



**HAL**  
open science

# A system of grabbing particles related to Galton-Watson trees

Jean Bertoin, Vladas Sidoravicius, Maria Eulalia Vares

► **To cite this version:**

Jean Bertoin, Vladas Sidoravicius, Maria Eulalia Vares. A system of grabbing particles related to Galton-Watson trees. 2009. hal-00270344v2

**HAL Id: hal-00270344**

**<https://hal.science/hal-00270344v2>**

Preprint submitted on 4 Jun 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A system of grabbing particles related to Galton-Watson trees

Jean Bertoin\*      Vladas Sidoravicius †      Maria Eulalia Vares ‡

## Abstract

We consider a system of particles with arms that are activated randomly to grab other particles as a toy model for polymerization. We assume that the following two rules are fulfilled: once a particle has been grabbed then it cannot be grabbed again, and an arm cannot grab a particle that belongs to its own cluster. We are interested in the shape of a typical polymer in the situation when the initial number of monomers is large and the numbers of arms of monomers are given by i.i.d. random variables. Our main result is a limit theorem for the empirical distribution of polymers, where limit is expressed in terms of a Galton-Watson tree.

## 1 Introduction

We consider a system of interacting particles which might be used as a toy model for the formation of certain polymers. Roughly speaking, each particle possesses a certain number of *arms* that are activated successively at random to grab other particles. When an arm becomes active, it creates an oriented edge from the particle with the activated arm to the particle which has been grabbed, so the system describes the evolution of clusters of vertices with some inactive arms which are connected by oriented edges. As time passes, the number of inactive arms diminishes whereas the connectivity in the system increases by the creation of new edges. We impose the rules that once a particle has been grabbed, then it cannot be grabbed again, and that each time an arm is activated, it grabs a particle which does not belong to its own cluster. The latter requirement impedes the formation of cycles; thus the structure that connects each

---

\*Laboratoire de Probabilités, UPMC, 175 rue du Chevaleret, 75013 Paris; and DMA, ENS Paris, France. Email: jean.bertoin@upmc.fr

†IMPA, Estr. Dona Castorina 110, Rio de Janeiro, Brasil; and CWI, Kruislaan 413, 1098 SJ P.O. Box 94079, 1090 GB Amsterdam The Netherlands. Email: vladas@impa.br

‡CBPF. Rua Xavier Sigaud, 150. 22290-180 Rio de Janeiro, R.J. Brasil. Email: eulalia@cbpf.br

cluster of particles is given by a tree and the terminal state obtained when all arms have been activated can be described as a random forest.

This grabbing particle system bears some resemblance with the so-called random configuration model, which has been introduced independently by Bollobás [2] and Wormald [9]; see also Chapter 3 in Durrett [3] for further references. The latter is a natural stochastic algorithm which aims at producing a random graph on a set of vertices with pre-described degrees (in general the resulting graph is not simple and may contain loops and multiple edges). Typically, a certain number of stubs is appended to each vertex, and one joins pairwise the stubs uniformly at random to create edges between vertices. Similarities and differences between the two models will be discussed at the end of this work.

We are mainly interested in the case when the number of particles is large, and more precisely, in the situation when the empirical distribution of the number of arms converges to some probability measure  $\mu$  on  $\mathbb{Z}_+$  as the number of particles goes to infinity. Thinking of these particles as monomers, our goal is to get information about the statistics of a typical polymer structure when the polymerization process is completed. In this direction, we shall consider the slightly different setting where the sequence of the number of arms is given by i.i.d. random variables with law  $\mu$ , which is simpler to analyze. Loosely speaking, we will show that when the number of particles goes to infinity, a typical tree (i.e. a tree chosen uniformly at random in the terminal forest) tends to be distributed as a Galton-Watson tree with reproduction law given by the distribution of the number of arms of a typical particle. This result is perhaps quite intuitive; however providing a rigorous argument is not so straightforward, and this motivates the present work.

The remainder of this paper is organized as follows. The grabbing particle system is presented formally in the next section, in the situation when the number of particles is finite and the number of arms of each particle is deterministic. In Section 3, we consider the case when the number of arms is random and given by a sequence of i.i.d. variables, and we make the connexion with Galton-Watson processes. More precisely, we shall show that conditionally on the total number of arms, the distribution of the terminal state is that of a Galton-Watson forest with a given number of trees, conditioned by its total size. Finally Section 4 is devoted to the asymptotic study of the empirical measure of trees in the terminal state when the number of particles tends to infinity.

## 2 The grabbing particle system

Consider  $n \geq 2$  particles which are labeled by  $1, \dots, n$ . Each particle, say  $i$ , consists in a vertex to which a certain number  $x_i \in \mathbb{Z}_+$  of arms are attached. Arms can be either active

or inactive. An active arm attached to  $i$  is an oriented edge  $i \rightarrow j$  linking  $i$  to another vertex  $j \neq i$ , whereas an inactive arm is an incomplete oriented edge  $i \rightarrow$  where only the left-end is specified. Initially, every arm is inactive, and once an arm has been activated, it remains active forever. We assume that the total number of arms is less than the number of particles, i.e.

$$x_1 + \dots + x_n = n - k \quad \text{for some integer } 1 \leq k \leq n. \quad (2.1)$$

We then enumerate these  $n - k$  arms uniformly at random, which specifies the order at which they are activated.

Specifically, suppose that the first arm to be activated belongs to the particle  $i$ . Then at time 1, this arm grabs a particle  $j \neq i$  chosen uniformly at random amongst the other  $n - 1$  particles. The system at time 1 thus consists in  $n - 2$  isolated particles with inactive arms, and a cluster of two particles,  $i$  and  $j$ , connected by an oriented edge  $i \rightarrow j$  which is labeled by 1. The particle  $i$  has still  $x_i - 1$  inactive arms, and the particle  $j$  has kept its  $x_j$  inactive arms; inactive arms are labeled by  $2, \dots, n - k$ . We iterate in an obvious independent manner, with the rules that every particle can be grabbed by at most one arm, and that each time an arm is activated, it grabs a particle from a different cluster.

It is convenient to depict the evolution of the system in the upper half-plane by introducing some further orderings which of course do not alter the dynamics. Specifically, imagine that initially, particles lie on the horizontal axis, such that the initial order from the left to the right is uniformly random. We may then think of the  $x_i$  arms attached to the particle  $i$  as vectors making an angle  $\ell\pi/(x_i + 1)$  for  $\ell = 1, \dots, x_i$  with the (directed) horizontal axis. If, as above, the first arm to be activated belongs to the particle  $i$  and grabs particle  $j$ , then at time 1, the particle  $j$  (together with its arms) is shifted at the right-end of the first arm. See Figure 1 for this first step of the evolution of the system (with  $n = 7$ ,  $k = 2$ ,  $i = 6$  and  $j = 5$ ).



*Figure 1: planar representation of a grabbing particle system at times  $t = 0$  (left) and  $t = 1$  (right)*

We then continue by an obvious iteration. Only particles lying on the horizontal axis can be grabbed, and each time such a particle is grabbed by the activation of an arm, it is shifted

together with its connected component to become the right-end of that arm. As time passes, we thus obtain a growing family of planar rooted trees, where vertices may bear some further inactive arms, and the roots are the vertices still lying on the horizontal axis, i.e. those which have not been grabbed so far. Edges and inactive arms are labeled by  $1, \dots, n - k$ , the label of an edge being always smaller than that of an inactive arm, and the vertices are labeled by  $1, \dots, n$ . The evolution reaches its final state at time  $n - k$  when all the arms have been activated; obviously we then have a forest of  $k$  planar rooted trees with  $n - k$  labeled edges on a set of  $n$  vertices, and the trees are ordered from the left to the right. See Figure 2 below. We stress that the terminal state retains all the essential<sup>1</sup> information about evolution of the system, as it suffices to prune the last  $n - k - \ell$  edges to recover the clusters of particles in the system and their structures at time  $\ell$ .

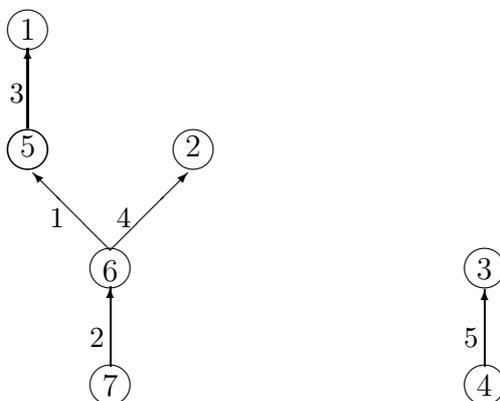


Figure 2: terminal state as a planar forest

We shall now describe the distribution of the terminal state for the grabbing particle system. In this direction, it is convenient to introduce first the space  $\Phi_{n,k}$  of planar forests with  $k$  rooted trees on a set of  $n$  labeled vertices, with labeled edges. The  $k$  roots lie on the horizontal axis and are thus naturally ordered from the left to the right. The choice of a root in a tree induces an orientation of the edges in that tree, by deciding that each edge points away from the root. In this setting, recall the notation  $i \rightarrow j$  for  $i, j \in \{1, \dots, n\}$  with  $i \neq j$  to denote an oriented edge from the vertex labeled by  $i$  to the vertex labeled by  $j$ . For any forest  $\varphi \in \Phi_{n,k}$  and  $i \in \{1, \dots, n\}$ , we write

$$d_i(\varphi) := \text{Card}\{j \in \{1, \dots, n\} \setminus \{i\} : i \rightarrow j \text{ is an edge of } \varphi\}$$

for the outer-degree of the vertex labeled by  $i$ , i.e. the number of oriented edges which left-end

---

<sup>1</sup>However the exact ordering of particles lying on the horizontal axis is lost in general, but this is not relevant as far as empirical distributions of the clusters and their structures are concerned.

is the  $i$ -th vertex.

Recall that  $x_1, \dots, x_n$  is a sequence of nonnegative integers such that (2.1) holds. We then denote by  $\Phi(x_1, \dots, x_n) \subset \Phi_{n,k}$  the subset of planar forests  $\varphi$  with  $k$  rooted trees, labeled edges and labeled vertices, such that the outer-degree of  $i$  is  $d_i(\varphi) = x_i$  for every  $i = 1, \dots, n$ .

**Lemma 1** *For every integer  $1 \leq k \leq n$  and every sequence of nonnegative integers  $x_1, \dots, x_n$  which fulfills (2.1), the distribution of the terminal state of the grabbing particle system is given by the uniform law on  $\Phi(x_1, \dots, x_n)$ .*

**Proof:** For  $\ell = 0, \dots, n - k$ , let  $\Phi_\ell(x_1, \dots, x_n)$  denote the set of planar configurations which can be obtained from the grabbing particle system at time  $\ell$ , when initially the particle  $i$  possesses  $x_i$  inactive arms for  $i = 1, \dots, n$ . In other words,  $\Phi_\ell(x_1, \dots, x_n)$  is the set of planar forests on a set of  $n$  labeled vertices, with  $n - \ell$  rooted trees lying on the horizontal axis, to which a total of  $n - k - \ell$  yet incomplete edges (i.e. with only a left-end) have been attached to certain vertices, in such a way that for every  $i = 1, \dots, n$ , the number of complete or incomplete edges with left-end  $i$  is  $x_i$ . Further, complete or incomplete edges are enumerated in such a way that complete edges are listed before the incomplete ones. In particular for  $\ell = n - k$ , in the notation above, there is the identity  $\Phi_{n-k}(x_1, \dots, x_n) = \Phi(x_1, \dots, x_n)$ .

Plainly, under the assumptions made in the claim, the initial state has the uniform distribution in  $\Phi_0(x_1, \dots, x_n)$ . We claim that the activation of the first arm yields a random state in  $\Phi_1(x_1, \dots, x_n)$  which has again the uniform distribution. Indeed, pick an arbitrary state  $\varphi \in \Phi_1(x_1, \dots, x_n)$ . There are exactly  $n$  states in  $\Phi_0(x_1, \dots, x_n)$  which may possibly yield  $\varphi$  by activating the first arm. More precisely, these are the states which are obtained from  $\varphi$  by pruning its unique complete edge at its right-end, and inserting this right-end (together with its inactive arms) on the horizontal axis relatively to the  $n - 1$  roots of  $\varphi$ . Further, starting from any of these  $n$  states, the probability of getting  $\varphi$  after activating their first arm is exactly the probability that this arm picks the right root amongst the  $n - 1$  possible ones, that is  $1/(n - 1)$ . This establishes our claim.

We then can iterate this argument, showing that the terminal state is uniformly distributed on  $\Phi_{n-k}(x_1, \dots, x_n) = \Phi(x_1, \dots, x_n)$ , as stated.  $\square$

### 3 Connexion with Galton-Watson forests

We now turn our attention to the situation when the number of arms of particles is given by an i.i.d. sequence of random variables. Specifically, we consider a probability measure  $\mu$  on  $\mathbb{Z}_+$  and assume that the number of arms of the  $i$ -th particle, which it is henceforth convenient to

denote by  $\xi_i$ , is random with law  $\mu$ , and that  $\xi_1, \dots, \xi_n$  are independent. We shall see that the terminal state of the system can then be described in terms of a Galton-Watson forest; in this direction, we first develop some material in this field.

Consider a Galton-Watson process with reproduction law  $\mu$ , so, roughly speaking, we have a population model with discrete non-overlapping generations, in which each individual gives birth at the next generation to a random number of children according to the law  $\mu$  and independently of the other individuals. We shall work on the event that the total descent of each individual is finite a.s. By ranking the ancestors and the progeny of each individual, we can represent the process by a planar forest with rooted trees, where roots correspond to the ancestors at the initial generation, and each tree describes a genealogy, see e.g. Section 6.2 in Pitman [6].

We denote by  $F_{n,k}$  the finite set of all planar forests with  $k$  rooted trees on a set of  $n$  vertices (vertices and edges are not yet labeled). Recall that vertices and edges in a forest in  $F_{n,k}$  can be listed by using, for instance, the depth-first search algorithm (cf. Figure 6.1 in [6] on its page 126). In this setting, assigning labels to the  $n$  vertices (respectively, to the  $n - k$  edges) is equivalent to choosing a permutation on  $\{1, \dots, n\}$  (respectively, a permutation on  $\{1, \dots, n - k\}$ ). Thus, if for every  $\ell \in \mathbb{N}$ ,  $\Sigma_\ell$  stands for the group of permutations on  $\{1, \dots, \ell\}$ , then there is the canonical identification as the product space

$$\Phi_{n,k} = F_{n,k} \times \Sigma_n \times \Sigma_{n-k}.$$

We then introduce a product probability measure on  $F_{n,k} \times \Sigma_n \times \Sigma_{n-k}$ , denoted by  $\mathbb{GW}_{k,n}^{(\mu)}$ , that can thus be viewed as the law of a triplet of independent variables. The first random variable with values in  $F_{n,k}$  is a Galton-Watson forest with reproduction law  $\mu$ , having  $k$  trees and conditioned to have a total size  $n$  (assuming implicitly that this event has positive probability), and the second and third are simply uniformly distributed on  $\Sigma_n$  and  $\Sigma_{n-k}$ , respectively. We may thus think of  $\mathbb{GW}_{k,n}^{(\mu)}$  as the law on  $\Phi_{n,k}$  of a random state which is obtained by picking a planar forest with  $k$  rooted trees on a set of  $n$  vertices according to a Galton-Watson process with reproduction law  $\mu$ , started with  $k$  ancestors and conditioned to have total size  $n$ , and independently enumerating the edges and the vertices uniformly at random.

**Theorem 1** *Let  $\xi_1, \dots$  be a sequence of i.i.d. variables with law  $\mu$ . Fix integers  $1 \leq k \leq n$  and work conditionally on  $\xi_1 + \dots + \xi_n = n - k$ , provided that the probability of this event is nonzero. Then the distribution of the terminal state of the grabbing particle system started with  $n$  particles and such that the initial number of inactive arms of the  $i$ -th particle is  $\xi_i$  for every  $i \in \{1, \dots, n\}$ , is the law  $\mathbb{GW}_{n,k}^{(\mu)}$  on  $\Phi_{n,k}$ .*

**Proof:** We first work with a Galton-Watson process with reproduction law  $\mu$  started with

$k$  ancestors; the notation  $\mathbb{P}_k$  will refer to the distribution of the process. Let  $(X_i)_{i \geq 0}$  be a sequence of i.i.d. variables with law  $\mu$ , introduce the downwards skip-free random walk  $S_\ell = X_1 + \dots + X_\ell - \ell$  and the first-passage time  $T := \min\{\ell \geq 1 : S_\ell = -k\}$ . By listing individuals in the Galton-Watson process according to the depth-first search algorithm, it is well-known that  $\mathbb{P}_k$  can be identified with the distribution of the stopped sequence  $(X_i)_{1 \leq i \leq T}$ . More precisely, in this setting,  $X_i$  represents the number of children of the  $i$ -th individual listed by depth-first search.

Next pick an arbitrary planar forest  $f \in F_{n,k}$  with  $k$  rooted trees and  $n$  vertices (vertices and edges are not labeled). The depth-first search algorithm enables us to encode  $f$  by some sequence of integers  $(y_i)_{1 \leq i \leq n}$  (more precisely,  $y_i$  is the outer-degree of the  $i$ -th vertex found by the depth-first search algorithm), and the probability that the genealogical forest induced by the  $k$  ancestors of the Galton-Watson process is given by  $f$  thus equals  $\mu(y_1) \cdots \mu(y_n)$ . As a consequence, for every permutations  $\sigma \in \Sigma_n$  and  $\varsigma \in \Sigma_{n-k}$ , we have

$$\mathbb{G}\mathbb{W}_{n,k}^{(\mu)}((f, \sigma, \varsigma)) = \frac{\mu(y_1) \cdots \mu(y_n)}{n!(n-k)!z}$$

with  $z = \mathbb{P}_k(T = n)$ .

Then pick a planar forest  $\varphi$  with  $k$  rooted trees on a set of  $n$  labeled vertices, with labeled edges, which belongs to the subset  $\Phi(x_1, \dots, x_n)$ . We can identify

$$\varphi = (f, \sigma, \varsigma)$$

for a unique  $f \in F_{n,k}$ ,  $\sigma \in \Sigma_n$  and  $\varsigma \in \Sigma_{n-k}$ . More precisely,  $f$  is encoded via depth-first search by the sequence  $(y_i)_{1 \leq i \leq n}$  where  $x_i = y_{\sigma(i)}$ , and therefore

$$\mathbb{G}\mathbb{W}_{n,k}^{(\mu)}(\varphi) = \frac{\mu(x_1) \cdots \mu(x_n)}{n!(n-k)!z}.$$

This shows that, provided that  $\mu(x_i) > 0$  for every  $i = 1, \dots, n$ , the conditional probability  $\mathbb{G}\mathbb{W}_{n,k}^{(\mu)}(\cdot \mid \Phi(x_1, \dots, x_n))$  is simply the uniform law on  $\Phi(x_1, \dots, x_n)$ . We deduce by comparison with Lemma 1 that for an arbitrary sequence of integers  $(x_i)_{1 \leq i \leq n}$  which fulfills (2.1) and  $\mu(x_i) > 0$ , the distribution of the terminal state in the grabbing particle system started with  $n$  particles and such that the number of arms of the  $i$ -th particle is  $x_i$  coincides with the conditional law  $\mathbb{G}\mathbb{W}_{n,k}^{(\mu)}(\cdot \mid \Phi(x_1, \dots, x_n))$ .

To conclude the proof, we just need to recall the following well-known consequence of the ballot theorem (see, for instance, Lemma 6.1 in Pitman [6]). Under the conditional law  $\mathbb{P}_k(\cdot \mid T = n)$  of the Galton-Watson process started with  $k$  ancestors and conditioned to have a total population with size  $n$ , the sequence of the number of children of individuals listed uniformly

at random is given by a sequence of  $n$  independent random variables with law  $\mu$  conditioned to add-up to  $n - k$ . Recall that the event that this sequence is given by  $(x_1, \dots, x_n)$  is precisely  $\Phi(x_1, \dots, x_n)$ . This completes the proof of the statement.  $\square$

## 4 Asymptotic behavior of the empirical measure

Our goal here is to investigate the asymptotic behavior of the empirical measure of the trees resulting from the grabbing particle system started with  $n$  i.i.d. particles as  $n \rightarrow \infty$ . More precisely, we shall work with the following setting.

Let  $\mu$  be a critical or sub-critical probability measure on  $\mathbb{Z}_+$ , i.e. with mean

$$m := \sum_{\ell=1}^{\infty} \ell \mu(\ell) \leq 1.$$

We shall also always assume that  $\mu(0) > 0$ . Let  $\xi_1, \dots$  be a sequence of i.i.d. variables with law  $\mu$ , and for each  $n \geq 2$ , set

$$k(n) := n - (\xi_1 + \dots + \xi_n).$$

We work conditionally on the event  $k(n) \geq 1$  and consider the grabbing particle system started with  $n$  particles such that the initial number of inactive arms of the  $i$ -th particle is  $\xi_i$  for every  $i \in \{1, \dots, n\}$ . We write  $\mathbf{P}_n$  for the distribution of its terminal state in  $\Phi_n := \bigcup_{1 \leq k \leq n} \Phi_{n,k}$ , and also denote by  $\tau_1, \dots, \tau_{k(n)}$  the sequence of planar rooted trees in this terminal state, ignoring labels on vertices and edges. That is to say that the sequence  $(\tau_1, \dots, \tau_{k(n)})$  is the image of the terminal state by the canonical projection  $\Phi_{n,k(n)} \rightarrow F_{n,k(n)}$ .

We also introduce the space of finite planar rooted trees with unlabeled edges and vertices,

$$\mathcal{T} := \bigcup_{\ell \geq 1} F_{\ell,1}.$$

We denote by  $\mathbb{T}^{(\mu)}$  the probability measure on  $\mathcal{T}$  which is induced by the genealogical tree of a single ancestor in a Galton-Watson process with reproduction law  $\mu$ . We are now able to state our main limit theorem.

**Theorem 2** *Assume that  $\mu$  is a critical or sub-critical probability measure on  $\mathbb{Z}_+$  with  $\mu(0) > 0$ . Then the proportion of tree structures identical to some fixed  $\mathbf{t} \in \mathcal{T}$  in the terminal state of the grabbing particle system converges to  $\mathbb{T}^{(\mu)}(\mathbf{t})$  in probability under  $\mathbf{P}_n = \mathbb{P}(\cdot \mid k(n) \geq 1)$  as*

$n \rightarrow \infty$ . That is, equivalently,

$$\lim_{n \rightarrow \infty} \mathbf{E}_n \left( \left| \frac{1}{k(n)} \sum_{\ell=1}^{k(n)} \mathbf{1}_{\{\tau_\ell = \mathbf{t}\}} - \mathbb{T}^{(\mu)}(\mathbf{t}) \right|^2 \right) = 0 \quad (4.2)$$

We shall reduce the proof of (4.2) to establishing the following two asymptotic behaviors under the conditional probabilities  $\mathbf{P}_n$ .

**Lemma 2** *The number of trees in the terminal state under the law  $\mathbf{P}_n$  converges in probability to  $\infty$ , that is*

$$\lim_{n \rightarrow \infty} \mathbf{P}_n(k(n) \geq k) = 1 \quad \text{for every } k \in \mathbb{N}.$$

**Proof:** Note that

$$\mathbb{P}(k(n) \geq k) = \mathbb{P}(\xi_1 + \cdots + \xi_n \leq n - k)$$

tends to 1 when  $\mu$  is subcritical (and then the statement is obvious), but may tend to 0 when  $\mu$  is critical. We therefore focus on the critical case. For simplicity, we assume that the support of  $\mu$  is not contained into a strict subgroup of  $\mathbb{Z}$ ; the modification needed to treat the opposite case is straightforward. Then  $\xi_1 + \cdots + \xi_n - n$  is an aperiodic, irreducible recurrent random walk on  $\mathbb{Z}$ , and we know from a ratio limit theorem in Spitzer [7] on its page 49 that for every fixed  $\ell \in \mathbb{Z}$

$$\mathbb{P}(\xi_1 + \cdots + \xi_n = n - \ell) \sim \mathbb{P}(\xi_1 + \cdots + \xi_n = n).$$

It follows

$$\mathbb{P}(\xi_1 + \cdots + \xi_n = n - \ell) = o(\mathbb{P}(\xi_1 + \cdots + \xi_n \leq n));$$

consequently

$$\mathbb{P}(\xi_1 + \cdots + \xi_n \leq n - \ell) \sim \mathbb{P}(\xi_1 + \cdots + \xi_n \leq n), \quad (4.3)$$

which yields our claim.  $\square$

**Lemma 3** *We have*

$$\lim_{n \rightarrow \infty} \mathbf{P}_n(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2, k(n) \geq 3) = \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2). \quad (4.4)$$

**Proof:** We denote by  $|\mathbf{t}|$  the size (number of vertices) of a tree  $\mathbf{t} \in \mathcal{T}$ . Using Theorem 1, we know that for each fixed  $n \geq 3$

$$\begin{aligned} & \mathbf{P}_n(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2, k(n) \geq 3) \\ &= c_n \sum_{k=3}^n \mathbb{P}(\xi_1 + \cdots + \xi_n = n - k) \mathbb{GW}_{n,k}^{(\mu)}(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2), \end{aligned} \quad (4.5)$$

with

$$c_n = 1/\mathbb{P}(\xi_1 + \cdots + \xi_n \leq n - 1).$$

Then recall the following celebrated identity of Dwass [4]: the first passage time process

$$T_k = \inf\{n \geq 0 : \xi_1 + \cdots + \xi_n = n - k\}$$

has i.i.d. increments, which have the same distribution as the sequence of the sizes of trees in a Galton-Watson forest with reproduction law  $\mu$ . In particular the step  $T_1$  is distributed as  $|\tau|$  under  $\mathbb{T}^{(\mu)}$ . Further the formula of Kemperman (see, for instance, (6.3) in [6] on its page 122) states:

$$\mathbb{P}(T_k = n) = \frac{k}{n} \mathbb{P}(\xi_1 + \cdots + \xi_n = n - k), \quad \text{for every } 1 \leq k < n.$$

By definition of the conditional law, we have thus for  $k \geq 3$  and  $t := |\mathbf{t}_1| + |\mathbf{t}_2|$

$$\begin{aligned} & \mathbb{G}\mathbb{W}_{n,k}^{(\mu)}(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2) \\ &= \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2) \frac{\mathbb{P}(T_{k-2} = n - t)}{\mathbb{P}(T_k = n)} \\ &= \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2) \frac{n(k-2)}{k(n-t)} \frac{\mathbb{P}(\xi_1 + \cdots + \xi_{n-t} = n - k + 2 - t)}{\mathbb{P}(\xi_1 + \cdots + \xi_n = n - k)}, \end{aligned}$$

so combining with (4.5) we arrive at

$$\begin{aligned} & \mathbf{P}_n(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2, k(n) \geq 3) \\ &= c_n \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2) \frac{n}{n-t} \sum_{k=3}^n \frac{k-2}{k} \mathbb{P}(\xi_1 + \cdots + \xi_{n-t} = n - k + 2 - t). \end{aligned} \quad (4.6)$$

In the subcritical case, we have  $c_n \rightarrow 1$  and also, by the weak law of large numbers, that the series in (4.6) converges to 1 as  $n \rightarrow \infty$ . This proves (4.4) in that case.

We henceforth focus on the critical case. In this direction, from (4.3) and the identity

$$\mathbb{P}(\xi_1 + \cdots + \xi_{n+1} \leq n + 1) = \sum_{k=0}^{\infty} \mu(k) \mathbb{P}(\xi_1 + \cdots + \xi_n \leq n + 1 - k)$$

we get by dominated convergence that

$$\mathbb{P}(\xi_1 + \cdots + \xi_{n+1} \leq n + 1) \sim \mathbb{P}(\xi_1 + \cdots + \xi_n \leq n). \quad (4.7)$$

On the one hand, combining (4.3), (4.7) and the obvious upper-bound

$$\sum_{k=3}^n \frac{k-2}{k} \mathbb{P}(\xi_1 + \cdots + \xi_{n-t} = n - k + 2 - t) \leq \mathbb{P}(\xi_1 + \cdots + \xi_{n-t} \leq n - 1 - t),$$

we deduce from (4.6) that

$$\limsup_{n \rightarrow \infty} \mathbf{P}_n(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2, k(n) \geq 3) \leq \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2).$$

On the other hand, pick  $0 < \eta < 1$  and let  $k_\eta$  be an integer such that  $(k-2)/k \geq 1 - \eta$  whenever  $k \geq k_\eta$ . Combining (4.3), (4.7) and the lower-bound

$$\sum_{k=3}^n \frac{k-2}{k} \mathbb{P}(\xi_1 + \cdots + \xi_{n-t} = n - k + 2 - t) \geq (1 - \eta) \mathbb{P}(\xi_1 + \cdots + \xi_{n-t} \leq n - k_\eta + 2 - t),$$

we deduce from (4.6) that

$$\liminf_{n \rightarrow \infty} \mathbf{P}_n(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2, k(n) \geq 3) \geq (1 - \eta) \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2).$$

Since  $\eta$  can be chosen arbitrarily small, this shows (4.4).  $\square$

We have now all the ingredients to establish Theorem 2 using a standard argument of propagation of chaos (see e.g. Sznitman [8]).

**Proof:** The obvious inequality

$$\sum_{\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{T}} \mathbf{P}_n(\tau_1 = \mathbf{t}_1, \tau_2 = \mathbf{t}_2, k(n) \geq 3) = \mathbf{P}_n(k(n) \geq 3) \leq 1 = \sum_{\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{T}} \mathbb{T}^{(\mu)}(\mathbf{t}_1) \mathbb{T}^{(\mu)}(\mathbf{t}_2)$$

combined with Scheffé's lemma show that the convergence (4.4) also holds in the space  $\ell^1(\mathcal{T} \times \mathcal{T})$  of summable families indexed by the countable set  $\mathcal{T} \times \mathcal{T}$ . It follows that

$$\lim_{n \rightarrow \infty} \mathbf{P}_n(\tau_1 = \mathbf{t}) = \mathbb{T}^{(\mu)}(\mathbf{t}),$$

and then by an argument based on exchangeability that

$$\lim_{n \rightarrow \infty} \mathbf{E}_n \left( \frac{1}{k(n)} \sum_{\ell=1}^{k(n)} \mathbf{1}_{\{\tau_\ell = \mathbf{t}\}} \right) = \mathbb{T}^{(\mu)}(\mathbf{t}).$$

Using again (4.4), exchangeability and Lemma 2, we get the asymptotic of the second moment

$$\lim_{n \rightarrow \infty} \mathbf{E}_n \left( \left( \frac{1}{k(n)} \sum_{\ell=1}^{k(n)} \mathbf{1}_{\{\tau_\ell = \mathbf{t}\}} \right)^2 \right) = \mathbb{T}^{(\mu)}(\mathbf{t})^2,$$

which establishes (4.2). □

Let us briefly discuss the super-critical case, i.e. when  $m > 1$  (recall that we always assume that  $\mu(0) > 0$ ). Then  $\mathbb{P}(k(n) \geq 1)$  tends to 0 as  $n \rightarrow \infty$ , so the conditioning on this event becomes singular. More precisely, it is readily seen that the distribution of  $k(n)$  under the conditional law  $\mathbf{P}_n = \mathbb{P}(\cdot \mid k(n) \geq 1)$  converges to some geometric law. In particular Lemma 2 fails and the situation differs crucially from that describes in Theorem 2. Nonetheless one can get an analogue of Theorem 2 in the supercritical case, provided that we condition on  $k(n) \geq cn$  for some  $0 < c < 1$ . Then standard arguments of large deviations can be incorporated to the proof of Theorem 2 and yield a similar limit theorem in which  $\mathbb{T}^{(\mu)}(\mathbf{t})$  has to be replaced by  $\mathbb{T}^{(\tilde{\mu})}(\mathbf{t})$ , where  $\tilde{\mu}$  is the unique law in the exponential family of  $\mu$  with mean  $1 - c$ . Details are left to the interested reader(s).

We now conclude this work by comparing the asymptotic behaviors for the present grabbing particle system and for the model of random configuration (see [1, 5]), starting in both cases from a large number of particles such that the sequence of the numbers of arms attached to those particles is i.i.d. with a fixed law  $\mu$ . Recall that in the latter, edges are formed by pairing arms uniformly at random, and that Molloy and Reed [5] have shown (under slightly different assumptions) that with high probability, the random configuration model does not possess a giant connected component if and only if  $\sum_{\ell=0}^{\infty} \ell(\ell - 2)\mu(\ell) \leq 0$ . Further, in that case, if we select an arm uniformly at random and independently of the pairing process, then the structure of the cluster containing that arm can be described in terms of two independent Galton-Watson trees with reproduction law  $\nu(k) = (k + 1)\mu(k + 1)/m$ ; see Theorem 1 in [1] for details. So, roughly speaking, both models produce random graphs with the same degree sequences, and for both models the structure of a typical cluster can be described in terms of Galton-Watson trees. We stress that the reproduction laws are distinct (except in the Poisson case), which may seem paradoxical as degree sequences are closely related to the latter. However one should keep in mind that selecting a cluster by picking an arm uniformly at random differs from selecting a cluster uniformly at random, and more precisely induces a bias by the number of arms in the random cluster. This explains the relation between the two reproduction laws.

**Acknowledgments :** We would like to thank two anonymous referees for their insightful comments on the first draft of this work. In particular, clever arguments due to a referee

have permitted to remove an irrelevant condition involving stable domains of attraction in our original version of Theorem 2.

J. B. thanks the programs ARCUS and France-Brésil for supporting his visit to IMPA at Rio de Janeiro during which the present joint work has been undertaken. M.E.V. acknowledges the support of CNPq, Brasil Grant number 302796/2002-9.

## References

- [1] J. Bertoin and V. Sidoravicius, *The structure of typical clusters in large sparse random configurations*, J. Stat. Phys. **135** (2009), 87-105.
- [2] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. Combin. **1** (1980), 311-316.
- [3] R. Durrett, *Random graph dynamics*. Cambridge University Press, Cambridge, 2007.
- [4] M. Dwass, *The total progeny in a branching process*, J. Appl. Probab. **6** (1969), 682-686.
- [5] M. Molloy and B. Reed, *A critical point for random graphs with a given degree sequence*, Random Struct. Algorithms **6** (1995), 161-179.
- [6] J. Pitman, *Combinatorial stochastic processes. École d'été de Probabilités de St-Flour*, Lect. Notes in Maths **1875**, Berlin, Springer, 2006. Available via :  
[http://stat-www.berkeley.edu /users/pitman/](http://stat-www.berkeley.edu/users/pitman/)
- [7] F. Spitzer, *Principles of random walks*. 2nd edition. Graduate Texts in Mathematics, Vol. 34. Springer-Verlag, New York-Heidelberg, 1976.
- [8] A.-S. Sznitman, *Topics in propagation of chaos. Ecole d'été de Probabilités de St-Flour XIX*, Lect. Notes in Maths **1464**, Springer, 1991.
- [9] N. C. Wormald, *Some Problems in the Enumeration of Labelled Graphs*. Doctoral thesis, Newcastle University, 1978.