The structure of typical clusters
in large sparse random configurations

Jean Bertoin*   Vladas Sidoravicius [†]

Abstract

The initial purpose of this work is to provide a probabilistic explanation of a recent result on a version of Smoluchowski's coagulation equations in which the number of aggregations is limited. The latter models the deterministic evolution of concentrations of particles in a medium where particles coalesce pairwise as time passes and each particle can only perform a given number of aggregations. Under appropriate assumptions, the concentrations of particles converge as time tends to infinity to some measure which bears a striking resemblance with the distribution of the total population of a Galton-Watson process started from two ancestors.

Roughly speaking, the configuration model is a stochastic construction which aims at producing a typical graph on a set of vertices with pre-described degrees. Specifically, one attaches to each vertex a certain number of stubs, and then join pairwise the stubs uniformly at random to create edges between vertices.

In this work, we use the configuration model as the stochastic counterpart of Smoluchowski's coagulation equations with limited aggregations. We establish a hydrodynamical type limit theorem for the empirical measure of the shapes of clusters in the configuration model when the number of vertices tends to ∞. The limit is given in terms of the distribution of a Galton-Watson process started with two ancestors.

1 Introduction

The motivation for this work stems from a recent study of a deterministic model for coagulation with limited number of aggregations. Specifically, in [5], one considers particles that are

---

*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

determined by a pair of integers (a, k) where $k \geq 1$ represents the size and $a \geq 0$ the number of aggregations that the particle can perform. In the model called symmetric, coagulations

$$\{(a, k), (a', k')\} \longrightarrow (a + a' - 2, k + k')$$

occurs at rate

$$aa' c_t(a, k) c_t(a', k'),$$

where $c_t(a, k)$ denotes the concentration of particles (a, k) at time t in the medium. Analytically, this means that the evolution of concentrations is governed by the following variation of Smoluchowski's coagulation equations (cf. the survey by Aldous [2]):

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} c_t(a, k) = \quad & \frac{1}{2} \sum_{a'=1}^{a+1} \sum_{k'=1}^{k-1} a'(a - a' + 2) c_t(a', k') c_t(a - a' + 2, k - k') \\
& - \sum_{a'=1}^{\infty} \sum_{k'=1}^{\infty} aa' c_t(a, k) c_t(a', k'),
\end{aligned}
\tag{1}
$$

where the first term in the right-hand side accounts for the creation of particles (a, k) as the result of coagulations of pairs $\{(a', k'), (a - a' + 2, k - k')\}$ and the second term for the disappearance of particles (a, k) after a coagulation with a particle (a', k').

One of the main results in [5] is that under appropriate conditions on the initial data that we shall recall later on, the concentrations $c_t(a, k)$ have a limit as time t tends to infinity which is given by

$$c_\infty(a, k) = \mathbf{1}_{\{a=0\}} \frac{1}{k(k-1)} \nu^{*k}(k - 2) \qquad \text{for } a \in \mathbb{N} \text{ and } k \geq 2. \tag{2}$$

Here, ν is a certain probability measure on $\mathbb{N}$ with $\sum_{n=0}^{\infty} n\nu(n) \leq 1$ that depends on the initial data, and $\nu^{*k} = \nu * \cdots * \nu$ denotes its k-th convolution power. The expression (2) bears a striking resemblance with a special case of the celebrated formula due to Dwass [9] who established that the total population $T_2(\nu)$ generated by a (sub)-critical Galton-Watson branching process with reproduction law ν and started from two ancestors is given by

$$\mathbb{P}(T_2(\nu) = k) = \frac{2}{k} \nu^{*k}(k - 2), \qquad k \geq 2. \tag{3}$$

This invites for a probabilistic explanation and provides the incentive for the present work.

Our approach for relating (2) to (3) stems from the fact that solutions to the classical Smoluchowski's coagulation equations (without restriction on the number of aggregations) appear as the hydrodynamical limit of certain stochastic coalescent models introduced by Marcus and

Lushnikov. In some loose sense, the latter describe the microscopic random dynamics of the particle system when the macroscopic evolution is governed by Smoluchowski's coagulation equations. This important feature has been established rigorously by Norris [18]. We also refer to Section 5.2.1 in [4] for an elementary approach in the special case of the multiplicative kernel, as the latter bears an obvious similarity with (1). On the other hand, it is well-known that the multiplicative coalescent is naturally related to the size of the connected components in the random graph model of Erdös and Rényi, see in particular the remarkable paper by Aldous [1]. This leads us to consider an extension of the random graph model where the sequence of degrees of vertices is given, and which is known as the *configuration model*. Loosely speaking, the configuration model is constructed by an elementary 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, in the sense that there may exist 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. This induces a natural partition of the set of vertices into clusters, i.e. connected components.

Since its introduction independently by Bollobás [7] and Wormald [20] (see also Bender and Canfield [3]), this model has been studied in the mathematical literature by many authors. We refer e.g. to [16] for an interesting review of applications of this and other random graph models to some real life network systems. The main known results chiefly concern asymptotics when the number of vertices is large and the empirical measure of the degrees of vertices converges. In particular, Molloy and Reed [14] have determined the critical parameter for the existence of a giant component; see also [15] and [17]. In different directions, van der Hofstad, Hooghiemstra and co-authors [10, 12, 13] have made deep contributions to the study of distances between vertices in such random graphs, while Britton *et al.* [8] used the configuration model to produce large random simple graphs with pre-described asymptotic degree distribution.

If we neglect the appearance of multiple edges, loops or cycles which do not contribute to aggregation of clusters, the configuration model may serve as a stochastic counterpart to the deterministic evolution of concentrations in the variant (1) of Smoluchowski's coagulation equations. This leads us to investigate the size of typical clusters, and more generally their combinatorial structures. Roughly speaking, the main result of this work is a hydrodynamical limit theorem for the empirical distribution of the shapes of clusters rooted at a generic stub. The limit is expressed in terms of a pair of Galton-Watson trees which are connected by an extra edge between the two roots. In particular, this yields the probabilistic explanation of the formal similarity between the solution (2) and Dwass formula (3).

Let us now present some heuristics which are close to some of those that have already been used in the literature on configuration models to relate the latter to Galton-Watson processes; see in particular [14] and [12]. Imagine that we pick a stub uniformly at random; the degree of

3

the vertex to which this stub is appended has then the size-biased law of the degree of a typical vertex. We then pick a second stub uniformly at random to create the first edge. Informally, when the number of vertices is large, the degree of the vertex to which the second stub is appended has again the size-biased law and is essentially independent of the first. These two vertices should be viewed as the ancestors of two growing populations, where, by induction, individuals beget independently and with a reproduction law given by the distribution of the *outer* degree of a size-biased vertex. When the reproduction law is critical or sub-critical, the Galton-Watson process eventually becomes extinct, and extinction occurs before any loop, multiple edge or cycle arises in that cluster of the configuration model. This suggests that the combinatorial structure of a typical (not too large) cluster could be described as a pair of independent Galton-Watson trees which are connected by an additional edge between the two roots. More precisely, the reproduction law should be given by the size-biased degree of a typical vertex, shifted by one unit, because the number of children corresponds to the outer-degree of the vertex.

The present work can be viewed as a companion to the recent paper [6], in which we also identify in terms of certain Galton-Watson trees the limiting empirical distribution of random structures that appear in a toy model for polymerization. More precisely, we consider in [6] a system of grabbing particles, where particles consist in monomers having a certain number of arms. Arms are activated successively uniformly at random, and each time an arm is activated, it grabs a particle uniformly at random amongst those which have not been previously grabbed and do not belong either to its own cluster. The main result of [6] is that when the initial number of particles is large and the numbers of arms are given by i.i.d. random variables with mean less than 1, then the empirical distribution of the shapes of polymers is closed to that induced by a Galton-Watson tree with a single ancestor and reproduction law given by the distribution of the number of arms of a typical monomer.

The plan of this work is as follows. The next section is devoted to preliminaries on configuration models, the combinatorial structure of planar rooted trees, and Galton-Watson processes. The emphasis is put on planar structures and their codings by the sequence of degrees via breadth-first search. The main result on the empirical distribution of the structures of rooted clusters in large random configurations is stated in Section 3 and then proved by explicit first and second moments estimates. Finally Section 4 is devoted to some applications. We shall point at certain invariance properties of Galton-Watson trees under random re-rooting, and conclude by explaining the striking resemblance between the formulas (2) and (3).

4

2 Preliminaries

2.1 Pairings, configurations and clusters

The aim of this work is to relate random configuration models to Galton-Watson trees, and as the latter have a natural planar structure, we shall introduce the former in planar setting which is tailored for our purposes. In this direction, we should imagine particles as planar star-shaped objects consisting in a vertex to which a certain number of stubs are appended.

Formally, we consider some finite set $\mathcal{V}$ of vertices and a map $d : \mathcal{V} \to \mathbb{N}^*$ where $d(v)$ represents the *degree* of the vertex v, that is number of stubs attached to v. We denote by $\mathcal{S} = \mathcal{S}(\mathcal{V}, d)$ the set of stubs and shall suppose for the sake of simplicity that the total number of stubs

$$S := \#\mathcal{S} = \sum_{v \in \mathcal{V}} d(v)$$

is even; otherwise we may always decide to add a new stub to some vertex (or to add a vertex with a single stub). We call a partition of $\mathcal{S}$ into $S/2$ pairs a *pairing of stubs* and write $\Pi(\mathcal{S})$ for the set of pairings of stubs. We first point at the following elementary facts.

Lemma 1 (i) *The cardinal of $\Pi(\mathcal{S})$ is given by*

$$\#\Pi(\mathcal{S}) = \frac{S!}{(S/2)!} 2^{-S/2} = \prod_{i=1}^{S/2} (S - 2i + 1) \,.$$

(ii) *Consider a partition of $\mathcal{V}$ into two subsets $\mathcal{V}_1, \mathcal{V}_2$ such that $S_1 := \sum_{v \in \mathcal{V}_1} d(v)$ and $S_2 := \sum_{v \in \mathcal{V}_2} d(v)$ are even numbers. Set $\mathcal{S}_1 := \mathcal{S}(\mathcal{V}_1, d)$ and $\mathcal{S}_2 := \mathcal{S}(\mathcal{V}_2, d)$. Then the map*

$$(\pi_1, \pi_2) \longrightarrow \pi_1 \sqcup \pi_2$$

is a bijection from $\Pi(\mathcal{S}_1) \times \Pi(\mathcal{S}_2)$ to the subset of $\Pi(\mathcal{S})$ consisting in pairings π such that there are no pairs $\{s_1, s_2\}$ in π formed by a stub s_1 attached to a vertex in $\mathcal{V}_1$ and a stub s_2 attached to a vertex in $\mathcal{V}_1$.

Proof: Indeed, a generic pairing can be obtained by enumerating the stubs by $\{1, \dots, S\}$ and then pairing the stubs according to the couples $(1, 2), (3, 4), \dots, (S - 1, S)$. There are $S!$ possible enumerations and the mapping is $(S/2)! 2^{S/2}$ on 1, where $(S/2)!$ accounts for the number of permutations of the $S/2$ couples $(2i - 1, 2i)$, and $2^{S/2}$ for the number of ways $S/2$ unordered pairs can be ordered into couples. This establishes the first claim. The second is obvious. □

We then form *edges* $e = \{v, v'\}$ with $v, v' \in \mathcal{V}$ by joining the tips of pairs of stubs $\{s, s'\}$, where s (respectively, s') is appended to v (respectively, to v'). We stress that an edge is unoriented, that it can be a *loop* (i.e. the two vertices v and v' defining an edge may coincide), and that the same edge may appear by joining different pairs of stubs. Each pairing of stubs π yields a *configuration* $\gamma(\pi)$, that is the family of the $S/2$ edges induced by the pairing. Note that there may be *multiple edges*; the same edge is repeated in $\gamma(\pi)$ as many times as it arises by joining different pairs of stubs in π. We also stress that the map $\pi \to \gamma(\pi)$ is not injective.

We view an edge which is not a loop as an elementary path connecting two different vertices, so a configuration $\gamma(\pi)$ on $(\mathcal{V}, d)$ naturally induces a partition of $\mathcal{V}$ into connected components. Endowing a given connected component with the restriction of $\gamma(\pi)$ to the set of edges formed by pairs of vertices in that component, we obtain a *cluster*.

2.2 Planar rooted trees and their structures

Lemma 1(ii) enables us to reduce the study of a given cluster to that of pairings $\pi \in \Pi(\mathcal{S})$ such that the entire set of vertices $\mathcal{V}$ is connected for the configuration $\gamma(\pi)$. We shall therefore focus on that case in this section. Recall that a *cycle* is a sequence of $\ell \geq 3$ distinct vertices, say $v_1, \ldots, v_\ell$, such that there exists an edge connecting v_j and v_{j+1} for every $j = 1, \ldots, \ell - 1$ and also an edge connecting v_ℓ and v_1.

A configuration $\gamma(\pi)$ that connects $\mathcal{V}$ is called a *tree* if it contains no loops, no multiple edges, and no cycle. Note that this can occur only when $S/2 = \#\mathcal{V} - 1$. Because particles (i.e. vertices and the stubs that are appended) can be viewed as planar objects, we may think of tree-configurations as planar structures, in the sense that they can be represented in the plane in such a way that edges are line segments which do no cross, by attributing lengths to the edges in an appropriate manner. Throughout this section, we assume that $\#\mathcal{V} = k$ and that the configuration $\gamma(\pi)$ is a tree; in particular $\gamma(\pi)$ consists in $k - 1$ edges and $S = 2(k - 1)$.

To describe precisely the shape, that is the combinatorial structure, of a tree, we need to specify an origin and an orientation. For this, we distinguish a stub s and call it the *root*. This stub is appended to a certain vertex v that we use as the origin. Distinguishing s also enables us to order all the stubs attached to v by deciding that the first stub is s and the next (if any) are ranked clockwise from that one. Further, for every vertex $v' \neq v$ in that tree, we distinguish the stub appended to v' that points at the origin v. This provides a natural order on the set of stubs appended to any given vertex of the tree, and thus enables the use of *breadth-first search* to enumerate the vertices of the tree.

Specifically, set $s_1 = s$ and $v_1 = v$, define s_{i+1} as the stub that is paired with the i-th stub appended to v_1 for $i = 1, \ldots, d(v_1)$, and write v_{i+1} for vertex to which s_{i+1} is appended.

We should think of $v_2, \ldots, v_{d(v_1)+1}$ as the children of v_1. The stub s_2 is chosen as the first of the stubs appended to v_2, thus it is the unique stub pointing at the origin and the other stubs attached to v_2 are ranked clockwise from s_2 and point at the children of v_2 (i.e. the vertices at distance 2 from the origin v_1 and at distance 1 from v_2). We denote these $d(v_2) - 1$ children by $v_{d(v_1)+2}, \ldots, v_{d(v_1)+d(v_2)+1}$, and continue with the next children $v_3, \ldots, v_{d(v_1)+1}$ of v_1 is an obvious way. Then we proceed with indexing the third generation of vertices, in the order which is naturally induced by the indexation of the second generation, and so on. See the figure below.
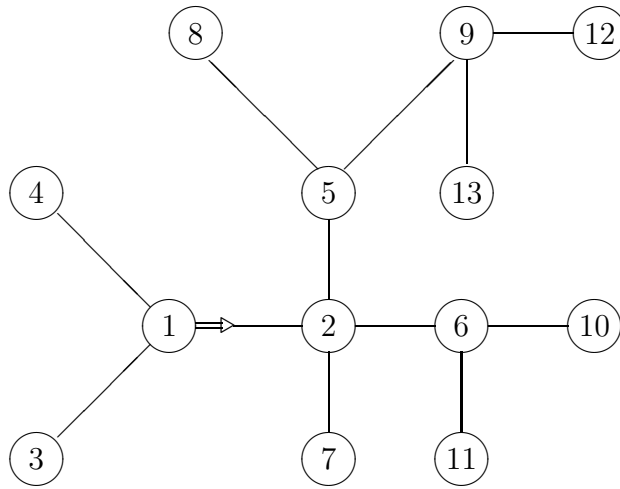

Figure 1 : *Enumeration by breadth-first search of the vertices of a planar tree rooted at the stub $\Rightarrow$. The degree sequence is $(3, 4, 1, 1, 3, 3, 1, 1, 3, 1, 1, 1, 1)$.*

We write d_i for the degree of the i-th vertex. We stress that for $2 \leq i \leq k$, the outer-degree of v_i, i.e. the number of stubs appended to v_i that point away from the origin, is $d_i - 1$. It is well-known that the sequence of degrees $\mathbf{d} = (d_1, \ldots, d_k)$ fulfills

$$\min\{j \geq 1 : d_1 + \cdots + d_j = 2(j - 1)\} = k, \tag{4}$$

and characterizes a unique planar rooted tree structure. Conversely, any finite sequence $\mathbf{d} = (d_1, \ldots, d_k)$ such that (4) holds encodes a planar rooted tree structure with k vertices. We write $\mathbf{D}$ for the set of sequences $\mathbf{d} = (d_1, \ldots, d_k)$ which fulfill (4), where the lenght $k \in \mathbb{N}^*$ is arbitrary, and think of the set $\mathbf{D}$ of sequences of degrees as the set of structures of planar rooted trees. We refer for instance to Section 6.2 in Pitman [19] for details.

We now summarize this discussion, introducing first some terminology for convenience. A bijection $\{1, \ldots, k\} \to \mathcal{V}$ can be represented as a sequence $\mathbf{v} = (v_1, \ldots, v_k)$ of distinct vertices

and will be referred to as an *enumeration* of $\mathcal{V}$. We also call a map $\varsigma : \mathcal{V} \to \mathcal{S}$ that associates to each vertex $v \in \mathcal{V}$ a stub $s \in \mathcal{S}$ appended to that vertex a *selection of stubs*. For every pairing $\pi \in \Pi(\mathcal{S})$ such that the configuration $\gamma(\pi)$ on $\mathcal{V}$ is a tree and every choice of a distinguished stub $s \in \mathcal{S}$, breadth first search yields a unique enumeration $\mathbf{v} = (v_1, \ldots, v_k)$ of $\mathcal{V}$ such that the sequence $d(\mathbf{v}) = (d(v_1), \ldots, d(v_k))$ belongs to $\mathbf{D}$ (i.e. fulfills (4)), and a unique a selection of stubs ς. The map

$$(\pi, s) \longrightarrow (\mathbf{v}, \varsigma)$$

is bijective. More precisely, we recover the pairing π and the root stub s by first constructing the planar rooted tree structure associated to $\mathbf{d} = (d(v_1), \ldots, d(v_k))$, and then placing the vertices $v_1, \ldots, v_k$ on this structure in the order induced by the breadth first search. The first stub appended to v_1 is $s = \varsigma(v)$, and for every $i = 2, \ldots, k$, $\varsigma(v_i)$ is the stub appended to v_i which points at the origin v_1. This determines the pairing π.

In order to record this analysis, it is convenient to introduce the multinomial coefficient

$$M(\mathcal{V}, d) := \binom{k}{\ell_1, \ldots, \ell_j} = \frac{k!}{\ell_1! \cdots \ell_j!}, \tag{5}$$

where j is the number of different values, say $x_1, \ldots, x_j$, occurring in the family $(d(v) : v \in \mathcal{V})$, and ℓ_i the number of occurrences of the value x_i in that family. For every structure $\mathbf{d} \in \mathbf{D}$, we say that $\mathbf{d}$ is *compatible* with $(\mathcal{V}, d)$ if there is at least an enumeration $\mathbf{v} = (v_1, \ldots, v_k)$ of $\mathcal{V}$ such that $\mathbf{d} = (d(v_1), \ldots, d(v_k))$, that is if and only if the sequence $\mathbf{d}$ takes the same values with the same multiplicity as the family $(d(v) : v \in \mathcal{V})$. The following statement should now be plain.

Lemma 2 *Suppose that $\#\mathcal{V} = k$ and $S = 2(k-1)$. Fix a rooted planar tree structure $\mathbf{d} = (d_1, \ldots, d_k) \in \mathbf{D}$. If $\mathbf{d}$ is compatible with $(\mathcal{V}, d)$, then the number of pairs $(\pi, s) \in \Pi(\mathcal{S}) \times \mathcal{S}$ for which the configuration $\gamma(\pi)$ is a tree with structure $\mathbf{d}$ when rooted at s equals*

$$M(\mathcal{V}, d) \prod_{v \in \mathcal{V}} d(v).$$

Otherwise (i.e. if $\mathbf{d}$ is not compatible), this number is 0.

We stress that all the rooted planar tree structures which are compatible with $(\mathcal{V}, d)$ are thus equally likely to occur if we choose the pair $(\pi, s) \in \Pi(\mathcal{S}) \times \mathcal{S}$ uniformly at random. In the same vein, it may be also interesting to point at the following simple formula, even though it will not be used in this paper .

Proposition 1 *Suppose that $\#\mathcal{V} = k$ and $S = 2(k-1)$. The number of pairings $\pi \in \Pi(\mathcal{S})$ for which the configuration $\gamma(\pi)$ is a tree, is*

$$(k-1)! \prod_{v \in \mathcal{V}} d(v),$$

Proof: To establish the formula, we simply need to calculate the number of enumerations $\mathbf{v}$ of $\mathcal{V}$ for which the sequence $(d(v_1), \ldots, d(v_k))$ corresponds to some rooted planar tree structure. Recall from the ballot theorem (see, e.g., Lemma 6.1 in [19]) that for each enumeration $\mathbf{v}$, there is a unique cyclic permutation σ of $\{1, \ldots, k\}$ such that $(d(v_{\sigma(1)}), \ldots, d(v_{\sigma(k)}))$ fulfills (4). This shows that this number is $(k-1)!$. □

2.3 Galton-Watson trees with two ancestors

We consider now a probability measure ν on $\mathbb{N}$ and associate to ν a measure on $\mathbf{D}$ by

$$\mathbb{GW}_2^\nu(\mathbf{d}) = \prod_{i=1}^{k} \nu(d_i - 1), \tag{6}$$

where $\mathbf{d} = (d_1, \ldots, d_k)$ denotes a generic rooted planar tree structure.

The measure $\mathbb{GW}_2^\nu$ has a simple interpretation in terms of Galton-Watson branching processes, and is in fact a sub-probability. More precisely, consider a Galton-Watson process with reproduction law ν and started from two ancestors. The process can be represented on the upper-half plane, where the individuals at generation $\ell \in \mathbb{N}$ lie on horizontal line $y = \ell$, in an order consistent with that of their respective parents, so that the edges (line-segments) linking parents to children do not cross each other. We further connect the two ancestors by an additional edge, and distinguish the stub attached to the left-most ancestor that thus points at the right-most ancestor. This enables us to list individuals (vertices) by breadth first search just as in the preceding section. Observe that the degree of the left-most ancestor (i.e. the origin) is distributed as $1 + \xi$ where ξ is a random variable with law ν, whereas the outer-degrees of the other individuals (i.e. their numbers of children) are given by independent copies of ξ.

The event when the total population is finite has probability one if and only if the reproduction law ν is critical or subcritical, i.e. $\sum_{i \in \mathbb{N}} i\nu(i) \leq 1$, and $\nu \neq \delta_1$. Restricting our attention to this event, the structure of this planar rooted tree is a random variable in $\mathbf{D}$ which has distribution $\mathbb{GW}_2^\nu$ and is defective in the supercritical case.

Remark. In the case when ν is the Poisson distribution with parameter $p \leq 1$, then it is easily checked that the law $\mathbb{GW}_2^\nu$ also describes the law of the genealogical tree of a Galton-Watson

process with reproduction law ν, started from a *single* ancestor, and conditioned to have size at least 2.

3 A limit theorem for typical rooted clusters

For each fixed integer n, we consider a set $\mathcal{V}_n$ of n vertices and a function $d_n : \mathcal{V}_n \to \mathbb{N}^*$ that specifies the number of stubs appended to each vertex. We introduce the empirical distribution of the number of stubs

$$\mu_n(i) := \frac{1}{n} \#\{v \in \mathcal{V}_n : d_n(v) = i\}, \qquad i \in \mathbb{N}^* .$$

We write

$$S_n := \sum_{v \in \mathcal{V}_n} d_n(v) = n \sum_{i=1}^{\infty} i \mu_n(i)$$

for the total number of stubs, assuming for simplicity that this quantity is even. Our basic assumption is that the limit

$$\lim_{n \to \infty} \mu_n(i) := \mu(i) \tag{7}$$

exists for every $i \geq 1$, and that the average number of stubs

$$n^{-1} S_n = \sum_{i=1}^{\infty} i \mu_n(i)$$

converges as $n \to \infty$ to the first moment of μ, i.e.

$$\lim_{n \to \infty} \sum_{i=1}^{\infty} i \mu_n(i) = \sum_{i=1}^{\infty} i \mu(i) := m < \infty . \tag{8}$$

We also denote by μ^* the probability measure on $\mathbb{N}^*$ which is obtained from μ by size-biased sampling, that is

$$\mu^*(i) := \frac{i \mu(i)}{m}, \qquad i \in \mathbb{N}^* .$$

A standard application of Scheffé's lemma shows that (7) can then be re-enforced to

$$\lim_{n \to \infty} i \mu_n(i) \frac{n}{S_n} = \mu^*(i) \qquad \text{in } L^1(\mathbb{N}^*). \tag{9}$$

Finally, we introduce the probability measure ν on $\mathbb{N}$ induced from μ^* by the shift $i \to i - 1$ from $\mathbb{N}^*$ to $\mathbb{N}$, viz.

$$\nu(i) = \mu^*(i + 1), \qquad i \geq 0 .$$

We write $\mathcal{S}_n$ for the set of stubs appended to vertices in $\mathcal{V}_n$. We pick a pairing $\pi \in \Pi(\mathcal{S}_n)$ uniformly at random, and denote by $\Gamma_n := \gamma(\pi)$ the resulting random configuration on $(\mathcal{V}_n, d_n)$. For every stub $s \in \mathcal{V}_n$, if the cluster of Γ_n which contains s is a tree, then T_s denotes the combinatorial structure which results from rooting that tree at the stub s (see Section 2.2), and otherwise, we decide that $T_s = \varnothing$.

We are interested in the random variable

$$\rho_n(\mathbf{d}) := \frac{1}{S_n} \#\{s \in \mathcal{S}_n : T_s = \mathbf{d}\}, \qquad \mathbf{d} \in \mathbf{D}$$

which counts the proportion of stubs s such that the cluster rooted at s induced by Γ_n is a tree with structure $\mathbf{d}$. Similarly, we write

$$\rho_n(\varnothing) := \frac{1}{S_n} \#\{s \in \mathcal{S}_n : T_s = \varnothing\}$$

for the proportion of stubs s such that the cluster containing s induced by Γ_n is not a tree. The collection $(\rho_n(\mathbf{d}) : \mathbf{d} \in \mathbf{D})$ should thus be viewed as a variant of the empirical measure of tree-clusters. We now able to state our main asymptotic result on large random configurations.

Theorem 1 *Assume that* (7) *and* (8) *hold. Then for every planar rooted tree configuration* $\mathbf{d} \in \mathbf{D}$, *the following limit holds in* $L^2(\mathbb{P})$:

$$\lim_{n \to \infty} \rho_n(\mathbf{d}) = \mathbb{GW}_2^\nu(\mathbf{d}) .$$

If we further suppose that

$$\sum_{i=1}^{\infty} i(i-2)\mu(i) \le 0 , \tag{10}$$

and also exclude the degenerate case when μ *is the Dirac point mass at 2, then*

$$\lim_{n \to \infty} \rho_n(\varnothing) = 0 \qquad in \ L^1(\mathbb{P}).$$

The condition (10) plays an important part for random configuration models. According to a well-known result due to Molloy and Reed [14], when (10) fails (assuming also some further technical conditions), then there is some constant $c > 0$ such that with probability one, the random configuration Γ_n contains almost surely a cluster of size at least cn when n is sufficiently large. The size of this giant component is estimated in [15]. At the opposite, when (10) holds with a strict inequality (again assuming some further technical conditions), Molloy and Reed [14] have shown that with probability one, the random configuration Γ_n contains at most $n^{1/4}$ cycles and no cluster of size at least $n^{1/4}$ whenever n is sufficiently large. Note that in the

11

critical case when (10) is an equality, Theorem 1 implies that the probability that there is a cluster of size at least εn tends to 0 for any $\varepsilon > 0$, because $\mathbb{GW}_2^\nu$ is a probability measure on **D**.

The proof of Theorem 1 relies on asymptotics for the first and second moments of $\rho_n(\mathbf{d})$. We first state:

Lemma 3 *We have*

$$\lim_{n \to \infty} \mathbb{E}(\rho_n(\mathbf{d})) = \mathbb{GW}_2^\nu(\mathbf{d})$$

for every $\mathbf{d} \in \mathbf{D}$.

Proof: Let the structure $\mathbf{d} = (d_1, \ldots, d_k)$ have size $k \geq 2$. We write $\mathcal{V}'$ for a generic subset of $\mathcal{V}_n$ with k vertices and $\mathcal{S}'$ for the set of stubs in $\mathcal{S}_n$ which are appended to vertices in $\mathcal{V}'$. There are two cases.

If the unordered families of degrees $\{d(v') : v' \in \mathcal{V}'\}$ and $\{d_i : 1 \leq i \leq k\}$ do no coincide (recall that in such families, numbers are repeated according to their multiplicity), then there is no pairing of stubs for which the vertices of $\mathcal{V}'$ are those of a tree-cluster with structure $\mathbf{d}$ when properly rooted. We say that $\mathcal{V}'$ is *bad*.

Otherwise, we say that $\mathcal{V}'$ is *good*. Introduce the set G' of couples $(s, \pi) \in \mathcal{S}' \times \Pi(\mathcal{S}_n)$ such that the cluster rooted at s induced by the configuration $\gamma(\pi)$ is a tree whose set of vertices coincides with $\mathcal{V}'$ and has structure $\mathbf{d}$. The cardinal of G' can then be computed by combining Lemmas 1 and 2. Since $\#\mathcal{S}' = 2(k-1)$, one gets

$$\#G' = M(\mathbf{d}) \frac{(S_n - 2(k-1))!}{(S_n/2 - k + 1)!} 2^{-S_n/2 + k - 1} \prod_{i=1}^{k} d_i \,, \tag{11}$$

where $M(\mathbf{d})$ denotes the multinomial coefficient

$$M(\mathbf{d}) := \binom{k}{\ell_1, \ldots, \ell_j} = \frac{k!}{\ell_1! \cdots \ell_j!} \,,$$

with j the number of different values in the sequence $\mathbf{d}$ and ℓ_i the number of occurrences in $\mathbf{d}$ of the i-th value for $1 \leq i \leq j$.

So it remains to estimate the number of good subsets $\mathcal{V}'$ with k vertices, and for this we use a probabilistic argument. We sample uniformly at random k vertices in $\mathcal{V}_n$, say, $v_1, \ldots, v_k$, successively and without replacement. It should be plain from the hypothesis (7) that when $n \to \infty$, the k-tuple of degrees $(d_n(v_1), \ldots d_n(v_k))$ converges in distribution to the k-tuple formed by i.i.d. variables with law μ; in particular the probability that $(d_n(v_1), \ldots, d_n(v_k)) = \mathbf{d}$

tends to $\prod_{i=1}^{k} \mu(d_i)$ as $n \to \infty$. We readily deduce that the probability that the (unordered) family $\{d(v_1), \ldots, d(v_k)\}$ is good converges as $n \to \infty$ to

$$\frac{k!}{M(\mathbf{d})} \prod_{i=1}^{k} \mu(d_i) \,.$$

As there are $n!/(n-k)! \sim n^k$ k-tuples of distinct vertices in $\mathcal{V}_n$ and as the map that transforms a k-tuple into an unordered set is $k!$ to 1, we conclude that the number of good subsets in $\mathcal{V}_n$ is equivalent for large n to

$$\frac{n^k}{M(\mathbf{d})} \prod_{i=1}^{k} \mu(d_i) \,. \tag{12}$$

Recall from Lemma 1(i) that

$$\# \left(\mathcal{S}_n \times \Pi(\mathcal{S}_n) \right) = S_n \frac{S_n!}{(S_n/2)!} 2^{-S_n/2} \,,$$

and that $d_i \mu(d_i) = m\nu(d_i - 1)$, by definition. Putting the pieces together, we find

$$
\begin{aligned}
\mathbb{E}(\rho_n(\mathbf{d})) &\sim \frac{n^k}{S_n} \frac{(S_n - 2(k-1))!(S_n/2)!}{S_n!(S_n/2 - k + 1)!} 2^{k-1} \prod_{i=1}^{k} (d_i \mu(d_i)) \\
&\sim \frac{n^k}{S_n} S_n^{-2(k-1)} (S_n/2)^{k-1} 2^{k-1} \prod_{i=1}^{k} (m\nu(d_i - 1)) \\
&= \frac{(nm)^k}{S_n^k} \prod_{i=1}^{k} \nu(d_i - 1) \,.
\end{aligned}
$$

By (6) and (8), this completes the proof. $\qquad \square$

Lemma 3 essentially means that if we pick a stub s uniformly at random in $\mathcal{S}_n$ and independently of the random configuration Γ_n, then the conditional distribution of the combinatorial structure of the random cluster rooted at s given the event that this cluster is a tree, converges weakly as $n \to \infty$ to the Galton-Watson law $\mathbb{GW}_2^{\nu}$. Theorem 1 is a much stronger statement that involves the empirical distribution of structures of clusters, and requires second moment estimates.

Lemma 4 *We have*

$$\lim_{n \to \infty} \mathbb{E}((\rho_n(\mathbf{d}))^2) = (\mathbb{GW}_2^{\nu}(\mathbf{d}))^2$$

for every $\mathbf{d} \in \mathbf{D}$.

Proof: The argument is similar to that of Lemma 3; in particular we shall use the same

notation and terminology. We start from the expression

$$\mathbb{E}((\rho_n(\mathbf{d}))^2) = S_n^{-2}\mathbb{E}\left(\#\{(s', s'') \in \mathcal{S}_n \times \mathcal{S}_n : T_{s'} = T_{s''} = \mathbf{d}\}\right).$$

Let $\mathcal{V}'$ and $\mathcal{V}''$ two generic subsets of $\mathcal{V}_n$, both with k vertices, and write $\mathcal{S}'$ (respectively, $\mathcal{S}''$) for the set of stubs in $\mathcal{S}_n$ which are appended to vertices in $\mathcal{V}'$ (respectively, $\mathcal{V}''$). Note that for every stubs $s' \in \mathcal{S}'$ and $s'' \in \mathcal{S}''$, the identity $T_{s'} = T_{s''} \neq \varnothing$ can occur only if $\mathcal{V}'$ and $\mathcal{V}''$ are both good and, either coincide or are disjoint.

We first consider the situation when $\mathcal{V}' = \mathcal{V}''$. Recall that for any $s' \in \mathcal{S}'$, if $T_{s'} = \mathbf{d}$, then $\mathcal{S}' = \mathcal{S}''$has exactly $2(k-1)$ stubs. It follows from the proof of Lemma 3 that the number of triplets $(s', s'', \pi) \in \mathcal{S}' \times \mathcal{S}'' \times \Pi(\mathcal{S}_n)$ such that the cluster rooted at s' induced by the configuration $\gamma(\pi)$ is a tree whose set of vertices coincides with $\mathcal{V}'$ and $T_{s'} = T_{s''} = \mathbf{d}$, is bounded from above by

$$2(k-1)M(\mathbf{d})\frac{(S_n - 2(k-1))!}{(S_n/2 - k + 1)!}2^{-S_n/2+k-1}\prod_{i=1}^{k}d_i\,;$$

see (11). Multiplying this by the number of good subsets $\mathcal{V}'$ in $\mathcal{V}_n$, that is approximatively by (12), we get a quantity which is small compared to

$$\#(\mathcal{S}_n \times \mathcal{S}_n \times \Pi(\mathcal{S}_n)) = S_n^2\frac{S_n!}{(S_n/2)!}2^{-S_n/2}$$

when $n \to \infty$. We conclude that in the evaluation of $\mathbb{E}((\rho_n(\mathbf{d}))^2)$, the contribution of pairs of stubs (s', s'') that belong to the same cluster becomes asymptotically negligible.

Next we consider the situation when $\mathcal{V}'$ and $\mathcal{V}''$ are good and disjoint. By calculations similar to those that yield (12) in the proof of Lemma 3, we get that the number of good disjoint pairs of subsets $(\mathcal{V}', \mathcal{V}'')$ in $\mathcal{V}_n$ is equivalent for large n to

$$\frac{n^{2k}}{M(\mathbf{d})^2}\left(\prod_{i=1}^{k}\mu(d_i)\right)^2.$$

It then follows from Lemmas 1 and 2 that the number of triplets $(s', s'', \pi) \in \mathcal{S}' \times \mathcal{S}'' \times \Pi(\mathcal{S}_n)$ such that $T_{s'} = T_{s''} = \mathbf{d}$ and the stubs s' and s'' belong to disjoint clusters is close to

$$n^{2k}\frac{(S_n - 4(k-1))!}{(S_n/2 - 2k + 2)!}2^{-S_n/2+2k-2}\left(\prod_{i=1}^{k}d_i\mu(d_i)\right)^2.$$

Putting the pieces together yields the estimate

$$\mathbb{E}((\rho_n(\mathbf{d}))^2 \quad \sim \quad \frac{n^{2k}}{S_n^2} \frac{(S_n - 4(k-1))!(S_n/2)!}{S_n!(S_n/2 - 2k + 2)!} 2^{2k-2} \left(\prod_{i=1}^{k} d_i \mu(d_i) \right)^2$$

$$\sim \quad \frac{n^{2k}}{S_n^2} S_n^{-4(k-1)} (S_n/2)^{2k-2} 2^{2k-2} \left(\prod_{i=1}^{k} (m\nu(d_i - 1)) \right)^2$$

$$\sim \quad \left(\prod_{i=1}^{k} \nu(d_i - 1) \right)^2 .$$

By (6), this shows our claim. $\qquad\square$

We are now able to establish Theorem 1.

Proof of Theorem 1: Combining Lemmas 3 and 4, we see that the variance of $\rho_n(\mathbf{d})$ tends to 0 as $n \to \infty$, which establishes the first claim. Assume now further that (10) holds and that $\mu \neq \delta_2$. Equivalently, this means that the reproduction law ν of the Galton-Watson process is critical or sub-critical, and is not the Dirac mass at 1. So extinction occurs a.s. and

$$\sum_{\mathbf{d} \in \mathbf{D}} \mathbb{GW}_2^\nu(\mathbf{d}) = 1 .$$

As

$$\rho_n(\varnothing) = 1 - \sum_{\mathbf{d} \in \mathbf{D}} \rho_n(\mathbf{d}) ,$$

Fatou lemma entails our second assertion. $\qquad\square$

4 Some applications

In this Section, we shall develop some consequences of our main result. Recall that Theorem 1 implies that if one selects a stub uniformly at random and independently of a large random configuration that fulfills the conditions there, then the structure of the cluster rooted at that stub has asymptotically the distribution $\mathbb{GW}_2^\nu$. This hints at an interesting property of invariance of such Galton-Watson trees under uniform random re-rooting. Recall the construction of the structure of a planar tree rooted at some stub as it has been presented in Section 2.2; Figure 2 below should explain better than words what is meant by re-rooting a rooted planar tree at some stub.

Corollary 1 *Suppose that ν is a critical or subcritical probability measure on $\mathbb{N}$ with $\nu \neq \delta_1$. Let D be a random rooted planar tree structure with distribution $\mathbb{GW}_2^\nu$. Conditionally on D,*

select one of the $2(|D|-1)$ stubs of D uniformly at random, and denote by D' the new structure obtained from D by re-rooting at that stub. Then D' has again the law $\mathbb{GW}_2^\nu$.
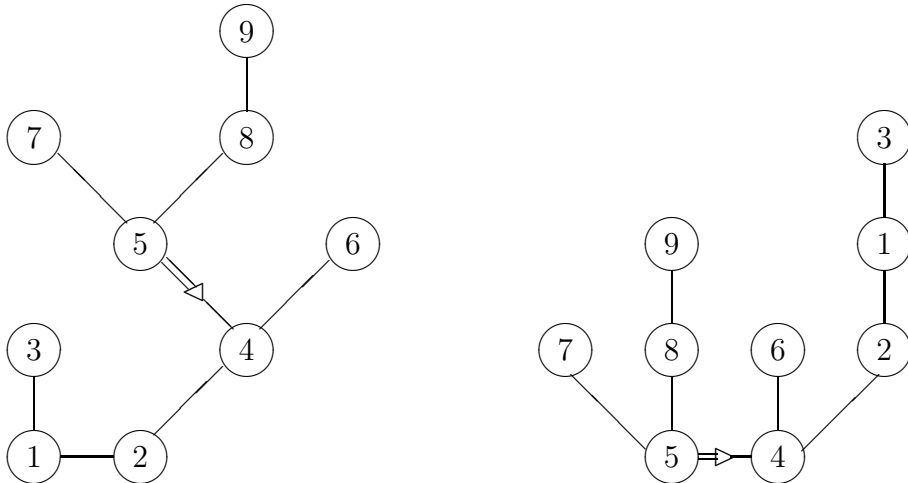

Figure 2 : *Two genealogical trees, both with two ancestors lying at the lowest level.*
The left-most ancestor serves as the origin, the root-stub pointing at the right-most ancestor.
The tree on the right is the image of the tree on the left by re-rooting at the stub $\Rightarrow$.
Vertices are labeled by breadth first order before re-rooting.

Proof: Re-rooting has no effect on the degree of a vertex, so we only need to verify the statement for the conditional law of the Galton-Watson genealogical tree with two ancestors given the unordered family of the degrees of vertices.

Fix some unordered family, say Δ, of k positive integers (with possible repetitions), which add up to $2(k-1)$ and such that $\nu(\delta-1) > 0$ for any integer δ in that family. Denote by $\mathbf{D}(\Delta)$ the subset of rooted planar tree structures corresponding to some ordering of Δ. We see from (6) that the conditional law $\mathbb{GW}_2^\nu(\cdot \mid \mathbf{D}(\Delta))$ is simply the uniform distribution on $\mathbf{D}(\Delta)$.

Next consider the random configuration on a set k vertices with degree family Δ that is induced by uniform random pairing, given that this configuration is a tree. Then root the configuration using some stub that is picked independently and uniformly at random. On the one hand, by construction, the law of the resulting combinatorial structure is obviously invariant by uniform random re-rooting. On the other hand, we see from Lemma 2 that it also coincides with the uniform distribution $\mathbf{D}(\Delta)$. This established our claim. $\qquad\square$

We also refer to the recent work by Haas *et al.* [11] and references therein for a different property of invariance under uniform re-rooting for certain classes of random continuous trees.

It may be interesting to point also at the following avatar of Corollary 1. A planar rooted tree is said *planted* if the degree of the origin, i.e. of the vertex to which the root-stub is appended, is 1. In other words, the combinatorial structure $\mathbf{d} = (d_1, \ldots, d_k)$ fulfills $d_1 = 1$. So a planted Galton-Watson tree describes the genealogy of a population where individuals beget independently with the same reproduction law, except the ancestor who has exactly one child. An easy consequence of Corollary 1 is that in the critical or sub-critical case, the structure of a planted Galton-Watson tree is statistically invariant under re-rooting at a leaf (i.e. a vertex with degree 1) chosen uniformly at random.

We next turn our attention to some quantitative consequences of Theorem 1, denoting for every $k \geq 2$ by $C_n(k)$ the number of clusters of size k in the random configuration Γ_n, i.e. the number of distinct connected components with k vertices in the partition of $\mathcal{V}_n$ induced by Γ_n.

Corollary 2 *Assume that* (7), (8) *and* (10) *hold, and exclude the case when* $\mu = \delta_2$. *We have*

$$\lim_{n \to \infty} \sum_{k=2}^{\infty} k \mathbb{E} \left(\left| n^{-1} C_n(k) - \frac{m}{k(k-1)} \nu^{*k}(k-2) \right| \right) = 0 \,.$$

Proof: Let us introduce first for every $k \geq 2$ the subset $\mathbf{D}_k$ of $\mathbf{D}$ consisting of structures of rooted planar trees $\mathbf{d} = (d_1, \ldots, d_k)$ of lenght k, and recall that according to Dwass [9],

$$\mathbb{GW}_2^{\nu}(\mathbf{D}_k) = \frac{2}{k} \nu^{*k}(k-2) \,,$$

where ν^{*k} stands for the k-th convolution power of ν. As a tree of size k has exactly $2(k-1)$ stubs and $\mathbf{D}_k$ is a finite set, we deduce from Theorem 1 that if we denote by $\tau_n(k)$ the number of clusters which are trees of size k, then

$$\lim_{n \to \infty} \frac{2(k-1)}{S_n} \tau_n(k) = \frac{2}{k} \nu^{*k}(k-2) \,,$$

where the convergence takes place in $L^2(\mathbb{P})$ and for every $k \geq 2$. Then we pick an arbitrary sequence of integers that tends to ∞, from which we can excerpt by a diagonal extraction procedure a subsequence such that with probability one,

$$\lim_{n \rightsquigarrow \infty} \frac{2(k-1)}{S_n} \tau_n(k) = \frac{2}{k} \nu^{*k}(k-2) \qquad \text{for all } k \geq 2,$$

where the notation $n \rightsquigarrow \infty$ means that n tends to infinity along that subsequence.

Then observe that for each n, there is the obvious inequality

$$\sum_{k \geq 2} 2(k-1)\tau_n(k) \leq S_n \,,$$

while

$$\sum_{k \geq 2} \frac{2}{k} \nu^{*k}(k-2) = \sum_{k \geq 2} \mathbb{GW}_2^\nu(\mathbf{D}_k) = 1 \,,$$

since the reproduction law ν of the Galton-Watson process is critical or sub-critical and $\nu \neq \delta_1$. A standard combination of Fatou and Scheffé lemmas entails that

$$\lim_{n \rightsquigarrow \infty} \sum_{k=2}^\infty \mathbb{E}\left(\left| \frac{2(k-1)}{S_n} \tau_n(k) - \frac{2}{k} \nu^{*k}(k-2) \right| \right) = 0 \,.$$

Next, note that $\tau_n(k) \leq C_n(k)$ and $\sum_{k \geq 2} 2(k-1)C_n(k) \leq S_n$ as at least $2(k-1)$ distinct stubs are needed to connect k vertices. It follows that

$$\sum_{k=2}^\infty \mathbb{E}\left(\left| \frac{2(k-1)}{S_n}(C_n(k) - \tau_n(k)) \right| \right)$$
$$= \mathbb{E}\left(\sum_{k=2}^\infty \frac{2(k-1)}{S_n} C_n(k) \right) - \mathbb{E}\left(\sum_{k=2}^\infty \frac{2(k-1)}{S_n} \tau_n(k)) \right)$$
$$\leq 1 - \mathbb{E}\left(\sum_{k=2}^\infty \frac{2(k-1)}{S_n} \tau_n(k)) \right) \,,$$

and we know from above that this quantity tends to 0 as $n \rightsquigarrow \infty$.

This shows that

$$\lim_{n \rightsquigarrow \infty} \sum_{k=2}^\infty \mathbb{E}\left(\left| \frac{2(k-1)}{S_n} C_n(k) - \frac{2}{k} \nu^{*k}(k-2) \right| \right) = 0 \,,$$

and since by the assumption (8), $S_n/n \to m$, we have thus proved that

$$\lim_{n \rightsquigarrow \infty} \sum_{k=2}^\infty k \mathbb{E}\left(\left| \frac{1}{n} C_n(k) - \frac{m}{k(k-1)} \nu^{*k}(k-2) \right| \right) = 0 \,.$$

As the sequence of integers tending to infinity that we started from is arbitrary, this establishes our claim. $\qquad\square$

Corollary 2 provides the explanation for the asymptotic behavior (2) that motivated this work. Specifically, we know from Theorem 1 that when the requirements (7), (8) and (10) are fulfilled, then, roughly speaking, multiple edges, loops or cycles are rare. Roughly speaking, this

means that almost all creation of edges correspond to aggregations of clusters and thus enables us to view the configuration model as a stochastic microscopic version of the terminal state of concentrations with a deterministic evolution governed by the variant (1) of Smoluchowski's coagulation equations. In [5], one assumes that initially all particles are monomers, i.e. consist in isolated vertices to which some stubs are appended. In the notation of the present work (beware that this differs from that in [5]!), the initial concentration of particles with $i \geq 1$ stubs is $m^{-1}\mu(i)$, which is a finite measure on $\mathbb{N}^*$ with unit first moment. In the framework of the random configuration model with n vertices, this corresponds to assuming that particles live in a volume mn and hence the the initial concentration of monomers with i stubs is given by

$$m^{-1}\mu_n(i) = \frac{1}{mn}\#\{v \in \mathcal{V}_n : d_n(v) = i\}, \qquad k \in \mathbb{N}^*.$$

After the random pairing, the concentration of polymers with size k (i.e. clusters with k vertices) is then $(mn)^{-1}C_n(k)$ and Corollary 2 shows that

$$\lim_{n\to\infty}\frac{1}{mn}C_n(k) := c_\infty(0, k) = \frac{1}{k(k-1)}\nu^{*k}(k-2).$$

One has thus recovered (2).

We also note that Corollary 2 solves a problem that has been addressed in Section II.C of [17] by analytic and numerical technics.

References

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

[2] Aldous, D. J. Deterministic and stochastic models for coalescence (aggregation, coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5** (1999), 3-48.

[3] Bender, E. and Canfield, E. The asymptotic number of labelled graphs with given degree sequences. *J. Comb. Theory, Ser. A* **24** (1978), 296-307.

[4] Bertoin, J. *Random Fragmentation and Coagulation Processes.* Cambridge University Press, Cambridge, 2006.

[5] Bertoin, J. Two solvable systems of coagulation equations with limited aggregations. Preprint available at `http://arxiv.org/abs/0806.3677`.

[6] Bertoin, J., Sidoravicius, V. and Vares, M. E. : A system of grabbing particles related to Galton-Watson trees. Preprint available at `http://fr.arXiv.org/abs/0804.0726`.

[7] Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** (1980), 311-316.

[8] Britton, T., Deijfen, M. and Martin-Löf, A. Generating simple random graphs with pre-scribed degree distribution. *J. Stat. Phys.* **124** (2006), 1377-1397.

[9] Dwass, M. The total progeny in a branching process. *J. Appl. Probab.* **6** (1969), 682-686.

[10] van den Esker, H., van der Hofstad, R. and Hooghiemstra, G. Universality for the distance in finite variance random graphs. *J. Stat. Phys.* **133** (2008), 169202.

[11] Haas, B., Pitman, J. and Winkel, M. Spinal partitions and invariance under re-rooting of continuum random trees. To appear in *Ann. Probab.*, available at `http://arxiv.org/abs/0705.3602`.

[12] van der Hofstad, R., Hooghiemstra, G. and van Mieghem, P.. Distances in random graphs with finite variance degrees. *Random Struct. Algorithms* **26** (2005), 76-123.

[13] van der Hofstad, R., Hooghiemstra, G. and Znamenski, D. Distances in random graphs with finite mean and infinite variance degrees. *Electronic J. Probab.* **12** (2007), 703-766.

[14] Molloy, M. and Reed, B. A critical point for random graphs with a given degree sequence. *Random Struct. Algorithms* **6** (1995), 161-179.

[15] Molloy, M. and Reed, B. The size of the giant component of a random graphs with a given degree sequence. *Comb. Probab. Comput.* **7** (1998), 295-305.

[16] Newman, M. E. J. The structure and function of complex networks. *SIAM Rev.* **45** (2003), 167-256.

[17] Newman, M. E. J., Strogatz, S. and Watts, D. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* 64:026118 (2001).

[18] Norris, J. R. Smoluchowski's coagulation equation: uniqueness, non-uniqueness and hy-drodynamic limit for the stochastic coalescent. *Ann. Appl. Probab.* **9** (1999), 78-109.

[19] Pitman, J. *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/`

[20] Wormald, N. C. *Some Problems in the Enumeration of Labelled Graphs*. Doctoral thesis, Newcastle University, 1978.