LIMIT THEOREM ON THE CREATION OF MULTICYCLIC COMPONENTS

ANNE-ELISABETH BAERT AND VLADY RAVELOMANANA

ABSTRACT. In this paper, we consider $(l+1)$-components, defined as connected components with $l+1$ edges more than vertices, of an evolving graph. We prove that *almost all* $(l+1)$-components, whose *last added edge* during the random graph process, forms a bridge (or a cut edge) between a p-component and an $(l-p)$-component are built by linking a unicyclic component to an l-component. We then obtain a limit theorem for the number of creations of $(l+1)$-components during the evolution of a random graph process. We also show that the random variable defined by the number of creations of l-components is concentrated about their means. These results are proved asymptotically for connected components with k vertices and $k+l+1$ edges whenever $k, l \to \infty$ but $l = o(k^{1/3})$.

RÉSUMÉ. Dans cet article, nous considérons les $(l+1)$-composantes, définies comme les composantes connexes ayant $l+1$ arêtes de plus que de sommets, dans un graphe aléatoire. Nous montrons que *presque toutes* les $(l+1)$-composantes dont *la dernière arête*, qui a été rajoutée lors de l'évolution du graphe aléatoire, forme un pont (ou un isthme) entre une p-composante et une $(l-p)$-composante sont construites en joignant une composante unicyclique à une l-composante. Un théorème limite est obtenu pour le nombre de créations de $(l+1)$-composantes pendant l'évolution du graphe. Nous montrons alors que la variable aléatoire définie par le nombre de créations de $(l+1)$-composantes est concentrée autour de sa moyenne. Nos résultats sont obtenus asymptotiquement pour toutes les composantes connexes avec k sommets et $k+l+1$ arêtes quand $k, l \to \infty$ et telle que $l = o(k^{1/3})$.

1. INTRODUCTION

Following Erdös and Rényi pioneering works around 1960 [8, 9], random graphs model have been the subject of intense studies for four decades. Topics on random graphs provide a large and particularly active body of research. We refer to the excellent books of Bollobas [5] and of Janson *et al.* [17] for precision on these subjects.

We consider here labelled graphs, i.e., graphs with labelled vertices and undirected edges without self-loops or multiple edges. These graphs have n vertices and let $V = \{1, 2, \ldots, n\}$ be the vertex set. The set of all such graphs is denoted by $\mathcal{G}^n$ and, a random graph is defined by a pair $(\mathcal{G}^n, P)$ where P is a probability distribution over $\mathcal{G}^n$. Let us introduce the three most encountered models of random graphs : $\{\mathbb{G}(n,M)\}_{0\le M\le \binom{n}{2}}$, $\{\mathbb{G}(n,p)\}_{0\le p\le 1}$ and $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$. The first model consists of all graphs with vertex set $V = \{1, 2, \ldots, n\}$ having M edges, in which one can randomly pick a graph with the same probability. Thus, with $N = \binom{n}{2}$, we have $0 \le M \le N$ and $\mathbb{G}(n,M)$ has $\binom{N}{M}$ elements and each element occurs with probability $\binom{N}{M}^{-1}$. In $\{\mathbb{G}(n,p)\}_{0\le p\le 1}$, we have $0 \le p \le 1$ and the model consists of all graphs with the same vertex set $V = \{1, 2, \ldots, n\}$ in which each of the N edges is drawn independently with probability p. The third model, $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$, may be constructed by letting each edge e, chosen amongst the N possible edges, appear at random time T_e, where

Key words and phrases. Random graphs; probabilistic/analytic combinatorics; asymptotic enumeration; labelled graphs; Wright's coefficients.

T_e are independent random variables uniformly distributed on $[0,1]$. $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$ is constructed with all edge e such that $T_e \le t$. The main difference between $\mathbb{G}(n,M)$ and $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$ is that in $\mathbb{G}(n,M)$, edges are added at fixed (slotted) times $1,2,\ldots,N$ so at any time T we obtain a random graph with n vertices and T edges, whereas in $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$ the edges are added at random times. At time $t=0$, we have a graph with n vertices and 0 edge, and as the time advances all edges e with r.v. T_e such that $T_e \le t$ (where t is the current time), are added to the graph until t reaches 1 in which case, one obtain the complete graph K_n.

We define the *excess* or the *complexity* of a connected graph to be the difference between its number of edges and its number of vertices. Throughout this paper, as the random graph process proceeds, we will often *fix* and study an arbitrary chosen connected component built with $k \le n$ vertices (where n is the total number of vertices) in the graph. For $l \ge -1$, a $(k, k+l)$ connected graph is one having k vertices and $k+l$ edges, thus its excess is l and we simply called it an l-*component*. In the models presented here, a random graph process begins with a set of n isolated vertices. Then, as evolution proceeds, edges are added at random (drawn without replacement) and at first, all components created are trees ((-1)-components), later 0-components (also called *unicyclic* components) will appear and eventually the first l-components are created, with $l > 0$. Usually, l-components are called *complex* whenever $l > 0$. As more edges are added, a complex component gradually swallowed up some other "simpler" components, and it is worth-noting that with probability non-zero, at least two components can co-exist as the random graphs evolves [14], [16].

In this paper, we consider the *continuous time* random graph model $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$, and we will study the creation of $(l+1)$-components ($l \ge 0$). We can observe that there are two manners to create a new $(l+1)$-component during the random graph process :

- either by adding an edge to an l-component,
- or by joining with the last added edge a p-component to an $(l-p)$-component, with $p \ge 0$.

We study the random variable $X_n^{(l)}$, defined as the number of creations of $(l+1)$-components during the evolution of the random graph. Following the author of [14], we denote, respectively, by $Y_n^{(l)}$ and $Z_n^{(l)}$ the number of $(l+1)$-components created by the two ways described above. More precisely, $Y_n^{(l)}$ equals the number of edges that are added inside an l-component creating an $(l+1)$-component, and $Z_n^{(l)}$ is the number of edges (bridges) added between a p-component and a $(l-p)$-component, for all $0 \le p$ during the evolution of the graph. Thus, we have $X_n^{(l)} = Y_n^{(l)} + Z_n^{(l)}$. Furthermore, $Z_n^{(l)}$ can be decomposed into $Z_{n,0}^{(l)}$, $Z_{n,1}^{(l)}$, $\cdots$ where $Z_{n,i}^{(l)}$, $i \le \lceil l/2 \rceil$ denotes the number of edges added between an i-component and an $(l-i)$-component for fixed $i \in [0, \lceil l/2 \rceil]$.

1.1. **Previous works.** In a former paper, Janson [14] obtained limit theorems for the number of *complex components*, i.e., components with more than one cycle, created during the evolution of the graph. In particular, Janson computed the probability that the process never contains more than one complex component is approximately 0.87 (as the number of vertices n tends to infinity). Thus, at least two complex components can co-exist in the random graph and there is not a zero-one law for this process. With the notations of our paper, Janson obtained limit laws for $X_n^{(l)}, Y_n^{(l)}$ and $Z_n^{(l)}$ for $l=1$ (see for instance [14] for precise statements of his results). Using enumerative and analytical methods, Janson, Knuth, Pittel and Luczak [16] obtained also the exact value $5\pi/18 = 0.872\cdots$ for the limit described above. In [3, 4], Bender *et al.* studied several properties of labelled graphs.

Working in the probability space of connected components, they obtained the asymptotic probability for a random chosen edge to be a bridge (see for instance [4, Section 5]).

1.2. **Our results.** We follow the probabilistic methods initiated by Janson and combine them with the enumerative/analytic methods to study the moments of the r.v. $X_n^{(l)}, Y_n^{(l)}$ and $Z_n^{(l)}$ described above, for values of l and n s.t. $l, n \to \infty$ but $l = o(n^{1/3})$. Precisely, to obtain the results presented here, methods of the probabilistic random graph model $\{\mathbb{G}(n,t)\}_{0\le t\le 1}$, studied in [14, 15], are combined with asymptotic enumeration methods, developed by Wright in [24] and by Bender, Canfield and McKay in [3, 4]. We will also use the analytical tools associated to the generating functions of Cayley's rooted trees [7], $T(z)$, which plays a central key role in the enumerative point of view of the general theory of random graphs. In particular, these methods were developed by Flajolet, Knuth and Pittel in [10] and extended in the "giant paper" of Janson, Knuth, Luczak and Pittel [16].

In fact, problems considered here are in essence combinatorial, and the following results are obtained along these lines.

• We prove that *almost all* $(l+1)$-components whose *last added edge* forms a bridge (or a cut edge) between a p-component and an $(l-p)$-component, for $0 \le p \le l$, are built by linking a unicyclic component to an l-component. In fact, we have

Theorem 1.1. *As $k, l \to \infty$ but $l = o(k^{1/3})$, the number of ways, $c'(k, k+l+1)$, to build an $(l+1)$-component of order k with a distinguished cut edge between a p-component and a $(l-p)$-component, $p \ge 0$, satisfies*

$$\begin{aligned} c'(k,k+l+1) &= \frac{1}{2}\sum_{p=0}^{l}\sum_{t=1}^{k-1}\binom{k}{t} t(k-t)\, c(t,t+p)\ c(k-t,k-t+l-p) \\ &= \sqrt{\frac{\pi}{12}}\,\frac{d}{l}\left(e/12l\right)^{l/2} k^{k+3l/2+3/2}\left(1+O(1/l)+O(\sqrt{l^3/k})\right). \qquad (1) \end{aligned}$$

Note here that our results differ from those in [4], since we are interested in edges whose additions during the random graph process, increase the complexity of some connected components (whereas in [4], all edges in a given connected component are considered with the same probability).

• We then use these results and obtain a limit theorem for the number of $(l+1)$-components, for $l \ge 0$, that are created during the evolution of the graph. More precisely, we have the following

Theorem 1.2. *Provided that the newly created $(k, k+l+1)$ component satisfies $l = o(k^{1/3})$, $l > 0$ and with $X_n^{(l)}$ defined as above, we have $X_n^{(l)} \xrightarrow{d} X_\infty^{(l)}$, as $n \to \infty$, where X_∞ is a positive integer valued random variable with the factorial moment*

$$\mathbb{E}(X_\infty)_m = \left(\frac{1+3l}{3l}\right)^m \left(\sqrt{\frac{3}{8}}\,\frac{d_l\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right)^m a_m^{(l)}\,, m \ge 0\,, \qquad (2)$$

where $d_l = \frac{1}{2\pi}(1+O(1/l))$ and

$$a_m^{(l)} = \int_0^\infty\cdots\int_0^\infty\int_{-\infty}^\infty\cdots\int_{-\infty}^\infty\left(\prod_{i=1}^m x_i^{3l/2+1}\right) \times \exp\Big(-\frac{1}{24}\Big(\sum_{i=1}^m x_i\Big)^3 - \frac{1}{2}\sum_{i=1}^m x_i {y_i}^2\Big)$$

$$\times\ \exp\Big(-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=i+1}^{j=m} x_i x_j |y_i-y_j|\Big)\ dx_1dx_2\cdots dx_m\ dy_1dy_2\cdots dy_m\ <\infty\ . \tag{3}$$

All moments of X_n converge to the corresponding moments of X_∞. Moreover, (Y_n, Z_n) converges in distribution, with all moments, to a two dimensional variable (Y_∞, Z_∞) with $Y_\infty + Z_\infty = X_\infty$,

$$\mathcal{L}\left(Y_\infty | X_\infty = r\right) = Bi\left(r, \frac{1}{3l+1}\right)\ \text{and}\ \mathcal{L}\left(Z_\infty | X_\infty = r\right) = Bi\left(r, \frac{3l}{3l+1}\right)\ .$$

The mixed factorial moments are given by

$$\mathbb{E}(Y_\infty)_p (Z_\infty)_q = \left(\frac{1}{3l}\right)^q \left(\sqrt{\frac{3}{8}}\frac{d_l\ \exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right)^{p+q} a^{(l)}_{p+q}\ ,\ p,\,q\geq 0\,. \tag{4}$$

2. Joining two complex components

In order to investigate $(Z_n^{(l)})_m$, i.e., the number of m-tuples of edges that are added between two complex components to build an $(l+1)$-component, we will use tools from enumerative/analytic methods. The enumeration of labelled trees goes back to Cayley. Denote by $T(z)$ the well-known exponential generating function (EGF) of Cayley's rooted trees [7], we have

$$T(z) = z\exp\left(T(z)\right) = \sum_{n=1}^{\infty}\frac{n^{n-1}z^n}{n!}\,, \tag{5}$$

where the variable z is associated to the labelled vertices. Next, Rényi [21] found the EGF W_0 of unicyclic graphs.

$$W_0(z) = -\frac{1}{2}\ln\left(1-T(z)\right) - \frac{T(z)}{2} - \frac{T(z)^2}{4}\,. \tag{6}$$

More generally, Wright [23] found a recurrence formula satisfied by the EGFs of l-components. Denote by $W_l(w,z)$ the bivariate EGF of l-components where the variable w marks the number of edges and the variable z the vertices. Thus, if $c(n, n+l)$ is the number of $(n, n+l)$ connected graphs with n vertices, we can write

$$W_l(w,z) = \sum_n c(n,n+l)w^{n+l}\frac{z^n}{n!} \tag{7}$$

and Wright's recurrence formula [23] can be stated as follow :

$$\vartheta_w W_{l+1} = w\Big(\frac{{\vartheta_z}^2-\vartheta_z}{2} - \vartheta_w\Big)W_l + \frac{w}{2}\left(\sum_{p=-1}^{l+1}(\vartheta_z W_p)(\vartheta_z W_{l-p})\right)\,, \tag{8}$$

where we denote by ϑ_w, resp. ϑ_z, the differential operator $w\frac{\partial}{\partial w}$, resp. $z\frac{\partial}{\partial z}$. Thus, the operator ϑ_w corresponds to marking an edge present in a graph. Similarly, ϑ_z corresponds to marking a vertex. The combinatorial pointing operator reflects the distinction of an object among all the others. For the use of pointing and marking, we refer to [12] and for general techniques concerning graphical enumerations we refer to [13]. All these EGFs are given and explained in details in [16]. In terms of coefficients, (8) reads

$$(9) \qquad \begin{aligned} (k+l+1)\, c(k,k+l+1) &= \left(\binom{k}{2} - k - l\right)\, c(k,k+l) \\ &+ \frac{1}{2}\sum_{t=1}^{k-1}\sum_{p=-1}^{l+1}\binom{k}{t} t(k-t)\, c(t,t+p)\; c(k-t,k-t+l-p)\,. \end{aligned}$$

Starting with the differential equation (8), Wright [23, 24] proved that each W_l can be written as :

$$(10) \qquad W_l(z) = \frac{b_l}{(1-T(z))^{3l}} - \frac{c_l}{(1-T(z))^{3l-1}} + \sum_{2\le s\le 3l-2} \frac{\omega_{l,s}}{(1-T(z))^{s}}\,,$$

where the coefficients (b_l) and (c_l) are rationals and more importantly, the summation is *finite*. The $(b_l)_{l\ge1}$ are called the Wright's constants of first order (also called Wright-Louchard-Takács constants, see [22]); $b_1 = \frac{5}{24}$ and for $l \ge 1$, b_l is defined recursively by

$$(11) \qquad 2(l+1)b_{l+1} = 3l(l+1)b_l + 3\sum_{p=1}^{l-1} t(l-p)b_p b_{l-p}\,.$$

Note that the sequence (c_l) satisfied also the following:

$$(12) \qquad 2(3l+2)c_{l+1} \;=\; 8(l+1)b_{l+1} + 3lb_l + (3l+2)(3l-1)c_l + 6\sum_{p=1}^{l-1} p(3l-3p-1)b_t c_{l-p}\,.$$

To study the asymptotic behaviour of the coefficients $c(k,k+l)$, Wright [24] established that:

$$(13) \qquad \frac{b_l}{(1-T(z))^{3l}} - \frac{c_l}{(1-T(z))^{3l-1}} \preceq W_l(z) \preceq \frac{b_l}{(1-T(z))^{3l}}\,,$$

which we shall call *Wright's inequalities*[1].

We are interested in the number of creation of $(l+1)$-components. In this Section, we will study edges that are added between a p-component and a $(l-p)$-component, with $p \ge 0$. Thus, we have to investigate the number of manners to build a component with a distinguished cut edge. The Theorem 1.1 gives an estimate of the number of such combinatorial structures. It be proved later since its proof involves the decomposition of the Wright's EGFs by means of inverse powers of $(1-T(z))$. In fact, Knuth and Pittel [18] studied combinatorially and analytically the polynomial $t_n(y)$ defined as follows

$$(14) \qquad t_n(y) = n!\,[z^n]\,\frac{1}{\big(1-T(z)\big)^{y}}\,,$$

which they call *tree polynomial*. The two authors observed that the analysis of these polynomials can be used to study random graphs analytically as shown in [10, 16]. For our purpose, a very

[1]Remark that if $A(z)$ and $B(z)$ are two formal power series, the notation $A(z) \preceq B(z)$ means that $\forall n$, $[z^n]\,A(z) \le [z^n]\,B(z)$.

similar formula can be defined :

$$t_{a,n}(y) = n!\,[z^n]\,\frac{T(z)^a}{\big(1-T(z)\big)^y}\,. \tag{15}$$

The lemma below is an application of the saddle point method [6, 11] to study the asymptotic behaviour of the coefficients $t_{a,n}(m) = n!\,[z^n]\,T(z)^a(1-T(z))^{-m(n)}$ as m, n tend to infinity but $m \equiv m(n) = o(n)$.

Lemma 2.1. *Let $\rho \equiv \rho(n)$ such that $\rho \to 0$ as $n \to \infty$ but $\frac{\rho\, n}{\ln n^2} \to \infty$, and let a and β be fixed numbers. Then, $t_{a,n}(\rho\, n+\beta)$ defined in (15) satisfies*

$$t_{a,n}(\rho\, n+\beta) = \frac{n!}{2\sqrt{\pi n}}\,\frac{\exp\,(n u_0)(1-u_0)^{(1-\beta)}}{{u_0}^n(1-u_0)^{\rho\, n}}\left(1+O\Big(\sqrt{\rho}\Big)+O\Big(\frac{\ln n}{\sqrt{\rho\, n}}\Big)\right) \tag{16}$$

where $u_0 = 1+\frac{\rho}{2}-\sqrt{\rho(1+\frac{\rho}{4})}$.

The proof is based upon the saddle-point method and is omitted in this extended abstract (due to place limitation).

Proof of Theorem 1.1. In term of EGFs, $c'(k, k+l+1)$ represents the coefficient

$$c'(k, k+l+1) = \frac{k!}{2}\,\left[z^k\right]\sum_{p=0}^{l}\Big(\vartheta_z W_p(z)\Big)\Big(\vartheta_z W_{l-p}(z)\Big)\,. \tag{17}$$

Applying Wright's inequalities, i.e. (13), to the EGF described in (17) above yields

$$L_l(z) \preceq \sum_k c'(k, k+l+1)\frac{z^k}{k!} \preceq R_l(z)\,,\ \ (l>0)\,, \tag{18}$$

where

$$R_l(z) = \frac{9}{2}\sum_{p=1}^{l-1}\frac{p(l-p)b_p b_{l-p} T(z)^2}{\big(1-T(z)\big)^{3l+4}} + \frac{3 l b_l T(z)^5}{2(1-T(z))^{3l+4}} \tag{19}$$

and

$$L_l(z) = R_l(z) - \left(\sum_{p=1}^{l-1}\frac{3(3p-1)(l-p)b_{l-p}c_p T(z)^2}{(1-T(z))^{3l+3}} + \frac{(3l-1)c_l T(z)^5}{2(1-T(z))^{3l+3}}\right)\,. \tag{20}$$

(We used $\vartheta_z T(z) = T(z)/(1-T(z))$.) Our aim is then to show that the difference between the coefficients of the right and left parts of (18), viz. $k!\,\left[z^k\right](R_l(z)-L_l(z))$ can be neglected in comparison to $k!\,\left[z^k\right]R_l(z)$ for $l = o(k^{1/3})$. To that extent, we use lemma 2.1, and several properties verified by the sequence (b_l). More precisely, lemma 2.1 tells us that in $R_l(z)$, the coefficients of z^k of $T(z)^2/(1-T(z))^{3l+4}$ and $T(z)^5/(1-T(z))^{3l+4}$ in (19) have the same order of magnitude for $l = o(k^{1/3})$. Next, using the definition of Wright's coefficients (11), we find

$$\frac{9}{2}\sum_{p=1}^{l-1}p(l-p)b_p b_{l-p} + \frac{3}{2}l b_l = 3(l+1)(b_{l+1}-\frac{3}{2}l b_l) + \frac{3}{2}l b_l\,. \tag{21}$$

We then have $b_{l+1} - \frac{3}{2} l b_l = (\frac{3}{2})^{l+1} l! (d_{l+1} - d_l)$ where we used $b_l = (3/2)^l (l-1)! d_l$ as studied in [24, eq. (1.4)] and in [16]. From the proof given by Meertens in [3, lemma 3.4], we have $0 < d_{l+1} - d_l = O(1/l^2)$. So,

$$\frac{9}{2} \sum_{p=1}^{l-1} p(l-p)\, b_p\, b_{l-p} + \frac{3}{2} l\, b_l = \left(\frac{3}{2}\right)^{l+1} l!\, d_l \Big(1 + O(1/l)\Big)\,. \tag{22}$$

On the other hand, the definition (12) of the sequence (c_l) tells us that the summation in (20) satisfies $\sum 3(3p-1)(l-p) b_{l-p} c_p = O(l c_l)$ and we know from [24] that $c_l = O(l b_l)$. Finally, lemma 2.1 suggests us that we have to find values of $l \equiv l(k)$ for which the coefficients of the difference $R_l - L_l$ satisfy $\left[z^k\right] \big(R_l(z) - L_l(z)\big) \ll \left[z^k\right] \big(R_l(z)\big)$. It comes $l = o(k^{1/3})$ which is the same range as in [24] and in [20] for connected graphs without prefixed (forbidden) configurations, the error terms being of order $O(1/l) + O(\sqrt{l^3/k})$. After a bit of algebra, we find (replacing $\rho = 3l/k$ in the saddle point u_0)

$$\begin{aligned} \frac{3}{2} l\, b_l\, t_{5,3l+4} &= \frac{3}{2} l\, b_l \frac{k^{k+3/2l+3/2}}{\sqrt{2}(3l)^{3/2l+3/2}} \exp{(3l/2)} \Big(1 + O(\sqrt{l^3/k})\Big) \\ &= \sqrt{\frac{\pi}{12}} \frac{d_l}{l} \left(\frac{e}{12l}\right)^{l/2} k^{k+3l/2+3/2} \Big(1 + O(1/l) + O(\sqrt{l^3/k})\Big)\,, \end{aligned} \tag{23}$$

which completes the proof of Theorem 1.1. □

Wright showed that the EGFs of all multicyclic components can be expressed in terms of the EGF of Cayley. In order to count the number of ways to label a complex component, one can repeatedly prune it by deleting recursively any vertex of degree 1. The graph obtained after removing all vertices of degree 1 is called a *smooth graph.*

Remark 2.2. *Theorem 1.1 tells us that asymptotically almost all $(l+1)$-components whose situation after smoothing contains a cut edge are built by linking a unicyclic component to another complex component. In fact, (1) reflects simply*

$$c'(k, k+l+1) \sim k! \left[z^k\right] \big(\vartheta_z W_0(z)\big) \big(\vartheta_z W_l(z)\big)\,, \ \textit{for } l = o(k^{1/3})\,. \tag{24}$$

3. Proof of Theorem 1.2

3.1. **Adding edges to l-components: moments computation.** As already said, proves given here follow (humbly) the works of Janson in [14, 15]. However, the main difference comes from the fact that our parameter, representing the complexity of the sparse components l, is no more fixed as in [15]. Turning to higher moments, we observe that $\mathbb{E}(Y_n^{(l)})_m$ is the number of m-tuples of edges that are added to a i−th l-component of order k_i during the evolution of the random graph process.

There are $\binom{n}{k_1 \ldots k_m} \prod_i c(k_i, k_i+l)$ manners to choose an l-component having respectively $k_1, \ldots, k_m$ vertices. There are $\prod_i \left(\binom{k_i}{2} - k_i - l\right)$ ways to choose the new edge. Furthermore, the probability that such possible component is one of $\{\mathbb{G}(n,t)\}_{0 \le t \le 1}$ is

$$\prod_i t_i^{k_i+l} (1-t_i)^{(n - \sum k_j) k_i + \binom{k_i}{2} - k_i - l} \prod_{i<j} (1 - t_i \vee t_j)^{k_i k_j}$$

and with the conditional probability $\frac{dt_i}{(1-t_i)}$ that a given edge is added during the interval (t_i, t_i+dt_i) and not earlier, integrating over all times, i.e. $t_i \in [0,\,1]$ and summing over k_i, we obtain

$$\begin{aligned}\mathbb{E}(Y_n^{(l)})_m \;&=\; \sum_{k_1=1}^{n}\cdots\sum_{k_m=1}^{n}\int_0^1\cdots\int_0^1 (n)_{k_*}\prod_i \frac{c(k_i,k_i+l)}{k_i!}\left(\binom{k_i}{2}-k_i-l\right)\\ &\qquad t_i^{k_i+l}(1-t_i)^{(n-k_i)k_j+\binom{k_i}{2}-k_i-l-1}\prod_{i<j}(1-t_i\vee t_j)^{k_ik_j}\,dt_1\ldots dt_m \end{aligned}\tag{25}$$

where $k_* = \sum_i k_i$. We remark here that $c(k_i, k_i+l) = 0$ for $k_i = 1, 2, \cdots \lceil(3+\sqrt{9+8l})/2\rceil - 1$. Rewriting the integrand in (25) as a function of k_i and t_i, viz. $\varphi_n(k_i,t_i) \equiv \varphi_n(k_1,\cdots,k_m,t_1,\cdots,t_m)$, with $\varphi_n(k_i,t_i) = 0$ if $\exists j \in [1,m]$ s.t. $k_j \le \lceil(3+\sqrt{9+8l})/2\rceil - 1$ or $k_j > n$ or $t_j \notin (0,1)$ and substituting $k_i = \lceil x_i n^{2/3}\rceil$ and $t_i = n^{-1} + u_i n^{-4/3}$, we have

$$\mathbb{E}(Y_n{}^{(l)})_m \;=\; \int_0^\infty\cdots\int_0^\infty\int_{-\infty}^\infty\cdots\int_{-\infty}^\infty \Psi_n^{(m)}(x_i,u_i)\,du_i\cdots du_m\,dx_i\cdots dx_m\,,\tag{26}$$

where

$$\Psi_n^{(m)}(x_i,u_i) \equiv \Psi_n^{(m)}(x_1,\cdots,x_m,\,u_1,\cdots,u_m) = \frac{\varphi_n\left(\lceil x_i n^{2/3}\rceil, \frac{1}{n}+\frac{u_i}{n^{4/3}}\right)}{n^{2m/3}}\,.\tag{27}$$

We shall now investigate the integrand of in (25). For this purpose, we consider each term of the products in this integrand and we assume that $x_i n^{2/3}$ are integers. In the following, for each factor, we use the substitutions $k_i = x_i n^{2/3}$ and $t_i = n^{-1} + u_i n^{-4/3}$ as done above. We then have (with $x_* = \sum x_i$)

$$(n)_{k_*} = n^{k_*}\exp\left(-\frac{x_*{}^2}{2}n^{1/3} - \frac{x_*{}^3}{6} + O\Big(\frac{x_*}{n^{1/3}}(1+x_*{}^3)\Big)\right),\tag{28}$$

$$\binom{k_i}{2} - k_i - l \;=\; \frac{x_i^2}{2}n^{4/3}\Big(1+O\Big(\frac{1}{x_i n^{2/3}}\Big)\Big)\,.\tag{29}$$

Using Stirling's formula and asymptotic formulae for $c(k_i, k_i+l)$ (see for instance [4, 24]), it yields

$$\begin{aligned}\frac{c(k_i,k_i+l)}{k_i!} \;&=\; \sqrt{\frac{3}{2}}d\,\exp\left(\frac{l}{2}+x_i n^{2/3}\right)x_i{}^{\frac{3}{2}l-1}n^{l-\frac{2}{3}}\frac{1}{(12\,l)^{l/2}}\\ &\qquad\times\left(1+O\Big(\frac{1}{l}\Big)+O\Big(\frac{1}{x_i n^{2/3}}\Big)+O\Big(\frac{l^{3/2}}{x_i{}^{1/2}n^{1/3}}\Big)\right)\,.\end{aligned}\tag{30}$$

(We use formulae $c(k_i, k_i+l)$ for $l = o(k_i^{1/3})$ and $d = \frac{1}{2\pi}$ as described in [4].) Also, after the same substitutions

$$t_i{}^{k_i+l} \;=\; \frac{1}{n^l n^{x_i n^{2/3}}}\exp\left(\frac{lu_i}{n^{1/3}} - \frac{lu_i}{2n^{2/3}} + x_iu_in^{1/3} - \frac{x_iu_i^2}{2} + O\Big(\frac{lu_i^3}{n}\Big) + O\Big(\frac{x_iu_i}{n^{1/3}}\Big)\right)\,,\tag{31}$$

and

$$(1-t_i\vee t_j)^{k_ik_j} \;=\; \exp\left(-x_ix_jn^{1/3} - x_ix_j(u_i\vee u_j) + O\Big(\frac{x_ix_j}{n^{2/3}}\Big)\right)\,.\tag{32}$$

Finally,

$$(33)\quad \begin{aligned}(1-t_i)^{(n-k_*)k_i+\binom{k_i}{2}-k_i-l-1} &= \exp\Big(-x_in^{2/3}-u_ix_in^{1/3}+x_*x_in^{1/3}+u_ix_ix_*-\frac{x_i^2}{2}n^{1/3}-\frac{x_i^2}{2}u_i\\ &\quad +\frac{3}{2}\frac{x_i}{n^{1/3}}+\frac{3}{2}\frac{u_ix_i}{n^{2/3}}+\frac{l}{n}+\frac{lu_i}{n^{4/3}}+\frac{1}{n}+\frac{u_i}{n^{4/3}}+O\Big(\frac{x_i}{n^{1/3}}\Big)\Big)\,.\end{aligned}$$

Using (28) – (33) above, the integrand in (26) reads

$$(34)\qquad \Psi_n^{(m)}(x_i,u_i)=A_m\exp(B_m)(1+\varepsilon)\,.$$

A bit of algebra gives A_m and B_m

$$(35)\qquad A_m=\left(\sqrt{\frac{3}{2}}\right)^m d^m\prod_{i=1}^m\frac{x_i^{3/2l+1}}{2}\,,$$

$$(36)\qquad B_m=\frac{ml}{2}\Big(1-\ln(12l)\Big)-\frac{{x_*}^3}{6}-\sum_{i=1}^m\frac{x_iu_i^2}{2}+x_*\sum_{i=1}^m x_iu_i-\frac{1}{2}\sum_{1\le i,\,j\le m}x_ix_j(u_i\vee u_j)\,.$$

The ε in (34) regroups all the big-Ohs produced by (28) – (33). In particular, if (x_i), (u_i) and $(1/x_i)$ are fixed, as $n\to\infty$, we have

$$\Psi_n^{(m)}(x_i,u_i)=A\exp(B)(1+o(1))\,.$$

So, if $x_i>0$, $u_i\in(-\infty,\infty)$ fixed, without restricting each $x_in^{2/3}$ to be an integer, we get

$$(37)\quad \begin{aligned}\Psi_n^{(m)}(x_i,u_i)\to&\left(\sqrt{\frac{3}{2}}\right)^m d^m\frac{\exp\left(\frac{ml}{2}\right)}{(12l)^{\frac{ml}{2}}}\left(\prod_{i=1}^m\frac{x_i^{3/2l+1}}{2}\right)\\ &\times\ \exp\left(-\frac{{x_*}^3}{6}-\sum_{i=1}^m\frac{x_iu_i^2}{2}+x_*\sum_{i=1}^m x_iu_i-\frac{1}{2}\sum_{1\le i,\,j\le m}x_ix_j(u_i\vee u_j)\right)\end{aligned}$$

as $n\to\infty$. Next, we use the estimate

$$\prod_{i<j}(1-t_i\vee t_j)^{k_ik_j}\le\prod_{i\ne j}(1-t_i)^{k_ik_j/2}\,,$$

to state that, there is a constant C_1 that may depend on m and l such that

$$(38)\qquad \Psi_n^{(m)}\le C_1(n)_{k_*}\prod_i\frac{c(k_i,k_i+l)}{k_i!}\left(\frac{k_i^2-3k_i}{2}-l-1\right){t_i}^{k_i}(1-t_i)^{k_i(n-2-k_*/2)}\,.$$

Then, using the bounds given in [14, eq. (2.12) – (2.18)] with (28) – (33), we get (the C_i below are constants and may depend on m and l)

$$(39)\quad \begin{aligned}\Psi_n^{(m)}\ \le\ & g_m(x_i,u_i)=C_2\exp(-\delta{x_*}^3)\prod_i x_i^{3l/2+1}\exp(-\delta x_iu_i^2)\\ &+\ C_3\exp(-\delta{x_*}^3)\prod_i x_i^{3l/2+1}\exp(-\delta x_iu_i)\\ &+\ C_4\exp(-\delta{x_*}^3)\prod_i\frac{1}{(1+u_i^2)}\,,\end{aligned}$$

valid for all n, x_i, u_i. Since $\int_0^\infty \cdots \int_0^\infty \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty g_m(x_i, u_i) dx_1 \cdots dx_m du_1 \cdots du_m < \infty$, (26), (37) and the use of Lebesgue dominated convergence yield

$$\mathbb{E}(Y_n^{(l)})_m \to \left(\sqrt{\frac{3}{8}} \frac{d \, \exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}} \right)^m a_m^{(l)} \,, \tag{40}$$

where

$$\begin{aligned} a_m^{(l)} &= \int_0^\infty \cdots \int_0^\infty \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty \left(\prod_{i=1}^m x_i^{3l/2+1} \right) \times \exp\Big(-\frac{1}{6} {x_*}^3 - \frac{1}{2} \sum_{i=1}^m x_i u_i^2 + x_* \sum_{i=1}^m x_i u_i \Big) \\ &\quad \times \exp\Big(-\frac{1}{2} \sum_{1 \le i, j \le m} x_i x_j (u_i \vee u_j) \Big) dx_1 \cdots dx_m du_1 \cdots du_m \,. \end{aligned} \tag{41}$$

3.2. **Creation of multicyclic components: Moments computation.** We observe that $(Z_n^{(l)})_m$ is the number of m-tuples of edges that are added between a p-component and, resp., a $(l-p)$-component of order k_i and, resp., k_j (with $-1 < p < l+1$). By Theorem 1.1, we find (for $l > 0$)

$$c'(k_i, k_i + l + 1) = \frac{1}{3l} \Big(k_i^2/2 - 3k_i/2 - l \Big) \, c(k_i, k_i + l) \Big(1 + O(1/l) + O(l^{3/2}/k_i^{1/2}) \Big) \tag{42}$$

which means that we can obtain expressions for $(Z_n^{(l)})$ by simply introducing a factor $1/3l$ in (25). Therefore,

$$\mathbb{E}(Z_n^{(l)})_m \to \left(\frac{1}{3l} \right)^m \left(\sqrt{\frac{3}{8}} \frac{d \, \exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}} \right)^m a_m^{(l)} \,, \tag{43}$$

3.3. **Mixed factorial moments.** Following (43), we have

$$\mathbb{E}(Y_n^{(l)})_p (Z_n^l)_q \to \left(\frac{1}{3l} \right)^q \left(\sqrt{\frac{3}{8}} \frac{d \, \exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}} \right)^{p+q} a_{p+q}^{(l)} \,, \tag{44}$$

as $n \to \infty$, $l > 0$ and $p, q \ge 0$. As for the mixed moments of $(Y_n^{(1)}, Z_n^{(1)})$ in [14], i.e. the convergence of all mixed factorial moments is equivalent to the convergence of all mixed moments , (44) implies that the sequence $(Y_n^{(l)}, Z_n^{(l)})$ converges in distribution to $(Y_\infty^{(l)}, Z_\infty^{(l)})$, with the factorial moments,

$$\mathbb{E}(Y_\infty^{(l)})_p \, (Z_\infty^l)_q = \left(\frac{1}{3l} \right)^q \left(\sqrt{\frac{3}{8}} \frac{d \, \exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}} \right)^{p+q} a_{p+q}^{(l)}$$

since (45) determine the distribution of $(Y_\infty^{(l)}, Z_\infty^{(l)})$. In order to show this, we note that (3) implies that $a_m^{(l)} \le a_1^m$, and thus the power series

$$A_l(z) = \sum_{m=0}^\infty a_m^{(l)} z^m / m!$$

converges for all z. (3) implies also that for all $t, u \ge 0$ (after a bit of algebra),

$$\mathbb{E} \Big((1+t)^{Y_\infty} \, (1+y)^{Z_\infty} \Big) \;=\; \mathbb{E} \sum_p \sum_q \binom{Y_\infty}{p} \binom{Z_\infty}{q} t^p u^q$$

$$(45) \qquad = A_l\left(\left(\left(\sqrt{\frac{3}{8}}\frac{d\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right) t + \left(\frac{1}{3l}\sqrt{\frac{3}{8}}\frac{d\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right) u\right)$$

Hence,

$$\mathbb{E}e^{tY_\infty + uZ_\infty} = A_l\left(\left(\left(\sqrt{\frac{3}{8}}\frac{d\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right)(e^t-1)\left(\frac{1}{3l}\sqrt{\frac{3}{8}}\frac{d\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right)(e^u-1)\right)\right) < \infty$$

for all $t, u \geq 0$, which implies that the distribution of (Y_∞, Z_∞) is determined by its moments.

This completes the proof that $(Y_n, Z_n) \to (Y_\infty, Z_\infty)$ in distribution. Particularly, $X_n = Y_n + Z_n \xrightarrow{d} Y_\infty + Z_\infty$, and if we define $X_\infty = Y_\infty + Z_\infty$, we have from (45)

$$\begin{aligned} \mathbb{E}(1+t)^{X_\infty} &= \mathbb{E}\left((1+t)^{Y_\infty}\,(1+t)^{Z_\infty}\right) \\ (46) \qquad &= \sum_{m=0} \frac{1}{m!}\,a_m^{(l)}\left(\frac{1+3l}{3l}\sqrt{\frac{3}{8}}\frac{d\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\,t\right)^m, \end{aligned}$$

which implies that $\mathbb{E}(X_\infty)_m = \left(\frac{1+3l}{3l}\sqrt{\frac{3}{8}}\frac{d\,\exp\left(\frac{l}{2}\right)}{(12l)^{\frac{l}{2}}}\right)^m a_m^{(l)}$. If Y is a r.v. with $\mathcal{L}\left(Y_\infty | X_\infty = r\right) = Bi\left(r, \left(\frac{3l}{3l+1}\right)\right)$ and $Z = X - Y$, we obtain

$$(47) \qquad \mathbb{E}\left((1+t)^Y (1+y)^Z\,|X_\infty = r\right) = \left(\frac{3l}{1+3l}(1+t) + \frac{1}{3l+1}(1+u)\right)^k$$

and finally, by (45) and (46)

$$\begin{aligned} \mathbb{E}\left((1+t)^Y(1+y)^Z\right) &= \mathbb{E}\left(\left(1+\frac{3l}{1+3l}\right)t + \left(\frac{1}{3l+1}\right)u\right)^{X_\infty} \\ (48) \qquad &= \mathbb{E}\left((1+t)^{Y_\infty}\,(1+y)^{Z_\infty}\right). \end{aligned}$$

Hence $(Y_n, Z_n) \stackrel{d}{=} (Y_\infty, Z_\infty)$, which justifies the description of the distribution (Y_∞, Z_∞) in the theorem.

Given the computation of $a_1^{(l)}$ below, the end of the proof of Theorem 1.2 can be completed using the same general lines as [14, eqs (2.33)–(2.37)]. By (3), we have for $l > 0$

$$\begin{aligned} a_1^{(l)} &= \int_0^\infty \int_{-\infty}^\infty x^{3l/2+1}\,\exp\left(-\frac{1}{24}x^3 - \frac{1}{2}xy^2\right) dy\,dx \\ (49) \qquad &= 2^{3l/2+2}\,3^{l/2-1/2}\,\pi^{1/2}\,\Gamma(l/2+1/2)\,. \end{aligned}$$

4. Conclusion

In this paper, we have studied the growths of complexity of connected components in an evolving graph. We have shown, using a combination of the methods from [14] and the theory of generating functions, how one can quantify asymptotically properties of the components growths. Amongst other things, we study complex components that increase their complexity by receiving new edges and/or by merging other complex components. As $l \to \infty$, our results show that whenever the second case occurs, almost all times, it is a unicyclic component that is swallowed.

Acknowledgments. We would like to thank Pr. Guy Louchard for encouraging us when working on this subject.

References

[1] Baert A. E., Ravelomanana, V., and Thimonier L., On the growth of components with non-fixed excesses. To appear in *Discrete Appl. Math.*
[2] Bender E. A. (1974). Asymptotic Methods in Enumeration, *SIAM Review*, 16:485–515.
[3] Bender, E. A., Canfield, E. R. and McKay B. D. (1990). The asymptotic number of labelled connected graphs with a given number of vertices and edges. *Random Structures and Algorithms*, 1:127–169.
[4] Bender, E. A., Canfield, E. R. and McKay B. D. (1992). Asymptotic properties of labeled connected graphs. *Random Structures and Algorithms*, 3:183–202.
[5] Bollobas, B. (1985). *Random Graphs.* Academic Press, London.
[6] De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis.* Dover, New-York.
[7] Cayley, A. (1889). A Theorem on Trees. *Quart. J. Math. Oxford Ser.*, 23:376–378.
[8] Erdös, P. and Rényi A. (1959). On random graphs. *Publ. Math. Debrecen*, 6:290–297.
[9] Erdös, P. and Rényi A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61.
[10] Flajolet, P., Knuth, D. E. and Pittel B. (1989). The first cycles in an evolving graph. *Discrete Math.*, 75:167–215.
[11] Flajolet, P. and Sedgewick, R. *Analytic Combinatorics.* To appear (chapters are avalaible as Inria research reports). See *http://algo.inria.fr/flajolet/Publications/books.html.*
[12] Goulden, I. P. and Jackson, D. M. (1983). *Combinatorial Enumeration.* Wiley, New York.
[13] Harary, F. and Palmer, E. (1973). *Graphical Enumeration.* Academic Press, New-York and London.
[14] Janson, S. (1993). Multicyclic components in random graphs process. *Random Structures and Algorithms*, 4:71–84.
[15] Janson, S. (2000). Growth of components in random graphs. *Random Structures and Algorithms*, 17:343-356.
[16] Janson, S., Knuth, D. E., Luczak, T. and Pittel B. (1993). The birth of the giant component. *Random Structures and Algorithms*, 4:233–358.
[17] Janson, S., Luczak, T. and Rucinski A. (2000). *Random Graphs.* John Wiley, New York.
[18] Knuth, D. E. and Pittel, B. (1989). A recurrence related to trees. *Proc. Am. Math. Soc.*, 105:335–349.
[19] Kolchin, V. F. (1999). *Random Graphs.* Encyclopedia of Mathematics and its Applications **53**. Cambridge University Press.
[20] Ravelomanana, V (2000). *Graphes multicycliques étiquetés : aspects combinatoires et probabilistes.* Phd Thesis – Université de Picardie, Amiens – France.
[21] Rényi, A. (1959). On connected graphs I. *Publ. Math. Inst. Hungarian Acad. Sci.* 4:385–388.
[22] J. Spencer (1997). Enumerating Graphs and Brownian Motion, *Communications on Pure and Applied Mathematics*, 50: 293–296.
[23] Wright, E. M. (1977). The Number of Connected Sparsely Edged Graphs. *Journal of Graph Theory*, 1:317–330.
[24] Wright, E. M. (1980). The Number of Connected Sparsely Edged Graphs. III. Asymptotic results *Journal of Graph Theory*, 4:393–407.

E-mail address: `baert@laria.u-picardie.fr, vlad@lipn.univ-paris13.fr`

Anne-Elisabeth BAERT, LaRIA EA-2083, Université de Picardie Jules-Verne, 80000 Amiens, France,

Vlady RAVELOMANANA, LIPN UMR-7030, Université de Paris-Nord, F 93430 Villetaneuse, France.