^{\ell/2-5/6}\Gamma(\ell/2+1/6)n^{1/3}.$$

This and Theorem 7 yields an extension of the second part of Theorem 1.

Theorem 8. *Let $\ell \ge 0$. Then the expected number of static ℓ-components is asymptotically $s_\ell n^{1/3}$, where*

$$s_\ell = 2^{3\ell/2+1/2}3^{\ell/2-5/6}\Gamma(\ell/2+1/6)\rho_\ell.$$

For large ℓ we can estimate s_ℓ using Stirling's formula and (12), or by the identity $s_\ell = 3^{-1/3}\bigl(\Gamma(\ell/2+1/6)/\Gamma(\ell/2+1/2)\bigr)u'_{\ell+1}$; this yields

$$s_\ell \sim (3\ell/2)^{-1/3} \qquad \text{as } \ell \to \infty.$$

Finally we consider the number of vertices in ℓ-components, generalizing the results for $\ell = 0$ in Section 4.

Theorem 9. *For any $\ell \geq 0$,*

$$\mathbb{E}\, V_\ell \sim v_\ell n^{2/3}$$

where

$$v_\ell = 2^{3\ell/2} 3^{\ell/2-2/3} \Gamma(\ell/2+1/3) \cdot \biggl(\rho_{\ell-1} + \sum_{m=0}^{\ell} \Bigl(\frac{2}{\pi}\Bigr)^{1/2} \rho_m \rho_{\ell-m-1} \frac{\Gamma(3m/2+1)\Gamma(3\ell/2-3m/2+1/2)}{\Gamma(3\ell/2+3/2)}\biggr). \tag{13}$$

In particular, $V_\ell = O_p(n^{2/3})$.

Proof. The expected number of vertices contributed by transitions $\ell-1 \to \ell$ is $\sum_k k\alpha(\ell-1;k)$, while the number of vertices contributed by a j-component ($-1 \leq j \leq \ell-1$) merging with an $(\ell-j-1)$-component is $\sum_{k_1,k_2} k_1\beta(j,\ell-j-1;k_1,k_2)$. Thus, with $m = \ell-j-1$,

$$\mathbb{E}\, V_\ell = \sum_k k\alpha(\ell-1;k) + \sum_{m=0}^{\ell} \sum_{k_1,k_2} k_1 \beta(\ell-m-1, m; k_1, k_2),$$

which yields the result using (4), (5) and Lemma 2 as above. $\square$

Corollary 10. *For any $\ell \geq 0$, $\mathbb{E}\, V_\ell^{\max} = O(n^{2/3})$, and thus $V_\ell^{\max} = O_p(n^{2/3})$.* $\square$

The proof of Theorem 9 shows that the different terms in (13) correspond to the different ways that an ℓ-component can be created or grow (the latter for $m = \ell$).

Example. For $\ell = 1$, Theorem 9 yields

$$\mathbb{E}\, V_1 \sim \tfrac{25}{16} 3^{-1/6} \pi^{1/2} \Gamma(5/6) n^{2/3} \approx 2.603\, n^{2/3}.$$

Of this asymptotical value, 16/25 comes from transitions $0 \to 1$, 4/25 from transitions $0 \oplus 0 \to 1$ and 5/25 from transitions $1 \oplus -1 \to 1$ (i.e. from growth of 1-components).

For large ℓ, straightforward estimates using Stirling's formula and (12) (or $u'_\ell \to 1$ and $u''_\ell \to 0$) show that the first term in (13) dominates, and we have

$$v_\ell \sim 2^{3\ell/2} 3^{\ell/2-2/3} \Gamma(\ell/2+1/3) \rho_{\ell-1} \sim (12\ell)^{1/3} \qquad \text{as } \ell \to \infty. \tag{14}$$

Remark 8. We have in this paper studied ℓ-components for ℓ fixed, but one can also let ℓ increase with n. It follows automatically from Theorem 9 and (14) that if $\ell = \ell(n) \to \infty$ slowly enough, then $\mathbb{E}\, V_\ell \sim (12\ell)^{1/3} n^{2/3}$; however, we do not know for which functions $\ell(n)$ this holds.

References

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

[2] G.N. Bagaev & E.F. Dmitriev, Enumeration of connected labeled bipartite graphs. (Russian.) *Dokl. Akad. Nauk BSSR* **28** (1984), 1061–1063, 1148.

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

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

[5] R.L. Graham, D.E. Knuth & O. Patashnik, *Concrete Mathematics.* 2nd ed., Addison-Wesley, Reading, Mass., 1994.

[6] S. Janson, Multicyclic components in a random graph process. *Rand. Struct. Alg.* **4** (1993), 71–84.

[7] S. Janson, D.E. Knuth, T. Łuczak & B. Pittel, The birth of the giant component. *Rand. Struct. Alg.* **4** (1993), 233–358.

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

[9] T. Łuczak, Component behavior near the critical point of the random graph process. *Rand. Struct. Alg.* **1** (1990), 287–310.

[10] T. Łuczak, The phase transition in a random graph. In *Combinatorics, Paul Erdős is Eighty, Vol. 2*, eds. D. Miklós, V.T. Sós & T. Szőnyi, Bolyai Soc. Math. Stud. 2, J. Bolyai Math. Soc., Budapest 1996, pp. 399–422.

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

[12] J. Spencer, Enumerating graphs and Brownian motion. *Commun. Pure Appl. Math.* **50** (1997), 291–294.

[13] E.M. Wright, The number of connected sparsely edged graphs. *J. Graph Th.* **1** (1977), 317–330.

Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se