THE WIENER INDEX OF SIMPLY GENERATED RANDOM TREES

SVANTE JANSON

Abstract. Asymptotics are obtained for the mean, variance and higher moments as well as for the distribution of the Wiener index of a random tree from a simply generated family (or, equivalently, a critical Galton–Watson tree). We also establish a joint asymptotic distribution of the Wiener index and the internal path length, as well as asymptotics for the covariance and other mixed moments. The limit laws are described using functionals of a Brownian excursion.

The methods include both Aldous' theory of the continuum random tree and analysis of generating functions.

1. Introduction

The *Wiener index* $W(G)$ of a connected graph G is the sum of all the distances between pairs of vertices of G. This index was introduced by the chemist Wiener [27] in the study of relations between the structure of organic compounds and their properties. It has since been studied extensively by both chemists and mathematicians, especially for trees; see the survey [8] for many results and references.

There has been comparatively little work on the Wiener index of random trees. A pioneering paper by Entringer, Meir, Moon and Székely [10] gives asymptotics for the mean $\mathbb{E}\, W(T_n)$ as $n \to \infty$, where T_n is a random tree of order n with the distribution given by a simply generated family of trees. (For some special simply generated families, exact expressions for $\mathbb{E}\, W(T_n)$ were derived too.) Recall that the simply generated families include several important families, for example binary trees, ordered trees and unordered labelled trees. Moreover, the simply generated random trees are the same as the conditioned Galton–Watson trees [2]. See further Section 2.

It was shown in [10] that for simply generated families of trees, $\mathbb{E}\, W(T_n) \sim Kn^{5/2}$ for a constant K depending on the family, see further Theorem 3.4 below. In other words, the average distance between two vertices in an average tree is of the order $n^{1/2}$.

More recently, Neininger [22] took a another pioneering step and obtained asymptotics for both the mean and variance as well as for the distribution of the Wiener index of two other types of random trees, viz. random recursive trees and binary search trees. These random trees are more compact than

Date: May 30, 2002.

the simply generated trees; they have average distances of the order $\log n$ and thus $\mathbb{E}\, W(T_n)$ is of the order $n^2 \log n$. Neininger [22] uses the contraction method and the limit distributions are characterized by a fixed point property from which moments can be pumped.

We are not aware of any paper besides [22] where distributional properties of the Wiener index is studied for some families of random trees. The purpose of the present paper is to prove corresponding results on variance, higher moments and asymptotic distribution for the simply generated families considered by [10].

The convergence in distribution is an almost immediate consequence of Aldous's theory of the continuum random tree [2, 3], see Sections 3 and 4 for details. The asymptotic moments are found in Section 5 by a generating function approach related to the arguments in [10], but expressed using subcritical Galton–Watson trees. Moment convergence follows immediately under some extra conditions, see Remark 3.5; we give a general proof in Section 6 by a related generating function method. (It might be possible to combine the arguments of Sections 5 and 6 and prove both results together, but we think that the result would be more complicated. Moreover, we find it interesting to compare the two related but differently formulated arguments in the two sections.) Finally, some possible extensions are mentioned in Section 7.

2. Simply generated trees

A *simply generated family of trees* is defined by a sequence φ_k, $k \geq 0$, of non-negative numbers (with $\varphi_0 > 0$); each ordered tree T is given a weight $\prod_v \varphi_{d(v)}$, where v ranges over the vertices of T and $d(v)$ is the outdegree (number of children) of v [21]. The corresponding *simply generated random tree* T_n is defined by choosing a tree of order n with probability proportional to its weight (provided that there is some tree of order n with positive weight).

It is well known [2] that the simply generated random trees obtained in this way are (except for some extreme cases usually not considered) the same as the random conditioned Galton–Watson trees, obtained as the family tree of a Galton–Watson process conditioned on a given total size. More precisely, suppose that the generating function $\varphi(z) := \sum_k \varphi_k z^k$ converges for some $z > 0$, and that X is an integer valued random variable with the distribution $\mathbb{P}(X = k) = \varphi_k z^k / \varphi(z)$. Then the conditioned Galton–Watson tree given by the offspring distribution X is easily seen to coincide with the simply generated random tree defined by (φ_k), regardless of the choice of z. (We make a slight abuse of notation here, since X is a random variable and not a distribution, but it is only the distribution of X that matters.) Conversely, given a conditioned Galton–Watson tree with an offspring distribution X, we can take $\varphi_k = \mathbb{P}(X = k)$.

Let $R \le \infty$ be the radius of convergence of $\varphi(z)$. Usually (for example in [10]), it is assumed that there exists a τ with $0 < \tau < R$ and $\tau\varphi'(\tau) = \varphi(\tau)$; taking $z = \tau$ above, this is equivalent to $\mathbb{E}\,X = 1$, so we are dealing with a critical Galton–Watson process. Moreover, then the variance of X, which appears as a scale factor in the results below, is given by $\sigma^2 := \operatorname{Var}(X) = \tau^2\varphi''(\tau)/\varphi(\tau)$. In this case we also have an exponential moment $\mathbb{E}\,e^{\alpha X} < \infty$ for some $\alpha > 0$. We are mainly interested in this case, but we will be somewhat more general and only demand a finite second moment. (This means in the simply generated setting that we allow $\tau = R$, provided $\varphi''(\tau)$ is finite.)

Many combinatorially interesting families are simply generated. Some examples to which our results apply are the following; for further examples see e.g. [2, 7].

(i) Ordered (=plane) trees. $\mathbb{P}(X = k) = 2^{-k-1}$; $\sigma^2 = 2$.
(ii) Unordered labelled trees (Cayley trees). $X \sim \operatorname{Po}(1)$; $\sigma^2 = 1$.
(iii) Binary trees. $X \sim \operatorname{Bi}(2, 1/2)$; $\sigma^2 = 1/2$.
(iv) Strict (full) binary trees. $\mathbb{P}(X = 0) = \mathbb{P}(X = 2) = 1/2$; $\sigma^2 = 1$.
(v) Unary-binary trees. $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \mathbb{P}(X = 2) = 1/3$; $\sigma^2 = 2/3$.
(vi) m-ary trees. $X \sim \operatorname{Bi}(m, 1/m)$; $\sigma^2 = 1 - 1/m$.

3. Results

The Wiener index is defined for unrooted trees; the trees studied in this paper are rooted, but the root is ignored. For rooted trees T we also define the *internal path length* $P(T) = P(T, o) := \sum_{v\in T} d(o, v)$, where o is the root, d is the distance in T and v ranges over the vertices in T. Note that

$$W(T) := \tfrac{1}{2} \sum_{v,w\in T} d(v, w) = \tfrac{1}{2} \sum_{w\in T} P(T, w), \tag{3.1}$$

so $(2/n)W(T)$ is the average of $P(T)$ after a random rerooting.

Returning to a fixed root o, let $v \wedge w$ denote the last common ancestor of two vertices v and w, i.e. the branch point where the paths from o to v and w diverge. Then the path from v to w consists of the parts of these paths beyond $v \wedge w$, and thus

$$d(v, w) = d(v, v\wedge w) + d(v\wedge w, w) = d(v, o) - d(v\wedge w, o) + d(w, o) - d(v\wedge w, o).$$

Consequently, if $n = |T|$,

$$W(T) := \tfrac{1}{2} \sum_{v,w\in T} d(v, w) = n \sum_{v\in T} d(v, o) - \sum_{v,w\in T} d(v \wedge w, o).$$

We define, for any rooted tree T with root o,

$$Q(T) := \sum_{v,w\in T} d(v \wedge w, o)$$

(summing over ordered pairs v, w). Thus, with $n = |T|$,

$$W(T) = nP(T) - Q(T). \tag{3.2}$$

Hence, in order to study the joint distribution of W and P, we may instead study the joint distribution of P and Q, which will be convenient below.

The internal path length P has been studied by many authors. In particular, Aldous [2, 3] has shown that $n^{-3/2}P$ converges in distribution to twice the Brownian excursion area. Our first theorem is a simple extension of this to include the Wiener index (and Q).

Theorem 3.1. *There exists a triple of positive random variables (ξ, η, ζ), with $\zeta = \xi - \eta$, such that if T_n is a conditioned Galton–Watson tree of order n defined by an offspring distribution X with $\mathbb{E}\, X = 1$ and $0 < \sigma^2 := \operatorname{Var} X < \infty$, then, as $n \to \infty$,*

$$\big(n^{-3/2}P(T_n), n^{-5/2}Q(T_n), n^{-5/2}W(T_n)\big) \xrightarrow{\mathrm{d}} (\sigma^{-1}\xi, \sigma^{-1}\eta, \sigma^{-1}\zeta).$$

The random variables (ξ, η, ζ) can be constructed from a normalized Brownian excursion $e(t)$, $0 \le t \le 1$, by

$$\begin{aligned}
\xi &= 2\int_0^1 e(t)\, dt, \\
\eta &= 4 \iint\limits_{0<s<t<1} \min_{s\le u\le t} e(u)\, ds\, dt, \\
\zeta &= \xi - \eta = 2 \iint\limits_{0<s<t<1} \big(e(s) + e(t) - 2\min_{s\le u\le t} e(u)\big)\, ds\, dt.
\end{aligned}$$

A proof following Aldous [2] is given in Section 4.

Remark 3.2. Alternatively, we can use another version of Aldous's limit result, which is intuitively simple but technically more complicated [3, Theorem 23, Remark 1]: There exist measure- and set-representations μ_n and $\mathcal{S}_n$ of rescaled T_n such that $(\mu_n, \mathcal{S}_n) \xrightarrow{\mathrm{d}} (\mu, \mathcal{S})$, where $(\mu, \mathcal{S})$ is the *Brownian continuum random tree* (alias *compact continuum random tree* [2]), see [3] for definitions. Since $\sigma n^{-1/2}P(T_n) = n\int d(0,x)\, d\mu_n(x)$ and $\sigma n^{-1/2}W(T_n) = \frac{1}{2}n^2 \iint d(x,y)\, d\mu_n(x)\, d\mu_n(y)$, it follows immediately that Theorem 3.1 holds with $\xi = \int d(0,x)\, d\mu(x)$ and $\zeta = \frac{1}{2}\iint d(x,y)\, d\mu(x)\, d\mu(y)$.

We may thus interpret ζ as the Wiener index of the continuum random tree, just as ξ is its internal path length.

The moments of the Brownian excursion area $\xi/2$ were found by Louchard [19] using stochastic calculus [18]. We have not been able to find the moments of η and ζ by similar methods, but encourage the adventurous reader to try to do so. Instead, we will in Section 5 use Theorem 3.1 and calculation for binary trees to obtain the following result. We use $\doteq$ for approximate equality.

Theorem 3.3. *The moments of* (ξ,η,ζ) *can be obtained from* $\zeta=\xi-\eta$ *and the formula*

$$\mathbb{E}\,\xi^k\eta^l=\frac{k!\,l!\,\sqrt{\pi}}{2^{(5k+7l-4)/2}\Gamma((3k+5l-1)/2)}\omega^*_{kl} \tag{3.3}$$

where the numbers ω^*_{kl} *are defined by* $\omega^*_{10}=\omega^*_{01}=1$ *and the recursion relation*

$$\omega^*_{k,l}=2(3k+5l-4)\omega^*_{k-1,l}+2(3k+5l-6)(3k+5l-4)\omega^*_{k,l-1}\\ +\sum\sum_{0<i+j<k+l}\omega^*_{i,j}\omega^*_{k-i,l-j}, \tag{3.4}$$

with $\omega^*_{kl}=0$ *when* $k<0$ *or* $l<0$.

In particular, the first and second moments are given by

$$\begin{aligned}
&\mathbb{E}\,\xi\ =\sqrt{\pi/2} &&\mathbb{E}\,\eta\ =\sqrt{\pi/8} &&\mathbb{E}\,\zeta\ =\sqrt{\pi/8}\\
&\mathbb{E}\,\xi^2=\frac{5}{3} &&\mathbb{E}\,\eta^2=\frac{7}{15} &&\mathbb{E}\,\zeta^2=\frac{2}{5}\\
&\mathbb{E}\,\xi\eta=\frac{13}{15} &&\mathbb{E}\,\xi\zeta=\frac{4}{5} &&\mathbb{E}\,\eta\zeta=\frac{2}{5}
\end{aligned}$$

and as a consequence

$$\begin{aligned}
\operatorname{Var}(\xi)&=\frac{10-3\pi}{6}\doteq 0.0959,\\
\operatorname{Var}(\eta)&=\frac{56-15\pi}{120}\doteq 0.0740,\\
\operatorname{Var}(\zeta)&=\frac{16-5\pi}{40}\doteq 0.0073,\\
\operatorname{Cov}(\xi,\zeta)&=\frac{16-5\pi}{20}\doteq 0.0146,\\
\operatorname{Corr}(\xi,\zeta)&=\sqrt{\frac{48-15\pi}{50-15\pi}}\doteq 0.5519.
\end{aligned}$$

This leads to moment asymptotics for P, Q and W. (For the mean, we recover the result by [10].)

Theorem 3.4. *Let* T_n *be as in Theorem 3.1. Then all mixed moments converge in Theorem 3.1, i.e., for any* $k,l,m\ge 0$

$$\mathbb{E}\big(P(T_n)^kQ(T_n)^lW(T_n)^m\big)\sim\sigma^{-(k+l+m)}n^{(3k+5l+5m)/2}\,\mathbb{E}(\xi^k\eta^l\zeta^m). \tag{3.5}$$

In particular,

$$\begin{aligned}
\mathbb{E}\,W(T_n)\ &\sim\sqrt{\frac{\pi}{8\sigma^2}}n^{5/2},\\
\mathbb{E}\,W(T_n)^2&\sim\frac{2}{5\sigma^2}n^5.
\end{aligned}$$

Remark 3.5. In view of Theorem 3.1, (3.5) is equivalent to uniform integrability of $n^{-(3k+5l+5m)/2}P(T_n)^k Q(T_n)^l W(T_n)^m$ for $n \geq 1$ and any fixed k, l, m. Since $Q(T_n) \leq nP(T_n)$ and $W(T_n) \leq nP(T_n)$ by (3.2), this follows if $n^{-3k/2}P(T_n)^k$, $n \geq 1$, is uniformly integrable for every fixed k. Consequently, (3.5) holds for all $k, l, m \geq 0$ if it holds for all $k \geq 0$ and $l = m = 0$. This is further, by a standard fact, equivalent to

$$\sup_n n^{-3k/2}\,\mathbb{E}\,P(T_n)^k < \infty \quad \text{for every } k \geq 1. \tag{3.6}$$

In the case that X has moments of all orders, Takács [24, 25] has proved (3.6) (and moment convergence), which thus implies (3.5). For completeness, we give a proof of (3.6) assuming only $\mathbb{E}\,X^2 < \infty$ in Section 6.

Remark 3.6. Since $P(T_n) \leq nH(T_n)$, where $H := \max_v d(o, v)$ is the height, it is evidently sufficient for (3.6) and thus for (3.5) that

$$\sup_n n^{-k/2}\,\mathbb{E}\,H(T_n)^k < \infty \quad \text{for every } k \geq 1. \tag{3.7}$$

Under the assumption $\sup_k \varphi_k < \infty$, this was proved (in the form of moment convergence) by Flajolet and Odlyzko [11]; see also [12, Theorem 2]. More generally, a proof of (3.7) when X has an exponential moment is given by Drmota and Marckert [9]. It seems to be unknown whether (3.7) holds assuming only a second moment of X.

Remark 3.7. The limit distributions found here are quite different from the ones found by Neininger [22] for random recursive trees and binary search trees. For one thing, the limit distributions in [22] are supported on the whole real line, while the ones found here are supported on a half-line.

Remark 3.8. The fact that $\mathbb{E}\,\zeta = \mathbb{E}\,\eta = \frac{1}{2}\,\mathbb{E}\,\xi$ follows easily by symmetry, as remarked in [10] for the corresponding statement there. Indeed, consider random unordered labelled trees. Since a random rerooting gives a tree with the same distribution, (3.1) yields in this case, for every n,

$$\mathbb{E}\,W(T) = \frac{n}{2}\,\mathbb{E}\,P(T),$$

and the moment convergence in Theorem 3.4 implies $\mathbb{E}\,\zeta = \frac{1}{2}\,\mathbb{E}\,\xi$.

Similarly, $\mathbb{E}\,\zeta^2 = \frac{1}{2}\,\mathbb{E}\,\xi\zeta$ and, more generally, $\mathbb{E}\,\zeta^{k+1} = \frac{1}{2}\,\mathbb{E}\,\xi\zeta^k$ for any $k \geq 0$.

Remark 3.9. It is well known that the moment generating function (Laplace transform) $\mathbb{E}\,e^{t\xi}$ is finite for every t. (For example, this follows easily from (3.8) below and (3.3).) Since $0 \leq \eta \leq \xi$ and $0 \leq \zeta \leq \xi$, also $\mathbb{E}\,e^{t\eta}$ and $\mathbb{E}\,e^{t\zeta}$ are finite for every t. It follows that the distributions of these variables are determined by their moments, and that the joint distribution is determined by the mixed moments in Theorem 3.3.

Remark 3.10. We have to define $\omega^*_{00} = -1/2$ in order for (3.3) to hold for $k = l = 0$ too. Apart from this exceptional case, the numbers ω^*_{kl} are

k, l	0	1	2	3	4
0	$-1/2$	1	49	9800	4412401
1	1	26	4308	1752652	1313146320
2	5	776	300966	217588128	252515984662
3	60	27052	20324608	23856758216	40646627470976
4	1105	1086576	1406019822	2510422982912	6022491449087070
5	27120	49568684	101869846464	263304392184360	860045720189315072

TABLE 1. The constants ω^*_{kl}.

positive integers by (3.4) and induction. These numbers are tabulated in Table 1 for small k and l. One can see that they grow fairly rapidly.

The formula for the moments can be written in several ways. For example, combining (3.3) and (3.4), we obtain a recursion formula for the moments directly, without the need for the constants ω^*_{kl}; this recursion formula is somewhat more complicated, however, and we leave it to the reader. We have chosen the version above because the recursion is fairly simple and the numbers are integers. Moreover, the numbers ω^*_{k0} are the same as ω^*_k in Flajolet, Poblete and Viola [13], where ξ is studied in another context. (They use the normalization $\sqrt{2}\xi$, called the Airy distribution in [13].) In particular, [13] gives recursions equivalent to the case $l = 0$ of (3.4). The Brownian excursion area $\frac{1}{2}\xi$ appears also in several other combinatorial problems, see again [13] and the references cited there. We do not know any similar other results involving η or ζ.

There is a simple connection, discovered by Spencer [23] and Aldous [4], between the moments of ξ and Wright's constants in the enumeration of connected graphs with n vertices and $n + k$ edges [28]. In fact, ρ_k and σ_k in [28] are given by $\rho_{k-1} = \mathbb{E}(\xi/2)^k/k!$ and $\omega^*_{k0} = 2^{2k}\sigma_{k-1}$ (the latter can be obtained by comparing the recursion for σ_k in [28] to (3.4)). The well-known asymptotics for Wright's constants, see [5], [6], [26], [14, §8], which can be expressed as, for $k \to \infty$,

$$\sigma_k \sim \frac{1}{2\pi}\Big(\frac{3}{2}\Big)^{k+1} k!$$

thus yields

$$\omega^*_{k0} \sim \frac{1}{2\pi} 6^k (k-1)!. \tag{3.8}$$

It would be interesting to find similar asymptotics for ω^*_{kl}, $l > 0$, as k or l or both tend to ∞, or asymptotics of $\mathbb{E}\,\zeta^k$. Numerical calculations suggest

$$\omega^*_{0l} \sim C 50^l (l-1)!^2,$$

with $C \doteq 0.01962 \doteq 1/50.9664$, but we have no proof of this, nor an identification of C (if it exists).

4. Convergence in distribution

Proof of Theorem 3.1. Consider a random simply generated tree T_n and its *depth-first walk* (or *search-depth process* [2, 3]) $f : \{0, 1, \dots, 2n\} \to [0, \infty)$ with $f(0) = f(2n) = 0$, and extend f to the interval $[0, 2n]$ by linear interpolation. Rescale f to $V(t) := \sigma n^{-1/2} f(2nt)$. This is a random function in $C[0, 1]$, and Aldous [3, Theorem 23 with Remark 2] has shown that, in $C[0, 1]$ with its usual topology,

$$V(t) \xrightarrow{\mathrm{d}} 2e(t). \tag{4.1}$$

Each vertex $v \in T_n$, except the root o, has an edge to its parent. This edge is traversed twice by the depth-first walk, say on the intervals $I_{1v} = [i_{1v} - 1, i_{1v}]$ and $I_{2v} = [i_{2v}, i_{2v} + 1]$ where $i_{1v} \le i_{2v}$; we have $f(i_{1v}) = f(i_{2v}) = d(v, o)$. Moreover, $f(x) > d(v, o)$ on (i_{1v}, i_{2v}); this is the interval where the depth-first walk visits the descendants of v. For $v = o$ we define $i_{1o} = 1$ and $i_{2o} = 2n - 1$ and have the same results. The intervals $\{I_{jv}\}_{j=0,1,\, v \in I_n}$ form a partition of $[0, 2n]$, and $\lceil f(x) \rceil = d(v, o)$ when $x \in I_{jv}$ (except at some endpoints).

Consider first $P(T_n)$, essentially repeating Aldous [2]. We have

$$\int_0^{2n} \lceil f(x) \rceil \, dx = \sum_{j=0}^{1} \sum_{v \in T_n} d(v, o) = 2P(T_n)$$

so

$$P(T_n) = \tfrac{1}{2} \int_0^{2n} \lceil f(x) \rceil \, dx = \tfrac{1}{2} \int_0^{2n} f(x) \, dx + O(n)$$

and thus

$$\sigma n^{-3/2} P(T_n) = \frac{1}{2n} \int_0^{2n} V\Bigl(\frac{x}{2n}\Bigr) \, dx + O(n^{-1/2}) = \int_0^1 V(t) \, dt + o(1).$$

Consequently, (4.1) yields the desired result

$$\sigma n^{-3/2} P(T_n) \xrightarrow{\mathrm{d}} \int_0^1 2e(t) \, dt. \tag{4.2}$$

Next, we turn to Q. In the remainder of the proof we, for convenience, define $[x, y] := [y, x]$ when $x > y$. In other words, $[x, y]$ always denotes the interval $[\min(x, y), \max(x, y)]$.

If u and v are two vertices in T_n, then for any $j, k \in \{1, 2\}$,

$$d(u \wedge v, o) = \min_{x \in [i_{ju}, i_{kv}]} f(x).$$

This is easily verified by separately considering the three cases (i) $u \wedge v = u$, i.e. v is a descendant of u; (ii) $u \wedge v = v$, i.e. u is a descendant of v; (iii) neither of these holds.

Hence, for any $j,k \in \{1,2\}$,

$$\iint_{I_{ju}\times I_{kv}} \min_{x\in[y,z]} f(x)\,dy\,dz = d(u\wedge v, o) + O(1)$$

and summing over all j,k,u,v we obtain

$$\int_0^{2n}\int_0^{2n} \min_{x\in[y,z]} f(x)\,dy\,dz = 4Q(T_n) + O(n^2)$$

and so

$$\begin{aligned}\sigma n^{-5/2}Q(T_n) &= (2n)^{-2}\int_0^{2n}\int_0^{2n} \min_{x\in[y,z]} V(2nx)\,dy\,dz + O(n^{-1/2})\\ &= \int_0^1\int_0^1 \min_{u\in[s,t]} V(u)\,ds\,dt + O(n^{-1/2})\\ &= 2\iint_{0<s<t<1} \min_{u\in[s,t]} V(u)\,ds\,dt + o(1).\end{aligned}$$

Hence (4.1) implies

$$\sigma n^{-5/2}Q(T_n) \xrightarrow{\mathrm{d}} 4\iint_{0<s<t<1} \min_{u\in[s,t]} e(u)\,ds\,dt.$$

Moreover, this holds jointly with (4.2), and Theorem 3.1 follows, using also (3.2). □

Remark 4.1. In the case $\mathbb{E}\, e^{\alpha X} < \infty$ for some $\alpha > 0$, a simpler proof of (4.2) has been given by Marckert and Mokkadem [20]. Moreover, [20] also shows convergence in this case of the *height process* to a Brownian excursion. More precisely, if $h(i) := d(v_i, o)$ where $v_1,\dots,v_n$ are the vertices of T_n in the order they appear in the depth-first walk, then $\sigma n^{-1/2}h(nt) \xrightarrow{\mathrm{d}} 2e(t)$.

We can argue as above using the height process instead of the depth-first walk; this is slightly simpler since each vertex corresponds to one interval instead of two. However, it seems to be unknown whether the height process converges assuming only a second moment on X, so this argument does not (yet?) prove Theorem 3.1 in full generality.

5. Moments

We calculate the moments of ζ by calculating the asymptotics of the moments of W for some simply generated family, applying Theorem 3.4. We may choose any family for which Theorem 3.4 applies; we find it most convenient to study full binary trees, i.e. trees where all internal nodes have degree 2; these are the conditioned Galton–Watson trees obtained with the offspring distribution $\mathbb{P}(X = 0) = \mathbb{P}(X = 2) = 1/2$. Note that $\mathbb{E}\,X = 1$

and $\sigma^2 = \operatorname{Var} X = 1$, and thus $\bigl(n^{-3/2}P(T_n), n^{-5/2}Q(T_n)\bigr) \xrightarrow{\mathrm{d}} (\xi,\eta)$ by Theorem 3.1. (Similar calculations can be done for e.g. binary trees or unordered trees.)

We obtain asymptotics of moments of $P(T_n)$ and $Q(T_n)$ by studying the corresponding subcritical Galton–Watson tree, defined by the offspring distribution $\mathbb{P}(X=2)=p$ and $\mathbb{P}(X=0)=1-p$ with $0<p<1/2$. We take $p=(1-\delta)/2$, with $0<\delta<1$, and let $\mathbb{P}_\delta$ and $\mathbb{E}_\delta$ denote probability and expectation for this Galton–Watson tree $T=T_\delta$.

Let b_n be the number of full binary trees of order n. Note that the trees all have odd order, and thus $b_n=0$ when n is even. Consequently, we restrict n to odd numbers in this section. (It is well known that b_{2k+1} is the Catalan number $(2k)!/(k!\,(k+1)!)$, see e.g. [15], but we do not need that.)

Let $N:=|T|$. A full binary tree with $n=2k+1$ vertices has k vertices with 2 children and $k+1$ with none. Hence

$$\begin{aligned}\mathbb{P}_\delta(N=2k+1) &= b_{2k+1}p^k(1-p)^{k+1} = 2^{-(2k+1)}b_{2k+1}(1-\delta)^k(1+\delta)^{k+1}\\ &= 2^{-(2k+1)}b_{2k+1}(1-\delta^2)^k(1+\delta).\end{aligned}$$

We let in this section $Z=Z(T)$ denote an arbitrary functional of the trees such that $|Z(T)|\le C|T|^k$ for some C and k (to guarantee that all sums below converge). If $z_n := \mathbb{E}\, Z(T_n)$ we have

$$\mathbb{E}_\delta Z = \sum_n \mathbb{P}_\delta(N=n)z_n = \sum_k 2^{-(2k+1)}b_{2k+1}z_{2k+1}(1-\delta^2)^k(1+\delta), \tag{5.1}$$

which is a generating function of $\{z_n\}$ (or rather of $\{b_n z_n\}$).

Lemma 5.1. *Let $f(\delta)=\mathbb{E}_\delta Z$. Then*

$$\mathbb{E}_\delta(NZ) = -\Bigl(\frac{1}{\delta}-\delta\Bigr)f'(\delta)+\frac{1}{\delta}f(\delta).$$

Proof. Differentiating (5.1) we obtain

$$\begin{aligned}\frac{d}{d\delta}\mathbb{E}_\delta Z &= \sum_k 2^{-(2k+1)}b_{2k+1}z_{2k+1}(-2k\delta)(1-\delta^2)^{k-1}(1+\delta)+\frac{1}{1+\delta}\mathbb{E}_\delta Z\\ &= -\frac{\delta}{1-\delta^2}\mathbb{E}_\delta\bigl(Z(N-1)\bigr)+\frac{1}{1+\delta}\mathbb{E}_\delta Z.\end{aligned}$$

Hence $(1-\delta^2)f'(\delta) = -\delta\,\mathbb{E}_\delta(ZN)+\delta\,\mathbb{E}_\delta Z+(1-\delta)\,\mathbb{E}_\delta Z$, and the result follows. ∎

Since the choice $Z:=1$ obviously gives $\mathbb{E}_\delta Z=1$, Lemma 5.1 yields

$$\mathbb{E}_\delta N = \delta^{-1}. \tag{5.2}$$

Another application of Lemma 5.1 then yields

$$\mathbb{E}_\delta N^2 = \delta^{-3}+\delta^{-2}-\delta^{-1}. \tag{5.3}$$

Repeating, we see that $\mathbb{E}_\delta N^k$ is a polynomial in δ^{-1} for each $k \geq 1$, but we have no need for exact formulas. We turn to asymptotics.

In the remainder of this section, we let $O(\delta^{-m})$ denote an unspecified *polynomial in* δ^{-1} of degree at most m. Lemma 5.1 has the following consequence.

Lemma 5.2. *If* $\mathbb{E}_\delta Z = a\delta^{-m} + O\big(\delta^{-(m-1)}\big)$, *where* $m \geq 1$ *and* $a \in \mathbb{R}$, *then* $\mathbb{E}_\delta(NZ) = ma\delta^{-(m+2)} + O\big(\delta^{-(m+1)}\big)$. □

The behaviour of the function $\mathbb{E}_\delta Z$ is translated into asymptotics of $\mathbb{E}\, Z(T_n)$ as follows.

Lemma 5.3. *If* $\mathbb{E}_\delta Z = a\delta^{-m} + O\big(\delta^{-(m-1)}\big)$, *where* $m \geq 1$ *and* $a \neq 0$, *then*

$$\mathbb{E}\, Z(T_n) \sim 2^{(1-m)/2} \frac{\Gamma(1/2)}{\Gamma(m/2)} a n^{(m+1)/2} \qquad \text{as } n \to \infty.$$

Proof. It suffices to consider the case $\mathbb{E}_\delta Z = \delta^{-m}$. In this case we have by (5.1), with $z_n := \mathbb{E}\, Z(T_n)$,

$$\sum_k 2^{-(2k+1)} b_{2k+1} z_{2k+1} (1-\delta^2)^{k+1} = (1-\delta)\, \mathbb{E}_\delta Z = \delta^{-m} - \delta^{-m-1}$$

and thus, taking $\delta = \sqrt{1-x}$,

$$\begin{aligned} 2^{-(2k+1)} b_{2k+1} z_{2k+1} &= [x^{k+1}]\big((1-x)^{-m/2} - (1-x)^{-(m-1)/2}\big) \\ &= \frac{\Gamma(m/2+k+1)}{\Gamma(m/2)(k+1)!} - \frac{\Gamma((m-1)/2+k+1)}{\Gamma((m-1)/2)(k+1)!} \\ &\sim \frac{1}{\Gamma(m/2)} k^{m/2-1} \qquad \text{as } k \to \infty. \end{aligned} \tag{5.4}$$

By (5.2), the choice $Z = N$ is of this type with $m = 1$; hence (5.4) yields, since then $z_n = n$,

$$2^{-(2k+1)} b_{2k+1}(2k+1) \sim \frac{1}{\Gamma(1/2)} k^{-1/2}. \tag{5.5}$$

(This also follows by well-known asymptotics of the Catalan numbers.) Returning to a general $m \geq 1$, we find by (5.4) and (5.5), when $\mathbb{E}_\delta Z = \delta^{-m}$,

$$\frac{z_{2k+1}}{2k+1} \sim \frac{\Gamma(1/2)}{\Gamma(m/2)} k^{m/2-1/2}.$$

The result follows in this case, and thus in general. □

Consider now the functionals P and Q on $T = T_\delta$. With probability $1-p$, T consists of the root only, and then $P(T) = Q(T) = 0$. With probability p, T consists of the root o and two subtrees T' and T''; these are independent and each has the same distribution as T. In this case, the definitions of P and Q yield (note that $x \wedge y = o$ when $x \in T'$ and $y \in T''$)

$$P(T) = P(T') + N(T') + P(T'') + N(T''), \tag{5.6}$$

$$Q(T) = Q(T') + N(T')^2 + Q(T'') + N(T'')^2. \tag{5.7}$$

Taking expectations we find

$$\begin{aligned}\mathbb{E}_\delta P &= p\,\mathbb{E}_\delta\big(P(T')+N(T')+P(T'')+N(T'')\big)\\ &= \frac{1-\delta}{2}\big(\mathbb{E}_\delta P+\mathbb{E}_\delta N+\mathbb{E}_\delta P+\mathbb{E}_\delta N\big)\\ &= (1-\delta)\,\mathbb{E}_\delta P+(1-\delta)\,\mathbb{E}_\delta N\end{aligned}$$

and similarly

$$\mathbb{E}_\delta Q = (1-\delta)\,\mathbb{E}_\delta Q+(1-\delta)\,\mathbb{E}_\delta N^2.$$

Hence, using (5.2) and (5.3),

$$\mathbb{E}_\delta P = \frac{1-\delta}{\delta}\,\mathbb{E}_\delta N = \delta^{-2}-\delta^{-1}, \tag{5.8}$$

$$\mathbb{E}_\delta Q = \frac{1-\delta}{\delta}\,\mathbb{E}_\delta N^2 = \delta^{-4}-2\delta^{-2}+\delta^{-1}. \tag{5.9}$$

More generally, for any $k\ge 0$, $l\ge 0$ with $k+l\ge 1$, by (5.6), (5.7) and the binomial theorem,

$$\begin{aligned}\mathbb{E}_\delta\big(P(T)^kQ(T)^l\big) &= \frac{1-\delta}{2}\sum_{i=0}^{k}\sum_{j=0}^{l}\binom{k}{i}\binom{l}{j}\,\mathbb{E}_\delta\Big((P(T')+N(T'))^i\\ &\qquad\cdot(P(T'')+N(T''))^{k-i}(Q(T')+N(T')^2)^j(Q(T'')+N(T'')^2)^{l-j}\Big)\\ &= \frac{1-\delta}{2}\sum_{i=0}^{k}\sum_{j=0}^{l}\binom{k}{i}\binom{l}{j}\,\mathbb{E}_\delta\big((P+N)^i(Q+N^2)^j\big)\,\mathbb{E}_\delta\big((P+N)^{k-i}(Q+N^2)^{l-j}\big)\end{aligned} \tag{5.10}$$

This can be used together with Lemma 5.1 to calculate $\mathbb{E}_\delta(P^kQ^l)$ recursively; we are only interested in the following asymptotical formula.

Lemma 5.4. *If $k\ge 0$ and $l\ge 0$ with $k+l\ge 1$, then*

$$\mathbb{E}_\delta(P^kQ^l) = a_{kl}\delta^{-(3k+5l-1)}+O\big(\delta^{-(3k+5l-2)}\big) \tag{5.11}$$

for some positive numbers a_{kl} satisfying $a_{10}=a_{01}=1$ and the recursion relation

$$\begin{aligned}a_{k,l} &= k(3k+5l-4)a_{k-1,l}+l(3k+5l-6)(3k+5l-4)a_{k,l-1}\\ &\qquad+\frac{1}{2}\sum\sum_{0<i+j<k+l}\binom{k}{i}\binom{l}{j}a_{i,j}a_{k-i,l-j}.\end{aligned} \tag{5.12}$$

Proof. Let $\mathcal{P}_0$ denote the set of all polynomials in δ^{-1} without constant term. Note that (5.8) and (5.9) show that (5.11) holds when $k+l=1$, with $\mathbb{E}_\delta(P^kQ^l)\in\mathcal{P}_0$ and $a_{10}=a_{01}=1$. We continue by induction on $k+l$, and assume that $K,L\ge 0$ with $K+L\ge 2$ are such that (5.11) holds and furthermore $\mathbb{E}_\delta(P^kQ^l) \in \mathcal{P}_0$, when $1 \le k+l < K+L$. For such k and l and any $m \ge 0$, Lemma 5.2 implies

$$\begin{aligned}\mathbb{E}_\delta(P^kQ^lN^m) &= a_{kl}\prod_{j=0}^{m-1}(3k+5l+2j-1)\cdot\delta^{-(3k+5l+2m-1)} + O\big(\delta^{-(3k+5l+2m-2)}\big)\\ &= O\big(\delta^{-(3k+5l+2m-1)}\big), \end{aligned}\tag{5.13}$$

and Lemma 5.1 shows further that $\mathbb{E}_\delta(P^kQ^lN^m) \in \mathcal{P}_0$. The same holds for $k=l=0$ and $m \ge 1$ too (with $a_{00} = -1$) by (5.2) and Lemmas 5.2 and 5.1.

Now consider (5.10) with $k=K$ and $l=L$. The two terms with $i=j=0$ or $i=K$, $j=L$ on the right hand side yield together, using the binomial theorem and (5.13),

$$\begin{aligned}(1-\delta)\,\mathbb{E}_\delta\big((P+N)^K(Q+N^2)^L\big) &= (1-\delta)\,\mathbb{E}_\delta(P^KQ^L) + K\,\mathbb{E}_\delta(P^{K-1}Q^LN)\\ &\quad + L\,\mathbb{E}_\delta(P^KQ^{L-1}N^2) + O\big(\delta^{-(3K+5L-3)}\big).\end{aligned}$$

Note that the fact $\mathbb{E}_\delta(P^kQ^lN^m) \in \mathcal{P}_0$ is used to guarantee that the error term is a polynomial in δ^{-1}, even though it contains terms with the factor δ or $1-\delta$.

Similarly, continuing with the right hand side of (5.10) for $k=K$ and $l=L$, a term with $1 \le i+j \le k+l-1$ yields

$$\begin{aligned}&\frac{(1-\delta)}{2}\binom{K}{i}\binom{L}{j}\big(\mathbb{E}_\delta(P^iQ^j) + O\big(\delta^{-(3i+5j-2)}\big)\big)\Big(\mathbb{E}_\delta(P^{K-i}Q^{L-j})\\ &\qquad\qquad + O\big(\delta^{-(3(K-i)+5(L-j)-2)}\big)\Big)\\ &= \frac{1}{2}\binom{K}{i}\binom{L}{j}\,\mathbb{E}_\delta(P^iQ^j)\,\mathbb{E}_\delta(P^{K-i}Q^{L-j}) + O\big(\delta^{-(3K+5L-3)}\big).\end{aligned}$$

Consequently, (5.10) yields, using the induction hypothesis (5.11) and (5.13),

$$\begin{aligned}\delta\,\mathbb{E}_\delta(P^KQ^L) &= K(3K+5L-4)a_{K-1,L}\delta^{-(3K+5L-2)}\\ &\quad + L(3K+5L-6)(3K+5L-4)a_{K,L-1}\delta^{-(3K+5L-2)}\\ &\quad + \sum\sum_{0<i+j<K+L}\frac{1}{2}\binom{K}{i}\binom{L}{j}a_{i,j}a_{K-i,L-j}\delta^{-(3K+5L-2)} + O\big(\delta^{-(3K+5L-3)}\big).\end{aligned}$$

This proves both the induction hypothesis and (5.12). □

Lemmas 5.3 and 5.4 now yield the sought moment asymptotics for full binary trees.

Lemma 5.5. *If $k \ge 0$ and $l \ge 0$ with $k+l \ge 1$, then*

$$\mathbb{E}(P(T_n)^kQ(T_n)^l) \sim \frac{\sqrt{\pi}}{2^{(3k+5l-2)/2}\Gamma((3k+5l-1)/2)}a_{kl}n^{(3k+5l)/2}. \tag{5.14}$$

where the numbers a_{kl} are defined in Lemma 5.4. □

Proof of Theorem 3.3. Note that (5.14) implies that (3.5) holds for the family of full binary trees, cf. Remark 3.5. Recall further that $\sigma^2 = 1$ for this family. Thus Lemma 5.5 yields

$$\mathbb{E}\,\xi^k\eta^l = \frac{\sqrt{\pi}}{2^{(3k+5l-2)/2}\Gamma((3k+5l-1)/2)}a_{kl}, \qquad k+l \geq 1.$$

We define $\omega^*_{kl} = 2^{k+l-1}a_{kl}/k!\,l!$ and obtain (3.3), while the recursion relation (3.4) follows from (5.12).

We have $\omega^*_{10} = \omega^*_{01} = 1$ and find from (3.4) $\omega^*_{20} = 5$, $\omega^*_{11} = 26$, $\omega^*_{02} = 49$, which yields the first and second moments by simple calculations. □

6. Moment convergence

Proof of Theorem 3.4. As said in Remark 3.5, it suffices to show (3.6) in order to obtain (3.5). The values of $\mathbb{E}\,\zeta$ and $\mathbb{E}\,\zeta^2$ obtained in Theorem 3.3 then complete the proof of Theorem 3.4.

Takács [24] has proved (3.6) under the assumption that $\mathbb{E}\,X^m < \infty$ for each m. (In fact, he proved $n^{-3/2}P(T_n) \xrightarrow{\mathrm{d}} \sigma^{-1}\xi$ by the method of moments.) We give here a proof by a variation of his method assuming only a second moment.

Remark 6.1. It should be possible to refine the argument below and obtain convergence of moments of $n^{-3/2}P(T_n)$ to the moments of $\sigma^{-1}\xi$ in this way as in [24], but we only need a bound (which simplifies the argument considerably), since convergence then follows using Theorem 3.1. It might even be possible to combine the argument below and the related one in Section 5 and obtain asymptotics for mixed moments of $P(T_n)$ and $Q(T_n)$ for general simply generated random trees directly.

Let $\operatorname{span}(X) := \operatorname{GCD}\{n : \mathbb{P}(X = n) > 0\}$. For simplicity we assume below that $\operatorname{span}(X) = 1$, and leave the minor modifications when $\operatorname{span}(X) = d > 1$ to the reader; the most important is that the estimates in Lemmas 6.2 and 6.3 hold in the sector $|\arg z| \leq \pi/d$, and that similar estimates hold in other sectors because $F(e^{2\pi i/d}z, w) = e^{2\pi i/d}F(z,w)$.

Let T be the Galton–Watson tree defined by X, and consider the bivariate probability generating function

$$F(z,w) := \mathbb{E}\big(z^{|T|}w^{P(T)}\big) \tag{6.1}$$

and the univariate probability generating function

$$G(z) := \mathbb{E}\,z^{|T|} = F(z,1).$$

Further let $\psi(z) := \mathbb{E}\,z^X$ be the probability generating function of X. (In the first setup of Section 2, $\psi(z) = \varphi(z\tau)/\varphi(\tau)$.)

Let $U = \{z : |z| < 1\}$ and $\overline{U} = \{z : |z| \leq 1\}$ be the open and the closed unit disk. As all probability generating functions, F, G and ψ are continuous on $\overline{U}\times\overline{U}$ or $\overline{U}$, respectively, and analytic in the interior $U \times U$ or U; further, $F(1,1) = G(1) = \psi(1) = 1$.

As is well known, see e.g. [24], and easily seen by conditioning on the number of children of the root,

$$F(z,w) = z\psi\bigl(F(zw,w)\bigr), \qquad z,w \in \overline{U}. \tag{6.2}$$

Lemma 6.2. *If* $\operatorname{span}(X) = 1$*, then, for some* $c_1, c_2 > 0$ *and every* $z \in U$*,*

$$|F(z,1)| = |G(z)| \le 1 - c_1|1-z|^{1/2}, \tag{6.3}$$

$$|1 - z\psi'\bigl(F(z,1)\bigr)| \ge c_2|1-z|^{1/2}, \tag{6.4}$$

and, for every fixed $m \ge 2$*,*

$$\psi^{(m)}\bigl(F(z,1)\bigr) = O\bigl(|1-z|^{-(m-2)/2}\bigr). \tag{6.5}$$

Proof. We begin with (6.3). This is well-known if the radius of convergence of ψ is greater than 1, and implicit in [24] in the present case; for completeness we give a simple proof.

Since $\operatorname{span}(X) = 1$, it is easily seen that $|G(z)| < 1$ for every $z \in \overline{U} \setminus \{1\}$. Hence, for any $\delta > 0$, by compactness, (6.3) holds for $z \in \overline{U}$ with $|1-z| \ge \delta$, provided c_1 is small enough. Hence it suffices to consider z close to 1.

Since $\mathbb{E}\, X^2 < \infty$, ψ is twice continuously differentiable in $\overline{U}$. Further, $G(1) = 1$, $\psi(1) = 1$, $\psi'(1) = \mathbb{E}\, X = 1$ and $\psi''(1) = \mathbb{E}\, X(X-1) = \sigma^2$. Hence, (6.2) and a Taylor expansion yield, as $z \to 1$ with $z \in \overline{U}$,

$$\begin{aligned} G(z) &= z\psi\bigl(G(z)\bigr) \\ &= z\Bigl(1 + \psi'(1)\bigl(G(z)-1\bigr) + \tfrac12\psi''(1)\bigl(G(z)-1\bigr)^2 + o\bigl(G(z)-1\bigr)^2\Bigr) \\ &= zG(z) + \tfrac12\sigma^2\bigl(G(z)-1\bigr)^2 + o\bigl(G(z)-1\bigr)^2 \end{aligned}$$

and thus $(1-z) \sim (1-z)G(z) \sim \frac12\sigma^2\bigl(G(z)-1\bigr)^2$ and

$$1 - G(z) = \Bigl(\frac{\sqrt2}{\sigma} + o(1)\Bigr)\sqrt{1-z}. \tag{6.6}$$

Since $|G(z)| \le 1$ and $\arg(1-z) \in (-\pi/2, \pi/2)$, we have here the principal branch of $\sqrt{1-z}$, and thus $|\arg\sqrt{1-z}| < \pi/4$, and it is easily seen that (6.6) implies (6.3) for z close to 1. Consequently (6.3) holds for all $z \in \overline{U}$ with a suitable $c_1 > 0$.

Another Taylor expansion yields

$$\begin{aligned} z\psi'\bigl(G(z)\bigr) &= \psi'\bigl(G(z)\bigr) + O(1-z) \\ &= \psi'(1) + \psi''(1)\bigl(G(z)-1\bigr) + O(1-z) + o\bigl(G(z)-1\bigr) \end{aligned}$$

and thus, by (6.6),

$$\begin{aligned} 1 - z\psi'\bigl(G(z)\bigr) &= \sigma^2\bigl(1-G(z)\bigr) + O(1-z) + o\bigl(1-G(z)\bigr) \\ &= \bigl(\sqrt2\sigma + o(1)\bigr)\sqrt{1-z}, \end{aligned}$$

which yields (6.4) for z close to 1 (again this is proved in [24] too). The general case follows by another compactness argument, using (6.3) and the fact that $|\psi'(w)| < 1$ when $|w| < 1$.

Finally, since $\psi''(z)$ is bounded and analytic in U, Cauchy's estimate gives

$$|\psi^{(m)}(w)| \le C_m(1-|w|)^{-(m-2)}, \qquad w \in U$$

(with $C_m = (m-2)!\,\sigma^2$), which by (6.3) yields (6.5). $\square$

Lemma 6.3. *For any fixed integers $k, l \ge 0$, the limit $\partial_z^k \partial_w^l F(z,1) := \lim_{U \ni w \to 1} \partial_z^k \partial_w^l F(z,w)$ exists for $z \in U$. If $\operatorname{span}(X) = 1$ and $k + l \ge 1$, then,*

$$\partial_z^k \partial_w^l F(z,1) = O\big(|1-z|^{-(2k+3l-1)/2}\big), \qquad z \in U. \tag{6.7}$$

Proof. The existence of the limit follows because $P(T) \le |T|^2$.

To prove (6.7), we use induction. Let $K, L \ge 0$ and assume that (6.7) holds when $1 \le k+l < K+L$, or when $k+l = K+L$ but $l < L$. For $z, w \in U$ we have by (6.2)

$$\partial_z^K \partial_w^L F(z,w) = \partial_z^K \partial_w^L \big(z\psi(F(zw,w))\big). \tag{6.8}$$

The right hand side can, by induction, be written as a linear combination of terms

$$z^{1-k_0}\psi^{(m)}\big(F(zw,w)\big)\prod_{i=1}^m \partial_z^{k_i}\partial_w^{l_i}\big(F(zw,w)\big) \tag{6.9}$$

where $0 \le m \le K+L$, $0 \le k_0 \le 1$, $k_i \ge 0$, $l_i \ge 0$, $k_i + l_i \ge 1$ for $i = 1, \dots, m$, and $\sum_0^m k_i = K$, $\sum_1^m l_i = L$. Now let $w \to 1$, keeping $z \in U$ fixed. Since, as is verified by induction,

$$\partial_w^l\big(F(zw,w)\big) = \sum_{j=0}^{l} \binom{l}{j} z^j \big(\partial_z^j \partial_w^{l-j} F\big)(zw,w), \tag{6.10}$$

the induction hypothesis easily implies, for $1 \le k+l < K+L$, or $k+l = K+L$ and $l < L$,

$$\partial_z^k \partial_w^l \big(F(zw,w)\big)\Big|_{w=1} = O\big(|1-z|^{-(2k+3l-1)/2}\big). \tag{6.11}$$

Using also (6.5), we see that a term (6.9) with $m \ge 2$ can, for $w = 1$, be estimated by

$$O\Big(|1-z|^{-\left(m-2+\sum_1^m(2k_i+3l_i-1)\right)/2}\Big) = O\big(|1-z|^{-(2K+3L-2)/2}\big).$$

There is no term (6.9) with $m = 0$ except for $K = 1$, $L = 0$, when there is a single term $\psi\big(F(zw,w)\big) = O(1) = O\big(|1-z|^{-(2K+3L-2)/2}\big)$. Finally, for $m = 1$, we have the terms

$$A_1(z,w) := z\psi'\big(F(zw,w)\big)\partial_z^K \partial_w^L\big(F(zw,w)\big), \tag{6.12}$$

which appears exactly once in (6.8), and (when $K \ge 1$)

$$A_2(z,w) := \psi'\big(F(zw,w)\big)\partial_z^{K-1} \partial_w^L\big(F(zw,w)\big). \tag{6.13}$$

Since ψ' is bounded in $\overline{U}$, we obtain from (6.11) (and directly for $K = 1$, $L = 0$)

$$A_2(z,1) = O\big(|1-z|^{-(2K+3L-2)/2}\big).$$

Moreover, expanding (6.12) by (6.10) and Leibniz' rule, collecting all terms coming from $j \geq 1$ in (6.10) into $A_3(z,w)$,

$$A_1(z,w) = z\psi'\bigl(F(zw,w)\bigr)w^k\bigl(\partial_z^K\partial_w^L F\bigr)(zw,w) + A_3(z,w)$$

where, again by (6.11),

$$A_3(z,1) = O\bigl(|1-z|^{-(2K+3L-2)/2}\bigr).$$

Consequently, letting $w \to 1$ in (6.8) yields

$$\partial_z^K\partial_w^L F(z,1) = z\psi'\bigl(F(z,1)\bigr)\partial_z^K\partial_w^L F(z,1) + O\bigl(|1-z|^{-(2K+3L-2)/2}\bigr)$$

or

$$\bigl(1 - z\psi'(F(z,1))\bigr)\partial_z^K\partial_w^L F(z,1) = O\bigl(|1-z|^{-(2K+3L-2)/2}\bigr),$$

which by (6.4) implies $\partial_z^K\partial_w^L F(z,1) = O\bigl(|1-z|^{-(2K+3L-1)/2}\bigr)$, thus completing the induction step and proving (6.7). □

We may now complete the proof of (3.6). Let $l \geq 1$. By (6.1),

$$\begin{aligned}\partial_w^l F(z,1) &= \mathbb{E}\bigl(z^{|T|}(P(T))_l\bigr) = \sum_{n=1}^{\infty}\mathbb{P}(|T|=n)\,\mathbb{E}\bigl(z^{|T|}(P(T))_l \bigm| |T|=n\bigr)\\ &= \sum_{n=1}^{\infty}\mathbb{P}(|T|=n)\,\mathbb{E}\bigl(P(T_n)_l\bigr)z^n.\end{aligned}$$

Cauchy's formula thus yields, for $0 < r < 1$,

$$\mathbb{P}(|T|=n)\,\mathbb{E}\bigl(P(T_n)_l\bigr) = \frac{1}{2\pi i}\int_{|z|=r} z^{-(n+1)}\partial_w^l F(z,1)\,dz.$$

Taking $r = 1 - 1/n$ (for $n \geq 2$), we find by (6.7), for $l \geq 2$ and some $C_1, \dots$ depending on l only,

$$\begin{aligned}\mathbb{P}(|T|=n)\,\mathbb{E}\bigl(P(T_n)_l\bigr) &\leq C_1\int_{-\pi}^{\pi}|\partial_w^l F(re^{i\theta},1)|\,d\theta\\ &\leq C_2\int_{-\pi}^{\pi}|1-re^{i\theta}|^{-(3l-1)/2}\,d\theta \leq C_3\int_{-\pi}^{\pi}\Bigl(\frac{1}{n}+|\theta|\Bigr)^{-(3l-1)/2}d\theta\\ &\leq 2C_3\int_{1/n}^{\infty}x^{-(3l-1)/2}\,dx \leq C_4\,n^{(3l-1)/2-1}.\end{aligned}$$

Moreover, it is well-known, see [16, Lemma 2.1.4], that $\mathbb{P}(|T|=n) \sim (2\pi\sigma^2)^{-1/2}n^{-3/2}$. Hence, for $l \geq 2$,

$$\mathbb{E}\bigl(P(T_n)_l\bigr) = O\bigl(n^{3l/2}\bigr),$$

which is equivalent to $\mathbb{E}\bigl(P(T_n)^l\bigr) = O\bigl(n^{3l/2}\bigr)$, $l \geq 2$. The case $l = 1$ follows from $\bigl(\mathbb{E}\,P(T_n)\bigr)^2 \leq \mathbb{E}\,P(T_n)^2$, and (3.6) is proved. □

7. Further comments

1. A generalization of the Wiener index that has been considered by chemists is

$$W_\lambda(T) := \tfrac{1}{2} \sum_{v,w\in T} d(v,w)^\lambda,$$

where λ is a positive constant, see [8] and the references given there. Both versions above (in Remark 3.2 and Section 4) of the proof of Theorem 3.1 generalize, and we obtain

$$n^{-(2+\lambda/2)} W_\lambda(T_n) \xrightarrow{\mathrm{d}} \sigma^{-\lambda}\zeta_\lambda$$

with

$$\zeta_\lambda = \frac{1}{2}\iint d(x,y)^\lambda \, d\mu(x)\, d\mu(y) = 2^\lambda \iint_{0<s<t<1} \bigl(e(s)+e(t)-2\min_{s\le u\le t} e(u)\bigr)^\lambda \, ds\, dt.$$

The expectation can be computed as follows. Arguing as in Remark 3.8 with random rerootings of random unordered labelled trees, we see that $\mathbb{E}\,\zeta_\lambda = \frac{1}{2}\mathbb{E}\,\xi_\lambda$, with $\xi_\lambda := \int d(0,x)^\lambda \, d\mu(x) = \int_0^1 (2e(t))^\lambda \, dt$. As shown by Lévy [17], the expected occupation density of e is $4xe^{-2x^2}$ (in other words, if U is uniform on $[0,1]$ and independent of e, then $e(U)$ has the Rayleigh distribution with this density), and thus

$$\begin{aligned}\mathbb{E}\,\zeta_\lambda &= \tfrac{1}{2}\mathbb{E}\,\xi_\lambda = 2^{\lambda-1}\,\mathbb{E}\int_0^1 e(t)^\lambda \, dt = 2^{\lambda-1}\int_0^\infty x^\lambda \cdot 4xe^{-2x^2}\, dx \\ &= 2^{\lambda/2-1}\Gamma(\lambda/2+1).\end{aligned}$$

Hence, at least when (3.7) holds (see Remark 3.6), for any $\lambda > 0$,

$$\mathbb{E}\,W_\lambda(T_n) \sim \sigma^{-\lambda} 2^{\lambda/2-1}\Gamma(\lambda/2+1) n^{2+\lambda/2}.$$

It should be possible to derive higher moments of ζ_λ, and thus asymptotics for the moments of $W_\lambda(T_n)$, by the method of Section 5, at least when λ is an integer, but we have not pursued this.

2. Define the weight $w(e)$ of an edge e a the tree T to be the product of the numbers of vertices in the two components of $T - e$. Thus $w(e)$ is the number of all paths in T that pass through e, and summing we find [27]

$$W(T) = \sum_e w(e). \tag{7.1}$$

Let us consider the individual terms in this representation of the Wiener index. Note that when $W(T)$ is written as the sum of all path lengths, as in the definition in Section 1, the individual terms are typically of the same order as the average, i.e. $n^{1/2}$. On the other hand, there are $n-1$ edges and thus the average edge weight in a typical tree is of order $n^{3/2}$, while a typical edge weight only is of order n. In fact, as a reformulation of Aldous [1, Lemma 9], if we take a random conditioned Galton–Watson tree T_n and a random edge $e \in T_n$, then the subtree which consists of the component of $T_n - e$ not containing the root has asymptotically the distribution of the corresponding (unconditioned) Galton–Watson tree T. Consequently, then $w(e)/n \xrightarrow{d} |T|$.

This shows that the main contribution to the sum (7.1) comes from the few extremal edges that split the tree into two large parts. It might be interesting to study the distribution of these large edge weights further, for example by studying $\sum_e w(e)^\lambda$ for constants $\lambda \neq 1$.

3. Takács [25] gives a limit theorem for $P(T_n)$ for *unlabelled* rooted trees too. It seems likely that there are versions of the results in this paper for unlabelled (rooted or unrooted) trees; cf. Aldous [2, Section 4], where also other types of random trees are discussed.

References

[1] D. Aldous, Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1** (1991), no. 2, 228–266.

[2] D. Aldous, The continuum random tree II: an overview. *Stochastic analysis (Proc., Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.

[3] D. Aldous, The continuum random tree III. *Ann. Probab.* **21**, no. 1, 248–289.

[4] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** (1997), no. 2, 812–854.

[5] G.N. Bagaev & E.F. Dmitriev, Enumeration of connected labeled bipartite graphs. (Russian) *Dokl. Akad. Nauk BSSR* **28** (1984), no. 12, 1061–1063, 1148.

[6] E.A. Bender, E.R. Canfield & B.D. McKay, The asymptotic number of labeled connected graphs with a given number of vertices and edges. *Random Structures Algorithms* **1** (1990), no. 2, 127–169.

[7] L. Devroye, Branching processes and their applications in the analysis of tree structures and tree algorithms. *Probabilistic methods for algorithmic discrete mathematics*, 249–314, eds. M. Habib et al., Algorithms Combin. 16, Springer, Berlin, 1998.

[8] A.A. Dobrynin, R. Entringer & I. Gutman, Wiener index of trees: theory and applications. *Acta Appl. Math.* **66** (2001), no. 3, 211–249.

[9] M. Drmota & J.-F. Marckert, Reinforced weak convergence of stochastic processes. Preprint, 2001. Available from `http://www.geometrie.tuwien.ac.at/drmota/`

[10] R.C. Entringer, A. Meir, J.W. Moon & L.A. Székely, The Wiener index of trees from certain families. *Australas. J. Combin.* **10** (1994), 211–224.

[11] P. Flajolet & A. Odlyzko, The average height of binary trees and other simple trees. *J. Comp. Sys. Sci.* **25** (1982), 171–213.

[12] P. Flajolet, Z. Gao, A. Odlyzko, & B. Richmond, The distribution of heights of binary trees and other simple trees. *Combin. Probab. Comput.* **2** (1993), 145-156.

[13] P. Flajolet, P. Poblete & A. Viola, On the analysis of linear probing hashing. *Algorithmica* **22** (1998), no. 4, 490–515.

[14] S. Janson, D.E. Knuth, T. Łuczak & B. Pittel, The birth of the giant component. *Random Structures Algorithms* **4** (1993), no. 3, 233–358.

[15] D.E. Knuth, *The Art of Computer Programming. Vol. 1: Fundamental algorithms.* 3rd ed., Addison–Wesley, Reading, Mass., 1997.

[16] V.F. Kolchin, *Random Mappings.* Nauka, Moscow, 1984 (Russian). English transl.: Optimization Software, New York, 1986.

[17] P. Lévy, *Processus Stochastiques et Mouvement Brownien.* Gauthier-Villars, Paris, 1948.

[18] G. Louchard, Kac's formula, Lévy's local time and Brownian excursion. *J. Appl. Probab.* **21** (1984), no. 3, 479–499.

[19] G. Louchard, The Brownian excursion area: a numerical analysis. *Comput. Math. Appl.* **10** (1984), no. 6, 413–417. Erratum: *Comput. Math. Appl. Part A* **12** (1986), no. 3, 375.

[20] J.-F. Marckert & A. Mokkadem, The depth first processes of Galton–Watson trees converge to the same Brownian excursion. Preprint, 2001. Available from `http://www.math.uvsq.fr/~marckert/`

[21] A. Meir & J.W. Moon, On the altitude of nodes in random trees. *Canad. J. Math.* **30** (1978), 997–1015.

[22] R. Neininger, Wiener index of random trees. Preprint, 2001. Available from `http://www.stochastik.uni-freiburg.de/homepages/neininger/`

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

[24] L. Takács, Conditional limit theorems for branching processes. *J. Appl. Math. Stochastic Anal.* **4** (1991), no. 4, 263–292.

[25] L. Takács, The asymptotic distribution of the total heights of random rooted trees. *Acta Sci. Math. (Szeged)* **57** (1993), no. 1-4, 613–625.

[26] V.A. Voblyĭ, Wright and Stepanov-Wright coefficients. (Russian) *Mat. Zametki* **42** (1987), no. 6, 854–862, 911. English transl.: *Math. Notes* **42** (1987), 969–974.

[27] H. Wiener, Structural determination of paraffin boiling points. *J. Amer. Chem. Soc.* **69** (1947), 17–20.

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

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

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

URL: `http://www.math.uu.se/~svante/`