CONDITIONED GALTON–WATSON TREES: THE SHAPE FUNCTIONAL, AND MORE ON THE SUM OF POWERS OF SUBTREE SIZES AND ITS MEAN

JAMES ALLEN FILL, SVANTE JANSON, AND STEPHAN WAGNER

Abstract. For a complex number α, we consider the sum of the αth powers of subtree sizes in Galton–Watson trees conditioned to be of size n. Limiting distributions of this functional $X_n(\alpha)$ have been determined for $\operatorname{Re}\alpha \neq 0$, revealing a transition between a complex normal limiting distribution for $\operatorname{Re}\alpha < 0$ and a non-normal limiting distribution for $\operatorname{Re}\alpha > 0$. In this paper, we complete the picture by proving a normal limiting distribution, along with moment convergence, in the missing case $\operatorname{Re}\alpha = 0$. The same results are also established in the case of the so-called shape functional $X_n'(0)$, which is the sum of the logarithms of all subtree sizes; these results were obtained earlier in special cases. Additionally, we prove convergence of all moments in the case $\operatorname{Re}\alpha < 0$, where this result was previously missing, and establish new results about the asymptotic mean for real $\alpha < 1/2$.

A novel feature for $\operatorname{Re}\alpha = 0$ is that we find joint convergence for several α to independent limits, in contrast to the cases $\operatorname{Re}\alpha \neq 0$, where the limit is known to be a continuous function of α. Another difference from the case $\operatorname{Re}\alpha \neq 0$ is that there is a logarithmic factor in the asymptotic variance when $\operatorname{Re}\alpha = 0$; this holds also for the shape functional.

The proofs are largely based on singularity analysis of generating functions.

1. Introduction and main results

This paper is a sequel to [5]. As there, we consider a conditioned Galton–Watson tree $\mathcal{T}_n$ and the random variables

$$X_n(\alpha) := F_\alpha(\mathcal{T}_n) := \sum_{v \in \mathcal{T}_n} |\mathcal{T}_{n,v}|^\alpha, \tag{1.1}$$

where $\mathcal{T}_{n,v}$ is the fringe subtree of $\mathcal{T}_n$ rooted at a vertex $v \in \mathcal{T}_n$, i.e., the subtree consisting of v and all its descendants. This is a special case of what is known as

Date: January 23, 2023.

2020 *Mathematics Subject Classification.* Primary: 05C05; Secondary: 60F05, 60C05.

Key words and phrases. Conditioned Galton–Watson tree, simply generated random tree, additive functional, tree recurrence, subtree sizes, shape functional, generating function, singularity analysis, Hadamard product of power series, method of moments, polylogarithm, Laplace transform.

Research of the first author supported by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

Research of the second and third authors supported by the Knut and Alice Wallenberg Foundation.

an additive functional: a functional associated with a rooted tree T that can be expressed in the form

$$F(T) = \sum_{v \in T} f(T_v) \tag{1.2}$$

for a certain toll function f. Thus, F_α is the additive functional on rooted trees defined by the toll function $f_\alpha(T) := |T|^\alpha$. For reasons discussed in [5], we allow the parameter α to be any complex number. (See further Section 2 for the notation used here and below.)

In [5], it is assumed that the conditioned Galton–Watson tree $\mathcal{T}_n$ is defined by some offspring distribution ξ with $\mathbb{E}\,\xi = 1$ and $0 < \sigma^2 := \operatorname{Var}\xi < \infty$. The main results are limit theorems showing that then the random variables $X_n(\alpha)$ converge in distribution after suitable normalization. The results differ for the two cases $\operatorname{Re}\alpha < 0$ and $\operatorname{Re}\alpha > 0$: Typical results are the following (here somewhat simplified), where

$$\widetilde{X}_n(\alpha) := X_n(\alpha) - \mathbb{E}\,X_n(\alpha). \tag{1.3}$$

For further related results, and references to previous work, see [5].

Theorem 1.1 ([5, Theorem 1.1])**.** *If* $\operatorname{Re}\alpha < 0$*, then*

$$n^{-1/2}\widetilde{X}_n(\alpha) \xrightarrow{\mathrm{d}} \widehat{X}(\alpha), \tag{1.4}$$

where $\widehat{X}(\alpha)$ *is a centered complex normal random variable with distribution depending on the offspring distribution* ξ*.*

Theorem 1.2 ([5, Theorem 1.2])**.** *If* $\operatorname{Re}\alpha > 0$*, then*

$$n^{-\alpha-\frac{1}{2}}\widetilde{X}_n(\alpha) \xrightarrow{\mathrm{d}} \sigma^{-1}\widetilde{Y}(\alpha), \tag{1.5}$$

where $\widetilde{Y}(\alpha)$ *is a centered random variable with a (non-normal) distribution that depends on* α *but does not depend on the offspring distribution* ξ*.*

Note the three differences between the two cases:

(i) the normalization is by different powers of n, with the power constant for $\operatorname{Re}\alpha < 0$ but not for $\operatorname{Re}\alpha > 0$;
(ii) the limit is normal for $\operatorname{Re}\alpha < 0$ but not for $\operatorname{Re}\alpha > 0$;
(iii) the limit distribution is universal for $\operatorname{Re}\alpha > 0$ in the sense that it depends on ξ only by the scale factor σ^{-1}, but for $\operatorname{Re}\alpha < 0$, the distribution seems to depend on the offspring distribution ξ in a more complicated way. (In the latter case, the distribution is complex normal, so it is determined by the covariance matrix of $\big(\operatorname{Re}\widehat{X}(\alpha), \operatorname{Im}\widehat{X}(\alpha)\big)$; a complicated formula for covariances is given in [5, Remark 5.1], but we do not know how to evaluate it for concrete examples, not even when $\alpha < 0$ is real and thus $\widehat{X}(\alpha)$ is a real random variable.)

The results above leave a gap: the case $\operatorname{Re}\alpha = 0$, and the main purpose of the present paper is to fill this gap, and to compare the results with the cases above.

The case $\alpha = 0$ is trivial, since $X_n(\alpha) = n$ is non-random. However, in this case we instead study the derivative

$$X'_n(0) = \sum_{v\in\mathcal{T}_n} \log|\mathcal{T}_{n,v}| = \log \prod_{v\in\mathcal{T}_n} |\mathcal{T}_{n,v}|, \tag{1.6}$$

which is known as the *shape functional.* This has earlier been studied in some special cases in e.g. [3; 15; 18; 7; 4; 1], see Section 3.

Another gap in [5] is that moment convergence was proved for $\operatorname{Re}\alpha > 0$ (Theorem 1.2) but not for $\operatorname{Re}\alpha < 0$ (Theorem 1.1). We fill that gap too.

For technical convenience, we assume throughout the paper the weak extra moment condition

$$\mathbb{E}\,\xi^{2+\delta} < \infty, \tag{1.7}$$

for some $\delta > 0$; we also continue to assume $\mathbb{E}\,\xi = 1$. We let $\mathcal{T}$ be an unconditioned Galton–Watson tree with offspring distribution ξ, and define, for complex α with $\operatorname{Re}\alpha < \frac{1}{2}$,

$$\mu(\alpha) := \mathbb{E}\,|\mathcal{T}|^\alpha = \sum_{n=1}^{\infty} n^\alpha\, \mathbb{P}(|\mathcal{T}| = n), \tag{1.8}$$

$$\mu' := \mu'(0) = \mathbb{E}\log|\mathcal{T}| = \sum_{n=1}^{\infty} \mathbb{P}(|\mathcal{T}| = n)\log n. \tag{1.9}$$

(The sum (1.8) converges for $\operatorname{Re}\alpha < \frac{1}{2}$, since $\mathbb{P}(|\mathcal{T}| = n) = O(n^{-3/2})$; see (2.25).)

Our main results are the following. Note that $X'_n(0)$ is a real random variable, while $X_n(\mathrm{i}t)$ and $X_n(\alpha)$ for $\alpha \notin \mathbb{R}$ are non-real except in trivial cases. As said above, special cases of Theorem 1.3 have been proved by Pittel [18], Fill and Kapur [7], and Caracciolo, Erba, and Sportiello [1].

Theorem 1.3. *Assume* (1.7) *with* $\delta > 0$. *Then,*

$$\frac{X'_n(0) - \mu' n}{\sqrt{n\log n}} \overset{\mathrm{d}}{\longrightarrow} N\bigl(0, 4(1-\log 2)\sigma^{-2}\bigr) \tag{1.10}$$

together with convergence of all moments.

Theorem 1.4. *Assume* (1.7) *with* $\delta > 0$. *Then, for any real* $t \neq 0$,

$$\frac{X_n(\mathrm{i}t) - \mu(\mathrm{i}t)n}{\sqrt{n\log n}} \overset{\mathrm{d}}{\longrightarrow} \zeta_{\mathrm{i}t} \tag{1.11}$$

together with convergence of all moments, where $\zeta_{\mathrm{i}t}$ *is a symmetric complex normal variable with variance*

$$\mathbb{E}\,|\zeta_{\mathrm{i}t}|^2 = \frac{1}{\sqrt{\pi}} \operatorname{Re} \frac{\Gamma(\mathrm{i}t - \frac{1}{2})}{\Gamma(\mathrm{i}t)} \sigma^{-2} > 0. \tag{1.12}$$

Theorem 1.5. *Assume* (1.7) *with* $\delta > 0$. *Then, for any complex* α *with* $\operatorname{Re}\alpha < 0$,

$$\frac{X_n(\alpha) - \mu(\alpha)n}{\sqrt{n}} \overset{\mathrm{d}}{\longrightarrow} \widehat{X}(\alpha) \tag{1.13}$$

together with convergence of all moments, where $\widehat{X}(\alpha)$ *is a centered complex normal random variable with positive variance and distribution depending on the offspring distribution* ξ*. Hence,* (1.4) *holds with convergence of all moments.*

Remark 1.6. By "convergence of all moments", we mean in the case of complex variables, Z_n say, convergence of all mixed moments of Z_n and $\overline{Z_n}$, which is equivalent to convergence of all mixed moments of $\operatorname{Re} Z_n$ and $\operatorname{Im} Z_n$. Since we have convergence in distribution, this is by a standard argument using uniform integrability also equivalent to convergence of all absolute moments.

Note that, conversely, by the method of moments applied to $(\operatorname{Re} Z_n, \operatorname{Im} Z_n)$, this implies convergence in distribution of Z_n, provided, as is the case here, the limit distribution is determined by its moments. Thus, our proof of moment convergence provides a new proof of Theorem 1.1, very different from the proof in [5]. □

Remark 1.7. Since the statements include convergence of the first moments (to 0), we may in Theorems 1.3–1.5 replace $\mu' n$, $\mu(\mathrm{i}t)n$, and $\mu(\alpha)n$ by the expectations $\mathbb{E}\, X'_n(0)$, $\mathbb{E}\, X_n(\mathrm{i}t)$, and $\mathbb{E}\, X_n(\alpha)$, respectively; in particular, this gives the last sentence in Theorem 1.5. More precise estimates of the expectations are given in (3.11), (3.20), (4.14), and (5.10). □

Theorems 1.3 and 1.4 combine some of the features found for $\operatorname{Re}\alpha < 0$ and $\operatorname{Re}\alpha > 0$ in (i)–(iii) above. First, the variances in Theorems 1.3 and 1.4 are of order $n \log n$. This might be a surprise since it is not what a naive extrapolation from either $\operatorname{Re}\alpha < 0$ in Theorem 1.1 or $\operatorname{Re}\alpha > 0$ in Theorem 1.2 would yield, where the variances are of order n ($\operatorname{Re}\alpha < 0$) and $n^{1+2\operatorname{Re}\alpha}$ ($\operatorname{Re}\alpha > 0$); however, it is not surprising that a logarithmic factor appears when the two different expressions meet. Secondly, the limits are normal, as heuristically would be expected by "continuity" from the left, see (ii). Thirdly, the limits are universal and depend only on σ as a scale factor, as heuristically would be expected by "continuity" from the right, see (iii).

The proofs in [5] use two different methods, which are combined to yield the full results: (1) methods using complex analysis and the fact that $X_n(\alpha)$ is an analytic function of α, and (2) analysis of moments for a fixed α using singularity analysis of generating functions based on results of Fill, Flajolet, and Kapur [4], also presented in [9, Section VI.10]. In the present paper, we will use only the second method. We follow the proofs in [5] with some variations (see also [6] and [7]). However, some new leading terms will appear in the singular expansions of the generating functions, which will dominate the terms that are leading in [5]; this explains both the logarithmic factors in the variance (and in higher moments) in Theorems 1.3 and 1.4, and the fact that these theorems yield normal limits while Theorem 1.2 does not.

After some preliminaries in Section 2, we first study the shape functional and prove Theorem 1.3 in Section 3; we then study the case of imaginary exponents and prove Theorem 1.4 in Section 4; after that, we consider the case $\operatorname{Re}\alpha < 0$ and prove Theorem 1.5 in Section 5. These three sections use the same method (from [7] and [5]), and are thus quite similar, but some details differ. The differences arise partly because $X'_n(0)$ is real, while $X_n(\mathrm{i}t)$ and (in general) $X_n(\alpha)$ are not; we will also see that the logarithmic factors in the first two cases appear in the moments in somewhat different ways, and that there is a cancellation of some leading terms in our induction for the first and third case, but not for $X_n(\mathrm{i}t)$. For this reason, we give complete arguments for all three cases, and we encourage the reader to compare them and see both similarities and differences.

In Section 6 we show how the centering functions (1.8) and (1.9) can be compared across variation in the offspring distribution when (real) α satisfies $\alpha < \frac{1}{2}$.

Remark 1.8. The results in [5] show also joint convergence for different α in Theorems 1.1 and 1.2, with limits $\widehat{X}(\alpha)$ and $\widetilde{Y}(\alpha)$ that are analytic, and in particular continuous, random functions of the parameter α in the half-planes $\operatorname{Re}\alpha < 0$ and $\operatorname{Re}\alpha > 0$, respectively. This does *not* extend to the imaginary axis $\operatorname{Re}\alpha = 0$; we will see in Theorem 4.2 that $X_n(\alpha)$ for different imaginary α are asymptotically independent (for $\operatorname{Im}\alpha > 0$), and thus it is not possible to have joint convergence to a continuous random function. □

Remark 1.9. Let $\alpha = s + \mathrm{i}t$, where t is real and fixed, and let $s \searrow 0$. (Thus $s > 0$ is real.) It is shown in [5, Appendix D] that if $t \neq 0$, then the limit $\widetilde{Y}(s + \mathrm{i}t)$ diverges (in probability, say) as $s \searrow 0$, and that $s^{1/2}\widetilde{Y}(s + \mathrm{i}t) \overset{\mathrm{d}}{\longrightarrow} \zeta$, where ζ is a symmetric complex normal variable with

$$\mathbb{E}\,|\zeta|^2 = \frac{1}{2\sqrt{\pi}} \operatorname{Re}\frac{\Gamma(\mathrm{i}t - \frac{1}{2})}{\Gamma(\mathrm{i}t)} > 0. \tag{1.14}$$

(However, unfortunately there is a typo in [5, (D.2)], see Appendix D.) Similarly, it is shown in [5, Appendix C] that $s^{-1/2}\widetilde{Y}(s) \overset{\mathrm{d}}{\longrightarrow} N(0, 2(1 - \log 2))$ as $s \searrow 0$; in particular $s^{-1}\widetilde{Y}(s)$ diverges.

These results may be compared to Theorems 1.3–1.4; note that the limits are the same, except that the variances in both cases differ by a factor 1/2 (which of course depends on the chosen normalizations). Both sets of results can be regarded as iterated limits of $\widetilde{X}_n(s + \mathrm{i}t)$, taking $n \to \infty$ and $s \searrow 0$ in different orders. The divergence of $\widetilde{Y}(s + \mathrm{i}t)$ as $s \searrow 0$ (for fixed $t \neq 0$) thus seems to be related to the fact that the asymptotic variance in Theorem 1.3 is of greater order than n, and similarly the divergence as $s \searrow 0$ of $s^{-1}\widetilde{Y}(s)$ (which loosely might be regarded as an approximation of $n^{-1/2}\widetilde{X}_n'(0)$) seems related to Theorem 1.4. However, we do not see why the factors $s^{\pm\frac{1}{2}}$ in these limits should correspond to the factor $(\log n)^{1/2}$ in Theorems 1.3 and 1.4 [or more precisely to the factor $(2\log n)^{1/2}$, to get exactly the same limit distributions]. □

We end with some problems suggested by the results and comments above.

Problem 1.10. Is there a simple explanation of the equality discussed in Remark 1.9 of iterated limits in different orders, but with different normalizations? Is this an instance of some general phenomenon? What happens if $s \searrow 0$ and $n \to \infty$ simultaneously, i.e., for $\widetilde{X}_n(s_n + \mathrm{i}t)$ where $s_n \searrow 0$ at some appropriate rate?

The asymptotic independence of $X(\mathrm{i}t)$ for $t > 0$ mentioned in Remark 1.8 suggests informally that the stochastic process $(\widetilde{X}_n(\mathrm{i}t) : t \geqslant 0)$ asymptotically looks something like white noise. This might be investigated further, for example as follows.

Problem 1.11. Consider the integrated process $\int_0^t \widetilde{X}_n(\mathrm{i}u)\,\mathrm{d}u$. What is the order of its variance? Does this process after normalization converge to a process with paths that are continuous in t?

The moment assumption (1.7) is used repeatedly to control error terms, but it seems convenient rather than necessary.

Problem 1.12. We conjecture that Theorems 1.3–1.5 hold also without the assumption (1.7). Prove (or disprove) this!

2. Notation and preliminaries

2.1. **General notation.** As said above, $\mathcal{T}$ is a Galton–Watson tree defined by an offspring distribution ξ with mean $\mathbb{E}\,\xi = 1$ and finite non-zero variance $\sigma^2 := \operatorname{Var}\xi < \infty$, and we assume (1.7) for some $\delta > 0$. Furthermore, the conditioned Galton–Watson tree $\mathcal{T}_n$ is defined as $\mathcal{T}$ conditioned on $|\mathcal{T}| = n$. We assume for simplicity that ξ has span 1; the general case follows by standard (and minor) modifications.

$\Gamma(z)$ denotes the Gamma function, $\psi(z) := \Gamma'(z)/\Gamma(z)$ is its logarithmic derivative, and $\gamma = -\psi(1)$ is Euler's constant.

A random variable ζ has a *complex normal distribution* if it takes values in $\mathbb{C}$ and $(\operatorname{Re}\zeta, \operatorname{Im}\zeta)$ is a 2-dimensional normal distribution (with arbitrary covariance matrix). In particular, ζ is *symmetric complex normal* if further $\mathbb{E}\,\zeta = 0$ and $(\operatorname{Re}\zeta, \operatorname{Im}\zeta)$ has covariance matrix $\begin{pmatrix} \varsigma^2/2 & 0 \\ 0 & \varsigma^2/2 \end{pmatrix}$ for some $\varsigma^2 = \mathbb{E}\,|\zeta|^2$, which is called the *variance*; equivalently, $\mathbb{E}\,\zeta = 0$, $\mathbb{E}\,\zeta^2 = 0$, and $\mathbb{E}\,|\zeta|^2 = \varsigma^2$. (See e.g. [10, Proposition 1.31].) A symmetric complex normal distribution with variance ς^2 is determined by the mixed moments of ζ and $\overline{\zeta}$, which are given by (see [10, p. 14])

$$\mathbb{E}\big[\zeta^\ell\,\overline{\zeta}^r\big] = \begin{cases} \varsigma^{2\ell}\ell!, & \ell = r, \\ 0, & \ell \neq r. \end{cases} \tag{2.1}$$

Unspecified limits are as $n \to \infty$. We let $\overset{\mathrm{d}}{\longrightarrow}$ denote convergence in distribution.

For real x and y, we denote $\min(x, y)$ by $x \wedge y$.

The semifactorial $\ell!!$ is defined for odd integers ℓ (the only case that we use) by

$$\ell!! := 1 \cdot 3 \cdot \cdots \cdot \ell = 2^{(\ell+1)/2}\Gamma\left(\tfrac{\ell}{2} + 1\right)/\sqrt{\pi}. \tag{2.2}$$

Note that $(-1)!! = 1!! = 1$.

ε denotes an arbitrarily small fixed number with $\varepsilon > 0$. (We will tacitly assume that ε is sufficiently small when necessary.)

C and c denote unimportant positive constants, possibly different each time; these may depend on the parameter α (or α_1, α_2 below). We sometimes use c with subscripts; these keep the same value within the same section.

2.2. **Δ-domains and singularity analysis.** A Δ-domain is a complex domain of the type

$$\{z : |z| < R,\, z \neq 1,\, |\arg(z-1)| > \theta\} \tag{2.3}$$

where $R > 1$ and $0 < \theta < \pi/2$, see [9, Section VI.3]. A function is Δ-*analytic* if it is analytic in some Δ-domain (or can be analytically continued to such a domain).

Our proofs are based on singularity analysis of various generating functions (see [9, Chapter VI]), using estimates as $z \to 1$ in a suitable Δ-domain; the domain may be different each time. All estimates below of analytic functions tacitly are valid in some Δ-domains (possibly different ones for different functions), even when that is not said explicitly.

2.3. **Polylogarithms.** $\mathrm{Li}_\alpha(z)$ and $\mathrm{Li}_{\alpha,r}(z)$ denote polylogarithms and generalized polylogarithms, respectively; they are defined for $\alpha \in \mathbb{C}$ and $r = 0, 1, \dots$ by the power series

$$\mathrm{Li}_\alpha(z) := \sum_{n=1}^{\infty} n^{-\alpha} z^n, \tag{2.4}$$

$$\mathrm{Li}_{\alpha,r}(z) := \sum_{n=1}^{\infty} (\log n)^r \frac{z^n}{n^\alpha} \tag{2.5}$$

for $|z| < 1$, and then extended analytically to $\mathbb{C} \setminus [0, \infty)$ (in particular they are Δ-analytic); see e.g. [9, Section VI.8]. Note that $\mathrm{Li}_{\alpha,0}(z) = \mathrm{Li}_\alpha(z)$. We will also use the notation

$$L(z) := -\log(1-z) = \sum_{n=1}^{\infty} \frac{z^n}{n} = \mathrm{Li}_1(z). \tag{2.6}$$

We will use singular expansions of polylogarithms and generalized polylogarithms into powers of $1 - z$, possibly including powers of $L(z)$. Infinite singular expansions of polylogarithms and generalized polylogarithms are given by Flajolet [8, Theorem 1] (also [9, Theorem VI.7]); we will mainly use only the following simple versions keeping only the main terms.

For any real a, let $\mathcal{P}_a$ be the set of all polynomials in z of degree $< a$. In particular, if $a \leqslant 0$, then $\mathcal{P}_a = \{0\}$. If $0 \leqslant a \leqslant 1$, then every polynomial in $\mathcal{P}_a$ is constant. These simple cases are the ones of most interest to us.

We then have, for each $\alpha \notin \{1, 2, \dots\}$,

$$\mathrm{Li}_\alpha(z) = \Gamma(1-\alpha)(1-z)^{\alpha-1} + P(z) + O\bigl(|1-z|^{\operatorname{Re}\alpha}\bigr), \tag{2.7}$$

for some $P(z) \in \mathcal{P}_{\operatorname{Re}\alpha}$.

Moreover, in our proofs we will often go back and forth between expansions in powers of $1 - z$ (including powers of $L(z)$) and expansions in (generalized) polylogarithms, using the following simple consequence of the singular expansions of generalized polylogarithms, proved in [7]. (Here slightly simplified.)

Lemma 2.1 ([7, Lemmas 2.5–2.6])**.** *Suppose that* $\operatorname{Re}\alpha < 1$*. Then, for each* $r \geqslant 0$*, in any fixed* Δ*-domain and for any* $\varepsilon > 0$*,*

$$\operatorname{Li}_{\alpha,r}(z) = \sum_{j=0}^{r} \rho_{r,j}(\alpha)(1-z)^{\alpha-1}L(z)^j + c_r(\alpha) + O\bigl(|1-z|^{\operatorname{Re}\alpha-\varepsilon}\bigr), \tag{2.8}$$

for some coefficients $\rho_{r,j}(\alpha)$ *and* $c_r(\alpha)$*, with leading coefficient*

$$\rho_{r,r}(\alpha) = \Gamma(1-\alpha). \tag{2.9}$$

Conversely,

$$(1-z)^{\alpha-1}L(z)^r = \sum_{j=0}^{r} \hat\rho_{r,j}(\alpha)\operatorname{Li}_{\alpha,j}(z) + \hat c_r(\alpha) + O\bigl(|1-z|^{\operatorname{Re}\alpha-\varepsilon}\bigr), \tag{2.10}$$

for some coefficients $\hat\rho_{r,j}(\alpha)$ *and* $\hat c_r(\alpha)$*, with*

$$\hat\rho_{r,r}(\alpha) = \rho_{r,r}(\alpha)^{-1} = \Gamma(1-\alpha)^{-1}. \tag{2.11}$$

Remark 2.2. The lemmas in [7] are stated for real α, but the proofs hold also for complex α. Moreover, the results extend to α with $\operatorname{Re}\alpha \geqslant 1$, assuming $\alpha \notin \{1,2,\dots\}$, provided the error terms $O\bigl(|1-z|^{\operatorname{Re}\alpha-\varepsilon}\bigr)$ are replaced by $O(|1-z|)$ when $\operatorname{Re}\alpha > 1$. ☐

2.4. **Hadamard products.** Recall that the *Hadamard product* $A(z)\odot B(z)$ of two power series $A(z)=\sum_{n=0}^\infty a_n z^n$ and $B(z)=\sum_{n=0}^\infty b_n z^n$ is defined by

$$A(z)\odot B(z) := \sum_{n=0}^{\infty} a_n b_n z^n. \tag{2.12}$$

As a simple example, for any complex α and β,

$$\operatorname{Li}_\alpha(z)\odot\operatorname{Li}_\beta(z) = \operatorname{Li}_{\alpha+\beta}(z), \tag{2.13}$$

and, more generally, by (2.5),

$$\operatorname{Li}_{\alpha,r}(z)\odot\operatorname{Li}_{\beta,s}(z) = \operatorname{Li}_{\alpha+\beta,r+s}(z). \tag{2.14}$$

We note also, for any constant c and power series $A(z)=\sum_{n=0}^\infty a_n z^n$, the trivial result

$$c\odot A(z) = ca_0. \tag{2.15}$$

For our error terms, we will use the following lemma; it is part of [5, Lemma 12.2] and taken from [4, Propositions 9 and 10(i)] and [9, Theorem VI.11 p. 423]. (Further related results are given in [4], [9, Section VI.10.2] and [5].)

Lemma 2.3 ([4; 9])**.** *If* $g(z)$ *and* $h(z)$ *are* Δ*-analytic, then* $g(z)\odot h(z)$ *is* Δ*-analytic. Moreover, suppose that* $g(z)=O(|1-z|^a)$ *and* $h(z)=O(|1-z|^b)$*, where* a *and* b *are real with* $a+b+1\notin\{0,1,2,\dots\}$*; then, as* $z\to1$ *in a suitable* Δ*-domain,*

$$g(z)\odot h(z) = P(z) + O\bigl(|1-z|^{a+b+1}\bigr), \tag{2.16}$$

for some $P(z)\in\mathcal{P}_{a+b+1}$*.*

2.5. **Generating functions for Galton–Watson trees.** Let $p_k := \mathbb{P}(\xi = k)$ denote the values of the probability mass function for the offspring distribution ξ, and let Φ be its probability generating function:

$$\Phi(z) := \mathbb{E}\, z^{\xi} = \sum_{k=0}^{\infty} p_k z^k. \tag{2.17}$$

Similarly, let $q_n := \mathbb{P}(|\mathcal{T}| = n)$, and let y denote the corresponding probability generating function:

$$y(z) := \mathbb{E}\, z^{|\mathcal{T}|} = \sum_{n=1}^{\infty} \mathbb{P}\big(|\mathcal{T}| = n\big) z^n = \sum_{n=1}^{\infty} q_n z^n. \tag{2.18}$$

As is well known, then

$$y(z) = z\Phi\big(y(z)\big). \tag{2.19}$$

Under our assumptions $\mathbb{E}\,\xi = 1$ and $0 < \operatorname{Var}\xi < \infty$, the generating function $y(z)$ extends analytically to a Δ-domain and is thus Δ-analytic; see [11, Lemma A.2] and [5, §12.1] (and under stronger assumptions [9, Theorem VI.6, p. 404]). Furthermore, see again [11, Lemma A.2], there exists a Δ-domain where $|y(z)| < 1$, and thus $\Phi(y(z))$ is Δ-analytic, as well as $\Phi^{(m)}\big(y(z)\big)$ for every $m \geqslant 1$.

We note some useful consequence of our extra moment assumption (1.7); we may without loss of generality assume $\delta \leqslant 1$. (Compare [5, (12.5), (12.30), and (12.31)] without the assumption (1.7) but with weaker error terms, and [9, Theorem VI.6] with stronger results under stronger assumptions.)

Lemma 2.4. *If* (1.7) *holds with* $0 < \delta \leqslant 1$, *then, for* z *in some* Δ*-domain,*

$$y(z) = 1 - \sqrt{2}\sigma^{-1}(1-z)^{1/2} + O\big(|1-z|^{\frac{1}{2}+\frac{\delta}{2}}\big), \tag{2.20}$$

$$y'(z) = 2^{-1/2}\sigma^{-1}(1-z)^{-1/2} + O\big(|1-z|^{-\frac{1}{2}+\frac{\delta}{2}}\big). \tag{2.21}$$

$$\frac{zy'(z)}{y(z)} = 2^{-1/2}\sigma^{-1}(1-z)^{-1/2} + O\big(|1-z|^{-\frac{1}{2}+\frac{\delta}{2}}\big), \tag{2.22}$$

In particular, all three functions are Δ*-analytic.*

Proof. That $y(z)$ is Δ-analytic was noted above, and the estimate (2.20) was shown in [5, Lemma 12.15]. A differentiation then yields (2.21) in a smaller Δ-domain, using Cauchy's estimates for a disc with radius $c|1-z|$ centered at z (see [9, Theorem VI.8 p. 419]).

Note that $zy'(z)/y(z)$ is analytic in any domain where y is defined and analytic with $|y(z)| < 1$, since then (2.19) holds in the domain and implies that $y(z) \neq 0$ for $z \neq 0$, and also that $z/y(z)$ is analytic at $z = 0$. Hence, also $zy'(z)/y(z)$ is Δ-analytic. Finally, (2.22) follows from (2.20) and (2.21). ∎

By (2.7), and using $\Gamma(-1/2) = -2\sqrt{\pi}$, we can rewrite (2.20) as

$$y(z) = -\frac{\sqrt{2}}{\Gamma(-\frac{1}{2})\sigma} \operatorname{Li}_{3/2}(z) + c + O\big(|1-z|^{\frac{1}{2}+\frac{\delta}{2}}\big)$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{Li}_{3/2}(z) + c + O\big(|1-z|^{\frac12+\frac{\delta}{2}}\big), \tag{2.23}$$

where (although we do not need it) $c = 1 - \zeta(3/2)/\sqrt{2\pi\sigma^2}$. Furthermore, by (2.4) and singularity analysis [9, Theorem VI.3, p. 390], (2.23) implies

$$q_n = \mathbb{P}(|\mathcal{T}| = n) = \frac{1}{\sqrt{2\pi}\sigma} n^{-3/2} + O\big(n^{-\frac32-\frac{\delta}{2}}\big) = \frac{1+O\big(n^{-\delta/2}\big)}{\sqrt{2\pi}\sigma} n^{-3/2}. \tag{2.24}$$

Remark 2.5. It is well known that the asymptotic formula

$$q_n = \mathbb{P}(\mathcal{T}| = n) = \frac{1+o(1)}{\sqrt{2\pi}\sigma} n^{-3/2} \qquad \text{as } n\to\infty, \tag{2.25}$$

with a weaker error bound than (2.24), holds assuming only $\operatorname{Var}\xi < \infty$ (and $\mathbb{E}\,\xi = 1$); see e.g. [17] (assuming an exponential moment), [14, Lemma 2.1.4] or [12, Theorem 18.11] (with $\tau = \Phi(\tau) = 1$) and the further references given there. □

Lemma 2.6. *Assume* (1.7) *with* $0 < \delta \leqslant 1$. *Then, for* z *in some* Δ*-domain,*

$$\Phi\big(y(z)\big) = 1 + O\big(|1-z|^{\frac12}\big), \tag{2.26}$$

$$\Phi'\big(y(z)\big) = 1 + O\big(|1-z|^{\frac12}\big), \tag{2.27}$$

$$\Phi''\big(y(z)\big) = \sigma^2 + O\big(|1-z|^{\frac{\delta}{2}}\big), \tag{2.28}$$

and, for each fixed $m \geqslant 3$,

$$\Phi^{(m)}\big(y(z)\big) = O\big(|1-z|^{\frac{\delta}{2}+1-\frac{m}{2}}\big). \tag{2.29}$$

Proof. The assumption (1.7) implies the estimate, see e.g. [5, Lemma 12.14],

$$\Phi(z) = z + \tfrac12\sigma^2(1-z)^2 + O\big(|1-z|^{2+\delta}\big), \qquad |z| \leqslant 1. \tag{2.30}$$

By differentiation of (2.30), for the remainder term using Cauchy's estimates for a disc with radius $(1-|z|)/2$ centered at z, we obtain for all z with $|z| < 1$, and each fixed $m \geqslant 3$,

$$\Phi'(z) = 1 - \sigma^2(1-z) + O\big(|1-z|^{2+\delta}/(1-|z|)\big), \tag{2.31}$$

$$\Phi''(z) = \sigma^2 + O\big(|1-z|^{2+\delta}/(1-|z|)^2\big), \tag{2.32}$$

$$\Phi^{(m)}(z) = O\big(|1-z|^{2+\delta}/(1-|z|)^m\big). \tag{2.33}$$

For z in a suitable Δ-domain we have (2.20), and as a consequence, if $|1-z|$ is small enough,

$$c|1-z|^{1/2} \leqslant 1 - |y(z)| \leqslant |1-y(z)| \leqslant C|1-z|^{1/2}. \tag{2.34}$$

The result follows by (2.30)–(2.34). □

Remark 2.7. In fact, (2.27) holds without the extra assumption (1.7), assuming only $\mathbb{E}\,\xi^2 < \infty$, because then Φ is twice continuously differentiable in the closed unit disc with $\Phi'(1) = 1$, and $y(z) = 1 - \sqrt{2}\sigma^{-1}(1-z)^{1/2} + o(|1-z|^{1/2})$ as is shown in [11, Lemma A.2], see also [5, (12.5)]. □

3. The shape functional

We consider here the shape functional $X'_n(0)$. Asymptotics for the mean and variance were found by Fill [3] in the case of uniform binary trees [the case $\xi \sim \mathrm{Bi}(2, \frac{1}{2})$]; this was generalized by Meir and Moon [15] to simply generated trees under a condition equivalent to our conditioned Galton–Watson trees with ξ having a finite exponential moment $\mathbb{E}\, e^{r\xi} < \infty$ for some $r > 0$. Pittel [18] showed asymptotic normality in the case of uniform labelled trees [the case $\xi \sim \mathrm{Po}(1)$] by estimating cumulants. Fill and Kapur [7] considered uniform binary trees [$\xi \sim \mathrm{Bi}(2, \frac{1}{2})$] and showed asymptotic normality by estimating moments by singularity analysis, see also Fill, Flajolet, and Kapur [4]. Asymptotic normality has recently been shown, by similar methods, also for uniformly random ordered trees [the case $\xi \sim \mathrm{Ge}(\frac{1}{2})$] by Caracciolo, Erba, and Sportiello [1], who further [personal communication] have extended the results to arbitrary offspring distributions ξ (with $\mathbb{E}\,\xi = 1$ as here), at least provided that ξ has a finite exponential moment $\mathbb{E}\, e^{r\xi} < \infty$ for some $r > 0$.

We will here extend these results to any offspring distribution ξ satisfying the standard condition $\mathbb{E}\,\xi = 1$ and the weak moment condition (1.7) for some $\delta > 0$. We assume without loss of generality that $0 < \delta \leqslant 1$. We will use singularity analysis to estimate moments, in the same way as [7; 4; 1].

In this section, we define (corresponding to [5, (12.46)])

$$b_n := \log n - \mu', \qquad n \geqslant 1, \tag{3.1}$$

where $\mu' = \mathbb{E} \log |\mathcal{T}| = \sum_{n=1}^{\infty} q_n \log n$ as in (1.9), and we let F be the additive functional defined by the toll function $f(T) := b_{|T|}$. Thus, by (1.6),

$$F(\mathcal{T}_n) = X'_n(0) - \mu' n. \tag{3.2}$$

The generating function of b_n is, by (2.4)–(2.5) and noting $\mathrm{Li}_0(z) = z/(1-z)$,

$$B(z) = \sum_{n=1}^{\infty} (\log n - \mu') z^n = \mathrm{Li}_{0,1}(z) - \mu' \,\mathrm{Li}_0(z). \tag{3.3}$$

Hence, by Lemma 2.1 (or [9, Figure VI.11, p. 410] with more terms),

$$B(z) = (1-z)^{-1} L(z) - c(1-z)^{-1} + O\bigl(|1-z|^{-\varepsilon}\bigr) \tag{3.4}$$

$$= O\bigl(|1-z|^{-1-\varepsilon}\bigr). \tag{3.5}$$

We define the generating functions, for $\ell \geqslant 1$,

$$M_\ell(z) := \mathbb{E}\bigl[F(\mathcal{T})^\ell z^{|\mathcal{T}|}\bigr] = \sum_{n=1}^{\infty} q_n\, \mathbb{E}[F(\mathcal{T}_n)^\ell] z^n. \tag{3.6}$$

These generating functions can be calculated recursively by the following formula (valid for any sequence b_n) from [5].

Lemma 3.1 ([5, Lemma 12.4])**.** *For every $\ell \geqslant 1$,*

$$M_\ell(z) = \frac{zy'(z)}{y(z)} \sum_{m=0}^{\ell} \frac{1}{m!} \sum\nolimits^{**} \binom{\ell}{\ell_0, \dots, \ell_m} B(z)^{\odot \ell_0} \odot \bigl[z M_{\ell_1}(z) \cdots M_{\ell_m}(z) \Phi^{(m)}\bigl(y(z)\bigr)\bigr], \tag{3.7}$$

where $\sum^{**}$ *is the sum over all* $(m+1)$*-tuples* $(\ell_0,\dots,\ell_m)$ *of non-negative integers summing to* ℓ *such that* $1 \leqslant \ell_1,\dots,\ell_m < \ell$.

Note that $B(z)$ is Δ-analytic by (3.3); furthermore, $zy'(z)/y(z)$ and $\Phi^{(m)}\bigl(y(z)\bigr)$ are also Δ-analytic, see Section 2.5. Hence, (3.7) and induction using Lemma 2.3 show that every $M_\ell(z)$ is Δ-analytic.

It will be convenient to denote the sum in (3.7) by $R_\ell(z)$. Thus,

$$M_\ell(z) = \frac{zy'(z)}{y(z)} R_\ell(z). \tag{3.8}$$

3.1. **The mean.** We begin with the mean $\mathbb{E}\, X'_n(0)$ and the corresponding generating function $M_1(z)$. The following result includes earlier results for special cases in [3; 15; 18; 7; 4; 1], but our error term is weaker [since we have the weaker moment assumption (1.7)]. Recall that $\psi(z) := \Gamma'(z)/\Gamma(z)$, and note that

$$\psi(-\tfrac12) = \psi(\tfrac12) + 2 = -\gamma - 2\log 2 + 2, \tag{3.9}$$

see [16, 5.5.2 and 5.4.13].

Lemma 3.2. *Assume* (1.7) *with* $0 < \delta \leqslant 1$. *Then, for any* $\varepsilon > 0$,

$$M_1(z) = -\sigma^{-2} L(z) + \frac{\mu' - \psi(-\frac12)}{\sigma^2} + O\bigl(|1-z|^{\frac{\delta}{2}-\varepsilon}\bigr) \tag{3.10}$$

and

$$\mathbb{E}\, X'_n(0) = \mu' n - \frac{\sqrt{2\pi}}{\sigma} n^{1/2} + O\bigl(n^{\frac12 - \frac{\delta}{2} + \varepsilon}\bigr). \tag{3.11}$$

Proof. For $\ell = 1$, the sums in (3.7) reduce to a single term with $m = 0$ and $\ell_0 = 1$, and thus, as in [5, (12.29)], using (2.19),

$$M_1(z) = \frac{zy'(z)}{y(z)} \cdot \bigl(B(z) \odot z\Phi\bigl(y(z)\bigr)\bigr) = \frac{zy'(z)}{y(z)} \cdot \bigl(B(z) \odot y(z)\bigr). \tag{3.12}$$

By (2.14), (2.15), (3.3) and (2.23) we obtain, using Lemma 2.3 and (3.5) for the error term,

$$B(z) \odot y(z) = \frac{1}{\sqrt{2\pi}\sigma} \operatorname{Li}_{3/2,1}(z) - \frac{\mu'}{\sqrt{2\pi}\sigma} \operatorname{Li}_{3/2}(z) + c_1 + O\bigl(|1-z|^{\frac12 + \frac{\delta}{2} - \varepsilon}\bigr). \tag{3.13}$$

Further, by our choice (1.9) of μ',

$$(B \odot y)(1) = \sum_{n=1}^\infty b_n q_n = \sum_{n=1}^\infty q_n (\log n - \mu') = \mu' - \mu' = 0. \tag{3.14}$$

By (2.7), we have

$$\operatorname{Li}_{3/2}(z) = \Gamma(-\tfrac12)(1-z)^{1/2} + c_2 + O\bigl(|1-z|\bigr). \tag{3.15}$$

Moreover, by [9, Theorem VI.7, p. 408] (or [8, Theorem 1]),

$$\operatorname{Li}_{3/2,1}(z) = \Gamma(-\tfrac12)(1-z)^{1/2} L(z) + \Gamma'(-\tfrac12)(1-z)^{1/2} + c_3 + O\bigl(|1-z|\bigr). \tag{3.16}$$

Hence, (3.13) and (3.15)–(3.16) yield, using (3.14) to see that the constant terms cancel,

$$\begin{aligned} B(z)\odot y(z) &= \frac{\Gamma(-\frac12)}{\sqrt{2\pi}\sigma}(1-z)^{1/2}L(z)+\frac{\Gamma'(-\frac12)-\mu'\Gamma(-\frac12)}{\sqrt{2\pi}\sigma}(1-z)^{1/2}+O\bigl(|1-z|^{\frac12+\frac{\delta}{2}-\varepsilon}\bigr)\\ &= -\frac{\sqrt2}{\sigma}(1-z)^{1/2}L(z)+\frac{\sqrt2\bigl(\mu'-\psi(-\frac12)\bigr)}{\sigma}(1-z)^{1/2}+O\bigl(|1-z|^{\frac12+\frac{\delta}{2}-\varepsilon}\bigr). \end{aligned} \tag{3.17}$$

Finally, (3.12), (2.22), and (3.17) yield (3.10).

Since $L(z)=\sum_{n=1}^\infty z^n/n$, (3.10) yields by standard singularity analysis, recalling the definition (3.6),

$$q_n\,\mathbb{E}\,F(\mathcal{T}_n)=-\sigma^{-2}n^{-1}+O\bigl(n^{-1-\frac{\delta}{2}+\varepsilon}\bigr). \tag{3.18}$$

Hence, using also (2.24),

$$\mathbb{E}\,F(\mathcal{T}_n)=-\frac{\sqrt{2\pi}}{\sigma}n^{1/2}+O\bigl(n^{\frac12-\frac{\delta}{2}+\varepsilon}\bigr) \tag{3.19}$$

and (3.11) follows by (3.2). ☐

Remark 3.3. Under stronger moment conditions on the offspring distribution ξ, we may in the same way obtain an expansion of the mean $\mathbb{E}\,X_n'(0)$ with further terms. For example, if $\mathbb{E}\,\xi^{3+\delta}<\infty$, then the same argument yields

$$\mathbb{E}\,X_n'(0)=\mu'n-\frac{\sqrt{2\pi}}{\sigma}n^{1/2}+\frac{\mathbb{E}[\xi(\xi-1)(\xi-2)]}{3\sigma^4}\log n+O(1). \tag{3.20}$$

In the special case of binary trees, this was given in [7, (4.2)]. Note that the coefficient of $\log n$ in (3.20) vanishes for binary trees, but not in general. ☐

3.2. **The second moment.**

Lemma 3.4. *Assume* (1.7) *with* $0<\delta\leqslant 1$. *Then, for any* $\varepsilon>0$,

$$M_2(z)=2^{3/2}(1-\log 2)\sigma^{-3}(1-z)^{-1/2}L(z)+c_4(1-z)^{-1/2}+O\bigl(|1-z|^{-\frac12+\frac{\delta}{2}-\varepsilon}\bigr). \tag{3.21}$$

Proof. We use Lemma 3.1 with the notation $R_\ell(z)$ as in (3.8). For $\ell=2$, Lemma 3.1 shows, using (2.19), that

$$R_2(z)=B(z)^{\odot 2}\odot y(z)+2B(z)\odot\bigl[zM_1(z)\Phi'(y(z))\bigr]+zM_1(z)^2\Phi''\bigl(y(z)\bigr). \tag{3.22}$$

We consider the three terms separately.

First, by (3.5), (2.20) and Lemma 2.3 (twice), we have

$$B(z)^{\odot 2}\odot y(z)=B(z)^{\odot 2}\odot\bigl(y(z)-1\bigr)=c_5+O\bigl(|1-z|^{\frac12-2\varepsilon}\bigr). \tag{3.23}$$

For the remaining two terms, we have to be more careful, since it will turn out that their main terms cancel.

For the second term, we note first that (3.10) implies $M_1(z)=O\bigl(|1-z|^{-\varepsilon}\bigr)$, and thus (2.27) yields

$$zM_1(z)\Phi'(y(z))=M_1(z)+O\bigl(|1-z|^{\frac12-\varepsilon}\bigr). \tag{3.24}$$

Hence, (3.5) and Lemma 2.3 yield

$$B(z) \odot \bigl[zM_1(z)\Phi'(y(z))\bigr] = B(z) \odot M_1(z) + c_6 + O\bigl(|1-z|^{\frac12-2\varepsilon}\bigr). \tag{3.25}$$

This implies, using (3.5), (3.10), and Lemma 2.3 again, followed by (3.3), and recalling $\mathrm{Li}_{0,1} \odot L(z) = \mathrm{Li}_{0,1} \odot \mathrm{Li}_{1,0}(z) = \mathrm{Li}_{1,1}(z)$,

$$\begin{aligned} B(z) \odot \bigl[zM_1(z)\Phi'(y(z))\bigr] &= -\sigma^{-2}B(z) \odot L(z) + c_7 + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr) \\ &= -\sigma^{-2}\bigl(\mathrm{Li}_{0,1}(z) \odot L(z) - \mu' L(z)\bigr) + c_7 + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr) \\ &= -\sigma^{-2}\,\mathrm{Li}_{1,1}(z) + \sigma^{-2}\mu' L(z) + c_7 + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr). \end{aligned} \tag{3.26}$$

We use the singular expansion of $\mathrm{Li}_{1,1}(z)$:

$$\mathrm{Li}_{1,1}(z) = \tfrac12 L^2(z) - \gamma L(z) + c_8 + O\bigl(|1-z|^{1-\varepsilon}\bigr), \tag{3.27}$$

which follows from [8, p. 380] and is given in [7, p. 96], except that the error term there should be $O(|(1-z)L(z)|)$, not $O(|1-z|)$. Consequently, (3.26) yields

$$B(z) \odot \bigl[zM_1(z)\Phi'(y(z))\bigr] = -\tfrac12\sigma^{-2}L^2(z) + \sigma^{-2}(\gamma+\mu')L(z) + c_9 + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr). \tag{3.28}$$

For the third term in (3.22), we have by (2.28) and (3.10), again using $M_1(z) = O\bigl(|1-z|^{-\varepsilon}\bigr)$,

$$\begin{aligned} zM_1(z)^2\Phi''\bigl(y(z)\bigr) &= \sigma^2 M_1(z)^2 + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr) \\ &= \sigma^{-2}L^2(z) - 2\frac{\mu' - \psi(-\frac12)}{\sigma^2}L(z) + c_{10} + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr). \end{aligned} \tag{3.29}$$

Finally, (3.22) yields, by summing (3.23), (3.28) (twice) and (3.29), recalling (3.9),

$$\begin{aligned} R_2(z) &= 2\frac{\gamma + \psi(-\frac12)}{\sigma^2}L(z) + c_{11} + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr) \\ &= 4(1-\log 2)\sigma^{-2}L(z) + c_{11} + O\bigl(|1-z|^{\frac{\delta}{2}-2\varepsilon}\bigr). \end{aligned} \tag{3.30}$$

The result (3.21) now follows by (3.30), (3.8), and (2.22), and replacing ε by $\varepsilon/2$ (as we may because ε is arbitrary). ☐

This gives the asymptotics for the second moment of the shape functional. Again, the result includes earlier results for special cases in [3; 15; 18; 7; 1]. Recall from (3.2) that $F(\mathcal{T}_n) = X_n'(0) - \mu' n$.

Lemma 3.5. *Assume* (1.7) *with* $\delta > 0$. *Then*

$$\mathbb{E}\bigl[\bigl(X_n'(0) - \mu' n\bigr)^2\bigr] = \mathbb{E}\bigl[F(\mathcal{T}_n)^2\bigr] = 4(1-\log 2)\sigma^{-2} n\log n + O\bigl(n\bigr), \tag{3.31}$$

and thus

$$\operatorname{Var} X_n'(0) = \operatorname{Var} F(\mathcal{T}_n) = 4(1-\log 2)\sigma^{-2} n\log n + O\bigl(n\bigr). \tag{3.32}$$

Proof. We may assume $\delta \leqslant 1$. The definition (3.6) and the singular expansion (3.21) yield by standard singularity analysis (using (2.10)–(2.11) or [9, Figure VI.5, p. 388])

$$q_n \,\mathbb{E}\big[F(\mathcal{T}_n)^2\big] = \frac{2^{3/2}(1-\log 2)}{\sqrt{\pi}}\sigma^{-3}n^{-1/2}\log n + O\big(n^{-\frac12}\big). \tag{3.33}$$

Hence, (3.31) follows by (2.24). Finally, (3.32) follows by (3.31) and (3.19). □

3.3. **Higher moments.** We extend the results above to higher moments, using the method used earlier for special cases in [7; 1]; see also [18] for a different method (in another special case).

We prove the following analogue of [5, Lemma 12.8]. Note that (3.34) is not true for $\ell = 1$, since the leading power of $L(z)$ in that case is $L(z)^1$ by (3.10). (Also (3.35) fails for $\ell = 1$ in general.)

Lemma 3.6. *Assume* (1.7) *with* $0 < \delta \leqslant 1$. *Then, for every* $\ell \geqslant 2$, $M_\ell(z)$ *is* Δ-*analytic, and, for any* $\varepsilon > 0$,

$$M_\ell(z) = \sigma^{-\ell-1}(1-z)^{(1-\ell)/2}\sum_{j=0}^{\lfloor \ell/2\rfloor}\kappa_{\ell,j}L(z)^j + O\big(|1-z|^{-\frac12\ell+\frac12+\frac{\delta}{2}-\varepsilon}\big) \tag{3.34}$$

$$= \sigma^{-\ell-1}\sum_{j=0}^{\lfloor \ell/2\rfloor}\widehat{\kappa}_{\ell,j}\,\mathrm{Li}_{(3-\ell)/2,j}(z) + O\big(|1-z|^{-\frac12\ell+\frac12+\frac{\delta}{2}-\varepsilon}\big), \tag{3.35}$$

for some coefficients $\kappa_{\ell,j}$ *and* $\widehat{\kappa}_{\ell,j}$. *The leading coefficients* $\kappa^*_{2k} := \kappa_{2k,k}$ *in the case that* $\ell = 2k$ *is even are given by the recursion*

$$\kappa^*_2 = 2^{3/2}(1-\log 2), \tag{3.36}$$

$$\kappa^*_{2k} = 2^{-3/2}\sum_{i=1}^{k-1}\binom{2k}{2i}\kappa^*_{2i}\kappa^*_{2(k-i)}, \qquad k \geqslant 2. \tag{3.37}$$

Furthermore,

$$\widehat{\kappa}_{2k,k} = \Gamma\big(k-\tfrac12\big)^{-1}\kappa_{2k,k} = \Gamma\big(k-\tfrac12\big)^{-1}\kappa^*_{2k}. \tag{3.38}$$

Proof. Note first that (3.34) and (3.35) are equivalent by Lemma 2.1, and that (3.38) follows using (2.11).

We use induction on ℓ. The base case $\ell = 2$ (including (3.36)) is Lemma 3.4, so we assume $\ell \geqslant 3$. We follow the proof of [5, Lemma 12.8], *mutatis mutandis.*

We first note that $L(z) = O\big(|1-z|^{-\varepsilon}\big)$. Hence, for every $\ell' < \ell$, the induction hypothesis and (for the case $\ell' = 1$) Lemma 3.2 show that

$$M_{\ell'}(z) = O\big(|1-z|^{-\frac12\ell'+\frac12-\varepsilon}\big). \tag{3.39}$$

(Here and in the sequel we replace without further comment, as we may, $c\varepsilon$ by ε, for any constant c.) Hence, using Lemma 2.6, for a typical term in (3.7) (with $m \geqslant 0$),

$$zM_{\ell_1}(z)\cdots M_{\ell_m}(z)\Phi^{(m)}\big(y(z)\big) = O\big(|1-z|^{-\frac12\sum_{i=1}^m \ell_i+\frac12 m-\varepsilon}\Phi^{(m)}\big(y(z)\big)\big)$$

$$= \begin{cases} O\bigl(|1-z|^{-\frac12(\ell-\ell_0)+\frac12 m-\varepsilon}\bigr), & m\leqslant 2,\\ O\bigl(|1-z|^{-\frac12(\ell-\ell_0)+1+\frac{\delta}{2}-\varepsilon}\bigr), & m\geqslant 3.\end{cases} \tag{3.40}$$

Since $\ell-\ell_0\geqslant m$, the exponent here is <0. Hence, (3.5) and Lemma 2.3 applied ℓ_0 times yield

$$B(z)^{\odot\ell_0}\odot\bigl[zM_{\ell_1}(z)\cdots M_{\ell_m}(z)\Phi^{(m)}\bigl(y(z)\bigr)\bigr] = \begin{cases} O\bigl(|1-z|^{-\frac12\ell+\frac12\ell_0+\frac12 m-\varepsilon}\bigr), & m\leqslant 2,\\ O\bigl(|1-z|^{-\frac12\ell+\frac12\ell_0+1+\frac{\delta}{2}-\varepsilon}\bigr), & m\geqslant 3.\end{cases} \tag{3.41}$$

If $m=0$, then $\ell_0=\ell\geqslant 3$, and if $m=1$, then $\ell_1<\ell$ and thus $\ell_0=\ell-\ell_1\geqslant 1$. Hence, except in the two cases (1) $m=1$ and $\ell_0=1$ and (2) $m=2$ and $\ell_0=0$, we have $m+\ell_0\geqslant 3$, and then the exponent in (3.41) is $\geqslant -\frac12\ell+1+\frac{\delta}{2}-\varepsilon$. Consequently, by (3.7)–(3.8),

$$R_l(z) = \ell B(z)\odot\bigl[zM_{\ell-1}(z)\Phi'\bigl(y(z)\bigr)\bigr] + \frac12\sum_{j=1}^{\ell-1}\binom{\ell}{j} zM_j(z)M_{\ell-j}(z)\Phi''\bigl(y(z)\bigr) + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr). \tag{3.42}$$

By (2.27), (3.39), (3.5) and Lemma 2.3, we have, similarly to (3.25),

$$B(z)\odot\bigl[zM_{\ell-1}(z)\Phi'(y(z))\bigr] = B(z)\odot M_{\ell-1}(z) + O\bigl(|1-z|^{-\frac12\ell+\frac32-\varepsilon}\bigr). \tag{3.43}$$

Hence, using also (2.28) and (again) (3.39), we can simplify (3.42) to

$$R_l(z) = \ell B(z)\odot M_{\ell-1}(z) + \frac{\sigma^2}{2}\sum_{j=1}^{\ell-1}\binom{\ell}{j}M_j(z)M_{\ell-j}(z) + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr). \tag{3.44}$$

In the remaining estimates we have to be more careful, in particular since there will be important cancellations. (This is as in the case $\ell=2$ treated earlier, but somewhat different.)

Consider first the Hadamard product in (3.44) (the case $m=1$ and $\ell_0=1$ above). We now use the induction hypothesis in the form (3.35) and obtain by (2.14) and (3.3), using again (3.5) and Lemma 2.3 for the error term, and finally rewriting by (2.8),

$$\begin{aligned} &B(z)\odot M_{\ell-1}(z)\\ &\quad= \sigma^{-\ell}\sum_{j=0}^{\lfloor(\ell-1)/2\rfloor}\widehat{\kappa}_{\ell-1,j}\bigl(\operatorname{Li}_{(4-\ell)/2,j+1}(z)-\mu'\operatorname{Li}_{(4-\ell)/2,j}(z)\bigr) + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr)\\ &\quad= \sigma^{-\ell}\sum_{k=0}^{\lfloor(\ell+1)/2\rfloor}c^{(1)}_{\ell,k}\operatorname{Li}_{(4-\ell)/2,k}(z) + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr)\\ &\quad= \sigma^{-\ell}(1-z)^{-\frac12\ell+1}\sum_{k=0}^{\lfloor(\ell+1)/2\rfloor}c^{(2)}_{\ell,k}L(z)^k + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr), \end{aligned} \tag{3.45}$$

where the leading coefficient in the sum is, using (2.9) and (2.11),

$$c^{(2)}_{\ell,\lfloor(\ell+1)/2\rfloor} = \Gamma(\ell/2-1)c^{(1)}_{\ell,\lfloor(\ell+1)/2\rfloor} = \Gamma(\ell/2-1)\widehat{\kappa}_{\ell-1,\lfloor(\ell-1)/2\rfloor} = \kappa_{\ell-1,\lfloor(\ell-1)/2\rfloor}. \tag{3.46}$$

The leading term in (3.45) is thus

$$\sigma^{-\ell}\kappa_{\ell-1,\lfloor(\ell-1)/2\rfloor}(1-z)^{-\frac12\ell+1}L(z)^{\lfloor(\ell+1)/2\rfloor}. \tag{3.47}$$

Consider now the terms with $j=1$ and $j=\ell-1$ in the sum in (3.44). By Lemma 3.2 and the induction hypothesis, we have

$$\sigma^2 M_1(z)M_{\ell-1}(z) = \sigma^{-\ell}(1-z)^{-\frac12\ell+1}\sum_{j=0}^{\lfloor(\ell-1)/2\rfloor}\kappa_{\ell-1,j}\bigl[-L(z)^{j+1}+cL(z)^j\bigr] + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr). \tag{3.48}$$

Note that the leading term in (3.48) cancels (3.47). Consequently, (3.45)–(3.48) yield

$$\ell B(z)\odot M_{\ell-1}(z) + \frac{\sigma^2}{2}\cdot 2\cdot\binom{\ell}{1}M_1(z)M_{\ell-1}(z) = (1-z)^{-\frac12\ell+1}\sum_{k=0}^{\lfloor(\ell-1)/2\rfloor}c^{(3)}_{\ell,k}L(z)^k + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr). \tag{3.49}$$

The remaining terms in (3.44) yield immediately, by the induction hypothesis,

$$\frac{\sigma^2}{2}\sum_{j=2}^{\ell-2}\binom{\ell}{j}M_j(z)M_{\ell-j}(z) = (1-z)^{-\frac12\ell+1}\sum_{k=0}^{\lfloor\ell/2\rfloor}c^{(4)}_{\ell,k}L(z)^k + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr). \tag{3.50}$$

Finally, (3.44) and (3.49)–(3.50) yield

$$R_\ell(z) = (1-z)^{-\frac12\ell+1}\sum_{j=0}^{\lfloor\ell/2\rfloor}c^{(5)}_{\ell,j}L(z)^j + O\bigl(|1-z|^{-\frac12\ell+1+\frac{\delta}{2}-\varepsilon}\bigr), \tag{3.51}$$

and (3.34) follows by (3.8) and (2.22), which completes the induction step.

It remains only to show the recursion (3.37) for the leading coefficients. If $\ell=2k$ is even, with $\ell\geqslant 4$, then (3.49) does not contribute to $c^{(5)}_{2k,k}$ nor thus to $\kappa_{2k,k}$, and neither do the terms in (3.50) with j odd. Hence, the argument above yields

$$c^{(5)}_{2k,k} = \frac12\sum_{i=1}^{k-1}\binom{2k}{2i}\sigma^{-2k}\kappa_{2i,i}\kappa_{2k-2i,k-i} \tag{3.52}$$

and thus, recalling again (2.22),

$$\kappa_{2k,k} = 2^{-3/2}\sum_{i=1}^{k-1}\binom{2k}{2i}\kappa_{2i,i}\kappa_{2k-2i,k-i}, \tag{3.53}$$

which is (3.37). □

The recursion (3.37) is the same as [5, (C.35)], and thus has the same solution [5, (C.40)], i.e.,

$$\kappa^*_{2k} = 2^{3/2}\frac{(2k)!\,(2k-2)!}{(k-1)!\,k!}d_1^k, \qquad k \geqslant 1, \tag{3.54}$$

with, see [5, (C.36)] and (3.36),

$$d_1 := 2^{-3/2}\kappa^*_2/2 = \tfrac{1}{2}(1-\log 2). \tag{3.55}$$

This is what we need to complete the proof of the asymptotic normality of $F(\mathcal{T}_n)$.

Proof of Theorem 1.3. If $\ell \geqslant 2$, then (3.6), the expansion (3.35), (2.5), and standard singularity analysis yield

$$q_n\,\mathbb{E}\big[F(\mathcal{T}_n)^\ell\big] = \sigma^{-\ell-1}\widehat{\kappa}_{\ell,\lfloor \ell/2\rfloor}n^{(\ell-3)/2}(\log n)^{\lfloor \ell/2\rfloor} + O\big(n^{(\ell-3)/2}(\log n)^{\lfloor \ell/2\rfloor-1}\big). \tag{3.56}$$

Hence, using (2.24),

$$\mathbb{E}\big[F(\mathcal{T}_n)^\ell\big] = \sigma^{-\ell}\sqrt{2\pi}\widehat{\kappa}_{\ell,\lfloor \ell/2\rfloor}n^{\ell/2}(\log n)^{\lfloor \ell/2\rfloor} + O\big(n^{\ell/2}(\log n)^{\lfloor \ell/2\rfloor-1}\big). \tag{3.57}$$

Consequently,

$$\frac{\mathbb{E}\big[F(\mathcal{T}_n)^\ell\big]}{(n\log n)^{\ell/2}} \to \begin{cases} 0, & \ell = 2k+1 \geqslant 3, \\ \sigma^{-2k}\sqrt{2\pi}\widehat{\kappa}_{2k,k}, & \ell = 2k \geqslant 2. \end{cases} \tag{3.58}$$

Furthermore, (3.58) holds also for $\ell = 1$ (with limit 0) by (3.19).

For even $\ell = 2k$, the limit in (3.58) is by (3.38), (3.54), and (3.55), cf. [5, (C.41)],

$$\begin{aligned}\sigma^{-2k}\frac{\sqrt{2\pi}}{\Gamma(k-\frac{1}{2})}\kappa^*_{2k} &= \sigma^{-2k}\frac{4\sqrt{\pi}}{\Gamma(k-\frac{1}{2})}\frac{(2k)!\,(2k-2)!}{(k-1)!\,k!}d_1^k = \sigma^{-2k}2^{2k}\frac{(2k)!}{k!}d_1^k \\ &= \big(8d_1\sigma^{-2}\big)^k\cdot(2k-1)!! = \big(4(1-\log 2)\sigma^{-2}\big)^k\cdot(2k-1)!!. \end{aligned}\tag{3.59}$$

Consequently, the limits appearing in (3.58) are the moments of a normal distribution $N\big(0, 4(1-\log 2)\sigma^{-2}\big)$, and thus (1.10) follows by the method of moments. (Recall that $F(\mathcal{T}_n) = X_n'(0) - \mu' n$ by (3.2).) ☐

4. Imaginary powers

In this section, we consider $X_n(\alpha)$ in (1.1) when the exponent α is purely imaginary, i.e., $\operatorname{Re}\alpha = 0$. We exclude the trivial case $\alpha = 0$, when $X_n(\alpha) = n$ is non-random. We assume throughout the section that $0 < \delta < 1$ and that (1.7) holds. As above, ε is an arbitrarily small positive number, and we replace $c\varepsilon$ by ε without comment.

We follow rather closely the argument for the case $0 < \operatorname{Re}\alpha < 1/2$ in [5, §12.4–6], but we will see new terms appearing that will lead to the dominating terms with logarithmic factors for the moments; this is very similar to the argument in Section 3, but we will see some differences. (Notably, there are no cancellations of leading terms like those in Section 3.)

As in [5, §12.4], we define

$$b_n := n^\alpha - \mu(\alpha), \tag{4.1}$$

with the following generating function (cf. [5, (12.44)] and (2.7), and note $\operatorname{Li}_0(z) = z(1-z)^{-1}$):

$$B(z) = B_\alpha(z) := \sum_{n=1}^{\infty} b_n z^n = \operatorname{Li}_{-\alpha}(z) - \mu(\alpha)\operatorname{Li}_0(z) \tag{4.2}$$

$$= \Gamma(1+\alpha)(1-z)^{-\alpha-1} - \mu(\alpha)(1-z)^{-1} + O(1) \tag{4.3}$$

$$= O\bigl(|1-z|^{-1}\bigr). \tag{4.4}$$

Let now $F(T) = F_\alpha(T)$ denote the additive functional defined by the toll function $f_\alpha(T) := b_{|T|}$. Thus,

$$F_\alpha(\mathcal{T}_n) = X_n(\alpha) - n\mu(\alpha). \tag{4.5}$$

4.1. **The mean.** For the mean, we define the generating function

$$M_\alpha(z) := \mathbb{E}\bigl[F_\alpha(\mathcal{T})z^{|\mathcal{T}|}\bigr] = \sum_{n=1}^{\infty} q_n \,\mathbb{E}[F_\alpha(\mathcal{T}_n)]z^n. \tag{4.6}$$

We then have, as in (3.12) and [5, (12.29)],

$$M_\alpha(z) = \frac{zy'(z)}{y(z)} \cdot \bigl(B_\alpha(z) \odot y(z)\bigr). \tag{4.7}$$

Thus $M_\alpha(z)$ is Δ-analytic. Further, we have by (2.13), (4.2), and (2.23), using (4.4) and Lemma 2.3 for the error term in (2.23), and then using for the second line (2.7) and $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$,

$$\begin{aligned} B_\alpha(z) \odot y(z) &= \frac{1}{\sqrt{2\pi}\sigma}\operatorname{Li}_{3/2-\alpha}(z) - \frac{\mu(\alpha)}{\sqrt{2\pi}\sigma}\operatorname{Li}_{3/2}(z) + c_1 + O\bigl(|1-z|^{\frac{1}{2}+\frac{\delta}{2}}\bigr) \\ &= \frac{\Gamma(\alpha-\frac{1}{2})}{\sqrt{2\pi}\sigma}(1-z)^{\frac{1}{2}-\alpha} + 2^{1/2}\sigma^{-1}\mu(\alpha)(1-z)^{1/2} + c_2 + O\bigl(|1-z|^{\frac{1}{2}+\frac{\delta}{2}}\bigr). \end{aligned} \tag{4.8}$$

Further, similarly to (3.14),

$$(B_\alpha \odot y)(1) = \sum_{n=1}^{\infty} b_n q_n = \sum_{n=1}^{\infty} q_n[n^\alpha - \mu(\alpha)] = \mathbb{E}\,|\mathcal{T}|^\alpha - \mu(\alpha) = 0. \tag{4.9}$$

Thus, letting $z \to 1$ in (4.8) shows that $c_2 = (B_\alpha \odot y)(1) = 0$.

Finally, (4.7), (2.22), and (4.8) yield, using (2.7) again,

$$M_\alpha(z) = \frac{\Gamma(\alpha-\frac{1}{2})}{2\sqrt{\pi}\sigma^2}(1-z)^{-\alpha} + \sigma^{-2}\mu(\alpha) + O\bigl(|1-z|^{\frac{\delta}{2}}\bigr) \tag{4.10}$$

$$= \frac{\Gamma(\alpha-\frac{1}{2})}{2\sqrt{\pi}\sigma^2\Gamma(\alpha)}\operatorname{Li}_{1-\alpha}(z) + c + O\bigl(|1-z|^{\frac{\delta}{2}}\bigr). \tag{4.11}$$

Singularity analysis now yields, from (4.6) and (4.11),

$$q_n\,\mathbb{E}[F_\alpha(\mathcal{T}_n)] = \frac{\Gamma(\alpha-\frac{1}{2})}{2\sqrt{\pi}\sigma^2\Gamma(\alpha)} n^{\alpha-1} + O\bigl(n^{-1-\frac{\delta}{2}}\bigr) \tag{4.12}$$

and thus, by (2.24),

$$\mathbb{E}[F_\alpha(\mathcal{T}_n)] = \frac{\Gamma(\alpha-\frac{1}{2})}{\sqrt{2}\sigma\Gamma(\alpha)} n^{\frac{1}{2}+\alpha} + O\bigl(n^{\frac{1}{2}-\frac{\delta}{2}}\bigr) \tag{4.13}$$

Hence, recalling (4.5),

$$\mathbb{E}\, X_n(\alpha) = \mu(\alpha)n + \frac{\Gamma(\alpha-\frac{1}{2})}{\sqrt{2}\sigma\Gamma(\alpha)} n^{\frac{1}{2}+\alpha} + O\bigl(n^{\frac{1}{2}-\frac{\delta}{2}}\bigr). \tag{4.14}$$

This agrees with [5, Theorem 1.7(ii)] (proved without (1.7), and by different methods), except that the error estimate here is smaller.

4.2. **Higher moments.** For higher moments, we need mixed moments for α and $\overline{\alpha} = -\alpha$. Thus, somewhat more generally, fix α_1 and α_2 with $\operatorname{Re}\alpha_1 = \operatorname{Re}\alpha_2 = 0$ but $\alpha_1 \neq 0 \neq \alpha_2$. We define, for integers $\ell_1, \ell_2 \geqslant 0$, the generating function

$$M_{\ell_1,\ell_2}(z) := \mathbb{E}\bigl[F_{\alpha_1}(\mathcal{T})^{\ell_1} F_{\alpha_2}(\mathcal{T})^{\ell_2} z^{|\mathcal{T}|}\bigr] = \sum_{n=1}^{\infty} q_n\, \mathbb{E}\bigl[F_{\alpha_1}(\mathcal{T}_n)^{\ell_1} F_{\alpha_2}(\mathcal{T}_n)^{\ell_2}\bigr] z^n. \tag{4.15}$$

Thus $M_{1,0} = M_{\alpha_1}$ and $M_{0,1} = M_{\alpha_2}$ are given by (4.7). The functions $M_{\ell,r}$ can then be found by the following recursion, given in [5, (12.75)], for every $\ell, r \geqslant 0$ with $\ell + r \geqslant 1$:

$$M_{\ell,r}(z) = \frac{zy'(z)}{y(z)} \sum_{m=0}^{\ell+r} \frac{1}{m!} \sum\nolimits^{**} \binom{\ell}{\ell_0,\dots,\ell_m}\binom{r}{r_0,\dots,r_m} B_{\alpha_1}(z)^{\odot \ell_0} \odot B_{\alpha_2}(z)^{\odot r_0} \odot \bigl[zM_{\ell_1,r_1}(z)\cdots M_{\ell_m,r_m}(z)\Phi^{(m)}\bigl(y(z)\bigr)\bigr], \tag{4.16}$$

where $\sum^{**}$ is the sum over all pairs of $(m+1)$-tuples $(\ell_0,\dots,\ell_m)$ and $(r_0,\dots,r_m)$ of non-negative integers that sum to ℓ and r, respectively, such that $1 \leqslant \ell_i + r_i < \ell + r$ for every $i \geqslant 1$. (Note that there are two typos in [5]: the lower summation limit should be $m = 0$, and the final qualification "$i \geqslant 1$" is missing there.) It follows by induction that every $M_{\ell,r}$ is Δ-analytic.

We define for convenience $R_{\ell,r}(z)$ as the sum in (4.16); thus

$$M_{\ell,r}(z) = \frac{zy'(z)}{y(z)} R_{\ell,r}(z). \tag{4.17}$$

Let us first consider second moments. Taking $\ell = r = 1$ in (4.16) yields, recalling (2.19),

$$R_{1,1}(z) = B_{\alpha_1}(z) \odot B_{\alpha_2}(z) \odot y(z) + B_{\alpha_1}(z) \odot [zM_{0,1}(z)\Phi'(y(z))] + B_{\alpha_2}(z) \odot [zM_{1,0}(z)\Phi'(y(z))] + zM_{1,0}(z)M_{0,1}(z)\Phi''(y(z)). \tag{4.18}$$

The first term is, by (4.4) and (4.8) together with Lemma 2.3,

$$O\bigl(|1-z|^{-1}\bigr) \odot O\bigl(|1-z|^{1/2}\bigr) = c_3 + O\bigl(|1-z|^{1/2}\bigr). \tag{4.19}$$

For the other terms in (4.18), we first note from (4.10) that $M_{1,0}(z) = M_{\alpha_1}(z) = O(1)$ and $M_{0,1}(z) = M_{\alpha_2}(z) = O(1)$. Thus, using also (2.27)–(2.28), (4.4) and Lemma 2.3, we may simplify to

$$R_{1,1}(z) = c_4 + B_{\alpha_1}(z) \odot M_{0,1}(z) + B_{\alpha_2}(z) \odot M_{1,0}(z) + M_{1,0}(z)M_{0,1}(z)\sigma^2 + O\bigl(|1-z|^{\delta/2}\bigr). \tag{4.20}$$

Furthermore, (4.10) yields

$$M_{1,0}(z)M_{0,1}(z) = c_5(1-z)^{-\alpha_1} + c_6(1-z)^{-\alpha_2} + c_7(1-z)^{-\alpha_1-\alpha_2} + c_8 + O\bigl(|1-z|^{\frac{\delta}{2}}\bigr). \tag{4.21}$$

We compute the Hadamard products in (4.20) by (2.13), (4.2) and (4.11), using again (4.4) and Lemma 2.3 for the error term. Together with (4.21), this yields from (4.20) a result that we write, using (2.7), as

$$\begin{aligned} R_{1,1}(z) = \Bigl(&\frac{\Gamma(\alpha_2-\frac12)}{2\sqrt{\pi}\sigma^2\Gamma(\alpha_2)} + \frac{\Gamma(\alpha_1-\frac12)}{2\sqrt{\pi}\sigma^2\Gamma(\alpha_1)}\Bigr)\operatorname{Li}_{1-\alpha_1-\alpha_2}(z) \\ &+ c_9(1-z)^{-\alpha_1} + c_{10}(1-z)^{-\alpha_2} + c_7(1-z)^{-\alpha_1-\alpha_2} + c_{11} \\ &+ O\bigl(|1-z|^{\delta/2}\bigr). \end{aligned} \tag{4.22}$$

If $\alpha_1 + \alpha_2 \neq 0$, we use (2.7) also on the first term and obtain

$$\begin{aligned} R_{1,1}(z) &= c_{12}(1-z)^{-\alpha_1-\alpha_2} + c_9(1-z)^{-\alpha_1} + c_{10}(1-z)^{-\alpha_2} + c_{13} \\ &\quad + O\bigl(|1-z|^{\delta/2}\bigr). \end{aligned} \tag{4.23}$$

On the other hand, if $\alpha_1 + \alpha_2 = 0$, we recall that $\operatorname{Li}_1(z) = L(z)$, and thus (4.22) yields

$$\begin{aligned} R_{1,1}(z) &= \frac{1}{\sqrt{\pi}\sigma^2}\operatorname{Re}\frac{\Gamma(\alpha_1-\frac12)}{\Gamma(\alpha_1)}\cdot L(z) + c_9(1-z)^{-\alpha_1} + c_{10}(1-z)^{-\alpha_2} + c_{14} \\ &\quad + O\bigl(|1-z|^{\delta/2}\bigr). \end{aligned} \tag{4.24}$$

We can now obtain $M_{1,1}(z)$ from (4.23)–(4.24) by (4.17) and (2.22). We do not state the result separately, but proceed immediately to a general formula.

Lemma 4.1. *Let $\alpha \neq 0$ with $\operatorname{Re}\alpha = 0$, and take $\alpha_1 = \alpha$ and $\alpha_2 = \overline{\alpha} = -\alpha$. Then, for each pair of integers $\ell, r \geqslant 0$ with $\ell + r \geqslant 2$, $M_{\ell,r}(z)$ is Δ-analytic and we have, for some coefficients $\varkappa_{\ell,r;j,k}$ and $\widehat{\varkappa}_{\ell,r;j,k}$, and every $\varepsilon > 0$,*

$$M_{\ell,r}(z) = \sum_{j,k} \varkappa_{\ell,r;j,k}(1-z)^{(1-\ell-r)/2+j\alpha}L(z)^k + O\bigl(|1-z|^{\frac12(1-\ell-r)+\frac{\delta}{2}-\varepsilon}\bigr) \tag{4.25}$$

$$= \sum_{j,k} \widehat{\varkappa}_{\ell,r;j,k}\operatorname{Li}_{(3-\ell-r)/2+j\alpha,k}(z) + O\bigl(|1-z|^{\frac12(1-\ell-r)+\frac{\delta}{2}-\varepsilon}\bigr), \tag{4.26}$$

where the sums are over integers j and k with $-\ell \leqslant j \leqslant r$ and $0 \leqslant k \leqslant \ell \wedge r$.

Furthermore, if $\ell + r = 1$, then (4.25) holds (but not (4.26)).

If $\ell = r$, then the only non-zero coefficients with $k = \ell = r$ are

$$\varkappa_{\ell,\ell;0,\ell} = \sigma^{-2\ell-1}\varkappa^*_\ell, \tag{4.27}$$

$$\widehat{\varkappa}_{\ell,\ell;0,\ell} = \Gamma\bigl(\ell-\tfrac12\bigr)^{-1}\varkappa_{\ell,\ell;0,\ell} = \frac{\sigma^{-2\ell-1}}{\Gamma\bigl(\ell-\frac12\bigr)}\varkappa^*_\ell, \tag{4.28}$$

where $\varkappa^*_\ell$ is given by the recursion

$$\varkappa^*_1 = \frac{1}{\sqrt{2\pi}}\operatorname{Re}\frac{\Gamma(\alpha-\frac12)}{\Gamma(\alpha)}, \tag{4.29}$$

$$\varkappa^*_\ell = 2^{-3/2}\sum_{i=1}^{\ell-1}\binom{\ell}{i}^2\varkappa^*_i\varkappa^*_{\ell-i}, \qquad \ell\geqslant 2. \tag{4.30}$$

Proof. Note first that for $\ell+r=1$, (4.25) follows from (4.10). (We see also from (4.11) that (4.26) would hold if we add a constant term; the problem is that $\operatorname{Li}_1(z)$ is $L(z)$ and not a constant.)

Assume in the rest of the proof that $\ell+r\geqslant 2$. Then the expansions (4.25) and (4.26) are equivalent by Lemma 2.1; furthermore, for the leading terms, (4.27) and (4.28) are equivalent by (2.11).

Consider next the case $\ell+r=2$. If $(\ell,r)=(2,0)$ or $(0,2)$, we use (4.23) with $\alpha_1=\alpha_2=\pm\alpha$ and obtain (4.25) by (4.17) and (2.22). (Now only terms with $k=0$ appear.)

If $\ell=r=1$, we similarly use (4.24), (4.17) and (2.22) and obtain (4.25) including a single term with $k=1$, viz. $\varkappa_{1,1;0,1}L(z)(1-z)^{-1/2}$ with $\varkappa_{1,1;0,1}$ given by (4.27) and (4.29).

For $\ell+r\geqslant 3$, we use induction on $\ell+r$. By the induction hypothesis (4.25) (including the case $\ell+r=1$ just proved by (4.10)), we have for every (ℓ',r') with $1\leqslant \ell'+r'<\ell+r$,

$$M_{\ell',r'}(z) = O\bigl(|1-z|^{-\frac12(\ell'+r')+\frac12-\varepsilon}\bigr). \tag{4.31}$$

Consequently, for a typical term in (4.16), as in (3.40) and using again Lemma 2.6,

$$\begin{aligned} zM_{\ell_1,r_1}(z)\cdots M_{\ell_m,r_m}(z)\Phi^{(m)}\bigl(y(z)\bigr) &= O\bigl(|1-z|^{-\frac12\sum_{i=1}^m(\ell_i+r_i)+\frac12 m-\varepsilon}\Phi^{(m)}\bigl(y(z)\bigr)\bigr)\\ &= \begin{cases} O\bigl(|1-z|^{-\frac12(\ell+r-\ell_0-r_0)+\frac12 m-\varepsilon}\bigr), & m\leqslant 2,\\ O\bigl(|1-z|^{-\frac12(\ell+r-\ell_0-r_0)+1+\frac{\delta}{2}-\varepsilon}\bigr), & m\geqslant 3.\end{cases} \end{aligned} \tag{4.32}$$

Again the exponent here is <0, and it follows by (4.4) and Lemma 2.3 that

$$\begin{aligned} &B_{\alpha_1}(z)^{\odot\ell_0}\odot B_{\alpha_2}(z)^{\odot r_0}\odot\bigl[zM_{\ell_1,r_1}(z)\cdots M_{\ell_m,r_m}(z)\Phi^{(m)}\bigl(y(z)\bigr)\bigr]\\ &\qquad= \begin{cases} O\bigl(|1-z|^{-\frac12(\ell+r)+\frac12(\ell_0+r_0)+\frac12 m-\varepsilon}\bigr), & m\leqslant 2,\\ O\bigl(|1-z|^{-\frac12(\ell+r)+\frac12(\ell_0+r_0)+1+\frac{\delta}{2}-\varepsilon}\bigr), & m\geqslant 3.\end{cases} \end{aligned} \tag{4.33}$$

As in the proof of Lemma 3.6, except in the two cases (1) $m = 1$ and $\ell_0 + r_0 = 1$ and (2) $m = 2$ and $\ell_0 = r_0 = 0$ we have $m + \ell_0 + r_0 \geqslant 3$, and then the exponent in (4.33) is $\geqslant -\frac{1}{2}(\ell + r) + 1 + \frac{\delta}{2} - \varepsilon$. Consequently, by (4.16)–(4.17),

$$\begin{aligned} R_{l,r}(z) &= \ell B_{\alpha_1}(z) \odot \bigl[zM_{\ell-1,r}(z)\Phi'\bigl(y(z)\bigr)\bigr] + rB_{\alpha_2}(z) \odot \bigl[zM_{\ell,r-1}(z)\Phi'\bigl(y(z)\bigr)\bigr] \\ &\quad + \frac{1}{2} \sum\sum_{0<i+j<\ell+r} \binom{\ell}{i}\binom{r}{j} zM_{i,j}(z)M_{\ell-i,r-j}(z)\Phi''\bigl(y(z)\bigr) \\ &\quad + O\bigl(|1-z|^{-\frac{1}{2}(\ell+r)+1+\frac{\delta}{2}-\varepsilon}\bigr). \end{aligned} \tag{4.34}$$

As in (3.42)–(3.44) and (4.18)–(4.20), this can be simplified, using (2.27)–(2.28), (4.31), (4.4) and Lemma 2.3, and we obtain

$$\begin{aligned} R_{l,r}(z) &= \ell B_{\alpha_1}(z) \odot M_{\ell-1,r}(z) + rB_{\alpha_2}(z) \odot M_{\ell,r-1}(z) \\ &\quad + \frac{\sigma^2}{2} \sum\sum_{0<i+j<\ell+r} \binom{\ell}{i}\binom{r}{j} M_{i,j}(z)M_{\ell-i,r-j}(z) + O\bigl(|1-z|^{-\frac{1}{2}(\ell+r)+1+\frac{\delta}{2}-\varepsilon}\bigr). \end{aligned} \tag{4.35}$$

By the induction hypothesis in the form (4.26) and (4.2), using as always Lemma 2.3 for the error term, we have

$$\begin{aligned} B_{\alpha_1}(z) \odot M_{\ell-1,r}(z) &= \sum_{j,k} \widehat{\varkappa}_{\ell-1,r;j,k} \operatorname{Li}_{(4-\ell-r)/2+j\alpha,k}(z) \odot \bigl(\operatorname{Li}_{-\alpha}(z) - \mu(\alpha)\operatorname{Li}_0(z)\bigr) \\ &\qquad + O\bigl(|1-z|^{-\frac{1}{2}(\ell+r)+1+\frac{\delta}{2}-\varepsilon}\bigr) \end{aligned} \tag{4.36}$$

summing over $-(\ell-1) \leqslant j \leqslant r$ and $0 \leqslant k \leqslant (\ell-1) \wedge r$. By (2.14), this can be rearranged as

$$\sum_{j,k} c^{(1)}_{\ell,r;j,k} \operatorname{Li}_{(4-\ell-r)/2+j\alpha,k}(z) + O\bigl(|1-z|^{-\frac{1}{2}(\ell+r)+1+\frac{\delta}{2}-\varepsilon}\bigr), \tag{4.37}$$

now summing over $-\ell \leqslant j \leqslant r$ and $0 \leqslant k \leqslant (\ell-1) \wedge r$. By Lemma 2.1, this can also be written

$$\sum_{j,k} c^{(2)}_{\ell,r;j,k} (1-z)^{(2-\ell-r)/2+j\alpha} L(z)^k + O\bigl(|1-z|^{-\frac{1}{2}(\ell+r)+1+\frac{\delta}{2}-\varepsilon}\bigr), \tag{4.38}$$

still summing over $-\ell \leqslant j \leqslant r$ and $0 \leqslant k \leqslant (\ell-1) \wedge r$.

By symmetry, $B_{\alpha_2}(z) \odot M_{\ell,r-1}(z)$ can also be written as (4.38) (with different coefficients $c^{(2)}_{\ell,r;j,k}$), now summing over $-\ell \leqslant j \leqslant r$ and $0 \leqslant k \leqslant \ell \wedge (r-1)$.

Finally, the double sum in (4.35) can by the induction hypothesis (4.25) also be written as (4.38), summing over $-\ell \leqslant j \leqslant r$ and $0 \leqslant k \leqslant \ell \wedge r$.

Consequently, (4.35) yields

$$R_{\ell,r}(z) = \sum_{j,k} c^{(3)}_{\ell,r;j,k} (1-z)^{(2-\ell-r)/2+j\alpha} L(z)^k + O\bigl(|1-z|^{-\frac{1}{2}(\ell+r)+1+\frac{\delta}{2}-\varepsilon}\bigr), \tag{4.39}$$

summing over $-\ell \leqslant j \leqslant r$ and $0 \leqslant k \leqslant \ell \wedge r$. By (4.17) and (2.22), this implies (4.25), which completes the induction proof of (4.25)–(4.26).

Now consider the case $\ell = r \geqslant 2$. We see that then the only terms above with $k = \ell = r$ come from the double sum in (4.35); moreover, they appear only for terms there with $i = j$, and we obtain by induction that the only non-zero coefficient in (4.39) with $k = \ell$ is, using (4.27),

$$c^{(3)}_{\ell,\ell;0,\ell} = \frac{\sigma^2}{2}\sum_{i=1}^{\ell-1}\binom{\ell}{i}^2 \varkappa_{i,i;0,i}\varkappa_{\ell-i,\ell-i;0,\ell-i} = \frac{1}{2}\sigma^{-2\ell}\sum_{i=1}^{\ell-1}\binom{\ell}{i}^2 \varkappa^*_i\varkappa^*_{\ell-i} \tag{4.40}$$

Hence, when deriving (4.25) from (4.39) by (4.17) and (2.22), we also find that the only non-zero coefficient with $k = \ell$ is

$$\varkappa_{\ell,\ell;0,\ell} = 2^{-1/2}\sigma^{-1}c^{(3)}_{\ell,\ell;0,\ell} = 2^{-3/2}\sigma^{-2\ell-1}\sum_{i=1}^{\ell-1}\binom{\ell}{i}^2\varkappa^*_i\varkappa^*_{\ell-i}. \tag{4.41}$$

This proves (4.27) and (4.30). □

The recursion (4.30) is the same as [5, (D.6)], and thus has the same solution [5, (D.10)]

$$\varkappa^*_\ell = 2^{3/2}\frac{\ell!\,(2\ell-2)!}{(\ell-1)!}d_1^\ell, \tag{4.42}$$

with, by [5, (D.9)] and (4.29),

$$d_1 := 2^{-3/2}\varkappa^*_1 = \frac{1}{4\sqrt{\pi}}\operatorname{Re}\frac{\Gamma(\alpha-\frac12)}{\Gamma(\alpha)}. \tag{4.43}$$

Proof of Theorem 1.4. We have $\alpha = it$. If $\ell + r \geqslant 2$, then (4.15), (4.26), (2.5), and singularity analysis yield

$$q_n\,\mathbb{E}\big[F_\alpha(\mathcal{T}_n)^\ell\,\overline{F_\alpha(\mathcal{T}_n)}^r\big] = q_n\,\mathbb{E}\big[F_\alpha(\mathcal{T}_n)^\ell F_{\overline{\alpha}}(\mathcal{T}_n)^r\big] = O\big(n^{(\ell+r-3)/2}(\log n)^{\ell\wedge r}\big). \tag{4.44}$$

When $\ell = r$, we find more precisely

$$q_n\,\mathbb{E}\big[F_\alpha(\mathcal{T}_n)^\ell\,\overline{F_\alpha(\mathcal{T}_n)}^\ell\big] = \widehat{\varkappa}_{\ell,\ell;0,\ell}n^{(2\ell-3)/2}(\log n)^\ell + O\big(n^{(2\ell-3)/2}(\log n)^{\ell-1}\big). \tag{4.45}$$

Hence, using (2.24) and (4.28),

$$\mathbb{E}\big[F_\alpha(\mathcal{T}_n)^\ell\,\overline{F_\alpha(\mathcal{T}_n)}^r\big] = \begin{cases} O\big(n^{(\ell+r)/2}(\log n)^{\ell\wedge r}\big), & \ell \neq r,\\ \sigma^{-2\ell}\frac{\sqrt{2\pi}}{\Gamma(l-\frac12)}\varkappa^*_\ell n^\ell(\log n)^\ell + O\big(n^\ell(\log n)^{\ell-1}\big), & \ell = r.\end{cases} \tag{4.46}$$

Consequently,

$$\frac{\mathbb{E}\big[F_\alpha(\mathcal{T}_n)^\ell\,\overline{F_\alpha(\mathcal{T}_n)}^r\big]}{(n\log n)^{(\ell+r)/2}} \to \begin{cases} 0, & \ell \neq r,\\ \sigma^{-2\ell}\frac{\sqrt{2\pi}}{\Gamma(l-\frac12)}\varkappa^*_\ell, & \ell = r \geqslant 1.\end{cases} \tag{4.47}$$

Furthermore, (4.47) holds also for $\ell + r = 1$ by (4.13).

For $\ell = r$, the limit in (4.47) is by (4.42) and (4.43), cf. [5, (D.11)],

$$\sigma^{-2\ell}\frac{\sqrt{2\pi}}{\Gamma(\ell-\frac12)}\varkappa^*_\ell = \sigma^{-2\ell}\frac{4\sqrt{\pi}}{\Gamma(\ell-\frac12)}\frac{\ell!\,(2\ell-2)!}{(\ell-1)!}d_1^\ell = \sigma^{-2\ell}2^{2\ell}\ell!\,d_1^\ell$$

$$= \bigl(4d_1\sigma^{-2}\bigr)^{\ell} \cdot \ell! = \Bigl(\frac{1}{\sqrt{\pi}\sigma^2} \operatorname{Re} \frac{\Gamma(\alpha-\frac12)}{\Gamma(\alpha)}\Bigr)^{\ell} \cdot \ell!. \tag{4.48}$$

Consequently, by (2.1), the limits in (4.47) are the moments of a symmetric complex normal distribution with variance (1.12), and thus (1.11) follows by the method of moments. (Recall that $F_\alpha(\mathcal{T}_n) = X_n(\alpha) - \mu(\alpha)n$ by (4.5).)

Finally, the claim in (1.12) that the variance is nonzero follows from the same claim in (1.14) (where the variance is the same up to a factor $\sigma^2/2$), which is shown in [5, Theorem D.1], with the correction in Appendix D below. ☐

4.3. **Joint distributions.** We can extend the arguments above to joint distributions of several $X_n(\alpha)$ with different imaginary α. Since we have $X_n(\overline{\alpha}) = \overline{X_n(\alpha)}$, it suffices to consider the case $\operatorname{Im}\alpha > 0$. In this case, different $X_n(\alpha)$ are asymptotically independent, as is stated more precisely in the following theorem.

Theorem 4.2. *For any finite set $t_1, \dots, t_n$ of distinct positive numbers, the complex random variables $\bigl(X_n(\mathrm{i}t_k) - \mu(\mathrm{i}t_k)n\bigr)/\sqrt{n\log n}$ converge, as $n \to \infty$, jointly in distribution to independent symmetric complex normal variables $\zeta_{\mathrm{i}t_k}$ with variances given by* (1.12).

This can be interpreted as joint convergence (in the product topology) of the entire family $\{X_n(it) : t > 0\}$ of random variables, after normalization, to an (uncountable) family of *independent* symmetric complex normal variables $\zeta_{\mathrm{i}t}$. As said in Remark 1.8, this behaviour is strikingly different from the cases $\operatorname{Re}\alpha < 0$ and $\operatorname{Re}\alpha > 0$, where we have joint convergence to analytic random functions of α.

Proof. We argue as above, using the method of moments and singularity analysis of generating functions, with mainly notational differences. We give only a sketch, leaving further details to the reader.

For a sequence of arbitrary non-zero imaginary numbers $\alpha_1, \dots, \alpha_\ell$ (allowing repetitions), define the generating function

$$M_{\alpha_1,\dots,\alpha_\ell}(z) := \mathbb{E}\bigl[F_{\alpha_1}(\mathcal{T}) \cdots F_{\alpha_\ell}(\mathcal{T}) z^{|\mathcal{T}|}\bigr] = \sum_{n=1}^{\infty} q_n \, \mathbb{E}\bigl[F_{\alpha_1}(\mathcal{T}_n) \cdots F_{\alpha_\ell}(\mathcal{T}_n)\bigr] z^n. \tag{4.49}$$

When $\ell = 1$ and 2, these are the same as $M_{\alpha_1}(z)$ or $M_{1,1}(z)$ in the notation used above. The recursion (4.16) extends as follows. We write again

$$M_{\alpha_1,\dots,\alpha_\ell}(z) = \frac{zy'(z)}{y(z)} R_{\alpha_1,\dots,\alpha_\ell}(z). \tag{4.50}$$

Then, by a straightforward extension of the proof of [5, Lemma 12.4], cf.(4.16),

$$R_{\alpha_1,\dots,\alpha_\ell}(z) = \sum_{m=0}^{\ell} \frac{1}{m!} \sum B_{\alpha_{i_1}}(z) \odot \cdots \odot B_{\alpha_{i_q}}(z) \odot \bigl[zM_{A_1}(z) \cdots M_{A_m}(z)\Phi^{(m)}\bigl(y(z)\bigr)\bigr] \tag{4.51}$$

where we sum over all partitions of $[\ell] := \{1, \dots, \ell\}$ into an ordered sequence of $m+1$ sets $I_0, \dots, I_m$ with $I_1, \dots, I_m$ neither empty nor equal to the full set $[\ell]$ (while

I_0 may be empty or equal to $[\ell]$), and i_j are defined by $I_0 = \{i_1, \dots, i_q\}$ and, for $1 \leqslant j \leqslant m$, A_j is the sequence $(\alpha_i : i \in I_j)$.

As in Lemma 4.1, it follows by induction that for any sequence $A = (\alpha_1, \dots, \alpha_\ell)$ of length $|A| = \ell \geqslant 2$,

$$M_A(z) = \sum_{\beta,k} \varkappa_{A;\beta,k} (1-z)^{(1-\ell)/2+\beta} L(z)^k + O\bigl(|1-z|^{\frac{1}{2}(1-\ell)+\frac{\delta}{2}-\varepsilon}\bigr) \tag{4.52}$$

$$= \sum_{\beta,k} \widehat{\varkappa}_{A;\beta,k} \operatorname{Li}_{(3-\ell)/2+\beta,k}(z) + O\bigl(|1-z|^{\frac{1}{2}(1-\ell)+\frac{\delta}{2}-\varepsilon}\bigr), \tag{4.53}$$

where we sum over $0 \leqslant k \leqslant \ell/2$ and all β such that $-\beta$ equals the sum of some subsequence of A. Moreover, as above, it is seen by induction that the only non-zero coefficients with $k = \ell/2$ have $\beta = 0$, and that they appear only when ℓ is even and A is balanced in the sense that it can be partitioned into $\ell/2$ pairs $\{\alpha_i, -\alpha_i\}$. We now write $\varkappa^*_A := \varkappa_{A;0,k}$ if A is balanced with $|A| = 2k$. (We let $\varkappa^*_A := 0$ if A is not balanced.) These leading terms come from the case $m = 2$ and $q = 0$ in (4.51), and we obtain the recurrence, for $|A| \geqslant 4$,

$$\varkappa^*_A = 2^{-3/2}\sigma \sum \varkappa^*_{A_1} \varkappa^*_{A_2}, \tag{4.54}$$

summing over all partitions of A into two nonempty sets A_1 and A_2 that both are balanced.

It follows by induction from (4.54) that if $|A| = 2k \geqslant 2$, then $\varkappa^*_A$ can be written as a sum

$$\varkappa^*_A = \bigl(2^{-3/2}\sigma\bigr)^{k-1} \sum \prod_{j=1}^{k} \varkappa^*_{A_j}, \tag{4.55}$$

where we sum over full binary trees with k leaves, where each leaf is labelled by a pair I_j of indices such that $I_1, \dots, I_k$ form a partition of $[2k]$, and furthermore the corresponding sets A_j are balanced, i.e., $\alpha_i + \alpha_{i'} = 0$ if $I_j = \{i, i'\}$.

Let $A = (\alpha_1, \dots, \alpha_{2k})$ consist of the numbers $\mathrm{i}t_j$ and $-\mathrm{i}t_j$ repeated k_j times each, for $j = 1, \dots, r$, where $t_1, \dots, t_r$ are distinct and positive; thus $|A| = 2k$ with $k = \sum_j k_j$. Then there are $\prod_j k_j!$ ways to partition A into balanced pairs, and for each binary tree with k leaves, these pairs can be assigned to the k leaves in $k!$ ways. Each tree and each assignment of balanced pairs A_i gives the same contribution to the sum (4.55), and we obtain, since there are $C_{k-1} = (2k-2)!/(k!(k-1)!)$ full binary trees with k leaves,

$$\varkappa^*_A = \bigl(2^{-3/2}\sigma\bigr)^{k-1} \frac{(2k-2)!}{(k-1)!} \prod_{j=1}^{r} \Bigl[(\varkappa^*_{\{\pm \mathrm{i}t_j\}})^{k_j} k_j! \Bigr]. \tag{4.56}$$

Let $\sigma^2_{\mathrm{i}t}$ be the variance of $\zeta_{\mathrm{i}t}$ in (1.12). For the case $A = \{\mathrm{i}t, -\mathrm{i}t\}$, Lemma 4.1 applies and we have by (4.27) and (4.29), in the present notation,

$$\varkappa^*_{\{\pm \mathrm{i}t\}} = 2^{-1/2}\sigma^{-1}\sigma^2_{\mathrm{i}t}. \tag{4.57}$$

Hence, (4.56) yields

$$\varkappa_A^* = 2^{-2k+\frac{3}{2}}\sigma^{-1}\frac{(2k-2)!}{(k-1)!}\prod_{j=1}^{r}\left(\sigma_{\mathrm{it}_j}^{2k_j}k_j!\right). \tag{4.58}$$

Since $\widehat{\varkappa}_{A;0,k} = \Gamma(k-\frac{1}{2})^{-1}\varkappa_A^*$, we finally obtain from (4.53), using (2.24), that

$$\begin{aligned} n^{-k}\,\mathbb{E}\big[F_{\alpha_1}(\mathcal{T}_n)\cdots F_{\alpha_{2k}}(\mathcal{T}_n)\big] &\to 2^{-2k+2}\sqrt{\pi}\frac{(2k-2)!}{\Gamma(k-\frac{1}{2})(k-1)!}\prod_{j=1}^{r}\left(\sigma_{\mathrm{it}_j}^{2k_j}k_j!\right) \\ &= \prod_{j=1}^{r}\left(\sigma_{\mathrm{it}_j}^{2k_j}k_j!\right), \end{aligned} \tag{4.59}$$

which equals the corresponding mixed moment $\mathbb{E}\big(\zeta_{\alpha_1}\cdots\zeta_{\alpha_{2k}}\big) = \prod_j \mathbb{E}\,|\zeta_{\mathrm{it}_j}|^{2k_j}$, see (2.1). Similarly, all mixed moments with unbalanced indices converge after normalization to 0. Hence, the result follows by the method of moments. □

Note that the combinatorial argument in the final part of the proof (restricted to the case $r = 1$) yields an alternative proof that the recursion (4.41) is solved by (4.42)–(4.43). Conversely, the argument above without detailed counting of possibilities shows that the left-hand side of (4.59) converges to c_k times the right-hand side, for some combinatorial constant c_k not depending on $k_1, \dots k_r$. Since (4.47) shows that the formula is correct for $r = 1$, we must have $c_k = 1$, and thus (4.59) holds.

5. Negative real part

In this section, we consider the case that α in (1.1) has negative real part. Applying the same approach as in previous sections, we prove convergence of all moments for the normalized random variable. As before, we assume throughout the section that (1.7) holds with $0 < \delta < 1$. Again, we set

$$b_n := n^\alpha - \mu(\alpha), \tag{5.1}$$

with the generating function

$$B(z) = B_\alpha(z) := \sum_{n=1}^{\infty} b_n z^n = \mathrm{Li}_{-\alpha}(z) - \mu(\alpha)\,\mathrm{Li}_0(z). \tag{5.2}$$

In contrast to Section 4, the term $\mu(\alpha)\,\mathrm{Li}_0(z) = \mu(\alpha)z(1-z)^{-1}$ now dominates. For later convenience, we let $\eta := \min(-\operatorname{Re}\alpha, \delta/2)$, and note that $0 < \eta < \frac{1}{2}$ (assuming again $\delta < 1$ as we may). Then (2.7) implies

$$B(z) = -\mu(\alpha)(1-z)^{-1} + O\big(|1-z|^{-1+\eta}\big). \tag{5.3}$$

This is even true for $\alpha \in \{-1, -2, \dots\}$, where logarithmic terms occur in the asymptotic expansion of $\mathrm{Li}_{-\alpha}$, due to the aforementioned fact that $\eta < \frac{1}{2}$.

Once again, we let $F(T) = F_\alpha(T)$ denote the additive functional defined by the toll function $f_\alpha(T) := b_{|T|}$, so that

$$F_\alpha(\mathcal{T}_n) = X_n(\alpha) - n\mu(\alpha). \tag{5.4}$$

5.1. **The mean.** We use the same notation for the generating function of the mean as in Section 4, i.e.,

$$M_\alpha(z) := \mathbb{E}\bigl[F_\alpha(\mathcal{T})z^{|\mathcal{T}|}\bigr] = \sum_{n=1}^{\infty} q_n\,\mathbb{E}[F_\alpha(\mathcal{T}_n)]z^n, \tag{5.5}$$

and note that (4.7) still holds:

$$M_\alpha(z) = \frac{zy'(z)}{y(z)}\cdot\bigl(B_\alpha(z)\odot y(z)\bigr). \tag{5.6}$$

Thus $M_\alpha(z)$ is still Δ-analytic. In analogy with (4.8), we now have

$$\begin{aligned} B_\alpha(z)\odot y(z) &= \frac{1}{\sqrt{2\pi}\sigma}\operatorname{Li}_{3/2-\alpha}(z) - \frac{\mu(\alpha)}{\sqrt{2\pi}\sigma}\operatorname{Li}_{3/2}(z) + c_1 + O\bigl(|1-z|^{\frac12+\frac{\delta}{2}}\bigr)\\ &= 2^{1/2}\sigma^{-1}\mu(\alpha)(1-z)^{1/2} + c_2 + O\bigl(|1-z|^{\frac12+\eta}\bigr). \end{aligned} \tag{5.7}$$

Moreover, (4.9) still holds, so $c_2 = 0$. Combining this with (2.22) now yields

$$M_\alpha(z) = \sigma^{-2}\mu(\alpha) + O\bigl(|1-z|^{\eta}\bigr). \tag{5.8}$$

Applying singularity analysis and (2.24), we find that

$$\mathbb{E}[F_\alpha(\mathcal{T}_n)] = O\bigl(n^{\frac12-\eta}\bigr) \tag{5.9}$$

or equivalently

$$\mathbb{E}\,X_n(\alpha) = \mu(\alpha)n + O\bigl(n^{\frac12-\eta}\bigr). \tag{5.10}$$

5.2. **Higher moments.** As in Section 4.2, we consider the mixed moments of $F_{\alpha_1}(\mathcal{T}_n)$ and $F_{\alpha_2}(\mathcal{T}_n)$ for two complex numbers α_1 and α_2 that are now both assumed to have negative real part. In particular, this includes the special case that $\alpha_2 = \overline{\alpha}_1$. We are thus interested in the generating function

$$M_{\ell_1,\ell_2}(z) := \mathbb{E}\bigl[F_{\alpha_1}(\mathcal{T})^{\ell_1}F_{\alpha_2}(\mathcal{T})^{\ell_2}z^{|\mathcal{T}|}\bigr] \tag{5.11}$$

for integers $\ell_1,\ell_2 \geqslant 0$, cf. (4.15). In particular, we have $M_{1,0} = M_{\alpha_1}$ and $M_{0,1} = M_{\alpha_2}$. Set $\eta := \min(-\operatorname{Re}\alpha_1, -\operatorname{Re}\alpha_2, \delta/2)$ (again noting that $\eta < \frac12$). Then by (5.8) we have

$$M_{1,0}(z) = \sigma^{-2}\mu(\alpha_1) + O\bigl(|1-z|^{\eta}\bigr) \text{ and } M_{0,1}(z) = \sigma^{-2}\mu(\alpha_2) + O\bigl(|1-z|^{\eta}\bigr). \tag{5.12}$$

In order to deal with higher moments, we make use of the recursion (4.16). Let us start with second-order moments: here, we obtain

$$\begin{aligned} M_{1,1}(z) = \frac{zy'(z)}{y(z)}\bigl[&B_{\alpha_1}(z)\odot B_{\alpha_2}(z)\odot y(z) + B_{\alpha_1}(z)\odot(zM_{0,1}(z)\Phi'(y(z)))\\ &+B_{\alpha_2}(z)\odot(zM_{1,0}(z)\Phi'(y(z))) + zM_{1,0}(z)M_{0,1}(z)\Phi''(y(z))\bigr]. \end{aligned} \tag{5.13}$$

In view of (5.8), (2.20), (2.27), and (2.28), the functions $y(z)$, $zM_{0,1}(z)\Phi'\bigl(y(z)\bigr)$, $zM_{1,0}(z)\Phi'\bigl(y(z)\bigr)$, and $zM_{1,0}(z)M_{0,1}(z)\Phi''\bigl(y(z)\bigr)$ are all of the form $c + O\bigl(|1-z|^{\eta}\bigr)$, and taking the Hadamard product with $B_{\alpha_1}(z)$ or $B_{\alpha_2}(z)$ does not change this property. Combining this with (2.22) we conclude that there is a constant $\varkappa_{1,1}$ such that

$$M_{1,1}(z) = 2^{-1/2}\sigma^{-1}\varkappa_{1,1}(1-z)^{-1/2} + O\bigl(|1-z|^{-\frac{1}{2}+\eta}\bigr), \tag{5.14}$$

which implies by virtue of singularity analysis and (2.24) that

$$\mathbb{E}[F_{\alpha_1}(\mathcal{T}_n)F_{\alpha_2}(\mathcal{T}_n)] = \varkappa_{1,1}\, n + O\bigl(n^{1-\eta}\bigr). \tag{5.15}$$

We can obtain the functions $M_{2,0}(z)$ and $M_{0,2}(z)$ as special cases of $M_{1,1}(z)$ where $\alpha_1 = \alpha_2$. Hence there are also constants $\varkappa_{2,0}$ and $\varkappa_{0,2}$ such that

$$M_{2,0}(z) = 2^{-1/2}\sigma^{-1}\varkappa_{2,0}(1-z)^{-1/2} + O\bigl(|1-z|^{-\frac{1}{2}+\eta}\bigr) \tag{5.16}$$

and

$$M_{0,2}(z) = 2^{-1/2}\sigma^{-1}\varkappa_{0,2}(1-z)^{-1/2} + O\bigl(|1-z|^{-\frac{1}{2}+\eta}\bigr), \tag{5.17}$$

and thus

$$\mathbb{E}[F_{\alpha_1}(\mathcal{T}_n)^2] = \varkappa_{2,0}\, n + O\bigl(n^{1-\eta}\bigr) \text{ and } \mathbb{E}[F_{\alpha_2}(\mathcal{T}_n)^2] = \varkappa_{0,2}\, n + O\bigl(n^{1-\eta}\bigr). \tag{5.18}$$

We will use these as the base case of an inductive proof of the following lemma.

Lemma 5.1. *Suppose that* $\operatorname{Re}\alpha_1 < 0$ *and* $\operatorname{Re}\alpha_2 < 0$, *and let*

$$\eta = \min(-\operatorname{Re}\alpha_1, -\operatorname{Re}\alpha_2, \delta/2) \tag{5.19}$$

be as above. Then, for all non-negative integers ℓ *and* r *with* $s = \ell + r \geqslant 1$, *the function* $M_{\ell,r}(z)$ *is* Δ*-analytic and we have*

$$M_{\ell,r}(z) = \widehat{\varkappa}_{\ell,r}(1-z)^{(1-s)/2} + O\bigl(|1-z|^{(1-s)/2+\eta}\bigr), \tag{5.20}$$

where $\widehat{\varkappa}_{1,0} = \sigma^{-2}\mu(\alpha_1)$, $\widehat{\varkappa}_{0,1} = \sigma^{-2}\mu(\alpha_2)$, *and, for* $s \geqslant 2$,

$$\widehat{\varkappa}_{\ell,r} = \frac{(s-3)!!}{\sigma 2^{(s-1)/2}} \sum_{\substack{j=0\\ j\equiv \ell \bmod 2}}^{\ell\wedge r} \binom{\ell}{j}\binom{r}{j} j!\,(\ell-j-1)!!\,(r-j-1)!!\,\varkappa_{1,1}^{j}\varkappa_{2,0}^{(\ell-j)/2}\varkappa_{0,2}^{(r-j)/2} \tag{5.21}$$

if s *is even, and* $\widehat{\varkappa}_{\ell,r} = 0$ *otherwise.*

Proof. We prove the statement by induction on $s = \ell + r$. Note that (5.12) as well as (5.14), (5.16), and (5.17) are precisely the cases $s = 1$ and $s = 2$, respectively.

For the induction step, we take $s \geqslant 3$ and use recursion (4.16). It follows immediately from this recursion that all $M_{\ell,r}$ are Δ-analytic, so we focus on the asymptotic behavior at 1. Let us first consider the product

$$zM_{\ell_1,r_1}(z)\cdots M_{\ell_m,r_m}(z)\Phi^{(m)}\bigl(y(z)\bigr), \tag{5.22}$$

where all ℓ_i and r_i are non-negative integers, $1 \leqslant \ell_i + r_i < s$ for every $i \geqslant 1$, $\ell_0 + \ell_1 + \cdots + \ell_m = \ell$, and $r_0 + r_1 + \cdots + r_m = r$. By the induction hypothesis, $M_{\ell_i,r_i}(z) = O\bigl(|1-z|^{(1-\ell_i-r_i)/2}\bigr)$ for all $i \geqslant 1$, which can be improved to $M_{\ell_i,r_i}(z) =$

$O\big(|1-z|^{(1-\ell_i-r_i)/2+\eta}\big)$ if $\ell_i + r_i$ is odd and greater than 1. Combining with (2.29), we obtain

$$zM_{\ell_1,r_1}(z)\cdots M_{\ell_m,r_m}(z)\Phi^{(m)}\big(y(z)\big) = O\big(|1-z|^{(m-\ell_1-\cdots-\ell_m-r_1-\cdots-r_m)/2+\frac{\delta}{2}+1-m/2}\big)$$
$$= O\big(|1-z|^{(\ell_0+r_0-\ell-r)/2+1+\eta}\big) \tag{5.23}$$

for $m \geqslant 3$. This estimate continues to hold after taking the Hadamard product with $B_{\alpha_1}(z)^{\odot \ell_0} \odot B_{\alpha_2}(z)^{\odot r_0}$, and the factor $\frac{zy'(z)}{y(z)}$ in (4.16) contributes $-\frac{1}{2}$ to the exponent by (2.22). Since ℓ_0 and r_0 are non-negative, it follows that the total contribution of all terms with $m \geqslant 3$ is $O\big(|1-z|^{(1-s)/2+\eta}\big)$ and thus negligible. We can therefore focus on the cases $m = 0$, $m = 1$, and $m = 2$. Here, $\Phi^{(m)}\big(y(z)\big)$ is $O(1)$ in all cases by (2.26)–(2.28), and we obtain

$$zM_{\ell_1,r_1}(z)\cdots M_{\ell_m,r_m}(z)\Phi^{(m)}\big(y(z)\big) = O\big(|1-z|^{(m-\ell_1-\cdots-\ell_m-r_1-\cdots-r_m)/2}\big)$$
$$= O\big(|1-z|^{(m+\ell_0+r_0-\ell-r)/2}\big). \tag{5.24}$$

Terms with $m + \ell_0 + r_0 \geqslant 3$ are negligible for the same reason as before. Likewise, terms with $m+\ell_0+r_0 = 2$ are negligible if at least one of the sums ℓ_i+r_i with $i \geqslant 1$ is odd and greater than 1, as we can then improve the bound on $M_{\ell_i,r_i}(z)$. Let us determine all remaining possibilities:

- $m = 0$ implies $m + \ell_0 + r_0 = \ell + r = s \geqslant 3$, so we have already accounted for this negligible case.
- $m = 1$ gives us $\ell_0 + \ell_1 = \ell$ and $r_0 + r_1 = r$ with $1 \leqslant \ell_1 + r_1 < \ell + r$, thus $\ell_0 + r_0 \geqslant 1$. So we have $(\ell_0, \ell_1, r_0, r_1) = (1, \ell-1, 0, r)$ and $(\ell_0, \ell_1, r_0, r_1) = (0, \ell, 1, r-1)$ as the only two relevant possibilities in this case.
- Finally, if $m = 2$, we must have $\ell_0 = r_0 = 0$ and $\ell_1 + \ell_2 = \ell$ and $r_1 + r_2 = r$.

Now we divide the argument into two subcases, according as $s = \ell + r$ is even or odd.

Odd $s \geqslant 3$. If $m = 2$, $\ell_0 = r_0 = 0$, and $\ell_1 + \ell_2 + r_1 + r_2 = \ell + r = s$, then either $\ell_1 + r_1$ or $\ell_2 + r_2$ is odd. Thus the corresponding term is asymptotically negligible unless $\ell_1 + r_1 = 1$ or $\ell_2 + r_2 = 1$. So in this case, there are only four terms that might be asymptotically relevant:

$$(\ell_1, \ell_2, r_1, r_2) \in \{(1, \ell-1, 0, r), (\ell-1, 1, r, 0), (0, \ell, 1, r-1), (\ell, 0, r-1, 1)\}. \tag{5.25}$$

In addition, $m = 1$ contributes with two terms as mentioned above. Thus we obtain

$$M_{\ell,r}(z) = \frac{zy'(z)}{y(z)}\Big[\ell B_{\alpha_1}(z) \odot \big(zM_{\ell-1,r}(z)\Phi'\big(y(z)\big)\big) + rB_{\alpha_2}(z) \odot \big(zM_{\ell,r-1}(z)\Phi'\big(y(z)\big)\big)$$
$$+\ell zM_{1,0}(z)M_{\ell-1,r}(z)\Phi''\big(y(z)\big) + rzM_{0,1}(z)M_{\ell,r-1}(z)\Phi''\big(y(z)\big)\Big]$$
$$+ O\big(|1-z|^{(1-s)/2+\eta}\big). \tag{5.26}$$

By the induction hypothesis, $M_{\ell-1,r}(z) = \widehat{\varkappa}_{\ell-1,r}(1-z)^{1-\frac{s}{2}} + O\big(|1-z|^{1-\frac{s}{2}+\eta}\big)$. Consequently, using (5.8), (2.27), and (2.28), we get

$$zM_{\ell-1,r}(z)\Phi'\big(y(z)\big) = \widehat{\varkappa}_{\ell-1,r}(1-z)^{1-\frac{s}{2}} + O\big(|1-z|^{1-\frac{s}{2}+\eta}\big), \tag{5.27}$$

$$zM_{1,0}(z)M_{\ell-1,r}(z)\Phi''\big(y(z)\big) = \mu(\alpha_1)\widehat{\varkappa}_{\ell-1,r}(1-z)^{1-\frac{s}{2}} + O\big(|1-z|^{1-\frac{s}{2}+\eta}\big). \tag{5.28}$$

Recall from (5.2) that $B_{\alpha_1}(z) = \mathrm{Li}_{-\alpha_1}(z) - \mu(\alpha_1)\,\mathrm{Li}_0(z)$. Applying the Hadamard product gives us, using (2.7), (2.13), and Lemma 2.3,

$$B_{\alpha_1}(z) \odot \bigl(zM_{\ell-1,r}(z)\Phi'\bigl(y(z)\bigr)\bigr) = -\mu(\alpha_1)\widehat{\varkappa}_{\ell-1,r}(1-z)^{1-\frac{s}{2}} + O\bigl(|1-z|^{1-\frac{s}{2}+\eta}\bigr), \tag{5.29}$$

so the first and third terms in (5.26) effectively cancel, and the same argument applies to the second and fourth terms. Hence we have proven the desired statement in the case that s is odd.

Even $s \geqslant 4$. In this case, we can neglect the terms with $m = 1$ and $\ell_1 + r_1 = \ell + r - 1 = s - 1$, since $s - 1$ is odd and greater than 1. Thus only terms with $m = 2$ and $\ell_0 = r_0 = 0$ matter. For the same reason, we can ignore all terms where $\ell_1 + r_1$ and $\ell_2 + r_2$ are odd: at least one of them has to be greater then 1, making all such terms asymptotically negligible. Hence we obtain

$$M_{\ell,r}(z) = \frac{zy'(z)}{y(z)} \cdot \frac{1}{2} \sum_{\substack{\ell_1,\ell_2,r_1,r_2 \\ \ell_1+\ell_2=\ell,\, r_1+r_2=r \\ \ell_i+r_i \text{ even and } >0}} \binom{\ell}{\ell_1}\binom{r}{r_1} zM_{\ell_1,r_1}(z)M_{\ell_2,r_2}(z)\Phi''\bigl(y(z)\bigr) + O\bigl(|1-z|^{(1-s)/2+\eta}\bigr). \tag{5.30}$$

Let us write $\sum^{\circ}$ for the sum in (5.30). Plugging in (2.22), (2.28), and the induction hypothesis, we obtain

$$M_{\ell,r}(z) = 2^{-3/2}\sigma \sum\nolimits^{\circ} \binom{\ell}{\ell_1}\binom{r}{r_1} \widehat{\varkappa}_{\ell_1,r_1}\widehat{\varkappa}_{\ell_2,r_2}(1-z)^{(1-s)/2} + O\bigl(|1-z|^{(1-s)/2+\eta}\bigr). \tag{5.31}$$

Thus we have completed the induction for (5.20) with

$$\widehat{\varkappa}_{\ell,r} = 2^{-3/2}\sigma \sum\nolimits^{\circ} \binom{\ell}{\ell_1}\binom{r}{r_1} \widehat{\varkappa}_{\ell_1,r_1}\widehat{\varkappa}_{\ell_2,r_2}. \tag{5.32}$$

In order to verify the formula (5.21) for $\widehat{\varkappa}_{\ell,r}$ given in the statement of the lemma, in light of (5.14), (5.16), and (5.17) we need only show that $\widehat{\varkappa}_{\ell,r}$ as defined in (5.21) satisfies the recursion (5.32). This is easy to achieve by means of generating functions, as follows. Set

$$\begin{aligned} K(x,y) &:= \sum_{\substack{s\geqslant 2 \\ s \text{ even}}} \sum_{\ell+r=s} \widehat{\varkappa}_{\ell,r} \frac{x^\ell}{\ell!}\frac{y^r}{r!} \\ &= \sum_{\substack{s\geqslant 2 \\ s \text{ even}}} \sum_{\ell+r=s} \frac{(s-3)!!}{\sigma 2^{(s-1)/2}} \sum_{\substack{j=0 \\ j\equiv \ell \bmod 2}}^{\ell\wedge r} \binom{\ell}{j}\binom{r}{j} j!\,(\ell-j-1)!!\,(r-j-1)!! \\ &\qquad \cdot \varkappa_{1,1}^{j}\varkappa_{2,0}^{(\ell-j)/2}\varkappa_{0,2}^{(r-j)/2}\frac{x^\ell}{\ell!}\frac{y^r}{r!}. \end{aligned} \tag{5.33}$$

Setting $\ell - j = 2a$ and $r - j = 2b$, this can be rewritten as

$$\begin{aligned} K(x,y) &= \sum_{\substack{s\geqslant 2\\ s \text{ even}}} \frac{(s-3)!!}{\sigma 2^{(s-1)/2}} \sum_{\substack{a,b,j\geqslant 0:\\ a+b+j=s/2}} \frac{\varkappa_{1,1}^j \varkappa_{2,0}^a \varkappa_{0,2}^b x^{j+2a} y^{j+2b}}{j!\, a!\, b!\, 2^{a+b}} \\ &= \sum_{\substack{s\geqslant 2\\ s \text{ even}}} \frac{(s-3)!!}{\sigma 2^{(s-1)/2}(s/2)!} \left(\frac{\varkappa_{2,0}\, x^2}{2} + \varkappa_{1,1}\, xy + \frac{\varkappa_{0,2}\, y^2}{2} \right)^{s/2} \\ &= \frac{\sqrt{2}}{\sigma} \sum_{t\geqslant 1} \frac{(2t-3)!!}{t!\, 2^{2t}} \left(\varkappa_{2,0}\, x^2 + 2\varkappa_{1,1}\, xy + \varkappa_{0,2}\, y^2 \right)^t \\ &= \frac{\sqrt{2}}{\sigma} - \frac{1}{\sigma}\sqrt{2 - (\varkappa_{2,0}\, x^2 + 2\varkappa_{1,1}\, xy + \varkappa_{0,2}\, y^2)}. \end{aligned} \tag{5.34}$$

The recursion (5.32) now follows by comparing coefficients of $x^\ell y^r$ in the identity

$$K(x,y) = 2^{-3/2}\sigma K(x,y)^2 + \frac{\varkappa_{2,0}\, x^2 + 2\varkappa_{1,1}\, xy + \varkappa_{0,2}\, y^2}{2^{3/2}\sigma}. \tag{5.35}$$

This completes the proof of the lemma. □

So the functions $M_{\ell,r}(z)$ are amenable to singularity analysis, and we obtain the following theorem as an immediate application.

Theorem 5.2. *Suppose that* $\operatorname{Re}\alpha_1 < 0$ *and* $\operatorname{Re}\alpha_2 < 0$. *Then there exist constants* $\varkappa_{2,0}$, $\varkappa_{1,1}$, *and* $\varkappa_{0,2}$ *such that, for all non-negative integers* ℓ *and* r,

$$\frac{\mathbb{E}[F_{\alpha_1}(\mathcal{T}_n)^\ell F_{\alpha_2}(\mathcal{T}_n)^r]}{n^{(\ell+r)/2}} \to \sum_{\substack{j=0\\ j\equiv \ell \bmod 2}}^{\ell\wedge r} \binom{\ell}{j}\binom{r}{j} j!\, (\ell-j-1)!!\, (r-j-1)!!\, \varkappa_{1,1}^j \varkappa_{2,0}^{(\ell-j)/2} \varkappa_{0,2}^{(r-j)/2} \tag{5.36}$$

as $n\to\infty$ *if* $\ell + r$ *is even, and* $\frac{\mathbb{E}[F_{\alpha_1}(\mathcal{T}_n)^\ell F_{\alpha_2}(\mathcal{T}_n)^r]}{n^{(\ell+r)/2}} \to 0$ *otherwise.*

Proof. In view of Lemma 5.1, singularity analysis gives us

$$[z^n]M_{\ell,r}(z) = \frac{\widehat{\varkappa}_{\ell,r}}{\Gamma((s-1)/2)} n^{(s-3)/2} + O\big(n^{(s-3)/2-\eta}\big) \tag{5.37}$$

for $s = \ell + r \geqslant 2$, so, using (2.24),

$$\mathbb{E}[F_{\alpha_1}(\mathcal{T}_n)^\ell F_{\alpha_2}(\mathcal{T}_n)^r] = \frac{[z^n]M_{\ell,r}(z)}{q_n} = \frac{\sqrt{2\pi}\sigma\widehat{\varkappa}_{\ell,r}}{\Gamma((s-1)/2)} n^{s/2} + O\big(n^{s/2-\eta}\big). \tag{5.38}$$

Since $\Gamma((s-1)/2) = 2^{1-(s/2)}\sqrt{\pi}(s-3)!!$ for even s (recall (2.2)), the statement follows immediately from the formula for $\widehat{\varkappa}_{\ell,r}$ in Lemma 5.1 for all $s \geqslant 2$ and from (5.9) for $s = 1$. □

The following lemma will be used in the proof of Theorem 1.5 to establish that the limiting variance is positive. Recall the notation (1.1) and $q_k = \mathbb{P}(|\mathcal{T}| = k)$.

Lemma 5.3. *Consider any complex α with $\operatorname{Re}\alpha \neq 0$. Then there exists k such that $q_k > 0$ and $F_\alpha(\mathcal{T}_k)$ is not deterministic.*

Proof. We know that $p_0 > 0$ and that $p_j > 0$ for some $j \geqslant 2$. Fix such a value j. Let $k = 3j + 1 \geqslant 7$. Consider two realizations of the random tree $\mathcal{T}_k$, each of which has positive probability. Tree 1 has j children of the root, and precisely two of those j children have j children each; the other $j - 2$ have no children. Tree 2 also has j children of the root; precisely one of those j children (call it child 1) has j children, while the other $j - 1$ have no children; precisely one of the children of child 1 has j children, while the others have no children.

Then the values of F_α for Tree 1 and Tree 2 are, respectively,

$$3j - 2 + 2(j+1)^\alpha + (3j+1)^\alpha \tag{5.39}$$

and

$$3j - 2 + (j+1)^\alpha + (2j+1)^\alpha + (3j+1)^\alpha. \tag{5.40}$$

These values can't be equal, because otherwise we would have $(j+1)^\alpha = (2j+1)^\alpha$; but the two numbers here have unequal absolute values. □

Proof of Theorem 1.5. The limit in (5.36) equals the mixed moment $\mathbb{E}\big[\zeta_1^\ell \zeta_2^r\big]$, where ζ_1 and ζ_2 have a joint complex normal distribution and $\mathbb{E}\,\zeta_1^2 = \varkappa_{2,0}$, $\mathbb{E}\,\zeta_1\zeta_2 = \varkappa_{1,1}$, and $\mathbb{E}\,\zeta_2^2 = \varkappa_{0,2}$; this follows by Wick's theorem [10, Theorem 1.28 or Theorem 1.36] by noting that the factor $\binom{\ell}{j}\binom{r}{j} j!\,(\ell - j - 1)!!\,(r - j - 1)!!$ in (5.36) is the number of perfect matchings of ℓ (labelled) copies of ζ_1 and r copies of ζ_2 such that there are j pairs (ζ_1, ζ_2).

Hence, Theorem 1.5, except for the assertion of positive variance addressed next, follows by the method of moments, taking $\alpha_1 := \alpha$ and $\alpha_2 := \overline{\alpha}$, cf. Remark 1.6.

We already know from Theorem 5.2 that $\operatorname{Var} F_\alpha(\mathcal{T}_n) = a\,n + o(n)$ for some $a \geqslant 0$; we need only show that $a > 0$. Fix k as in Lemma 5.3. Write $v_k > 0$ for the variance of $F_\alpha(\mathcal{T}_k)$. Let $N_{n,k}$ denote the number of fringe subtrees of size k in $\mathcal{T}_n$. It follows from [13, Theorem 1.5(i)] that

$$\mathbb{E}\, N_{n,k} \sim q_k n \tag{5.41}$$

as $n \to \infty$. If for $\mathcal{T}_n$ we condition on $N_{n,k} = m$ and all of $\mathcal{T}_n$ except for fringe subtrees of size k, then the conditional variance of $F_\alpha(\mathcal{T}_n)$ is the variance of the sum of m independent copies of $F_\alpha(\mathcal{T}_k)$, namely, $m v_k$. Thus

$$\operatorname{Var} F_\alpha(\mathcal{T}_n) \geqslant v_k\, \mathbb{E}\, N_{n,k} \geqslant (1 + o(1)) v_k q_k n, \tag{5.42}$$

so the constant a mentioned at the start of this paragraph satisfies $a \geqslant v_k q_k > 0$. □

Remark 5.4. We recall that asymptotic normality of $X_n(\alpha)$, or equivalently of $F_\alpha(\mathcal{T}_n)$, is already proven in [5, Theorem 1.1]. Furthermore, [5, Section 5] shows joint asymptotic normality for several α with $\operatorname{Re}\alpha < 0$, which for the case of two values α_1 and α_2 is consistent with (5.36) (by the argument in the proof of Theorem 1.5 above). It would certainly be possible to generalize the moment convergence results in this section to convergence of mixed moments for combinations of several α_i, similarly to Section 4.3, including also the possibility $\operatorname{Re}\alpha_i \geqslant 0$ for some values of i. However, this would require a lengthy case distinction (depending on the signs of the values $\operatorname{Re}\alpha_i$), so we did not perform these calculations explicitly. Instead we just note that if we consider only the case $\operatorname{Re}\alpha_i < 0$, then convergence of all mixed moments follows from the joint convergence in (1.4) for several α_i shown in [5, Section 5] together with the uniform integrability of $|n^{-1/2}[X_n(\alpha) - \mu(\alpha)n]|^r$ for arbitrary $r > 0$ that follows from Theorem 1.5 (see Remark 1.6). □

6. Fractional moments (mainly of negative order) of tree-size: comparisons across offspring distributions

Recall from [5, Theorem 1.7] that the αth moment $\mu(\alpha) = \mathbb{E}\,|\mathcal{T}|^\alpha$ of tree size defined at (1.8) is the slope in the lead-order linear approximation $\mu(\alpha)n$ of $\mathbb{E}\,X_n(\alpha)$ whenever $\operatorname{Re}\alpha < \frac{1}{2}$; and from Theorem 1.5 that this linear approximation suffices as a centering for $X_n(\alpha)$ in order to obtain a normal limit distribution when $\operatorname{Re}(\alpha) < 0$. (See also Remark 1.7.) It is therefore of interest to compute $\mu(\alpha)$ and, similarly, the constant $\mu' = \mathbb{E}\log|\mathcal{T}|$ defined at (1.9), which serves as the centering slope in Theorem 1.3.

In [5, Appendix A] it is noted that although $\mu(\alpha)$ can be evaluated numerically, no exact values for important examples of Galton–Watson trees are known in any simple form except in the case that α is a negative integer. This section is motivated by our having noticed that for all such values (for small k) reported for four examples in that appendix, $\mu(-k)$ is smallest for binary trees [5, Example A.3], second smallest for labelled trees [5, Example A.1], second largest for full binary trees [5, Example A.4], and largest for ordered trees [5, Example A.2]. We wanted to understand why this ordering occurs and whether any such ordering could be predicted for the values μ' defined at (1.9).

In Section 6.1 we give a sufficient condition ((6.25) in Theorem 6.8) for such (strict) orderings that is fairly easy to check. In Section 6.2 we give a class of examples extending the four in [5, Appendix A] where this condition is met. In Section 6.3 we discuss numerical computation of μ', which we carry out for the four examples in [5, Appendix A] and some additional examples.

The results of this Section 6 do not require (1.7).

6.1. Comparison theory.

The main results of this section are in Theorem 6.8. Working toward those results, we begin by recalling from [5, (A.6)] (where y is called "g" and (1.7) is not required) that for $\operatorname{Re}\alpha < \frac{1}{2}$ we have

$$\mu(\alpha) = \frac{1}{\Gamma(1-\alpha)} \int_0^1 (\log \tfrac{1}{t})^{-\alpha} y'(t)\,\mathrm{d}t. \tag{6.1}$$

To utilize (6.1) directly, even merely to obtain inequalities across models for real α, one needs to compute the derivative of the tree-size probability generating function y, or at least to compare the functions y' for the compared models. This is nontrivial, since explicit computation of y' (or y) is difficult or even infeasible in examples such as m-ary trees and full m-ary trees when $m > 2$. Fortunately, according to (6.3) (and similarly (6.4) in regard to μ') to follow, one need only treat the simpler offspring probability generating function(s) Φ.

Before proceeding to our main results, we present a simple lemma, a recasting of (6.1), and a definition.

Lemma 6.1. *The function $t \mapsto t/\Phi(t)$ is the inverse function of $y : [0,1] \to [0,1]$, and it increases strictly from 0 to 1 for $t \in [0,1]$.*

Proof. It is obvious from (2.18) that $y(z)$ is continuous and strictly increasing for $z \in [0,1]$ with $y(0) = 0$ and $y(1) = 1$. Hence its inverse is also strictly increasing from 0 to 1 on $[0,1]$. Finally, (2.19) shows that the inverse is $t \mapsto t/\Phi(t)$. $\square$

We will henceforth write

$$R(\eta) := \frac{1}{y^{-1}(\eta)} = \frac{\Phi(\eta)}{\eta} \in [1,\infty), \quad \eta \in (0,1]; \tag{6.2}$$

this strictly decreasing function R will appear on several occasions in the sequel, especially in Appendix A.2.

It follows from (6.1), Lemma 6.1, a change of variables from t to $\eta = y(t)$, and (2.19) that

$$\mu(\alpha) = \frac{1}{\Gamma(1-\alpha)} \int_0^1 [\log R(\eta)]^{-\alpha}\, d\eta. \tag{6.3}$$

Further, differentiation with respect to α at $\alpha = 0$ gives

$$\mu' = -\gamma - \int_0^1 [\log\log R(\eta)]\, d\eta. \tag{6.4}$$

For the remainder of Section 6 we focus on real α and utilize the following notation.

Definition 6.2. For two real-valued functions g_1 and g_2 defined on $(0,1)$, write $g_1 \leqslant g_2$ to mean that $g_1(t) \leqslant g_2(t)$ for all $t \in (0,1)$; write $g_1 < g_2$ to mean that $g_1 \leqslant g_2$ but $g_2 \nleqslant g_1$ (equivalently, that $g_1(t) \leqslant g_2(t)$ for all $t \in (0,1)$, with strict inequality for at least one value of t); and write $g_1 \prec g_2$ to mean that $g_1(t) < g_2(t)$ for all $t \in (0,1)$.

Consider two Galton–Watson trees, $\mathcal{T}^{(1)}$ and $\mathcal{T}^{(2)}$, with respective offspring distributions ξ_1 and ξ_2. Denote the trees' respective Φ-functions by Φ_1 and Φ_2, and use similarly subscripted notation for other functions associated with the trees.

We note in passing that, as a simple consequence of Lemma 6.1 whose proof is left to the reader,

$$\Phi_1 \leqslant \Phi_2 \quad \text{if and only if} \quad y_1 \leqslant y_2, \tag{6.5}$$

and hence also

$$\Phi_1 < \Phi_2 \quad \text{if and only if} \quad y_1 < y_2. \tag{6.6}$$

The result (6.5) is perhaps of some independent interest but is used in the sequel mainly in the proof of Theorem 6.5.

Theorem 6.3. *Consider two Galton–Watson trees, $\mathcal{T}^{(1)}$ and $\mathcal{T}^{(2)}$. Suppose*

$$\Phi_1 \leqslant \Phi_2. \tag{6.7}$$

(i) *If $\alpha < 0$, then*

$$\mu_1(\alpha) \leqslant \mu_2(\alpha). \tag{6.8}$$

(ii) *If* $0 < \alpha < \frac{1}{2}$, *then*

$$\mu_1(\alpha) \geqslant \mu_2(\alpha). \tag{6.9}$$

(iii) *The centering constants for the corresponding shape functionals satisfy*

$$\mu_1' \geqslant \mu_2'. \tag{6.10}$$

Proof. This is immediate from (6.3) and (6.4). □

Note that, by considering difference quotients, each of (i) and (ii) in Theorem 6.3 implies (iii) there; one doesn't need the stronger hypothesis (6.7) for this conclusion.

Remark 6.4. The conclusions in Theorem 6.3 do not always extend from $\mu(\alpha)$ to $\mathbb{E}\, X_n(\alpha)$ for finite n. A counterexample with $n = 3$ is provided by taking $\xi_1 \sim 2\,\mathrm{Bi}(1, \frac{1}{2})$ (corresponding to uniform full binary trees, with $X_3(\alpha)$ concentrated at $2+3^\alpha$) and $\xi_2 \sim \mathrm{Ge}(\frac{1}{2})$ (corresponding to ordered trees, with $\mathbb{E}\, X_3(\alpha) = \frac{3}{2}+\frac{1}{2}2^\alpha+3^\alpha$). As shown in Lemma 6.11, we have $\Phi_1 \leqslant \Phi_2$, but Theorem 6.3(i)–(ii) with $\mathbb{E}\, X_3(\alpha)$ in place of of $\mu(\alpha)$ fails for *every* value of α, as does (6.10). □

The converse to Theorem 6.3 fails. That is, Theorem 6.3(i)–(ii) do not imply that (6.7) does, too. A counterexample is provided in Appendix A.2. However, as the next theorem shows, (6.7) has for $\alpha < 0$ a stronger consequence than Theorem 6.3(i), and this stronger consequence yields a converse result:

Theorem 6.5. *We have*

$$\mathbb{E}(|\mathcal{T}_1| - 1 + t)^\alpha \leqslant \mathbb{E}(|\mathcal{T}_2| - 1 + t)^\alpha \text{ for all integers } \alpha < 0 \text{ and all } t \in (0, \infty) \tag{6.11}$$

if and only if (6.7) *holds, in which case the inequality in* (6.11) *also holds for all real* $\alpha < 0$ *and all* $t \in (0, \infty)$.

Proof. Setting $\alpha = -k$ in (6.3) and summing over positive integers k, for complex z in the open unit disk define the function H as at [5, (A.7)]:

$$H(z) := \mathbb{E}\Bigl(1 - \frac{z}{|\mathcal{T}|}\Bigr)^{-1} = \sum_{k=0}^{\infty} \mu(-k) z^k = \int_0^1 \exp\left[z \log R(\eta)\right] \mathrm{d}\eta. \tag{6.12}$$

Changing variables (back) from η to $t = y^{-1}(\eta) = 1/R(\eta)$, we then find

$$H(z) = \int_0^1 t^{-z} y'(t)\,\mathrm{d}t = 1 + z \int_0^1 t^{-z-1} y(t)\,\mathrm{d}t, \tag{6.13}$$

with the last equality, resulting from integration by parts, as noted at [5, (A.9)]; thus

$$\mathbb{E}(|\mathcal{T}| - z)^{-1} = z^{-1}(H(z) - 1) = \int_0^1 t^{-z-1} y(t)\,\mathrm{d}t. \tag{6.14}$$

Since both the first and third expressions in (6.14) are analytic for all z with $\operatorname{Re} z < 1$, they are equal in this halfplane. Changing variables, we then find, for $\operatorname{Re} z > -1$, that

$$\mathbb{E}(|\mathcal{T}| + z)^{-1} = \int_0^\infty e^{-zx} y(e^{-x})\,\mathrm{d}x. \tag{6.15}$$

In particular, if (6.7) holds, then (recalling (6.5))

$$\begin{aligned}\mathbb{E}(|\mathcal{T}_1|-1+t)^{-1} &= \int_0^\infty e^{-tx}e^{x}y_1(e^{-x})\,\mathrm{d}x \\ &\leqslant \int_0^\infty e^{-tx}e^{x}y_2(e^{-x})\,\mathrm{d}x = \mathbb{E}(|\mathcal{T}_2|-1+t)^{-1} \qquad (6.16)\end{aligned}$$

for real $t>0$.

But more is true. Let $\Delta(z) := e^{z}[y_2(e^{-z})-y_1(e^{-z})]$. Then for $t>0$ we have that

$$h(t) := \mathbb{E}(|\mathcal{T}_2|-1+t)^{-1} - \mathbb{E}(|\mathcal{T}_1|-1+t)^{-1} = \int_0^\infty e^{-tx}\Delta(x)\,\mathrm{d}x \qquad (6.17)$$

is the Laplace transform of the bounded continuous function Δ on $(0,\infty)$. It follows from the Bernstein–Widder theorem (e.g., [2, Theorem XIII.4.1a]) that h satisfies the (weak) complete monotonicity inequalities (6.11), i.e.,

$$(-1)^r h^{(r)}(t) \geqslant 0 \text{ for all integers } r \geqslant 0 \text{ and all } t \in (0,\infty), \qquad (6.18)$$

if and only if $\Delta(x) \geqslant 0$ for a.e. $x>0$, which in turn is true if and only if $y_1 \leqslant y_2$, or (by (6.5)) equivalently (6.7), holds.

Next, if (6.7) holds, then for real $\alpha<0$ and $t \in (0,1]$ we have

$$\begin{aligned}\mathbb{E}(|\mathcal{T}_1|-1+t)^{\alpha} &= \sum_{k=0}^{\infty}\binom{|\alpha|+k-1}{k}\mu_1(\alpha-k)\,(1-t)^k \\ &\leqslant \sum_{k=0}^{\infty}\binom{|\alpha|+k-1}{k}\mu_2(\alpha-k)\,(1-t)^k \\ &= \mathbb{E}(|\mathcal{T}_2|-1+t)^{\alpha}, \qquad (6.19)\end{aligned}$$

where the inequality holds by Theorem 6.3(i).

Finally, if (6.7) holds, then for real $\alpha<0$ and $t>1$, Theorem B.1 in Appendix B implies that for $j \in \{1,2\}$ we have

$$\mathbb{E}(|\mathcal{T}_j|-1+t)^{\alpha} = (t-1)^{\alpha}\int_0^1\left[1-\frac{c_j(\eta)}{\Gamma(-\alpha)}\right]\mathrm{d}\eta, \qquad (6.20)$$

where c_j is the incomplete gamma function value

$$c_j(\eta) = \int_{(t-1)\log R_j(\eta)}^{\infty} w^{-\alpha-1}e^{-w}\,\mathrm{d}w, \qquad (6.21)$$

and from (6.20) it is evident that $\mathbb{E}(|\mathcal{T}_1|-1+t)^{\alpha} \leqslant \mathbb{E}(|\mathcal{T}_2|-1+t)^{\alpha}$. □

Remark 6.6. This remark concerns sufficient conditions for (6.7) (equivalently, by (6.5), for $y_1 \leqslant y_2$).

(a) The condition

$$|\mathcal{T}^{(1)}| \geqslant |\mathcal{T}^{(2)}| \text{ stochastically} \qquad (6.22)$$

is stronger than $y_1 \leqslant y_2$ and is of course equivalent to the condition that

$$\mathbb{E}\,g(|\mathcal{T}^{(1)}|) \geqslant \mathbb{E}\,g(|\mathcal{T}^{(2)}|) \qquad (6.23)$$

for every nonnegative nondecreasing function g defined on the positive integers. In particular, (6.22) implies the conclusions of Theorem 6.3 and (6.11) in Theorem 6.5.

Note, however, that (6.22) is *strictly* stronger than $y_1 \leqslant y_2$. While the stronger condition (6.22) holds for some of the comparisons in Section 6.2 (for example, binary trees vs. labelled trees, for which there is monotone likelihood ratio (MLR); and full binary trees vs. ordered trees, for which there is no MLR but still stochastic ordering), an example satisfying (6.7) (see Lemma 6.11 for a proof) but not (6.22) is $\xi_1 \sim \mathrm{Po}(1)$ (labelled trees) and $\xi_2 \sim 2\,\mathrm{Bi}(1, \frac{1}{2})$ (full binary trees), because $\mathbb{P}(\xi_1 \leqslant 1) = \frac{2}{e} > \frac{1}{2} = \mathbb{P}(\xi_2 \leqslant 1)$.

(b) Similarly, the condition

$$\xi_1 \geqslant \xi_2 \text{ stochastically} \tag{6.24}$$

is stronger than (6.7); indeed, it's even stronger than (6.22). But this stochastic ordering of offspring distributions can only hold if ξ_1 and ξ_2 have the same distribution, because $\mathbb{E}\,\xi_1 = \mathbb{E}\,\xi_2 = 1$. □

Remark 6.7. This remark concerns necessary conditions for (6.7).

(a) If (6.7) holds, then by a Taylor expansion near $t = 1$ [11, (A.6)] (or, alternatively, recalling (6.5), by $y_1 \leqslant y_2$ and [11, (A.5)]), $\sigma_1^2 \leqslant \sigma_2^2$. (This does not require the assumption (1.7); when (1.7) holds, we can also use [5, Lemma 12.14] or (2.20).)

(b) More generally, and by similar reasoning, if (6.7) holds and for some integer $r \geqslant 2$ we have $\mathbb{E}\,\xi_1^j = \mathbb{E}\,\xi_2^j \leqslant \infty$ for $j = 1, \dots, r-1$, then $(-1)^r\,\mathbb{E}\,\xi_1^r \leqslant (-1)^r\,\mathbb{E}\,\xi_2^r \leqslant \infty$. See Appendix C for details.

(c) We can also consider a Taylor expansion near $t = 0$. Thus if (6.7) holds, then $\mathbb{P}(\xi_1 = 0) \leqslant \mathbb{P}(\xi_2 = 0)$. More generally, if for some integer $r \geqslant 0$ we have $\mathbb{P}(\xi_1 = j) = \mathbb{P}(\xi_2 = j)$ for $j = 0, \dots, r-1$, then $\mathbb{P}(\xi_1 = r) \leqslant \mathbb{P}(\xi_2 = r)$. □

We next address the question of a stronger condition than (6.7) under which the inequalities in (6.8)–(6.10) and (6.11) are all strict. Recall the meaning of $g_1 < g_2$ described in Definition 6.2.

Theorem 6.8. *Consider two Galton–Watson trees, $\mathcal{T}^{(1)}$ and $\mathcal{T}^{(2)}$. Suppose*

$$\Phi_1 < \Phi_2. \tag{6.25}$$

(i) *If $\alpha < 0$, then*

$$\mu_1(\alpha) < \mu_2(\alpha). \tag{6.26}$$

(ii) *If $0 < \alpha < \frac{1}{2}$, then*

$$\mu_1(\alpha) > \mu_2(\alpha). \tag{6.27}$$

(iii) *We have*

$$\mu_1' > \mu_2'. \tag{6.28}$$

Proof. If (6.25) holds, then (by continuity of Φ_1 and Φ_2) strict inequality $\Phi_1(t) < \Phi_2(t)$ holds over some interval of positive length. The inequalities (6.26)–(6.28) are then immediate from (6.3)–(6.4). □

Theorem 6.9. *We have*

$$\mathbb{E}(|\mathcal{T}_1| - 1 + t)^{-m} < \mathbb{E}(|\mathcal{T}_2| - 1 + t)^{-m} \text{ for all integers } m \geqslant 0 \text{ and all } t \in (0, \infty) \tag{6.29}$$

if and only if (6.25) *holds.*

Proof. The forward direction (6.29) $\Longrightarrow$ (6.25) follows from Theorem 6.5.

For the opposite direction, use the representation (6.17) and take derivatives with respect to t. □

Remark 6.10. For all the comparison examples in Section 6.2 where the condition (6.25) holds, we in fact have the stronger condition that $\Phi_1 \prec \Phi_2$. When (6.25) holds, we can't have $\Phi_1 = \Phi_2$ over a nondegenerate interval because $\Phi_2(z) - \Phi_1(z)$ is analytic for z in the open unit disk. But it *is* possible to have $\Phi_1(t) = \Phi_2(t)$ for some values of $t \in (0, 1)$. For an example with one such value, namely, $t = 1/6$, use the notation of Appendix A.1 and take $\Phi_1 = \Phi$ and $\Phi_2 = \widetilde{\Phi}_0$. □

6.2. **Comparison examples.** In this subsection we consider the following important examples of critical Galton–Watson trees, and we fix the subscripting notation in (6.30)–(6.34) for the remainder of Section 6:

$$m\text{-ary trees: } \xi_{1,m} \sim \mathrm{Bi}(m, \tfrac{1}{m}) \ (m \geqslant 2); \tag{6.30}$$

$$\text{labelled trees: } \xi_2 \sim \mathrm{Po}(1); \tag{6.31}$$

$$\text{full binary trees: } \xi_3 \sim 2\,\mathrm{Bi}(1, \tfrac{1}{2}); \tag{6.32}$$

$$\text{ordered trees: } \xi_4 \sim \mathrm{Ge}(\tfrac{1}{2}); \tag{6.33}$$

$$\text{full } m\text{-ary trees: } \xi_{5,m} \sim m\,\mathrm{Bi}(1, \tfrac{1}{m}) \ (m \geqslant 3). \tag{6.34}$$

Observe that

$$\sigma_{1,m}^2 = 1 - \tfrac{1}{m} \uparrow \text{ strictly as } m \uparrow, \tag{6.35}$$

that

$$\sigma_{5,m}^2 = m - 1 \uparrow \text{ strictly as } m \uparrow, \tag{6.36}$$

and that, for any $m \geqslant 2$, we have

$$\sigma_{1,m}^2 < \sigma_2^2 = \sigma_3^2 < \sigma_4^2 = \sigma_{5,3}^2. \tag{6.37}$$

Further,

$$\mathbb{E}\,\xi_2^3 = 5 > 4 = \mathbb{E}\,\xi_3^3 \tag{6.38}$$

and

$$\mathbb{E}\,\xi_4^3 = 13 > 9 = \mathbb{E}\,\xi_{5,3}^3. \tag{6.39}$$

According to Remark 6.7(a)–(b) and (6.35)–(6.39), the only possible Φ-orderings in the order $<$ among the trees listed in (6.30)–(6.34) are

$$\Phi_{1,m} \uparrow \text{ strictly as } m \uparrow, \tag{6.40}$$

$$\Phi_{5,m} \uparrow \text{ strictly as } m \uparrow, \tag{6.41}$$

and, for any $m \geqslant 2$,

$$\Phi_{1,m} < \Phi_2 < \Phi_3 < \Phi_4 < \Phi_{5,3}. \tag{6.42}$$

Alternatively, we can note that

$$\mathbb{P}(\xi_{1,m} = 0) = (1 - \tfrac{1}{m})^m \uparrow \text{ strictly as } m \uparrow \tag{6.43}$$

(see (6.51) below with $t = 0$); that

$$\mathbb{P}(\xi_{5,m} = 0) = 1 - \tfrac{1}{m} \uparrow \text{ strictly as } m \uparrow; \tag{6.44}$$

that, for any $m \geqslant 2$, we have

$$\mathbb{P}(\xi_{1,m} = 0) < e^{-1} = \mathbb{P}(\xi_2 = 0) < \mathbb{P}(\xi_3 = 0) = \mathbb{P}(\xi_4 = 0) < \mathbb{P}(\xi_{5,3} = 0); \tag{6.45}$$

and, further, that

$$\mathbb{P}(\xi_3 \leqslant 1) = \tfrac{1}{2} < \tfrac{3}{4} = \mathbb{P}(\xi_4 \leqslant 1) \tag{6.46}$$

to conclude again, now using Remark 6.7(c), that the only possible Φ-orderings in the order $<$ for (6.30)–(6.34) are (6.40)–(6.42).

Remarkably, all the inequalities in (6.40)–(6.42) are true, and in fact there is strict inequality at *every* argument.

Lemma 6.11. *For every $t \in (0,1)$ we have*

$$\Phi_{1,m}(t) \uparrow \text{ strictly as } m \uparrow, \tag{6.47}$$

$$\Phi_{5,m}(t) \uparrow \text{ strictly as } m \uparrow; \tag{6.48}$$

and, for any $m \geqslant 2$,

$$\Phi_{1,m} \prec \Phi_2 \prec \Phi_3 \prec \Phi_4 \prec \Phi_{5,3}. \tag{6.49}$$

Proof. The proof is a collection of simple exercises in calculus.

Proof of (6.47). Fix $m \geqslant 2$ and $t \in (0,1)$. Observe that

$$\Phi_{1,m}(t) = (\tfrac{m-1}{m} + \tfrac{1}{m}t)^m = [1 - \tfrac{1}{m}(1-t)]^m. \tag{6.50}$$

Thus

$$\begin{aligned}
&\log \Phi_{1,m+1}(t) - \log \Phi_{1,m}(t)\\
&= (m+1)\log\Bigl(1 - \frac{1-t}{m+1}\Bigr) - m\log\Bigl(1 - \frac{1-t}{m}\Bigr)\\
&= -\left[(1-t) + \frac{(1-t)^2}{2(m+1)} + \frac{(1-t)^3}{3(m+1)^2} + \cdots\right] + \left[(1-t) + \frac{(1-t)^2}{2m} + \frac{(1-t)^3}{3m^2} + \cdots\right]\\
&> 0.
\end{aligned} \tag{6.51}$$

Proof of (6.48). Fix $m \geqslant 3$. Consider $t \in (0,1]$ and observe that

$$\Phi_{5,m}(t) = \tfrac{1}{m}(m - 1 + t^m) \tag{6.52}$$

Let $f(t) := \Phi_{5,m+1}(t) - \Phi_{5,m}(t)$. We have $f(1) = 1 - 1 = 0$ and

$$f'(t) = t^m - t^{m-1} = -t^{m-1}(1-t) < 0 \tag{6.53}$$

for $t \in (0,1)$. Thus $f(t) > 0$ for $t \in (0,1)$.

Proof of $\Phi_{1,m} \prec \Phi_2$ for $2 \leqslant m < \infty$. From (6.50) we see that

$$\Phi_{1,\infty}(t) := \lim_{m\to\infty} \Phi_{1,m}(t) = e^{t-1} = \Phi_2(t). \tag{6.54}$$

The result follows.

Proof of $\Phi_2 \prec \Phi_3$. Consider $t \in (0,1]$ and let

$$f(t) := \ln \Phi_3(t) - \ln \Phi_2(t) = \ln(1+t^2) - \ln 2 - (t-1). \tag{6.55}$$

We have $f(1) = 0$ and

$$f'(t) = 2t(1+t^2)^{-1} - 1 = -(1-t)^2(1+t^2)^{-1} < 0 \tag{6.56}$$

for $t \in (0,1)$. Thus $f(t) > 0$ for $t \in (0,1)$.

Proof of $\Phi_3 \prec \Phi_4$. Consider $t \in [0,1]$ and let

$$f(t) := \Phi_4(t) - \Phi_3(t) = \tfrac{1}{2}(1-\tfrac{1}{2}t)^{-1} - \tfrac{1}{2}(1+t^2) = \tfrac{1}{4}t(1-t)^2(1-\tfrac{1}{2}t)^{-1}. \tag{6.57}$$

Clearly, $f(t) > 0$ for $t \in (0,1)$.

Proof of $\Phi_4 \prec \Phi_{5,3}$. Consider $t \in (0,1)$ and let

$$f(t) := \frac{\Phi_{5,3}(t)}{\Phi_4(t)} = \frac{\frac{2}{3} + \frac{1}{3}t^3}{\frac{1}{2}(1-\frac{1}{2}t)^{-1}}. \tag{6.58}$$

Then

$$f(t) = 1 + \tfrac{1}{3}(1 - 2t + 2t^3 - t^4) = 1 + \tfrac{1}{3}(1-t)^3(1+t) > 1, \tag{6.59}$$

as desired. □

Theorem 6.12. (i) *If* $\alpha < 0$, *then*

$$\mu_{1,m}(\alpha) \uparrow \text{ strictly as } m \uparrow, \tag{6.60}$$

$$\mu_{5,m}(\alpha) \uparrow \text{ strictly as } m \uparrow; \tag{6.61}$$

and, for any $m \geqslant 2$,

$$\mu_{1,m}(\alpha) < \mu_2(\alpha) < \mu_3(\alpha) < \mu_4(\alpha) < \mu_{5,3}(\alpha). \tag{6.62}$$

(ii) *The orders in* (i) *are all reversed for* $0 < \alpha < \frac{1}{2}$ *and for* μ'.

Proof. The theorem is immediate from Lemma 6.11 and Theorem 6.8. □

Remark 6.13. The only two examples among (6.30)–(6.34) for which $\xi \leqslant 2$ a.s. are binary trees with $\Phi_{1,2}(t) = \frac{1}{4}(1+t)^2$ and full binary trees for which $\Phi_3(t) = \frac{1}{2}(1+t^2)$. These are two examples ($c = \frac{1}{2}$ and $c = 1$, respectively) of the most general critical Galton–Watson offspring distribution $\xi_{(c)}$ to satisfy $\xi_{(c)} \leqslant 2$ a.s., with $0 < c \leqslant 1$ and

$$\mathbb{P}(\xi_{(c)} = 0) = \mathbb{P}(\xi_{(c)} = 2) = \tfrac{1}{2}c, \qquad \mathbb{P}(\xi_{(c)} = 1) = 1 - c. \tag{6.63}$$

Generalizing $\Phi_{1,2} \prec \Phi_3$ from (6.49) in Lemma 6.11, we claim that $\Phi_{(c)}$ is strictly increasing in the order $\prec$. Indeed, for $t \in (0,1)$ we have

$$\Phi_{(c)}(t) = t + \tfrac{1}{2}c(1-t)^2, \tag{6.64}$$

which is clearly strictly increasing in $c \in (0,1]$. □

Remark 6.14. Despite a suggestion to the contrary provided by Lemma 6.11 and Remark 6.13, the partial order $\leqslant$ on tree-size probability generating functions is *not* a linear order. An example of incomparable Φ and $\widetilde{\Phi}$ is provided in Appendix A.1 (taking $\varepsilon \in (0,1]$ in the notation there). For a simpler counterexample, which shows that $\leqslant$ does not even linearly order cubic probability generating functions, let

$$\Phi(t) := \Phi_{1,2}(t) = \tfrac{1}{4}(1+t)^2 \tag{6.65}$$

correspond to binary trees, as at (6.30); and let

$$\widetilde{\Phi}(t) := \tfrac{6}{32} + \tfrac{23}{32}t + \tfrac{3}{32}t^3. \tag{6.66}$$

If $t_0 = \frac{1}{2}$, then

$$\Phi(t_0) = \tfrac{9}{16} = \tfrac{144}{256} > \tfrac{143}{256} = \widetilde{\Phi}(t_0); \tag{6.67}$$

while if $t_1 = \frac{7}{8}$, then

$$\Phi(t_1) = \tfrac{225}{256} = \tfrac{14400}{16384} < \tfrac{14405}{16384} = \widetilde{\Phi}(t_1). \tag{6.68}$$

□

6.3. **Numerical computation of μ'.** In this subsection we will compute the constant μ' for several examples of critical Galton–Watson trees. First, to set the stage for what to expect, we consider in the next remark the possible values of μ' as ξ ranges over all critical offspring distributions.

Recall (6.4). For the next remark, we find it convenient to break the integral into two pieces, using the notation $x^+ := \max\{x, 0\}$ and $x^- := \max\{-x, 0\}$:

$$\begin{aligned}\mu' &= -\gamma - \int_{t\in(0,1)} [\log\log R(t)]^+ \,\mathrm{d}t + \int_{t\in(0,1)} [\log\log R(t)]^- \,\mathrm{d}t \\ &= -\gamma - J_+ + J_-, \end{aligned} \tag{6.69}$$

say.

Remark 6.15. In this remark we argue that there is no finite upper bound, nor positive lower bound, on μ' over all Galton–Watson trees.

(a) Referring to Remark 6.13, observe that

$$\Phi_{(c)}(t) \searrow t \tag{6.70}$$

for each $t \in (0,1)$ as $c \searrow 0$. By the dominated convergence theorem (DCT), $J_+ \searrow 0$ as $c \searrow 0$. By the monotone convergence theorem (MCT), $J_- \nearrow \infty$ as $c \searrow 0$. Thus, as $c \searrow 0$ we have

$$\mu'_{(c)} \nearrow \infty. \tag{6.71}$$

Indeed, it can be shown that $\mu'_{(c)} = \log \frac{2}{c} + 1 - \gamma + o(1)$ as $c \to 0$.

(b) For the offspring distributions $\xi_{5,m}$, we have as $m \to \infty$ that

$$\mathbb{P}(\xi_{5,m} = 0) = 1 - \tfrac{1}{m} \to 1, \tag{6.72}$$

and thus $\xi_{5,m} \xrightarrow{\mathrm{p}} 0$, which implies convergence of the probability generating functions for every $t \in (0,1)$; hence,

$$\Phi_{5,\infty}(t) := \lim_{m\to\infty} \Phi_{5,m}(t) = 1, \tag{6.73}$$

which is otherwise obvious by direct calculation (showing also that the limit is an increasing one). By the MCT applied to J_+ and the DCT applied to J_-, we find

$$\mu'_{5,m} \searrow -\gamma - \int_0^1 (\log\log \tfrac{1}{t})\,\mathrm{d}t = 0 \tag{6.74}$$

as $m \nearrow \infty$. Indeed, it can be shown that $\mu'_{5,m} \sim m^{-1}\ln m$ as $m \to \infty$.

(c) We claim that the image of μ' over Galton–Watson tree models is in fact $(0,\infty)$. To see this, we first note that $\mu'_{(c)}$ is continuous in c, with $\mu'_{(1)} = \mu'_3$, which by (a) implies that the image contains $[\mu'_3,\infty)$. Further, by considering the offspring probability generating functions

$$\Phi_{\lambda,m} := \lambda\Phi_{5,m} + (1-\lambda)\Phi_3 \tag{6.75}$$

with $\lambda \in [0,1]$, one can show (by consideration of large m) that the image also contains $(0,\mu'_3)$; we omit the details.

(d) Similarly as for (c), for each fixed value of $\alpha < 0$ the image of $\mu(\alpha)$ over all Galton–Watson tree models is $(0,1)$, and for each fixed value of $\alpha \in (0,\frac{1}{2})$ the image is $(1,\infty)$. ☐

Example 6.16. The constant $\mu'_{1,2}$ is computed to 50 digits in [4, Section 5.2] using the alternative form

$$\mu' = -\gamma - \int_0^1 (\log\log \tfrac{1}{t})y'(t)\,\mathrm{d}t \tag{6.76}$$

of (6.4), explicit calculation of

$$y_{1,2}(t) = \frac{2-t-2\sqrt{1-t}}{t} = t(1+\sqrt{1-t})^{-2} \tag{6.77}$$

and thence its derivative

$$y'_{1,2}(t) = (1-t)^{-1/2}(1+\sqrt{1-t})^{-2}, \tag{6.78}$$

and numerical integration. But it is easier to use (6.4) for (high-precision) computation of μ', especially for the values $\mu'_{1,m}$ and $\mu'_{5,m}$.

As examples, we find, rounded to five digits,

$$\mu'_{1,2} = 2.0254, \qquad \mu'_2 = 1.5561, \qquad \mu'_3 = 1.4414, \qquad \mu'_4 = 1.1581. \tag{6.79}$$

Note that

$$\infty > \mu'_{1,2} > \mu'_2 > \mu'_3 > \mu'_4 > 0, \tag{6.80}$$

as guaranteed by Lemma 6.11 and Theorem 6.8(iii); see also Remark 6.15 concerning the *a priori* lack of an upper bound on $\mu'_{1,2}$ and a positive lower bound on μ'_4.

As other examples, we find, rounded to five digits,

$$\mu'_{1,2} = 2.0254, \qquad \mu'_{1,3} = 1.8224, \qquad \mu'_{1,10^3} = 1.5567; \tag{6.81}$$

$$\mu'_{5,3} = 1.0164, \qquad \mu'_{5,4} = 0.80800, \qquad \mu'_{5,10^6} = 1.5372 \times 10^{-5}; \tag{6.82}$$

and, in the notation of Remark 6.13,

$$\mu'_{(10^{-6})} = 14.931, \qquad \mu'_{(1/2)} = \mu'_{1,2} = 2.0254, \qquad \mu'_{(1-10^{-2})} = 1.4496; \tag{6.83}$$

□

Appendix A. Comparison counterexamples

A.1. **A framework for comparison counterexamples.** In this Appendix A.1 we establish a framework for various counterexamples involving comparisons of offspring distributions in Section 6. The idea is to set up two offspring distributions, say ξ and $\tilde{\xi}$, with respective probability generating functions Φ and $\widetilde{\Phi}$, such that, for (real) $t \in [0,1)$, the difference $\Delta(t) := \widetilde{\Phi}(t) - \Phi(t)$ satisfies $\Delta(t) > 0$ for most values of t, and $\Delta(t) \leqslant 0$ (but not by much) for t very near $\frac{1}{6}$ (with this value somewhat arbitrarily chosen).

Let ξ have the following probability mass function satisfying $\mathbb{E}\,\xi = 1$, as required for a critical offspring distribution:

$$p_0 := \mathbb{P}(\xi = 0) = \tfrac{1}{4} + 3e^{-3} + 5e^{-11} > 0, \tag{A.1}$$

$$p_1 := \mathbb{P}(\xi = 1) = \tfrac{1}{2} > 0, \tag{A.2}$$

$$p_2 := \mathbb{P}(\xi = 2) = \tfrac{1}{4} - 4e^{-3} + 36e^{-11} > 0, \tag{A.3}$$

$$p_k := \mathbb{P}(\xi = k) = e^{-11}\frac{8^k}{k!} > 0 \text{ for } k \geqslant 3. \tag{A.4}$$

We denote its probability generating function by Φ.

Let $\varepsilon \geqslant 0$ and for $t \in [0,1]$ define

$$g_\varepsilon(t) := \tfrac{1}{2}\left[1 - \cos\big((4\pi)(\tfrac{3}{5}t + \tfrac{2}{5})\big)\right] - \varepsilon(1-t)^3. \tag{A.5}$$

Note that

$$g_\varepsilon(1) = g'_\varepsilon(1) = 0; \tag{A.6}$$

moreover, for every $t \in [0,1]$ we have

$$-\varepsilon \leqslant g_\varepsilon(t) \leqslant 1 \tag{A.7}$$

(in particular, $g_0(t) \geqslant 0$), and one can verify for small $\varepsilon > 0$ that the set $\{t \in [0,1) : g_\varepsilon(t) < 0\}$ is an open interval of length $O(\varepsilon^{1/2})$ containing $\frac{1}{6}$.

Because $12\pi/5 < 8$, it's easy to check that there exists $c_1 > 0$ such that for all $\varepsilon \in [0,1]$ the function

$$\widetilde{\Phi}_\varepsilon(t) := \Phi(t) + c_1 g_\varepsilon(t) \tag{A.8}$$

has a power series expansion about the origin with nonnegative coefficients. From (A.6) it now follows that $\widetilde{\Phi}_\varepsilon$ is the probability generating function of a random variable $\tilde{\xi}$ with $\mathbb{E}\,\tilde{\xi} = 1$.

As we have now discussed, the difference function $\Delta_\varepsilon(t) := \widetilde{\Phi}_\varepsilon(t) - \Phi(t)$ is nonnegative when $\varepsilon = 0$; and when $\varepsilon > 0$ is small, the set

$$I_\varepsilon := \{t \in [0,1) : \Delta_\varepsilon(t) < 0\} = \{t \in [0,1) : g_\varepsilon(t) < 0\} \tag{A.9}$$

is an open interval of length $O(\varepsilon^{1/2})$ containing $\frac{1}{6}$.

Although not needed anywhere in Section 6 nor in this Appendix A, we note in passing that both ξ and $\tilde{\xi}_\varepsilon$ have moment generating functions that are finite everywhere and probability generating functions that are entire; in particular, both satisfy (1.7).

A.2. **The converse to Theorem 6.3 fails.** In this subsection we show that (i) and (ii) of Theorem 6.3 together do not imply (6.7). In fact, not even the strict inequalities in (i)–(iii) of Theorem 6.8 do.

Example A.1. In the notation of Appendix A.1, take Φ_1 to be Φ and Φ_2 to be the probability generating function $\widetilde{\Phi}_\varepsilon$ of (A.8). We do not have $\Phi_1 \leqslant \Phi_2$. But we claim that for all sufficiently small $\varepsilon > 0$ (not depending on α, to be clear), Theorem 6.8(i)–(iii) hold.

To establish the desired inequalities about $\mu(\alpha)$, we will utilize (6.3). For this we apply the mean value theorem to the function $x \mapsto (\log x)^{-\alpha}$, $x \in (1, \infty)$, as follows. Let $1 < x_1 \leqslant x_2 < \infty$. If $\alpha \leqslant -1$, then for some point $x \in [x_1, x_2]$ we have

$$\begin{aligned}(\log x_2)^{-\alpha} &- (\log x_1)^{-\alpha}\\ &= (-\alpha)x^{-1}(\log x)^{-\alpha-1}(x_2 - x_1)\\ &\in [(-\alpha)x_2^{-1}(\log x_1)^{-\alpha-1}(x_2 - x_1),\ (-\alpha)x_1^{-1}(\log x_2)^{-\alpha-1}(x_2 - x_1)]. \qquad \text{(A.10)}\end{aligned}$$

Similarly, if $\alpha \in (-1, 0)$, then

$$\begin{aligned}(\log x_2)^{-\alpha} &- (\log x_1)^{-\alpha}\\ &\in [(-\alpha)x_2^{-1}(\log x_2)^{-\alpha-1}(x_2 - x_1),\ (-\alpha)x_1^{-1}(\log x_1)^{-\alpha-1}(x_2 - x_1)]; \qquad \text{(A.11)}\end{aligned}$$

and if $\alpha > 0$, then

$$\begin{aligned}(\log x_1)^{-\alpha} &- (\log x_2)^{-\alpha}\\ &\in [\alpha x_2^{-1}(\log x_2)^{-\alpha-1}(x_2 - x_1),\ \alpha x_1^{-1}(\log x_1)^{-\alpha-1}(x_2 - x_1)]. \qquad \text{(A.12)}\end{aligned}$$

For $t \in (0, 1) \setminus I_\varepsilon$ we have $\widetilde{\Phi}_\varepsilon(t) \geqslant \Phi(t)$ and thus $\tilde{R}_\varepsilon(t) \geqslant R(t)$; hence it follows from (A.10)–(A.12) that

$$[\log \tilde{R}_\varepsilon(t)]^{-\alpha} - [\log R(t)]^{-\alpha} \geqslant |\alpha| \frac{1}{\tilde{R}_\varepsilon(t)} [\log R(t)]^{-\alpha-1} \frac{c_1 g_\varepsilon(t)}{t} \qquad \text{if } \alpha \leqslant -1; \qquad \text{(A.13)}$$

$$[\log \tilde{R}_\varepsilon(t)]^{-\alpha} - [\log R(t)]^{-\alpha} \geqslant |\alpha| \frac{1}{\tilde{R}_\varepsilon(t)} [\log \tilde{R}_\varepsilon(t)]^{-\alpha-1} \frac{c_1 g_\varepsilon(t)}{t} \qquad \text{if } \alpha \in (-1, 0); \qquad \text{(A.14)}$$

$$[\log R(t)]^{-\alpha} - [\log \tilde{R}_\varepsilon(t)]^{-\alpha} \geqslant \alpha \frac{1}{\tilde{R}_\varepsilon(t)} [\log \tilde{R}_\varepsilon(t)]^{-\alpha-1} \frac{c_1 g_\varepsilon(t)}{t} \qquad \text{if } \alpha > 0. \qquad \text{(A.15)}$$

Denote the interval I_ε defined at (A.9) by $(a_\varepsilon, b_\varepsilon)$. Consider $t \in I_\varepsilon$ for the next three displays; thus $\Phi(t) > \widetilde{\Phi}_\varepsilon(t)$ and $R(t) > \tilde{R}_\varepsilon(t)$. If $\alpha \leqslant -1$ we have, recalling Lemma 6.1,

$$[\log \tilde{R}_\varepsilon(t)]^{-\alpha} - [\log R(t)]^{-\alpha} \geqslant -|\alpha| \frac{1}{\tilde{R}_\varepsilon(t)} [\log R(t)]^{-\alpha-1} \frac{c_1 |g_\varepsilon(t)|}{t}$$

$$= -|\alpha| \frac{1}{\widetilde{\Phi}_\varepsilon(t)} [\log R(t)]^{-\alpha-1} c_1 |g_\varepsilon(t)|$$
$$\geqslant -|\alpha| \frac{1}{\widetilde{\Phi}_\varepsilon(a_\varepsilon)} [\log R(a_\varepsilon)]^{-\alpha-1} c_1 \varepsilon, \tag{A.16}$$

where we have used (A.7) at the last inequality; similarly, if $\alpha \in (-1, 0)$ we have

$$[\log \tilde{R}_\varepsilon(t)]^{-\alpha} - [\log R(t)]^{-\alpha} \geqslant -|\alpha| \frac{1}{\widetilde{\Phi}_\varepsilon(t)} [\log \tilde{R}_\varepsilon(t)]^{-\alpha-1} c_1 |g_\varepsilon(t)|$$
$$\geqslant -|\alpha| \frac{1}{\widetilde{\Phi}_\varepsilon(a_\varepsilon)} [\log \tilde{R}_\varepsilon(b_\varepsilon)]^{-\alpha-1} c_1 \varepsilon; \tag{A.17}$$

and if $\alpha > 0$ we have

$$[\log R(t)]^{-\alpha} - [\log \tilde{R}_\varepsilon(t)]^{-\alpha} \geqslant -\alpha \frac{1}{\widetilde{\Phi}_\varepsilon(t)} [\log \tilde{R}_\varepsilon(t)]^{-\alpha-1} c_1 |g_\varepsilon(t)|$$
$$\geqslant -\alpha \frac{1}{\widetilde{\Phi}_\varepsilon(a_\varepsilon)} [\log \tilde{R}_\varepsilon(b_\varepsilon)]^{-\alpha-1} c_1 \varepsilon. \tag{A.18}$$

We continue by assessing the contribution to the difference $\operatorname{sgn}(\alpha) \cdot [\mu(\alpha) - \tilde{\mu}_\varepsilon(\alpha)]$ of integrals from $t \in I_\varepsilon$, with asserted inequalities valid for all small $\varepsilon > 0$. If $\alpha \leqslant -1$, the contribution is, using (A.16), bounded below by

$$-|\alpha| c_1 \varepsilon \frac{1}{\widetilde{\Phi}_\varepsilon(a_\varepsilon)} [\log R(a_\varepsilon)]^{-\alpha-1} (b_\varepsilon - a_\varepsilon) \geqslant -|\alpha| C_1 \varepsilon^{3/2} \bigl[\log\bigl(R(1/6) + C_2 \varepsilon^{1/2}\bigr)\bigr]^{-\alpha-1}$$
$$\geqslant -|\alpha| C_1 \varepsilon^{3/2} [\log R(1/7)]^{-\alpha-1}, \tag{A.19}$$

where we have used Lemma 6.1 and where the constants C_1 and C_2 do not depend on α. Similarly, for $\alpha \in (-1, \frac{1}{2})$ the contribution is bounded below by

$$-|\alpha| C_3 \varepsilon^{3/2} [\log R(1/5)]^{-\alpha-1}. \tag{A.20}$$

Next we similarly assess the contribution to $\operatorname{sgn}(\alpha) \cdot [\mu(\alpha) - \tilde{\mu}_\varepsilon(\alpha)]$ from values $t \in (0,1) \setminus I_\varepsilon$. For all small $\varepsilon > 0$, for $\alpha \leqslant -1$ the contribution is, using (A.13), at least

$$\int_{1/9}^{1/8} |\alpha| \frac{1}{\widetilde{\Phi}_\varepsilon(t)} [\log R(t)]^{-\alpha-1} c_1 g_\varepsilon(t) \, \mathrm{d}t \geqslant |\alpha| \, c_2 \, [\log R(1/8)]^{-\alpha-1}, \tag{A.21}$$

where the constant c_2 does not depend on α. Similarly, for $\alpha \in (-1, 1/2)$ the contribution is, using (A.14)–(A.15), at least

$$|\alpha| c_3 \int_{7/24}^{1/3} [\log \tilde{R}_\varepsilon(t)]^{-\alpha-1} \, \mathrm{d}t \geqslant |\alpha| c_4 \, [\log R(1/4)]^{-\alpha-1}. \tag{A.22}$$

Summarizing, for $\alpha \leqslant -1$ we have, using (A.21) and (A.19),

$$\tilde{\mu}_\varepsilon(\alpha) - \mu(\alpha) \geqslant |\alpha| c_2 [\log R(1/8)]^{-\alpha-1} - |\alpha| C_1 \varepsilon^{3/2} [\log R(1/7)]^{-\alpha-1}; \tag{A.23}$$

and for $\alpha \in (-1, \frac{1}{2})$ we have, using (A.22) and (A.20),

$$\operatorname{sgn}(\alpha) \cdot [\mu(\alpha) - \tilde{\mu}_\varepsilon(\alpha)] \geqslant |\alpha| c_4 [\log R(1/4)]^{-\alpha-1} - |\alpha| C_3 \varepsilon^{3/2} [\log R(1/5)]^{-\alpha-1}. \tag{A.24}$$

Since by Lemma 6.1

$$R(1/8) > R(1/7) \quad \text{and} \quad R(1/4) < R(1/5), \tag{A.25}$$

for sufficiently small $\varepsilon \leqslant (\min\{c_2/C_1, c_4/C_3\})^{2/3}$ the desired strict inequalities all follow. Our calculations also demonstrate that

$$\mu' - \tilde{\mu}'_\varepsilon \geqslant c_4[\log R(1/4)]^{-1} - C_3\varepsilon^{3/2}[\log R(1/5)]^{-1}, \tag{A.26}$$

which is (strictly) positive for sufficiently small $\varepsilon \leqslant (c_4/C_3)^{2/3}$. □

Appendix B. Negative moments of affine functions of tree size

The representation (6.3) of $\mu(\alpha)$ as an integral in terms of the offspring probability generating function Φ and the consequent ordering of μ-values exhibited in Theorem 6.3(i) can be extended to treat means of more general functions of the Galton–Watson tree-size. We illustrate this with the following theorem, used in the proof of Theorem 6.5.

Theorem B.1. *For real $\alpha < 0$ and $t > 1$, we have*

$$\mathbb{E}(|\mathcal{T}| - 1 + t)^\alpha = (t-1)^\alpha \int_0^1 \left[1 - \frac{c(\eta; -\alpha, t)}{\Gamma(-\alpha)}\right] \mathrm{d}\eta, \tag{B.1}$$

where $c(\eta; -\alpha, t)$ is the incomplete gamma function value

$$c(\eta; -\alpha, t) = \int_{(t-1)\log R(\eta)}^\infty v^{-\alpha-1}e^{-v}\,\mathrm{d}v. \tag{B.2}$$

Proof. Let $f(s) := (s - 1 + t)^\alpha$. Observe that $s \mapsto f(s)/s$ for $s > 0$ is the Laplace transform of the (strictly) increasing function g mapping $x > 0$ to

$$g(x) := (t-1)^\alpha \left[1 - \frac{\gamma((t-1)x; -\alpha)}{\Gamma(-\alpha)}\right] \in (0, (t-1)^\alpha), \tag{B.3}$$

where here $\gamma(\cdot; -\alpha)$ is the incomplete gamma function

$$\int_\cdot^\infty v^{-\alpha-1}e^{-v}\,\mathrm{d}v. \tag{B.4}$$

Therefore

$$\begin{aligned}
\mathbb{E}(|\mathcal{T}| - 1 + t)^\alpha &= \mathbb{E}\, f(|\mathcal{T}|) \\
&= \sum_{n=1}^\infty \mathbb{P}(|\mathcal{T}| = n) f(n) = \sum_{n=1}^\infty n\,\mathbb{P}(|\mathcal{T}| = n) \int_0^\infty e^{-nx} g(x)\,\mathrm{d}x \\
&= \int_0^\infty g(x) y'(e^{-x}) e^{-x}\,\mathrm{d}x = \int_0^1 g(-\log u) y'(u)\,\mathrm{d}u \\
&= \int_0^1 g\bigl(\log(R(\eta))\bigr)\,\mathrm{d}\eta,
\end{aligned} \tag{B.5}$$

again changing variables by $u = y^{-1}(\eta) = 1/R(\eta)$, and (B.1) follows by (B.3). □

Appendix C. Comparisons allowing infinite offspring moments

Remark 6.7(b) follows quickly from the following theorem concerning Laplace transforms.

Theorem C.1. *Let ξ be a (not necessarily integer-valued) nonnegative random variable with Laplace transform f and moments*

$$m_j := \mathbb{E}\,\xi^j \leqslant \infty, \quad j = 0, 1, 2, \dots \tag{C.1}$$

(with $m_0 := 0$). For a given positive integer r, suppose that $m_{r-1} < \infty$. Then

$$g(t) := (-1)^r t^{-r} r! \left[f(t) - \sum_{j=0}^{r-1} (-1)^j m_j \frac{t^j}{j!} \right] \tag{C.2}$$

is nonnegative for $t > 0$ and increases (weakly) to $m_r \leqslant \infty$ as $t \searrow 0$.

We will prove Theorem C.1 using the following calculus lemma.

Lemma C.2. *Let r be a fixed positive integer, and define*

$$h(x) := (-1)^r x^{-r} \left[e^{-x} - \sum_{j=0}^{r-1} (-1)^j \frac{x^j}{j!} \right], \quad x > 0. \tag{C.3}$$

Then h is (strictly) positive and (strictly) decreasing, with limit $1/r!$ as $x \searrow 0$.

Proof. The lemma is immediate from the claim that

$$h(x) = \frac{1}{(r-1)!} \int_0^1 v^{r-1} e^{-x(1-v)}\,\mathrm{d}v. \tag{C.4}$$

We offer two proofs of this claim.

Proof #1 of (C.4). By Taylor's theorem with remainder in integral form,

$$h(x) = \frac{x^{-r}}{(r-1)!} \int_0^x (x-u)^{r-1} e^{-u}\,\mathrm{d}u. \tag{C.5}$$

Now simply change the variable of integration from u to $v = 1 - \frac{u}{x}$.

Proof #2 of (C.4). Let B denote Euler's beta function. Then the right side of (C.4) equals

$$\begin{aligned} \frac{1}{(r-1)!} \sum_{k=0}^{\infty} (-1)^k \frac{x^k}{k!} \int_0^1 v^{r-1} (1-v)^k\,\mathrm{d}v &= \frac{1}{(r-1)!} \sum_{k=0}^{\infty} (-1)^k \frac{x^k}{k!} B(r, k+1) \\ &= (-1)^r x^{-r} \sum_{j=r}^{\infty} (-1)^j \frac{x^j}{j!} = h(x). \end{aligned} \tag{C.6}$$

□

Proof of Theorem C.1. For $t > 0$ we have

$$\begin{aligned} g(t) &= \mathbb{E}\left[(-1)^r t^{-r} r! \left(e^{-t\xi} - \sum_{j=0}^{r-1} (-1)^j \frac{(t\xi)^j}{j!}\right)\right] \\ &= \mathbb{E}\left[(-1)^r t^{-r} r! \left(e^{-t\xi} - \sum_{j=0}^{r-1} (-1)^j \frac{(t\xi)^j}{j!}\right); \xi > 0\right] \\ &= \mathbb{E}\left[r!\, h(t\xi)\, \xi^r;\ \xi > 0\right]. \end{aligned} \tag{C.7}$$

By Lemma C.2, the nonnegative random variables $r!\, h(t\xi)\, \xi^r \mathbf{1}(\xi > 0)$ increase (weakly) to $\xi^r \mathbf{1}(\xi > 0) = \xi^r$ as $t \searrow 0$. Thus, by the MCT, $g(t) \nearrow m_r \leqslant \infty$ as $t \searrow 0$. □

Appendix D. Corrigendum to [5]

As said in Remark 1.9, there is a typo in [5, Theorem D.1]; the variance given in (D.2) there is incorrect and should be

$$\mathbb{E}\,|\zeta|^2 = \frac{1}{2\sqrt{\pi}} \operatorname{Re} \frac{\Gamma(\mathrm{i}t - \frac{1}{2})}{\Gamma(\mathrm{i}t)},$$

as stated in (1.14).

The formula (D.2) in [5] has, incorrectly, $\Gamma(\mathrm{i}t - 1)$ in the denominator, which comes from (D.5) which has the same error. Formula (D.8) in the proof is correct, with denominator $\Gamma(\mathrm{i}t)$, and yields (D.5) and (D.2) with the same denominator, i.e., (1.14).

Theorem D.1 in [5] also claims that $\mathbb{E}\,|\zeta|^2 > 0$. The proof is based on the incorrect formula given there, but luckily the same proof applies also to the correct formula. In (D.14) we obtain $\Gamma(1 - \mathrm{i}t)$ instead of $\Gamma(2 - \mathrm{i}t)$ (and an immaterial change of sign); hence we have to show that $\Gamma(1 - \mathrm{i}t)/\Gamma(\frac{3}{2} - \mathrm{i}t)$ is not real for $t \neq 0$. Thus, in (D.15), we should have $-\operatorname{Im} \int_1^{3/2} \psi(s - \mathrm{i}t)\, \mathrm{d}s$. We use (D.18) as before, and now see that if $t < 0$, then $0 > \arg\bigl(\Gamma(1 - \mathrm{i}t)/\Gamma(\frac{3}{2} - \mathrm{i}t)\bigr) > -\pi/4$, which completes the proof that the variance in (1.14) (and the display above) is nonzero.

References

[1] Sergio Caracciolo, Vittorio Erba, & Andrea Sportiello. The p-Airy distribution. Preprint, 2020. `arXiv:2010.14468v1`

[2] William Feller. *An Introduction to Probability Theory and its Applications, Volume II.* 2nd ed., Wiley, New York, 1971.

[3] James Allen Fill. On the distribution of binary search trees under the random permutation model. *Random Structures Algorithms* **8** (1996), no. 1, 1–25.

[4] James Allen Fill, Philippe Flajolet, & Nevin Kapur. Singularity analysis, Hadamard products, and tree recurrences. *J. Comput. Appl. Math.* **174** (2005), no. 2, 271–313.

[5] James Allen Fill & Svante Janson. The sum of powers of subtree sizes for conditioned Galton-Watson trees. *Electron. J. Probab.* **27** (2022), Paper No. 114, 77 pp.

[6] James Allen Fill & Nevin Kapur. An invariance principle for simply generated families of trees. Unpublished manuscript, 2003.
[7] James Allen Fill & Nevin Kapur. Limiting distributions for additive functionals on Catalan trees. *Theoret. Comput. Sci.* **326** (2004), no. 1-3, 69–102.
[8] Philippe Flajolet. Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.* **215** (1999), no. 1-2, 371–381.
[9] Philippe Flajolet & Robert Sedgewick. *Analytic Combinatorics*. Cambridge Univ. Press, Cambridge, UK, 2009.
[10] Svante Janson. *Gaussian Hilbert Spaces*. Cambridge Univ. Press, Cambridge, UK, 1997.
[11] Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms* **29** (2006), no. 2, 139–179.
[12] Svante Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probability Surveys* **9** (2012), 103–252.
[13] Svante Janson. Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton–Watson trees. *Random Struct. Alg.* **48** (2016), no. 1, 57–101.
[14] Valentin F. Kolchin. *Random Mappings*. Nauka, Moscow, 1984 (Russian). English transl.: Optimization Software, New York, 1986.
[15] A. Meir & J. W. Moon. On the log-product of the subtree-sizes of random trees. *Random Structures Algorithms* **12** (1998), no. 2, 197–212.
[16] *NIST Handbook of Mathematical Functions*. Edited by Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert & Charles W. Clark. Cambridge Univ. Press, 2010.
Also available as *NIST Digital Library of Mathematical Functions*, http://dlmf.nist.gov/
[17] Richard Otter. The multiplicative process. *Ann. Math. Statistics* **20** (1949), 206–224.
[18] Boris Pittel. Normal convergence problem? Two moments and a recurrence may be the clues. *Ann. Appl. Probab.* **9** (1999), no. 4, 1260–1302.

Department of Applied Mathematics and Statistics, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218-2682 USA
Email address: jimfill@jhu.edu
URL: http://www.ams.jhu.edu/~fill/

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
Email address: svante.janson@math.uu.se
URL: http://www2.math.uu.se/~svante/

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
Email address: stephan.wagner@math.uu.se