GENERAL EDGEWORTH EXPANSIONS WITH APPLICATIONS TO PROFILES OF RANDOM TREES: EXTENDED VERSION

ZAKHAR KABLUCHKO, ALEXANDER MARYNYCH, AND HENNING SULZBACH

Abstract. We prove an asymptotic Edgeworth expansion for the profiles of certain random trees including binary search trees, random recursive trees and plane-oriented random trees, as the size of the tree goes to infinity. All these models can be seen as special cases of the one-split branching random walk for which we also provide an Edgeworth expansion. These expansions lead to new results on mode, width and occupation numbers of the trees, settling several open problems raised in Devroye and Hwang [*Ann. Appl. Probab.* 16(2): 886–918, 2006], Fuchs, Hwang and Neininger [*Algorithmica*, 46 (3–4): 367–407, 2006], and Drmota and Hwang [*Adv. in Appl. Probab.*, 37 (2): 321–341, 2005]. The aforementioned results are special cases and corollaries of a general theorem: an Edgeworth expansion for an arbitrary sequence of random or deterministic functions $\mathbb{L}_n : \mathbb{Z} \to \mathbb{R}$ which converges in the mod-ϕ-sense. Applications to Stirling numbers of the first kind will be given in a separate paper.

1. Introduction

1.1. **Introduction.** The aim of this paper is to study asymptotic properties of profiles for certain families of random trees when the size of the tree goes to infinity. The profile is the function $k \mapsto \mathbb{L}_n(k)$, where $\mathbb{L}_n(k)$ is the number of nodes at depth k in the tree of size n. These numbers are also called "occupation numbers". We shall mainly be interested in the following families of random trees:

(1) binary search trees (BSTs) and, more generally, D-ary recursive trees;
(2) random recursive trees (RRTs);
(3) plane-oriented recursive trees (PORTs) and, more generally, p-oriented trees.

BSTs, RRTs and PORTs have been much studied in the literature; see [7, 9, 13, 35, 20, 14, 16]. Mahmoud's book [32] and Drmota's monograph [15] contain further pointers to relevant literature. It is well known that BSTs are intimately connected to the Quicksort algorithm.

Our main result is an asymptotic expansion for the profile which is somewhat similar to the classical Chebyshev–Edgeworth–Cramér expansion for sums of independent identically distributed (i.i.d.) integer-valued random variables. As a consequence of our asymptotic expansion, we derive limit theorems for several functionals

2010 *Mathematics Subject Classification.* Primary, 60G50; secondary, 60F05, 60J80, 60J85, 60F10, 60F15.

Key words and phrases. Biggins martingale, branching random walk, central limit theorem, Edgeworth expansion, mod-ϕ-convergence, mode, profile, random analytic function, random tree, width.

of the profile such as the mode, the width, and the occupation numbers, thus answering a number of open questions on these quantities. Many known results on the profiles such as the (local) central limit theorem or limit theorems for occupation numbers on the scale of large deviations can be recovered in a unified way as corollaries of our expansion.

The scope of our method is by no means restricted to random trees. In Section 2 we shall state and prove a very general asymptotic expansion (Theorem 2.1) which can be applied to any sequence of random or deterministic functions $\mathbb{L}_n : \mathbb{Z} \to \mathbb{R}$, $n \in \mathbb{N}$, provided certain natural conditions are satisfied. Recently, a closely related expansion was derived by Féray et al. [19] in the framework of the mod-ϕ-convergence. It has been observed by Nikeghbali and collaborators that mod-ϕ-convergence, see Remark 2.9 for the definition, is a common phenomenon in probability, combinatorics, number theory and statistical mechanics; see [11, 19, 25, 29, 30, 33]. In this work, we show how mod-ϕ-convergence[1] can be applied to the analysis of random trees of logarithmic height.

The paper is organized as follows. The general asymptotic expansion is stated in Section 2. In Section 3 we apply this expansion to the profile of the *one-split branching random walk*, a model which contains profiles of all random trees listed above as special cases. Since these results are quite general and require heavy notation, we motivate and prepare the reader in the next Section 1.2 by formulating the results in the special case of binary search trees. In Section 3.2 we shall explain how to formulate similar results for other random trees including RRTs and PORTs. Proofs are given in Sections 4 and 5.

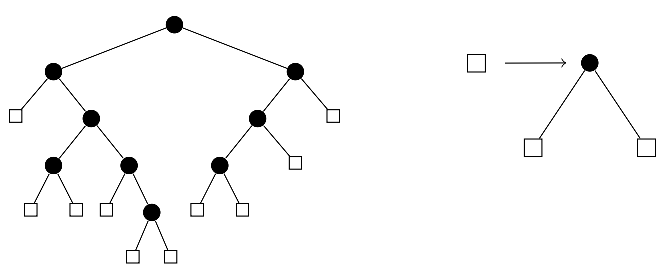


FIGURE 1. Left: Sample realization of the BST. Right: Construction rule for the BST

1.2. **Results for binary search trees.** For our purposes, the following construction of BSTs is most convenient. There are nodes of two types: the external ones (denoted by $\square$) and the internal ones (denoted by $\bullet$); see Figure 1 (left). At time $n = 0$ start with one external node (the root of the tree). At any step of the construction, pick one of the existing external nodes uniformly at random, declare it to be internal, and connect it to two new external nodes according to the rule shown on Figure 1 (right). After n steps, one obtains a random tree T_n having n internal and $n+1$ external nodes; see Figure 1 (left) for a sample realization.

The *depth* of an external node is its distance to the root. Denote by $\mathbb{L}_n(k)$ the number of external nodes of T_n at depth $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and let $\mathbb{L}_n(k) = 0$ for $k < 0$. The random function $k \mapsto \mathbb{L}_n(k)$, $k \in \mathbb{Z}$, is called the (external) *profile* of

[1] In all examples mentioned in Section 1.1 it is, in fact, mod-Poisson convergence.

the tree T_n. Denote by $x_{1,n},\dots,x_{n+1,n}$ the depths of the (arbitrarily enumerated) external nodes of T_n so that

$$\mathbb{L}_n(k)=\#\{1\le i\le n+1\colon x_{i,n}=k\},\quad k\in\mathbb{Z}.$$

The BST profile has been much studied; see [7, 9, 14, 16, 17, 20], and [15, Section 6.5] for a survey. In the following, we provide a short overview of known results. Let $\beta_-\approx -1.678$ and $\beta_+\approx 0.768$ be the solutions to the equation $2\mathrm{e}^{\beta}(1-\beta)=1$. The numbers $2\mathrm{e}^{\beta_-}\approx 0.373$ and $2\mathrm{e}^{\beta_+}\approx 4.311$ are called the *fill-up level constant* and the *height constant* of the BST because of the classical results

$$\frac{1}{\log n}\min_{i=1,\dots,n+1} x_{i,n}\xrightarrow[n\to\infty]{a.s.}2\mathrm{e}^{\beta_-},\quad \frac{1}{\log n}\max_{i=1,\dots,n+1} x_{i,n}\xrightarrow[n\to\infty]{a.s.}2\mathrm{e}^{\beta_+},$$

going back to Devroye [12], see also [5]. Define the following random function

$$W_n(\beta)=\frac{1}{n^{(2\mathrm{e}^{\beta}-1)}}\sum_{i=1}^{n+1}\mathrm{e}^{\beta x_{i,n}},\quad \beta\in\mathbb{C},$$

which is the normalized moment generating function of the random counting measure $\sum_{k\in\mathbb{Z}}\mathbb{L}_n(k)\delta_k=\sum_{i=1}^{n+1}\delta_{x_{i,n}}$ where δ_x denotes a Dirac measure at point x. The basic fact underlying all further arguments is that W_n converges as $n\to\infty$ to a random analytic function W_∞ with probability 1. More precisely, it is known [9] that there is an open set $\mathscr{D}\subset\mathbb{C}$ containing the interval (β_-,β_+) and a random analytic function W_∞ defined on $\mathscr{D}$ such that

$$\sup_{\beta\in K}|W_n(\beta)-W_\infty(\beta)|\xrightarrow[n\to\infty]{a.s.}0$$

for every compact set $K\subset\mathscr{D}$. Note that $W_\infty(0)=1$ since $W_n(0)=(n+1)/n$ for all $n\in\mathbb{N}$.

It is useful to keep in mind the following principle: $k\mapsto\frac{1}{n}\mathbb{L}_n(k)$ is "close" to the probability mass function of the Poisson distribution with intensity $2\log n$. The moment generating function of the latter distribution is $\beta\mapsto n^{2\mathrm{e}^{\beta}-2}$, and the general philosophy of mod-ϕ-convergence [25, 29] suggests to view the limit function W_∞ as a quantification of the "difference" between $\frac{1}{n}\mathbb{L}_n$ and the Poisson distribution with parameter $2\log n$. Note that in our case, this function is random.

An important role will be played by the derivatives and logarithmic derivatives of W_∞ (the latter are called random cumulants). In particular, we shall frequently encounter the random variables

$$\chi_1(0):=(\log W_\infty)'(0)=W_\infty'(0),\quad \chi_2(0):=(\log W_\infty)''(0)=W_\infty''(0)-(W_\infty'(0))^2.$$

It will be shown in Section 3.4 that they can also be represented as the a.s. limits

$$\chi_1(0)=\lim_{n\to\infty}(\log W_n)'(0)=\lim_{n\to\infty}\left(\frac{1}{n+1}\sum_{i=1}^{n+1}x_{i,n}-2\log n\right),\tag{1}$$

$$\begin{aligned}\chi_2(0)&=\lim_{n\to\infty}(\log W_n)''(0)\\&=\lim_{n\to\infty}\left(\frac{1}{n+1}\sum_{i=1}^{n+1}x_{i,n}^2-\left(\frac{1}{n+1}\sum_{i=1}^{n+1}x_{i,n}\right)^2-2\log n\right).\end{aligned}\tag{2}$$

It is useful to think of $\chi_1(0)$ (whose distribution is the celebrated Quicksort law [36, 37], modulo an additive constant) as a parameter describing the random shift of the BST profile with respect to (w.r.t.) the location $2\log n$. The random variable

$\chi_2(0)$ describes the random deviation of the empirical variance of the profile from the value $2\log n$ and seems to appear for the first time.

We now proceed to the asymptotic properties of the BST profile $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$, as $n\to\infty$. The following (local) central limit theorem was proved by Chauvin et al. [7]:

$$\sup_{k\in\mathbb{Z}}\left|\frac{1}{n}\mathbb{L}_n(k)-\frac{1}{\sqrt{4\pi\log n}}\mathrm{e}^{-\frac{(k-2\log n)^2}{4\log n}}\right| = O\left(\frac{1}{\log n}\right) \quad \text{a.s.} \tag{3}$$

As a special case of our results, we shall derive an asymptotic expansion complementing (3).

Theorem 1.1. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the external profile of a binary search tree with $n+1$ external nodes. For every $r\in\mathbb{N}_0$, we have*

$$(\log n)^{\frac{r+1}{2}}\sup_{k\in\mathbb{Z}}\left|\frac{1}{n}\mathbb{L}_n(k)-\frac{\mathrm{e}^{-\frac{(k-2\log n)^2}{4\log n}}}{\sqrt{4\pi\log n}}\sum_{j=0}^{r}G_j\left(\frac{k-2\log n}{\sqrt{2\log n}};0\right)\frac{1}{(\log n)^{j/2}}\right|\xrightarrow[n\to\infty]{a.s.}0,$$

where $G_j(x;0)$ is a polynomial in x of degree at most $3j$ whose coefficients can be linearly expressed through the derivatives $W'_\infty(0),\dots,W_\infty^{(j)}(0)$. For example,

$$G_0(x;0)=1,\quad G_1(x;0)=\frac{1}{\sqrt{2}}\left(W'_\infty(0)x+\frac{x^3-3x}{6}\right);$$

see (51), (52) below for an explicit general formula.

The above results deal with the profile around level $2\log n$, where "most" external nodes are located. The shape of the profile at levels around $c\log n$, $2\mathrm{e}^{\beta_-}<c<2\mathrm{e}^{\beta_+}$, is described by the following result due to Chauvin et al. [9], compare also [7, 17, 24] for weaker formulations and [20] for pointwise convergence theorems:

$$\sup_{k\in\mathbb{Z}\cap(\log n)L}\left|\frac{nk!}{(2\log n)^k}\mathbb{L}_n(k)-W_\infty\left(\log\left(\frac{k}{2\log n}\right)\right)\right|\xrightarrow[n\to\infty]{a.s.}0 \tag{4}$$

for every compact set $L\subset(2\mathrm{e}^{\beta_-},2\mathrm{e}^{\beta_+})$. We can derive an asymptotic expansion complementing (4).

Theorem 1.2. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the external profile of a binary search tree with $n+1$ external nodes. For every $r\in\mathbb{N}_0$ and every compact set $L\subset(2\mathrm{e}^{\beta_-},2\mathrm{e}^{\beta_+})$ we have*

$$(\log n)^{r+1}\sup_{k\in\mathbb{Z}\cap(\log n)L}\left|n\left(\frac{k}{2\mathrm{e}\log n}\right)^k\mathbb{L}_n(k)-\frac{1}{\sqrt{2\pi k}}\sum_{j=0}^{r}\frac{F_{2j}(0;\beta_n(k))}{(\log n)^j}\right|\xrightarrow[n\to\infty]{a.s.}0,$$

where $\beta_n(k)$ is the solution[2] of $2\mathrm{e}^{\beta_n(k)}=k/\log n$ and $F_{2j}(0;\beta):=W_\infty(\beta)G_{2j}(0;\beta)$ is a linear combination of $1,W_\infty(\beta),\dots,W_\infty^{(2j)}(\beta)$; see (51), (52) below for an explicit formula. For example,

$$F_0(0;\beta)=W_\infty(\beta),\quad F_2(0;\beta)=\frac{1}{4}(W'_\infty(\beta)-W''_\infty(\beta))-\frac{1}{24}.$$

[2] Obviously, we have $\beta_n(k)=\log k-\log\log n-\log 2$, however we prefer to define $\beta_n(k)$ implicitly to conform with the definition of $\beta_n(k)$ in general case; see formula (26) below.

Remark 1.3. Theorems 1.1 and 1.2 are special cases (with $\beta = 0$ and $\beta = \beta_n(k)$, respectively) of Theorem 3.11 (see also Theorem 3.15) which deals with general one-split branching random walks and which we shall state and prove below. Similar results can easily be obtained for RRTs and PORTs; see Section 3.2 for details.

Remark 1.4. Our techniques yield analogous expansions for the mean profile: Theorems 1.1 and 1.2 remain valid upon replacing $\mathbb{L}_n(k)$ by $\mathbb{E}[\mathbb{L}_n(k)]$ and the random polynomials G_j, F_{2j} by their expectations.

The above expansions can be used to answer a number of open questions on the BST profile. The *mode* u_n and the *width* M_n of a binary search tree are defined by

$$u_n = \arg\max_{k\in\mathbb{N}_0} \mathbb{L}_n(k), \quad M_n = \max_{k\in\mathbb{N}_0} \mathbb{L}_n(k).$$

These quantities were studied by Chauvin et al. [7], Drmota and Hwang [16] and Devroye and Hwang [14]. In particular, Devroye and Hwang [14, Theorem 4.1] proved that u_n is concentrated near $2\log n$ in the sense that for every $B > 0$ there is $C_0 = C_0(B)$ such that

$$\mathbb{P}[|u_n - 2\log n| > T] \le C_0 T^{-B}, \quad n \in \mathbb{N}, \quad T \ge 1.$$

We show that, starting from some random almost surely finite time K, the mode u_n attains only one of two possible explicitly given values. This result is a special case of Theorem 3.17.

Theorem 1.5. *There is an a.s. finite random variable K such that for $n > K$, the mode u_n of the BST with $n+1$ external nodes is equal to one of the numbers*

$$\lfloor 2\log n + \chi_1(0) - 1/2 \rfloor \ \text{or}\ \lceil 2\log n + \chi_1(0) - 1/2 \rceil,$$

where $\lfloor\cdot\rfloor$, $\lceil\cdot\rceil$ denote the floor and the ceiling functions, respectively, and $\chi_1(0)$ is the Quicksort-distributed random variable defined in (1).

Remark 1.6. Using Proposition 3.19 we can say more about the behavior of the mode. More precisely, the following statements hold with probability 1:

- there are arbitrarily long intervals of consecutive n's for which the mode u_n is unique and $u_n = \lceil 2\log n + \chi_1(0) - 1/2 \rceil$;
- similarly, there are arbitrarily long intervals of consecutive n's for which u_n is unique and $u_n = \lfloor 2\log n + \chi_1(0) - 1/2 \rfloor$;
- the set of $n \in \mathbb{N}$ such that u_n is the integer closest to $2\log n + \chi_1(0) - 1/2$ has asymptotic density one (see (55) for the definition of the asymptotic density).

As for the BST's width M_n, it is known [7, 14, 16] that

$$\mathbb{E}[M_n] = \frac{n}{\sqrt{4\pi\log n}}\left(1 + O\left(\frac{1}{\log n}\right)\right), \quad \frac{M_n\sqrt{4\pi\log n}}{n} \xrightarrow[n\to\infty]{a.s.} 1.$$

Both Devroye and Hwang [14] as well as Drmota and Hwang [16, Section 5] asked for the limit distribution of M_n (if it exists). The next result (which holds a.s. rather than in distribution) settles this issue, and is a consequence of Theorem 3.21 and the remark following it.

Theorem 1.7. *Let M_n be the width of a binary search tree with $n+1$ external nodes. With probability* 1*, the set of subsequential limits of the sequence*

$$\tilde{M}_n := 4\log n\left(1 - \frac{\sqrt{4\pi\log n}\, M_n}{n}\right), \quad n \in \mathbb{N},$$

is the interval $[\chi_2(0)-1/12, \chi_2(0)+1/6]$ *with* $\chi_2(0)$ *as in* (2). *Furthermore, with* $\theta_n = \min_{k\in\mathbb{Z}} |2\log n + \chi_1(0) - 1/2 - k|$ *we have*

$$\tilde{M}_n - \theta_n^2 \xrightarrow[n\to\infty]{a.s.} \chi_2(0) - \frac{1}{12}. \tag{5}$$

Remark 1.8. Let us stress that the centering θ_n^2 in (5) is random since it involves $\chi_1(0)$. In order to obtain a non-random centering, we have to pass to a subsequence. If $(n_j)_{j\in\mathbb{N}} \subset \mathbb{N}$ is any sequence with $n_j \to +\infty$ and $\{2\log n_j\} \to \alpha \in [0,1]$ as $j\to\infty$ (where $\{\cdot\}$ denotes the fractional part), then $\lim_{j\to\infty}\theta_{n_j} = |\{\alpha+\chi_1(0)\}-1/2|$ and we obtain

$$\tilde{M}_{n_j} \xrightarrow[j\to\infty]{a.s.} \chi_2(0) - \frac{1}{12} + \left(\{\alpha+\chi_1(0)\} - \frac{1}{2}\right)^2.$$

Since the set of accumulation points of the sequence $(\{2\log n\})_{n\in\mathbb{N}}$ is the interval $[0,1]$, we obtain for $\tilde{M}_n$ a family of subsequential limit distributions indexed by $\alpha\in[0,1]$ with values $\alpha=0$ and $\alpha=1$ corresponding to the same limit.

In the next theorem we describe the asymptotic behavior of the "occupation numbers" $\mathbb{L}_n(k_n)$, where $(k_n)_{n\in\mathbb{N}}$ is a deterministic sequence with sufficiently regular behavior. These quantities were the main object of study in Fuchs et al. [20]; see also [7, 9] and (for results on lattice branching random walks) [10, 21, 26]. It is known [9, 20] (and not difficult to deduce from (4)) that if $k_n = 2e^\beta \log n + \alpha\sqrt{2e^\beta\log n} + o(\sqrt{\log n})$ for some $\beta\in(\beta_-,\beta_+)$ and $\alpha\in\mathbb{R}$, then

$$\frac{\sqrt{2e^\beta \log n}}{n^{2e^\beta - 1}} e^{\beta k_n} \mathbb{L}_n(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{W_\infty(\beta)}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha^2}. \tag{6}$$

Furthermore, the convergence of moments of order κ, for any $\kappa\in(1,2)\cup\{2,3,\dots\}$ with $\mathbb{E}[W_\infty^\kappa(\beta)]<\infty$, was proved in [20]. Another consequence of (4) is that for $k_n = 2e^\beta\log n + c_n$, where $\beta\in(\beta_-,\beta_+)$, $c_n = o(\log n)$ we have

$$\frac{\sqrt{2e^\beta\log n}}{e^{c_n} n^{2e^\beta-1}} \left(\frac{k_n}{2\log n}\right)^{k_n} \mathbb{L}_n(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{W_\infty(\beta)}{\sqrt{2\pi}}. \tag{7}$$

For $\beta=0$, the limit random variable in (6), (7) is degenerate because $W_\infty(0)=1$, and a more refined analysis is needed to obtain a non-degenerate limit law. It turns out that all such results hold also in the almost sure sense.

Theorem 1.9. *Let* $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ *be the external profile of a BST with* $n+1$ *external nodes. Put* $\mathbb{L}_n^\circ(k) := \mathbb{L}_n(k) - \mathbb{E}[\mathbb{L}_n(k)]$ *and let* $(k_n)_{n\in\mathbb{N}}$ *be a deterministic integer sequence.*

(a) *If* $k_n = 2\log n + \alpha\sqrt{2\log n} + o(\sqrt{\log n})$ *for some* $\alpha\in\mathbb{R}$, *then*

$$\frac{\log n}{n}\mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\chi_1(0)-\mathbb{E}[\chi_1(0)]}{2\sqrt{2\pi}} \alpha e^{-\frac{1}{2}\alpha^2}.$$

(b) *If* $k_n = 2\log n + c_n$, *where* $c_n = o(\log n)$ *and* $\lim_{n\to\infty}|c_n| = \infty$, *then*

$$\frac{(\log n)^{3/2}}{n c_n e^{c_n}} \left(\frac{k_n}{2\log n}\right)^{k_n} \mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\chi_1(0)-\mathbb{E}[\chi_1(0)]}{4\sqrt{\pi}}.$$

In particular, if $c_n = o(\sqrt{\log n})$ *and* $\lim_{n\to\infty}|c_n|=\infty$, *then*

$$\frac{(\log n)^{3/2}}{n c_n}\mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\chi_1(0)-\mathbb{E}[\chi_1(0)]}{4\sqrt{\pi}}.$$

(c) *If* $k_n = 2\log n + c_n$*, where* c_n *is bounded, then*

$$\frac{(\log n)^{3/2}}{n}\mathbb{L}_n^\circ(k_n) - \frac{\chi_1(0)-\mathbb{E}[\chi_1(0)]}{4\sqrt{\pi}}\left(c_n+\frac{1}{2}\right) \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} -\frac{W''_\infty(0)-\mathbb{E}[W''_\infty(0)]}{8\sqrt{\pi}}.$$

Remark 1.10. More specifically, if in case (c) we have $k_n = \lfloor 2\log n\rfloor + a$ for $a\in\mathbb{Z}$, then the set of subsequential limits of the sequence $\left(\frac{1}{n}(\log n)^{3/2}\mathbb{L}_n^\circ(k_n)\right)_{n\in\mathbb{N}}$ equals, with probability 1, the closed interval

$$\left\{\frac{\chi_1(0)-\mathbb{E}[\chi_1(0)]}{4\sqrt{\pi}}(a+y) - \frac{W''_\infty(0)-\mathbb{E}[W''_\infty(0)]}{8\sqrt{\pi}} :\ -\frac{1}{2}\le y\le\frac{1}{2}\right\}.$$

A subsequential limit of this form is attained along any subsequence $(n_j)_{n\in\mathbb{N}}$ with $\{2\log n_j\}\to\frac{1}{2}-y$ as $j\to\infty$.

Remark 1.11. Theorem 1.9 and Equations (6), (7) are special cases of more general Theorems 3.23, 3.25 which deal with one-split branching random walks. As will be explained in Section 3.2, analogous results can be obtained for RRTs and PORTs simply by inserting suitable parameters into the theorems listed.

In cases (a) and (b), distributional convergence (and, in fact, convergence of all moments) was proved by Fuchs et al. [20]. Our approach (which is very different from the method of moments and the contraction method used in [20]) yields a.s. convergence. In case (c), Fuchs et al. [20] showed that $\mathbb{L}_n(k_n)$, centered by its expectation and normalized by its standard deviation, has no non-degenerate limit law. Our result identifies all possible weak (and, in fact, even a.s.) subsequential limits of the appropriately normalized $\mathbb{L}_n(k_n)$. One may also ask for multivariate limit laws for the occupation numbers. For example, in case (c) it is natural to investigate the joint limit distribution of the random vector $(\mathbb{L}_n(\lfloor 2\log n\rfloor + a))_{a=-K,\dots,K}$ where $K\in\mathbb{N}_0$ is fixed. Since our results are a.s., they automatically yield such multivariate limit theorems, whereas the moment method and the contraction method seem less convenient to treat such multivariate problems. Finally, let us mention that there is one more case in which $W_\infty(\beta)$ is a.s. constant, namely $\beta = -\log 2$; see Section 3.8 for a detailed analysis of this case.

2. The general Edgeworth expansion

2.1. **Assumptions on the profiles.** Consider a sequence $\mathbb{L}_1, \mathbb{L}_2, \dots$ such that each $\mathbb{L}_n = (\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ is a real-valued stochastic process defined on the integer lattice $\mathbb{Z}$. We assume that all $\mathbb{L}_1, \mathbb{L}_2, \dots$ are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We shall consider the random function

$$\mathbb{L}_n : \mathbb{Z}\to\mathbb{R},\quad k\mapsto \mathbb{L}_n(k)$$

as a "*random profile*". As has already been mentioned, in our applications to random trees, $\mathbb{L}_n(k)$ is the number of nodes of depth k in a random tree at time n. Our aim is to prove that under appropriate assumptions, $\mathbb{L}_n$ satisfies an Edgeworth-type asymptotic expansion with probability 1. Let us state these assumptions.

Assumption A1: There is an open, non-empty interval $(\beta_-, \beta_+)\subset\mathbb{R}$ containing 0 such that for every $n\in\mathbb{N}$ and every $\beta\in(\beta_-,\beta_+)$,

$$\sum_{k\in\mathbb{Z}} |\mathbb{L}_n(k)|\mathrm{e}^{\beta k} < \infty \quad \text{a.s.} \tag{8}$$

The interval (β_-, β_+) need not be bounded. For example, Assumption A1 is satisfied on whole $\mathbb{R}$ if for every $n \in \mathbb{N}$ the profile support $\{k \in \mathbb{Z} \colon \mathbb{L}_n(k) \neq 0\}$ is a finite set with probability 1.

The next assumption essentially states that the Laplace transform of the profile given by

$$\beta \mapsto \sum_{k\in\mathbb{Z}} \mathbb{L}_n(k) \mathrm{e}^{\beta k}$$

converges, after an appropriate normalization, to a random analytic function on some domain $\mathscr{D}$ in the complex plane which contains the interval (β_-, β_+). To state this assumption we need the following ingredients:

- a sequence $(w_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ such that $\lim_{n\to\infty} w_n = +\infty$;
- an open, connected set $\mathscr{D} \subset \{\beta \in \mathbb{C} \colon \beta_- < \operatorname{Re}\beta < \beta_+\}$ such that $\mathscr{D} \cap \mathbb{R} = (\beta_-, \beta_+)$;
- a (deterministic) analytic function $\varphi : \mathscr{D} \to \mathbb{C}$ such that, for real $\beta \in (\beta_-, \beta_+)$, we have $\varphi(\beta) \in \mathbb{R}$ and $\varphi''(\beta) > 0$.

It follows from Assumption A1 that, with probability 1, the normalized Laplace transform

$$W_n(\beta) := \mathrm{e}^{-\varphi(\beta) w_n} \sum_{k\in\mathbb{Z}} \mathbb{L}_n(k) \mathrm{e}^{\beta k}, \quad \beta \in \mathscr{D}, \tag{9}$$

is a random analytic function on $\mathscr{D}$.

Assumption A2: With probability 1, the sequence of random analytic functions $(W_n)_{n\in\mathbb{N}}$ converges locally uniformly on $\mathscr{D}$, as $n \to \infty$, to some random analytic function W_∞ such that $\mathbb{P}[W(\beta) \neq 0 \text{ for all } \beta \in (\beta_-, \beta_+)] = 1$.

Moreover, we require that the speed of convergence is superpolynomial in w_n.

Assumption A3: For every compact set $K \subset \mathscr{D}$ and $r \in \mathbb{N}$ we can find an a.s. finite random variable $C_{K,r}$ such that for all $n \in \mathbb{N}$,

$$\sup_{\beta\in K} |W_n(\beta) - W_\infty(\beta)| < C_{K,r} w_n^{-r}. \tag{10}$$

The last assumption is of technical nature. In the classical Edgeworth expansion for sums of i.i.d. integer-valued variables, it corresponds to the assumption that $\mathbb{Z}$ is the minimal lattice on which the distribution is concentrated.

Assumption A4: For every compact set $K \subset (\beta_-, \beta_+)$, every $a > 0$ and $r \in \mathbb{N}_0$, we have

$$\sup_{\beta\in K} \left[\mathrm{e}^{-\varphi(\beta) w_n} \int_a^\pi \left| \sum_{k\in\mathbb{Z}} \mathbb{L}_n(k) \mathrm{e}^{k(\beta+iu)} \right| \mathrm{d}u \right] = o(w_n^{-r}) \quad \text{a.s. as } n \to \infty. \tag{11}$$

2.2. **Statement of the general Edgeworth expansion.** Consider a sequence of profiles $\mathbb{L}_1, \mathbb{L}_2, \dots$ satisfying Assumptions A1–A4. We are going to state an Edgeworth expansion for $\mathbb{L}_n$ as $n \to \infty$. In fact, we shall obtain an expansion of the "tilted" profile $k \mapsto \mathrm{e}^{\beta k - \varphi(\beta) w_n} \mathbb{L}_n(k)$ which is uniform as long as β stays in a certain range.

We shall see that the following parameters $\mu(\beta)$ and $\sigma(\beta)$ play the role of the "drift" and the "standard deviation" of the tilted profile:

$$\mu(\beta) = \varphi'(\beta), \quad \sigma^2(\beta) = \varphi''(\beta). \tag{12}$$

Introduce the normalized coordinate

$$x_n(k) = x_n(k;\beta) = \frac{k-\mu(\beta)w_n}{\sigma(\beta)\sqrt{w_n}}, \quad k \in \mathbb{Z}. \tag{13}$$

Define the "deterministic cumulants" $\kappa_j(\beta)$ and the "random cumulants" $\chi_j(\beta)$ by

$$\kappa_j(\beta) = \varphi^{(j)}(\beta), \quad \chi_j(\beta) = (\log W_\infty)^{(j)}(\beta). \tag{14}$$

The general Edgeworth expansion for the profile $\mathbb{L}_n$ reads as follows.

Theorem 2.1. *Let $\mathbb{L}_1, \mathbb{L}_2, \dots$ be a sequence of random profiles satisfying Assumptions A1–A4. Fix $r \in \mathbb{N}_0$ and a compact set $K \subset (\beta_-, \beta_+)$. Then,*

$$w_n^{\frac{r+1}{2}} \sup_{k\in\mathbb{Z}} \sup_{\beta\in K} \left| \mathrm{e}^{\beta k - \varphi(\beta) w_n} \mathbb{L}_n(k) - \frac{W_\infty(\beta)\mathrm{e}^{-\frac{1}{2}x_n^2(k)}}{\sigma(\beta)\sqrt{2\pi w_n}} \sum_{j=0}^{r} \frac{G_j(x_n(k);\beta)}{w_n^{j/2}} \right| \xrightarrow[n\to\infty]{a.s.} 0. \tag{15}$$

Here, $G_j(x) = G_j(x;\beta)$, $j \in \mathbb{N}_0$, is a polynomial of degree at most $3j$ given by

$$G_j(x) = \frac{(-1)^j}{j!} \mathrm{e}^{\frac{1}{2}x^2} B_j(D_1, \dots, D_j) \mathrm{e}^{-\frac{1}{2}x^2}, \tag{16}$$

where B_j is the j-th Bell polynomial (defined in Remark 2.2 below) and $D_1, D_2, \dots$ are linear differential operators (with random coefficients) given by

$$D_j = \frac{\varphi^{(j+2)}(\beta)}{(j+1)(j+2)} \left(\frac{1}{\sigma(\beta)} \frac{\mathrm{d}}{\mathrm{d}x}\right)^{j+2} + \chi_j(\beta) \left(\frac{1}{\sigma(\beta)} \frac{\mathrm{d}}{\mathrm{d}x}\right)^{j}. \tag{17}$$

Remark 2.2. The (complete) *Bell polynomials* $B_j(z_1, \dots, z_j)$ are defined by the formal identity

$$\exp\left\{\sum_{j=1}^{\infty} \frac{x^j}{j!} z_j\right\} = \sum_{j=0}^{\infty} \frac{x^j}{j!} B_j(z_1, \dots, z_j). \tag{18}$$

It follows that $B_0 = 1$ and for $j \in \mathbb{N}$,

$$B_j(z_1, \dots, z_j) = \sum{}' \frac{j!}{i_1! \dots i_j!} \left(\frac{z_1}{1!}\right)^{i_1} \dots \left(\frac{z_j}{j!}\right)^{i_j}, \tag{19}$$

where the sum $\sum'$ is taken over all $i_1, \dots, i_j \in \mathbb{N}_0$ satisfying $1i_1 + 2i_2 + \dots + ji_j = j$. We shall need the first three Bell polynomials which are given by

$$B_0 = 1, \quad B_1(z_1) = z_1, \quad B_2(z_1, z_2) = z_1^2 + z_2. \tag{20}$$

Remark 2.3. It follows from (16), (17), (20) that G_0, G_1, G_2 are given by

$$G_0(x) = 1, \tag{21}$$

$$G_1(x) = \frac{\chi_1(\beta)}{\sigma(\beta)} x + \frac{\kappa_3(\beta)}{6\sigma^3(\beta)} \mathrm{He}_3(x), \tag{22}$$

$$G_2(x) = \frac{\chi_1^2(\beta) + \chi_2(\beta)}{2\sigma^2(\beta)} \mathrm{He}_2(x) + \left(\frac{\kappa_4(\beta)}{24\sigma^4(\beta)} + \frac{\kappa_3(\beta)\chi_1(\beta)}{6\sigma^4(\beta)}\right) \mathrm{He}_4(x) + \frac{\kappa_3^2(\beta)}{72\sigma^6(\beta)} \mathrm{He}_6(x), \tag{23}$$

where $\mathrm{He}_n(x)$ denotes the n-th "probabilist" *Hermite polynomial*:

$$\mathrm{He}_n(x) = \mathrm{e}^{\frac{1}{2}x^2}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\right)^n \mathrm{e}^{-\frac{1}{2}x^2}.$$

The first few Hermite polynomials relevant to us are

$$\mathrm{He}_1(x) = x, \quad \mathrm{He}_2(x) = x^2 - 1, \quad \mathrm{He}_3(x) = x^3 - 3x, \tag{24}$$

$$\mathrm{He}_4(x) = x^4 - 6x^2 + 3, \quad \mathrm{He}_6(x) = x^6 - 15x^4 + 45x^2 - 15. \tag{25}$$

Remark 2.4. We have $G_j(-x) = (-1)^j G_j(x)$ for all $j \in \mathbb{N}_0$. In particular, $G_j(0) = 0$ for odd j. Indeed, by (17), D_k is a linear combination of the differential operators $(\mathrm{d}/\mathrm{d}x)^k$ and $(\mathrm{d}/\mathrm{d}x)^{k+2}$. It follows from (19) that $B_j(D_1, \dots, D_j)$ is a linear combination of the differential operators of the form

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{m_1 i_1} \cdots \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{m_j i_j} = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{m_1 i_1 + \ldots + m_j i_j},$$

where each m_k is either k or $k+2$, so that $m_1 i_1 + \ldots + m_j i_j$ has the same parity as j because of the relation $1i_1 + 2i_2 + \ldots + j i_j = j$. Hence, by (16), $G_j(x)$ is a linear combination of Hermite polynomials $\mathrm{He}_k(x)$, where k has the same parity as j. The statement follows from the relation $\mathrm{He}_k(-x) = (-1)^k \mathrm{He}(x)$.

Remark 2.5. In Section 4.2 we will show that $F_j(x;\beta) := W_\infty(\beta) G_j(x;\beta)$ is a polynomial in x (which is evident) whose coefficients are *linear combinations* (rather than rational functions) of $1, W_\infty(\beta), \dots, W_\infty^{(j)}(\beta)$ (which is not evident). For example,

$$W_\infty(\beta)\chi_1(\beta) = W'_\infty(\beta), \quad W_\infty(\beta)(\chi_1^2(\beta) + \chi_2(\beta)) = W''_\infty(\beta),$$

thus proving the above claim for $G_2(x;\beta)$; see (23).

Taking $r = 0$ and $\beta = 0$ in Theorem 2.1 we obtain the following local limit theorem for the profile $\mathbb{L}_n$.

Theorem 2.6. *Let $\mathbb{L}_1, \mathbb{L}_2, \dots$ be a sequence of random profiles satisfying Assumptions A1–A4. Then,*

$$\sqrt{w_n} \sup_{k \in \mathbb{Z}} \left| \mathrm{e}^{-\varphi(0) w_n} \mathbb{L}_n(k) - \frac{W_\infty(0)}{\sigma(0)\sqrt{2\pi w_n}} \exp\left\{ -\frac{1}{2}\left(\frac{k - \mu(0) w_n}{\sigma(0)\sqrt{w_n}}\right)^2 \right\} \right| \xrightarrow[n\to\infty]{a.s.} 0.$$

Theorem 2.1 contains one free parameter β which can be chosen as a function of k and n. With $\beta = 0$ we obtain an asymptotic expansion complementing Theorem 2.6. On the other hand, it is natural to choose $\beta = \beta_n(k)$ as the solution to

$$\varphi'(\beta_n(k)) = \frac{k}{w_n}, \quad \frac{k}{w_n} \in \varphi'((\beta_-, \beta_+)), \tag{26}$$

where the strict monotonicity of φ' has to be recalled. Then, $x_n(k) = 0$ by definition (13), and we obtain the following result.

Theorem 2.7. *Let $\mathbb{L}_1, \mathbb{L}_2, \dots$ be a sequence of random profiles satisfying Assumptions A1–A4. Then, for all $r \in \mathbb{N}_0$ and any compact set $K \subset (\beta_-, \beta_+)$,*

$$w_n^{r+1} \sup_{k \in \mathbb{Z} \cap w_n \varphi'(K)} \left| \frac{\mathrm{e}^{k\beta_n(k)}}{\mathrm{e}^{\varphi(\beta_n(k)) w_n}} \mathbb{L}_n(k) - \frac{W_\infty(\beta_n(k))}{\sigma(\beta_n(k))\sqrt{2\pi w_n}} \sum_{j=0}^{r} \frac{G_{2j}(0;\beta_n(k))}{w_n^j} \right| \xrightarrow[n\to\infty]{a.s.} 0. \tag{27}$$

Remark 2.8. Note that only half-integer powers of w_n are present in the sum in (27) because $G_{2j-1}(0) = 0$ for $j \in \mathbb{N}$; see Remark 2.4. In particular, with $r = 0$ we obtain a precise large deviations asymptotics

$$w_n \sup_{k \in \mathbb{Z} \cap w_n \varphi'(K)} \left| \mathrm{e}^{k\beta_n(k) - \varphi(\beta_n(k)) w_n} \mathbb{L}_n(k) - \frac{W_\infty(\beta_n(k))}{\sigma(\beta_n(k))\sqrt{2\pi w_n}} \right| \xrightarrow[n\to\infty]{a.s.} 0.$$

Remark 2.9 (On mod-ϕ-convergence). Let ϕ be a non-degenerate infinitely divisible distribution with cumulant generating function $\eta(\beta) = \log \int_{\mathbb{R}} \mathrm{e}^{\beta x} \phi(\mathrm{d}x)$. Féray et al. [19] called a sequence of real random variables $(X_n)_{n \in \mathbb{N}}$ *mod-ϕ convergent* with speed $w_n \to +\infty$ if

$$\lim_{n\to\infty} \frac{\mathbb{E}[\mathrm{e}^{\beta X_n}]}{\mathrm{e}^{\eta(\beta) w_n}} = \Psi_\infty(\beta) \tag{28}$$

locally uniformly on some strip $\{\beta \in \mathbb{C} \colon \delta_- < \operatorname{Re}\beta < \delta_+\}$, where $\Psi_\infty(\beta)$ is an analytic function which does not vanish on (δ_-, δ_+). Variations of this definition can be found in [11, 25, 29, 30, 33]. Assuming that (28) holds with speed $O(w_n^{-r})$, for all $r \in \mathbb{N}$, they obtained Edgeworth expansions for both lattice and non-lattice X_n. In particular, Theorem 3.4 of Féray et al. [19] is closely related to expansion (27). In our setting, the distribution of X_n, namely the function $k \mapsto \mathbb{P}[X_n = k]$, is replaced by the profile $k \mapsto \mathbb{L}_n(k)$ (which may be random). More importantly, the analogue of (28) given in Assumptions A2 and A3 holds in some open neighborhood $\mathscr{D}$ of (β_-, β_+), but it fails to hold in the strip $\{\beta \in \mathbb{C} \colon \beta_- < \operatorname{Re}\beta < \beta_+\}$ in our applications to random trees. The function W_∞ replacing Ψ_∞ may be random in our setting. Also note that the expansion in Theorem 2.1 is uniform in the "tuning" parameter β and its terms are given explicitly using Bell and Hermite polynomials, which is convenient in applications.

2.3. **Mode and width.** Using the Edgeworth expansion stated in Theorem 2.1 we can obtain limit theorems for the *width* M_n and the *mode* u_n of the profile $k \mapsto \mathbb{L}_n(k)$. These are defined by

$$M_n = \max_{k \in \mathbb{Z}} \mathbb{L}_n(k), \quad u_n = \arg\max_{k \in \mathbb{Z}} \mathbb{L}_n(k). \tag{29}$$

Theorem 2.10. *Consider a sequence of random profiles $\mathbb{L}_1, \mathbb{L}_2, \dots$ satisfying Assumptions A1–A4. There is an a.s. finite random variable K such that for $n > K$, the mode u_n is equal to $\lfloor u_n^* \rfloor$ or $\lceil u_n^* \rceil$, where*

$$u_n^* = \varphi'(0) w_n + \chi_1(0) - \frac{\kappa_3(0)}{2\sigma^2(0)}. \tag{30}$$

Remark 2.11. The uniqueness of the arg max is a rather subtle question and is not discussed here (see, e.g., [18] where uniqueness of the mode is proved for Stirling numbers of the first kind). In the case when the arg max is non-unique Theorem 2.10 has to be understood as follows: for $n > K$ there are at most two maximizers of $\mathbb{L}_n(k)$ and they belong to the set $\{\lfloor u_n^* \rfloor, \lceil u_n^* \rceil\}$.

The next result on the width $M_n = \mathbb{L}_n(u_n)$ is not surprising in view of the local limit Theorem 2.6.

Theorem 2.12. *Consider a sequence of random profiles $\mathbb{L}_1, \mathbb{L}_2, \dots$ satisfying Assumptions A1–A4. Then the width M_n satisfies*

$$\sigma(0)\sqrt{2\pi w_n} \mathrm{e}^{-\varphi(0) w_n} M_n \xrightarrow[n\to\infty]{a.s.} W_\infty(0). \tag{31}$$

In our applications to random trees, the limiting random variable $W_\infty(0)$ is a.s. constant. It is therefore natural to ask whether it is possible to obtain a more refined result with a non-degenerate limit.

Theorem 2.13. *Consider a sequence of random profiles $\mathbb{L}_1, \mathbb{L}_2, \ldots$ satisfying Assumptions A1–A4. Let*

$$\tilde{M}_n := 2\sigma^2(0) w_n \left(1 - \frac{\sqrt{2\pi w_n}\,\sigma(0)\,M_n}{W_\infty(0)\mathrm{e}^{\varphi(0)w_n}}\right).$$

If $\theta_n := \min_{k\in\mathbb{Z}} |u_n^* - k|$ denotes the distance between u_n^* and the nearest integer, then

$$\tilde{M}_n - \theta_n^2 \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} \chi_2(0) + \frac{\kappa_3^2(0)}{6\sigma^4(0)} - \frac{\kappa_4(0)}{4\sigma^2(0)}.$$

We conclude this section with several generalizations of our main results, all of which are consequences of the proof of Theorem 2.1.

Remark 2.14. Let $(\mathbb{L}_t)_{t\geq 0}$ be a continuous-time profile satisfying the obvious continuous-time formulations of Assumptions A1–A4 for some real-valued function $(w_t)_{t\geq 0}$ with $w_t \to +\infty$ as $t\to\infty$. Then, all results in this section apply analogously to the profile $(\mathbb{L}_t)_{t\geq 0}$.

Remark 2.15. All results remain valid if the sequence $w_n, n\in\mathbb{N}$, is random and $w_n \to +\infty$ almost surely.

Remark 2.16. Under Assumptions A1 and A4, if there exists a sequence of random analytic functions $\tilde{W}_n, n\in\mathbb{N}$, on $\mathscr{D}$ such that the convergence (10) holds with $\tilde{W}_n$ instead of W_∞, then Theorems 2.1, 2.6 and 2.7 hold with W_∞ replaced by $\tilde{W}_n$.

Remark 2.17. Let Assumptions A1 and A2 be fulfilled and assume that the convergence (10) in Assumption A3 holds for some real $r_1 \geq 1/2$ (rather than for all $r\in\mathbb{N}_0$), and the convergence (11) for some real $r_2 > 1/2$. Then, for any $r \leq 2r_1 - 1$, $r\in\mathbb{N}_0$ and $r\leq \alpha < \min\{1+r, 2r_2-1\}$, we have

$$w_n^{\frac{\alpha+1}{2}} \sup_{k\in\mathbb{Z}} \sup_{\beta\in K} \left| \mathrm{e}^{\beta k - \varphi(\beta)w_n}\mathbb{L}_n(k) - \frac{W_\infty(\beta)\mathrm{e}^{-\frac{1}{2}x_n^2(k)}}{\sigma(\beta)\sqrt{2\pi w_n}} \sum_{j=0}^{r} \frac{G_j(x_n(k);\beta)}{w_n^{j/2}} \right| \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} 0.$$

Remark 2.18. Similarly to the previous remark, take Assumptions A1 and A2 for granted and assume that (11) holds for some real $r > 1/2$. Further, in the notation of Assumption A3, instead of (10), impose that

$$\sup_{\beta\in K} |W_n(\beta) - W_\infty(\beta)| \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} 0.$$

Then,

$$\sup_{k\in\mathbb{Z}} \sup_{\beta\in K} \left| \mathrm{e}^{\beta k - \varphi(\beta)w_n}\mathbb{L}_n(k) - \frac{W_\infty(\beta)\mathrm{e}^{-\frac{1}{2}x_n^2(k)}}{\sigma(\beta)\sqrt{2\pi w_n}} \right| \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} 0.$$

2.4. **Example: The classical Edgeworth expansion.** In this and the subsequent section we consider two examples of deterministic profiles. Let $Z_1, Z_2, \ldots$ be i.i.d. integer-valued random variables with mean $\mu := \mathbb{E}Z_1$, variance $\sigma^2 := \operatorname{Var} Z_1 \neq 0$, and cumulant generating function

$$\varphi(\beta) := \log \mathbb{E}\mathrm{e}^{\beta Z_1} \tag{32}$$

which is finite on some interval (β_-, β_+) containing zero. Consider a sequence of deterministic profiles in which $\mathbb{L}_n$ is defined as the probability mass function of the sum $Z_1 + \ldots + Z_n$, that is

$$\mathbb{L}_n(k) = \mathbb{P}[Z_1 + \ldots + Z_n = k], \quad k \in \mathbb{Z}.$$

Then, Assumptions A1–A3 are satisfied with φ as in (32), $w_n = n$ and $W_\infty(\beta) = W_n(\beta) = 1$. Hence, the cumulants χ_k vanish. Assumption A4 is also satisfied if we additionally assume that the minimal step of the distribution of Z_1 is 1. In other words, there is no non-trivial sublattice $a\mathbb{Z} + b$, with $a \in \{2, 3, \ldots\}$ and $b \in \mathbb{Z}$, such that $\mathbb{P}[Z_1 \in a\mathbb{Z} + b] = 1$.

Applying Theorem 2.1 with $\beta = 0$ we obtain the classical Chebyshev–Edgeworth–Cramér asymptotic expansion for sums of i.i.d. lattice random variables, see Theorem 13 in Petrov [34, Ch. VII, p. 205]:

$$\lim_{n\to\infty} n^{\frac{r+1}{2}} \sup_{k\in\mathbb{N}} \left| \mathbb{P}[Z_1 + \ldots + Z_n = k] - \frac{\mathrm{e}^{-\frac{1}{2}x_n^2(k)}}{\sigma\sqrt{2\pi n}} \sum_{j=0}^{r} \frac{q_j(x_n(k))}{n^{j/2}} \right| = 0, \tag{33}$$

where q_j is a polynomial of degree at most $3j$ whose coefficients can be expressed through the cumulants $\kappa_2, \ldots, \kappa_{j+2}$. To obtain q_j, remove in G_j (see (16) for its definition) all terms involving the χ_k's. The first three terms in the expansion are given by

$$q_0(x) = 1, \quad q_1(x) = \frac{\kappa_3}{6\sigma^3}\mathrm{He}_3(x), \quad q_2(x) = \frac{\kappa_4}{24\sigma^4}\mathrm{He}_4(x) + \frac{\kappa_3^2}{72\sigma^6}\mathrm{He}_6(x). \tag{34}$$

Applying Theorem 2.1 with arbitrary β one can obtain asymptotic expansions for large deviation probabilities; see [21]. Note, however, that the moment condition which we imposed on Z_1 can be relaxed; see Theorem 13 in Petrov [34, Ch. VII, p. 205].

2.5. **Example: Stirling numbers of the first kind.** The (unsigned) *Stirling numbers* of the first kind are defined by the formula

$$\theta^{(n)} := \theta(\theta+1)\ldots(\theta+n-1) = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} \theta^k. \tag{35}$$

The following sequence of deterministic profiles given by the probability mass function of the *Ewens distribution* with parameter $\theta > 0$

$$\mathbb{L}_n(k) = \frac{\theta^k}{\theta^{(n)}} \begin{bmatrix} n \\ k \end{bmatrix} \mathbb{1}_{\{k\in\{1,\ldots,n\}\}} \tag{36}$$

can be shown to satisfy Assumptions A1–A4. Applications to Stirling numbers and the Ewens distribution will be studied in a separate paper [27].

3. Edgeworth expansions for random trees

3.1. **One-split branching random walk.** Consider a system of particles on $\mathbb{Z}$ which evolves in discrete time as follows. At time 0, we have a single particle located at 0. In each step *one* of the particles is chosen uniformly at random. This particle is replaced by a random cluster of particles whose displacements w.r.t. the original particle are described by a point process $\zeta = \sum_{i=1}^{N} \delta_{Z_i}$ (where N, the number of particles, is a.s. finite) on $\mathbb{Z}$. In other words, if the original particle is located at x, its descendants are located at $x+Z_1,\dots,x+Z_N$. All random mechanisms involved in this definition are independent.

Remark 3.1. The difference between this model and the usual discrete-time, many splits BRW (for which the Edgeworth expansion was obtained in [21]) is that in the one-split BRW, only *one* particle (chosen uniformly at random) is allowed to split, whereas in the many-split BRW *all* particles split at the same time. We shall see that there are many differences between these models.

Denote by S_n the number of particles after n splitting events, and let their positions be $x_{1,n},\dots,x_{S_n,n}$. Let us denote by $\mathbb{L}_n(k)$ the number of particles at site $k\in\mathbb{Z}$ after n splitting events:

$$\mathbb{L}_n(k) = \#\{1\le j\le S_n\colon x_{j,n}=k\}. \tag{37}$$

We are interested in the function $k\mapsto \mathbb{L}_n(k)$ which is called the *profile* of the one-split BRW.

We are going to state our assumptions on the one-split BRW. Denote by ν_k the expected number of particles at site $k\in\mathbb{Z}$ in the cluster process ζ:

$$\nu_k = \mathbb{E}\zeta(\{k\}) = \mathbb{E}\left[\sum_{i=1}^{N}\mathbb{1}_{\{Z_i=k\}}\right],\quad k\in\mathbb{Z}. \tag{38}$$

The first assumption states that non-zero jumps are possible with positive probability and thus excludes the case in which all particles stay at 0. The second assumption requires the one-split BRW to be supercritical and excludes the possibility that it can become extinct.

Assumption B1: We have $\nu_k>0$ for at least one $k\in\mathbb{Z}\backslash\{0\}$.

Assumption B2: The cluster point process ζ is a.s. non-empty, and the probability that it has at least 2 particles is positive. In other words, $N\ge 1$ a.s. and $\mathbb{P}[N=1]\neq 1$.

Remark 3.2. It is possible to replace this assumption by a weaker one requiring that $\mathbb{E}N>1$ (supercriticality), in which case all results would hold a.s. on the survival event.

Denote by $m(\beta)$ the moment generating function of the intensity of the cluster point process ζ minus 1:

$$m(\beta) = \sum_{k\in\mathbb{Z}} \mathrm{e}^{\beta k}\nu_k - 1 = \mathbb{E}\left[\sum_{i=1}^{N}\mathrm{e}^{\beta Z_i}\right]-1. \tag{39}$$

The expected number of particles at time n is $\mathbb{E}S_n = 1+m(0)n$, where, by Assumption B2,

$$m(0) = \sum_{k\in\mathbb{Z}}\nu_k - 1 = \mathbb{E}N-1>0. \tag{40}$$

Assumption B3: The function m is finite on some non-empty open interval $\mathscr{I}$ containing 0.

It follows that the function m is well-defined for $\beta\in\{z\in\mathbb{C}:\operatorname{Re} z\in\mathscr{I}\}$ and strictly convex and infinitely differentiable on $\mathscr{I}$. We shall need the function

$$\varphi(\beta) = \frac{m(\beta)}{m(0)},\quad \operatorname{Re}\beta\in\mathscr{I}.$$

Denote by $(\beta_-, \beta_+) \subset \mathscr{I}$ the open interval on which $\varphi'(\beta)\beta < \varphi(\beta)$:

$$\beta_- = \inf\{\beta \in \mathscr{I} : \varphi'(\beta)\beta < \varphi(\beta)\}, \tag{41}$$

$$\beta_+ = \sup\{\beta \in \mathscr{I} : \varphi'(\beta)\beta < \varphi(\beta)\}. \tag{42}$$

The interval (β_-, β_+) is non-empty because it contains 0. The endpoints of the intervals $\mathscr{I}$ and (β_-, β_+) are allowed to be infinite.

The (normalized) moment-generating function of the one-split BRW profile is defined, for $\operatorname{Re}\beta \in \mathscr{I}$, by

$$W_n(\beta) = \frac{1}{n^{\varphi(\beta)}} \sum_{i=1}^{S_n} \mathrm{e}^{\beta x_{i,n}}. \tag{43}$$

The following aperiodicity condition plays an important role in the verification of Assumption A4. Here, and subsequently, we denote by $\nu = \sum_{k\in\mathbb{Z}} \nu_k \delta_k$ the intensity measure of the point process ζ.

Assumption B4: ν is not concentrated on any proper additive subgroup of $\mathbb{Z}$. In other words, $\nu(\mathbb{Z}\backslash a\mathbb{Z}) \neq 0$ for all $a \in \{2, 3, \dots\}$.

Assumption B4 can be imposed without loss of generality: if $\nu(a^*\mathbb{Z}) = 1$ for some $a^* \geq 2$ (chosen to be maximal with this property), then we can rescale the jumps by a^* and work equivalently with the one-split BRW governed by the intensity measure ν^*, where $\nu^*(\{k\}) = \nu(\{k/a^*\})$. Note that this contrasts the situation in the many-split BRW [21] and in Section 2.4, where it was necessary to exclude measures ν concentrated on lattices of the form $a\mathbb{Z} + b$.

Finally, we also need the following moment condition which supplements Assumption B3.

Assumption B5: For any $\beta \in (\beta_-, \beta_+)$ there is $\gamma = \gamma(\beta) > 1$ such that

$$\mathbb{E}\left[\left(\sum_{i=1}^{N} \mathrm{e}^{\beta Z_i}\right)^{\gamma}\right] < \infty.$$

Remark 3.3. This is easily shown to be equivalent to the following assumption: For every compact set $K \subset (\beta_-, \beta_+)$ there is $\gamma = \gamma(K) > 1$ such that the above expectation is bounded uniformly in $\beta \in K$.

The next theorem states that the sequence of the one-split BRW profiles satisfies Assumptions A2 and A3 with $w_n = \log n$.

Theorem 3.4. *Under Assumptions B1–B3 and B5, there is an open neighborhood $\mathscr{D}$ of the interval (β_-, β_+) in the complex plane such that, with probability* 1, *W_n converges to some random analytic function W_∞ locally uniformly on $\mathscr{D}$. Moreover, for every compact set $K \subset \mathscr{D}$ and $r \in \mathbb{N}$ we can find an a.s. finite random variable $C_{K,r}$ such that for all $n \in \mathbb{N}$,*

$$\sup_{\beta\in K} |W_n(\beta) - W_\infty(\beta)| < C_{K,r}(\log n)^{-r}. \tag{44}$$

With probability 1, *the function W_∞ has no zeros on the interval (β_-, β_+).*

The proof of the theorem will be given in Section 5.1 and uses an embedding into a continuous-time BRW in conjunction with results of Biggins [3] (see also [41]). The explicit form of the neighborhood $\mathscr{D}$ plays no role in the sequel. However, let us stress that we cannot take $\mathscr{D}$ to be the strip $\{\beta \in \mathbb{C}\colon \beta_- < \operatorname{Re}\beta < \beta_+\}$. In the case of the BSTs, the exact shape of $\mathscr{D}$ can be found in [9]: it is a bounded set. For this reason, the asymptotic expansion obtained by Féray et al. [19] does not apply directly.

Remark 3.5. Note that $\varphi(0) = 1$ and by the law of large numbers,

$$W_\infty(0) = \lim_{n\to\infty} W_n(0) = \lim_{n\to\infty} \frac{S_n}{n} = m(0) \quad \text{a.s.}$$

3.2. **Random trees and one-split BRWs.** We can identify profiles of random trees and profiles of the one-split BRW as follows: Particles correspond to (external or internal) nodes, and positions of particles correspond to the depths of the nodes. In the following, we describe the cluster point process, and give explicit formulas for the quantities $m(0)$, $\varphi(\beta)$, $\mu(0) = \varphi'(0)$, $\sigma^2(0) = \varphi''(0)$ and $\kappa_j(0) = \varphi^{(j)}(0)$, $j \in \mathbb{N}$, which will be relevant in our limit theorems.

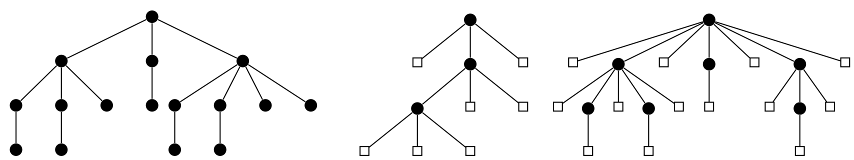


FIGURE 2. Sample realizations of random trees. Left: RRT. Middle: D-ary recursive tree with $D = 3$. Right: PORT.

(i) External profiles of BSTs defined in Section 1.2 correspond to the one-split BRW with the deterministic displacement point process $\zeta = 2\delta_1$ because at any step of the construction an external node at depth k is replaced by two new external nodes at depth $k+1$; see Figure 1 (right). We have

$$\varphi(\beta) = 2\mathrm{e}^\beta - 1, \quad m(0) = 1, \quad \mu(0) = \sigma^2(0) = \kappa_j(0) = 2, \quad j \in \mathbb{N}.$$

The constants $\beta_- \approx -1.678$ and $\beta_+ \approx 0.768$ are the solutions of $2\mathrm{e}^\beta(1-\beta) = 1$.

(ii) Random recursive trees (RRTs), see Figure 2 (left), can be defined as follows. At time $n = 0$ start with one node (denoted by $\bullet$) at level 0. At any step, pick one of the existing nodes (say, x) uniformly at random and connect it to a new node one level deeper than x; see Figure 3 (left). Let $\mathbb{L}_n(k)$ be the number of nodes at depth k in a RRT with $n+1$ nodes. RRTs correspond to the one-split BRW with the deterministic displacement point process $\zeta = \delta_0 + \delta_1$. In particular,

$$\varphi(\beta) = \mathrm{e}^\beta, \quad m(0) = 1, \quad \mu(0) = \sigma^2(0) = \kappa_j(0) = 1, \quad j \in \mathbb{N}.$$

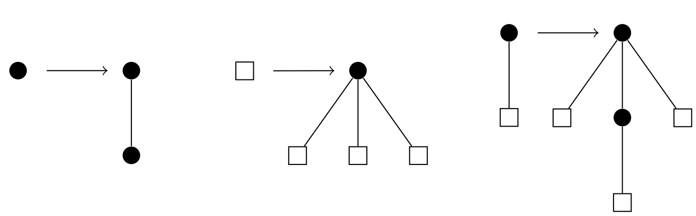


FIGURE 3. Construction rules for random trees. Left: RRT. Middle: D-ary recursive tree with $D = 3$. Right: PORT.

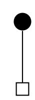
We have $\beta_- = -\infty$ and $\beta_+ = 1$. For results on the profile of RRTs, we refer to [14, 16, 20, 38], see also [15, Section 6.3] for a detailed discussion of the three main methods applied in this context: the martingale method, the method of moments and the contraction method.

(iii) D-ary recursive trees[3] with $D \in \{2, 3, \ldots\}$ are a special case of so-called increasing trees introduced by Bergeron, Flajolet and Salvy [2], see also [15, Sections 1.3.3 and 6.5] and [38] for results on the profile. The model reduces to BSTs for $D = 2$; see Figure 2 (middle). These trees can be constructed in a similar manner as BSTs with the only difference that at each step D new external nodes are attached; see Figure 3 (middle). The external profile of D-ary recursive trees correspond to the one-split BRW with the displacement point process $\zeta = D\delta_1$. We have

$$\varphi(\beta) = \frac{De^{\beta} - 1}{D - 1}, \ \ m(0) = D - 1, \ \ \mu(0) = \sigma^2(0) = \kappa_j(0) = \frac{D}{D-1}, \ \ j \in \mathbb{N}.$$

The constants $\beta_- < 0$ and $\beta_+ > 0$ are the solutions of $De^{\beta}(1 - \beta) = 1$.

(iv) Plane-oriented recursive trees (PORTs), see Figure 2 (right) for a sample realization and [15, Section 1.3.2] for a discussion of this model, are constructed in the following way. At time 0 start with an internal node $\bullet$ at level 0 connected to an external node $\square$ at level 1:



At each step choose one external node uniformly at random, declare it internal and add 3 new external nodes as shown on Figure 3 (right). After n steps we obtain a tree with $2n + 1$ external nodes. As opposed to BSTs, RRTs and D-ary recursive trees, the external profiles of PORTs follow the dynamics of a one-split BRW initiated with one particle (external node) at position *one* (short: initiated at one) at time zero. The displacement point process is $\zeta = 2\delta_0 + \delta_1$. We obtain

$$\varphi(\beta) = \frac{1}{2}(e^{\beta} + 1), \ \ m(0) = 2, \ \ \mu(0) = \sigma^2(0) = \kappa_j(0) = \frac{1}{2}, \ \ j \in \mathbb{N}.$$

We have $\beta_- = -\infty$, whereas $\beta_+ \approx 1.278$ is the solution of $e^{\beta}(\beta - 1) = 1$. The profile of PORTs was studied in [23, 28, 38, 39].

(v) p-oriented trees (which reduce to PORTs if $p = 2$) correspond to the one-split BRW initiated at one with $\zeta = p\delta_0 + \delta_1$, where $p \in \{2, 3, \ldots\}$. They also fall under the general model introduced in [2]. We have

$$\varphi(\beta) = \frac{1}{p}(e^{\beta} + p - 1), \ \ m(0) = p, \ \ \mu(0) = \sigma^2(0) = \kappa_j(0) = \frac{1}{p}, \ \ j \in \mathbb{N}.$$

We have $\beta_- = -\infty$, whereas β_+ is the solution of $e^{\beta}(\beta - 1) = p - 1$. For further information on p-oriented trees, we refer to Sections 1.3.3 and 6.5 in Drmota's monograph [15].

Remark 3.6. Writing $(\mathbb{L}_n(k))_{k \in \mathbb{N}}$ for the external profile of PORTs (or p-oriented trees), and $(\mathbb{L}_n^*(k))_{k \in \mathbb{N}_0}$, for the profile of the corresponding standard one-split BRW initiated at zero, we can identify $\mathbb{L}_n(k) = \mathbb{L}_n^*(k-1)$, for $n \in \mathbb{N}_0, k \in \mathbb{N}$. Denoting by $W_\infty^*(\beta)$ the almost sure limit in Theorem 3.4 for the profile $(\mathbb{L}_n^*(k))_{k \in \mathbb{N}_0}$, the limiting

[3] Not to be confused with m-ary *search* trees, which is a different model; see [15, Section 1.4.2].

process $W_\infty(\beta)$ for the profile $(\mathbb{L}_n(k))_{k\in\mathbb{N}}$ is equal to $\mathrm{e}^\beta W^*_\infty(\beta)$. In particular, for the random cumulants, we have $\chi_1(\beta) = 1+\chi_1^*(\beta)$ and $\chi_k(\beta) = \chi_k^*(\beta)$ for all $k \geq 2$.

Remark 3.7. In all examples listed above the displacement point process ζ is concentrated on $\{0,1\}$ and therefore we have $\varphi(\beta) = 1+\varphi'(0)(\mathrm{e}^\beta - 1)$ (since $\varphi(0) = 1$ by definition) and hence, almost surely

$$n^{\varphi(\beta)} = n \cdot \mathrm{e}^{(\varphi(\beta)-1)\log n} = n \cdot \mathrm{e}^{\varphi'(0)(\mathrm{e}^\beta-1)\log n} \sim \frac{S_n}{m(0)}\mathrm{e}^{\varphi'(0)(\mathrm{e}^\beta-1)\log n},\ n \to \infty.$$

Thus, Theorem 3.4 states that the sequence $(S_n^{-1}\sum_{k\in\mathbb{N}_0}\mathbb{L}_n(k)\delta_k)_{n\in\mathbb{N}}$ of random probability measures on $\mathbb{Z}$ converges in the mod-Poisson sense with probability 1; see [29] and Remark 2.9.

(vi) So far we considered “horizontally projected profiles”. The *vertically projected external profile* of a binary search tree can be defined as follows. At time 0, assign to the root of the BST the abscissa 0. During the construction of the BST, if some external node with abscissa i is chosen, then its two descendants are assigned abscissas $i-1$ and $i+1$. The abscissa of a node describes its so-called *left-right imbalance* since it measures the difference between the number of times the path from the root to the node turns right rather than left. Denote by $\mathbb{L}_n(k)$ the number of external nodes with abscissa $k \in \mathbb{Z}$ in a BST with $n+1$ external nodes. This profile corresponds to the one-split BRW with $\zeta = \delta_{-1} + \delta_{+1}$ and we have

$$\varphi(\beta) = \mathrm{e}^\beta + \mathrm{e}^{-\beta} - 1,\ m(0) = 1,\ \mu(0) = \kappa_{2j-1}(0) = 0, \sigma^2(0) = \kappa_{2j}(0) = 2,\ j \in \mathbb{N}.$$

The constants $\beta_+ \approx 0.9071$ and $\beta_- = -\beta_+$ are the solutions of the equation $(\mathrm{e}^\beta - \mathrm{e}^{-\beta})\beta = \mathrm{e}^\beta + \mathrm{e}^{-\beta} - 1$. The left-right imbalance of nodes and the corresponding path length were studied by Kuba and Panholzer [31], the profile by Schopp [38].

3.3. **Jabbour martingale.** In all models listed in the previous section, the number of descendants of any particle in the one-split BRW is deterministic. Recall that this number equals $m(0)+1$, so that for BSTs, RRTs and PORTs we have $m(0) = 1, 1, 2$, respectively. In this case, it turns out that the Laplace transform of the particle positions divided by its expectation is a martingale. In the case of BSTs, this martingale has been introduced by Jabbour-Hattab [24]; see also Chauvin et al. [7]. The next theorem generalizes Jabbour's martingale to general one-split BRWs with a deterministic number of descendants.

Theorem 3.8. *Consider a one-split branching random walk in which the number of descendants of every particle is deterministic and equals* $m(0)+1 \in \mathbb{N}$*. Assume that the function* $m(\beta)$ *defined by* (39) *is finite on some interval* $\mathscr{I}$ *containing* 0*. Then, for all* $\beta \in \{z \in \mathbb{C}\colon \operatorname{Re} z \in \mathscr{I}\}$*, the sequence* $(J_n(\beta))_{n\in\mathbb{N}_0}$ *defined by*

$$J_n(\beta) := \frac{1}{\alpha_n(\beta)}\sum_{i=1}^{1+m(0)n}\mathrm{e}^{\beta x_{i,n}}, \quad \alpha_n(\beta) = \prod_{k=0}^{n-1}\left(1 + \frac{m(\beta)}{1+m(0)k}\right),$$

is a martingale w.r.t. the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$*, where* $\mathcal{F}_n$ *is a* σ*-algebra generated by the first* n *generations of the one-split BRW. Also,* $\mathbb{E}J_n(\beta) = 1$*.*

Proof. Note that the number of particles at time n is $S_n = 1+m(0)n$. Let $Z_n(\beta) = \sum_{i=1}^{1+m(0)n}\mathrm{e}^{\beta x_{i,n}}$, where $x_{1,n},\ldots,x_{S_n,n}$ are the positions of the particles at time n.

Denoting by $\zeta_n = \sum_{j=1}^{m(0)+1} \delta_{Z_{j,n}}$ the point process of descendants used to pass from generation n to generation $n+1$, we have

$$\begin{aligned}\mathbb{E}[Z_{n+1}(\beta)|\mathcal{F}_n] &= \frac{1}{1+m(0)n} \sum_{i=1}^{1+m(0)n} \mathbb{E}\left[Z_n(\beta) - \mathrm{e}^{\beta x_{i,n}} + \sum_{j=1}^{m(0)+1} \mathrm{e}^{\beta(x_{i,n}+Z_{j,n})} \middle| \mathcal{F}_n \right] \\ &= Z_n(\beta) + \left[\sum_{i=1}^{1+m(0)n} \mathrm{e}^{\beta x_{i,n}} \left(\mathbb{E} \sum_{j=1}^{m(0)+1} \mathrm{e}^{\beta Z_{j,n}} - 1 \right) \right] \frac{1}{1+m(0)n} \\ &= Z_n(\beta) \left(1 + \frac{m(\beta)}{1+m(0)n} \right),\end{aligned}$$

where we used that $m(\beta) = \mathbb{E} \sum_{j=1}^{m(0)+1} \mathrm{e}^{\beta Z_{j,n}} - 1$; see (39). It follows that $J_n(\beta)$ is a martingale. Since $J_0(\beta) = 1$, we have $\mathbb{E} J_n(\beta) = 1$ for all $n \in \mathbb{N}_0$. □

For the function W_n introduced in (43) we obtain (in the case of deterministic number of descendants)

$$\mathbb{E} W_n(\beta) = \frac{\alpha_n(\beta)}{n^{\varphi(\beta)}} = n^{-\frac{m(\beta)}{m(0)}} \frac{\left(\frac{m(\beta)+1}{m(0)} \right)^{(n)}}{\left(\frac{1}{m(0)} \right)^{(n)}}, \quad \operatorname{Re} \beta \in \mathscr{I}, \tag{45}$$

where $z^{(n)} := z(z+1)\dots(z+n-1)$ is the rising factorial. For $\operatorname{Re}\beta \in \mathscr{I}$, using the formula $z^{(n)} \sim n^z \Gamma(n)/\Gamma(z)$ as $n \to \infty$, we obtain

$$\lim_{n\to\infty} \mathbb{E} W_n(\beta) = \frac{\Gamma\left(\frac{1}{m(0)} \right)}{\Gamma\left(\frac{m(\beta)+1}{m(0)} \right)}. \tag{46}$$

Further, for $\beta \in (\beta_-, \beta_+)$ we shall show that $\mathbb{E} W_\infty(\beta) = \lim_{n\to\infty} \mathbb{E} W_n(\beta)$; see Section 5.3, below.

3.4. **Cumulants of the profile.** Recall that $x_{1,n}, \dots, x_{S_n,n}$ denote the positions of the particles in a one-split BRW after n splits. For $\beta \in \mathscr{I}$ consider

$$\chi_{k,n}(\beta) = \left(\frac{\mathrm{d}}{\mathrm{d}\beta} \right)^k \log \sum_{i=1}^{S_n} \mathrm{e}^{\beta x_{i,n}}.$$

It is easy to see that $\chi_{k,n}(\beta)$ is the k-th cumulant of the Gibbs probability measure assigning to each point $x_{i,n}$ the weight proportional to $\mathrm{e}^{\beta x_{i,n}}$.

Remark 3.9. The most interesting case is $\beta = 0$. Then, $\chi_{k,n} := \chi_{k,n}(0)$ is the k-th cumulant of the empirical measure assigning to each particle $x_{i,n}$ the same weight $1/S_n$. For example,

$$\chi_{1,n} = \frac{1}{S_n} \sum_{i=1}^{S_n} x_{i,n}, \quad \chi_{2,n} = \frac{1}{S_n} \sum_{i=1}^{S_n} (x_{i,n} - \chi_{1,n})^2, \quad \chi_{3,n} = \frac{1}{S_n} \sum_{i=1}^{S_n} (x_{i,n} - \chi_{1,n})^3$$

are the empirical mean, the empirical variance and the empirical central third moment of the particle positions in the one-split BRW. In the context of random trees, $S_n \chi_{1,n}$ is the external path length of the tree. Specifically, in the BST case, $(n+1)\chi_{1,n} - 2n$ can be interpreted as the number of key comparisons used by the Quicksort algorithm to sort n randomly ordered items.

Theorem 3.10. *Consider a one-split BRW satisfying Assumptions B1–B3 and B5. Uniformly on any compact set $K \subset (\beta_-, \beta_+)$ we have*

$$\left(\frac{\mathrm{d}}{\mathrm{d}\beta}\right)^k \log W_n(\beta) = \chi_{k,n}(\beta) - \varphi^{(k)}(\beta)\log n \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} \chi_k(\beta), \tag{47}$$

where the limiting random variable $\chi_k(\beta)$ is given by

$$\chi_k(\beta) = \left(\frac{\mathrm{d}}{\mathrm{d}\beta}\right)^k \log W_\infty(\beta).$$

Proof. The equality in (47) follows from the definition of W_n; see (43). We prove the convergence. Let $\mathscr{D}$ be an open neighborhood of (β_-, β_+) as in Theorem 3.4. In the probability space on which the one-split BRW is defined, consider some outcome ω and let $\mathscr{D}' := \mathscr{D}'(\omega) \subseteq \mathscr{D}$ be an open subset with $K \subseteq \mathscr{D}'$ such that W_∞ is almost surely non-zero on the closure of $\mathscr{D}'$, and the analytic functions $W_n(\cdot;\omega)$ converge, as $n \to \infty$, to $W_\infty(\cdot;\omega)$ in $\mathcal{H}(\mathscr{D}')$, the set of analytic functions on $\mathscr{D}'$ with the topology of locally uniform convergence. The set of such outcomes has probability 1. By Theorem 3.4,

$$\log W_n(\beta) \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} \log W_\infty(\beta) \quad \text{in } \mathcal{H}(\mathscr{D}').$$

Observe that the logarithm can be defined continuously since W_∞ and W_n do not vanish for sufficiently large n. By the Cauchy formula, taking the k-th derivative is a continuous map from $\mathcal{H}(\mathscr{D}')$ to $\mathcal{H}(\mathscr{D}')$. Consequently,

$$\left(\frac{\mathrm{d}}{\mathrm{d}\beta}\right)^k \log W_n(\beta) \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} \left(\frac{\mathrm{d}}{\mathrm{d}\beta}\right)^k \log W_\infty(\beta) \quad \text{in } \mathcal{H}(\mathscr{D}'),$$

and hence uniformly in $\beta \in K$. This concludes the proof. □

3.5. **Edgeworth expansion for one-split BRW.** We are going to state an Edgeworth expansion for the profile of the one-split BRW. We recall the parameters $\mu(\beta)$ and $\sigma(\beta)$ from (12),

$$\mu(\beta) = \varphi'(\beta), \quad \sigma^2(\beta) = \varphi''(\beta), \tag{48}$$

as well as the deterministic cumulants $\kappa_j(\beta) = \varphi^{(j)}(\beta)$, $j \in \mathbb{N}$. As in (13) (with $w_n = \log n$), we introduce the normalized coordinate

$$x_n(k) = x_n(k;\beta) = \frac{k - \mu(\beta)\log n}{\sigma(\beta)\sqrt{\log n}}, \quad k \in \mathbb{Z}. \tag{49}$$

Theorem 3.11. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the profile at time n of a one-split branching random walk satisfying Assumptions B1–B5. Fix $r \in \mathbb{N}_0$ and a compact set $K \subset (\beta_-, \beta_+)$. Then,*

$$(\log n)^{\frac{r+1}{2}} \sup_{k\in\mathbb{Z}} \sup_{\beta\in K} \left| \frac{\mathrm{e}^{\beta k}\mathbb{L}_n(k)}{n^{\varphi(\beta)}} - \frac{W_\infty(\beta)\mathrm{e}^{-\frac{1}{2}x_n^2(k)}}{\sigma(\beta)\sqrt{2\pi\log n}} \sum_{j=0}^{r} \frac{G_j(x_n(k);\beta)}{(\log n)^{j/2}} \right| \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} 0. \tag{50}$$

Here, $G_j(x) = G_j(x;\beta)$ is a polynomial of degree at most $3j$ given by

$$G_j(x) = \frac{(-1)^j}{j!}\,\mathrm{e}^{\frac{1}{2}x^2} B_j(D_1, \ldots, D_j)\mathrm{e}^{-\frac{1}{2}x^2}, \tag{51}$$

where B_j is the j-th Bell polynomial (see Remark 2.2) and $D_1, D_2, \dots$ are differential operators (with random coefficients) given by

$$D_j = \frac{\varphi^{(j+2)}(\beta)}{(j+1)(j+2)}\left(\frac{1}{\sigma(\beta)}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{j+2} + \chi_j(\beta)\left(\frac{1}{\sigma(\beta)}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{j} \tag{52}$$

with $\chi_j(\beta)$ as in Theorem 3.10.

Remark 3.12. The expressions for the first three terms in the expansion have the same form as in (21), (22), (23).

Remark 3.13. An Edgeworth expansion for the profile of a *many-split* BRW was obtained in [21]. Theorem 2.1 from the present paper can be applied to the many-split BRW, but both the representation of the terms of the expansion and the proof given in [21] differ from ours. In Section 4.2 we provide an alternative representation for the terms in Theorem 2.1 which allows to derive the many-split BRW expansion of [21]. There are many differences between the one-split and many-split BRW cases. For example, in the former case the expansions are in negative powers of $\sqrt{\log n}$, whereas in the latter case negative powers of $\sqrt{n}$ appear.

Taking $\beta = 0$ and $r = 0$ in Theorem 3.11, and recalling that $W_\infty(0) = m(0)$ and $\varphi(0) = 1$, we obtain the following local limit theorem for the one-split BRW.

Theorem 3.14. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the profile at time n of a one-split branching random walk satisfying Assumptions B1–B5. Then,*

$$\sqrt{\log n}\,\sup_{k\in\mathbb{Z}}\left|\frac{\mathbb{L}_n(k)}{n} - \frac{m(0)}{\sigma(0)\sqrt{2\pi\log n}}\exp\left\{-\frac{1}{2}\left(\frac{k-\mu(0)\log n}{\sigma(0)\sqrt{\log n}}\right)^2\right\}\right| \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} 0. \tag{53}$$

More terms can be obtained by taking $\beta = 0$ and arbitrary $r \in \mathbb{N}_0$. Another possibility is to take $\beta = \beta_n(k)$ as in (26), that is $\varphi'(\beta_n(k)) = k/\log n$. Then, $x_n(k) = 0$ and we obtain the following expansion containing only half-integer powers of $\log n$ (c.f. Theorem 2.7):

Theorem 3.15. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the profile at time n of a one-split branching random walk satisfying Assumptions B1–B5. Then, for all $r \in \mathbb{N}_0$ and any compact set $K \subset (\beta_-, \beta_+)$,*

$$(\log n)^{r+1}\sup_{k\in\mathbb{Z}\cap(\log n)\varphi'(K)}\left|\frac{\mathrm{e}^{k\beta_n(k)}}{n^{\varphi(\beta_n(k))}}\mathbb{L}_n(k) - \frac{1}{\sqrt{2\pi\log n}}\sum_{j=0}^{r}\frac{F_{2j}(0;\beta_n(k))}{\sigma(\beta_n(k))(\log n)^j}\right| \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} 0,$$

where $F_{2j}(0;\beta) := W_\infty(\beta)G_{2j}(0;\beta)$ is a linear combination of $1, W_\infty(\beta), \dots, W_\infty^{(2j)}(\beta)$ (see Section 4.2 for the proof of the latter claim).

Our results easily imply an expansion similar to Theorem 3.11 for the *mean* of the profile when the number of particles in the first generation is almost surely constant.

Theorem 3.16. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the profile at time n of a one-split branching random walk with the deterministic number of descendants and satisfying Assumptions B1–B5. Fix $r \in \mathbb{N}_0$ and a compact set $K \subset (\beta_-, \beta_+)$. Then,*

$$(\log n)^{\frac{r+1}{2}}\sup_{k\in\mathbb{Z}}\sup_{\beta\in K}\left|\frac{\mathrm{e}^{\beta k}\mathbb{E}[\mathbb{L}_n(k)]}{n^{\varphi(\beta)}} - \frac{\mathbb{E}[W_\infty(\beta)]\,\mathrm{e}^{-\frac{1}{2}x_n^2(k)}}{\sigma(\beta)\sqrt{2\pi\log n}}\sum_{j=0}^{r}\frac{\tilde{G}_j(x_n(k);\beta)}{(\log n)^{j/2}}\right| \underset{n\to\infty}{\longrightarrow} 0.$$

Here, $\tilde{G}_j(x;\beta)$ is defined by the same formulas (51), (52) *as $G_j(x;\beta)$, but with $\chi_j(\beta)$ replaced by its deterministic analogue*

$$\tilde{\chi}_j(\beta) = \left(\frac{\mathrm{d}}{\mathrm{d}\beta}\right)^j \log \mathbb{E} W_\infty(\beta) = -\left(\frac{\mathrm{d}}{\mathrm{d}\beta}\right)^j \log \Gamma\left(\frac{m(\beta)+1}{m(0)}\right).$$

Again, it is natural to choose β as in (26). Then, $x_n(k) = 0$ and one obtains an expansion containing half-integer powers of $\log n$ only.

3.6. **Width and mode of the one-split BRW.** Recall the definitions of the *width* M_n and the *mode* u_n of a one-split BRW at time n in (29). In the setting of random trees, the mode is the level having the largest number of nodes, while the width is the maximal number of nodes at a level. From Theorems 2.10, 2.12 and 2.13 we obtain the following results for the one-split BRW.

Theorem 3.17. *Consider a one-split BRW satisfying Assumptions B1–B5. There is an a.s. finite random variable K such that for $n > K$, the mode u_n is equal to one of the numbers $\lfloor u_n^* \rfloor$ or $\lceil u_n^* \rceil$, where*

$$u_n^* = \varphi'(0)\log n + \chi_1(0) - \frac{\kappa_3(0)}{2\sigma^2(0)}. \tag{54}$$

Remark 3.18. In fact, one can provide more information on which of the two values, $\lfloor u_n^* \rfloor$ or $\lceil u_n^* \rceil$, is the mode. Let $\mathrm{nint}(u_n^*) = \arg\min_{k\in\mathbb{Z}} |u_n^* - k|$ be the integer closest to u_n^* with convention that $\mathrm{nint}(u_n^*) = \lfloor u_n^* \rfloor$ if u_n^* is a half-integer. The proof of Theorem 2.10, see formula (89) below, shows that, for every $\varepsilon > 0$, we can find an a.s. finite random variable $K(\varepsilon)$ such that for all $n > K(\varepsilon)$ satisfying $\min_{k\in\mathbb{Z}} |u_n^* - k - \frac{1}{2}| > \varepsilon$, the mode u_n is unique and equals $\mathrm{nint}(u_n^*)$.

Case 1: $\varphi'(0) = 0$ (meaning that the one-split BRW has no drift, which applies to Example (vi) of Section 3.2). If the random variable $\chi_1(0)$ has no atoms, then $\chi_1(0) - \frac{1}{2}\kappa_3(0)/\sigma^2(0)$ is not a half-integer with probability 1. It follows that there is an a.s. finite random variable K_1 such that

$$u_n = \mathrm{nint}\left(\chi_1(0) - \frac{\kappa_3(0)}{2\sigma^2(0)}\right)$$

for all $n > K_1$. Absolute continuity of $\chi_1(0)$ in Example (vi) follows from the fixed point equation derived in [31].

Case 2: $\varphi'(0) \neq 0$ (which is true in examples (i)–(v) of Section 3.2). The arithmetic properties of the sequence $(\{\varphi'(0)\log n\})_{n\in\mathbb{N}}$, with $\{\cdot\}$ denoting the fractional part, allow us to deduce an additional information compared to the general result given by Theorem 2.10. Here, we say that a set $A \subset \mathbb{N}$ has *asymptotic density* $\alpha \in [0,1]$ if

$$\lim_{n\to\infty} \frac{\#(A \cap \{1,\dots,n\})}{n} = \alpha. \tag{55}$$

Proposition 3.19. *Consider a one-split BRW satisfying Assumptions B1–B5 with $\varphi'(0) \neq 0$. Then, almost surely,*

(i) *there are arbitrary long intervals of consecutive n's for which u_n is unique and $u_n = \lfloor u_n^* \rfloor$; and, similarly, arbitrary long intervals for which u_n is unique and $u_n = \lceil u_n^* \rceil$;*

(ii) *the asymptotic density of the set $A = \{n \in \mathbb{N} : u_n \text{ is unique and } u_n = \mathrm{nint}(u_n^*)\}$ equals one.*

The next two results on the width M_n are special cases of Theorems 2.12 and 2.13.

Theorem 3.20. *Consider a one-split BRW satisfying Assumptions B1–B5. Then, the width M_n satisfies*

$$\frac{\sqrt{2\pi \log n}\,\sigma(0)\,M_n}{m(0)\,n} \xrightarrow[n\to\infty]{a.s.} 1. \tag{56}$$

Theorem 3.21. *Consider a one-split BRW satisfying Assumptions B1–B5. Let*

$$\tilde{M}_n := 2\sigma^2(0)\log n \left(1 - \frac{\sqrt{2\pi \log n}\,\sigma(0)\,M_n}{m(0)\,n}\right).$$

If $\theta_n := \min_{k\in\mathbb{Z}} |u_n^* - k|$ denotes the distance between u_n^* and the nearest integer, then

$$\tilde{M}_n - \theta_n^2 \xrightarrow[n\to\infty]{a.s.} \chi_2(0) + \frac{\kappa_3^2(0)}{6\sigma^4(0)} - \frac{\kappa_4(0)}{4\sigma^2(0)}.$$

Remark 3.22. Again, the details depend on whether $\varphi'(0)$ vanishes or not.

Case 1: Suppose that $\varphi'(0) = 0$. Then, θ_n does not depend on n and we obtain

$$\tilde{M}_n \xrightarrow[n\to\infty]{a.s.} \chi_2(0) + \frac{\kappa_3^2(0)}{6\sigma^4(0)} - \frac{\kappa_4(0)}{4\sigma^2(0)} + \min_{k\in\mathbb{Z}} \left|\chi_1(0) - \frac{\kappa_3(0)}{2\sigma^2(0)} - k\right|^2.$$

Case 2: $\varphi'(0) \neq 0$. The sequence $(\{\varphi'(0)\log n\})_{n\in\mathbb{N}}$ is dense in $[0,1]$. This implies that the set of subsequential limits of the sequence θ_n is equal to $[0,1/2]$. It follows that

$$\liminf_{n\to\infty} \tilde{M}_n = \chi_2(0) + \frac{\kappa_3^2(0)}{6\sigma^4(0)} - \frac{\kappa_4(0)}{4\sigma^2(0)} \quad \text{a.s.},$$
$$\limsup_{n\to\infty} \tilde{M}_n = \chi_2(0) + \frac{\kappa_3^2(0)}{6\sigma^4(0)} - \frac{\kappa_4(0)}{4\sigma^2(0)} + \frac{1}{4} \quad \text{a.s.},$$

and every point between the lim inf and lim sup is a.s. a subsequential limit of $\tilde{M}_n$. Thus, we have infinitely many different a.s. (and hence, weak) subsequential limits of $\tilde{M}_n$. If $\chi_2(0)$ is non-degenerate, it follows from the convergence of types lemma that the random variable M_n cannot be normalized by an affine transformation to converge (in the weak sense) to a non-degenerate limit law. This agrees with the observation of Fuchs et al. [20, Theorem 2].

3.7. **Occupation numbers in the one-split BRW.** Consider a one-split BRW with profiles $\mathbb{L}_1, \mathbb{L}_2, \ldots$. In this section we shall state limit theorems on the "occupation numbers" $\mathbb{L}_n(k_n)$, where k_n is a (deterministic) integer sequence with some regular type of behavior. These limit theorems can be applied to random trees (including BSTs, RRTs and PORTs; see Section 3.2) and improve on the results of Fuchs et al. [20]. In these applications $\mathbb{L}_n(k_n)$ is interpreted as the number of nodes at depth k_n in a random tree. To prove these theorems, we shall use a suitable number of terms in the Edgeworth expansion of $\mathbb{L}_n$ stated in Theorem 3.11. Our aim is to find a non-degenerate limit distribution for $\mathbb{L}_n(k_n)$, but it turns out that our results hold even in the sense of a.s. convergence. As in (26), define $\beta_n = \beta_n(k)$ to be the solution of

$$\varphi'(\beta_n(k)) = \frac{k}{\log n}, \quad \frac{k}{\log n} \in \varphi'((\beta_-, \beta_+)). \tag{57}$$

Theorem 3.23. *Consider a one-split BRW satisfying Assumptions B1–B5. Let k_n be an integer sequence such that, for some $\beta \in (\beta_-, \beta_+)$, we have $k_n = \varphi'(\beta)\log n + o(\log n)$. Then, with $\beta_n = \beta_n(k_n)$ as in (57), we have*

$$\frac{\sqrt{\log n}}{n^{\varphi(\beta_n)-\beta_n\varphi'(\beta_n)}}\,\mathbb{L}_n(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{W_\infty(\beta)}{\sqrt{2\pi}\sigma(\beta)}. \tag{58}$$

If, for some $\alpha \in \mathbb{R}$,

$$k_n = \varphi'(\beta)\log n + \alpha\sigma(\beta)\sqrt{\log n} + o(\sqrt{\log n}), \quad n \to \infty, \tag{59}$$

then

$$\frac{\sqrt{\log n}}{n^{\varphi(\beta)}}\,\mathrm{e}^{\beta k_n}\mathbb{L}_n(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\mathrm{e}^{-\frac{1}{2}\alpha^2}}{\sqrt{2\pi}\sigma(\beta)}W_\infty(\beta). \tag{60}$$

Proof of Theorem 3.23. From Theorem 3.11 with $r = 0$ and $K \subseteq (\beta_-, \beta_+)$ compact, we have

$$\sqrt{\log n}\sup_{\beta'\in K}\sup_{k\in\mathbb{Z}}\left|\frac{\mathrm{e}^{\beta' k}\mathbb{L}_n(k)}{n^{\varphi(\beta')}} - \frac{W_\infty(\beta')\mathrm{e}^{-\frac{1}{2}\left(\frac{k-\varphi'(\beta')\log n}{\sigma(\beta')\sqrt{\log n}}\right)^2}}{\sigma(\beta')\sqrt{2\pi\log n}}\right| \xrightarrow[n\to\infty]{a.s.} 0. \tag{61}$$

From here, (58) follows readily upon taking $k = k_n$, $\beta' = \beta_n(k_n)$ as in (57) (which converges to β, as $n \to \infty$), recalling the continuity of W_∞ and σ and noting that the term in the exponent vanishes. Formula (60) follows from (59) upon choosing $\beta' = \beta$ in (61) and using the observation

$$x_n(k_n) = \frac{k_n - \varphi'(\beta)\log n}{\sigma(\beta)\sqrt{\log n}} \underset{n\to\infty}{\longrightarrow} \alpha.$$

The proof is complete. □

Remark 3.24. If, in addition to the conditions stated in the theorem, we assume that the BRW has a deterministic number of descendants, then the convergences (58) and (60) also hold in L_1 sense.

Theorem 3.23 is applicable in the case $\beta = 0$, however, $W_\infty(0) = m(0)$ is a.s. constant (see Remark 3.5) meaning that the limits in (58) and (60) are degenerate. It is therefore natural to ask whether non-degenerate limits can be obtained by choosing a more refined normalization of $\mathbb{L}_n(k_n)$. Denote by $\mathbb{L}_n^\circ(k)$ the profile centered by its expectation:

$$\mathbb{L}_n^\circ(k) = \mathbb{L}_n(k) - \mathbb{E}[\mathbb{L}_n(k)], \quad k \in \mathbb{Z}.$$

In the following we assume that the integer sequence k_n can be represented in the form

$$k_n = \varphi'(0)\log n + c_n,$$

where c_n is a sequence on which we impose various growth conditions. While Theorem 3.23 can be derived from the first term in the Edgeworth expansion (meaning that $r = 0$), the following more refined theorem requires more terms (meaning that $r = 1$ or $r = 2$).

Theorem 3.25. *Consider a one-split BRW with deterministic number of descendants and satisfying Assumptions B1–B5. Let $(k_n)_{n\in\mathbb{N}}$ be an integer sequence.*

(a) *If $k_n = \varphi'(0)\log n + \alpha\sigma(0)\sqrt{\log n} + o(\sqrt{\log n})$ for some $\alpha \in \mathbb{R}$, then*

$$\frac{\log n}{n}\,\mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{m(0)\alpha e^{-\frac{1}{2}\alpha^2}}{\sqrt{2\pi}\sigma^2(0)}\left(\chi_1(0) - \mathbb{E}\chi_1(0)\right). \tag{62}$$

(b) *If $k_n = \varphi'(0)\log n + c_n$ with $\lim_{n\to\infty}|c_n| = \infty$ and $c_n = o(\log n)$, then, with β_n as in* (57)*,*

$$\frac{(\log n)^{\frac{3}{2}}}{c_n n^{\varphi(\beta_n)-\beta_n\varphi'(\beta_n)}}\,\mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{m(0)\left(\chi_1(0) - \mathbb{E}\chi_1(0)\right)}{\sqrt{2\pi}\sigma^3(0)}. \tag{63}$$

In particular, if $\lim_{n\to\infty}|c_n| = \infty$ and $c_n = o(\sqrt{\log n})$, then

$$\frac{(\log n)^{\frac{3}{2}}}{nc_n}\mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{m(0)(\chi_1(0) - \mathbb{E}\chi_1(0))}{\sqrt{2\pi}\sigma^3(0)}. \tag{64}$$

(c) *If $k_n = \varphi'(0)\log n + c_n$ where c_n is bounded, then*

$$\frac{(\log n)^{\frac{3}{2}}}{n}\mathbb{L}_n^\circ(k_n) - R^\circ(c_n) \xrightarrow[n\to\infty]{a.s.} 0, \tag{65}$$

where $R^\circ(c) = R(c) - \mathbb{E}R(c)$ and $R(c)$ is a random variable given by

$$R(c) := \frac{m(0)}{\sqrt{2\pi}\sigma^3(0)}\left(\chi_1(0)\left(c + \frac{\kappa_3(0)}{2\sigma^2(0)}\right) - \frac{\chi_1^2(0) + \chi_2(0)}{2}\right), \quad c \in \mathbb{R}. \tag{66}$$

Proof of Theorem 3.25. Taking $\beta = 0$ and $r = 1$ in Theorem 3.11 and using formula (22), we obtain

$$\frac{\mathbb{L}_n(k_n)}{n} = \frac{m(0)e^{-\frac{1}{2}x_n^2(k_n)}}{\sigma(0)\sqrt{2\pi\log n}}\left(1 + \frac{\frac{\chi_1(0)}{\sigma(0)}x_n(k_n) + \frac{\kappa_3(0)}{6\sigma^3(0)}\mathrm{He}_3(x_n(k_n))}{\sqrt{\log n}}\right) + o\left(\frac{1}{\log n}\right)$$

almost surely. By Theorem 3.16, we have an analogous expansion for the expectation of $\mathbb{L}_n(k_n)$. Subtracting both expansions, we obtain

$$\frac{\mathbb{L}_n^\circ(k_n)}{n} = \frac{m(0)e^{-\frac{1}{2}x_n^2(k_n)}}{\sigma(0)\sqrt{2\pi\log n}} \cdot \frac{\chi_1(0) - \mathbb{E}\chi_1(0)}{\sigma(0)\sqrt{\log n}}x_n(k_n) + o\left(\frac{1}{\log n}\right) \quad \text{a.s.}, \tag{67}$$

where

$$x_n(k_n) = \frac{k_n - \varphi'(0)\log n}{\sigma(0)\sqrt{\log n}} = \frac{c_n}{\sigma(0)\sqrt{\log n}}.$$

To prove (62), it is enough to notice that $\lim_{n\to\infty} x_n(k_n) = \alpha$. Inserting this into (67), we obtain (62).

For the remaining results, we need to apply the Edgeworth expansion with $r = 2$. First, choosing $\beta = 0$ in Theorems 3.11, 3.16, using (22), (23), and subtracting the expansions for $\mathbb{L}_n(k_n)$ and $\mathbb{E}[\mathbb{L}_n(k_n)]$, we obtain

$$\frac{1}{n}\mathbb{L}_n^{\circ}(k_n) = \frac{m(0)\mathrm{e}^{-\frac{1}{2}x_n^2(k_n)}}{\sigma(0)\sqrt{2\pi\log n}}\Bigg(\frac{\chi_1(0)-\mathbb{E}\chi_1(0)}{\sigma(0)\sqrt{\log n}}x_n(k_n) + \kappa_3(0)\frac{\chi_1(0)-\mathbb{E}\chi_1(0)}{6\sigma^4(0)\log n}\mathrm{He}_4(x_n(k_n)) + \frac{\chi_1^2(0)+\chi_2(0)-\mathbb{E}[\chi_1^2(0)+\chi_2(0)]}{2\sigma^2(0)\log n}\mathrm{He}_2(x_n(k_n))\Bigg) + o\left(\frac{1}{(\log n)^{\frac{3}{2}}}\right) \quad \text{a.s.}$$

Multiplying both sides of the last display by $(\log n)^{3/2}/c_n$ yields (64) because $\lim_{n\to\infty} x_n(k_n) = 0$ and $\mathrm{He}_2(x) = -1 + o(1), \mathrm{He}_4(x) = 3 + o(1)$ as $x \to 0$. For the proof of (65) use the same expansion as above and note that $x_n(k_n) = O(1/\sqrt{\log n})$.

It remains to show that (63) holds. Here, we use the Edgeworth expansion with $r = 2$ and β_n as in (57). First, note that, by a simple Taylor expansion, we have

$$\beta_n = \frac{c_n(1+o(1))}{\sigma^2(0)\log n}. \tag{68}$$

Next, by Theorem 3.11 and (21), (23),

$$\frac{\mathrm{e}^{\beta_n k_n}}{n^{\varphi(\beta_n)}}\mathbb{L}_n(k_n) = \frac{1}{\sigma(\beta_n)\sqrt{2\pi\log n}}\Bigg(W_\infty(\beta_n)+\frac{3}{\log n}\left(\frac{\kappa_4(\beta_n)}{24\sigma^4(\beta_n)}+\frac{\kappa_3(\beta_n)W_\infty'(\beta_n)}{6\sigma^4(\beta_n)}\right) - \frac{W_\infty''(\beta_n)}{2\sigma^2(\beta_n)\log n} - \frac{15\kappa_3^2(\beta_n)}{72\sigma^6(\beta_n)\log n}\Bigg) + o\left(\frac{1}{(\log n)^{\frac{3}{2}}}\right) \quad \text{a.s.},$$

and similarly, by Theorem 3.16 with $\beta = \beta_n$,

$$\frac{\mathrm{e}^{\beta_n k_n}}{n^{\varphi(\beta_n)}}\mathbb{E}[\mathbb{L}_n(k_n)] = \frac{1}{\sigma(\beta_n)\sqrt{2\pi\log n}}\Bigg(\mathbb{E}W_\infty(\beta_n)+\frac{3}{\log n}\left(\frac{\kappa_4(\beta_n)}{24\sigma^4(\beta_n)}+\frac{\kappa_3(\beta_n)\mathbb{E}W_\infty'(\beta_n)}{6\sigma^4(\beta_n)}\right) - \frac{\mathbb{E}W_\infty''(\beta_n)}{2\sigma^2(\beta_n)\log n} - \frac{15\kappa_3^2(\beta_n)}{72\sigma^6(\beta_n)\log n}\Bigg) + o\left(\frac{1}{(\log n)^{\frac{3}{2}}}\right).$$

Since $|c_n| \to \infty$, taking the difference of both expansions yields

$$\frac{(\log n)^{3/2}}{c_n}\frac{\mathrm{e}^{\beta_n k_n}}{n^{\varphi(\beta_n)}}\mathbb{L}_n^{\circ}(k_n) = \frac{(W_\infty(\beta_n)-\mathbb{E}W_\infty(\beta_n))\log n}{c_n\sigma(\beta_n)\sqrt{2\pi}} + o(1) \quad \text{a.s.} \tag{69}$$

Since $\beta_n \to 0$, we have, almost surely, $W_\infty(\beta_n) = m(0) + W_\infty'(0)\beta_n + o(\beta_n)$. Since $\mathbb{E}W_\infty(\beta)$ is analytic in a neighbourhood of $\beta = 0$, the analogous expansion holds for the mean. The assertion now follows from (69) together with (68). (Note that, even though the higher order terms in the Edgeworth expansion appearing in the proof are asymptotically irrelevant, we cannot obtain the result using the expansion for $r = 1$). □

Remark 3.26. In the setting of Theorem 3.25, part (c), the lim sup and the lim inf of the sequence $\frac{1}{n}(\log n)^{3/2}\mathbb{L}_n^{\circ}(k_n)$ are a.s. finite (but not necessarily equal to each other). Whether or not this sequence has an a.s. limit depends on the value of $\varphi'(0)$.

Case 1: $\varphi'(0) = 0$ (which applies to Example (vi) of Section 3.2). It is natural to take $k_n = a \in \mathbb{Z}$. Then, $c_n = a$ and we obtain

$$\frac{(\log n)^{\frac{3}{2}}}{n} \mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} R^\circ(a).$$

Case 2: $\varphi'(0) \neq 0$ (which applies to Examples (i)–(v) of Section 3.2). It is natural to take $k_n = \lfloor \varphi'(0) \log n \rfloor + a$, where $a \in \mathbb{Z}$, which means that $c_n = a - \{\varphi'(0) \log n\}$. The set of accumulation points of the sequence c_n is the interval $[a-1, a]$. Hence, we can parametrize the set of all a.s. subsequential limits of $\frac{1}{n}(\log n)^{3/2}\mathbb{L}_n^\circ(k_n)$ as follows:

$$\{R^\circ(a - z) \colon z \in [0, 1]\}. \tag{70}$$

Remark 3.27. Also we point out that in Theorem 3.25 the assumption that a BRW has deterministic number of descendants is used only to derive the Edgeworth expansion for the centering $\mathbb{E}[\mathbb{L}_n(k)]$ by using Theorem 3.16. If such an expansion holds *a priori*, the results of the above theorems remain valid without this constraint.

3.8. Profile of binary search trees around level $\log n$.

Applying the results of Section 3.7 to the special case of BSTs we obtain Equations (6), (7) and Theorem 1.9 stated in the introduction. In Equations (6), (7) (which deal with levels near $2e^\beta \log n$, $\beta \in \mathbb{R}$), the limit random variable is a multiple of $W_\infty(\beta)$. For $\beta = 0$, the limit $W_\infty(0) = 1$ is degenerate, and we collected more precise results describing the behavior of the profile around level $\varphi'(0)\log n = 2\log n$ in Theorem 1.9.

However, there is one more value of β for which $W_\infty(\beta)$ is degenerate, namely $\beta = -\log 2 \approx -0.693$. By construction of the BSTs, we have $W_n(-\log 2) = 1 = W_\infty(-\log 2)$ for all $n \in \mathbb{N}$. The value $\beta = -\log 2$ corresponds to the behavior of the BST profile around level $\varphi'(-\log 2)\log n = \log n$. We conclude this section with a discussion of this case. Similarly to Theorem 3.25, Fuchs et al. [20, Theorem 6] showed that the scaling behaviour of $\mathbb{L}_n(k_n)$ with $k_n = \log n + c_n$ depends drastically on whether $|c_n| \to \infty$ or $c_n = O(1)$. The next theorem is proved along the same lines as Theorem 3.25.

Theorem 3.28. *Let $(\mathbb{L}_n(k))_{k\in\mathbb{Z}}$ be the profile of a random binary search tree with $n+1$ external nodes. Let $(k_n)_{n\in\mathbb{N}}$ be an integer sequence.*

(a) *If $k_n = \log n + \alpha\sqrt{\log n} + o(\sqrt{\log n})$ with $\alpha \in \mathbb{R}$, then*

$$\frac{\log n}{2^{k_n}} \mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\chi_1(-\log 2) - \mathbb{E}[\chi_1(-\log 2)]}{\sqrt{2\pi}} \alpha e^{-\frac{1}{2}\alpha^2}.$$

(b) *If $k_n = \log n + c_n$, with $\lim_{n\to\infty} |c_n| = \infty$ and $c_n = o(\log n)$, then, with β_n as in* (57)*,*

$$\frac{(\log n)^{3/2}}{c_n n^{2e^{\beta_n}(1-\beta_n)-1}} \mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\chi_1(-\log 2) - \mathbb{E}[\chi_1(-\log 2)]}{\sqrt{2\pi}}.$$

In particular, if $\lim_{n\to\infty} |c_n| = \infty$ but $c_n = o(\sqrt{\log n})$, then

$$\frac{(\log n)^{3/2}}{c_n 2^{k_n}} \mathbb{L}_n^\circ(k_n) \xrightarrow[n\to\infty]{a.s.} \frac{\chi_1(-\log 2) - \mathbb{E}[\chi_1(-\log 2)]}{\sqrt{2\pi}}.$$

(c) *If* $k_n = \log n + c_n$, *where* c_n *is bounded, then*

$$\frac{(\log n)^{\frac{3}{2}}}{2^{k_n}} \mathbb{L}_n^\circ(k_n) - R_*^\circ(c_n) \underset{n\to\infty}{\overset{a.s.}{\longrightarrow}} 0,$$

where $R_*^\circ(c) = R_*(c) - \mathbb{E}R_*(c)$ *and* $R_*(c)$ *is a random variable given by*

$$R_*(c) := \frac{1}{\sqrt{2\pi}} \left(\chi_1(-\log 2) \left(c + \frac{1}{2} \right) - \frac{\chi_1^2(-\log 2) + \chi_2(-\log 2)}{2} \right), \quad c \in \mathbb{R}.$$

Remark 3.29. The random variable $\chi_1(-\log 2)$ is not almost surely constant: in the space of distributions with zero mean and finite variance, $\chi_1(-\log 2)$ is uniquely characterized by the stochastic fixed-point equation

$$\chi_1(-\log 2) \overset{d}{=} \frac{1}{2}\chi_1^{(1)}(-\log 2) + \frac{1}{2}\chi_1^{(2)}(-\log 2) + 1 + \frac{1}{2}\left(\log U + \log(1-U)\right), \tag{71}$$

where $\chi_1^{(1)}(-\log 2), \chi_1^{(2)}(-\log 2)$ are distributional copies of $\chi_1(-\log 2)$, U is uniformly distributed on $[0,1]$, and all three variables are independent. This follows from the arguments on page 35 in [20], see also display (35) in [9] for a less explicit variant of (71).

Let us finally mention that the random variable $W_\infty(\beta)$ is non-degenerate for all $\beta \in (\beta_-, \beta_+)$ except $\beta = 0$ and $\beta = -\log 2$. Indeed, we have the stochastic fixed point equation (see, e.g., (97) below)

$$\mathrm{e}^{-\beta} W_\infty(\beta) \overset{d}{=} U^{2\mathrm{e}^\beta - 1} W_{1,\infty}(\beta) + (1-U)^{2\mathrm{e}^\beta - 1} W_{2,\infty}(\beta),$$

where $W_{1,\infty}(\beta), W_{2,\infty}(\beta)$ are distributional copies of $W_\infty(\beta)$, U is uniformly distributed on $[0,1]$, and all three variables are independent. A constant random variable $W_\infty(\beta) = c > 0$ satisfies this equation if and only if $2\mathrm{e}^\beta - 1 \in \{0,1\}$. This corresponds to $\beta \in \{0, -\log 2\}$.

4. Proof of the general Edgeworth expansion

4.1. **Proof of Theorem 2.1.** The proof is based on studying the characteristic function of the profile. For notational reasons, we shall use μ and σ^2 as shorthands for $\mu(\beta)$ and $\sigma(\beta)$. Consider the following signed measure on $\mathbb{R}$: for $\beta \in (\beta_-, \beta_+)$,

$$\mu_n := \mu_n(\beta) := \sum_{k\in\mathbb{Z}} \mathrm{e}^{\beta k - \varphi(\beta) w_n} \mathbb{L}_n(k) \delta\left(\frac{k - \mu w_n}{\sigma\sqrt{w_n}} \right). \tag{72}$$

Here, $\delta(z)$ is the Dirac delta-measure at $z \in \mathbb{R}$. The characteristic function of μ_n has the form

$$\psi_n(s) := \psi_n(s,\beta) := \int_{\mathbb{R}} \mathrm{e}^{ist} \mu_n(\mathrm{d}t) = \mathrm{e}^{-\varphi(\beta) w_n - is\frac{\mu w_n}{\sigma\sqrt{w_n}}} \sum_{k\in\mathbb{Z}} \mathbb{L}_n(k) \mathrm{e}^{k\left(\beta + \frac{is}{\sigma\sqrt{w_n}}\right)}. \tag{73}$$

Fix some $\beta_0 \in (\beta_-, \beta_+)$ and *random* $\varepsilon_0 > 0$ such that $\mathbb{D}_{3\varepsilon_0}(\beta_0) \subset \mathscr{D}$ and W_∞ is non-zero on $\mathbb{D}_{3\varepsilon_0}(\beta_0)$. Here, $\mathbb{D}_r(\beta_0) = \{z \in \mathbb{C} \colon |z - \beta_0| < r\}$ denotes an open disk with radius r centered at β_0. For any $\beta \in \mathscr{I}_0 := (\beta_0 - \varepsilon_0, \beta_0 + \varepsilon_0)$, we have $\mathbb{D}_{2\varepsilon_0}(\beta) \subset \mathscr{D}$. In the following, all estimates are going to be uniform in $\beta \in \mathscr{I}_0$. Since any compact set $K \subset (\beta_-, \beta_+)$ can be covered by finitely many such intervals $\mathscr{I}_0$, the uniformity in $\beta \in K$ follows. After recalling the definition of W_n, see (9), we obtain that, for all $\beta \in \mathscr{I}_0$, as long as the variable $s \in \mathbb{R}$ satisfies

$$\left| \frac{s}{\sigma\sqrt{w_n}} \right| < \varepsilon_0,$$

the function ψ_n is well-defined and can be written in the form

$$\psi_n(s) = \mathrm{e}^{-\varphi(\beta)w_n - is\frac{\mu w_n}{\sigma\sqrt{w_n}} + w_n\varphi\left(\beta+\frac{is}{\sigma\sqrt{w_n}}\right)} W_n\left(\beta + \frac{is}{\sigma\sqrt{w_n}}\right). \tag{74}$$

Our aim is to derive an asymptotic expansion of $\psi_n(s)$ in powers of $w_n^{-1/2}$. Consider a modification of $\psi_n(s)$ in which W_n is replaced by W_∞ and $w_n^{-1/2}$ is replaced by a new variable u. For any fixed $s \in \mathbb{R}$ and $\beta \in \mathscr{I}_0$, the function

$$\tilde{\psi}(s;u) = \exp\left\{-\frac{\varphi(\beta)}{u^2} - is\frac{\mu}{\sigma u} + \frac{1}{u^2}\varphi\left(\beta + \frac{isu}{\sigma}\right) + \log W_\infty\left(\beta + \frac{isu}{\sigma}\right)\right\} \tag{75}$$

is well-defined and analytic in u in the disk $|u| < \sigma\varepsilon_0/|s|$. Note that $\log W_\infty$ is defined as an analytic function because W_∞ does not vanish on $\mathbb{D}_{3\varepsilon_0}(\beta_0)$. Thus, as long as $\left|\frac{su}{\sigma}\right| < \varepsilon_0$,

$$\log\tilde{\psi}(s;u) = \sum_{k=0}^{\infty}\frac{a_k(s)}{k!}u^k,$$

where

$$a_k(s) = a_k(s,\beta) := \frac{\varphi^{(k+2)}(\beta)}{(k+2)(k+1)}\left(\frac{is}{\sigma}\right)^{k+2} + \chi_k(\beta)\left(\frac{is}{\sigma}\right)^k. \tag{76}$$

Recall from the definition of Bell polynomials, see (18), that there is a formal identity

$$\exp\left\{\sum_{k=1}^{\infty}\frac{a_k}{k!}x^k\right\} = \sum_{k=0}^{\infty}\frac{B_k(a_1,\dots,a_k)}{k!}x^k.$$

It follows that the following holds (not only formally!) for $|u| < \varepsilon_0\sigma/|s|$:

$$\tilde{\psi}(s;u) = W_\infty(\beta)\mathrm{e}^{-\frac{s^2}{2}}\sum_{k=0}^{\infty}\frac{B_k(a_1(s),\dots,a_k(s))}{k!}u^k. \tag{77}$$

To see that (77) holds not only formally, note that $\tilde{\psi}(s;u)$, being an analytic function of u in the disk $|u| < \varepsilon_0\sigma/|s|$, has a convergent Taylor expansion. But in order to compute the coefficients of this expansion, we can use formal series. We shall need a uniform estimate for the remainder term in (77).

Lemma 4.1. *Recall that $a_k(s)$ is given by* (76)*. There exists an a.s. finite random variable $M > 0$ such that, for all $\beta \in \mathscr{I}_0$,*

$$\left|\frac{a_k(s)}{k!}\right| \le M^k(|s|+1)^{k+2}$$

for all $s \in \mathbb{R}$ and $k \in \mathbb{N}$.

Proof. Since the functions φ and $\log W_\infty$ are analytic on the disk $\mathbb{D}_{2\varepsilon_0}(\beta_0)$, the Cauchy formula implies that, for $\beta \in \mathscr{I}_0$,

$$\left|\frac{\varphi^{(k+2)}(\beta)}{(k+2)!}\right| \le \sup_{\gamma\in\mathbb{D}_{\varepsilon_0}(\beta_0)}|\varphi(\gamma)|\varepsilon_0^{-k-2}, \quad \left|\frac{\chi_k(\beta)}{k!}\right| \le \sup_{\gamma\in\mathbb{D}_{\varepsilon_0}(\beta_0)}|\log W_\infty(\gamma)|\varepsilon_0^{-k},$$

for all $k \in \mathbb{N}$. With $M' = \max\{1, \sup_{\gamma\in\mathbb{D}_{\varepsilon_0}(\beta_0)}|\varphi(\gamma)|, \sup_{\gamma\in\mathbb{D}_{\varepsilon_0}(\beta_0)}|\log W_\infty(\gamma)|\}$, and $C = \max(1, \sup_{\gamma\in\mathbb{D}_{\varepsilon_0}(\beta_0)}(\varepsilon_0\sigma(\gamma))^{-1})$, it follows from (76) that

$$\left|\frac{a_k(s)}{k!}\right| \le \left|\frac{s}{\sigma}\right|^{k+2}\left|\frac{\varphi^{(k+2)}(\beta)}{(k+2)!}\right| + \left|\frac{s}{\sigma}\right|^k\left|\frac{\chi_k(\beta)}{k!}\right| \le M'(|s|^{k+2}C^{k+2} + |s|^kC^k)$$

which yields the desired estimate choosing $M = M'C^3$. □

Lemma 4.2. *There is an a.s. finite random variable $M_1 > 0$ such that, for all $\beta \in \mathscr{I}_0$,*

$$\frac{1}{k!}\,|B_k(a_1(s),\dots,a_k(s))| \le M_1^k(|s|+1)^{3k}$$

for all $k \in \mathbb{N}$ and $s \in \mathbb{R}$.

Proof. By definition of the Bell polynomial B_k, see (19),

$$\begin{aligned}\frac{1}{k!}\,|B_k(a_1(s),\dots,a_k(s))| &\le \sum\nolimits' \frac{1}{j_1!\dots j_k!}\left|\frac{a_1(s)}{1!}\right|^{j_1}\dots\left|\frac{a_k(s)}{k!}\right|^{j_k}\\ &\le \sum\nolimits' \frac{1}{j_1!\dots j_k!} M^{1j_1+\dots+kj_k}(|s|+1)^{\sum_{m=1}^k (m+2)j_m},\end{aligned}$$

where the sum $\sum'$ is taken over all $j_1,\dots,j_k \in \mathbb{N}_0$ satisfying $1j_1+2j_2+\dots+kj_k = k$. Using that $1j_1+\dots+kj_k = k$ (and consequently $j_1+\dots+j_k \le k$) and the inequality $\sum \frac{1}{j_1!\dots j_k!} \le \mathrm{e}^k$, we obtain the required estimate choosing $M_1 = \mathrm{e}M$. □

Lemma 4.3. *Fix $r \in \mathbb{N}_0$. There exist a.s. finite random variables $U > 0$ and $M_2 > 0$ such that for all $\beta \in \mathscr{I}_0, u \in (-U,U)$ and $s \in \mathbb{R}$ with $1+|s| < u^{-1/4}$, we have*

$$\left|\tilde\psi(s;u) - W_\infty(\beta)\mathrm{e}^{-\frac12 s^2}\sum_{k=0}^r \frac{B_k(a_1(s),\dots,a_k(s))}{k!}u^k\right| \le M_2\mathrm{e}^{-\frac12 s^2}(1+|s|)^{3r+3}|u|^{r+1}.$$

Proof. Using formula (77) for $\tilde\psi(s;u)$ and then Lemma 4.2 we obtain

$$\begin{aligned}\text{LHS} &\le |W_\infty(\beta)|\mathrm{e}^{-\frac12 s^2}\sum_{k=r+1}^\infty \frac{|B_k(a_1(s),\dots,a_k(s))|}{k!}|u|^k\\ &\le |W_\infty(\beta)|\mathrm{e}^{-\frac12 s^2}\sum_{k=r+1}^\infty M_1^k(|s|+1)^{3k}|u|^k\\ &\le \frac{M_2}{2}\mathrm{e}^{-\frac12 s^2}(|s|+1)^{3r+3}|u|^{r+1}\sum_{k=0}^\infty M_1^k(|s|+1)^{3k}|u|^k,\end{aligned}$$

where $M_2 = 2M_1^{r+1}\sup_{\gamma\in\mathbb{D}_{\varepsilon_0}(\beta_0)}|W_\infty(\gamma)|$. The sum on the right-hand side can be estimated using the assumptions $1+|s| < u^{-1/4}$ and $|u| < U$ as follows:

$$\sum_{k=0}^\infty M_1^k(|s|+1)^{3k}|u|^k \le \sum_{k=0}^\infty M_1^k|u|^{-\frac34 k}|u|^k \le \sum_{k=0}^\infty M_1^k U^{k/4} = 2,$$

where the last step holds if we choose $U = (16M_1^4)^{-1}$. □

We are now able to state the expansion for the characteristic function ψ_n with an estimate for the remainder term. Let

$$V_{r,n}(s) = W_\infty(\beta)\mathrm{e}^{-\frac12 s^2}\sum_{k=0}^r \frac{B_k(a_1(s),\dots,a_k(s))}{k!}w_n^{-\frac{k}{2}} \tag{78}$$

Lemma 4.4. *There exist a.s. finite numbers $K > 0$ and $M_3 > 0$ such that*

$$|\psi_n(s) - V_{r,n}(s)| \le M_3\mathrm{e}^{-\frac12 s^2}(|s|+1)^{3r+3}w_n^{-\frac{r+1}{2}}.$$

for all $\beta \in \mathscr{I}_0, n > K$ and $s \in \mathbb{R}$ satisfying $1+|s| < w_n^{1/8}$.

Proof. We have

$$\text{LHS} \leq \left|\tilde{\psi}(s; w_n^{-\frac{1}{2}}) - V_{r,n}(s)\right| + \left|\psi_n(s) - \tilde{\psi}(s; w_n^{-\frac{1}{2}})\right|. \tag{79}$$

We estimate the terms on the right-hand side in two steps.

STEP 1. We start with the first term on the RHS in (79). By Lemma 4.3 with $u = w_n^{-1/2}$, the estimate

$$\left|\tilde{\psi}(s; w_n^{-\frac{1}{2}}) - V_{r,n}(s)\right| \leq M_2 \mathrm{e}^{-\frac{1}{2}s^2}(|s|+1)^{3k+3} w_n^{-\frac{r+1}{2}} \tag{80}$$

holds provided that $w_n^{-1/2} < U$ and $1 + |s| < w_n^{1/8}$. Since $\lim_{n\to\infty} w_n = +\infty$, we can choose a random variable K such that $w_n^{-1/2} < U$ for $n > K$.

STEP 2. We estimate the second term on the RHS in (79). Let $z_n = \frac{is}{\sigma\sqrt{w_n}}$ so that for sufficiently large n, we have $|z_n| < \varepsilon_0$. With this notation, we have

$$\left|\psi_n(s) - \tilde{\psi}(s; w_n^{-\frac{1}{2}})\right| = \left|\mathrm{e}^{w_n\left(\varphi(\beta+z_n) - \varphi(\beta) - \varphi'(\beta)z_n\right)}\right| \left|W_\infty(\beta+z_n) - W_n(\beta+z_n)\right|.$$

By Assumption A3, see (10), we have, for some a.s. finite number M' depending on β_0 and ε_0 but not on β,

$$|W_\infty(\beta+z_n) - W_n(\beta+z_n)| \leq \sup_{z\in\mathbb{D}_{2\varepsilon_0}(\beta_0)} |W_\infty(z) - W_n(z)| < M' w_n^{-\frac{r+1}{2}}.$$

By the Taylor expansion of φ at β, we obtain the following estimate in which the O-term is uniform as long as $|z_n| < \varepsilon_0$ and $\beta \in \mathscr{I}_0$:

$$\begin{aligned} w_n\left(\varphi(\beta+z_n) - \varphi(\beta) - \varphi'(\beta)z_n\right) &= \left(\frac{\sigma^2}{2}z_n^2 + O(z_n^3)\right) w_n \\ &= -\frac{s^2}{2} + O\left(\frac{s^3}{\sqrt{w_n}}\right) \leq -\frac{s^2}{2} + O(w_n^{-1/8}), \end{aligned} \tag{81}$$

where in the last step we used the restriction $1 + |s| < w_n^{1/8}$. Combining the above estimates we obtain

$$\left|\psi_n(s) - \tilde{\psi}(s; w_n^{-\frac{1}{2}})\right| \leq M' w_n^{-\frac{r+1}{2}} \left(\mathrm{e}^{-\frac{1}{2}s^2 + O(w_n^{-1/8})}\right). \tag{82}$$

Taking (80) and (82) together, we obtain the statement of the lemma. □

In order to obtain the Edgeworth expansion for $\mathbb{L}_n(k)$ we shall apply Fourier inversion to the expansion for ψ_n established above. Recall formula (73) for the characteristic function ψ_n. It follows by Fourier inversion that

$$\sigma\sqrt{w_n}\mathrm{e}^{\beta k - \varphi(\beta)w_n}\mathbb{L}_n(k) = \frac{1}{2\pi}\int_{-\pi\sigma\sqrt{w_n}}^{\pi\sigma\sqrt{w_n}} \psi_n(s)\mathrm{e}^{-isx_n(k)}\mathrm{d}s,$$

where $x_n(k)$ was defined in (13).

Lemma 4.5. *Recall from* (78) *the definition of* $V_{r,n}$. *For every fixed* $r \in \mathbb{N}_0$,

$$w_n^{\frac{r}{2}} \sup_{k\in\mathbb{Z}} \sup_{\beta\in\mathscr{I}_0} \left| \int_{-\pi\sigma\sqrt{w_n}}^{\pi\sigma\sqrt{w_n}} \psi_n(s)\mathrm{e}^{-isx_n(k)}\mathrm{d}s - \int_{\mathbb{R}} V_{r,n}(s)\mathrm{e}^{-isx_n(k)}\mathrm{d}s \right| \xrightarrow[n\to\infty]{a.s.} 0.$$

Proof. Step 1. We show that

$$w_n^{\frac{r}{2}} \sup_{\beta\in\mathscr{I}_0} \int_{-w_n^{1/9}}^{w_n^{1/9}} |\psi_n(s) - V_{r,n}(s)|\,\mathrm{d}s \xrightarrow[n\to\infty]{a.s.} 0.$$

Indeed, we know from Lemma 4.4 that, for all $\beta \in \mathscr{I}_0$,

$$|\psi_n(s) - V_{r,n}(s)| \le M_3 \mathrm{e}^{-\frac{1}{2}s^2}(|s|+1)^{3r+3} w_n^{-\frac{r+1}{2}}$$

for $n > K$, $1 + |s| < w_n^{1/8}$. Integrating this, we obtain the required estimate.

Step 2. We show that there is an $a > 0$ such that

$$w_n^{\frac{r}{2}} \sup_{\beta\in\mathscr{I}_0} \int_{|w_n|^{1/9}<|s|<a\sqrt{w_n}} |\psi_n(s)|\mathrm{d}s \xrightarrow[n\to\infty]{a.s.} 0. \tag{83}$$

Let $z_n = \frac{is}{\sigma\sqrt{w_n}}$. We can choose $a > 0$ so small that $|z_n| < \varepsilon_0$ provided that $|s| < a\sqrt{w_n}$. From the uniform convergence of W_n to W_∞ on $\mathbb{D}_{2\varepsilon_0}(\beta_0)$ and from the Taylor series for φ we infer

$$|\psi_n(s)| = \left|\mathrm{e}^{w_n\left(\varphi(\beta+z_n)-\varphi(\beta)-\varphi'(\beta)z_n\right)}\right| |W_n(\beta + z_n)| \le M' \mathrm{e}^{-\frac{1}{2}s^2},$$

for some a.s. finite $M' > 0$ depending on β_0 and ε_0 but not on β. It follows that

$$\sup_{\beta\in\mathscr{I}_0} \int_{|w_n|^{1/9}<|s|<a\sqrt{w_n}} |\psi_n(s)|\mathrm{d}s \le M' \int_{|w_n|^{1/9}<|s|<a\sqrt{w_n}} \mathrm{e}^{-\frac{1}{2}s^2}\mathrm{d}s = o(w_n^{-\frac{r}{2}}) \quad \text{a.s.}$$

This completes the proof of (83).

Step 3. We prove that for every $a > 0$,

$$w_n^{\frac{r}{2}} \sup_{\beta\in\mathscr{I}_0} \int_{a\sqrt{w_n}<|s|<\sigma\pi\sqrt{w_n}} |\psi_n(s)|\mathrm{d}s \xrightarrow[n\to\infty]{a.s.} 0. \tag{84}$$

In this case, $z_n = \frac{is}{\sigma\sqrt{w_n}}$ need not satisfy $|z_n| \le \varepsilon_0$ so that W_n need not converge (nor even be well-defined) and the estimate from Step 2 does not work. Instead, we shall use Assumption A4. Using the definition of ψ_n, see (73),

$$\begin{aligned}\int_{a\sqrt{w_n}<|s|<\sigma\pi\sqrt{w_n}} |\psi_n(s)|\mathrm{d}s &= \mathrm{e}^{-\varphi(\beta)w_n} \int_{a\sqrt{w_n}<|s|<\sigma\pi\sqrt{w_n}} \left|\sum_{k\in\mathbb{Z}} \mathbb{L}_n(k)\mathrm{e}^{k\left(\beta+\frac{is}{\sigma\sqrt{w_n}}\right)}\right| \mathrm{d}s \\ &= \mathrm{e}^{-\varphi(\beta)w_n}\sigma\sqrt{w_n} \int_{a/\sigma<|u|<\pi} \left|\sum_{k\in\mathbb{Z}} \mathbb{L}_n(k)\mathrm{e}^{k(\beta+iu)}\right| \mathrm{d}u,\end{aligned}$$

so that (84) is implied by Assumption A4 since σ is bounded on $\mathscr{I}_0$.

The same estimates as (83) and (84), but with $V_{r,n}(s)$ instead of $\psi_n(s)$, hold since $V_{r,n}$ is a product of $\mathrm{e}^{-s^2/2}$ and a polynomial in s. Combining pieces together we obtain the claim of the lemma. □

To complete the proof of Theorem 2.1 it remains to show that

$$\int_{\mathbb{R}} V_{r,n}(s)\mathrm{e}^{-isz}\mathrm{d}s = \sqrt{2\pi}W_\infty(\beta)\mathrm{e}^{-\frac{1}{2}z^2} \sum_{k=0}^{r} \frac{G_k(z)}{w_n^{k/2}}, \quad z \in \mathbb{R},$$

which, in turn, amounts to

$$\frac{1}{k!}\int_{\mathbb{R}} B_k(a_1(s),\dots,a_k(s))\mathrm{e}^{isx}\mathrm{e}^{-\frac{1}{2}s^2}\mathrm{d}s = \sqrt{2\pi}G_k(-x)\mathrm{e}^{-\frac{1}{2}x^2}, \quad x\in\mathbb{R}, \tag{85}$$

for every $k\in\mathbb{N}_0$. To check (85) note that

$$B_k(a_1(s),\dots,a_k(s))\mathrm{e}^{isx} = B_k(D_1,\dots,D_k)(\mathrm{e}^{isx}), \quad s\in\mathbb{R},$$

where the differential operators $D_1,\dots,D_k$ are given by (17). This yields

$$\begin{aligned}\int_{\mathbb{R}} B_k(a_1(s),\dots,a_k(s))\mathrm{e}^{isx}\mathrm{e}^{-\frac{1}{2}s^2}\mathrm{d}s &= B_k(D_1,\dots,D_k)\left(\int_{\mathbb{R}}\mathrm{e}^{isx}\mathrm{e}^{-\frac{1}{2}s^2}\mathrm{d}s\right)\\ &= \sqrt{2\pi}B_k(D_1,\dots,D_k)\mathrm{e}^{-\frac{1}{2}x^2}\\ &= \sqrt{2\pi}(-1)^k k!\mathrm{e}^{-\frac{1}{2}x^2}G_k(x).\end{aligned}$$

Formula (85) now follows from the observation $(-1)^kG_k(x)=G_k(-x)$, $k\in\mathbb{N}_0$, see Remark 2.4. The proof of Theorem 2.1 is complete.

4.2. **Alternative expression for $G_j(x;\beta)$.** In this section we show that upon multiplying by $W_\infty(\beta)$ the functions $x\mapsto G_j(x;\beta)$, $j\in\mathbb{N}_0$, become polynomials in x whose coefficients are linear combinations of $1, W_\infty(\beta),\dots,W_\infty^{(j)}(\beta)$. Write $D=\frac{1}{\sigma(\beta)}\frac{\mathrm{d}}{\mathrm{d}x}$ and recall from (17) that

$$D_k = \frac{\varphi^{(k+2)}(\beta)}{(k+1)(k+2)}D^{k+2} + \chi_k(\beta)D^k.$$

From now on we consider D as a formal variable rather than a differential operator. Formula (16) shows that $W_\infty(\beta)G_j(x;\beta)$ can be obtained from the expression $(-1)^j/j!W_\infty(\beta)B_j(D_1,\dots,D_j)$ by replacing each D^k-term by the Hermite polynomial $(-1)^k\mathrm{He}_k(x)$. Therefore, it suffices to show that this expression is a polynomial in D whose coefficients are linear combinations of $1, W_\infty(\beta),\dots,W_\infty^{(j)}(\beta)$. By the definition of the Bell polynomials, see (18), we have

$$\frac{1}{j!}B_j(D_1,\dots,D_j) = [y^j]\exp\left\{\sum_{k=1}^{\infty}\frac{y^k}{k!}D_k\right\},$$

where $[y^j]f(y)$ denotes the coefficient of y^j in the formal power series $f(y)$. It follows that, in the sense of formal power series,

$$\frac{1}{j!}B_j(D_1,\dots,D_j) = [y^j]\left(\exp\left\{\sum_{k=1}^{\infty}\frac{y^k\varphi^{(k+2)}(\beta)}{(k+2)!}D^{k+2}\right\}\exp\left\{\sum_{k=1}^{\infty}\frac{(yD)^k}{k!}\chi_k(\beta)\right\}\right).$$

Multiplying both sides by $W_\infty(\beta)=\mathrm{e}^{\chi_0(\beta)}$ and observing that, by Taylor's expansion,

$$W_\infty(\beta+yD) = \exp\left\{\sum_{k=0}^{\infty}\frac{(yD)^k}{k!}\chi_k(\beta)\right\},$$

we obtain

$$\begin{aligned}
&\frac{W_\infty(\beta)}{j!}B_j(D_1,\dots,D_j)\\
&=[y^j]\left(\exp\left\{\sum_{k=1}^{\infty}\frac{y^k\varphi^{(k+2)}(\beta)}{(k+2)!}D^{k+2}\right\}W_\infty(\beta+yD)\right)\\
&=[y^j]\left(\exp\left\{\sum_{k=1}^{\infty}\frac{y^k\varphi^{(k+2)}(\beta)}{(k+2)!}D^{k+2}\right\}\left(\sum_{k=0}^{\infty}\frac{(yD)^k}{k!}W_\infty^{(k)}(\beta)\right)\right).
\end{aligned}$$

Clearly, the right-hand side is a polynomial in D whose coefficients are linear combinations of $1, W_\infty(\beta),\dots,W_\infty^{(j)}(\beta)$.

4.3. **Proofs of Theorems 2.10, 2.12, 2.13 and Proposition 3.19.** Our proof runs along the same lines as the proof of Theorem 2.17 in Grübel and Kabluchko [21]. In order to keep this paper self-contained we present all details.

The aim is to search for $k\in\mathbb{Z}$ maximizing the profile $\mathbb{L}_n(k)$. Write $k=k_n(a)=\varphi'(0)w_n+a$, where $a\in\mathbb{Z}-\varphi'(0)w_n$. We use Theorem 2.1 with $\beta=0$ and $r=2$. Inserting this k into (15) we obtain

$$\begin{aligned}
&\sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(k)=W_\infty(0)\mathrm{e}^{-\frac{a^2}{2\sigma^2(0)w_n}}\times\\
&\left(1+\frac{1}{\sqrt{w_n}}G_1\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right)+\frac{1}{w_n}G_2\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right)+o\left(\frac{1}{w_n}\right)\right),
\end{aligned}\tag{86}$$

where the o-term is uniform in a. This and similar expansions below hold a.s.

Step 1. Let us assume that $|a|<w_n^{1/4-\varepsilon}$ for some small $\varepsilon>0$, say $\varepsilon=1/100$. Since $x_n(k)=a/(\sigma(0)\sqrt{w_n})$ and $|a|<w_n^{1/4-\varepsilon}$ we have

$$x_n(k)=o(1)\quad\text{and}\quad x_n^3(k)=o\left(\frac{1}{\sqrt{w_n}}\right).$$

Hence, by (22),

$$G_1\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right)=\frac{a}{\sqrt{w_n}}\left(\frac{\chi_1(0)}{\sigma^2(0)}-\frac{\kappa_3(0)}{2\sigma^4(0)}\right)+o\left(\frac{1}{\sqrt{w_n}}\right),$$

and by (23),

$$\begin{aligned}
G_2\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right)&=-\frac{\chi_1^2(0)+\chi_2(0)}{2\sigma^2(0)}+3\left(\frac{\kappa_4(0)}{24\sigma^4(0)}+\frac{\kappa_3(0)\chi_1(0)}{6\sigma^4(0)}\right)-\frac{15\kappa_3^2(0)}{72\sigma^6(0)}+o(1)\\
&=:C+o(1).
\end{aligned}$$

By a standard Taylor expansion, we get

$$\mathrm{e}^{-\frac{a^2}{2\sigma^2(0)w_n}}=1-\frac{a^2}{2\sigma^2(0)w_n}+o\left(\frac{1}{w_n}\right).$$

Inserting these expansions into (86), we arrive at

$$\begin{aligned}
&\sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(k)\\
&=W_\infty(0)\left(1-\left(\frac{a^2}{2\sigma^2(0)}-\left(\frac{\chi_1(0)}{\sigma^2(0)}-\frac{\kappa_3(0)}{2\sigma^4(0)}\right)a-C\right)\frac{1}{w_n}\right)+o\left(\frac{1}{w_n}\right).
\end{aligned}\tag{87}$$

Differentiation with respect to a shows that the maximum is attained at

$$a_* = \chi_1(0) - \frac{\kappa_3(0)}{2\sigma^2(0)}. \tag{88}$$

Using (87) and (88), we obtain

$$\sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}(\mathbb{L}_n(k+1)-\mathbb{L}_n(k)) = \frac{W_\infty(0)}{\sigma^2(0)w_n}\left(a_* - \frac{1}{2} - a\right) + o\left(\frac{1}{w_n}\right). \tag{89}$$

Put $u_n^* = \varphi'(0)w_n + a_*$. From (89) it is clear that, for all sufficiently large n, the value u_n maximizing $\mathbb{L}_n(\cdot)$ in the region $|a| < w_n^{1/4-\varepsilon}$ is either $\lfloor u_n^* \rfloor$ or $\lceil u_n^* \rceil$.

Step 2. Note that $a_* = O(1)$ a.s. If $\lfloor u_n^* \rfloor$ or $\lceil u_n^* \rceil$ is indeed the mode u_n of $\mathbb{L}_n(\cdot)$ (over the whole range $k \in \mathbb{Z}$), then, from (87), we deduce

$$\sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(u_n) = W_\infty(0) + O\left(\frac{1}{w_n}\right),$$

and Theorem 2.12 follows. To complete the proof of Theorem 2.10 it remains to show that (for n large enough) the mode cannot lie in the region $|a| \geq w_n^{1/4-\varepsilon}$. To this end we shall show that for every $B > 0$ there exists an a.s. finite $K \in \mathbb{N}$ such that

$$W_\infty(0) - \sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(k) > \frac{B}{w_n} \tag{90}$$

for all $n > K$ and $|a| \geq w_n^{1/4-\varepsilon}$.

Step 3. Let $|a| > \sigma(0)\sqrt{w_n}$. By Theorem 2.6,

$$W_\infty(0) - \sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(k) = W_\infty(0)\left(1 - \mathrm{e}^{-\frac{a^2}{\sigma^2(0)w_n}}\right) + o(1).$$

Since $W_\infty(0) > 0$ a.s., the expression on the right-hand side is larger than $W_\infty(0)/10$ for all n sufficiently large. Hence, there exists an a.s. finite $K_1 \in \mathbb{N}$ such that (90) holds for $|a| > \sigma(0)\sqrt{w_n}$ and $n > K_1$.

Step 4. Let $\sigma(0)w_n^{3/8+\varepsilon} < |a| \leq \sigma(0)\sqrt{w_n}$ for some small $\varepsilon > 0$. By Theorem 2.1 with $r = 1$ and $\beta = 0$, we have

$$\begin{aligned}&\sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(k)\\ &\quad= W_\infty(0)\mathrm{e}^{-\frac{a^2}{2\sigma^2(0)w_n}}\left(1 + \frac{1}{\sqrt{w_n}}G_1\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right)\right) + o\left(\frac{1}{\sqrt{w_n}}\right) \quad \text{a.s..}\end{aligned}$$

Since $|a| \leq \sigma(0)\sqrt{w_n}$ we have $G_1\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right) = O(1)$ a.s. Therefore,

$$\begin{aligned}W_\infty(0) - \sigma(0)\sqrt{2\pi w_n}\mathrm{e}^{-\varphi(0)w_n}\mathbb{L}_n(k) &= W_\infty(0)\left(1 - \mathrm{e}^{-\frac{a^2}{2\sigma^2(0)w_n}}\right) + O\left(\frac{1}{\sqrt{w_n}}\right)\\ &\geq W_\infty(0)\left(1 - \mathrm{e}^{-\frac{1}{2}w_n^{-1/4+2\varepsilon}}\right) + O\left(\frac{1}{\sqrt{w_n}}\right).\end{aligned}$$

From the elementary inequality

$$1 - \mathrm{e}^{-y/2} > y/3, \quad y \in [0,1], \tag{91}$$

it now easily follows that there exist an a.s. finite $K_2 \in \mathbb{N}$ such that (90) is satisfied for $\sigma(0)w_n^{3/8+\varepsilon} < |a| \le \sigma(0)\sqrt{w_n}$ and $n > K_2$.

STEP 5. Finally, let $\sigma(0)w_n^{1/4-2\varepsilon} < |a| \le \sigma(0)w_n^{3/8+\varepsilon}$. Using (86) and noting that

$$\frac{1}{\sqrt{w_n}} G_1\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right) = O(w_n^{-\frac{5}{8}+\varepsilon}), \quad \frac{1}{w_n} G_2\left(\frac{a}{\sigma(0)\sqrt{w_n}}\right) = O\left(\frac{1}{w_n}\right) \quad \text{a.s.,}$$

we obtain

$$\begin{aligned} W_\infty(0) - \sigma(0)\sqrt{2\pi w_n} \mathrm{e}^{-\varphi(0)w_n} \mathbb{L}_n(k) &= W_\infty(0)\left(1 - \mathrm{e}^{-\frac{a^2}{2\sigma^2(0)w_n}}\right) + O\left(w_n^{-\frac{5}{8}+\varepsilon}\right) \\ &\ge \frac{1}{3} W_\infty(0) w_n^{-1/2-4\varepsilon} + O\left(w_n^{-\frac{5}{8}+\varepsilon}\right), \end{aligned}$$

where the second inequality follows from (91) and holds for all n sufficiently large. Hence, upon choosing ε sufficiently small, there exists an a.s. finite $K_3 \in \mathbb{N}$ such that (90) holds for all $n > K_3$ and $\sigma(0)w_n^{1/4-2\varepsilon} < |a| \le \sigma(0)w_n^{3/8+\varepsilon}$. The proof of Theorem 2.10 is complete.

STEP 6. It remains to prove Theorem 2.13. From what we have already proved, it follows that, for all n sufficiently large, the mode u_n can be written as

$$u_n = u_n^* + \gamma_n = \varphi'(0)w_n + a_* + \gamma_n, \quad |\gamma_n| < 1,$$

where γ_n equals either $\lfloor u_n^* \rfloor - u_n^*$ or $\lceil u_n^* \rceil - u_n^*$. Inserting u_n into (87), we obtain

$$\begin{aligned} &\sigma(0)\sqrt{2\pi w_n} \mathrm{e}^{-\varphi(0)w_n} \mathbb{L}_n(u_n) \\ &\quad = W_\infty(0)\left(1 - \left(\frac{(a_* + \gamma_n)^2}{2\sigma^2(0)} - \frac{a_*(a_* + \gamma_n)}{\sigma^2(0)} - C\right)\frac{1}{w_n}\right) + o\left(\frac{1}{w_n}\right) \\ &\quad = W_\infty(0)\left(1 + \frac{a_*^2 - \gamma_n^2}{2\sigma^2(0)w_n} + \frac{C}{w_n}\right) + o\left(\frac{1}{w_n}\right) \quad \text{a.s.} \end{aligned}$$

Since u_n is the mode and hence maximizes the left-hand side, γ_n in the above relation can be replaced by the $\theta_n := \min\{u_n^* - \lfloor u_n^* \rfloor, \lceil u_n^* \rceil - u_n^*\}$. Upon rearranging the terms and recalling the notation $\tilde{M}_n$ from Theorem 2.13, this becomes

$$\tilde{M}_n - \theta_n^2 = -a_*^2 - 2\sigma^2(0)C + o(1) = \chi_2(0) + \frac{\kappa_3^2(0)}{6\sigma^4(0)} - \frac{\kappa_4(0)}{4\sigma^2(0)} + o(1) \quad \text{a.s.,}$$

where the last equality follows from (88) and the definition of C. This concludes the proof of Theorem 2.13.

Proof of Proposition 3.19. Both assertions follow from properties of the logarithm. The first claim (i) follows immediately from the fact that, for every fixed $L > 0$, we have $\log(n + L) - \log n \to 0$ as $n \to \infty$. To show (ii), it is sufficient to verify that, almost surely,

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\#\{1 \le k \le n : \operatorname{dist}(u_k^*, \mathbb{Z} + 1/2) < \varepsilon\}}{n} = 0.$$

Since $\varphi'(0) \ne 0$, using the explicit expression (54), the claim follows if, for any $\alpha > 0$ and $\beta \in \mathbb{R}$,

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\#\{1 \le k \le n : \operatorname{dist}(\log k, \alpha\mathbb{Z} + \beta) < \varepsilon\}}{n} = 0.$$

Let us show how to estimate the numerator of the last fraction. We have, assuming that $\varepsilon < \alpha/2$,

$$\begin{aligned}\#\{1 \le k \le n : \mathrm{dist}(\log k, \alpha\mathbb{Z} + \beta) < \varepsilon\} &= \sum_{k=1}^{n} \#\{j \in \mathbb{Z} : \mathrm{dist}(\log k, \alpha j + \beta) < \varepsilon\} \\ &= \sum_{j\in\mathbb{Z}} \#\{1 \le k \le n : \mathrm{e}^{\alpha j+\beta-\varepsilon} < k < \mathrm{e}^{\alpha j+\beta+\varepsilon}\} \\ &\le \sum_{j\in\mathbb{Z}} \#\{k \in \mathbb{N} : \mathrm{e}^{\alpha j+\beta-\varepsilon} \vee 1 \le k \le \mathrm{e}^{\alpha j+\beta+\varepsilon} \wedge n\}.\end{aligned}$$

The inner sum on the right-hand side is the number of integers in the interval $[\mathrm{e}^{\alpha j+\beta-\varepsilon} \vee 1, \mathrm{e}^{\alpha j+\beta+\varepsilon} \wedge n]$ (which is empty if either $\mathrm{e}^{\alpha j+\beta-\varepsilon} > n$ or $\mathrm{e}^{\alpha j+\beta+\varepsilon} < 1$) and hence is bounded from above by $\left(\mathrm{e}^{\alpha j+\beta+\varepsilon} \wedge n - \mathrm{e}^{\alpha j+\beta-\varepsilon} \vee 1 + 1\right)_+$. Therefore,

$$\#\{1 \le k \le n : \mathrm{dist}(\log k, \alpha\mathbb{Z} + \beta) < \varepsilon\}| \le \sum_{j\in\mathbb{Z}} \left(\mathrm{e}^{\alpha j+\beta+\varepsilon} \wedge n - \mathrm{e}^{\alpha j+\beta-\varepsilon} \vee 1 + 1\right)_+ .$$

The relation

$$\limsup_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{\sum_{j\in\mathbb{Z}} \left(\mathrm{e}^{\alpha j+\beta+\varepsilon} \wedge n - \mathrm{e}^{\alpha j+\beta-\varepsilon} \vee 1 + 1\right)_+}{n} = 0$$

can be checked by direct calculations. We omit further details; see [27, Proof of Theorem 1.4 (iii)]. □

5. Proofs for random trees

5.1. **Embedding the one-split BRW into a continuous-time BRW.** Continuous-time embeddings of discrete-time Markov chains in the study of random discrete structures go back at least to Athreya and Karlin [1] in the context of Pólya urn models. In the framework of random trees, Pittel [35] was the first to use a continuous-time embedding in the analysis of the height of BSTs. In the study of the profile of BSTs, the idea was introduced in a series of works by Chauvin and collaborators; see [6, 8, 9]. More recent works crucially relying on this technique are, among others, [22, 38, 39] and [40]. Start with a one-split BRW as described in Section 3.1. Consider a *continuous-time* BRW which starts with a single particle at the origin at time $\tau_0 := 0$ and in which any particle splits, with intensity 1, into a cluster of particles described by the same point process ζ as in the one-split BRW. The particles do not move between the splits. Denote the split times by $\tau_1 < \tau_2 < \dots$ and write N_t for the number of particles in the process at time $t \ge 0$. Note that $(N_t)_{t\ge 0}$ is a Galton–Watson process in continuous time. Further, we let $z_{1,t}, \dots, z_{N_t,t}$ be the positions of the particles and

$$\mathcal{L}_t(k) = \#\{1 \le j \le N_t \colon z_{j,t} = k\}, \quad k \in \mathbb{Z},$$

be the corresponding profile at time $t \ge 0$. We have the following correspondence:

$$S_n = N_{\tau_n}, \qquad x_{i,n} = z_{i,\tau_n},\ 1 \le i \le S_n, \qquad \mathbb{L}_n(k) = \mathcal{L}_{\tau_n}(k),\ k \in \mathbb{Z}.$$

For $\beta \in \mathbb{C}, \operatorname{Re}\beta \in \mathscr{I}$ (see Assumption B3) consider the Biggins martingale:

$$\mathcal{W}_t(\beta) = \mathrm{e}^{-m(\beta)t} \sum_{k=1}^{N_t} \mathrm{e}^{\beta z_{k,t}}, \quad t \ge 0.$$

Set $H_n(\beta) := \mathrm{e}^{m(\beta)\tau_n - \varphi(\beta)\log n}$, and note that $W_n(\beta)$ (see formula (43)) and $\mathcal{W}_{\tau_n}(\beta)$ are connected via the relation

$$W_n(\beta) = \mathcal{W}_{\tau_n}(\beta) H_n(\beta). \tag{92}$$

Our aim is to show that W_n converges, with speed $(\log n)^{-r}$, to a random analytic function W_∞, thus verifying Assumptions A2 and A3 of Theorem 2.1. Let us analyze the factors on the right-hand side of (92). Let $m^*(\beta) = \mathrm{e}^{m(\beta)}$. For $\gamma \in (1, 2]$, define the open sets

$$\Omega^1_\gamma = \operatorname{int}\left\{\beta \in \mathbb{C} : \operatorname{Re}\beta \in \mathscr{I}, \mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}} \mathrm{e}^{(\operatorname{Re}\beta)k} \mathcal{L}_1(k)\right)^\gamma\right] < \infty\right\},$$
$$\Omega^2_\gamma = \left\{\beta \in \mathbb{C} : \gamma \operatorname{Re}\beta \in \mathscr{I}, \frac{m^*(\gamma \operatorname{Re}\beta)}{|m^*(\beta)|^\gamma} < 1\right\},$$

and let

$$\mathscr{D} = \bigcup_{\gamma\in(1,2]} (\Omega^1_\gamma \cap \Omega^2_\gamma) \subset \mathbb{C}.$$

Note that the set $\mathscr{D}$ is open. Biggins [3] proved that, with probability 1, $\mathcal{W}_t$ converges locally uniformly on $\mathscr{D}$, as $t \to \infty$. The next proposition is a slight extension of this classical result adapted to our needs.

Proposition 5.1. *Under Assumptions B1–B3 and B5, there exists a random analytic function $\mathcal{W}_\infty$ on $\mathscr{D}$ such that, for all compact sets $K \subset \mathscr{D}$, there exists $0 < r = r(K) < 1$ with*

$$r^{-t} \sup_{\beta\in K} |\mathcal{W}_t(\beta) - \mathcal{W}_\infty(\beta)| \xrightarrow[t\to\infty]{a.s.} 0. \tag{93}$$

It holds that $(\beta_-, \beta_+) \subset \mathscr{D}$. Finally, for $\gamma \in (1, 2]$ and $\beta \in \Omega^1_\gamma \cap \Omega^2_\gamma$, we have

$$\lim_{t\to\infty} \mathbb{E}|\mathcal{W}_t(\beta) - \mathcal{W}_\infty(\beta)|^\gamma = 0. \tag{94}$$

Proof. Fix $\varepsilon > 0$ small enough. By compactness we can assume that $K = \overline{\mathbb{D}}_\varepsilon(z_0) \subset \Omega^1_\gamma \cap \Omega^2_\gamma$ for some $\gamma \in (1, 2]$ and some $z_0 \in \mathscr{D}$, and, moreover, $\overline{\mathbb{D}}_{2\varepsilon}(z_0) \subset \Omega^1_\gamma \cap \Omega^2_\gamma$.

Note that, for $\operatorname{Re}\beta \in \mathscr{I}$, the process $(\mathcal{W}_n(\beta))_{n\in\mathbb{N}_0}$ is the discrete-time Biggins martingale corresponding to a standard (many-split) discrete-time branching random walk whose point process $\zeta^* = \sum_{k\in\mathbb{Z}} \mathcal{L}_1(k)\delta_k$ has moment generating function $m^*(\beta) = \mathrm{e}^{m(\beta)}$. Hence, following the proof of Theorem 2 in [3], there exists a random analytic function $\mathcal{W}_\infty$ on $\mathscr{D}$ and $r_1 = r_1(z_0, \varepsilon) \in (0, 1)$ such that

$$r_1^{-n} \sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(z_0)} |\mathcal{W}_n(\beta) - \mathcal{W}_\infty(\beta)| \xrightarrow[n\to\infty]{a.s.} 0.$$

To prove (93) it remains to show that there exist $r_2 \in (0, 1)$ such that

$$r_2^{-n} \sup_{t\in[n,n+1]} \sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(z_0)} |\mathcal{W}_t(\beta) - \mathcal{W}_n(\beta)| \xrightarrow[n\to\infty]{a.s.} 0. \tag{95}$$

By Cauchy's integral formula, c.f. [3, Lemma 3]

$$\sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(z_0)} |\mathcal{W}_t(\beta) - \mathcal{W}_n(\beta)| \le \frac{1}{\pi} \int_0^{2\pi} \left|\mathcal{W}_t(z_0 + 2\varepsilon \mathrm{e}^{i\phi}) - \mathcal{W}_n(z_0 + 2\varepsilon \mathrm{e}^{i\phi})\right| \mathrm{d}\phi$$

and, using Hölder's inequality for integrals,

$$\sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(z_0)} |\mathcal{W}_t(\beta)-\mathcal{W}_n(\beta)|^\gamma \le \frac{1}{\pi^\gamma}\left(\int_0^{2\pi} \left|\mathcal{W}_t(z_0+2\varepsilon \mathrm{e}^{i\phi})-\mathcal{W}_n(z_0+2\varepsilon \mathrm{e}^{i\phi})\right| \mathrm{d}\phi\right)^\gamma$$
$$\le \frac{2^{\gamma-1}}{\pi}\int_0^{2\pi} \left|\mathcal{W}_t(z_0+2\varepsilon \mathrm{e}^{i\phi})-\mathcal{W}_n(z_0+2\varepsilon \mathrm{e}^{i\phi})\right|^\gamma \mathrm{d}\phi.$$

Therefore,

$$\mathbb{E}\left[\sup_{t\in[n,n+1]}\sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(z_0)} |\mathcal{W}_t(\beta)-\mathcal{W}_n(\beta)|^\gamma\right]$$
$$\le \frac{2^{\gamma-1}}{\pi}\int_0^{2\pi}\mathbb{E}\left[\sup_{t\in[n,n+1]}\left|\mathcal{W}_t(z_0+2\varepsilon \mathrm{e}^{i\phi})-\mathcal{W}_n(z_0+2\varepsilon \mathrm{e}^{i\phi})\right|^\gamma\right]\mathrm{d}\phi.$$

Since $(\mathcal{W}_t(\beta)-\mathcal{W}_n(\beta))_{t\ge n}$ is a martingale, by Doob's inequality, there exists a universal constant $C>0$ such that, for all $n\in\mathbb{N}$ and $\beta\in\mathscr{D}$,

$$\mathbb{E}\left[\sup_{t\in[n,n+1]}|\mathcal{W}_t(\beta)-\mathcal{W}_n(\beta)|^\gamma\right]\le C\mathbb{E}|\mathcal{W}_{n+1}(\beta)-\mathcal{W}_n(\beta)|^\gamma.$$

By [3, Lemma 2(i)], there is $M>0$ such that for all $\beta\in\mathscr{D}$,

$$\mathbb{E}|\mathcal{W}_{n+1}(\beta)-\mathcal{W}_n(\beta)|^\gamma\le M\left(\frac{m^*(\gamma\operatorname{Re}\beta)}{|m^*(\beta)|^\gamma}\right)^n.$$

Combining pieces together, we obtain

$$\mathbb{E}\left[\sup_{t\in[n,n+1]}\sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(z_0)}|\mathcal{W}_t(\beta)-\mathcal{W}_n(\beta)|^\gamma\right]\le 2^\gamma MCr_3^n,$$

where

$$r_3:=\sup_{\beta\in\overline{\mathbb{D}}_{2\varepsilon}(z_0)}\frac{m^*(\gamma\operatorname{Re}\beta)}{|m^*(\beta)|^\gamma}<1.$$

By the Borel–Cantelli lemma and the Markov inequality, (95) holds with arbitrary $r_2\in(r_3,1)$.

Equation (94) follows analogously upon noting that, by [3, Theorem 1], for $\gamma\in(1,2]$ and $\beta\in\Omega_\gamma^1\cap\Omega_\gamma^2$, it holds

$$\lim_{n\to\infty}\mathbb{E}|\mathcal{W}_n(\beta)-\mathcal{W}_\infty(\beta)|^\gamma=0.$$

□

Proposition 5.2. *Under Assumptions B1–B3 and B5, almost surely, the function $\mathcal{W}_\infty$ has no zeros on the interval (β_-,β_+).*

Proof. For any fixed $\beta\in(\beta_-,\beta_+)$, it is known that

$$\mathbb{P}[\mathcal{W}_\infty(\beta)=0]=0 \tag{96}$$

since the extinction probability of our BRW equals zero by Assumption B2; see Theorem 1 in [4]. We use a well-known fact that the limit process $(\mathcal{W}_\infty(\beta))_{\beta\in\mathscr{D}}$, satisfies a stochastic fixed-point equation. Let τ_1^* be the time of the first *non-trivial* split, that is the first time the number of particles in the BRW becomes at least 2. For $i\in\mathbb{N}$, denote by $(\mathcal{W}_t^{(i)})_{t\ge 0}$ the Biggins martingale corresponding to the continuous-time BRW initiated by the i-th individual at time τ_1^*. With probability 1,

$$\mathcal{W}_t(\beta) = \mathbb{1}_{\{t\geq \tau_1^*\}} \sum_{i=1}^{N_{\tau_1^*}} \mathrm{e}^{-m(\beta)\tau_1^*} \mathrm{e}^{\beta z_{i,\tau_1^*}} \mathcal{W}^{(i)}_{t-\tau_1^*}(\beta) + \mathbb{1}_{\{t<\tau_1^*\}} \mathrm{e}^{-m(\beta)t} \mathrm{e}^{\beta z_{1,t}}, \quad \beta \in \mathscr{D}.$$

Thus, sending t to $+\infty$, we obtain

$$\mathcal{W}_\infty(\beta) = \sum_{i=1}^{N_{\tau_1^*}} \mathrm{e}^{-m(\beta)\tau_1^*} \mathrm{e}^{\beta z_{i,\tau_1^*}} \mathcal{W}^{(i)}_\infty(\beta), \quad \beta \in \mathscr{D}. \tag{97}$$

Here, the martingale limits $(\mathcal{W}^{(i)}_\infty(\beta))_{\beta\in\mathscr{D}}$, $i \in \mathbb{N}$, are independent and have the same law as the process $(\mathcal{W}_\infty(\beta))_{\beta\in\mathscr{D}}$. Also, these processes are independent of the number and the positions of the offspring of the initial particle, but not independent of $(\mathcal{W}_\infty(\beta))_{\beta\in\mathscr{D}}$.

If the function $\mathcal{W}_\infty$ has a zero at some $\beta \in (\beta_-, \beta_+)$, then it follows from (97) that the processes $\mathcal{W}^{(i)}_\infty$, $1 \leq i \leq N_{\tau_1^*}$, have zeros at the same point β. Since $N_{\tau_1^*} \geq 2$, we infer

$$\begin{aligned}\mathbb{P}[\mathcal{W}_\infty(\beta) = 0 \text{ for some } \beta \in (\beta_-, \beta_+)]& \\ \leq \mathbb{P}[\mathcal{W}^{(1)}_\infty(\beta) = \mathcal{W}^{(2)}_\infty(\beta) = 0 &\text{ for some } \beta \in (\beta_-, \beta_+)].\end{aligned}$$

Since $(\mathcal{W}^{(1)}_\infty(\beta))_{\beta\in\mathscr{D}}$ is a random analytic function on $\mathscr{D}$ (which is not identically zero, with probability 1; see (96)), its zeros form a point process on (β_-, β_+) and we can construct a sequence of random variables $X_1, X_2, \dots$ (which depend only on $\mathcal{W}^{(1)}_\infty$) such that all zeros of $\mathcal{W}^{(1)}_\infty$ are contained in this list. Since $(\mathcal{W}^{(1)}_\infty(\beta))_{\beta\in\mathscr{D}}$ and $(\mathcal{W}^{(2)}_\infty(\beta))_{\beta\in\mathscr{D}}$ are independent, we infer

$$\begin{aligned}\mathbb{P}[\mathcal{W}_\infty(\beta) = 0 \text{ for some } \beta \in (\beta_-, \beta_+)] &\leq \sum_{i=1}^{\infty} \mathbb{P}[\mathcal{W}^{(2)}_\infty(X_i) = 0] \\ &= \sum_{i=1}^{\infty} \int_{(\beta_-,\beta_+)} \mathbb{P}[\mathcal{W}^{(2)}_\infty(x) = 0]\mathbb{P}[X_i \in \mathrm{d}x],\end{aligned}$$

which vanishes because $\mathbb{P}[\mathcal{W}^{(2)}_\infty(x) = 0] = 0$ for every $x \in (\beta_-, \beta_+)$ by (96). □

From Proposition 5.1 we can easily obtain the following a.s. asymptotics for the n-th split time τ_n.

Lemma 5.3. *There exists a deterministic $\varepsilon > 0$ such that, as $n \to \infty$,*

$$\tau_n = \frac{\log n}{m(0)} + \frac{\log m(0) - \log \mathcal{W}_\infty(0)}{m(0)} + o(n^{-\varepsilon}) \quad a.s. \tag{98}$$

Proof. From Proposition 5.1 with $\beta = 0$ we obtain

$$\mathcal{W}_t(0) \xrightarrow[t\to\infty]{a.s.} \mathcal{W}_\infty(0).$$

The continuous-time Galton–Watson process $(N_t)_{t\geq 0}$ does not explode because the expected number of particles in the cluster ζ is finite by Assumption B3. This means that $\tau_n \to \infty$ a.s., as $n \to \infty$, and the last display implies

$$\mathcal{W}_{\tau_n}(0) = \mathrm{e}^{-m(0)\tau_n} S_n \xrightarrow[n\to\infty]{a.s.} \mathcal{W}_\infty(0).$$

From Remark 3.5 we know that $S_n/n \to m(0)$ a.s., which yields

$$\tau_n = \frac{\log n}{m(0)} + O(1) \quad \text{a.s.} \tag{99}$$

Using Proposition 5.1 with $\beta = 0$ and $t = \tau_n$ gives:

$$\mathrm{e}^{-m(0)\tau_n} S_n - \mathcal{W}_\infty(0) = o(r^{\tau_n}) \quad \text{a.s.}$$

From (99) we deduce $r^{\tau_n} = o(n^{-\varepsilon_1})$ a.s., as $n \to \infty$, for every $\varepsilon_1 < |\log r|/m(0)$. The variance of S_1 is finite by Assumption B3. By the law of iterated logarithm, for every $\delta > 0$, as $n \to \infty$,

$$S_n = m(0)n + o(n^{1/2+\delta}) \quad \text{a.s.}$$

Combining the estimates, we see that (98) holds for $\varepsilon < (|\log r|/m(0) \wedge 1/2)$. □

Recall that $H_n(\beta) = \mathrm{e}^{m(\beta)\tau_n - \varphi(\beta)\log n}$. Lemma 5.3 immediately yields

Lemma 5.4. *For $\beta \in \mathscr{D}$, let $H_\infty(\beta) = (\mathcal{W}_\infty(0))^{-\varphi(\beta)} m(0)^{\varphi(\beta)}$. For any compact set $K \subset \mathscr{D}$, there exists $\varepsilon = \varepsilon(K) > 0$ such that*

$$n^\varepsilon \sup_{\beta \in K} |H_n(\beta) - H_\infty(\beta)| \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} 0.$$

Proof of Theorem 3.4. Recall from (92) that $W_n(\beta) = \mathcal{W}_{\tau_n}(\beta) H_n(\beta)$. Define

$$W_\infty(\beta) = \mathcal{W}_\infty(\beta) H_\infty(\beta) = \mathcal{W}_\infty(\beta)(\mathcal{W}_\infty(0))^{-\varphi(\beta)} m(0)^{\varphi(\beta)}.$$

By Proposition 5.1, Lemmas 5.3 and 5.4 and the triangle inequality, $W_n(\beta)$ converges to $W_\infty(\beta)$ locally uniformly on $\mathscr{D}$, with probability 1 and speed $(\log n)^{-r}$. Since $H_\infty(\beta) > 0$ for real β and, by Proposition 5.2, the function $\mathcal{W}_\infty$ has no zeros on the interval (β_-, β_+) (with probability 1), the same is true for the function W_∞. □

5.2. **Proof of Theorem 3.11.** Consider a one-split BRW satisfying Assumptions B1–B5. We are going to verify Assumptions A1–A4 of Theorem 2.1 for the sequence of its profiles $\mathbb{L}_1, \mathbb{L}_2, \ldots$ and $w_n = \log n$. Assumption A1 is fulfilled because the number of particles in the one-split BRW is finite at any time and hence, the function $\mathbb{L}_n$ has a.s. finite support. Assumptions A2 and A3 were verified in Theorem 3.4.

The next proposition verifies an analogue of Assumption A4 for the continuous-time BRW. The result is essentially shown in [3] without rate of convergence and only in the non-lattice case. For the sake of completeness, we include the proof here.

Proposition 5.5. *For any compact set $K \subset (\beta_-, \beta_+)$ and $0 < a < \pi$, under Assumptions B1–B5, there exists $0 < r = r(K, a) < 1$ such that*

$$r^{-t} \sup_{\beta \in K} \sup_{a \le \eta \le \pi} \mathrm{e}^{-m(\beta)t} \left| \sum_{k=1}^{N_t} \mathrm{e}^{(\beta + i\eta) z_{k,t}} \right| \overset{a.s.}{\underset{t\to\infty}{\longrightarrow}} 0. \tag{100}$$

Proof. Set $\psi(\beta) = |m^*(\beta)|/m^*(\operatorname{Re}\beta)$ and note that, for $\beta \in \mathbb{C}, \operatorname{Re}\beta \in \mathscr{I}$,

$$(m^*(\operatorname{Re}\beta))^{-t} \left| \sum_{k=1}^{N_t} \mathrm{e}^{\beta z_{k,t}} \right| = (\psi(\beta))^t |\mathcal{W}_t(\beta)|.$$

Therefore, (100) is equivalent to

$$r^{-t} \sup_{\beta \in G} \psi^t(\beta) |\mathcal{W}_t(\beta)| \xrightarrow[t\to\infty]{a.s.} 0, \tag{101}$$

where $G := \{\beta \in \mathbb{C} : \operatorname{Re}\beta \in K,\ \operatorname{Im}\beta \in [a, \pi]\}$. By compactness, it is enough to check that, for any $\beta_0 \in G$, there exists $\varepsilon > 0$ such that

$$r^{-t} \sup_{\beta \in \overline{\mathbb{D}}_\varepsilon(\beta_0)} \psi^t(\beta) |\mathcal{W}_t(\beta)| \xrightarrow[t\to\infty]{a.s.} 0. \tag{102}$$

By the Borel–Cantelli lemma, (102) follows from summability of the sequence

$$r^{-n} \mathbb{E}\left[\sup_{t \in [n,n+1]} \sup_{\beta \in \overline{\mathbb{D}}_\varepsilon(\beta_0)} \psi^n(\beta) |\mathcal{W}_t(\beta)| \right], \quad n \in \mathbb{N}. \tag{103}$$

By Cauchy's integral formula, c.f. [3, Lemma 3],

$$\sup_{\beta \in \overline{\mathbb{D}}_\varepsilon(\beta_0)} |\mathcal{W}_t(\beta)| \le \frac{1}{\pi} \int_0^{2\pi} |\mathcal{W}_t(\beta_0 + 2\varepsilon e^{i\phi})| d\phi,$$

whence, for $\gamma > 1$,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t \in [n,n+1]} \sup_{\beta \in \overline{\mathbb{D}}_\varepsilon(\beta_0)} |\mathcal{W}_t(\beta)| \right] &\le \frac{1}{\pi} \int_0^{2\pi} \mathbb{E}\left[\sup_{t \in [n,n+1]} |\mathcal{W}_t(\beta_0 + 2\varepsilon e^{i\phi})| \right] d\phi \\
&\le \frac{1}{\pi} \int_0^{2\pi} \left(\mathbb{E}\left[\sup_{t \in [n,n+1]} |\mathcal{W}_t(\beta_0 + 2\varepsilon e^{i\phi})|^\gamma \right] \right)^{1/\gamma} d\phi \\
&\le \frac{\gamma}{\pi(\gamma - 1)} \int_0^{2\pi} \left(\mathbb{E}|\mathcal{W}_{n+1}(\beta_0 + 2\varepsilon e^{i\phi})|^\gamma \right)^{1/\gamma} d\phi,
\end{aligned}$$

having utilized Doob's inequality in the last passage. Choose $\varepsilon > 0$ small enough such that there exists $\gamma > 1$ with $\overline{\mathbb{D}}_{2\varepsilon}(\operatorname{Re}\beta_0) \subseteq \Omega^1_\gamma \cap \Omega^2_\gamma$. We fix this γ in the remainder of the proof. Further, let

$$\kappa(\beta) := \frac{m^*(\gamma \operatorname{Re}\beta)}{|m^*(\beta)|^\gamma}.$$

From our choice of γ, it follows that $\mathbb{E}|\mathcal{W}_1(\beta)|^\gamma < \infty$ for all $\beta \in \overline{\mathbb{D}}_{2\varepsilon}(\beta_0)$. Using [3, Lemma 2(ii)], we obtain

$$\begin{aligned}
\mathbb{E}\left[\sup_{t \in [n,n+1]} \sup_{\beta \in \overline{\mathbb{D}}_\varepsilon(\beta_0)} |\mathcal{W}_t(\beta)| \right] &\le C \int_0^{2\pi} \left(\sum_{j=0}^n \kappa^j(\beta_0 + 2\varepsilon e^{i\phi}) \right)^{1/\gamma} d\phi, \\
&\le C(n+1)^{1/\gamma} \int_0^{2\pi} \left(\kappa(\beta_0 + 2\varepsilon e^{i\phi}) \vee 1 \right)^{n/\gamma} d\phi
\end{aligned}$$

for some $C > 0$. Therefore,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[n,n+1]}\sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(\beta_0)}\psi^n(\beta)|\mathcal{W}_t(\beta)|\right]\\
&\qquad\le\; C(n+1)^{1/\gamma}\left(\sup_{\beta\in\overline{\mathbb{D}}_\varepsilon(\beta_0)}\psi^n(\beta)\right)\int_0^{2\pi}\left(\kappa(\beta_0+2\varepsilon \mathrm{e}^{i\phi})\vee 1\right)^{n/\gamma}\mathrm{d}\phi\\
&\qquad\le\; 2C\pi(n+1)^{1/\gamma}\left(\left(\sup_{\beta\in\overline{\mathbb{D}}_{2\varepsilon}(\beta_0)}\psi^\gamma(\beta)\right)\left(\sup_{\beta\in\overline{\mathbb{D}}_{2\varepsilon}(\beta_0)}(\kappa(\beta)\vee 1)\right)\right)^{n/\gamma}.
\end{aligned}$$

Note that both $\beta\mapsto\psi^\gamma(\beta)$ and $\beta\mapsto\kappa(\beta)\vee 1$ are continuous in some neighborhood of β_0. Hence, for $\delta > 0$, upon possibly decreasing $\varepsilon > 0$, we obtain

$$\sup_{\beta\in\overline{\mathbb{D}}_{2\varepsilon}(\beta_0)}\psi^\gamma(\beta)\le(1+\delta)\psi^\gamma(\beta_0)\qquad\text{and}\qquad\sup_{\beta\in\overline{\mathbb{D}}_{2\varepsilon}(\beta_0)}(\kappa(\beta)\vee 1)\le(1+\delta)(\kappa(\beta_0)\vee 1).$$

(103) now follows from these bounds by a suitable choice of δ upon verifying that

$$\psi^\gamma(\beta_0)(\kappa(\beta_0)\vee 1)=\frac{m^*(\gamma\operatorname{Re}\beta_0)\vee|m^*(\beta_0)|^\gamma}{(m^*(\operatorname{Re}\beta_0))^\gamma}<1.$$

First, on the one hand, since $\operatorname{Re}\beta_0\in\Omega^1_\gamma\cap\Omega^2_\gamma$ we have

$$\frac{m^*(\gamma\operatorname{Re}\beta_0)}{(m^*(\operatorname{Re}\beta_0))^\gamma}<1.$$

On the other hand, since $a\le\operatorname{Im}\beta_0\le\pi$, we have

$$\frac{|m^*(\beta_0)|}{m^*(\operatorname{Re}\beta_0)}=\exp\left\{\sum_{k\in\mathbb{Z}}\nu_k\mathrm{e}^{k\operatorname{Re}\beta_0}(\cos(k\operatorname{Im}\beta_0)-1)\right\}<1,$$

having utilized Assumptions B1 and B4. The proof of Proposition 5.5 is complete. □

Now we can pass back to the one-split BRW. By combining Lemma 5.3 and Proposition 5.5, one deduces that, for any compact set $K\subset(\beta_-,\beta_+)$ and $0<a<\pi$, there exists $\varepsilon=\varepsilon(K,a)>0$ such that

$$n^\varepsilon\sup_{\beta\in K}\sup_{a<\eta\le\pi}n^{-\varphi(\beta)}\left|\sum_{k=1}^{S_n}\mathrm{e}^{(\beta+i\eta)x_{k,n}}\right|\xrightarrow[n\to\infty]{a.s.}0. \tag{104}$$

Assumption A4 follows readily. Theorem 3.11 now follows from Theorem 2.1.

5.3. **Proof of Theorem 3.16.** We apply Theorem 2.1 to the (deterministic) profile function $\tilde{\mathbb{L}}_n(k):=\mathbb{E}[\mathbb{L}_n(k)]$. Obviously, the corresponding moment generating function $\tilde{W}_n(\beta)$ is simply $\mathbb{E}W_n(\beta)$. Its limit $\tilde{W}_\infty(\beta)$ was calculated in (46): for any $\beta\in\mathbb{C}$ with $\operatorname{Re}\beta\in(\beta_-,\beta_+)$,

$$\tilde{W}_\infty(\beta):=\lim_{n\to\infty}\mathbb{E}W_n(\beta)=\frac{\Gamma\left(\frac{1}{m(0)}\right)}{\Gamma\left(\frac{m(\beta)+1}{m(0)}\right)}. \tag{105}$$

Using the explicit formula (45), a direct application of Stirling's formula shows that Assumption A3 is satisfied. Similarly, Assumption A4 is easily verified using (45) and Assumption B4. To conclude the proof, it remains to show that

$\tilde{W}_\infty(\beta) = \mathbb{E}W_\infty(\beta)$ for real $\beta \in (\beta_-, \beta_+)$ which is true if (and only if) the sequence $(W_n(\beta))_{n\in\mathbb{N}}$ is uniformly integrable.

Proposition 5.6. *Consider a one-split BRW with deterministic number of descendants and satisfying Assumptions B1–B3 and B5. Then, for every $\beta \in (\beta_-, \beta_+)$, the sequence $(W_n(\beta))_{n\in\mathbb{N}_0}$ is bounded in L_γ, for some $\gamma = \gamma(\beta) > 1$.*

Proof. For $\beta = 0$ and $\gamma = 2$, the relevant argument is given in the proof of Proposition 6 in [40]. Fix $\beta \in (\beta_-, \beta_+)$ and $\gamma \in (1, 2]$ such that $\beta \in \Omega^1_\gamma \cap \Omega^2_\gamma$. Note that $W_n(\beta)$ and $H_n(\beta)$ are independent and $\mathcal{W}_{\tau_n}(\beta) = H_n(\beta)W_n(\beta)$. By the optional stopping theorem, $(\mathcal{W}_{\tau_n}(\beta))_{n\in\mathbb{N}}$ is a martingale with mean 1 and bounded in L_γ. By indepedence, $\mathbb{E}\mathcal{W}_{\tau_n}(\beta) = \mathbb{E}H_n(\beta)\mathbb{E}W_n(\beta)$. Since $\mathbb{E}W_n(\beta)$ converges to a non-zero limit, see (105), it follows that $\mathbb{E}H_n(\beta)$ is bounded away from zero. Thus, by independence and Jensen's inequality,

$$\mathbb{E}\mathcal{W}^\gamma_{\tau_n}(\beta) = \mathbb{E}H^\gamma_n(\beta) \cdot \mathbb{E}W^\gamma_n(\beta) \geq \mathbb{E}W^\gamma_n(\beta)(\mathbb{E}H_n(\beta))^\gamma.$$

It follows that $\sup_{n\geq 0} \mathbb{E}W^\gamma_n(\beta) < \infty$ which completes the proof. □

Acknowledgement. Zakhar Kabluchko is grateful to Rudolf Grübel for useful discussions. The work of Alexander Marynych was supported by a Humboldt Research Fellowship of the Alexander von Humboldt Foundation. The work of Henning Sulzbach was supported by a Feodor Lynen Research Fellowship of the Alexander von Humboldt Foundation.

References

[1] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.
[2] F. Bergeron, P. Flajolet, and B. Salvy. Varieties of increasing trees. In *CAAP '92 (Rennes, 1992)*, volume 581 of *Lecture Notes in Comput. Sci.*, pages 24–48. Springer, Berlin, 1992.
[3] J. D. Biggins. Uniform convergence of martingales in the branching random walk. *Ann. Probab.*, 20(1):137–151, 1992.
[4] J. D. Biggins and D. R. Grey. Continuity of limit random variables in the branching random walk. *J. Appl. Probab.*, 16(4):740–749, 1979.
[5] J. D. Biggins and D. R. Grey. A note on the growth of random trees. *Statist. Probab. Lett.*, 32(4):339–342, 1997.
[6] B. Chauvin and A. Rouault. Connecting Yule Process, Bisection and Binary Search Tree via Martingales. *J. Iran. Statist. Soc.*, 3(2):89–116, 2004.
[7] B. Chauvin, M. Drmota, and J. Jabbour-Hattab. The profile of binary search trees. *Ann. Appl. Probab.*, 11(4):1042–1062, 2001.
[8] B. Chauvin, T. Klein, J.-F. Marckert, and A. Rouault. Martingales, embedding and tilting of binary trees. *Preprint*, 2003.
[9] B. Chauvin, T. Klein, J.-F. Marckert, and A. Rouault. Martingales and profile of binary search trees. *Elect. J. Probab.*, 10:420–435, 2005.
[10] X. Chen. Exact convergence rates for the distribution of particles in branching random walks. *Ann. Appl. Probab.*, 11(4):1242–1262, 2001.
[11] F. Delbaen, E. Kowalski, and A. Nikeghbali. Mod-φ convergence. *Int. Math. Res. Not. IMRN*, (11):3445–3485, 2015.

[12] L. Devroye. A note on the height of binary search trees. *J. Assoc. Comput. Mach.*, 33(3):489–498, 1986.
[13] L. Devroye. Branching processes in the analysis of the heights of trees. *Acta Inf.*, 24:277–298, 1987.
[14] L. Devroye and H.-K. Hwang. Width and mode of the profile for some random trees of logarithmic height. *Ann. Appl. Probab.*, 16(2):886–918, 2006.
[15] M. Drmota. *Random trees: An interplay between combinatorics and probability.* Springer–Verlag Wien, 2009.
[16] M. Drmota and H.-K. Hwang. Profiles of random trees: correlation and width of random recursive trees and binary search trees. *Adv. Appl. Probab.*, 37(2): 321–341, 2005.
[17] M. Drmota, S. Janson, and R. Neininger. A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.*, 18(1):288–333, 2008.
[18] P. Erdös. On a conjecture of Hammersley. *J. London Math. Soc.*, 28:232–236, 1953.
[19] V. Féray, P.-L. Méliot, and A. Nikeghbali. *Mod–ϕ Convergence: Normality Zones and Precise Deviations.* SpringerBriefs in Probability and Mathematical Statistics. Springer International Publishing, 2016.
[20] M. Fuchs, H.-K. Hwang, and R. Neininger. Profiles of random trees: limit theorems for random recursive trees and binary search trees. *Algorithmica*, 46 (3-4):367–407, 2006.
[21] R. Grübel and Z Kabluchko. Edgeworth expansions for profiles of lattice branching random walks. 2015. Ann. Inst. H. Poincaré, to appear. Preprint at `http://arxiv.org/abs/1503.04616`.
[22] R. Grübel and Z. Kabluchko. A functional central limit theorem for branching random walks, almost sure weak convergence and applications to random trees. *Ann. Appl. Probab.*, 26(6):3659–3698, 2016.
[23] H.-K. Hwang. Profiles of random trees: plane-oriented recursive trees. *Random Structures Algorithms*, 30(3):380–413, 2007.
[24] J. Jabbour-Hattab. Martingales and large deviations for binary search trees. *Random Structures Algorithms*, 19(2):112–127, 2001.
[25] J. Jacod, E. Kowalski, and A. Nikeghbali. Mod-Gaussian convergence: new limit theorems in probability and number theory. *Forum Math.*, 23(4):835–873, 2011.
[26] Z. Kabluchko. Distribution of levels in high-dimensional random landscapes. *Ann. Appl. Probab.*, 22(1):337–362, 2012.
[27] Z. Kabluchko, A. Marynych, and H. Sulzbach. Mode and Edgeworth expansion for the Ewens distribution and the Stirling numbers. *J. Integer Seq.*, 19, 2016. Article 16.8.8.
[28] Z. Katona. Width of a scale-free tree. *J. Appl. Probab.*, 42(3):839–850, 2005.
[29] E. Kowalski and A. Nikeghbali. Mod-Poisson convergence in probability and number theory. *Int. Math. Res. Not. IMRN*, (18):3549–3587, 2010.
[30] E. Kowalski and A. Nikeghbali. Mod-Gaussian convergence and the value distribution of $\zeta(\frac{1}{2}+it)$ and related quantities. *J. Lond. Math. Soc. (2)*, 86 (1):291–319, 2012.
[31] M. Kuba and A. Panholzer. The left-right-imbalance of binary search trees. *Theoret. Comput. Sci.*, 370(1-3):265–278, 2007.

[32] H. M. Mahmoud. *Evolution of random search trees.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1992.

[33] P.-L. Méliot and A. Nikeghbali. Mod-Gaussian convergence and its applications for models of statistical mechanics. In *In memoriam Marc Yor—Séminaire de Probabilités XLVII*, volume 2137 of *Lecture Notes in Math.*, pages 369–425. Springer, Cham, 2015.

[34] V. V. Petrov. *Sums of independent random variables.* Springer–Verlag, New York-Heidelberg, 1975. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[35] B. Pittel. On growing random binary trees. *J. Math. Anal. Appl.*, 103(2): 461–480, 1984.

[36] M. Régnier. A limiting distribution for quicksort. *RAIRO Inform. Théor. Appl.*, 23(3):335–343, 1989.

[37] U. Rösler. A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, 25 (1):85–100, 1991.

[38] E.-M. Schopp. A functional limit theorem for the profile of b-ary trees. *Ann. Appl. Probab.*, 20(3):907–950, 2010.

[39] H. Sulzbach. A functional limit law for the profile of plane-oriented recursive trees. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 339–350. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.

[40] H. Sulzbach. On martingale tail sums for the path length in random trees. *Random Structures Algorithms*, 50(3):493–508, 2017.

[41] K. Uchiyama. Spatial growth of a branching process of particles living in $\mathbf{R}^d$. *Ann. Probab.*, 10(4):896–918, 1982.

Zakhar Kabluchko, Institut für Mathematische Statistik, Universität Münster, Orléans–Ring 10, 48149 Münster, Germany

E-mail address: zakhar.kabluchko@uni-muenster.de

Alexander Marynych, Faculty of Cybernetics, Taras Shevchenko National University of Kyiv, 01601 Kyiv, Ukraine

E-mail address: marynych@unicyb.kiev.ua

Henning Sulzbach, School of Computer Science, McGill University, 3480 University Street H3A 0E9 Montréal, QC, Canada

E-mail address: henning.sulzbach@gmail.com