

# RANDOM TREE-WEIGHTED GRAPHS

LOUIGI ADDARIO-BERRY AND JORDAN BARRETT

ABSTRACT. For each  $n \geq 1$ , let  $d^n = (d^n(i), 1 \leq i \leq n)$  be a sequence of positive integers with even sum  $\sum_{i=1}^n d^n(i) \geq 2n$ . Let  $(G_n, T_n, \Gamma_n)$  be uniformly distributed over the set of simple graphs  $G_n$  with degree sequence  $d^n$ , endowed with a spanning tree  $T_n$  and rooted along an oriented edge  $\Gamma_n$  of  $G_n$  which is not an edge of  $T_n$ . Under a finite variance assumption on degrees in  $G_n$ , we show that, after rescaling,  $T_n$  converges in distribution to the Brownian continuum random tree as  $n \rightarrow \infty$ . Our main tool is a new version of Pitman's additive coalescent [18], which can be used to build both random trees with a fixed degree sequence, and random tree-weighted graphs with a fixed degree sequence. As an input to the proof, we also derive a Poisson approximation theorem for the number of loops and multiple edges in the superposition of a fixed graph and a random graph with a given degree sequence sampled according to the configuration model; we find this to be of independent interest.

## 1. Introduction

By a *rooted tree* we mean a labeled tree  $t = (v(t), e(t))$ , with a distinguished root node denoted  $r(t)$ . A *tree-rooted graph* is a pair  $(g, t, \gamma)$  where  $g = (v(g), e(g))$  is a labeled graph,  $t = (v(t), e(t))$  is a spanning tree of  $g$ , and  $\gamma = uv$  is a distinguished oriented edge with  $\{u, v\} \in e(g) \setminus e(t)$ . We view  $t$  as a rooted tree by setting  $r(t) = u$ .

Throughout this work, we allow our graphs to have multiple edges and loops; in tree-rooted graphs, the root edge is allowed to be a loop. We say  $(g, t, \gamma)$  is *simple* if  $g$  is simple, i.e., if  $g$  contains no multiple edges or loops.

For a node  $u$  of a rooted tree  $t$ , we write  $c_t(u)$  for the number of children of  $u$  in  $t$ . Given a rooted tree  $t$  with  $v(t) = [n] := \{1, 2, \dots, n\}$ , the *child sequence* of  $t$  is the sequence  $c_t = (c_t(i), 1 \leq i \leq n)$ . Similarly, given a tree-rooted graph  $(g, t, \gamma)$  with vertex set  $v(g) = [n]$ , the *degree sequence* of  $(g, t, \gamma)$  is the sequence  $(d_g(i), 1 \leq i \leq n)$ , where  $d_g(i)$  is the number of endpoints of edges incident to  $i$  in  $g$ ; here loops are counted twice.

For any sequence  $d = (d(1), \dots, d(n))$  of non-negative integers, we define the *degree distribution*  $p_d = (p_d(k), k \geq 1)$  of  $d$  by letting  $p_d(k) = \#\{i \in [n] : d(i) = k\}/n$ .

The following theorem contains our main result, which is an invariance principle for the spanning trees in random tree-rooted graphs with a fixed degree sequence. To state it, two further bits of notation are needed. Given a finite graph  $g = (v, e)$  and a constant  $c > 0$ , we write  $cg$  for the measured metric space  $(v, \text{dist}, \pi)$  whose points are the elements of  $v$ , with  $\text{dist}(x, y) := c \cdot \text{dist}_g(x, y)$ , where  $\text{dist}_g(x, y)$  denotes graph distance in  $g$ , and with  $\pi$  the uniform measure on  $v$ . Also, for a sequence  $p = (p(k), k \geq 1)$  of real numbers, we write  $\mu_1(p) := \sum_{k \geq 1} kp(k)$  and  $\mu_2(p) := \sum_{k \geq 1} k^2 p(k)$ .

**Theorem 1.1.** *For each  $n \geq 1$  let  $d^n = (d^n(i), 1 \leq i \leq n)$  be a degree sequence with  $\min_{1 \leq i \leq n} d^n(i) \geq 1$ , with  $\sum_{i \in [n]} d^n(i) \geq 2n$  and with  $\sum_{i \in [n]} d^n(i)$  even. Let  $p^n$  be the degree distribution of  $d^n$ . Suppose that there exists a probability distribution  $p = (p(k), k \geq 1)$  such that (a)  $p^n \rightarrow p$  pointwise, and (b)  $\mu_2(p^n) \rightarrow \mu_2(p) \in (0, \infty)$ . Then there exists  $\sigma = \sigma(p) \in (0, \infty)$  such that the following holds.*

---

Date: August 27, 2020.

2010 *Mathematics Subject Classification.* Primary: 60C05, 05C80; Secondary: 05C05, 60F05.

For  $n \geq 1$  let  $(G_n, T_n, \Gamma_n)$  be chosen uniformly at random among all simple tree-rooted graphs with vertex set  $[n]$  and degree sequence  $d^n$ . Then

$$\frac{\sigma}{n^{1/2}} T_n \xrightarrow{d} \mathcal{T}$$

as  $n \rightarrow \infty$  with respect to the Gromov-Hausdorff-Prokhorov topology, where  $\mathcal{T}$  is the Brownian continuum random tree.

We refer the reader to [1] for a good discussion of the Gromov-Hausdorff-Prokhorov topology aimed at probabilists. The technical insight underlying the proof of Theorem 1.1 is the fact that Pitman's additive coalescent [18] can be modified to yield a simple construction procedure for random tree-weighted graphs with a given degree sequence. We anticipate that this procedure has further interesting features to be plumbed.

**1.1. Related work.** The enumerative combinatorics of tree-rooted maps was developed in the 1960's and 1970's [17, 20]. The area has seen renewed attention over the last decade or so [3–5]. Random tree-rooted maps can be interpreted as samples from a Fortuin-Kastelyn model at zero temperature, and are an active object of study in the planar probability community (see, e.g., [2, 7, 10–13, 15]).

There has also been some work on the typical number of spanning trees in uniformly random graphs [8, 9, 14] with given degree sequences. (In such models, the underlying graph is sampled uniformly at random from some set of allowed graphs; in our model, it is the tree-weighted graph which is uniformly random, which means the underlying measure on graphs is biased in favour of graphs with a greater number of spanning trees.)

Except in the setting of graphs on surfaces, we have not found any previous work on tree-weighted graphs, random or otherwise.

**1.2. Overview of the proof.** We begin with a small number of facts and definitions that are required for the overview. We say a sequence  $c = (c(i), 1 \leq i \leq n)$  of non-negative integers is a child sequence if it is the child sequence of some tree. Note that  $c$  is a child sequence if and only if  $\sum_{1 \leq i \leq n} c(i) = n - 1$ , in which case

$$\#\{\text{rooted trees } t : c_t = c\} = \binom{n-1}{c(1), \dots, c(n)} = \frac{1}{n} \frac{n!}{\prod_{i=1}^n c(i)!}; \quad (1)$$

see [16], Section 3.3.

Given any sequence  $c = (c(i), 1 \leq i \leq n)$  of non-negative integers, for  $k \geq 1$  we write  $Q_c(k) = \#\{i \in [n] : c(i) = k\}$ . We call  $Q_c = (Q_c(k), k \geq 0)$  the *child statistics vector* of  $c$ . For a tree  $t$  with child sequence  $c_t$ , we will sometimes write  $Q_t = Q_{c_t}$  for succinctness.

Given a graph  $g$ , for an edge  $e \in e(g)$  we write  $m_g(e)$  for the multiplicity of edge  $e$  in  $g$ . Given a degree sequence  $d = (d(1), \dots, d(n))$ , the classical *configuration model* [19, Chapter 7] produces a random graph  $G$  such that for any fixed graph  $g$  with degree sequence  $d$ ,

$$\mathbf{P}\{G = g\} \propto \frac{1}{\prod_{i=1}^n 2^{m_g(ii)} \prod_{e \in e(g)} m_g(e)!}. \quad (2)$$

In Section 2, we define a sampling procedure, inspired by the configuration model and by Pitman's additive coalescent [18], which produces a random tree-weighted graph  $(G, T, \Gamma)$  with the property that for any fixed tree-weighted graph  $(g, t, \gamma)$  with degree sequence  $d$ ,

$$\mathbf{P}\{(G, T, \Gamma) = (g, t, \gamma)\} \propto \frac{2^{\mathbf{1}[\gamma \text{ is a loop}] \cdot m_{g-t}(\gamma)}}{\prod_{i=1}^n 2^{m_{g-t}(ii)} \cdot \prod_{e \in e(g)} m_{g-t}(e)!}, \quad (3)$$

where  $g - t$  is the graph with the same vertex set as  $g$  and with edge multiplicities given by

$$m_{g-t}(e) = \begin{cases} m_g(e) & \text{if } e \notin e(t) \\ m_g(e) - 1 & \text{if } e \in e(t). \end{cases}$$

We call a random tree-weighted graph  $(G, T, \Gamma)$  with distribution given by (3) a *random tree-weighted graph with degree sequence*  $d$ . Note that in this case, conditionally given that  $G$  is simple,  $(G, T, \Gamma)$  is uniformly distributed over simple tree-rooted graphs with degree sequence  $d$ .

The sampling procedure we use has enough exchangeability that, conditional on its child sequence, the resulting spanning tree  $T$  is uniformly distributed; that is, for any fixed child sequence  $c = (c(i), 1 \leq i \leq n)$ , and any tree  $t$  with  $c_t = c$ ,

$$\mathbf{P}\{T = t \mid c_T = c\} = \left( \binom{n-1}{c(1), \dots, c(n)} \right)^{-1}.$$

Now let  $(d^n, n \geq 1)$  be a sequence of degree sequences satisfying the conditions of Theorem 1.1; for each  $n$  let  $(G(d^n), T(d^n), \Gamma(d^n))$  be a random tree-weighted graph with degree sequence  $d^n$ . We prove (see Proposition 3.1) that there is a probability distribution  $q = (q(i), i \geq 0)$  with  $\mu_2(q) < \infty$  such that the child statistics vector  $Q_{T(d^n)}$  satisfies that  $n^{-1}Q_{T(d^n)}(a) \rightarrow q(a)$  in probability for all  $a \geq 0$ , and moreover that  $\mu_2(n^{-1}Q_{T(d^n)}) \rightarrow \mu_2(q)$  in probability. It then follows from a result of Broutin and Marckert [6] that

$$\frac{\sigma}{n^{1/2}}T(d^n) \xrightarrow{d} \mathcal{T} \quad (4)$$

in the Gromov-Hausdorff-Prokhorov sense, where  $\sigma^2 = \mu_2(q) - 1$  and  $\mathcal{T}$  is the Brownian continuum random tree.

This is not quite the convergence claimed in Theorem 1.1, because  $(G(d^n), T(d^n), \Gamma(d^n))$  is a random tree-weighted graph with degree sequence  $d^n$ , whereas Theorem 1.1 concerns random *simple* tree-weighted graphs. To obtain Theorem 1.1 from (4), we show that there is  $\alpha \in (0, 1]$  such that as  $n \rightarrow \infty$ ,

$$\mathbf{P}\{G(d^n) \text{ is simple} \mid T(d^n)\} \rightarrow \alpha \quad (5)$$

in probability. The value of (5), informally, is that it implies that conditioning  $G(d^n)$  to be simple has an asymptotically negligible effect on the law of  $T(d^n)$ . Since the law of  $(G(d^n), T(d^n), \Gamma(d^n))$ , conditional on the simplicity of  $G(d^n)$ , is uniform over simple tree-weighted graphs with degree sequence  $d^n$ , we can then conclude straightforwardly.

We finish the overview with a brief discussion of how we prove (5). Our procedure for constructing random tree-weighted graphs with a given degree sequence first constructs the tree  $T(d^n)$ , then randomly pairs the remaining half-edges as in the standard configuration model. Viewing  $T(d^n)$  as fixed, this leads us to the following more general question. Let  $G = (V, E)$  be a random graph with a given degree sequence generated according to the configuration model, and let  $T = (V, E')$  be a fixed, simple graph with the same vertex set. What is the probability that the union of  $G$  and  $T$  forms a simple graph (i.e. that  $G$  is a simple graph and that  $E$  and  $E'$  are disjoint)? We provide a partial answer to this question by proving a fairly general Poisson approximation theorem for the number of loops and multiple edges in the superposition of a fixed graph and a random graph drawn from the configuration model (see Theorem 4.1). In order to apply Theorem 4.1, we need that the joint degree statistics in  $G(d^n)$  and in  $T(d^n)$  are sufficiently well-behaved; proving this is the task of Section 3. We then state and prove Theorem 4.1 in Section 4, and finally put all the pieces together to prove Theorem 1.1 in Section 5.

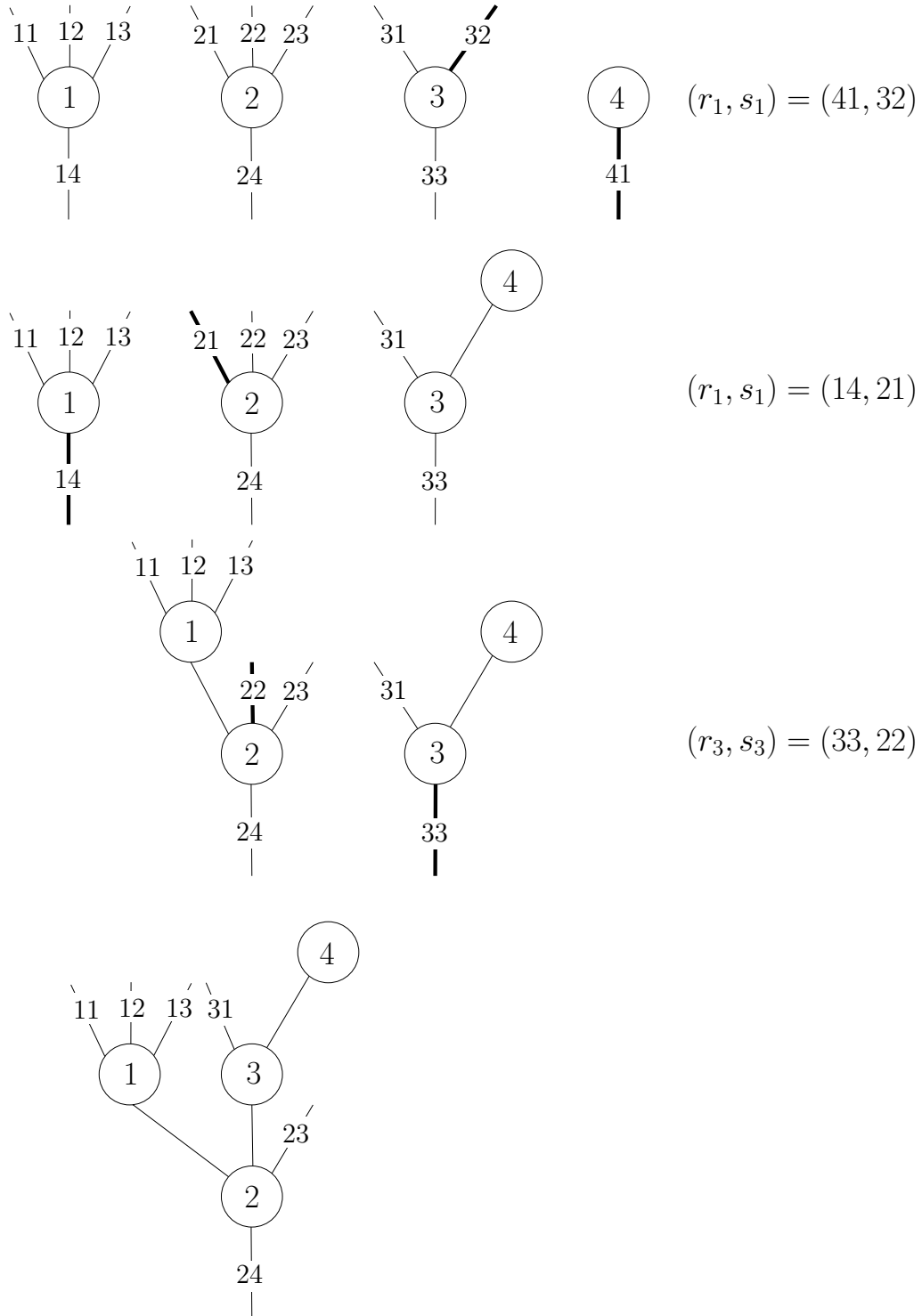


FIGURE 1. An example of an execution path of Pitman's additive coalescent. The forests  $F_1, F_2, F_3$  and  $F_4$  are displayed in successive rows.

## 2. PITMAN'S ADDITIVE COALESCENT WITH A FIXED DEGREE SEQUENCE

**2.1. The sampling process.** Let  $\mathbf{d} = (d(1), \dots, d(n))$  be a degree sequence, which is to say that  $d(1), \dots, d(n)$  are non-negative integers. To be well-defined, the next process requires that  $\sum_{1 \leq i \leq n} d(i) \geq 2n - 1$  and that  $d(i) \geq 1$  for all  $1 \leq i \leq n$ .

**Pitman's additive coalescent.** The process has  $n - 1$  steps, and at the start of step  $k$  consists of a rooted forest  $F_k(d) = \{T_1^k(d), \dots, T_{n+1-k}^k(d)\}$  with  $n + 1 - k$  trees. At the start of step 1, these trees are isolated vertices with labels  $1, \dots, n$ . Vertex  $i$  has  $d(i)$  half-edges  $(i1, i2, \dots, id(i))$  attached to it, and  $id(i)$  is distinguished as the *root half-edge*.

*Step  $k$ :*

Choose a uniformly random pair  $(r_k, s_k)$ , where  $r_k$  is a root half-edge which is not paired in  $F_k(d)$  and  $s_k$  is a non-root half-edge which is not paired in  $F_k(d)$  and additionally belongs to a different tree of  $F_k(d)$  from  $r_k$ .

Pair the half-edges  $r_k$  and  $s_k$  to create an edge  $e_k$  connecting their endpoints; this merges two trees of  $F_k(d)$ . The root of the new tree is the same as the root of the tree of  $F_k(d)$  containing  $s_k$ . In the new tree, the vertex incident to  $r_k$  is the child of the vertex incident to  $s_k$ .

Define  $F_{k+1}(d)$  to be the forest consisting of the new tree thus created, together with the remaining  $n - k - 1$  unaltered trees of  $F_k(d)$ .

An example is shown in Figure 1. Write  $T(d) = T_1^n(d)$  for the single tree in the random forest  $F_n(d)$ . Attached to the tree  $T(d)$  there is a single pendant (unpaired) root half-edge which is incident to the root of  $T(d)$ , and if  $\sum_{i=1}^n d(i) > 2n - 1$  then there are also other pendant half-edges. By ignoring pendant half-edges, we may view  $T(d)$  as a random rooted tree with vertex set  $[n]$ .

It will be useful to additionally define two edge labellings of  $T(d)$ , denoted  $K$  and  $H$ . We define  $K(e)$  to be the step at which edge  $e$  was added; so  $K(e_k) := k$ . We define  $H(e)$  to be the non-root half-edge used in creating  $e$ ; so  $H(e_k) = s_k$ . Note that  $K$  is a bijection between  $e(T(d))$  and  $[n - 1]$ . Also, if  $i \in [n]$  has  $c_{T(d)}(i) = c$ , then  $H$  assigns  $c$  distinct half-edges from the set  $\{i1, \dots, id(i)\}$  to the edges between  $i$  and its children in  $T(d)$ .

We use the phrase “execution path” to mean a sequence of pairs  $(r_1, s_1), \dots, (r_{n-1}, s_{n-1})$  which may conceivably appear as the ordered sequence of pairs of half-edges added during the course of Pitman's coalescent.

The next proposition fully describes the joint distribution of  $T(d)$ ,  $K$ , and  $H$ . In its proof, and in what follows, for a rooted tree  $t$  and a node  $u \in v(t) \setminus \{r(t)\}$  we write  $\text{par}(u)$  for the parent of  $u$  in  $t$ . Also, we use the falling factorial notation  $(k)_\ell := k(k-1)\dots(k-\ell+1) = k!/(k-\ell)!$ .

**Proposition 2.1.** *Let  $d = (d(1), \dots, d(n))$  be a degree sequence with  $d(i) \geq 1$  for all  $i \in [n]$  and with  $\sum_{i=1}^n d(i) \geq 2(n-1)$ , and write  $m = \frac{1}{2} \sum_{i=1}^n d(i)$ . Then the following properties all hold.*

- (1) *For any fixed rooted tree  $t$  with vertex set  $[n]$ ,*

$$\mathbf{P}\{T(d) = t\} = \frac{1}{(2m-n)_{n-1}} \prod_{i=1}^n (d(i)-1)_{c_t(i)}.$$

- (2) *Fix any set  $\mathcal{H} \subset \bigcup_{i=1}^n \{i1, \dots, i(d_i-1)\}$  with  $|\mathcal{H}| = n-1$ . Conditionally given that  $\{s_1, \dots, s_{n-1}\} = \mathcal{H}$ , the triple  $(T(d), K, H)$  is uniformly distributed over the  $((n-1)!)^2$  triples which are consistent with the event  $\{s_1, \dots, s_{n-1}\} = \mathcal{H}$ .*
- (3) *The sequence  $(s_1, \dots, s_{n-1})$  of non-root half-edges, added by Pitman's coalescent, is uniformly distributed over the set of sequences of  $(n-1)$  distinct elements of  $\bigcup_{1 \leq i \leq n} \{i1, \dots, i(d(i)-1)\}$ . Consequently,  $\{s_1, \dots, s_{n-1}\}$  is a uniformly random size- $(n-1)$  subset of  $\bigcup_{1 \leq i \leq n} \{i1, \dots, i(d(i)-1)\}$ .*
- (4) *Finally, conditionally given that  $T(d) = t$  and given the set  $\{s_1, \dots, s_{n-1}\}$  of non-root half-edges added by Pitman's coalescent, the ordering  $(e_1, \dots, e_{n-1})$  of  $e(t)$  is uniformly distributed over the  $(n-1)!$  possible orderings of  $e(t)$ .*

*Proof.* At step  $i$  of the process, there are  $n + 1 - i$  components and  $m - n + 1 - i$  unpaired non-root half-edges. We may specify the pair  $(r_i, s_i)$  by first revealing the non-root half-edge  $s_i$ , then revealing  $r_i$ . Whatever the choice of  $s_i$ , there are  $n - i$  possibilities for  $r_i$ , so the number of distinct choices for the pair  $(r_i, s_i)$  is  $(2m - n + 1 - i)(n - i)$ . Thus, the total number of possible execution paths for the process is

$$\prod_{i=1}^{n-1} (2m - n + 1 - i)(n - i) = (n - 1)!(2m - n)_{n-1}. \quad (6)$$

The execution path followed by the process is uniquely determined by the tree  $T(d)$  and the functions  $K : e(T(d)) \rightarrow [n - 1]$  and  $H : e(T(d)) \rightarrow \bigcup_{i=1}^n \{i1, \dots, i(d_i - 1)\}$ . To see this, fix any  $k \in [n - 1]$ . Then the edge  $e_k$  created at step  $k$  of Pitman's coalescent may be recovered as  $e_k = K^{-1}(k)$ ; and, if  $e_k = uv$  with  $v = \text{par}(u)$  then the half-edges paired to create  $e_k$  are the root half-edge  $vd(v)$  incident to  $v$  and the half-edge  $H^{-1}(e_k)$ .

Now, fix any tree  $t$  with degree sequence  $d$ , any bijection  $k : e(t) \rightarrow [n - 1]$ , and any function  $h : e(t) \rightarrow \mathbb{N}$  which, for all  $i \in [n]$ , assigns  $c_t(i)$  distinct values from the set  $\{1, \dots, (d(i) - 1)\}$  to the edges between  $i$  and its children in  $t$ . Together with (6), the observation of the preceding paragraph implies that

$$\mathbf{P} \{T(d) = t, K = k, H = h\} = \frac{1}{(n - 1)!(2m - n)_{n-1}}.$$

Having fixed the tree  $t$ , the number of possible values for  $K$  is  $(n - 1)!$  and the number of possible values for  $H$  is  $\prod_{i \in [n]} (d(i) - 1)_{c_t(i)}$ . It follows that

$$\mathbf{P} \{T(d) = t\} = \frac{(n - 1)! \cdot \prod_{i \in [n]} (d(i) - 1)_{c_t(i)}}{(n - 1)!(2m - n)_{n-1}} = \frac{\prod_{i \in [n]} (d(i) - 1)_{c_t(i)}}{(2m - n)_{n-1}},$$

which proves the first claim of the proposition.

Next, fix  $\mathcal{H}$  as in the second assertion of the proposition, and any ordering of  $\mathcal{H}$  as  $(h_1, \dots, h_{n-1})$ . Then the number of execution paths which yield that  $s_k = h_k$  for  $k \in [n - 1]$  is precisely  $(n - 1)!$ . To see this, note that if  $s_j = h_j$  for  $1 \leq j \leq k$  then, whatever the choices of the root half-edges  $(r_j, 1 \leq j \leq k)$ , the forest  $F_k^n$  has  $n + 1 - k$  component trees so there are  $n - k$  unpaired root half-edges in components different from that of  $s_k$ ; any such root half-edge may be chosen as  $r_k$ . Since there are also  $(n - 1)!$  possible orderings of  $\mathcal{H}$ , the second assertion of the proposition follows.

To prove the third statement, fix a set  $\mathcal{H}$  and an ordering  $(h_1, \dots, h_{n-1})$  of its elements, as in the previous paragraph. For each  $1 \leq k < n - 1$ , given that  $s_j = h_j$  for  $1 \leq k < i$ , whatever the choices of  $(r_j, 1 \leq j < k)$  may be, there are  $n - k$  ways to choose  $r_k$  in a distinct tree from  $h_k$ . It follows that there are  $\prod_{k=1}^{n-2} (n - k) = (n - 1)!$  execution paths with the property that  $s_k = h_k$  for each  $1 \leq k \leq n - 1$ . Since this number does not depend on  $\mathcal{H}$ , it follows that each size- $(n - 1)$  subset of  $\bigcup_{1 \leq i \leq n} \{i1, \dots, i(d(i) - 1)\}$  is equally likely.

Finally, fix both the tree  $t$  and an unordered set  $\mathcal{H}$  of non-root half-edges with  $|\mathcal{H} \cap \{i1, \dots, i(d(i) - 1)\}| = c_t(i)$  for all  $i \in [n]$ . We consider the number of execution paths which yield  $T(d) = t$  and  $\{s_1, \dots, s_n\} = \mathcal{H}$ . The number of choices of an ordering function  $k : e(t) \rightarrow [n - 1]$  consistent with these constraints is still  $(n - 1)!$ . Moreover, whatever the choice of  $k$ , under the further constraint  $K = k$ , the number of possibilities for  $H$  is  $\prod_{i \in [n]} c_t(i)!$ . To see this, note that for each  $i \in [n]$ , the constraints precisely imply that  $\mathcal{H} \cap \{i1, \dots, i(d(i) - 1)\} = \{s_1, \dots, s_{n-1}\} \cap \{i1, \dots, i(d(i) - 1)\}$ , and  $H$  is fixed once we additionally specify which of these  $c_t(i)$  half-edges is matched to which child of  $i$ , for each  $i \in [n]$ . It follows that the number of execution paths which yield that  $T(d) = t$ , that  $K = k$  and that  $\{s_1, \dots, s_n\} = \mathcal{H}$  is

$$\prod_{i=1}^n c_t(i)!.$$

As this quantity doesn't depend on the choice of the ordering function  $k$ , the final assertion of the proposition follows.  $\square$

We state a corollary of the above proposition, for later use.

**Corollary 2.2.** *The tree  $T(d)$  is a uniformly random rooted tree with child sequence  $c_T$ .*

The corollary follows since the formula for  $\mathbf{P}\{T(d) = t\}$  from Proposition 2.1 only depends on  $t$  through  $c_t$ .

We now assume that  $\sum_{i=1}^n d(i) \geq 2n$  and that  $\sum_{i=1}^n d(i)$  is even, and define a random tree-rooted graph  $(G, T, \Gamma) = (G(d), T(d), \Gamma(d))$  as follows: First, let  $T = T(d)$  be the random tree built by Pitman's coalescent, and let  $\Gamma^+$  be its root half-edge. We refer to  $T$  as the *spanning tree-elect* of a to-be-constructed tree-rooted graph. Next, choose a uniformly random matching of the  $2m - 2(n - 1)$  pendant half-edges attached to  $T$ , and pair the half-edges according to this matching to create  $G = G(d)$ . Then let  $\Gamma$  be the edge containing  $\Gamma^+$ , oriented so that  $\Gamma^+$  is at the head. We call  $(G(d), T(d), \Gamma(d))$ , or any other graph with the same distribution, a *random tree-rooted graph with degree sequence  $d$* . The tree  $T$  has now taken office.

The next proposition describes the distribution of  $(G(d), T(d), \Gamma(d))$ . For a tree-rooted graph  $(g, t, \gamma)$ ,

**Proposition 2.3.** *Let  $d = (d(1), \dots, d(n))$  be a degree sequence with  $d(i) \geq 1$  for all  $i \in [n]$ , and write  $m = \frac{1}{2} \sum_{i=1}^n d(i)$ . Fix a tree-rooted graph  $(g, t, \gamma)$  where  $g$  is a graph with degree sequence  $d$ . Then*

$$\mathbf{P}\{(G(d), T(d), \Gamma(d)) = (g, t, \gamma)\} \propto \frac{2^{\mathbf{1}_{[\gamma \text{ is a loop}]}} \cdot m_{g-t}(\gamma)}{\prod_{i=1}^n 2^{m_{g-t}(ii)} \cdot \prod_{e \in e(g)} m_{g-t}(e)!}.$$

*Proof.* Proposition 2.1 gives us a formula for  $\mathbf{P}\{T(d) = t\}$ . We next focus on computing

$$\mathbf{P}\{G(d) = g \mid T(d) = t\}.$$

Write  $r$  for the root of  $t$ , and  $\gamma = qr$  for the oriented root edge of  $g$ . Given that  $T(d) = t$ , each  $i \in [n]$  with  $i \neq r(t)$  has  $d'(i) := d(i) - c_t(i) - 1$  pendant half-edges attached to it, and  $r$  has  $d'(r) := d(r) - c_t(r)$  half-edges attached to it. Conditionally given that  $T(d) = t$ , the graph  $G(d) - T(d)$  is distributed as  $CM(d')$ , a random graph with degree sequence  $d' = (d'(1), \dots, d'(n))$  sampled according to the configuration model (so with distribution as in (2)). Writing  $m' := m - (n - 1) = \frac{1}{2} \sum_{i=1}^n d'(i)$  and  $g' = (v(g), e(g) \setminus e(t))$ , it follows that

$$\begin{aligned} \mathbf{P}\{G(d) = g \mid T(d) = t\} &= \mathbf{P}\{CM(d') = g'\} \\ &= \frac{2^{m'}(m')!}{(2m')!} \frac{\prod_{i=1}^n d'(i)!}{\prod_{i=1}^n 2^{m_{g'}(ii)} \cdot \prod_{e \in e(g')} m_{g'}(e)!} \\ &= \frac{2^{m'}(m')!}{(2m')!} \frac{\prod_{i=1}^n d'(i)!}{\prod_{i=1}^n 2^{m_{g-t}(ii)} \cdot \prod_{e \in e(g)} m_{g-t}(e)!}. \end{aligned}$$

In the last equality, we use that  $m_{g'}(ii) = m_{g-t}(ii)$  since  $t$  is a tree so contains no loops, and that  $m_{g'}(e) = m_{g-t}(e)$  by definition when  $e \in e(g')$ .

Given that  $T(d) = t$  and that  $G(d) = g$ , in order to have  $(G(d), T(d), \Gamma(d)) = (g, t, \gamma)$  it is necessary and sufficient that  $\Gamma(d) = \gamma$ . This occurs precisely if  $\gamma^+$ , the half-edge of  $\gamma$  incident to  $r$ , was matched with some half-edge incident to  $q$ . Since the matching of half-edges in  $G(d) - T(d)$  is chosen uniformly at random, by symmetry the conditional probability that this occurred is  $m_{g-t}(\gamma)/d'(r)$  if  $\gamma$  is not a loop, and is  $2m_{g-t}(\gamma)/d'(r)$  if  $\gamma$  is a loop. We may unify these two formulas by writing

$$\mathbf{P}\{\Gamma(d) = \gamma \mid T(d) = t, G(d) = g\} = \frac{2^{\mathbf{1}_{[\gamma \text{ is a loop}]}} m_{g-t}(\gamma)}{d'(r)}.$$

Combined with the formula for  $\mathbf{P}\{T(d) = t\}$  from Proposition 2.1, this gives

$$\begin{aligned} & \mathbf{P}\{(G(d), T(d), \Gamma(d)) = (g, t, \gamma)\} \\ &= \frac{1}{(2m-n)_{n-1}} \prod_{i=1}^n (d(i)-1)_{c_t(i)} \\ & \quad \cdot \frac{2^{m'}(m')!}{(2m')!} \frac{\prod_{i=1}^n d'(i)!}{\prod_{i=1}^n 2^{m_{g-t}(ii)} \cdot \prod_{e \in e(g)} m_{g-t}(e)!} \\ & \quad \cdot \frac{2^{\mathbf{1}_{[\gamma \text{ is a loop}]}} m_{g-t}(\gamma)}{d'(r)} \\ &= \prod_{i=1}^n (d(i)-1)! \cdot \frac{2^{m-(n-1)}(m-(n-1))!}{(2m-n)!} \cdot \frac{2^{\mathbf{1}_{[\gamma \text{ is a loop}]}} m_{g-t}(\gamma)}{\prod_{i=1}^n 2^{m_{g-t}(ii)} \cdot \prod_{e \in e(g)} m_{g-t}(e)!}. \end{aligned}$$

In the second equality we have used that  $(2m-n)_{n-1}(2m')! = (2m-n)!$ , that  $(d(i)-1)_{c_t(i)}d'(i)! = (d(i)-1)!$  for  $i \neq r$ , and that  $(d(r)-1)_{c_t(r)}d'(r)! = d'(r)(d(r)-1)!$ . The first two terms on the final line do not depend on the triple  $(g, t, \gamma)$ , so the result follows.  $\square$

### 3. CONCENTRATION OF DEGREES

Throughout this section, let  $(d^n, n \geq 1)$  be a sequence of degree sequences satisfying the conditions of Theorem 1.1, and also let  $p^n$  and  $p$  be as in Theorem 1.1. Next, for  $n \geq 1$  let  $T(d^n)$  be the tree built by Pitman's additive coalescent applied to the degree sequence  $d^n = (d^n(i), 1 \leq i \leq n)$ . Let  $c^n = (c^n(i), 1 \leq i \leq n)$  be the child sequence of  $T(d^n)$ , and recall that  $Q_{c^n} = (Q_{c^n}(a), a \geq 0)$  is the child statistics vector of  $c^n$ . Also, for  $0 \leq a < b$ , let  $P_{b,a}^n = \#\{1 \leq i \leq n : d^n(i) = b, c^n(i) = a\}$ . Finally, let  $\rho := 1/(\mu_1(p) - 1)$ .

**Proposition 3.1.** *For  $a \geq 0$  let*

$$q(a) := \sum_{b=a+1}^{\infty} p(b) \cdot \mathbf{P}\{\text{Bin}(b-1, \rho) = a\}.$$

*Then  $\mu_2(q) < \infty$  and  $\mu_2(n^{-1}Q_{c^n}) \rightarrow \mu_2(q)$  in probability as  $n \rightarrow \infty$ . Moreover, for all  $0 \leq a < b$ ,  $n^{-1}P_{b,a}^n \xrightarrow{\text{prob}} p(b) \cdot \mathbf{P}\{\text{Bin}(b-1, \rho) = a\}$ , and  $n^{-1}Q_{c^n}(a) \xrightarrow{\text{prob}} q(a)$ , in both cases as  $n \rightarrow \infty$ .*

Let  $(G(d^n), T(d^n), \Gamma(d^n))$  be a random tree-weighted graph with degree sequence  $d^n$ . Using Proposition 3.1, together with existing results from the literature, it is fairly straightforward to establish that  $(\sigma n^{-1/2})T(d^n) \xrightarrow{d} \mathcal{T}$ , with  $\sigma = \mu_2(q) - 1$ , where  $\mathcal{T}$  is the Brownian continuum random tree. However, in order to show that such convergence holds for the corresponding random *simple* tree-weighted graphs, we additionally need the next proposition, which establishes that the number of pairs of tree-adjacent vertices in  $T(d^n)$  with given fixed degrees is well-concentrated around its expected values. This will be used in order to show that  $G(d^n)$  is simple given  $T(d^n)$  asymptotically behaves like a constant.

Write  $G_-(d^n) = G(d^n) - T(d^n)$  and let  $d_-^n = (d_-^n(i), 1 \leq i \leq n)$  be the degree sequence of  $G_-(d^n)$ . For integers  $k, \ell \geq 0$ , let

$$\alpha(k, \ell) = \sum_{a_1, a_2 \geq 0} a_2 p(\ell + a_2 + 1) \mathbf{P}\{\text{Bin}(\ell + a_2, \rho) = a_2\} \cdot p(k + a_1 + 1) \mathbf{P}\{\text{Bin}(k + a_1, \rho) = a_1\}. \quad (7)$$

**Proposition 3.2.** *For integers  $k, \ell \geq 0$  let*

$$A^n(k, \ell) = |\{uv \in e(T(d^n)) : d_-^n(u) = k, d_-^n(v) = \ell\}|.$$



Then for all  $k, \ell \geq 0$ ,

$$\frac{1}{n} A^n(k, \ell) \xrightarrow{\text{prob}} \alpha(k, \ell)$$

as  $n \rightarrow \infty$ , and also

$$\frac{1}{n} \sum_{k, \ell \geq 0} k \ell A^n(k, \ell) \xrightarrow{\text{prob}} \sum_{k, \ell \geq 0} k \ell \alpha(k, \ell).$$

The proofs of Propositions 3.1 and 3.2 appear in Appendix A.

To conclude the section, we observe that  $\alpha(k, \ell)$  defines a probability distribution on pairs of non-negative integers. Indeed,

$$\begin{aligned} \sum_{k \geq 0} \sum_{a_1 \geq 0} p(k + a_1 + 1) \mathbf{P} \{ \text{Bin}(k + a_1, \rho) = a \} &= \sum_{m \geq 0} \sum_{k+a=m} p(m + 1) \mathbf{P} \{ \text{Bin}(m, \rho) = a \} \\ &= \sum_{m \geq 0} p(m + 1) = 1 - p(0) = 1, \end{aligned}$$

and

$$\begin{aligned} &\sum_{\ell \geq 0} \sum_{a_2 \geq 0} a_2 p(\ell + a_2 + 1) \mathbf{P} \{ \text{Bin}(\ell + a_2, \rho) = a_2 \} \\ &= \sum_{m \geq 0} \sum_{\ell+a=m} a p(m + 1) \mathbf{P} \{ \text{Bin}(m, \rho) = a \} \\ &= \sum_{m \geq 0} p(m + 1) \cdot m \rho = (\mu_1(p) - (1 - p(0))) \rho = (\mu_1(p) - 1) \rho = 1, \end{aligned}$$

so by factorizing  $\sum_{k, \ell \geq 0} \alpha(k, \ell)$  we obtain

$$\sum_{k, \ell \geq 0} \alpha(k, \ell) = \left( \sum_{m \geq 0} p(m + 1) \right) \cdot \left( \sum_{m \geq 0} \sum_{m \geq 0} p(m + 1) \cdot m \rho \right) = 1;$$

the fact that  $\sum_{k, \ell \geq 0} \alpha(k, \ell) = 1$  will be used in the proof of Proposition 3.2. A similar computation shows that

$$\sum_{k, \ell \geq 0} k \ell \cdot \alpha(k, \ell) \leq \left( \sum_{m \geq 0} m p(m + 1) \right) \cdot \left( \sum_{m \geq 0} m^2 p(m + 1) \rho \right) = \mu_2(p) - 2\mu_1(p) + 1 < \infty, \quad (8)$$

a fact we will use in bounding the probability of simplicity of  $G(d^n)$ .

#### 4. POISSON APPROXIMATION FOR GRAPH SUPERPOSITIONS.

In this section we state a Poisson approximation theorem for the number of loops and multiple edges in the superposition of a fixed simple graph and a random graph with a fixed degree sequence; this in particular allows us to control the probability that such a superposition yields a simple graph.

Let  $H$  be a simple graph with vertex set  $v(H) = [n]$ . Fix a degree sequence  $d = (d(1), \dots, d(n))$  whose sum of degrees is even, and let  $G$  be a random graph with degree sequence  $d$  sampled according to the configuration model. For vertices  $u, v \in [n]$  and  $i \in [d(u)], j \in [d(v)]$ , let  $\mathbf{1}_{[ui, vj]}$  be the indicator of the event that half-edge  $ui$  is matched

with half-edge  $vj$  in  $G$ . Now write

$$\begin{aligned}\mathcal{L} &= \mathcal{L}(G) = \{(ui, uj) : u \in [n], i, j \in [d(u)], i < j\} \\ \mathcal{M} &= \mathcal{M}(G, H) = \{((ui_1, vj_1), (ui_2, vj_2)) : u, v \in [n], uv \notin e(H), \\ &\quad i_1, i_2 \in [d(u)], j_1, j_2 \in [d(v)], u < v, i_1 < i_2, j_1 \neq j_2\}, \text{ and} \\ \mathcal{N} &= \mathcal{N}(G, H) = \{(ui, vj) : uv \in e(H), i \in [d(u)], j \in [d(v)]\},\end{aligned}$$

and let

$$\begin{aligned}L &= L(G) = \sum_{(ui,uj) \in \mathcal{L}} \mathbf{1}_{[ui,uj]}, \\ M &= M(G, H) = \sum_{((ui_1,vj_1),(ui_2,vj_2)) \in \mathcal{M}} \mathbf{1}_{[(ui_1,vj_1)]} \mathbf{1}_{[(ui_2,vj_2)]}, \text{ and} \\ N &= N(G, H) = \sum_{(ui,vj) \in \mathcal{C}} \mathbf{1}_{[ui,vj]}.\end{aligned}$$

Note that the graph with edge set  $e(G) \cup e(H)$  is simple precisely if  $L + M + N = 0$ .

**Theorem 4.1.** *Fix a sequence of graphs  $(h_n, n \geq 1)$  with  $v(h_n) = [n]$  and  $\max_{v \in [n]} \{\deg_{h_n}(v)\} = o(n)$  for all  $n \geq 1$ . For each  $n \geq 1$  let  $d^n = (d^n(v), 1 \leq v \leq n)$  be a degree sequence and let  $p^n$  be the degree distribution of  $d^n$ . Suppose that there exists a probability distribution  $p = (p(k), k \geq 0)$  with  $\mu_2(p) \in [0, \infty)$  and  $p(0) < 1$  such that the following holds.*

*First,  $p^n \rightarrow p$  pointwise and  $\mu_2(p^n) \rightarrow \mu_2(p)$ . Second, there are non-negative numbers  $(\alpha(a, b), a, b \geq 0)$  such that for any  $a, b \geq 0$*

$$\alpha^n(a, b) := \frac{1}{n} |\{uv \in e(h_n) : d^n(u) = a, d^n(v) = b\}| \rightarrow \alpha(a, b),$$

and

$$\sum_{k, \ell \geq 0} kl \alpha^n(k, \ell) \rightarrow \sum_{k, \ell \geq 0} kl \alpha(k, \ell) < \infty \quad (9)$$

*For  $n \geq 1$  let  $G_n$  be distributed according to the configuration model on graphs with vertex set  $[n]$  and degree sequence  $d^n$ . Then with  $L_n = L(G_n)$ ,  $M_n = M(G_n, h_n)$  and  $N_n = N(G_n, h_n)$ , we have*

$$\|\text{Dist}(L_n, M_n, N_n) - \text{Poi}(\nu/2) \otimes \text{Poi}(\nu^2/4) \otimes \text{Poi}(\eta)\|_{\text{TV}} \rightarrow 0$$

*as  $n \rightarrow \infty$ , where  $\nu = \mu_2(p)/\mu_1(p) - 1$  and  $\eta = \frac{1}{\mu_1(p)} \sum_{i, j \geq 1} ij \alpha(i, j)$ .*

In the statement of Theorem 4.1 we have introduced the notation  $\deg_{h_n}(v)$  for the degree of vertex  $v$  in  $h_n$ , and the notation  $\|\mu - \nu\|_{\text{TV}}$  for the total variation distance between probability measures. The proof of Theorem 4.1 appears in Appendix B. This theorem has the following consequence for random tree-weighted graphs, which we will use in the next section.

**Corollary 4.2.** *Let  $(d^n, n \geq 1)$  and  $(p^n, n \geq 1)$  be as in Theorem 1.1, and for  $n \geq 1$  let  $(G(d^n), T(d^n), \Gamma(d^n))$  be a random tree-weighted graph with degree sequence  $d$ . Then*

$$\mathbf{P}\{G(d^n) \text{ simple} \mid T(d^n)\} \xrightarrow{\text{prob}} \exp(-\nu/2 - \nu^2/4 - \eta),$$

as  $n \rightarrow \infty$ .

*Proof.* First, the fact that  $\sum_{k, \ell \geq 0} kl \cdot \alpha(k, \ell) < \infty$  appears in (8), above, from which it is immediate that  $\eta < \infty$ . Next, let  $G_-(d^n)$  be the subgraph of  $G(d^n)$  with edge set  $e(G(d^n)) \setminus e(T(d^n))$ , and let  $d_-^n$  be the degree sequence of  $G_-(d^n)$ , as defined in Section 3 previous to Proposition 3.2. Finally, write  $L_n = L(G_-(d^n))$ ,  $M_n = M(G_-(d^n), T(d^n))$ , and  $N_n = N(G_-(d^n), T(d^n))$ .

Conditionally given  $T(d^n)$ , the graph  $G_-(d^n)$  is a random graph with degree sequence  $d_-^n$ . By Proposition 3.2 we know that  $\alpha^n(k, l) \xrightarrow{\text{prob}} \alpha(k, l)$  for all  $k, l \geq 0$ , and that

$$\sum_{k, l \geq 0} kl \alpha^n(k, l) \xrightarrow{\text{prob}} \sum_{k, l \geq 0} kl \alpha(k, l).$$

By Theorem 4.1 applied with  $h_n = T(d^n)$  and  $G_n = G_-(d^n) = G(d^n) - T(d^n)$ , it now follows that conditionally given  $T(d^n)$ ,

$$\|\text{Dist}(L_n, M_n, N_n) - \text{Poi}(\nu/2) \otimes \text{Poi}(\nu^2/4) \otimes \text{Poi}(\eta)\|_{\text{TV}} \xrightarrow{\text{prob}} 0$$

as  $n \rightarrow \infty$ . If  $(L, M, N)$  is  $\text{Poi}(\nu/2) \otimes \text{Poi}(\nu^2/4) \otimes \text{Poi}(\eta)$ -distributed, then we have  $\mathbf{P}\{L = M = N = 0\} = \exp(-\nu/2 - \nu^2/4 - \eta)$ ; since  $G(d^n)$  is simple if and only if  $L_n = M_n = N_n = 0$ , it follows that

$$\begin{aligned} \mathbf{P}\{G(d^n) \text{ simple} \mid T(d^n)\} &= \mathbf{P}\{L_n = M_n = N_n = 0 \mid T(d^n)\} \\ &\xrightarrow{\text{prob}} \mathbf{P}\{L = M = N = 0\} = \exp(-\nu/2 - \nu^2/4 - \eta), \end{aligned}$$

as required.  $\square$

## 5. PROOF OF THEOREM 1.1

Let  $(d^n, n \geq 1)$  be a sequence of degree sequences satisfying the conditions of Theorem 1.1. For  $n \geq 1$  let  $T(d^n)$  be the tree built by Pitman's additive coalescent applied to degree sequence  $d^n$ , and let  $c^n$  be the child sequence of  $T(d^n)$ . By Proposition 2.1 (1), conditionally given  $c^n$ , the tree  $T(d^n)$  is uniformly distributed over the set of trees with child sequence  $c^n$ .

By Proposition 3.1, the child statistics vectors  $(Q_{c^n}, n \geq 1)$  satisfy that, as  $n \rightarrow \infty$ , for all  $a \geq 0$ ,

$$n^{-1} Q_{c^n}(a) \xrightarrow{\text{prob}} q(a),$$

and moreover that  $\mu_2(n^{-1} Q_{c^n}(a)) \rightarrow \mu_2(q)$ . Here  $q = (q(a), a \geq 0)$  is as in Proposition 3.1, and in particular satisfies  $\mu_2(q) < \infty$ . Writing  $\sigma = \mu_2(q) - 1$ , it then follows by Theorem 1 of [6] that

$$\bar{T}(d^n) := \frac{\sigma}{n^{1/2}} T(d^n) \xrightarrow{d} \mathcal{T},$$

in the Gromov-Hausdorff-Prokhorov sense.<sup>1</sup>

We aim to prove the same statement with  $\bar{T}(d^n)$  replaced by  $\bar{T}_n := (\sigma/n^{1/2})T_n$ , where  $(G_n, T_n, \Gamma_n)$  is a uniformly random simple tree-rooted graph with degree sequence  $d^n$ . To accomplish this, we use that the law of  $(G_n, T_n, \Gamma_n)$  is precisely the conditional law of  $(G(d^n), T(d^n), \Gamma(d^n))$  given that  $G(d^n)$  is a simple graph.

Writing  $\mathbb{K}$  for Gromov-Hausdorff-Prokhorov space as in [1], for any bounded continuous function  $f : \mathbb{K} \rightarrow \mathbb{R}$  we have

$$\begin{aligned} \mathbf{E}(f(\bar{T}(d^n)) \cdot \mathbf{1}_{[G(d^n) \text{ simple}]}) &= \mathbf{E}\left(\mathbf{E}(f(\bar{T}(d^n)) \cdot \mathbf{1}_{[G(d^n) \text{ simple}]} \mid T(d^n))\right) \\ &= \mathbf{E}(f(\bar{T}(d^n)) \cdot \mathbf{P}\{G(d^n) \text{ simple} \mid T(d^n)\}). \end{aligned}$$

<sup>1</sup>Theorem 1 of [6] is stated for plane trees with a fixed degree sequence, rather than labeled trees with a fixed degree sequence. However, as noted by Broutin and Marckert [6, page 295], a straightforward combinatorial argument shows that the same result holds for labeled trees. Also, as stated, the theorem only yields convergence in the Gromov-Hausdorff sense; but the proof proceeds by establishing convergence distributional of coding functions. As explained in [1, Section 3], such proofs immediately yield the stronger Gromov-Hausdorff-Prokhorov convergence.

Since  $\mathbf{E}f(\bar{T}(d^n)) \rightarrow \mathbf{E}f(\mathcal{T})$ , and  $\mathbf{P}\{G(d^n) \text{ simple} \mid T(d^n)\} \xrightarrow{\text{prob}} \exp(-\nu/2 - \nu^2/4 - \eta)$  by Corollary 4.2, it follows that

$$\mathbf{E}(f(\bar{T}(d^n)) \cdot \mathbf{1}_{[G(d^n) \text{ simple}]}) \rightarrow \exp(-\nu/2 - \nu^2/4 - \eta)\mathbf{E}(f(\mathcal{T})).$$

Furthermore,

$$\mathbf{P}\{G(d^n) \text{ simple}\} = \mathbf{E}(\mathbf{P}\{G(d^n) \text{ simple} \mid T(d^n)\}) \rightarrow \exp(-\nu/2 - \nu^2/4 - \eta),$$

and therefore

$$\mathbf{E}(f(\bar{T}(d^n)) \mid G(d^n) \text{ simple}) = \frac{\mathbf{E}(f(\bar{T}(d^n)) \cdot \mathbf{1}_{[G(d^n) \text{ simple}]})}{\mathbf{P}\{G(d^n) \text{ simple}\}} \rightarrow \mathbf{E}(f(\mathcal{T})).$$

Since

$$\mathbf{E}(f(\bar{T}_n)) = \mathbf{E}(f(\bar{T}(d^n)) \mid G(d^n) \text{ simple}),$$

the fact that  $\bar{T}_n \xrightarrow{d} \mathcal{T}$  now follows by the Portmanteau theorem.  $\square$

#### APPENDIX A. PROOFS OF PROPOSITIONS 3.1 AND 3.2

Before beginning the proofs in earnest, we record the following fact, which will be used multiple times below, and which follows straightforwardly from the assumptions of Theorem 1.1.

**Fact A.1.** *Let  $(d^n, n \geq 1)$  be a sequence of degree sequences satisfying the conditions of Theorem 1.1. Then  $\max_{1 \leq i \leq n} d^n(i) = o(n^{1/2})$ .*

*Proof.* If  $p^n \rightarrow p$  pointwise and  $\mu_2(p) \in (0, \infty)$ , then for all  $\epsilon > 0$  there is  $M$  such that

$$\liminf_{n \rightarrow \infty} \sum_{k=1}^M k^2 p^n(k) \geq \mu_2(p) - \epsilon,$$

so  $\sup_{M \geq 1} \liminf_{n \rightarrow \infty} \sum_{k=1}^M k^2 p^n(k) \geq \mu_2(p)$ . If also there is  $\delta > 0$  such that  $\max_{1 \leq i \leq n} d^n(i) \geq \delta n^{1/2}$  for infinitely many  $n$ , then

$$\limsup_{n \rightarrow \infty} \mu_2(p^n) \geq \delta^2 + \sup_{M \geq 1} \liminf_{n \rightarrow \infty} \sum_{k=1}^M k^2 p^n(k) > \mu_2(p),$$

so  $\mu_2(p^n) \not\rightarrow \mu_2(p)$ .  $\square$

To prove Proposition 3.1, we will also make use of the following lemma, which uses the second moment method to control how subsampling affects degree distributions. The proof of the proposition immediately follows that of the lemma.

**Lemma A.2.** *For any integer  $b \geq 1$  there exists  $n_0$  such that for all  $n \geq n_0$  the following holds. Let  $d = (d(1), \dots, d(n))$  be a degree sequence with  $d(i) \geq 1$  for all  $i \in [n]$  and with  $\sum_{i=1}^n d(i) \geq 2n - 1$ , set  $S = \bigcup_{i=1}^n \{i1, \dots, i(d(i) - 1)\}$  and write  $s = |S|$ . Let  $U$  be a uniformly random subset of  $S$  with  $|U| = n - 1$ , and for  $1 \leq i \leq n$  write  $U_i = \#\{1 \leq j < d(i) : (i, j) \in U\}$ . For  $0 \leq a < b$ , write  $P_{b,a} = \#\{1 \leq i \leq n : (d(i), U_i) = (b, a)\}$ . Then for all  $\epsilon > 0$ ,*

$$\mathbf{P}\{|P_{b,a} - \mathbf{E}P_{b,a}| > \epsilon \mathbf{E}P_{b,a}\} \leq \frac{2b^2}{\epsilon^2 s}.$$

*Proof.* We fix  $0 \leq a \leq b$  and compute the first and second moments of  $P_{b+1,a}$ ; this makes the calculations slightly easier to read than they would be for  $P_{b,a}$ .

Fix indices  $k$  and  $\ell$  with  $d(k) = d(\ell) = b + 1$ . Since  $U$  is a uniformly random subset of  $S$ , by symmetry we have

$$\mathbf{P}\{|U_k| = a\} = \mathbf{P}\{|U_\ell| = a\} = \binom{b}{a} \cdot \binom{s-b}{n-1-a} \binom{s}{n-1}^{-1},$$

and

$$\mathbf{P}\{|U_k| = |U_\ell| = a\} = \binom{b}{a}^2 \cdot \binom{s-2b}{n-1-2a} \binom{s}{n-1}^{-1},$$

so writing  $n_{b+1} = \#\{1 \leq i \leq n : d(i) = b+1\}$ , we have

$$\mathbf{E}P_{b+1,a} = n_{b+1} \binom{b}{a} \cdot \binom{s-b}{n-1-a} \binom{s}{n-1}^{-1} \quad (10)$$

and

$$\begin{aligned} & \mathbf{Var}\{P_{b+1,a}\} \\ &= n_{b+1}(n_{b+1}-1) \binom{b}{a}^2 \left( \binom{s-2b}{n-1-2a} \binom{s}{n-1}^{-1} - \binom{s-b}{n-1-a} \binom{s}{n-1}^{-2} \right) \\ &= n_{b+1}(n_{b+1}-1) \binom{b}{a}^2 \left( \frac{(n-1)_{2a} (s-(n-1))_{2(b-a)}}{(s)_{2b}} - \frac{(n-1)_a^2 (s-(n-1))_{b-a}^2}{(s)_b^2} \right) \end{aligned}$$

The ratio of the first and the second term in the final parenthesis is

$$\frac{(n-1)_{2a} (s-(n-1))_{2(b-a)}}{(n-1)_a^2 (s-(n-1))_{b-a}^2} \frac{(s)_b^2}{(s)_{2b}} \leq \frac{(s)_b^2}{(s)_{2b}} \leq \left(1 + \frac{b}{s-2b}\right)^b \leq 1 + \frac{2b^2}{s},$$

the last bound holding for  $b$  fixed and  $s$  large. This gives

$$\mathbf{Var}\{P_{b+1,a}\} \leq n_{b+1}(n_{b+1}-1) \binom{b}{a}^2 \frac{2b^2}{s} \binom{s-b}{n-1-a}^2 \binom{s}{n-1}^{-2} < \frac{2b^2}{s} (\mathbf{E}P_{b+1,a})^2,$$

and the lemma follows by Chebyshev's inequality.  $\square$

*Proof of Proposition 3.1.* We first bound  $\mu_2(q)$  by writing

$$\begin{aligned} \mu_2(q) &= \sum_{a \geq 0} a^2 q(a) \\ &= \sum_{a \geq 0} a^2 \sum_{b > a} p(b) \cdot \mathbf{P}\{\text{Bin}(b-1, \rho) = a\} \\ &\leq \sum_{b > 0} b^2 p(b) \cdot \sum_{0 \leq a < b} \mathbf{P}\{\text{Bin}(b-1, \rho) = a\} \\ &= \mu_2(p)^2 < \infty. \end{aligned}$$

Next, since  $p^n \rightarrow p$  pointwise and  $\mu_2(p^n) \rightarrow \mu_2(p) < \infty$ , for any  $\epsilon > 0$  there is  $k$  such that  $\sum_{d \geq k} d^2 p^n(d) < \epsilon$  and  $\sum_{d \geq k} d^2 p(d) < \epsilon$ . If node  $i$  has  $a$  children in  $T(d^n)$  then  $b^n(i) \geq a+1$ , so it follows that

$$\begin{aligned} \sum_{a \geq k} k^2 \frac{Q_{c^n}(k)}{n} &= \sum_{a \geq k} \sum_{b > a} a^2 \frac{\#\{i \leq n : c^n(i) = a, d^n(i) = b\}}{n} \\ &\leq \sum_{b > k} b^2 \sum_{a < b} \frac{\#\{i \leq n : c^n(i) = a, d^n(i) = b\}}{n} \\ &= \sum_{b > k} b^2 p^n(b) < \epsilon. \end{aligned}$$

To complete the proof it thus suffices to show that  $n^{-1}P_{b,a}^n \rightarrow p(b) \cdot \mathbf{P}\{\text{Bin}(b-1, \rho) = a\}$  in probability for all  $0 \leq a < b$  and that  $n^{-1}Q_{c^n} \rightarrow q$  pointwise in probability; the fact that  $\mu_2(n^{-1}Q_{c^n}) \rightarrow \mu_2(q)$  in probability then immediately follows.

By the third statement of Proposition 2.1, the set of non-root half-edges in  $T(d^n)$  is a uniformly random size- $(n-1)$  subset of the set  $S^n := \bigcup_{1 \leq i \leq n} \{i1, \dots, i(d^n(i)-1)\}$ . We

will apply Lemma A.2 to control the numbers of nodes with a given number of children in  $T(d^n)$ . To make the coming applications of that lemma transparent, we write  $s^n := |S^n| = \sum_{1 \leq i \leq n} (i-1)d^n(i)$ .

We handle the cases  $\mu_1(p) = 2$  and  $\mu_1(p) > 2$  separately. If  $\mu_1(p) = 2$  then  $|S^n| = \sum_{1 \leq i \leq n} (d^n(i) - 1) = (1 + o(1))n$  as  $n \rightarrow \infty$ . Note that in this case  $r = 1/(\mu_1(p) - 1) = 1$  so  $q(a) = p(a+1)$  for all  $a \geq 0$ . For any  $a \geq 0$ , by (10) we then have

$$\mathbf{E}P_{a+1,a}^n = (1 - o(1))np^n(a+1) \binom{a}{a} \binom{|S^n| - 1 - a}{n - 1 - a} \binom{|S^n|}{n - 1}^{-1} = (1 - o(1))np^n(a+1),$$

so by Lemma A.2,

$$\frac{P_{a+1,a}^n}{n} \xrightarrow{\text{prob}} p(a+1) = q(a).$$

It follows that  $\sum_{a \geq 0} P_{a+1,a}^n/n \rightarrow 1$  in probability. This implies that  $\sum_{b > a+1} P_{b,a}^n/n \rightarrow 0$  in probability, so

$$\frac{Q_{c^n}(a)}{n} = \frac{1}{n} \sum_{b > a} P_{b,a}^n = \frac{P_{a+1,a}^n}{n} + \sum_{b > a+1} \frac{P_{b,a}^n}{n} \xrightarrow{\text{prob}} q(a),$$

and that for all  $b > a+1$ ,  $P_{b,a}^n/n \rightarrow 0 = p(b) \cdot \mathbf{P}\{\text{Bin}(b-1, \rho) = a\}$  in probability, as required.

We now assume  $\mu_1(p) > 2$ , so that  $r = 1/(\mu_1(p) - 1) < 1$ . Since  $p = (p_k, k \geq 1)$  is supported on the positive integers,

$$\sum_{a \geq 0} q(a) = \sum_{a \geq 0} \sum_{b=a+1}^{\infty} p(b) \mathbf{P}\{\text{Bin}(b-1, \rho) = a\} = \sum_{b \geq 1} p(b) = 1.$$

Recalling that  $Q_{c^n}(a) = \sum_{b > a} P_{b,a}^n$ , to show that  $n^{-1}Q_{c^n}(a) \rightarrow q(a)$  in probability, it therefore suffices to prove that  $P_{b+1,a}^n/n \rightarrow p(b+1) \cdot \mathbf{P}\{\text{Bin}(b, \rho) = a\}$  for all  $0 \leq a \leq b$ , and we now turn to this.

Since  $\mu_1(p^n) \rightarrow \mu_1(p)$ , it follows that  $|\sum_{i=1}^n d(i)^n - \mu_1(p)n| = n|\mu_1(p^n) - \mu_1(p)| = o(n)$  as  $n \rightarrow \infty$ , so  $s^n = (1 + o(1))n(\mu_1(p) - 1)$ . Thus, for any  $b \geq 1$  and  $0 \leq a \leq b$  we have

$$\begin{aligned} \binom{b}{a} \cdot \binom{s^n - b}{n - 1 - a} \binom{s^n}{n - 1}^{-1} &= \binom{b}{a} \frac{(n-1)_a (s^n - (n-1))_{b-a}}{(s^n)_b} \\ &= (1 - o(1)) \binom{b}{a} \frac{n^a ((\mu_1(p) - 2)n)^{b-a}}{((\mu_1(p) - 1)n)^b} \\ &= (1 - o(1)) \binom{b}{a} \frac{(\mu_1(p) - 2)^{b-a}}{(\mu_1(p) - 1)^b} = (1 - o(1)) \mathbf{P}\{\text{Bin}(b, \rho) = a\}. \end{aligned}$$

Using (10) we thus have

$$\mathbf{E}P_{b+1,a}^n = (1 - o(1))np^n(b+1) \mathbf{P}\{\text{Bin}(b, \rho) = a\} = (1 - o(1))np(b+1) \mathbf{P}\{\text{Bin}(b, \rho) = a\},$$

so by Lemma A.2, as  $n \rightarrow \infty$ ,

$$\frac{P_{b+1,a}^n}{n} \xrightarrow{\text{prob}} p(b+1) \mathbf{P}\{\text{Bin}(b, \rho) = a\},$$

as required.  $\square$

We now turn to controlling the joint degrees of pairs of tree-adjacent vertices in tree-weighted graphs. Given a degree sequence  $\mathbf{d} = (d(1), \dots, d(n))$  and a tree  $t$  with  $v(t) = [n]$ ,

for integers  $b_1, b_2, a_1, a_2$  let

$$\begin{aligned} R_{b_1, b_2, a_1, a_2}(t, d) &= \#\{u \in v(t) \setminus \{r(t)\} : d(u) = b_1, d(\text{par}(u)) = b_2, c_t(u) = a_1, c_t(\text{par}(u)) = a_2\} \\ &= \sum_{u \in v(t) \setminus \{r(t)\}} [d(u) = b_1, c_t(u) = a_1] \cdot [d(\text{par}(u)) = b_2, c_t(\text{par}(u)) = a_2]. \end{aligned}$$

If  $(g, t, \gamma)$  is a tree-rooted graph and  $g$  has degree sequence  $d$ , then  $R_{b_1, b_2, a_1, a_2}(t, d)$  counts the number of edges  $xy$  of  $t$  with  $y = \text{par}(x)$  such that  $c_t(x) = a_1$ ,  $c_t(y) = a_2$  and  $d_g(x) = b_1$ ,  $d_g(y) = b_2$ .

**Proposition A.3.** *Under the assumptions of Theorem 1.1, for any integers  $0 \leq a_1 < b_1$  and  $0 \leq a_2 < b_2$ , as  $n \rightarrow \infty$ ,*

$$\frac{R_{b_1, b_2, a_1, a_2}(T(d^n), d^n)}{n} \rightarrow a_2 p(b_2) \mathbf{P}\{\text{Bin}(b_2 - 1, \rho) = a_2\} \cdot p(b_1) \mathbf{P}\{\text{Bin}(b_1 - 1, \rho) = a_1\}$$

in probability, where  $\rho = 1/(\mu_1(p) - 1)$ .

We introduce two pieces of notation before beginning the proof. For a half-edge  $h$  we write  $v(h)$  for the vertex incident to  $h$ . Also, for  $r \in \mathbb{R}$  we write  $r_+ := \max(r, 0)$ .

*Proof.* First, if  $a_2 = 0$  then the right-hand side is zero, and also  $R_{b_1, b_2, a_1, a_2}(T(d^n), d^n) = 0$ , since if  $v = \text{par}(u) \in T(d^n)$  then  $c_{T(d^n)}(v) \geq 1$ . The result thus holds trivially when  $a_2 = 0$ , and we assume hereafter that  $a_2 \geq 1$ . For the remainder of the proof we write  $R_{b_1, b_2, a_1, a_2} = R_{b_1, b_2, a_1, a_2}(T(d^n), d^n)$  for succinctness.

Let  $\mathcal{H}$  be a fixed, size- $(n-1)$  subset of  $\mathcal{S}^n := \bigcup_{1 \leq i \leq n} \{i1, \dots, i(d^n(i) - 1)\}$ . Write  $\mathcal{S}^n = \{s_1, \dots, s_{n-1}\}$  for the (unordered) set of non-root half-edges of  $T(d^n)$ . We now show that for any half edge  $h \in \mathcal{H}$  and any root half-edge  $r$  not incident to the same vertex as  $h$ , for all  $1 \leq i \leq n-1$ ,

$$\mathbf{P}\{(r_i, s_i) = (r, h) \mid \mathcal{S}^n = \mathcal{H}\} = \mathbf{P}\{(r_1, s_1) = (r, h) \mid \mathcal{S}^n = \mathcal{H}\}. \quad (11)$$

To see this, note that by the second assertion of Proposition 2.1, the number of execution paths with  $\mathcal{S}^n = \mathcal{H}$  is  $((n-1)!)^2$ . We claim that for any  $i \in [n-1]$ , the number of execution paths with  $\mathcal{S}^n = \mathcal{H}$  which additionally satisfy that  $(r_i, s_i) = (r, h)$  is  $((n-2)!)^2$ . As this number does not depend on  $i \in [n-1]$ , the displayed identity follows from this claim.

To prove the claim, simply note that there are  $(n-2)!$  possible orderings of  $\mathcal{H}$  consistent with the constraint that  $s_i = h$ . Having fixed such an ordering  $(h_1, \dots, h_{n-1})$ , for each  $j \in [n-1]$  with  $j \neq i$ , if  $s_k = h_k$  for  $1 \leq k < j$  then, excluding  $r_i$  there are  $n-j-1_{[j < i]}$  unpaired root half-edges in components different from that of  $s_j$ , and any such root half-edge may be chosen as  $r_j$ . Thus the number of execution paths with  $\mathcal{S}^n = \mathcal{H}$  and such that  $(r_i, s_i) = (r, h)$  is  $(n-2)! \cdot \prod_{j \in [n-1] \setminus \{i\}} (n-j-1_{[j < i]}) = ((n-2)!)^2$ .

Now fix a second non-root half-edge  $h' \neq h$  and a second root half-edge  $r' \neq r$  not incident to the same vertex as  $h'$ . Then a similar argument to the one leading to (11) shows that for any  $1 \leq i < j \leq n$ ,

$$\begin{aligned} &\mathbf{P}\{(r_i, s_i) = (r, h), (r_j, s_j) = (r', s') \mid \mathcal{S}^n = \mathcal{H}\} \\ &= \mathbf{P}\{(r_1, s_1) = (r, h), (r_2, s_2) = (r', s') \mid \mathcal{S}^n = \mathcal{H}\}. \end{aligned} \quad (12)$$

In the current case, the number of execution paths leading to the events in both the left- and right-hand probabilities is  $((n-3)!)^2$ .

We will next use the above identities in order to perform first and second moment computations. For any set  $H \subset \mathcal{S}^n$ , for  $0 \leq a < b$  let  $V_{b,a}^n(H) = \{i \in [n] : d^n(i) = b, |H \cap \{i1, \dots, i(d^n(i) - 1)\}| = a\}$ . Note that  $V_{b,a}^n(\mathcal{S}^n)$  is simply the set of nodes with degree  $b$  in  $G(d^n)$  and with  $a$  children in  $T(d^n)$ ; so  $P_{b,a}^n = |V_{b,a}^n(\mathcal{S}^n)|$ .

Fix a non-root node  $u \in T(d^n)$ , and let  $m \in [n-1]$  be such that  $e_m = \{\text{par}(u), u\}$ . Then  $v(r_m) = u$  and  $v(s_m) = \text{par}(u)$ , so  $u \in R_{b_1, b_2, a_1, a_2}$  if and only if  $v(r_m) \in V_{b_1, a_1}^n(\mathcal{S}^n)$  and  $v(s_m) \in V_{b_2, a_2}^n(\mathcal{S}^n)$ . By (11), it follows that

$$\begin{aligned} & \mathbf{E}(R_{b_1, b_2, a_1, a_2} \mid \mathcal{S}^n = \mathcal{H}) \\ &= (n-1) \mathbf{P} \{v(r_1) \in V_{b_1, a_1}^n(\mathcal{S}^n), v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}\}. \end{aligned} \quad (13)$$

Likewise, by the second of the two displayed identities,

$$\begin{aligned} & \mathbf{E} \left( \binom{R_{b_1, b_2, a_1, a_2}}{2} \mid \mathcal{S}^n = \mathcal{H} \right) \\ &= \binom{n-1}{2} \mathbf{P} \{v(r_1), v(r_2) \in V_{b_1, a_1}^n(\mathcal{S}^n), v(s_1), v(s_2) \in V_{b_2, a_2}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}\}. \end{aligned} \quad (14)$$

We develop the latter two identities in turn.

For integers  $0 \leq a < b$ , the number of non-root half-edges  $h \in \mathcal{S}^n$  with  $v(h) \in V_{b, a}^n(\mathcal{S}^n)$  is  $a \cdot |V_{b, a}^n(\mathcal{S}^n)|$ , and the number of root half-edges  $h$  with  $v(h) \in V_{b, a}^n(\mathcal{S}^n)$  is just  $|V_{b, a}^n(\mathcal{S}^n)|$ . Conditionally given that  $\mathcal{S}^n = \mathcal{H}$ , the half-edge  $s_1$  is a uniformly random element of  $\mathcal{H}$ , so

$$\mathbf{P} \{v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}\} = \frac{a_2 |V_{b_2, a_2}^n(\mathcal{H})|}{|\mathcal{H}|} = \frac{a_2 |V_{b_2, a_2}^n(\mathcal{H})|}{n-1}.$$

Having chosen  $s_1$ , if  $v(s_1) = v$  then  $v(r_1)$  is a uniformly random element of  $[n] \setminus \{v\}$ , so

$$\mathbf{P} \{v(r_1) \in V_{b_1, a_1}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}, v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n)\} = \frac{(|V_{b_1, a_1}^n(\mathcal{H})| - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+}{n-1}.$$

Using these two identities in (13), it follows that

$$(n-1) \mathbf{E}(R_{b_1, b_2, a_1, a_2} \mid \mathcal{S}^n = \mathcal{H}) = a_2 |V_{b_2, a_2}^n(\mathcal{H})| (|V_{b_1, a_1}^n(\mathcal{H})| - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+,$$

so since  $|V_{b, a}^n(\mathcal{S}^n)| = P_{b, a}^n$  for all  $0 \leq a < b$ , by Proposition 3.1 we have

$$\begin{aligned} \mathbf{E} \left( \frac{R_{b_1, b_2, a_1, a_2}}{n} \mid \mathcal{S}^n \right) &= \frac{1}{n(n-1)} a_2 P_{b_2, a_2}^n (P_{b_1, a_1}^n - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]}) \\ &\xrightarrow{\text{prob}} a_2 p(b_2) \mathbf{P} \{\text{Bin}(b_2 - 1, \rho) = a_2\} \cdot p(b_1) \mathbf{P} \{\text{Bin}(b_1 - 1, \rho) = a_1\}. \end{aligned} \quad (15)$$

For the second moment calculation, we need to additionally compute

$$\mathbf{P} \{v(r_2) \in V_{b_1, a_1}^n(\mathcal{S}^n), v(s_2) \in V_{b_2, a_2}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}, v(r_1) \in V_{b_1, a_1}^n(\mathcal{S}^n), v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n)\}.$$

Under the conditioning in the preceding displayed probability, the number of non-root half-edges  $h \in \mathcal{S}^n \setminus \{s_1\}$  with  $v(h) \in V_{b_2, a_2}^n(\mathcal{H})$  is  $(a_2 \cdot |V_{b_2, a_2}^n(\mathcal{H})| - 1)_+$ , so

$$\begin{aligned} & \mathbf{P} \{v(s_2) \in V_{b_2, a_2}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}, v(r_1) \in V_{b_1, a_1}^n(\mathcal{S}^n), v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n)\} \\ &= \frac{(a_2 \cdot |V_{b_2, a_2}^n(\mathcal{H})| - 1)_+}{n-2}. \end{aligned}$$

Now suppose that  $\mathcal{S}^n = \mathcal{H}$ ,  $v(r_1) \in V_{b_1, a_1}^n(\mathcal{S}^n)$ ,  $v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n)$ , and that  $v(s_2) \in V_{b_2, a_2}^n(\mathcal{H})$ , and consider the number of possible values for  $r_2$ . We claim that the number of unpaired root half-edges  $h$  with  $v(h) \in V_{b_1, a_1}^n(\mathcal{S}^n)$  such that  $v(h)$  is in a component different from  $v(s_2)$  is

$$(|V_{b_1, a_1}^n(\mathcal{H})| - 1 - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+.$$

To see this, note that if  $(b_1, a_1) = (b_2, a_2)$  and either  $v(s_2) = v(s_1)$  or  $v(s_2) = v(r_1)$ , then we are precisely constrained to choose  $h$  so that  $v(h) \in V_{b_1, a_1}^n(\mathcal{H}) \setminus \{v(r_1), v(s_1)\}$ . On the other hand, if  $(b_1, a_1) = (b_2, a_2)$  and  $v(s_2) \notin \{v(r_1), v(s_1)\}$  then we are constrained to choose  $h$  so that  $v(h) \in V_{b_1, a_1}^n(\mathcal{H}) \setminus \{v(r_1), v(s_2)\}$ . Both cases agree with the above



formula. When  $(b_1, a_1) = (b_2, a_2)$ , the claim is straightforward, since in that case we are only constrained to choose  $h$  so that  $v(h) \in V_{b_1, a_1}^n(\mathcal{H}) \setminus \{v(r_1)\}$ . It follows that

$$\begin{aligned} & \mathbf{P} \left\{ v(r_2) \in V_{b_1, a_1}^n(\mathcal{S}^n) \mid \mathcal{S}^n = \mathcal{H}, v(s_2) \in V_{b_2, a_2}^n(\mathcal{S}^n), v(r_1) \in V_{b_1, a_1}^n(\mathcal{S}^n), v(s_1) \in V_{b_2, a_2}^n(\mathcal{S}^n) \right\} \\ &= \frac{(|V_{b_1, a_1}^n(\mathcal{H})| - 1 - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+}{n - 2}. \end{aligned}$$

Combining the above identities with (14) yields that

$$\begin{aligned} & 2(n-1)(n-2) \mathbf{E} \left( \binom{R_{b_1, b_2, a_1, a_2}}{2} \mid \mathcal{S}^n = \mathcal{H} \right) \\ &= a_2 |V_{b_2, a_2}^n(\mathcal{H})| (a_2 |V_{b_2, a_2}^n(\mathcal{H})| - 1)_+ \\ &\quad \cdot (|V_{b_1, a_1}^n(\mathcal{H})| - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+ (|V_{b_1, a_1}^n(\mathcal{H})| - 1 - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+, \end{aligned}$$

so since  $a_2 \geq 1$ , Proposition 3.1 implies that

$$\begin{aligned} & \mathbf{E} \left( \frac{R_{b_1, b_2, a_1, a_2} (R_{b_1, b_2, a_1, a_2} - 1)}{n^2} \mid \mathcal{S}^n \right) \\ &= \frac{a_2 P_{b_2, a_2}^n (a_2 P_{b_2, a_2}^n - 1)_+}{n(n-1)} \cdot \frac{(P_{b_1, a_1}^n - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+ (P_{b_1, a_1}^n - 1 - \mathbf{1}_{[(b_1, a_1) = (b_2, a_2)]})_+}{n(n-2)} \\ &\xrightarrow{\text{prob}} (a_2 p(b_2) \mathbf{P} \{ \text{Bin}(b_2 - 1, \rho) = a_2 \} \cdot p(b_1) \mathbf{P} \{ \text{Bin}(b_1 - 1, \rho) = a_1 \})^2. \end{aligned}$$

Also, (15) implies that  $\mathbf{E} (n^{-2} R_{b_1, b_2, a_1, a_2} \mid \mathcal{S}^n) \xrightarrow{\text{prob}} 0$ , which with the preceding asymptotic implies that

$$\mathbf{E} \left( \frac{R_{b_1, b_2, a_1, a_2}^2}{n^2} \mid \mathcal{S}^n \right) \xrightarrow{\text{prob}} (a_2 p(b_2) \mathbf{P} \{ \text{Bin}(b_2 - 1, \rho) = a_2 \} \cdot p(b_1) \mathbf{P} \{ \text{Bin}(b_1 - 1, \rho) = a_1 \})^2.$$

Combining this with (15) gives that

$$\mathbf{E} \left( \left( \frac{R_{b_1, b_2, a_1, a_2}}{n} \right)^2 \mid \mathcal{S}^n \right) - \left( \mathbf{E} \left( \frac{R_{b_1, b_2, a_1, a_2}}{n} \mid \mathcal{S}^n \right) \right)^2 \xrightarrow{\text{prob}} 0;$$

the conditional Chebyshev's inequality then gives that for all  $\epsilon > 0$ ,

$$\mathbf{P} \left\{ \left| \frac{R_{b_1, b_2, a_1, a_2}}{n} - \mathbf{E} \left( \frac{R_{b_1, b_2, a_1, a_2}}{n} \mid \mathcal{S}^n \right) \right| > \epsilon \mid \mathcal{S}^n \right\} \xrightarrow{\text{prob}} 0.$$

Taking expectations on the left of the previous inequality to remove the conditioning, and again using (15), this time to replace the term  $\mathbf{E} \left( \frac{R_{b_1, b_2, a_1, a_2}}{n} \mid \mathcal{S}^n \right)$  in the probability by the constant  $C := a_2 p(b_2) \mathbf{P} \{ \text{Bin}(b_2 - 1, \rho) = a_2 \} \cdot p(b_1) \mathbf{P} \{ \text{Bin}(b_1 - 1, \rho) = a_1 \}$ , we obtain that

$$\mathbf{P} \left\{ \left| \frac{R_{b_1, b_2, a_1, a_2}}{n} - C \right| > \epsilon \right\} \rightarrow 0,$$

as required.  $\square$

*Proof of Proposition 3.2.* We may reexpress  $A^n(k, \ell)$  as

$$\begin{aligned} A^n(k, \ell) &= \sum_{a_1, a_2 \geq 0} \sum_{u \in v(T(d^n))} \mathbf{1}_{[r(T(d^n)) \notin \{u, \text{par}(u)\}]} \\ &\quad \cdot \mathbf{1}_{[d^n(u) = k + a_1 + 1, c^n(u) = a_1]} \\ &\quad \cdot \mathbf{1}_{[d^n(\text{par}(u)) = \ell + a_2 + 1, c^n(\text{par}(u)) = a_2]} \\ &+ \sum_{a_1, a_2 \geq 0} \sum_{u \in v(T(d^n))} \mathbf{1}_{[\text{par}(u) = r(T(d^n))]} \\ &\quad \cdot \mathbf{1}_{[d^n(u) = k + a_1, c^n(u) = a_1]} \\ &\quad \cdot \mathbf{1}_{[d^n(\text{par}(u)) = \ell + a_2, c^n(\text{par}(u)) = a_2]} \end{aligned}$$

For fixed  $a_1, a_2 \geq 0$ , if we replace  $\mathbf{1}_{[r(T(d^n)) \notin \{u, \text{par}(u)\}]}$  by  $\mathbf{1}_{[u \neq r(T(d^n))]}$  in the first double sum, then the inner sum is simply  $R_{k+a_1+1, \ell+a_2+1, a_1, a_2}$ . It follows that

$$\begin{aligned} A^n(k, \ell) &= \sum_{a_1, a_2 \geq 0} R_{k+a_1+1, \ell+a_2+1, a_1, a_2} \\ &+ \sum_{a_1, a_2 \geq 0} \sum_{u \in v(T(d^n))} \mathbf{1}_{[\text{par}(u) = r(T(d^n))]} \\ &\quad \cdot \mathbf{1}_{[d^n(u) = k + a_1, c^n(u) = a_1]} \\ &\quad \cdot \mathbf{1}_{[d^n(\text{par}(u)) = \ell + a_2, c^n(\text{par}(u)) = a_2]} \\ &- \sum_{a_1, a_2 \geq 0} \sum_{u \in v(T(d^n))} \mathbf{1}_{[\text{par}(u) = r(T(d^n))]} \\ &\quad \cdot \mathbf{1}_{[d^n(u) = k + a_1 + 1, c^n(u) = a_1]} \\ &\quad \cdot \mathbf{1}_{[d^n(\text{par}(u)) = \ell + a_2 + 1, c^n(\text{par}(u)) = a_2]} \end{aligned}$$

But each of the last two double sums is bounded by  $c^n(r(T(d^n)))$ , since they both count each child of the root at most once. Under the assumptions of Theorem 1.1, by Fact A.1 we have  $c^n(r(T(d^n))) \leq \max_{1 \leq i \leq n} d^n(i) = o(n^{1/2})$ , so the preceding identity gives

$$\left| A^n(k, \ell) - \sum_{a_1, a_2 \geq 0} R_{k+a_1+1, \ell+a_2+1, a_1, a_2} \right| = o(n^{1/2}).$$

Since

$$n^{-1} R_{k+a_1+1, \ell+a_2+1, a_1, a_2}$$

$$\xrightarrow{\text{prob}} a_2 p(\ell + a_2 + 1) \mathbf{P}\{\text{Bin}(\ell + a_2, \rho) = a_2\} \cdot p(k + a_1 + 1) \mathbf{P}\{\text{Bin}(k + a_1, \rho) = a_1\}$$

by Proposition A.3, and summing the right-hand side of the last expression over  $k, l \geq 0$  gives  $\alpha(k, l)$ , it follows that for any  $\epsilon > 0$ ,

$$\mathbf{P}\{A^n(k, l)/n \geq \alpha(k, l) - \epsilon\} \rightarrow 1.$$

But also  $n^{-1} \sum_{k, l \geq 0} A^n(k, l) = |e(T(d^n))| = (n-1)/n \rightarrow 1$ ; so since  $\sum_{k, l \geq 0} \alpha(k, l) = 1$ , we must in fact have that

$$\frac{A^n(k, l)}{n} \xrightarrow{\text{prob}} \alpha(k, l)$$

for all  $k, l \geq 0$ , as required.

It remains to show that  $n^{-1} \sum_{k, l \geq 0} kl A^n(k, l) \xrightarrow{\text{prob}} \sum_{k, l \geq 0} kl \alpha(k, l)$ . For this we will exploit the exchangeability of Pitman's additive coalescent. Recall the notation  $v(h)$  for

the vertex incident to half-edge  $h$ . Note that for any  $M \in \mathbb{N}$  we have

$$\begin{aligned} \sum_{k,l \geq 0} klA^n(k,l) - \sum_{0 \leq k,l \leq M} klA^n(k,l) &= \sum_{uv \in e(T^n)} d_-^n(u)d_-^n(v) \mathbf{1}_{[\max(d_-^n(u), d_-^n(v)) > M]} \\ &\leq \sum_{uv \in e(T^n)} d^n(u)d^n(v) \mathbf{1}_{[\max(d^n(u), d^n(v)) > M]} \\ &= \sum_{i=1}^{n-1} d^n(v(r_i))d^n(v(s_i)) \mathbf{1}_{[\max(d^n(v(r_i)), d^n(v(s_i))) > M]}. \end{aligned}$$

Now, by Proposition 2.1 (3) and the identity (11), for any  $1 \leq i \leq n-1$  we have

$$\begin{aligned} &\mathbf{E} \left( d^n(v(r_i))d^n(v(s_i)) \mathbf{1}_{[\max(d^n(v(r_i)), d^n(v(s_i))) > M]} \right) \\ &= \mathbf{E} \left( d^n(v(r_1))d^n(v(s_1)) \mathbf{1}_{[\max(d^n(v(r_1)), d^n(v(s_1))) > M]} \right), \end{aligned}$$

and by the definition of Pitman's additive coalescent we have

$$\begin{aligned} &\mathbf{E} \left( d^n(v(r_1))d^n(v(s_1)) \mathbf{1}_{[\max(d^n(v(r_1)), d^n(v(s_1))) > M]} \right) \\ &= \sum_{u \in [n]} \sum_{v \in [n]} d^n(u)d^n(v) \mathbf{1}_{[\max(d^n(u), d^n(v)) > M]} \mathbf{P} \{v(s_1) = u, v(r_1) = v\} \\ &= \sum_{u \in [n]} \sum_{v \in [n]} d^n(u)d^n(v) \mathbf{1}_{[\max(d^n(u), d^n(v)) > M]} \cdot \frac{d^n(u)}{n\mu_1(p^n)} \cdot \frac{1}{n-1}, \end{aligned}$$

where we have used that  $\sum_{i \in [n]} d^n(i) = n\mu_1(p^n)$ . Next,

$$\begin{aligned} &\sum_{u \in [n]} \sum_{v \in [n]} (d^n(u))^2 d^n(v) \mathbf{1}_{[\max(d^n(u), d^n(v)) > M]} \\ &\leq \left( \left( \sum_{u \in [n]: d^n(u) > M} (d^n(u))^2 \right) \cdot \sum_{v \in [n]} d^n(v) + \left( \sum_{v \in [n]: d^n(v) > M} d^n(v) \right) \cdot \sum_{u \in [n]} (d^n(u))^2 \right), \\ &= n\mu_1(p^n) \cdot \sum_{u \in [n]: d^n(u) > M} (d^n(u))^2 + n\mu_2(p^n) \cdot \sum_{v \in [n]: d^n(v) > M} d^n(v). \end{aligned}$$

Since  $p^n \rightarrow p$ ,  $\mu_1(p^n) \rightarrow \mu_1(p)$  and  $\mu_2(p^n) \rightarrow \mu_2(p)$ , for any  $\delta > 0$  we may choose  $M = M(\delta)$  sufficiently large so that  $\sum_{u \in [n]: d^n(u) > M} (d^n(u))^2 < \delta n$  and  $\sum_{v \in [n]: d^n(v) > M} d^n(v) < \delta n$ , for all  $n \geq 1$ . For such  $M$ , the previous bound and the two identities which precede it yield that

$$\mathbf{E} \left( \sum_{k,l \geq 0} klA^n(k,l) - \sum_{0 \leq k,l \leq M} klA^n(k,l) \right) \leq \frac{1}{\mu_1(p^n)} (\delta n \mu_1(p^n) + \delta n \mu_2(p^n)).$$

By Markov's inequality, it follows that for all  $\epsilon > 0$  there is  $M \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$ ,

$$\mathbf{P} \left\{ \sum_{k,l \geq 0} klA^n(k,l) - \sum_{0 \leq k,l \leq M} klA^n(k,l) > \epsilon n \right\} < \epsilon.$$

Finally, since  $n^{-1}A^n(k,l) \xrightarrow{\text{prob}} \alpha(k,l)$ , it follows that for all  $M \in \mathbb{N}$  we have

$$\frac{1}{n} \sum_{0 \leq k,l \leq M} klA^n(k,l) \xrightarrow{\text{prob}} \sum_{0 \leq k,l \leq M} kl\alpha(k,l),$$

so the preceding probability bound implies that

$$\frac{1}{n} \sum_{k,l \geq 0} klA^n(k,l) \xrightarrow{\text{prob}} \lim_{M \rightarrow \infty} \sum_{0 \leq k,l \leq m} kl\alpha(k,l) = \sum_{k,l \geq 0} kl\alpha(k,l),$$

as required.  $\square$

### APPENDIX B. PROOF OF THEOREM 4.1

Let  $H$  be a simple graph with vertex set  $v(H) = [n]$ , and let  $G$  be a random graph with degree sequence  $\mathbf{d} = (d(1), \dots, d(n))$  sampled according to the configuration model. Recall the definitions of  $\mathcal{L}(G), \mathcal{M}(G, H), \mathcal{N}(G, H)$  and  $L(G), M(G, H), N(G, H)$  from Section 4. The first subsection will provide a quantitative approximation result for mixed moments of  $L, M$  and  $N$ . In the second subsection, we will use this approximation to prove Theorem 4.1.

**B.1. Deterministic bounds on loops and multi-edges.** Our arguments in this section are based on and fairly closely parallel those from [19, Chapter 7]. We recall the falling factorial notation  $(x)_\ell = x(x-1)\dots(x-\ell+1)$ . In what follows, it is convenient to define  $(x)_\ell = 1$  if  $\ell = 0$ , and  $(x)_\ell = 0$  if  $\ell < 0$ .

**Proposition B.1.** *Write  $m = \frac{1}{2} \sum_{i=1}^n d(i)$ , and write  $d_{\max} = \max\{d(1), \dots, d(n)\}$ . For any positive integers  $q, r, s \in \mathbb{N}$ ,*

$$\left| \mathbf{E}((L)_q(M)_r(N)_s) - \frac{(|\mathcal{L}|)_q(|\mathcal{M}|)_r(|\mathcal{N}|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i} \right| \leq C(S_1 + S_2)$$

where  $C$  is a constant depending only on  $q, r$  and  $s$ ,  $S_1$  is defined by the following identity,

$$\begin{aligned} S_1 \prod_{i=0}^{q+2r+s-1} (2m-1-2i) &= (|\mathcal{L}|)_{q-2}(|\mathcal{M}|)_r(|\mathcal{N}|)_s \sum_{1 \leq u \leq n} d(u)^3 \\ &\quad + (|\mathcal{L}|)_{q-1}(|\mathcal{M}|)_{r-1}(|\mathcal{N}|)_s \sum_{1 \leq u \neq v \leq n} d(u)^3 d(v)^2 \\ &\quad + (|\mathcal{L}|)_{q-1}(|\mathcal{M}|)_r(|\mathcal{N}|)_{s-1} \sum_{uv \in e(H)} d(u)^2 d(v) \\ &\quad + (|\mathcal{L}|)_q(|\mathcal{M}|)_{r-2}(|\mathcal{N}|)_s \sum_{1 \leq u \neq v_1, v_2 \leq n} d(u)^3 d(v_1)^2 d(v_2)^2 \\ &\quad + (|\mathcal{L}|)_q(|\mathcal{M}|)_{r-1}(|\mathcal{N}|)_{s-1} \sum_{uv \in e(H)} d(u)^2 d(v)^3 \\ &\quad + (|\mathcal{L}|)_q(|\mathcal{M}|)_r(|\mathcal{N}|)_{s-2} \sum_{uv_1, uv_2 \in e(H)} d(u) d(v_1) d(v_2), \end{aligned}$$

and  $S_2$  is defined by

$$S_2 = (|\mathcal{L}|)_q(|\mathcal{N}|)_s \sum_{k=1}^{r-1} (|\mathcal{M}|)_{r-k} \sum_{\ell=0}^k d_{\max}^{2\ell} \prod_{i=0}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i}.$$

*Proof.* Throughout the proof, write

$$(x, y, z) = ((x_1, \dots, x_q), (y_1, \dots, y_r), (z_1, \dots, z_s)) \in \mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s$$

to denote a generic element of  $\mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s$ . For  $(x, y, z) \in \mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s$ , write

$$\mathbf{1}_{[x]} = \prod_{i=1}^q \mathbf{1}_{[x_i]}, \quad \mathbf{1}_{[y]} = \prod_{i=1}^r \mathbf{1}_{[y_i]}, \quad \mathbf{1}_{[z]} = \prod_{i=1}^s \mathbf{1}_{[z_i]}.$$

We say  $(x, y, z)$  is *non-repeating* if  $x_1, x_2, \dots, x_q$  are pairwise distinct,  $y_1, y_2, \dots, y_r$  are pairwise distinct, and  $z_1, z_2, \dots, z_s$  are pairwise distinct. In what follows, write

$$\sum^* := \sum_{\substack{(x,y,z) \in \mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s \\ (x,y,z) \text{ is non-repeating}}},$$

and, for  $S \subset \mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s$ , write

$$\sum_S^* := \sum_{\substack{(x,y,z) \in S \\ (x,y,z) \text{ is non-repeating}}}.$$

Note that  $(L)_q(M)_r(N)_s = \sum^* \mathbf{1}_{[x]} \mathbf{1}_{[y]} \mathbf{1}_{[z]}$ , so

$$\mathbf{E}((L)_q(M)_r(N)_s) = \sum^* \mathbf{P} \{ \mathbf{1}_{[x]} \mathbf{1}_{[y]} \mathbf{1}_{[z]} = 1 \}. \quad (16)$$

We say  $(x, y, z)$  is *non-conflicting* if the  $2q + 4r + 2s$  half-edges appearing in  $x, y$  and  $z$  are pairwise distinct, and otherwise we say  $(x, y, z)$  is conflicting. Since half-edges in  $G$  are paired uniformly at random, for non-conflicting  $(x, y, z)$  we have

$$\mathbf{P} \{ \mathbf{1}_{[x]} \mathbf{1}_{[y]} \mathbf{1}_{[z]} = 1 \} = \prod_{i=0}^{q+2r+s-1} \frac{1}{2m-1-2i}, \quad (17)$$

Now, for a given  $(x, y, z)$ , let  $er(x, y, z)$  be defined as follows:

$$er(x, y, z) := \mathbf{P} \{ \mathbf{1}_{[x]} \mathbf{1}_{[y]} \mathbf{1}_{[z]} = 1 \} - \prod_{i=0}^{q+2r+s-1} \frac{1}{2m-1-2i}.$$

By (17), if  $(x, y, z)$  is non-conflicting then  $er(x, y, z) = 0$ , so

$$\begin{aligned} \mathbf{E}((L)_q(M)_r(N)_s) &= \sum^* \prod_{i=0}^{q+2r+s-1} \frac{1}{2m-1-2i} + \sum^* er(x, y, z) \\ &= \frac{(|\mathcal{L}|)_q (|\mathcal{M}|)_r (|\mathcal{N}|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i} + \sum_{(x,y,z) \text{ is conflicting}}^* er(x, y, z). \end{aligned}$$

By the triangle inequality this implies

$$\left| \mathbf{E}((L)_q(M)_r(N)_s) - \frac{(|\mathcal{L}|)_q (|\mathcal{M}|)_r (|\mathcal{N}|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i} \right| \leq \sum_{(x,y,z) \text{ is conflicting}}^* |er(x, y, z)|. \quad (18)$$

To bound the error terms, we must make a distinction between two types of conflicts. This distinction is most easily understood by way of an example. On the one hand, suppose  $x_1 = (ui, uj_1)$  and  $x_2 = (ui, uj_2)$  for  $u \in V(G)$  and distinct  $i, j_1, j_2 \in [d(u)]$ . Then  $\mathbf{1}_{[x_1]} \mathbf{1}_{[x_2]} = 1$  is the event that the half edge  $ui$  is joined to  $uj_1$  and  $uj_2$  simultaneously, and  $\mathbf{P} \{ \mathbf{1}_{[x_1]} \mathbf{1}_{[x_2]} = 1 \} = 0$ . On the other hand, suppose  $y_1 = (ui_1, vj_1), (ui_2, vj_2)$  and  $y_2 = (ui_1, vj_1), (ui_3, vj_3)$  for distinct  $u, v \in V(G)$ , distinct  $i_1, i_2, i_3 \in [d(u)]$ , and distinct  $j_1, j_2, j_3 \in [d(v)]$ . Then  $\mathbf{1}_{[y_1]} \mathbf{1}_{[y_2]} = 1$  is the event that  $u$  and  $v$  are connected by a triple edge, and  $\mathbf{P} \{ \mathbf{1}_{[y_1]} \mathbf{1}_{[y_2]} = 1 \} > 0$ .

We say a conflicting triple  $(x, y, z)$  is a *bad conflict* if  $\mathbf{P} \{ \mathbf{1}_{[x]} \mathbf{1}_{[y]} \mathbf{1}_{[z]} = 1 \} = 0$ , and otherwise we say  $(x, y, z)$  is a *good conflict*. In the above examples, the first is a bad conflict and the second is a good conflict. Let  $\mathcal{B} = \mathcal{B}(q, r, s) \subseteq \mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s$  and  $\mathcal{G} = \mathcal{G}(q, r, s) \subseteq \mathcal{L}^q \times \mathcal{M}^r \times \mathcal{N}^s$  be the collections of bad conflicts and good conflicts respectively. The rest of this proof is dedicated to bounding  $|\mathcal{B}|, |\mathcal{G}|$ , and  $er(x, y, z)$  for  $(x, y, z) \in \mathcal{B} \cup \mathcal{G}$ .

(*Bounding  $|\mathcal{B}|$* ): If  $(x, y, z)$  is a bad conflict then one of the following must hold:

- (1) There exists  $1 \leq a < b \leq q$ ,  $u \in V(G)$  and distinct  $i_1, i_2, i_3 \in [d(u)]$  such that  $x_a = (ui_1, ui_2)$  and  $x_b = (ui_1, ui_3)$ . Write  $\mathcal{B}_{xx}$  for the set of  $(x, y, z) \in \mathcal{B}$  that contain a pair  $(x_a, x_b)$  of this form.
- (2) There exists  $1 \leq a \leq q$ ,  $1 \leq b \leq r$ , distinct  $u, v \in V(G)$ , distinct  $i_1, i_2, i_3 \in [d(u)]$ , and distinct  $j_1, j_2 \in [d(v)]$  such that  $x_a = (ui_1, ui_3)$  and  $y_b = (ui_1, vj_1), (ui_2, vj_2)$ . Write  $\mathcal{B}_{xy}$  for the set of  $(x, y, z) \in \mathcal{B}$  that contain a pair  $(x_a, y_b)$  of this form.
- (3) There exists  $1 \leq a \leq q$ ,  $1 \leq b \leq s$ ,  $uv \in e(H)$ , distinct  $i_1, i_2 \in [d(u)]$ , and  $j \in [d(v)]$  such that  $x_a = (ui_1, ui_2)$  and  $z_b = (ui_1, vj)$ . Write  $\mathcal{B}_{xz}$  for the set of  $(x, y, z) \in \mathcal{B}$  that contain a pair  $(x_a, z_b)$  of this form.
- (4) There exists  $1 \leq a < b \leq r$ , distinct  $u, v_1, v_2 \in V(G)$ , distinct  $i_1, i_2, i_3 \in [d(u)]$ , distinct  $j_1, j_2 \in [d(v_1)]$ , and distinct  $k_1, k_2 \in [d(v_2)]$  such that  $y_a = (ui_1, v_1j_1), (ui_2, v_1j_2)$  and  $y_b = (ui_1, v_2k_1), (ui_3, v_2k_2)$ . Write  $\mathcal{B}_{yy}$  for the set of  $(x, y, z) \in \mathcal{B}$  that contain a pair  $(y_a, y_b)$  of this form.
- (5) There exists  $1 \leq a \leq r$ ,  $1 \leq b \leq s$  and distinct  $u, v_1, v_2 \in V(G)$  such that  $uv_2 \in e(H)$ , distinct  $i_1, i_2 \in [d(u)]$ , distinct  $j_1, j_2 \in [d(v_1)]$ , and  $k \in [d(v_2)]$  such that  $y_a = (ui_1, v_1j_1), (ui_2, v_1j_2)$  and  $z_b = (ui_1, v_2k)$ . Write  $\mathcal{B}_{yz}$  for the set of  $(x, y, z) \in \mathcal{B}$  that contain a pair  $(y_a, z_b)$  of this form.
- (6) There exists  $1 \leq a < b \leq s$ , distinct  $uv_1, uv_2 \in e(H)$ ,  $i \in [d(u)]$ ,  $j \in [d(v_1)]$ , and  $k \in [d(v_2)]$  such that  $z_a = (ui, v_1j)$  and  $z_b = (ui, v_2k)$ . Write  $\mathcal{B}_{zz}$  for the set of  $(x, y, z) \in \mathcal{B}$  that contain a pair  $(z_a, z_b)$  of this form.

We next turn to bounding the sizes of each set, starting with  $\mathcal{B}_{xx}$ . The number of choices for  $a$  and  $b$  is  $q(q-1)/2$ . Having chosen these, for each possible choice of  $u \in V(G)$ , there are less than  $d(u)^3$  choices for  $i_1, i_2$  and  $i_3$ . Then, having chosen these, we must choose one of  $i_1, i_2$  and  $i_3$  to be repeated in  $x_a$  and  $x_b$ . Lastly, we must choose the remaining  $q-2$  entries for  $x$ ,  $r$  entries for  $y$ , and  $s$  entries for  $z$ . Hence,

$$\begin{aligned} |\mathcal{B}_{xx}| &\leq (|\mathcal{L}|)_{q-2}(|\mathcal{M}|)_r(|\mathcal{N}|)_s q(q-1)/2 \sum_{1 \leq u \leq n} 3d(u)^3 \\ &= C_1(|\mathcal{L}|)_{q-2}(|\mathcal{M}|)_r(|\mathcal{N}|)_s \sum_{1 \leq u \leq n} d(u)^3, \end{aligned}$$

where  $C_1 = C_1(q) = 3q(q-1)/2$ . Note that  $\mathcal{B}_{xx}$  is empty if  $q \leq 1$ , so  $|\mathcal{B}_{xx}| = 0$ . In this case the right hand side is also zero by our convention that  $\binom{k}{\ell} = 0$  for  $\ell < 0$ . Therefore, the bound also holds for  $q \leq 1$ . The subsequent bounds can likewise be seen to hold when the right hand side is zero, though we do not explicitly verify this in every case.

When building an element of  $\mathcal{B}_{xy}$ , the number of ways to choose  $a$  and  $b$  is  $qr$ . Having chosen these, for each pair  $u, v \in V(G)$ , there are less than  $d(u)^3 d(v)^2$  choices for  $i_1, i_2, i_3 \in [d(u)]$  and  $j_1, j_2 \in [d(v)]$ . Then, there are a constant number of ways to arrange the half-edges in  $x_a$  and  $y_b$ . Lastly, we must choose the remaining entries for  $x, y$  and  $z$ . Hence,

$$|\mathcal{B}_{xy}| \leq C_2(|\mathcal{L}|)_{q-1}(|\mathcal{M}|)_{r-1}(|\mathcal{N}|)_s \sum_{1 \leq u \neq v \leq n} d(u)^3 d(v)^2,$$

where  $C_2$  depends only on  $q$  and  $r$ .

For an element of  $\mathcal{B}_{xz}$ , for each  $uv \in e(H)$ , there are less than  $d(u)^2 d(v)$  ways to choose  $i_1, i_2 \in [d(u)]$  and  $j \in [d(v)]$ . Hence,

$$|\mathcal{B}_{xz}| \leq C_3(|\mathcal{L}|)_{q-1}(|\mathcal{M}|)_r(|\mathcal{N}|)_{s-1} \sum_{uv \in e(H)} d(u)^2 d(v).$$

For  $\mathcal{B}_{yy}$ , for each  $u, v_1, v_2 \in V(G)$ , there are less than  $d(u)^3 d(v_1)^2 d(v_2)^2$  ways to choose  $i_1, i_2, i_3 \in [d(u)]$ ,  $j_1, j_2 \in [d(v_1)]$ , and  $k_1, k_2 \in [d(v_2)]$ . Hence,

$$|\mathcal{B}_{yy}| \leq C_4(|\mathcal{L}|)_q(|\mathcal{M}|)_{r-2}(|\mathcal{N}|)_s \sum_{1 \leq u \neq v_1 \neq v_2 \leq n} d(u)^3 d(v_1)^2 d(v_2)^2.$$

For  $\mathcal{B}_{yz}$ , for each  $u, v_1, v_2 \in V(G)$  such that  $uv_2 \in e(H)$ , there are less than  $d(u)^2 d(v_1)^2 d(v_2)$  ways to choose  $i_1, i_2 \in [d(u)]$ ,  $j_1, j_2 \in [d(v_1)]$ , and  $k \in [d(v_2)]$ . Hence,

$$|\mathcal{B}_{yz}| \leq C_5 (|\mathcal{L}|)_q (|\mathcal{M}|)_{r-1} (|\mathcal{N}|)_{s-1} \sum_{\substack{1 \leq u \neq v_1 \neq v_2 \leq n \\ uv_2 \in e(H)}} d(u)^2 d(v_1)^2 d(v_2).$$

Lastly, for  $\mathcal{B}_{zz}$ , for each  $uv_1, uv_2 \in e(H)$ , there are less than  $d(u)d(v_1)d(v_2)$  ways to choose  $i \in [d(u)]$ ,  $j \in [d(v_1)]$  and  $k \in [d(v_2)]$ . Hence,

$$|\mathcal{B}_{zz}| \leq C_6 (|\mathcal{L}|)_q (|\mathcal{M}|)_r (|\mathcal{N}|)_{s-2} \sum_{uv_1, uv_2 \in e(H)} d(u)d(v_1)d(v_2).$$

Note that the values of  $C_1, C_2, C_3, C_4, C_5$  and  $C_6$  depend only on  $q, r$  and  $s$ .

(Bounding  $|er(x, y, z)|$  for  $(x, y, z) \in \mathcal{B}$ ): If  $(x, y, z) \in \mathcal{B}$  then  $\mathbf{P}\{\mathbf{1}_{[x]}\mathbf{1}_{[y]}\mathbf{1}_{[z]} = 1\} = 0$ , meaning

$$|er(x, y, z)| = \prod_{i=0}^{q+2r+s-1} \frac{1}{2m-1-2i} \quad (19)$$

(Bounding  $|\mathcal{G}|$ ): Suppose  $(x, y, z) \in \mathcal{G}$ . Then  $(x, y, z)$  is conflicting, meaning a half-edge appears more than once in  $x \cup y \cup z$ . However, since  $\mathbf{P}\{\mathbf{1}_{[x]}\mathbf{1}_{[y]}\mathbf{1}_{[z]}\} > 0$  for  $(x, y, z) \in \mathcal{G}$ , it cannot be the case where  $\mathbf{1}_{[x]}\mathbf{1}_{[y]}\mathbf{1}_{[z]}$  contains the event that a half-edge is paired to two different half-edges simultaneously. Hence, there must be a half-edge pair that appears more than once in  $x \cup y \cup z$ . Furthermore, this half-edge pair must appear more than once in  $y$ , since the edges in  $y$  and  $z$  are disjoint. It follows that any  $(x, y, z) \in \mathcal{G}$  can be constructed in the following way:

- (1) Choose  $x$  and  $z$  arbitrarily. The number of choices here is  $(|\mathcal{L}|)_q (|\mathcal{N}|)_s$ .
- (2) Choose a set of indices  $1 \leq a_1 < a_2 < \dots < a_k \leq r$  and arbitrarily choose the elements  $y_a \in y$  such that  $a \neq a_i$  for all  $1 \leq i \leq k$ . The number of choices here is  $(|\mathcal{M}|)_{r-k}$  times a constant in terms of  $r$ .
- (3) Choose  $1 \leq b_1 \leq \dots \leq b_\ell \leq r$  such that  $\{b_1, \dots, b_\ell\} \subseteq \{a_1, \dots, a_k\}$ . For each  $1 \leq i \leq \ell$ , build  $y_{b_i}$  by choosing a half-edge pair already in  $y$ , then choosing the other half-edge pair arbitrarily. The number of choices for each  $i$  is less than  $d_{\max}^2$  times a constant in terms of  $r$ .
- (4) For each  $a \in \{a_1, \dots, a_k\} \setminus \{b_1, \dots, b_\ell\}$ , build  $y_a$  by choosing two half-edge pairs already in  $y$ . The number of choices here is a constant in terms of  $r$ .

Since every element of  $\mathcal{G}$  can be constructed in this way, we get

$$|\mathcal{G}| \leq C(r) (|\mathcal{L}|)_q (|\mathcal{N}|)_s \sum_{k=1}^{r-1} (|\mathcal{M}|)_{r-k} \sum_{\ell=0}^k d_{\max}^{2\ell}.$$

(Bounding  $er(x, y, z)$  for  $(x, y, z) \in \mathcal{G}$ ): Let  $(x, y, z) \in \mathcal{G}$  and suppose we can construct  $(x, y, z)$  as above with a particular  $k$  and  $\ell$ . Then there are  $2k - \ell$  half-edge pairs that are redundant when calculating  $\mathbf{P}\{\mathbf{1}_{[x]}\mathbf{1}_{[y]}\mathbf{1}_{[z]} = 1\}$ . Hence,

$$\mathbf{P}\{\mathbf{1}_{[x]}\mathbf{1}_{[y]}\mathbf{1}_{[z]} = 1\} = \prod_{i=0}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i}.$$

Therefore, for such  $(x, y, z) \in \mathcal{G}$ ,

$$\begin{aligned} |er(x, y, z)| &= \left| \prod_{i=0}^{q+2r+s-1} \frac{1}{2m-1-2i} - \prod_{i=0}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i} \right| \\ &\leq \prod_{i=0}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i}. \end{aligned} \quad (20)$$

Finally, by (18), we know that

$$\left| \mathbf{E}((L)_q(M)_r(N)_s) - \frac{(|\mathcal{L}|)_q(|\mathcal{M}|)_r(|\mathcal{N}|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i} \right| \leq \sum_{(x,y,z) \in \mathcal{B}}^* |er(x, y, z)| + \sum_{(x,y,z) \in \mathcal{G}}^* |er(x, y, z)|,$$

from which the result now follows by using the bounds on  $|\mathcal{B}_{xx}|, |\mathcal{B}_{xy}|, |\mathcal{B}_{xz}|, |\mathcal{B}_{yy}|, |\mathcal{B}_{yz}|, |\mathcal{B}_{zz}|$  and on  $|\mathcal{G}|$ , together with (19) and (20)  $\square$

**B.2. The probability of simplicity for a random superposition of graphs.** Before proving Theorem 4.1, it will be useful to show some auxiliary bounds.

**Lemma B.2.** *Under the assumptions of Theorem 4.1, we have*

$$\sum_{v \in [n]} d^n(v) = O(n), \quad (21)$$

$$\sum_{v \in [n]} (d^n(v))^2 = O(n), \quad (22)$$

$$\sum_{uv \in e(h_n)} d^n(u)d^n(v) = O(n), \quad (23)$$

and

$$\sup_{u \in [n]} \sum_{v: uv \in e(h_n)} d^n(v) = o(n), \quad (24)$$

*Proof.* Equations (21) and (22) follow from the fact that  $\mu_1(p^n) \rightarrow \mu_1(p) < \infty$  and  $\mu_2(p^n) \rightarrow \mu_2(p) < \infty$ . Indeed, we have

$$\sum_{v \in [n]} d^n(v) = n \sum_{k \geq 1} k p^n(k) = n \mu_1(p^n) = O(n),$$

and

$$\sum_{v \in [n]} (d^n(v))^2 = n \sum_{k \geq 1} k^2 p^n(k) = n \mu_2(p^n) = O(n).$$

Equation (23) follows from the convergence of  $\sum_{i,j \geq 1} ij \alpha^n(i, j)$ . Notice that, by the definition of  $\alpha^n$ ,

$$\sum_{uv \in e(h_n)} d^n(u)d^n(v) = \sum_{i,j \geq 1} \left( \sum_{uv \in e(h_n): d^n(u)=i, d^n(v)=j} ij \right) = \sum_{i,j \geq 1} ij (n \alpha^n(i, j)).$$

Hence,

$$\frac{1}{n} \sum_{uv \in e(h_n)} d^n(u)d^n(v) \rightarrow \sum_{i,j \geq 1} ij \alpha(i, j) < \infty,$$

implying that

$$\sum_{uv \in e(h_n)} d^n(u)d^n(v) = O(n).$$



We will prove the fourth bound by contradiction. To this end, suppose (24) fails. Then we can find  $c > 0$  and a sequence of vertices  $(u_n, n \geq 1)$  with  $u_n \in v(h_n)$  such that for all  $n$  sufficiently large,

$$\sum_{v:u_nv \in e(h_n)} d^n(v) \geq cn. \quad (25)$$

Write  $\deg(u_n) = \deg_{h_n}(u_n)$  and list the neighbours of  $u_n$  in  $h_n$  as  $N_{h_n}(u_n) = \{v_1^n, \dots, v_{\deg(u_n)}^n\}$  so that  $d^n(v_i^n) \geq d^n(v_{i+1}^n)$  for all  $1 \leq i < \deg(u_n)$ .

Next, fix  $D \in \mathbb{N}$  and let  $k = k(n) = \max\{i : d^n(v_i^n) \geq D\}$ . Then

$$\sum_{i=k+1}^{\deg(u_n)} d^n(v_i^n) \leq (\deg(u_n) - k)(D - 1) = o(n),$$

the last bound holding since  $\deg(u_n) = o(n)$  by assumption. Thus,

$$\begin{aligned} \sum_{v:u_nv \in e(h_n)} d^n(v)^2 &= \sum_{i=1}^{\deg(u_n)} d^n(v_i^n)^2 \\ &\geq \sum_{i=1}^k d^n(v_i^n)^2 \\ &\geq D \sum_{i=1}^k d^n(v_i^n) \\ &= D \left( \sum_{v:u_nv \in e(h_n)} d^n(v) - o(n) \right) \\ &\geq D(c - o(1))n, \end{aligned}$$

the last bound holding by (25). Since  $D \in \mathbb{N}$  was arbitrary, it follows that

$$\sum_{v \in [n]} d^n(v)^2 \geq \sum_{v:u_nv \in e(h_n)} d^n(v)^2 = \omega(n),$$

contradicting (22).  $\square$

The next lemma is the last ingredient needed, and also assumes  $p(0) + p(1) < 1$ , i.e. an asymptotically non-zero proportion of the degrees in  $G_n$  are 2 or greater. We will show later that Theorem 4.1 is straightforward when  $p(0) + p(1) = 1$ .

**Lemma B.3.** *Suppose that  $p(0) + p(1) < 1$ . Then*

$$|\mathcal{L}(G_n)| = \Theta(n), \quad (26)$$

and

$$|\mathcal{M}(G_n, h_n)| = \Theta(n^2), \quad (27)$$

*Proof.* Let  $\mathcal{L}_n = \mathcal{L}(G_n)$ ,  $\mathcal{M}_n = \mathcal{M}(G_n, h_n)$ , and  $\mathcal{N}_n = \mathcal{N}(G_n, h_n)$ . By their definitions, we have

$$\begin{aligned} |\mathcal{L}_n| &= \sum_{v \in [n]} \frac{d^n(v)(d^n(v) - 1)}{2}, \text{ and} \\ |\mathcal{M}_n| &= \sum_{\{u < v: uv \notin e(h_n)\}} \frac{d^n(u)(d^n(u) - 1)}{2} d^n(v)(d^n(v) - 1). \end{aligned}$$

For the upper bounds, by Lemma B.2 we have

$$|\mathcal{L}_n| = \sum_{v \in [n]} \frac{d^n(v)(d^n(v) - 1)}{2} \leq \sum_{v \in [n]} d^n(v)^2 = O(n), \text{ and}$$

$$|\mathcal{M}_n| = \sum_{\{u < v : uv \notin e(h_n)\}} \frac{d^n(u)(d^n(u) - 1)}{2} d^n(v)(d^n(v) - 1) \leq \left( \sum_{v \in [n]} d^n(v)^2 \right)^2 = O(n^2).$$

For the lower bounds, first notice that

$$\begin{aligned} |\mathcal{L}_n| &= \sum_{v \in [n]} \frac{d^n(v)(d^n(v) - 1)}{2} \\ &= \sum_{v \in [n] : d^n(v) > 1} \frac{d^n(v)(d^n(v) - 1)}{2} \\ &\geq |\{v \in [n] : d^n(v) > 1\}| \\ &= n(1 - p^n(0) - p^n(1)) \end{aligned}$$

Since  $p^n(0) + p^n(1) \rightarrow p(0) + p(1) < 1$  by assumption, this implies  $|\{v \in [n] : d^n(v) > 1\}| = \Theta(n)$ , and so

$$|\mathcal{L}_n| \geq |\{v \in [n] : d^n(v) > 1\}| = \Theta(n).$$

Similarly, we have

$$|\mathcal{M}_n| \geq |\{(u, v) : u < v, uv \notin e(h_n) \text{ and } d^n(u), d^n(v) > 1\}|.$$

Notice also that since  $|e(h_n)| = o(n^2)$ , we have

$$\begin{aligned} &|\{(u, v) : u < v, uv \notin e(h_n) \text{ and } d^n(u), d^n(v) > 1\}| \\ &= |\{(u, v) : u < v, d^n(u), d^n(v) > 1\}| - o(n^2), \end{aligned}$$

and writing  $k = k(n) = |\{v \in [n] : d^n(v) > 1\}|$ , we have

$$|\{(u, v) : u < v, d^n(u), d^n(v) > 1\}| = \binom{k}{2} = \Theta(k^2) = \Theta(n^2),$$

and hence,  $|\mathcal{M}_n| = \Theta(n^2)$ .  $\square$

*Proof of Theorem 4.1.* Let  $m = m(n) = \frac{1}{2} \sum_{v \in [n]} d^n(v)$ , and let  $q, r$  and  $s$  be positive integers. Also, in what follows write  $d_{\max}^n = \max_{1 \leq i \leq n} d^n(i)$ ; by Fact A.1 we know that  $d_{\max}^n = o(n^{1/2})$ .

Assume for the time being that  $p(0) + p(1) < 1$  and that  $\eta > 0$ . Notice that when  $\eta > 0$  and  $\mu_1(p) > 0$  we have

$$\frac{|\mathcal{N}_n|}{n} = \frac{1}{n} \sum_{uv \in e(h_n)} d^n(u)d^n(v) = \sum_{i, j \geq 1} ij\alpha^n(i, j) \rightarrow \sum_{i, j \geq 1} ij\alpha(i, j) = \mu_1(p)\eta > 0,$$

meaning  $|\mathcal{N}_n| = \Theta(n)$ .

We first claim that

$$\mathbf{E}((L_n)_q(M_n)_r(N_n)_s) = \frac{(|\mathcal{L}_n|)_q(|\mathcal{M}_n|)_r(|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} 2m - 1 - 2i} (1 + o(1)).$$

From Proposition B.1 we know that

$$\left| \mathbf{E}((L_n)_q(M_n)_r(N_n)_s) - \frac{(|\mathcal{L}_n|)_q(|\mathcal{M}_n|)_r(|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} 2m - 1 - 2i} \right| \leq C(S_1 + S_2),$$

where  $S_1$  is defined by the relationship

$$\begin{aligned}
& S_1 \cdot \prod_{i=0}^{q+2r+s-1} (2m-1-2i) \\
&= (|\mathcal{L}_n|)_{q-2} (|\mathcal{M}_n|)_r (|\mathcal{N}_n|)_s \sum_{v \in [n]} (d^n(v))^3 \\
&+ (|\mathcal{L}_n|)_{q-1} (|\mathcal{M}_n|)_{r-1} (|\mathcal{N}_n|)_s \sum_{1 \leq u \neq v \leq n} (d^n(u))^3 (d^n(v))^2 \\
&+ (|\mathcal{L}_n|)_{q-1} (|\mathcal{M}_n|)_r (|\mathcal{N}_n|)_{s-1} \sum_{uv \in e(h_n)} (d^n(u))^2 d^n(v) \\
&+ (|\mathcal{L}_n|)_q (|\mathcal{M}_n|)_{r-2} (|\mathcal{N}_n|)_s \sum_{1 \leq u \neq v_1, v_2 \leq n} (d^n(u))^3 (d^n(v_1))^2 (d^n(v_2))^2 \\
&+ (|\mathcal{L}_n|)_q (|\mathcal{M}_n|)_{r-1} (|\mathcal{N}_n|)_{s-1} \sum_{uv \in e(h_n)} (d^n(u))^2 (d^n(v))^3 \\
&+ (|\mathcal{L}_n|)_q (|\mathcal{M}_n|)_r (|\mathcal{N}_n|)_{s-2} \sum_{uv_1, uv_2 \in e(h_n)} d^n(u) d^n(v_1) d^n(v_2),
\end{aligned}$$

and  $S_2$  is defined by

$$S_2 = (|\mathcal{L}_n|)_q (|\mathcal{N}_n|)_s \sum_{k=1}^{r-1} (|\mathcal{M}_n|)_{r-k} \sum_{\ell=0}^k (d_{\max}^n)^{2\ell} \prod_{i=0}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i};$$

recall that we set  $(k)_\ell = 0$  if  $\ell < 0$ . We now show that

$$S_1 = o\left(\frac{(|\mathcal{L}_n|)_q (|\mathcal{M}_n|)_r (|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} (2m-1-2i)}\right) \text{ and } S_2 = o\left(\frac{(|\mathcal{L}_n|)_q (|\mathcal{M}_n|)_r (|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} (2m-1-2i)}\right). \quad (28)$$

We will start by bounding  $S_1$ . Since  $|\mathcal{L}_n|, |\mathcal{M}_n|, |\mathcal{N}_n| \rightarrow \infty$  as  $n \rightarrow \infty$ , to prove the first bound in (28) it suffices to establish the following bounds:

$$\begin{aligned}
& \sum_{v \in [n]} (d^n(v))^3 = o(|\mathcal{L}_n|^2), \\
& \sum_{1 \leq u \neq v \leq n} (d^n(u))^3 (d^n(v))^2 = o(|\mathcal{L}_n| |\mathcal{M}_n|), \\
& \sum_{uv \in e(h_n)} (d^n(u))^2 d^n(v) = o(|\mathcal{L}_n| |\mathcal{N}_n|), \\
& \sum_{1 \leq u \neq v_1, v_2 \leq n} (d^n(u))^3 (d^n(v_1))^2 (d^n(v_2))^2 = o(|\mathcal{M}_n|^2), \\
& \sum_{uv \in e(h_n)} (d^n(u))^2 (d^n(v))^3 = o(|\mathcal{M}_n| |\mathcal{N}_n|), \text{ and} \\
& \sum_{uv_1, uv_2 \in e(h_n)} d^n(u) d^n(v_1) d^n(v_2) = o(|\mathcal{N}_n|^2).
\end{aligned} \quad (29)$$

Using Lemmas B.2 and B.3, together with the fact that  $d_{\max}^n = o(n^{1/2})$ , we get the following results:

$$\sum_{v \in [n]} (d^n(v))^3 \leq d_{\max}^n \sum_{v \in [n]} (d^n(v))^2 = o(n^{1/2}) \cdot O(n) = o(n^2) = o(|\mathcal{L}_n|^2),$$

$$\begin{aligned}
\sum_{1 \leq u \neq v \leq n} (d^n(u))^3 (d^n(v))^2 &\leq d_{\max}^n \sum_{1 \leq u \neq v \leq n} (d^n(u))^2 (d^n(v))^2 \\
&\leq d_{\max}^n \left( \sum_{v \in [n]} (d^n(v))^2 \right)^2 = o(n^3) = o(|\mathcal{L}_n| |\mathcal{M}_n|), \\
\sum_{uv \in e(h_n)} (d^n(u))^2 d^n(v) &\leq d_{\max}^n \sum_{uv \in e(h_n)} d^n(u) d^n(v) = d_{\max}^n |\mathcal{N}_n| = o(|\mathcal{L}_n| |\mathcal{N}_n|), \\
\sum_{1 \leq u \neq v_1, v_2 \leq n} (d^n(u))^3 (d^n(v_1))^2 (d^n(v_2))^2 &\leq d_{\max}^n \sum_{1 \leq u \neq v_1, v_2 \leq n} (d^n(u))^2 (d^n(v_1))^2 (d^n(v_2))^2 \\
&\leq d_{\max}^n \left( \sum_{v \in [n]} (d^n(v))^2 \right)^3 = o(n^4) = o(|\mathcal{M}_n|^2), \\
\sum_{uv \in e(h_n)} (d^n(u))^2 (d^n(v))^3 &\leq (d_{\max}^n)^3 \sum_{uv \in e(h_n)} d^n(u) d^n(v) = o(n^{3/2}) |\mathcal{N}_n| = o(|\mathcal{M}_n| |\mathcal{N}_n|),
\end{aligned}$$

and

$$\begin{aligned}
\sum_{uv_1, uv_2 \in e(h_n)} d^n(u) d^n(v_1) d^n(v_2) &= \sum_{uv \in e(h_n)} d^n(u) d^n(v) \cdot \sum_{w: uw \in e(h_n)} d^n(w) \\
&\leq \sum_{uv \in e(h_n)} d^n(u) d^n(v) \cdot \left( \sup_{u \in [n]} \sum_{w: uw \in e(h_n)} d^n(w) \right) \\
&= |\mathcal{N}_n| \cdot o(n) \\
&= o(|\mathcal{N}_n|^2),
\end{aligned}$$

the last bound holding as  $\mathcal{N}_n = \Theta(n)$ .

To prove the bound on  $S_2$  from (28), first notice that

$$\begin{aligned}
S_2 &= (|\mathcal{L}_n|)_q (|\mathcal{N}_n|)_s \sum_{k=1}^{r-1} (|\mathcal{M}_n|)_{r-k} \sum_{\ell=0}^k (d_{\max}^n)^{2\ell} \prod_{i=0}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i} \\
&= \frac{(|\mathcal{L}_n|)_q (|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+s-1} 2m-1-2i} \sum_{k=1}^{r-1} (|\mathcal{M}_n|)_{r-k} \sum_{\ell=0}^k (d_{\max}^n)^{2\ell} \prod_{q+s}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i}.
\end{aligned}$$

Hence, it suffices to show the following:

$$\sum_{k=1}^{r-1} (|\mathcal{M}_n|)_{r-k} \sum_{\ell=0}^k (d_{\max}^n)^{2\ell} \prod_{q+s}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i} = o\left( \frac{(|\mathcal{M}_n|)_r}{\prod_{q+s}^{q+2r+s-1} 2m-1-2i} \right).$$

Since  $r$  is fixed, we need only show that for arbitrary  $k \in [1, r-1]$  and  $\ell \in [0, k]$ ,

$$(|\mathcal{M}_n|)_{r-k} (d_{\max}^n)^{2\ell} \prod_{q+s}^{q+2r+s-1-(2k-\ell)} \frac{1}{2m-1-2i} = o\left( \frac{(|\mathcal{M}|)_r}{\prod_{q+s}^{q+2r+s-1} 2m-1-2i} \right).$$

By cancelling out some terms, this follows if we can show that

$$(d_{\max}^n)^{2\ell} = o\left( \frac{(|\mathcal{M}_n| - (r-k))_k}{\prod_{q+2r+s-(2k-\ell)}^{q+2r+s-1} 2m-1-2i} \right).$$

Now since  $m = \Theta(n)$ ,  $|\mathcal{M}_n| = \Theta(n^2)$ , and  $r$  is a constant, this in turn holds, provided that

$$\begin{aligned} (d_{\max}^n)^{2\ell} &= o\left(\frac{n^{2k}}{n^{2k-\ell}}\right) \\ &= o\left(n^\ell\right), \end{aligned}$$

which holds since  $d_{\max}^n = o(n^{1/2})$ .

Therefore,

$$S_1 + S_2 = o\left(\frac{(|\mathcal{L}_n|)_q(|\mathcal{M}_n|)_r(|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i}\right),$$

which proves that

$$\mathbf{E}((L_n)_q(M_n)_r(N_n)_s) = \frac{(|\mathcal{L}_n|)_q(|\mathcal{M}_n|)_r(|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i} (1 + o(1)).$$

Furthermore, since  $q, r$  and  $s$  are fixed, we obtain that

$$\begin{aligned} &\frac{(|\mathcal{L}_n|)_q(|\mathcal{M}_n|)_r(|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} 2m-1-2i} \\ &= \frac{|\mathcal{L}_n|^q |\mathcal{M}_n|^r |\mathcal{N}_n|^s}{(2m)^{q+2r+s}} (1 + o(1)) \\ &= \left(\frac{|\mathcal{L}_n|}{\sum_{v \in [n]} d^n(v)}\right)^q \left(\frac{|\mathcal{M}_n|}{\left(\sum_{v \in [n]} d^n(v)\right)^2}\right)^r \left(\frac{|\mathcal{N}_n|}{\sum_{v \in [n]} d^n(v)}\right)^s (1 + o(1)) \end{aligned}$$

Next, we claim that

$$\frac{|\mathcal{L}_n|}{\sum_{v \in [n]} d^n(v)} \rightarrow \nu/2, \quad \frac{|\mathcal{M}_n|}{\left(\sum_{v \in [n]} d^n(v)\right)^2} \rightarrow \nu^2/4, \quad \text{and} \quad \frac{|\mathcal{N}_n|}{\sum_{v \in [n]} d^n(v)} \rightarrow \eta.$$

For the first of these three claims, we have

$$\begin{aligned} \frac{|\mathcal{L}_n|}{\sum_{v \in [n]} d^n(v)} &= \frac{\frac{1}{2} \sum_{v \in [n]} d^n(v)(d^n(v)-1)}{\sum_{v \in [n]} d^n(v)} = \frac{1}{2} \frac{\sum_{v \in [n]} (d^n(v))^2 - \sum_{v \in [n]} d^n(v)}{\sum_{v \in [n]} d^n(v)} \\ &= \frac{1}{2} \left(\frac{\mu_2(p^n) - \mu_1(p^n)}{\mu_1(p^n)}\right) \rightarrow \nu/2. \end{aligned}$$

For the second, we have

$$\begin{aligned} &\frac{|\mathcal{M}_n|}{\left(\sum_{v \in [n]} d^n(v)\right)^2} \\ &= \frac{\frac{1}{2} \sum_{\{u < v: uv \notin e(h_n)\}} d^n(u)(d^n(u)-1)d^n(v)(d^n(v)-1)}{\left(\sum_{v \in [n]} d^n(v)\right)^2} \\ &= \frac{1}{4 \left(\sum_{v \in [n]} d^n(v)\right)^2} \left( \left[ \sum_{v \in [n]} d^n(v)(d^n(v)-1) \right]^2 - \sum_{v \in [n]} [d^n(v)(d^n(v)-1)]^2 \right. \\ &\quad \left. - \sum_{uv \in e(h_n)} d^n(u)(d^n(u)-1)d^n(v)(d^n(v)-1) \right). \end{aligned}$$

The second and third terms vanish in the limit since  $\mu_1(p) > 0$  and so

$$\left( \sum_{v \in [n]} d^n(v) \right)^2 = \Omega(n^2),$$

and

$$\sum_{v \in [n]} [d^n(v) (d^n(v) - 1)]^2 \leq (d_{\max}^n)^2 \sum_{v \in [n]} (d^n(v))^2 = (d_{\max}^n)^2 n \mu_2(p^n) = o(n^2),$$

and

$$\begin{aligned} \sum_{uv \in e(h_n)} d^n(u) (d^n(u) - 1) d^n(v) (d^n(v) - 1) &\leq (d_{\max}^n)^2 \sum_{uv \in e(h_n)} d^n(u) d^n(v) \\ &= (d_{\max}^n)^2 |\mathcal{N}_n| \\ &= o(n^2). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{|\mathcal{M}_n|}{\left( \sum_{v \in [n]} d^n(v) \right)^2} &= \frac{1}{4} \frac{\left[ \sum_{v \in [n]} d^n(v) (d^n(v) - 1) \right]^2}{\left( \sum_{v \in [n]} d^n(v) \right)^2} (1 + o(1)) \\ &\rightarrow \frac{1}{4} \frac{(\mu_2(p) - \mu_1(p))^2}{(\mu_1(p))^2} \\ &= \nu^2/4. \end{aligned}$$

For the third claim, we have

$$\begin{aligned} \frac{|\mathcal{N}_n|}{\sum_{v \in [n]} d^n(v)} &= \frac{\sum_{uv \in e(h_n)} d^n(u) d^n(v)}{\sum_{v \in [n]} d^n(v)} = \frac{\sum_{i,j \geq 1} ij \alpha^n(i, j)}{\mu_1(p^n)} \\ &\rightarrow \frac{\sum_{i,j \geq 1} ij \alpha(i, j)}{\mu_1(p)} = \eta. \end{aligned}$$

Therefore,

$$\begin{aligned} &\lim_{n \rightarrow \infty} \mathbf{E}((L_n)_q (M_n)_r (N_n)_s) \\ &= \lim_{n \rightarrow \infty} \frac{(|\mathcal{L}_n|)_q (|\mathcal{M}_n|)_r (|\mathcal{N}_n|)_s}{\prod_{i=0}^{q+2r+s-1} 2m - 1 - 2i} (1 + o(1)) \tag{30} \\ &= \lim_{n \rightarrow \infty} \left( \frac{|\mathcal{L}_n|}{\sum_{v \in [n]} d^n(v)} \right)^q \left( \frac{|\mathcal{M}_n|}{\left( \sum_{v \in [n]} d^n(v) \right)^2} \right)^r \left( \frac{|\mathcal{N}_n|}{\sum_{v \in [n]} d^n(v)} \right)^s (1 + o(1)) \\ &= (\nu/2)^q (\nu^2/4)^r (\eta)^s. \end{aligned}$$

It then follows, by Theorem 2.6 of [19], that the random variables  $L_n$ ,  $M_n$  and  $N_n$  converge to independent Poisson random variables with parameters  $\nu/2$ ,  $\nu^2/4$  and  $\eta$  respectively. This proves the theorem in the case that  $p(0) + p(1) < 1$  and  $\eta > 0$ .

Lastly we will deal with the cases that arise if  $p(0) + p(1) = 1$  or if  $\eta = 0$ . First, if  $\eta = 0$  then  $\lim_{n \rightarrow \infty} \sum_{i,j \geq 1} ij \alpha^n(i, j) = \sum_{i,j \geq 1} ij \alpha(i, j) = 0$ . Recall that for any two half-edges  $ui$  and  $vj$  with  $u, v \in [n]$ ,  $i \in [d^n(u)]$  and  $j \in [d^n(v)]$ ,  $\mathbf{P} \{ \mathbf{1}_{[ui, vj]} \} = \frac{1}{2m-1}$ . So by (16),

since  $m = m(n) = n\mu_1(p^n)/2 = \Theta(n)$ , we have that

$$\begin{aligned} \mathbf{E}(N_n) &= \sum_{uv \in e(h_n)} \sum_{i \in [d^n(u)]} \sum_{j \in [d^n(v)]} \mathbf{P}\{\mathbf{1}_{[ui, vj]}\} = \frac{\sum_{uv \in e(h_n)} d^n(u)d^n(v)}{2m-1} \\ &= \frac{n}{2m-1} \sum_{i, j \geq 1} ij\alpha^n(i, j) = o(1). \end{aligned}$$

Hence,  $\lim_{n \rightarrow \infty} \mathbf{E}(N_n) = 0$ . In this case, if  $p(0) + p(1) < 1$  then a reprise of the argument leading to (30) gives that for all  $q, r \geq 1$ ,

$$\lim_{n \rightarrow \infty} \mathbf{E}((L_n)_q (M_n)_r) = (\nu/2)^q (\nu^2/4)^r,$$

and therefore that  $L_n$  and  $M_n$  converge to independent Poisson random variables with parameters  $\nu/2$  and  $\nu^2/4$  respectively.

Lastly, we deal with the case when  $p(0) + p(1) = 1$ ; In this case,  $\mu_2(p) = \mu_1(p)$ , so  $\nu = 0$ . Furthermore,

$$\mathbf{E}(L_n) = \frac{\frac{1}{2} \sum_{v \in [n]} d^n(v)(d^n(v) - 1)}{2m - 1} = \frac{n}{4m - 2} (\mu_2(p^n) - \mu_1(p^n)).$$

Since  $\mu_2(p^n) - \mu_1(p^n) \rightarrow \mu_2(p) - \mu_1(p) = 0$ , we get that  $\lim_{n \rightarrow \infty} \mathbf{E}(L_n) = 0$ , and an analogous argument shows that  $\lim_{n \rightarrow \infty} \mathbf{E}(M_n) = 0$ . In this case, another reprise of the argument leading to (30) gives that for all  $s \geq 1$ ,

$$\lim_{n \rightarrow \infty} \mathbf{E}((N_n)_s) = \eta^s,$$

so  $N_n$  is asymptotically Poisson( $\eta$ ) distributed.  $\square$

#### REFERENCES

- [1] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18:no. 14, 21, 2013. doi: 10.1214/EJP.v18-2116. URL <https://arxiv.org/abs/1202.5464>.
- [2] Nathanaël Berestycki, Benoît Laslier, and Gourab Ray. Critical exponents on Fortuin-Kasteleyn weighted planar maps. *Comm. Math. Phys.*, 355(2):427–462, 2017. ISSN 0010-3616. doi: 10.1007/s00220-017-2933-7.
- [3] Olivier Bernardi. Bijective counting of tree-rooted maps and shuffles of parenthesis systems. *Electron. J. Combin.*, 14(1):Research Paper 9, 36, 2007. URL [http://www.combinatorics.org/Volume\\_14/Abstracts/v14i1r9.html](http://www.combinatorics.org/Volume_14/Abstracts/v14i1r9.html).
- [4] Olivier Bernardi. Tutte polynomial, subgraphs, orientations and sandpile model: new connections via embeddings. *Electron. J. Combin.*, 15(1):Research Paper 109, 53, 2008. URL [http://www.combinatorics.org/Volume\\_15/Abstracts/v15i1r109.html](http://www.combinatorics.org/Volume_15/Abstracts/v15i1r109.html).
- [5] Mireille Bousquet-Mélou and Julien Courtiel. Spanning forests in regular planar maps. *J. Combin. Theory Ser. A*, 135:1–59, 2015. ISSN 0097-3165. doi: 10.1016/j.jcta.2015.04.002. URL <https://arxiv.org/abs/1306.4536>.
- [6] Nicolas Broutin and Jean-François Marckert. Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures Algorithms*, 44(3):290–316, 2014. ISSN 1042-9832. doi: 10.1002/rsa.20463. URL <https://arxiv.org/abs/1110.5203>.
- [7] Jason Miller Ewain Gwynne and Scott Sheffield. The Tutte embedding of the mated-CRT map converges to Liouville quantum gravity. arXiv:1705.11161 [math.PR], May 2017. URL <https://arxiv.org/abs/1705.11161>.
- [8] Catherine Greenhill, Matthew Kwan, and David Wind. On the number of spanning trees in random regular graphs. *Electron. J. Combin.*, 21(1):Paper 1.45, 26, 2014. doi: 10.37236/3752. URL <https://www.combinatorics.org/ojs/index.php/eljc/article/view/v21i1p45>.
- [9] Catherine Greenhill, Mikhail Isaev, Matthew Kwan, and Brendan D. McKay. The average number of spanning trees in sparse graphs with given degrees. *European J. Combin.*, 63:6–25,

2017. ISSN 0195-6698. doi: 10.1016/j.ejc.2017.02.003. URL <https://arxiv.org/abs/1606.01586>.
- [10] Ewain Gwynne, Adrien Kassel, Jason Miller, and David B. Wilson. Active spanning trees with bending energy on planar maps and SLE-decorated Liouville quantum gravity for  $\kappa > 8$ . *Comm. Math. Phys.*, 358(3):1065–1115, 2018. ISSN 0010-3616. doi: 10.1007/s00220-018-3104-1. URL <https://arxiv.org/abs/1603.09722>.
- [11] Ewain Gwynne, Nina Holden, and Xin Sun. A mating-of-trees approach for graph distances in random planar maps. *Probab. Theory Related Fields*, 177(3-4):1043–1102, 2020. ISSN 0178-8051. doi: 10.1007/s00440-020-00969-8. URL <https://arxiv.org/abs/1711.00723>.
- [12] Nina Holden and Xin Sun. SLE as a mating of trees in Euclidean geometry. *Comm. Math. Phys.*, 364(1):171–201, 2018. ISSN 0010-3616. doi: 10.1007/s00220-018-3149-1. URL <https://arxiv.org/abs/1610.05272>.
- [13] Yiting Li, Xin Sun, and Samuel S. Watson. Schnyder woods, SLE(16), and Liouville quantum gravity. arXiv:1705.03573, May 2017. URL <https://arxiv.org/abs/1705.03573>.
- [14] Brendan D. McKay. Spanning trees in random regular graphs. In *Proceedings of the Third Caribbean Conference on Combinatorics and Computing (Bridgetown, 1981)*, pages 139–143. Univ. West Indies, Cave Hill Campus, Barbados, 1981.
- [15] Jason Miller and Scott Sheffield. Liouville quantum gravity spheres as matings of finite-diameter trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(3):1712–1750, 2019. ISSN 0246-0203. doi: 10.1214/18-aihp932. URL <https://arxiv.org/abs/1506.03804>.
- [16] J. W. Moon. *Counting labelled trees*, volume 1969 of *From lectures delivered to the Twelfth Biennial Seminar of the Canadian Mathematical Congress (Vancouver)*. Canadian Mathematical Congress, Montreal, Que., 1970.
- [17] R. C. Mullin. On the enumeration of tree-rooted maps. *Canadian J. Math.*, 19:174–183, 1967. ISSN 0008-414X. doi: 10.4153/CJM-1967-010-x. URL <https://doi.org/10.4153/CJM-1967-010-x>.
- [18] Jim Pitman. Coalescent random forests. *J. Combin. Theory Ser. A*, 85(2):165–193, 1999. ISSN 0097-3165. doi: 10.1006/jcta.1998.2919. URL <https://www.stat.berkeley.edu/~pitman/457.pdf>.
- [19] Remco van der Hofstad. *Random graphs and complex networks. Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017. ISBN 978-1-107-17287-6. doi: 10.1017/9781316779422. URL <https://www.win.tue.nl/~rhofstad/NotesRGCN.pdf>.
- [20] T. Walsh and A. B. Lehman. Counting rooted maps by genus. II. *J. Combinatorial Theory Ser. B*, 13:122–141, 1972. ISSN 0095-8956. doi: 10.1016/0095-8956(72)90049-4. URL <https://www.sciencedirect.com/science/article/pii/0095895672900494>.

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCGILL UNIVERSITY, MONTRÉAL, CANADA

Email address: [louigi.addario@mcgill.ca](mailto:louigi.addario@mcgill.ca)

Email address: [jordan.barrett@mail.mcgill.ca](mailto:jordan.barrett@mail.mcgill.ca)

URL: <http://problab.ca/louigi/>