Combinatorics, Probability and Computing (2024), **33**, pp. 518–553
doi:10.1017/S0963548324000099

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

The distribution of the maximum protection number in simply generated trees

Clemens Heuberger[1], Sarah J. Selkirk[1] and Stephan Wagner[2,3]

[1]Institut für Mathematik, Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria, [2]Department of Mathematics, Uppsala Universitet, Uppsala, Sweden, and [3]Institute of Discrete Mathematics, TU Graz, Graz, Austria
Corresponding author: Clemens Heuberger; Email: clemens.heuberger@aau.at

(Received 30 May 2023; revised 13 February 2024; accepted 19 February 2024; first published online 12 April 2024)

Abstract

The protection number of a vertex v in a tree is the length of the shortest path from v to any leaf contained in the maximal subtree where v is the root. In this paper, we determine the distribution of the maximum protection number of a vertex in simply generated trees, thereby refining a recent result of Devroye, Goh, and Zhao. Two different cases can be observed: if the given family of trees allows vertices of outdegree 1, then the maximum protection number is on average logarithmic in the tree size, with a discrete double-exponential limiting distribution. If no such vertices are allowed, the maximum protection number is doubly logarithmic in the tree size and concentrated on at most two values. These results are obtained by studying the singular behaviour of the generating functions of trees with bounded protection number. While a general distributional result by Prodinger and Wagner can be used in the first case, we prove a variant of that result in the second case.

Keywords: Protection number; simply generated trees; generating functions

2020 MSC Codes: Primary: 05C05; Secondary: 05A15, 05A16, 05C80

1. Introduction

1.1 Simply generated trees

Simply generated trees were introduced by Meir and Moon [22], and owing to their use in describing an entire class of trees, have created a general framework for studying random trees. A simply generated family of rooted trees is characterised by a sequence of weights associated with the different possible outdegrees of a vertex. Specifically, for a given sequence of nonnegative real numbers w_j ($j \geq 0$), one defines the weight of a rooted ordered tree to be the product $\prod_v w_{d(v)}$ over all vertices of the tree, where $d(v)$ denotes the outdegree (number of children) of v. Letting $\Phi(t) = \sum_{j\geq 0} w_j t^j$ be the weight generating function and $Y(x)$ the generating function in which the coefficient of x^n is the sum of the weights over all n-vertex rooted ordered trees, one has the fundamental relation

$$Y(x) = x\Phi(Y(x)). \tag{1}$$

Common examples of simply generated trees are: plane trees with weight generating function $\Phi(t) = 1/(1-t)$; binary trees ($\Phi(t) = 1 + t^2$); pruned binary trees ($\Phi(t) = (1 + t)^2$); and labelled trees ($\Phi(t) = e^t$). In the first three examples, $Y(x)$ becomes an ordinary generating function with the total weight being the number of trees in the respective family, while $Y(x)$ can be seen as an exponential generating function in the case of labelled trees.

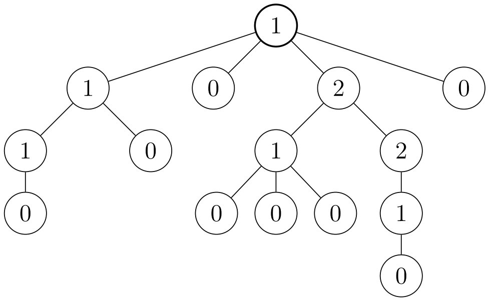

Figure 1. A plane tree with 15 vertices and the protection number of each vertex indicated. The maximum protection number of this tree is 2.

In addition to the fact that the notion of simply generated trees covers many important examples, there is also a strong connection to the probabilistic model of Bienaymé–Galton–Watson trees: here, one fixes a probability distribution on the set of nonnegative integers. Next, a random tree is constructed by starting with a root that produces offspring according to the given distribution. In each subsequent step, all vertices of the current generation also produce offspring according to the same distribution, all independent of each other and independent of all previous generations. The process stops if none of the vertices of a generation have children. If the weights in the construction of a simply generated family are taken to be the corresponding probabilities of the offspring distribution, then one verifies easily that the distribution of a random n-vertex tree from that family (with probabilities proportional to the weights) is the same as that of the Bienaymé–Galton–Watson process, conditioned on the event that the final tree has n vertices.

Conversely, even if the weight sequence of a simply generated family does not represent a probability measure, it is often possible to determine an equivalent probability measure that produces the same random tree distribution. For example, random plane trees correspond to a geometric distribution while random rooted labelled trees correspond to a Poisson distribution. We refer to [8] and [17] for more background on simply generated trees and Bienaymé–Galton–Watson trees.

1.2 Protection numbers in trees

Protection numbers in trees measure the distance to the nearest leaf successor. Formally, this can be expressed as follows.

Definition 1.1 (Protection number). *The protection number of a vertex v is the length of the shortest path from v to any leaf contained in the maximal subtree where v is the root.*

Alternatively, the protection number can be defined recursively: a leaf has protection number 0, the parent of a leaf has protection number 1, and generally the protection number of an interior vertex is the minimum of the protection numbers of its children plus 1. In this paper, we will be particularly interested in the *maximum protection number* of a tree, which is the largest protection number among all vertices. Fig. 1 shows an example of a tree along with the protection numbers of all its vertices.

The study of protection numbers in trees began with Cheon and Shapiro [4] considering the average number of vertices with protection number of at least 2 (called 2-protected) in ordered trees. Several other authors contributed to knowledge in this direction, by studying the number of 2-protected vertices in various types of trees: k-ary trees [21]; digital search trees [9]; binary search trees [20]; ternary search trees [15]; tries and suffix trees [11]; random recursive trees [19]; and general simply generated trees from which some previously known cases were also obtained [7].

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Generalising the concept of a vertex being 2-protected, k-protected vertices – when a vertex has protection number at least k – also became a recent topic of interest. Devroye and Janson [7] proved convergence of the probability that a random vertex in a random simply generated tree has protection number k. Copenhaver gave a closed formula for the number of k-protected vertices in all unlabelled rooted plane trees on n vertices along with expected values [5], and these results were extended by Heuberger and Prodinger [14]. A study of k-protected vertices in binary search trees was done by Bóna [2] and Bóna and Pittel [3]. Holmgren and Janson [16] proved general limit theorems for fringe subtrees and related tree functionals, applications of which include a normal limit law for the number of k-protected vertices in binary search trees and random recursive trees.

Moreover, the protection number of the root of families of trees has also been studied. In [14], Heuberger and Prodinger derived the probability of a plane tree having a root that is k-protected, the probability distribution of the protection number of the root of recursive trees is determined by Gołębiewski and Klimczak in [13]. The protection number of the root in simply generated trees, Pólya trees, and unlabelled non-plane binary trees was studied by Gittenberger, Gołębiewski, Larcher, and Sulkowska in [12], where they also obtained results relating to the protection number of a randomly chosen vertex.

Very recently, Devroye et al. [6] studied the maximum protection number in Bienaymé–Galton–Watson trees, referring to it as the leaf-height. Specifically, they showed the following: if X_n is the maximum protection number in a Bienaymé–Galton–Watson tree conditioned on having n vertices, then $\frac{X_n}{\log n}$ converges in probability to a constant if there is a positive probability that a vertex has exactly one child. If this is not the case, then $\frac{X_n}{\log\log n}$ converges in probability to a constant.

Our aim in this paper is to refine the result of Devroye, Goh and Zhao by providing the full limiting distribution of the maximum protection number. For our analytic approach, the framework of simply generated trees is more natural than the probabilistic setting of Bienaymé–Galton–Watson trees, though as mentioned earlier the two are largely equivalent.

1.3 Statement of results

As was already observed by Devroye et al. in [6], there are two fundamentally different cases to be considered, depending on whether or not vertices of outdegree 1 are allowed (have nonzero weight) in the given family of simply generated trees. If such vertices can occur, then we find that the maximum protection number of a random tree with n vertices is on average of order $\log n$, with a discrete double-exponential distribution in the limit. On the other hand, if there are no vertices of outdegree 1, then the maximum protection number is on average of order $\log\log n$. There is an intuitive explanation for this phenomenon. If outdegree 1 is allowed, it becomes easy to create vertices with high protection number: if the subtree rooted at a vertex is an $(h+1)$-vertex path, then this vertex has protection number h. On the other hand, if outdegree 1 is forbidden, then the smallest possible subtree rooted at a vertex of protection number h is a complete binary tree with $2^{h+1}-1$ vertices. An illustration of the two cases is given in Fig. 2.

In the case where vertices of outdegree 1 can occur, the limiting distribution turns out to be a discrete double-exponential distribution that also occurs in many other combinatorial examples, and for which general results are available – see Section 2.2. These results are adapted in Section 5.2 to the case where there are no vertices of outdegree 1.

In the following results, we make a common technical assumption, stating formally that there is a positive real number τ, less than the radius of convergence of Φ, such that $\Phi(\tau) = \tau\Phi'(\tau)$ (see Section 2.1 for further details). This is equivalent to the offspring distribution of the associated Bienaymé–Galton–Watson process having a finite exponential moment, which is the case for all the examples mentioned earlier (plane trees, binary trees, pruned binary trees, and labelled trees). This assumption is crucial for the analytic techniques that we are using, which are based on an

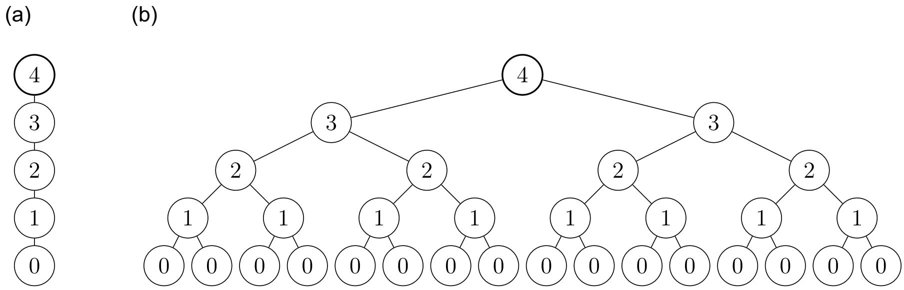

Figure 2. Smallest examples where a tree may (a) or may not (b) have exactly one child and the root has protection number 4.

asymptotic analysis of generating functions. However, it is quite likely that our main results remain valid under somewhat milder conditions.

Theorem 1.2. *Given a family of simply generated trees with $w_1 = \Phi'(0) \neq 0$, the proportion of trees of size n whose maximum protection number is at most h is asymptotically given by*

$$\exp\left(-\kappa n d^{-h}\right)(1 + o(1))$$

as $n \to \infty$ and $h = \log_d(n) + O(1)$, where κ (given in (55)) and $d = (\rho\Phi'(0))^{-1} > 1$ are positive constants, with ρ as defined in (3). Moreover, the expected value of the maximum protection number in trees with n vertices is

$$\log_d(n) + \log_d(\kappa) + \frac{\gamma}{\log(d)} + \frac{1}{2} + \psi_d(\log_d(\kappa n)) + o(1),$$

where γ denotes the Euler–Mascheroni constant and ψ_d is the 1-periodic function that is defined by the Fourier series

$$\psi_d(x) = -\frac{1}{\log(d)} \sum_{k \neq 0} \Gamma\left(-\frac{2k\pi i}{\log(d)}\right) e^{2k\pi i x}. \tag{2}$$

In the case where vertices of outdegree 1 are excluded, we show that the maximum protection number is strongly concentrated. In fact, with high probability it only takes on one of at most two different values (depending on the size of the tree). The precise result can be stated as follows.

Theorem 1.3. *Given a family of simply generated trees with $w_1 = \Phi'(0) = 0$, set $r = \min\{i \in \mathbb{N} \colon i \geq 2 \text{ and } w_i \neq 0\}$ and $D = \gcd\{i \in \mathbb{N} \colon w_i \neq 0\}$. The proportion of trees of size n whose maximum protection number is at most h is asymptotically given by*

$$\exp\left(-\kappa n d^{-r^h}(1 + o(1)) + o(1)\right)$$

as $n \to \infty$, $n \equiv 1 \pmod{D}$, and $h = \log_r(\log_d(n)) + O(1)$, where $\kappa = \frac{w_r \lambda_1^r}{\Phi(\tau)}$ and $d = \mu^{-r} > 1$ are positive constants with λ_1 and μ defined in (61) and (62) respectively (see Lemma 5.5). Moreover, there is a sequence of positive integers h_n such that the maximum protection number of a tree with n vertices is h_n or $h_n + 1$ with high probability (i.e., probability tending to 1 as $n \to \infty$) where $n \equiv 1 \pmod{D}$.

Specifically, with $m_n = \log_r \log_d(n)$ and $\{m_n\}$ denoting its fractional part, one can set

$$h_n = \begin{cases} \lfloor m_n \rfloor & \text{if } \{m_n\} \leq \frac{1}{2}, \\ \lceil m_n \rceil & \text{if } \{m_n\} > \frac{1}{2}. \end{cases}$$

If we restrict to those values of n for which $\{m_n\} \in [\varepsilon, 1 - \varepsilon]$, where $\varepsilon > 0$ is fixed, then with high probability X_n is equal to $\lceil m_n \rceil$.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

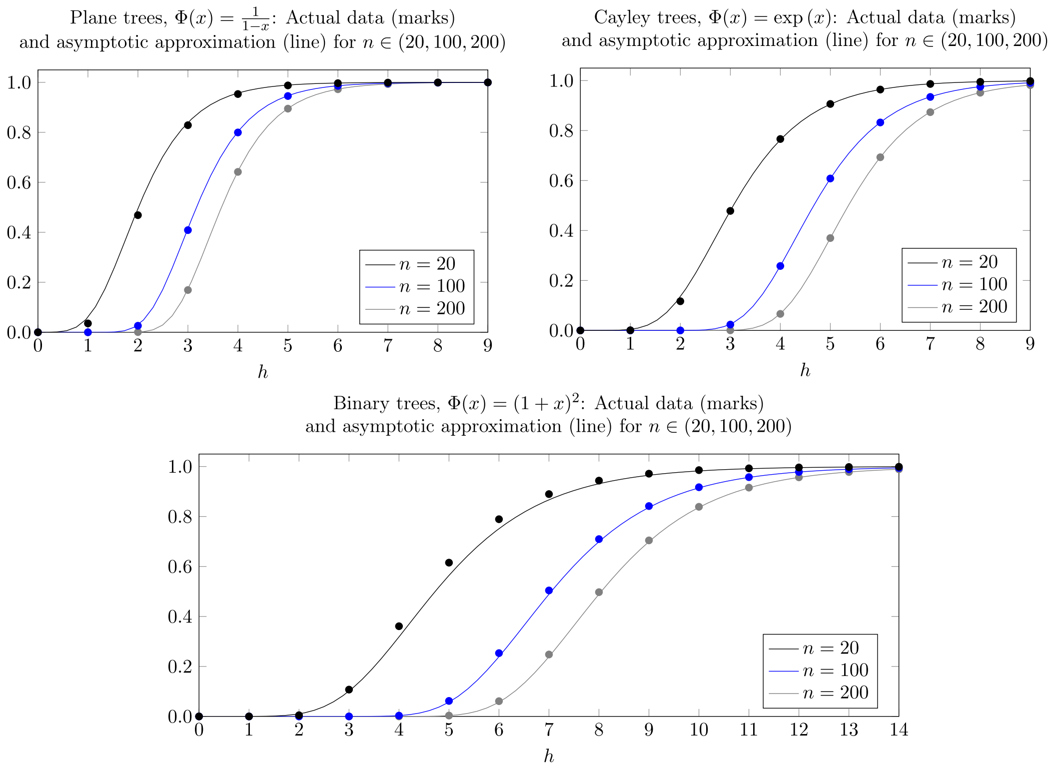

Figure 3. The asymptotic cumulative distribution function plotted against calculated values for plane, binary, and Cayley trees.

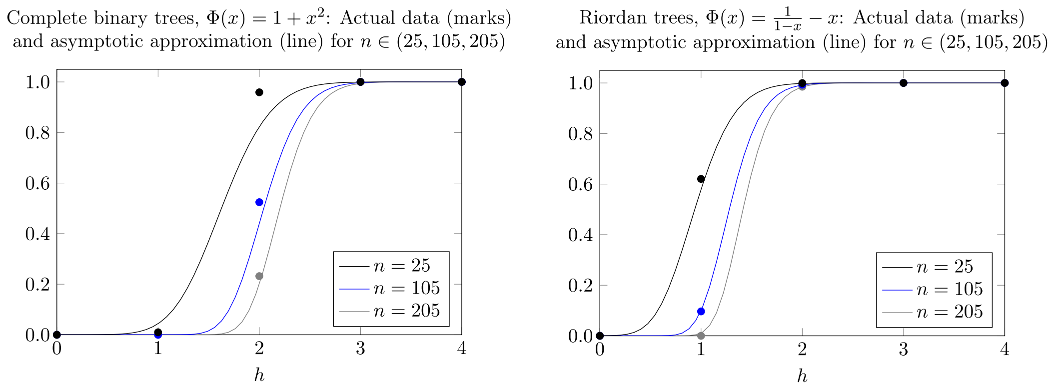

Figure 4. The asymptotic cumulative distribution function plotted against calculated values for complete binary, and Riordan trees [18].

Note that in the setting of Theorem 1.3, it is easy to see that there are no trees of size n if $n \not\equiv 1 \pmod{D}$. In the setting of Theorem 1.2, we have $\gcd\{i \in \mathbb{N} \colon w_i \neq 0\} = 1$ because $w_1 \neq 0$. Theorem 1.2 is illustrated in Fig. 3, while Theorem 1.3 is illustrated in Fig. 4.

The proof of Theorem 1.2 relies on a general distributional result provided in [23], see Theorem 2.1. For the proof of Theorem 1.3, however, we will need a variant for doubly exponential convergence of the dominant singularities. The statement and proof are similar to the original and we expect that this variant will be useful in other contexts, too.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Theorem 1.4. *Let $Y_h(x) = \sum_{n\geq 0} y_{h,n}x^n$ ($h \geq 0$) be a sequence of generating functions with nonnegative coefficients such that $y_{h,n}$ is nondecreasing in h and (coefficientwise)*

$$\lim_{h\to\infty} Y_h(x) = Y(x) = \sum_{n\geq 0} y_n x^n,$$

and let X_n denote the sequence of random variables with support $\mathbb{N}_0$ defined by

$$\mathbb{P}(X_n \leq h) = \frac{y_{h,n}}{y_n}.$$

Assume that each generating function Y_h has a singularity at $\rho_h \in \mathbb{R}$ such that

(1) *$\rho_h = \rho(1 + \kappa\zeta^{r^h} + o(\zeta^{r^h}))$ as $h \to \infty$ for some constants $\rho > 0$, $\kappa > 0$, $\zeta \in (0, 1)$, and $r > 1$.*
(2) *$Y_h(x)$ can be continued analytically to the domain*

$$\{x \in \mathbb{C} : |x| \leq (1 + \delta)|\rho_h|, |\mathrm{Arg}(x/\rho_h - 1)| > \phi\}$$

for some fixed $\delta > 0$ and $\phi \in (0, \pi/2)$, and

$$Y_h(x) = U_h(x) + A_h(1 - x/\rho_h)^\alpha + o((1 - x/\rho_h)^\alpha)$$

holds within this domain, uniformly in h, where $U_h(x)$ is analytic and uniformly bounded in h within the aforementioned region, $\alpha \in \mathbb{R} \setminus \mathbb{N}_0$, and A_h is a constant dependent on h such that $\lim_{h\to\infty} A_h = A \neq 0$. Finally,

$$Y(x) = U(x) + A(1 - x/\rho)^\alpha + o((1 - x/\rho)^\alpha)$$

in the region

$$\{x \in \mathbb{C} : |x| \leq (1 + \delta)|\rho|, |\mathrm{Arg}(x/\rho - 1)| > \phi\}$$

for a function $U(x)$ that is analytic within this region.

Then the asymptotic formula

$$\mathbb{P}(X_n \leq h) = \frac{y_{h,n}}{y_n} = \exp\left(-\kappa n\zeta^{r^h}(1 + o(1)) + o(1)\right)$$

holds as $n \to \infty$ and $h = \log_r(\log_d(n)) + O(1)$, where $d = \zeta^{-1}$.

Note that here we have $\rho_h = \rho(1 + \kappa\zeta^{r^h} + o(\zeta^{r^h}))$, while in Theorem 2.1 we have the exponential case $\rho_h = \rho(1 + \kappa\zeta^h + o(\zeta^h))$.

In the next theorem, we show that the consequences of this distributional result are quite drastic.

Theorem 1.5. *Assume the conditions of Theorem 1.4. There is a sequence of nonnegative integers h_n such that X_n is equal to h_n or $h_n + 1$ with high probability. Specifically, with $m_n = \log_r \log_d(n)$ and $\{m_n\}$ denoting its fractional part, one can set*

$$h_n = \begin{cases} \lfloor m_n \rfloor & \text{if } \{m_n\} \leq \frac{1}{2}, \\ \lceil m_n \rceil & \text{if } \{m_n\} > \frac{1}{2}. \end{cases}$$

If we restrict to those values of n for which $\{m_n\} \in [\varepsilon, 1 - \varepsilon]$, where $\varepsilon > 0$ is fixed, then with high probability X_n is equal to $\lceil m_n \rceil$.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

2. Preliminaries

2.1 Basic facts about simply generated trees

For our purposes, we will make the following typical technical assumptions: first, we assume without loss of generality that $w_0 = 1$ or equivalently $\Phi(0) = 1$. In other words, leaves have an associated weight of 1, which can be achieved by means of a normalising factor if necessary. Moreover, to avoid trivial cases in which the only possible trees are paths, we assume that $w_j > 0$ for at least one $j \geq 2$. Finally, we assume that there is a positive real number τ, less than the radius of convergence of Φ, such that $\Phi(\tau) = \tau\Phi'(\tau)$. As mentioned earlier, this is equivalent to the offspring distribution having exponential moments.

It is well known (see e.g. [8, Section 3.1.4]) that if such a τ exists, it is unique, and the radius of convergence ρ of Y can be expressed as

$$\rho = \tau/\Phi(\tau) = 1/\Phi'(\tau), \tag{3}$$

which is equivalent to ρ and τ satisfying the simultaneous equations $y = x\Phi(y)$ and $1 = x\Phi'(y)$ (which essentially mean that the implicit function theorem fails at the point (ρ, τ)). Moreover, Y has a square root singularity at ρ with $\tau = Y(\rho)$, with a singular expansion of the form

$$Y(x) = \tau + a\Big(1 - \frac{x}{\rho}\Big)^{1/2} + b\Big(1 - \frac{x}{\rho}\Big) + c\Big(1 - \frac{x}{\rho}\Big)^{3/2} + O\Big((\rho - x)^2\Big). \tag{4}$$

The coefficients a, b, c can be expressed in terms of Φ and τ. In particular, we have

$$a = -\Big(\frac{2\Phi(\tau)}{\Phi''(\tau)}\Big)^{1/2}.$$

In fact, there is a full Newton–Puiseux expansion in powers of $(1 - x/\rho)^{1/2}$. If the weight sequence is *aperiodic*, i.e., $\gcd\{j\colon w_j \neq 0\} = 1$, then ρ is the only singularity on the circle of convergence of Y, and for sufficiently small $\varepsilon > 0$ there are no solutions to the simultaneous equations $y = x\Phi(y)$ and $1 = x\Phi'(y)$ with $|x| \leq \rho + \varepsilon$ and $|y| \leq \tau + \varepsilon$ other than $(x, y) = (\rho, \tau)$. Otherwise, if this gcd is equal to D, there are D singularities at $\rho e^{2k\pi i/D}$ ($i \in \{0, 1, \ldots, D-1\}$), all with the same singular behaviour. In the following, we assume for technical simplicity that the weight sequence is indeed aperiodic, but the proofs are readily adapted to the periodic setting, see Remarks 3.17 and 5.9.

By means of singularity analysis [10, Chapter VI], the singular expansion (4) yields an asymptotic formula for the coefficients of Y: we have

$$y_n = [x^n]Y(x) \sim \frac{-a}{2\sqrt{\pi}} n^{-3/2} \rho^{-n}.$$

If the weight sequence corresponds to a probability distribution, then y_n is the probability that an *unconditioned* Bienaymé–Galton–Watson tree has exactly n vertices when the process ends. For other classes such as plane trees or binary trees, y_n represents the number of n-vertex trees in the respective class.

2.2 A general distributional result

The discrete double-exponential distribution in Theorem 1.2 has been observed in many other combinatorial instances, for example the longest run of zeros in a random 0-1-string, the longest horizontal segment in Motzkin paths or the maximum outdegree in plane trees. This can often be traced back to the behaviour of the singularities of associated generating functions. The following general result [23], similar to Theorem 1.4 but with an exponential instead of doubly exponential rate of convergence of the dominant singularity, will be a key tool for us.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Theorem 2.1 (see [23, Theorem 1]). *Let $Y_h(x) = \sum_{n\geq 0} y_{h,n}x^n$ $(h \geq 0)$ be a sequence of generating functions with nonnegative coefficients such that $y_{h,n}$ is nondecreasing in h and (coefficientwise)*

$$\lim_{h\to\infty} Y_h(x) = Y(x) = \sum_{n\geq 0} y_n x^n,$$

and let X_n denote the sequence of random variables with support $\mathbb{N}_0$ defined by

$$\mathbb{P}(X_n \leq h) = \frac{y_{h,n}}{y_n}. \tag{5}$$

Assume, moreover, that each generating function Y_h has a singularity $\rho_h \in \mathbb{R}$, such that

(1) *$\rho_h = \rho(1 + \kappa\zeta^h + o(\zeta^h))$ as $h \to \infty$ for some constants $\rho > 0$, $\kappa > 0$ and $\zeta \in (0, 1)$.*
(2) *$Y_h(x)$ can be continued analytically to the domain*

$$\{x \in \mathbb{C} : |x| \leq (1 + \delta)|\rho_h|, |\mathrm{Arg}(x/\rho_h - 1)| > \phi\} \tag{6}$$

for some fixed $\delta > 0$ and $\phi \in (0, \pi/2)$, and

$$Y_h(x) = U_h(x) + A_h(1 - x/\rho_h)^\alpha + o((1 - x/\rho_h)^\alpha)$$

holds within this domain, uniformly in h, where $U_h(x)$ is analytic and uniformly bounded in h within the aforementioned region, $\alpha \in \mathbb{R} \setminus \mathbb{N}_0$, and A_h is a constant depending on h such that $\lim_{h\to\infty} A_h = A \neq 0$. Finally,

$$Y(x) = U(x) + A(1 - x/\rho)^\alpha + o((1 - x/\rho)^\alpha)$$

in the region

$$\{x \in \mathbb{C} : |x| \leq (1 + \delta)|\rho|, |\mathrm{Arg}(x/\rho - 1)| > \phi\}$$

for a function $U(x)$ that is analytic within this region.

Then the asymptotic formula

$$\mathbb{P}(X_n \leq h) = \frac{y_{h,n}}{y_n} = \exp\left(-\kappa n\zeta^h\right)(1 + o(1))$$

holds as $n \to \infty$ and $h = \log_d(n) + O(1)$, where $d = \zeta^{-1}$. Hence the shifted random variable $X_n - \log_d(n)$ converges weakly to a limiting distribution if n runs through a subset of the positive integers such that the fractional part $\{\log_d(n)\}$ of $\log_d(n)$ converges.

As we will see, the conditions of this theorem hold for the random variable X_n given by the maximum protection number of a random n-vertex tree from a simply generated family that satisfies our technical assumptions. Under slightly stronger assumptions, which also hold in our case, one has the following theorem on the expected value of the random variable X_n.

Theorem 2.2 (see [23, Theorem 2]). *In the setting of Theorem 2.1, assume additionally that*

(1) *There exists a constant K such that $y_{h,n} = y_n$ for $h > Kn$,*
(2) *$\sum_{h\geq 0} |A - A_h| < \infty$,*
(3) *the asymptotic expansions of Y_h and Y around their singularities are given by*

$$Y_h(x) = U_h(x) + A_h(1 - x/\rho_h)^\alpha + B_h(1 - x/\rho_h)^{\alpha+1} + o((1 - x/\rho_h)^{\alpha+1}),$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

uniformly in h, and

$$Y(x) = U(x) + A(1 - x/\rho_h)^{\alpha} + B(1 - x/\rho_h)^{\alpha+1} + o((1 - x/\rho_h)^{\alpha+1}),$$

respectively, such that $\lim_{h\to\infty} B_h = B$.

Then the mean of X_n *satisfies*

$$\mathbb{E}(X_n) = \log_d(n) + \log_d(\kappa) + \frac{\gamma}{\log(d)} + \frac{1}{2} + \psi_d(\log_d(\kappa n)) + o(1),$$

where γ *denotes the Euler–Mascheroni constant and* ψ_d *is given by (2).*

2.3 A system of functional equations

As a first step of our analysis, we consider a number of auxiliary generating functions and derive a system of functional equations that is satisfied by these generating functions. The family of simply generated trees and the associated weight generating function Φ are regarded fixed throughout. Let h be a positive integer and k an integer with $0 \le k \le h$. Consider trees with the following two properties:

P1. No vertex has a protection number greater than h.

P2. The root is k-protected (but also has protection number at most h).

Let $Y_{h,k}(x)$ be the associated generating function, where x marks the number of vertices. Note in particular that when $k = 0$, we obtain the generating function for trees where the maximum protection number is at most h. Hence we can express the probability that the maximum protection number of a random n-vertex tree (from our simply generated family) is at most h as the quotient

$$\frac{[x^n]Y_{h,0}(x)}{[x^n]Y(x)}.$$

This is precisely the form of (5), and indeed our general strategy will be to show that the generating functions $Y_{h,0}$ satisfy the technical conditions of Theorem 2.1. Compared to the examples given in [23], this will be a rather lengthy technical task. However, we believe that the general method, in which a sequence of functional equations is shown to converge uniformly in a suitable region, is also potentially applicable to other instances and therefore interesting in its own right.

Let us now derive a system of functional equations, using the standard decomposition of a rooted tree into the root and its branches. Clearly, if a tree has property P1, then this must also be the case for all its branches. Moreover, property P2 is satisfied for $k > 0$ if and only if the root of each of the branches is $(k-1)$-protected, but not all of them are h-protected (as this would make the root $(h+1)$-protected). Thus, for $1 \le k \le h$, we have

$$Y_{h,k}(x) = x\Phi(Y_{h,k-1}(x)) - x\Phi(Y_{h,h}(x)). \tag{7}$$

Note that the only case in which the root is only 0-protected is when the root is the only vertex. Hence we have

$$Y_{h,0}(x) = Y_{h,1}(x) + x. \tag{8}$$

The analytic properties of the system of functional equations given by (7) and (8) will be studied in the following section, culminating in Proposition 3.16, which shows that Theorem 2.1 is indeed applicable to our problem.

3. Analysis of the functional equations

3.1 Contractions and implicit equations

This section is devoted to a detailed analysis of the generating functions $Y_{h,k}$ that satisfy the system of equations given by (7) and (8). The first step will be to reduce it to a single implicit equation satisfied by $Y_{h,1}$ that is then shown to converge to the functional equation (1) in a sense that will be made precise. This is then used to infer information on the region of analyticity of $Y_{h,1}$ as well as its behaviour around the dominant singularity, which is also shown to converge to the dominant singularity of Y. This information is collected in Proposition 3.16 at the end of the section.

In the following, we will prove various statements for sufficiently small $\varepsilon > 0$. In several, but finitely many, steps it might be necessary to decrease ε; we tacitly assume that ε is always small enough to ensure validity of all statements up to the given point. In order to avoid ambiguities, we will always assume that $\varepsilon < 1$. Let us remark that ε and other constants as well as all implied O-constants that occur in this section depend on the specific simply generated family of trees (in particular the weight generating function Φ and therefore ρ and τ), but nothing else.

Recall that ρ is the dominant singularity of the generating function Y of our simply generated family of trees. Moreover, $\tau = Y(\rho)$ is characterised by the equation $\tau\Phi'(\tau) = \Phi(\tau)$ (see (3)) and satisfies $\tau = \rho\Phi(\tau)$. Since Φ is increasing and $\Phi(0) = 1$, we also have $\tau = \rho\Phi(\tau) > \rho\Phi(0) = \rho$.

Let us write $D_\delta(w) := \{z \in \mathbb{C} : |z - w| < \delta\}$ for open disks. For $\varepsilon > 0$, we define

$$\begin{aligned}\Xi_\varepsilon^{(1)} &:= D_{\rho+\varepsilon}(0),\\ \Xi_\varepsilon^{(2)} &:= D_{\tau-\rho+\varepsilon}(0),\\ \Xi_\varepsilon^{(3)} &:= D_\varepsilon(0).\end{aligned}$$

For $1 \le j < k \le 3$, we set $\Xi_\varepsilon^{(j,k)} := \Xi_\varepsilon^{(j)} \times \Xi_\varepsilon^{(k)}$, and we also set $\Xi_\varepsilon := \Xi_\varepsilon^{(1,2,3)} := \Xi_\varepsilon^{(1)} \times \Xi_\varepsilon^{(2)} \times \Xi_\varepsilon^{(3)}$. As τ is less than the radius of convergence of Φ by our assumptions, we may choose $\varepsilon > 0$ sufficiently small such that $\tau + 2\varepsilon$ is still smaller than the radius of convergence of Φ.

Consider the function defined by $f_{x,z}(y) = x(\Phi(y) - \Phi(z))$. We can rewrite the functional equation (7) in terms of this function as

$$Y_{h,k}(x) = f_{x,Y_{h,h}(x)}(Y_{h,k-1}(x)) \tag{9}$$

for $1 \le k \le h$. For $j \ge 0$, we denote the jth iterate of $f_{x,z}$ by $f_{x,z}^{(j)}$, i. e., $f_{x,z}^{(0)}(y) = y$ and $f_{x,z}^{(j+1)}(y) = f_{x,z}^{(j)}(f_{x,z}(y))$ for $j \ge 0$. Iterating (9) then yields

$$Y_{h,k}(x) = f_{x,Y_{h,h}(x)}(Y_{h,k-1}(x)) = \cdots = f_{x,Y_{h,h}(x)}^{(k-1)}(Y_{h,1}(x))$$

for $1 \le k \le h$ and therefore

$$Y_{h,h}(x) = f_{x,Y_{h,h}(x)}^{(h-1)}(Y_{h,1}(x)). \tag{10}$$

Plugging (8) into (7) for $k = 1$ yields

$$Y_{h,1}(x) = x\big(\Phi(Y_{h,1}(x) + x) - \Phi(Y_{h,h}(x))\big). \tag{11}$$

This means that (10) and (11) are a system of two functional equations for $Y_{h,1}(x)$ and $Y_{h,h}(x)$. We intend to solve (10) for $Y_{h,h}(x)$ and then plug the solution into (11). As a first step towards this goal, we show that $f_{x,z}$ represents a contraction on a suitable region.

Lemma 3.1. *For sufficiently small $\varepsilon > 0$, we have $|f_{x,z}(y)| < \tau - \rho$ for all $(x, y, z) \in \Xi_\varepsilon$.*

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Proof. By the triangle inequality, definition of Ξ_ε, non-negativity of the coefficients of Φ, and $\Phi(0)=1$, we have

$$\begin{aligned}|f_{x,z}(y)| &= \big|x\big((\Phi(y)-1)-(\Phi(z)-1)\big)\big|\\ &\le (\rho+\varepsilon)(|\Phi(y)-1|+|\Phi(z)-1|)\\ &\le (\rho+\varepsilon)((\Phi(|y|)-1)+(\Phi(|z|)-1))\\ &\le (\rho+\varepsilon)(\Phi(\tau-\rho+\varepsilon)-1+\Phi(\varepsilon)-1).\end{aligned}$$

For $\varepsilon\to 0$, the upper bound converges to $\rho\Phi(\tau-\rho)-\rho$ because we are assuming that $\Phi(0)=1$. As $\rho\Phi(\tau-\rho)-\rho<\rho\Phi(\tau)-\rho=\tau-\rho$ by (3), the assertion of the lemma holds for sufficiently small $\varepsilon>0$. □

Lemma 3.2. *For sufficiently small $\varepsilon>0$ and $(x,y,z)\in\Xi_\varepsilon$, we have $|f'_{x,z}(y)|=|x\Phi'(y)|\le\lambda$ for some constant $\lambda<1$.*

Proof. For any triple $(x,y,z)\in\Xi_\varepsilon$,

$$|f'_{x,z}(y)|=|x\Phi'(y)|\le(\rho+\varepsilon)\Phi'(\tau-\rho+\varepsilon).$$

For $\varepsilon\to 0$, the upper bound converges to $\rho\Phi'(\tau-\rho)$, which is less than $\rho\Phi'(\tau)=1$ (by (3)). □

For the remainder of this section, λ will be defined as in Lemma 3.2.

Lemma 3.3. *For sufficiently small $\varepsilon>0$ and $(x,z)\in\Xi_\varepsilon^{(1,3)}$, $f_{x,z}$ maps $\Xi_\varepsilon^{(2)}$ to itself and is a contraction with Lipschitz constant λ.*

Proof. The fact that $f_{x,z}$ maps $\Xi_\varepsilon^{(2)}$ to itself for sufficiently small $\varepsilon>0$ is a direct consequence of Lemma 3.1.

Making use of Lemma 3.2, the contraction property now follows by a standard argument: For $y_1,y_2\in\Xi_\varepsilon^{(2)}$, we have

$$|f_{x,z}(y_2)-f_{x,z}(y_1)|\le\int_{[y_1,y_2]}|f'_{x,z}(y)|\,|dy|\le\lambda|y_2-y_1|.$$

□

For sufficiently small ε and $(x,z)\in\Xi_\varepsilon^{(1,3)}$, Banach's fixed point theorem together with Lemma 3.3 implies that $f_{x,z}$ has a unique fixed point in $\Xi_\varepsilon^{(2)}$. This fixed point will be denoted by $g(x,z)$, i. e.,

$$g(x,z)=f_{x,z}(g(x,z))=x(\Phi(g(x,z))-\Phi(z)). \tag{12}$$

If we plug in 0 for z, we see that (12) holds for $g(x,0)=0$, so uniqueness of the fixed point implies that

$$g(x,0)=0 \tag{13}$$

for $x\in\Xi_\varepsilon^{(1)}$.

Lemma 3.4. *For sufficiently small $\varepsilon>0$, $g\colon\Xi_\varepsilon^{(1,3)}\to\Xi_\varepsilon^{(2)}$ is an analytic function, and $\frac{\partial}{\partial z}g(x,z)$ is bounded.*

Proof. Note that using Lemma 3.2, we have that $|\frac{\partial}{\partial y}(y-f_{x,z}(y))|=|1-f'_{x,z}(y)|\ge 1-|f'_{x,z}(y)|\ge 1-\lambda$ is bounded away from zero for sufficiently small $\varepsilon>0$ and $(x,y,z)\in\Xi_\varepsilon$. Thus the analytic implicit function theorem shows that g as defined by (12) is analytic and has bounded partial derivative $\frac{\partial}{\partial z}g(x,z)$ on $\Xi_\varepsilon^{(1,3)}$ for sufficiently small $\varepsilon>0$. □

We now intend to solve (10) for $Y_{h,h}(x)$. Therefore, we consider the equation

$$z=f_{x,z}^{(h-1)}(y) \tag{14}$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

and attempt to solve it for z. For large h, $f_{x,z}^{(h-1)}(y)$ will be close to the fixed point $g(x,z)$ of $f_{x,z}$ by the Banach fixed point theorem.

Therefore, we define Λ_h as the difference between the two: $\Lambda_h(x,y,z) := f_{x,z}^{(h-1)}(y) - g(x,z)$. So (14) can be rewritten as

$$z = g(x,z) + \Lambda_h(x,y,z). \tag{15}$$

We first establish bounds on Λ_h.

Lemma 3.5. *For sufficiently small $\varepsilon > 0$,*

$$\Lambda_h(x,y,z) = O(\lambda^h) \text{ and} \tag{16}$$

$$\frac{\partial}{\partial z}\Lambda_h(x,y,z) = O(\lambda^h) \tag{17}$$

hold uniformly for $(x,y,z) \in \Xi_\varepsilon$.

Proof. Since g is defined as the fixed point of $f_{x,z}$ and $f_{x,z}$ is a contraction with Lipschitz constant λ, we have

$$|\Lambda_h(x,y,z)| = |f_{x,z}^{(h-1)}(y) - f_{x,z}^{(h-1)}(g(x,z))| \le \lambda^{h-1}|y - g(x,z)| = O(\lambda^h)$$

for $(x,y,z) \in \Xi_\varepsilon$, so we have shown (16).

For $(x,y,z) \in \Xi_{\varepsilon/3}$, Cauchy's integral formula yields

$$\frac{\partial}{\partial z}\Lambda_h(x,y,z) = \frac{1}{2\pi i}\oint_{|\zeta - z| = \varepsilon/3} \frac{\Lambda_h(x,y,\zeta)}{(\zeta - z)^2}\, d\zeta.$$

By (16), we can bound the integral by $O(\lambda^h)$. Thus replacing ε by $\varepsilon/3$ yields (17). $\square$

In order to apply the analytic implicit function theorem to the implicit equation (10) for $Y_{h,h}$, we will need to show that the derivative of the difference of the two sides of (15) with respect to z is nonzero. The derivative of the second summand on the right-hand side of (15) is small by (17), so we first consider the remaining part of the equation.

Lemma 3.6. *There is a $\delta > 0$ such that for sufficiently small $\varepsilon > 0$, we have*

$$\left|\frac{\partial}{\partial z}(z - g(x,z))\right| > \delta \tag{18}$$

for $(x,z) \in \Xi_\varepsilon^{(1,3)}$.

Proof. To compute $\frac{\partial}{\partial z}g(x,z)$, we differentiate (12) with respect to z and obtain

$$\frac{\partial}{\partial z}g(x,z) = x\Phi'(g(x,z))\frac{\partial}{\partial z}g(x,z) - x\Phi'(z),$$

which leads to

$$\frac{\partial}{\partial z}g(x,z) = -\frac{x\Phi'(z)}{1 - x\Phi'(g(x,z))}.$$

Note that the denominator is nonzero for $(x,z) \in \Xi_\varepsilon^{(1,3)}$ by Lemma 3.2. We obtain

$$\left|\frac{\partial}{\partial z}(z - g(x,z))\right| = \left|\frac{1 + x(\Phi'(z) - \Phi'(g(x,z)))}{1 - x\Phi'(g(x,z))}\right| \ge \frac{1 - (\rho + \varepsilon)\,\left|\Phi'(z) - \Phi'(g(x,z))\right|}{1 + (\rho + \varepsilon)\,\left|\Phi'(g(x,z))\right|}. \tag{19}$$

By Lemma 3.4, $\frac{\partial g(x,z)}{\partial z}$ is analytic and bounded for $(x,z) \in \Xi_\varepsilon^{(1,3)}$, and by (13), it follows that

$$g(x,z) = g(x,z) - g(x,0) = \int_{[0,z]} \frac{\partial g(x,\zeta)}{\partial \zeta}\, d\zeta = O(|z|) = O(\varepsilon)$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

for $\varepsilon \to 0$, uniformly in x. Therefore, we have

$$\Phi'(z) - \Phi'(g(x,z)) = (\Phi'(z) - \Phi'(0)) - (\Phi'(g(x,z)) - \Phi'(0)) = O(\varepsilon)$$

and $|\Phi'(g(x,z))| = \Phi'(0) + O(\varepsilon)$ for $\varepsilon \to 0$. So (19) yields

$$\left|\frac{\partial}{\partial z}(z - g(x,z))\right| \geq \frac{1 - (\rho+\varepsilon)O(\varepsilon)}{1 + (\rho+\varepsilon)(\Phi'(0) + O(\varepsilon))} = \frac{1}{1+\rho\Phi'(0)} + O(\varepsilon)$$

for $\varepsilon \to 0$. Setting $\delta := \frac{1}{2}\frac{1}{1+\rho\Phi'(0)}$ and choosing ε small enough yields the result. □

We need bounds for z such that we remain in the region where our previous results hold. In fact, (13) shows that $z = 0$ would be a solution when the summand Λ_h (which is $O(\lambda^h)$) is removed from the implicit equation, so we expect that the summand Λ_h does not perturb z too much. This is shown in the following lemma.

Lemma 3.7. *Let $\varepsilon > 0$ be sufficiently small and $(x, y, z) \in \Xi_\varepsilon$ such that (15) holds. Then*

$$z = O(\lambda^h). \tag{20}$$

Proof. In view of (15) and (16), we have

$$g(x,z) - z = O(\lambda^h). \tag{21}$$

By definition, $g(x,z) \in \Xi_\varepsilon^{(2)}$. The implicit equation (12) for $g(x,z)$ and (21) imply

$$g(x,z) = x(\Phi(g(x,z)) - \Phi(z)) = x\int_{[z,g(x,z)]}\Phi'(\zeta)\,d\zeta = O(|g(x,z) - z|) = O(\lambda^h).$$

Inserting this into (21) leads to (20). □

Lemma 3.8. *There exists an $\varepsilon > 0$ such that for sufficiently large h, there is a unique analytic function $q_h\colon \Xi_\varepsilon^{(1,2)} \to \mathbb{C}$ such that*

$$q_h(x,y) = f_{x,q_h(x,y)}^{(h-1)}(y) \tag{22}$$

and $q_h(x,0) = 0$ for $(x,y) \in \Xi_\varepsilon^{(1,2)}$; furthermore, $q_h(x,y) = O(\lambda^h)$ holds uniformly in x and y.

Proof. We choose h sufficiently large such that (17) implies

$$\left|\frac{\partial}{\partial z}\Lambda_h(x,y,z)\right| \leq \frac{\delta}{2} \tag{23}$$

for $(x,y,z) \in \Xi_\varepsilon$, where δ is taken as in Lemma 3.6, and such that (20) implies

$$|z| \leq \frac{\varepsilon}{2} \tag{24}$$

for all $(x,y,z) \in \Xi_\varepsilon$ for which (15) holds.

By definition of f, we have $f_{x,0}(0) = 0$ and therefore $f_{x,0}^{(h-1)}(0) = 0$ for every $x \in \Xi_\varepsilon^{(1)}$, so $z = 0$ is a solution of (14) for $y = 0$. By (18) and (23), we have

$$\frac{\partial}{\partial z}(f_{x,z}^{(h-1)}(y) - z) \neq 0 \tag{25}$$

for $(x,y,z) \in \Xi_\varepsilon$. The analytic implicit function theorem thus implies that, for every $x \in \Xi_\varepsilon^{(1)}$, there is an analytic function q_h defined in a neighbourhood of $(x, 0)$ such that (22) holds there and such that $q_h(x,0) = 0$. Next we show that this extends to the whole region $\Xi_\varepsilon^{(1,2)}$.

For $x_0 \in \Xi_\varepsilon^{(1)}$, let $r(x_0)$ be the supremum of all $r < \tau - \rho + \varepsilon$ for which there is an analytic extension of $y \mapsto q_h(x_0, y)$ from the open disk $D_r(0)$ to $\Xi_\varepsilon^{(3)}$. Suppose for contradiction that $r(x_0) <$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

$\tau - \rho + \varepsilon$. Consider a point y_0 with $|y_0| = r(x_0)$, and take a sequence $y_n \to y_0$ such that $|y_n| < r(x_0)$. Note that $\left|q_h(x_0, y_n)\right| \le \frac{\varepsilon}{2}$ by (24). Without loss of generality, we can assume that $q_h(x_0, y_n)$ converges to some q_0 with $|q_0| \le \frac{\varepsilon}{2}$ as $n \to \infty$ (by compactness). By continuity, we have $q_0 = f_{x_0,q_0}^{(h-1)}(y_0)$. Since $(x_0, y_0, q_0) \in \Xi_\varepsilon$, we can still use the analytic implicit function theorem together with (25) to conclude that there is a neighbourhood of (x_0, y_0, q_0) where the equation $f_{x,z}^{(h-1)}(y) = z$ has exactly one solution z for every x and y, and an analytic function $\tilde{q}_h(x, y)$ such that $\tilde{q}_h(x, y) = f_{x,\tilde{q}_h(x,y)}^{(h-1)}(y)$ and $\tilde{q}_h(x_0, y_0) = q_0$. We assume the neighbourhood to be chosen small enough such that $\tilde{q}_h(x, y) \in \Xi_\varepsilon^{(3)}$ for all (x, y) in the neighbourhood. For large enough n, this neighbourhood contains $(x_0, y_n, q_h(x_0, y_n))$, so we must have $q_h(x_0, y_n) = \tilde{q}_h(x_0, y_n)$ for all those n. This implies that $\tilde{q}_h$ is an analytic continuation of q_h in a neighbourhood of (x_0, y_0) with values in $\Xi_\varepsilon^{(3)}$. Since y_0 was arbitrary, we have reached the desired contradiction.

So we conclude that there is indeed such an analytic function q_h defined on all of $\Xi_\varepsilon^{(1,2)}$, with values in $\Xi_\varepsilon^{(3)}$. The fact that $q_h(x, y) = O(\lambda^h)$ finally follows from Lemma 3.7. □

3.2 Location of the dominant singularity

Let us summarise what has been proven so far. By (10) and Lemma 3.8, for sufficiently large h we can express $Y_{h,h}$ in terms of $Y_{h,1}$ as

$$Y_{h,h}(x) = q_h(x, Y_{h,1}(x))$$

at least in a neighbourhood of 0, which we can plug into (11) to get

$$Y_{h,1}(x) = x\big(\Phi(Y_{h,1}(x) + x) - \Phi(q_h(x, Y_{h,1}(x)))\big).$$

Setting

$$F_h(x, y) = x(\Phi(y + x) - \Phi(q_h(x, y))),$$

this can be rewritten as

$$Y_{h,1}(x) = F_h(x, Y_{h,1}(x)).$$

The function F_h is analytic on $\Xi_\varepsilon^{(1,2)}$ by Lemma 3.8 and the fact that Φ is analytic for these arguments. Note also that

$$\lim_{h\to\infty} F_h(x, y) = x\big(\Phi(y + x) - 1\big) =: F_\infty(x, y)$$

pointwise for $(x, y) \in \Xi_\varepsilon^{(1,2)}$. By the estimate on q_h in Lemma 3.8, we also have

$$F_h(x, y) = F_\infty(x, y) + O(\lambda^h), \tag{26}$$

uniformly for $(x, y) \in \Xi_\varepsilon^{(1,2)}$. Using the same argument as in Lemma 3.5, we can also assume (redefining ε if necessary) that

$$\frac{\partial}{\partial y} F_h(x, y) = \frac{\partial}{\partial y} F_\infty(x, y) + O(\lambda^h) \tag{27}$$

and analogous estimates for any finite number of partial derivatives hold as well. Having reduced the original system of equations to a single equation for $Y_{h,1}(x)$, we now deduce properties of its dominant singularity. Since $Y_{h,1}(x)$ has a power series with nonnegative coefficients, by Pringsheim's theorem it must have a dominant positive real singularity that we denote by ρ_h. Since the coefficients of $Y_{h,1}(x)$ are bounded above by those of $Y(x)$, we also know that $\rho_h \ge \rho$.

Lemma 3.9. *For every sufficiently large h, $\rho_h \le \rho + \lambda^{h/2}$. Moreover, $\eta_{h,1} := Y_{h,1}(\rho_h) = \tau - \rho + O(\lambda^{h/2})$.*

Proof. Note first that $Y_{h,1}(x)$ is an increasing function of x for positive real $x < \rho_h$. Let $\tilde{\rho} = \min(\rho_h, \rho + \frac{\varepsilon}{2})$. Suppose first that $\lim_{x\to\tilde{\rho}^-} Y_{h,1}(x) \geq \tau - \rho + \frac{\varepsilon}{2}$. If h is large enough, this implies together with (27) that

$$\begin{aligned}\lim_{x\to\tilde{\rho}^-} \frac{\partial F_h}{\partial y}(x, Y_{h,1}(x)) &= \lim_{x\to\tilde{\rho}^-} \frac{\partial F_\infty}{\partial y}(x, Y_{h,1}(x)) + O(\lambda^h)\\ &\geq \frac{\partial F_\infty}{\partial y}\Big(\rho, \tau - \rho + \frac{\varepsilon}{2}\Big) + O(\lambda^h)\\ &= \rho\Phi'\Big(\tau + \frac{\varepsilon}{2}\Big) + O(\lambda^h) > \rho\Phi'(\tau) = 1.\end{aligned}$$

On the other hand, we also have

$$\begin{aligned}\frac{\partial F_h}{\partial y}(\rho/2, Y_{h,1}(\rho/2)) &= \frac{\partial F_\infty}{\partial y}(\rho/2, Y_{h,1}(\rho/2)) + O(\lambda^h)\\ &\leq \frac{\partial F_\infty}{\partial y}(\rho/2, Y(\rho/2) - \rho/2) + O(\lambda^h)\\ &< \rho\Phi'(\tau) = 1,\end{aligned}$$

so by continuity there must exist some $x_0 \in (\rho/2, \tilde{\rho})$ such that

$$\frac{\partial F_h}{\partial y}(x_0, Y_{h,1}(x_0)) = 1.$$

Moreover, if h is large enough we have

$$\frac{\partial^2 F_h}{\partial y^2}(x_0, Y_{h,1}(x_0)) = \frac{\partial^2 F_\infty}{\partial y^2}(x_0, Y_{h,1}(x_0)) + O(\lambda^h) > 0$$

as x_0 and thus also $Y_{h,1}(x_0)$ are bounded below by positive constants, and analogously $\frac{\partial F_h}{\partial x}(x_0, Y_{h,1}(x_0)) > 0$. But this would mean that $Y_{h,1}$ has a square root singularity at $x_0 < \rho_h$ (compare the discussion in Section 3.3 later), and we reach a contradiction. Hence we can assume that

$$\lim_{x\to\tilde{\rho}^-} Y_{h,1}(x) < \tau - \rho + \frac{\varepsilon}{2}. \tag{28}$$

Assume next that $\rho_h > \rho + \lambda^{h/2}$. Now for $x_1 = \rho + \lambda^{h/2} < \tilde{\rho}$ (the inequality holds if h is large enough to make $\lambda^{h/2} < \frac{\varepsilon}{2}$), $u_1 = Y_{h,1}(x_1) + x_1$ satisfies

$$u_1 = x_1\Phi(u_1) + O(\lambda^h), \tag{29}$$

since $F_h(x, y) = F_\infty(x, y) + O(\lambda^h) = x(\Phi(y + x) - 1) + O(\lambda^h)$. Note here that $u_1 \leq \tau + \frac{\varepsilon}{2} + \lambda^{h/2}$ by (28), thus u_1 is in the region of analyticity of Φ (again assuming h to be large enough). However, since $u \leq \rho\Phi(u)$ for all positive real u for which $\Phi(u)$ is well-defined (the line $u \mapsto \frac{u}{\rho}$ is a tangent to the graph of the convex function Φ at τ), for sufficiently large h the right-hand side in (29) is necessarily greater than the left, and we reach another contradiction. So it follows that $\rho_h \leq \rho + \lambda^{h/2}$, and in particular $\tilde{\rho} = \rho_h < \rho + \frac{\varepsilon}{2}$ if h is large enough. Since we know that $\lim_{x\to\tilde{\rho}^-} Y_{h,1}(x) < \tau - \rho + \frac{\varepsilon}{2}$, we also have $\eta_{h,1} := Y_{h,1}(\rho_h) < \tau - \rho + \frac{\varepsilon}{2}$. We conclude that $(\rho_h, \eta_{h,1}) \in \Xi_\varepsilon^{(1,2)}$, i.e., $(\rho_h, \eta_{h,1})$ lies within the region of analyticity of F_h. So the singularity at ρ_h must be due to the implicit function theorem failing at this point:

$$\eta_{h,1} = F_h(\rho_h, \eta_{h,1}) \text{ and } 1 = \frac{\partial F_h}{\partial y}(\rho_h, \eta_{h,1}).$$

The second equation in particular gives us

$$\rho_h\Phi'(\eta_{h,1} + \rho_h) = 1 + O(\lambda^h)$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

by (27). Since Φ' is increasing for positive real arguments and we know that $\rho\Phi'(\tau) = 1$ and $\rho_h = \rho + O(\lambda^{h/2})$, we can conclude from this that $\eta_{h,1} = \tau - \rho + O(\lambda^{h/2})$. □

As we have established that $\eta_{h,1} \to \tau - \rho$ as $h \to \infty$, we will use the abbreviation $\eta_1 := \tau - \rho$ in the following. This will later be generalised to $\eta_{h,k} := Y_{h,k}(\rho_h) \to \eta_k$, see Sections 4 and 5. For our next step, we need a multidimensional generalisation of Rouché's theorem:

Theorem 3.10 (see [1, p. 20, Theorem 2.5]). *Let Ω be a bounded domain in $\mathbb{C}^n$ whose boundary $\partial\Omega$ is piecewise smooth. Suppose that $u, v : \overline{\Omega} \to \mathbb{C}^n$ are analytic functions, and that the boundary of Ω does not contain any zeros of u. Moreover, assume that for every $z \in \partial\Omega$, there is at least one coordinate j for which $|u_j(z)| > |v_j(z)|$ holds. Then u and $u + v$ have the same number of zeros in Ω.*

Lemma 3.11. *If ε is chosen sufficiently small and h sufficiently large, then the pair $(\rho_h, \eta_{h,1})$ is the only solution to the simultaneous equations $F_h(x, y) = y$ and $\frac{\partial}{\partial y}F_h(x, y) = 1$ with $(x, y) \in \Xi_\varepsilon^{(1,2)}$.*

Proof. Note that (ρ, η_1) is a solution to the simultaneous equations $F_\infty(x, y) = x(\Phi(x + y) - 1) = y$ and $\frac{\partial}{\partial y}F_\infty(x, y) = x\Phi'(x + y) = 1$, and that there is no other solution with $|x| \le \rho + \varepsilon$ and $|y| \le \eta_1 + \varepsilon$ if ε is chosen sufficiently small by our assumptions on the function Φ (see Section 2.1). We take $\Omega = \Xi_\varepsilon^{(1,2)}$ in Theorem 3.10 and set

$$u(x, y) = \Big(F_\infty(x, y) - y, \frac{\partial}{\partial y}F_\infty(x, y) - 1\Big).$$

Moreover, take

$$v(x, y) = \Big(F_h(x, y) - F_\infty(x, y), \frac{\partial}{\partial y}F_h(x, y) - \frac{\partial}{\partial y}F_\infty(x, y)\Big).$$

Note that both coordinates of v are $O(\lambda^h)$ by (26) and (27). Since the boundary $\partial\Omega$ contains no zeros of u, if we choose h sufficiently large, then the conditions of Theorem 3.10 are satisfied. Consequently, u and $u + v$ have the same number of zeros in Ω, namely 1. Solutions to the simultaneous equations $F_h(x, y) = y$ and $\frac{\partial}{\partial y}F_h(x, y) = 1$ are precisely zeros of $u + v$, so this completes the proof. □

At this point, it already follows from general principles (see the discussion in [10, Chapter VII.4]) that for every sufficiently large h, $Y_{h,1}$ has a dominant square root singularity at ρ_h, and is otherwise analytic in a domain of the form (6). As we will need uniformity of the asymptotic expansion and a uniform bound for the domain of analyticity, we will make this more precise in the following section.

3.3 Asymptotic expansion and area of analyticity

Lemma 3.12. *Let $\varepsilon > 0$ be such that all previous lemmata hold. There exist $\delta_1, \delta_2 > 0$, some positive number h_0, and analytic functions R_h on $D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1})$ and S_h on $D_{\delta_2}(\eta_{h,1})$ for $h \ge h_0$ such that $\delta_2 < \varepsilon$, $D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1}) \subseteq \Xi_\varepsilon^{(1,2)}$ and*

$$F_h(x, y) - y = (x - \rho_h)R_h(x, y) + (y - \eta_{h,1})^2 S_h(y) \tag{30}$$

holds for $(x, y) \in D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1})$ and $h \ge h_0$ and such that $|R_h|$ is bounded from above and below by positive constants on $D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1})$ for $h \ge h_0$ (uniformly in h) and $|S_h|$ is bounded from above and below by positive constants on $D_{\delta_2}(\eta_{h,1})$ for $h \ge h_0$ (uniformly in h).

Furthermore, the sequences R_h and S_h converge uniformly to some analytic functions R and S, respectively. The same holds for their partial derivatives.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Proof. Recall that we can approximate partial derivatives of F_h by those of F_∞ with an exponential error bound (as in (27)), giving us

$$\begin{aligned}\frac{\partial}{\partial x}F_h(x,y) &= \frac{\partial}{\partial x}F_\infty(x,y) + O(\lambda^h)\\ &= \frac{\partial F_\infty}{\partial x}(\rho,\eta_1) + O(\lambda^h) + O(x-\rho) + O(y-\eta_1)\\ &= \Phi(\tau) + O(\lambda^h) + O(x-\rho) + O(y-\eta_1),\end{aligned}$$

as well as

$$\begin{aligned}\frac{\partial^2}{\partial y^2}F_h(x,y) &= \frac{\partial^2}{\partial y^2}F_\infty(x,y) + O(\lambda^h)\\ &= \frac{\partial^2 F_\infty}{\partial y^2}(\rho,\eta_1) + O(\lambda^h) + O(x-\rho) + O(y-\eta_1)\\ &= \rho\Phi''(\tau) + O(\lambda^h) + O(x-\rho) + O(y-\eta_1)\end{aligned}$$

for (x,y) in a neighbourhood of (ρ,η_1) contained in $\Xi_\varepsilon^{(1,2)}$ and $h\to\infty$.

Using Lemma 3.9, we choose $\delta_1>0$ and $\delta_2>0$ small enough and h_0 large enough such that $|x-\rho_h|\le\delta_1$, $|y-\eta_{h,1}|\le\delta_2$, and $h\ge h_0$ imply that

$$\left|\frac{\partial}{\partial x}F_h(x,y)-\Phi(\tau)\right| \le \frac{1}{2}\Phi(\tau) \tag{31}$$

and

$$\left|\frac{\partial^2}{\partial y^2}F_h(x,y)-\rho\Phi''(\tau)\right| \le \frac{1}{2}\rho\Phi''(\tau), \tag{32}$$

and such that $\overline{D_{\delta_1}(\rho_h)\times D_{\delta_2}(\eta_{h,1})}\subseteq \Xi_\varepsilon^{(1,2)}$. By Lemma 3.11, we have

$$F_h(\rho_h,\eta_{h,1}) = \eta_{h,1}, \tag{33}$$

$$\frac{\partial F_h}{\partial y}(\rho_h,\eta_{h,1}) = 1. \tag{34}$$

We now define

$$S_h(y) := \frac{F_h(\rho_h,y)-y}{(y-\eta_{h,1})^2}$$

for $y\in\overline{D_{\delta_2}(\eta_{h,1})}\setminus\{\eta_{h,1}\}$. By (33) and (34), S_h has a removable singularity at $\eta_{h,1}$. Therefore it is analytic on $D_{\delta_2}(\eta_{h,1})$. By (33), we have

$$\begin{aligned}F_h(\rho_h,y)-y &= (F_h(\rho_h,y)-y)-(F_h(\rho_h,\eta_{h,1})-\eta_{h,1})\\ &= \int_{\eta_{h,1}}^{y}\left(\frac{\partial}{\partial w}F_h(\rho_h,w)-1\right)dw.\end{aligned}$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

By (34), this can be rewritten as

$$\begin{aligned}
F_h(\rho_h, y) - y &= \int_{\eta_{h,1}}^{y} \Big(\Big(\frac{\partial F_h}{\partial y}(\rho_h, w) - 1 \Big) - \Big(\frac{\partial F_h}{\partial y}(\rho_h, \eta_{h,1}) - 1 \Big) \Big)\, dw \\
&= \int_{\eta_{h,1}}^{y} \int_{\eta_{h,1}}^{w} \frac{\partial^2 F_h}{\partial y^2}(\rho_h, v)\, dv\, dw \\
&= \int_{\eta_{h,1}}^{y} \int_{\eta_{h,1}}^{w} \rho \Phi''(\tau)\, dv\, dw + \int_{\eta_{h,1}}^{y} \int_{\eta_{h,1}}^{w} \Big(\frac{\partial^2 F_h}{\partial y^2}(\rho_h, v) - \rho \Phi''(\tau) \Big)\, dv\, dw \\
&= \frac{1}{2} \rho \Phi''(\tau)(y - \eta_{h,1})^2 + \int_{\eta_{h,1}}^{y} \int_{\eta_{h,1}}^{w} \Big(\frac{\partial^2 F_h}{\partial y^2}(\rho_h, v) - \rho \Phi''(\tau) \Big)\, dv\, dw.
\end{aligned}$$

Rearranging and using the definition of $S_h(y)$ as well as (32) yields

$$\left| S_h(y) - \frac{1}{2} \rho \Phi''(\tau) \right| \leq \frac{1}{4} \rho \Phi''(\tau)$$

for all $y \in \overline{D_{\delta_2}(\eta_{h,1})}$ and $h \geq h_0$. Thus $|S_h(y)|$ is bounded from below and above by positive constants for every such y and h.

We now define $R_h(x, y)$ such that (30) holds, which is equivalent to

$$R_h(x, y) := \frac{F_h(x, y) - F_h(\rho_h, y)}{x - \rho_h}$$

for $x \in \overline{D_{\delta_1}(\rho_h)} \setminus \{\rho_h\}$ and $y \in \overline{D_{\delta_2}(\eta_{h,1})}$. We have

$$\begin{aligned}
F_h(\rho_h, y) - F_h(x, y) &= \int_{x}^{\rho_h} \frac{\partial F_h}{\partial x}(w, y)\, dw \\
&= \Phi(\tau)(\rho_h - x) + \int_{x}^{\rho_h} \Big(\frac{\partial F_h}{\partial x}(w, y) - \Phi(\tau) \Big)\, dw.
\end{aligned}$$

Rearranging and using the definition of $R_h(x, y)$ yields

$$\big| R_h(x, y) - \Phi(\tau) \big| \leq \frac{1}{2} \Phi(\tau)$$

by (31) for $x \in \overline{D_{\delta_1}(\rho_h)} \setminus \{\rho_h\}$ and $y \in \overline{D_{\delta_2}(\eta_{h,1})}$ and $h \geq h_0$. In other words, $|R_h(x, y)|$ is bounded from below and above by positive constants for these (x, y) and h.

To prove analyticity of R_h, we use Cauchy's formula to rewrite it as

$$R_h(x, y) = \frac{1}{2\pi i} \oint_{|\zeta - \rho_h| = \delta_1} \frac{F_h(\zeta, y) - F_h(\rho_h, y)}{\zeta - \rho_h} \frac{d\zeta}{\zeta - x}$$

for $x \neq \rho_h$ (note that the integrand has a removable singularity at $\zeta = \rho_h$ in this case). The integral is also defined for $x = \rho_h$ and clearly defines an analytic function on $D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1})$ whose absolute value is bounded from above and below by a constant.

To see uniform convergence of R_h, we use Cauchy's formula once more and get

$$R_h(x, y) = \frac{1}{(2\pi i)^2} \oint_{|\zeta - \rho_h| = \delta_1} \oint_{|\eta - \eta_{h,1}| = \delta_2} \frac{F_h(\zeta, \eta) - F_h(\rho_h, \eta)}{\zeta - \rho_h} \frac{d\eta}{\eta - y} \frac{d\zeta}{\zeta - x} \tag{35}$$

for $x \in D_{\delta_1}(\rho_h)$ and $y \in D_{\delta_2}(\eta_{h,1})$. Without loss of generality, h_0 is large enough such that $|\rho_h - \rho| < \delta_1/4$ and $|\eta_{h,1} - \eta_1| < \delta_2/4$. By Cauchy's theorem, we can change the contour of integration such that (35) implies

$$R_h(x, y) = \frac{1}{(2\pi i)^2} \oint_{|\zeta - \rho| = \delta_1/2} \oint_{|\eta - \eta_1| = \delta_2/2} \frac{F_h(\zeta, \eta) - F_h(\rho_h, \eta)}{\zeta - \rho_h} \frac{d\eta}{\eta - y} \frac{d\zeta}{\zeta - x}$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

for $x \in D_{\delta_1/4}(\rho)$ and $y \in D_{\delta_2/4}(\eta_1)$, as the deformation is happening within the region of analyticity of the integrand. Using (26) and the fact that the denominator of the integrand is bounded away from zero shows that

$$R_h(x,y) = \frac{1}{(2\pi i)^2} \oint_{|\zeta-\rho|=\delta_1/2} \oint_{|\eta-\eta_1|=\delta_2/2} \frac{F_\infty(\zeta,\eta)-F_\infty(\rho_h,\eta)}{\zeta-\rho_h} \frac{d\eta}{\eta-y} \frac{d\zeta}{\zeta-x} + O(\lambda^h)$$

for $x \in D_{\delta_1/4}(\rho)$ and $y \in D_{\delta_2/4}(\eta_1)$. By Lemma 3.9, replacing the remaining occurrences of ρ_h by ρ induces another error term of $O(\lambda^{h/2})$, so that we get

$$R_h(x,y) = R(x,y) + O(\lambda^{h/2})$$

with

$$R(x,y) := \frac{1}{(2\pi i)^2} \oint_{|\zeta-\rho|=\delta_1/2} \oint_{|\eta-\eta_1|=\delta_2/2} \frac{F_\infty(\zeta,\eta)-F_\infty(\rho,\eta)}{\zeta-\rho} \frac{d\eta}{\eta-y} \frac{d\zeta}{\zeta-x}$$

for $x \in D_{\delta_1/4}(\rho)$ and $y \in D_{\delta_2/4}(\eta_1)$. Of course, the O constants do not depend on x and y; therefore, we have uniform convergence. Analogously, we get

$$\begin{aligned} S_h(y) &= \frac{1}{2\pi i} \oint_{|\eta-\eta_{h,1}|=\delta_2} \frac{F_h(\rho_h,\eta)-\eta}{(\eta-\eta_{h,1})^2} \frac{d\eta}{\eta-y} && (36)\\ &= S(y) + O(\lambda^{h/2}) && (37) \end{aligned}$$

with

$$S(y) := \frac{1}{2\pi i} \oint_{|\eta-\eta_1|=\delta_2/2} \frac{F_h(\rho,\eta)-\eta}{(\eta-\eta_1)^2} \frac{d\eta}{\eta-y},$$

for $y \in D_{\delta_2/4}(\eta_1)$. Analogous results hold for partial derivatives.

We replace δ_1 by $\delta_1/4$ and δ_2 by $\delta_2/4$ to get the result as stated in the lemma. □

Lemma 3.13. *The constants δ_1, δ_2 and h_0 in Lemma 3.12 can be chosen such that whenever $y = F_h(x,y)$ for some $(x,y) \in D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1})$ and some $h \geq h_0$, we have $|y-\eta_{h,1}| < \delta_2/2$.*

Proof. We first choose δ_1 and δ_2 as in Lemma 3.12. Then $y = F_h(x,y)$ and Lemma 3.12 imply that

$$|y-\eta_{h,1}| = \sqrt{|x-\rho_h| \left|\frac{R_h(x,y)}{S_h(y)}\right|}.$$

The fraction on the right-hand side is bounded by some absolute constant according to Lemma 3.12. So by decreasing δ_1 if necessary, the right-hand side is at most $\delta_2/2$. □

Lemma 3.14. *Let $\varepsilon > 0$ be such that the previous lemmata hold. There exists $\delta_0 > 0$ such that, for all sufficiently large h, the asymptotic formula*

$$Y_{h,1}(x) = \eta_{h,1} + a_h\Big(1-\frac{x}{\rho_h}\Big)^{1/2} + b_h\Big(1-\frac{x}{\rho_h}\Big) + c_h\Big(1-\frac{x}{\rho_h}\Big)^{3/2} + O\Big((\rho_h-x)^2\Big) \quad (38)$$

holds for $x \in D_{\delta_0}(\rho_h)$ with $|\mathrm{Arg}(x-\rho_h)| \geq \pi/4$ and certain sequences a_h, b_h and c_h. The O-constant is independent of h, and a_h, b_h, c_h converge to the coefficients a, b, c in (4) at an exponential rate as $h \to \infty$. Additionally, $|Y_{h,1}(x)-\eta_1| < \varepsilon/2$ for all these x.

Proof. By (30), the function $Y_{h,1}$ is determined by the implicit equation

$$0 = F_h(x, Y_{h,1}(x)) - Y_{h,1}(x) = (x-\rho_h)R_h(x, Y_{h,1}(x)) + (Y_{h,1}(x)-\eta_{h,1})^2 S_h(Y_{h,1}(x)). \quad (39)$$

For $r > 0$, set $C(r) := \{x \in D_r(\rho_h) \colon |\mathrm{Arg}(x-\rho_h)| \geq \pi/4\}$ and $\widetilde{C}(r) := \{x \in \mathbb{C} \colon |x-\rho_h| = r$ and $|\mathrm{Arg}(x-\rho_h)| \geq \pi/4\}$. Choose δ_1, δ_2, h_0 as in Lemma 3.13. For some $h \geq h_0$, let r_h be the

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

supremum of all $r \le \delta_1$ such that $Y_{h,1}$ can be continued analytically to $C(r)$ with values in $D_{\delta_2/2}(\eta_{h,1})$. We claim that $r_h = \delta_1$.

Suppose for contradiction that $r_h < \delta_1$ and let $x_\infty \in \widetilde{C}(r_h)$. Choose a sequence of elements $x_n \in C(r_h)$ converging to x_∞ for $n \to \infty$ and set $y_n := Y_{h,1}(x_n)$ for all n. By assumption, we have $|y_n - \eta_{h,1}| \le \delta_2/2$. By replacing the sequence x_n by a subsequence if necessary, we may assume that the sequence y_n is convergent to some limit y_∞. Note that $|y_\infty - \eta_{h,1}| \le \delta_2/2$. By continuity of F_h, we also have $y_\infty = F_h(x_\infty, y_\infty)$. As $(x_\infty, y_\infty) \in \Xi_\varepsilon^{(1,2)}$ with $x_\infty \ne \rho_h$, Lemma 3.11 and the analytic implicit function theorem imply that $Y_{h,1}$ can be continued analytically in a suitable open neighbourhood of x_∞. This neighbourhood can be chosen small enough such that the inequality $|Y_{h,1}(x) - \eta_{h,1}| \le \delta_2$ holds for all x in this neighbourhood. However, Lemma 3.13 implies that we then actually have $|Y_{h,1}(x) - \eta_{h,1}| \le \delta_2/2$ for all such x.

The set of these open neighbourhoods associated with all $x_\infty \in \widetilde{C}(r_h)$ covers the compact set $\widetilde{C}(r_h)$, so a finite subset of these open neighbourhoods can be selected. Thus we find an analytic continuation of $Y_{h,1}$ to $C(\widetilde{r}_h)$ for some $\widetilde{r}_h \in (r_h, \delta_1)$ with values still in $D_{\delta_2/2}(\eta_{h,1})$, which is a contradiction to the choice of r_h.

Thus we have $r_h = \delta_1$. In particular, choosing h large enough that $|\eta_{h,1} - \eta_1| < (\varepsilon - \delta_2)/2$ gives $|Y_{h,1}(x) - \eta_1| \le |Y_{h,1}(x) - \eta_{h,1}| + |\eta_{h,1} - \eta_1| < \delta_2/2 + (\varepsilon - \delta_2)/2 = \varepsilon/2$ for all $x \in C(\delta_1)$.

Rearranging (39) yields

$$(\eta_{h,1} - Y_{h,1}(x))^2 = \left(\rho_h \frac{R_h(x, Y_{h,1}(x))}{S_h(Y_{h,1}(x))}\right)\left(1 - \frac{x}{\rho_h}\right). \tag{40}$$

We know from Lemma 3.12 that R_h is bounded above and S_h is bounded below on $D_{\delta_1}(\rho_h) \times D_{\delta_2}(\eta_{h,1})$ and $D_{\delta_2}(\eta_{h,1})$, respectively. Therefore, the absolute value of the first factor on the right-hand side of (40) is bounded above and below by positive constants for $x \in D_{\delta_1}(\rho_h)$. For $x < \rho_h$, we have that the factor $(1 - x/\rho_h)$ is trivially positive and that $\eta_{h,1} > Y_{h,1}(x)$ because $Y_{h,1}$ is strictly increasing on $(0, \rho_h)$, so the first factor on the right-hand side of (40) must be positive. Thus we may take the principal value of the square root to rewrite (40) as

$$\eta_{h,1} - Y_{h,1}(x) = \sqrt{\rho_h \frac{R_h(x, Y_{h,1}(x))}{S_h(Y_{h,1}(x))}}\left(1 - \frac{x}{\rho_h}\right)^{1/2} \tag{41}$$

for $x \in C(\delta_1)$. The above considerations also show that the radicand in (41) remains positive in the limit $x \to \rho_h^-$ (i.e., as x approaches ρ_h from the left) and then for $h \to \infty$.

As we just observed that the first factor on the right-hand side of (41) is bounded, (41) implies

$$Y_{h,1}(x) - \eta_{h,1} = O\big((x - \rho_h)^{1/2}\big), \tag{42}$$

with an O-constant that is independent of h. We can now iterate this argument: using Taylor expansion along with the fact that partial derivatives of R_h and S_h are uniformly bounded above while S_h is also uniformly bounded below, we obtain

$$\begin{aligned}\frac{R_h(x, Y_{h,1}(x))}{S_h(Y_{h,1}(x))} &= \frac{R_h(\rho_h, \eta_{h,1}) + O(x - \rho_h) + O(Y_{h,1}(x) - \eta_{h,1})}{S_h(\eta_{h,1}) + O(Y_{h,1}(x) - \eta_{h,1})} \\ &= \frac{R_h(\rho_h, \eta_{h,1})}{S_h(\eta_{h,1})} + O\big((x - \rho_h)^{1/2}\big).\end{aligned}$$

Plugging this into (41) yields

$$\eta_{h,1} - Y_{h,1}(x) = \sqrt{\rho_h \frac{R_h(\rho_h, \eta_{h,1})}{S_h(\eta_{h,1})}}\left(1 - \frac{x}{\rho_h}\right)^{1/2} + O(x - \rho_h),$$

still with an O-constant that is independent of h. This can be continued arbitrarily often to obtain further terms of the expansion and an improved error term (for our purposes, it is enough to stop

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

at $O((x-\rho_h)^2)$). Indeed it is well known (cf. [10, Lemma VII.3]) that an implicit equation of the form (39) has a solution as a power series in $(1-x/\rho_h)^{1/2}$. In particular, (38) follows with an error term that is uniform in h. The coefficients a_h, b_h, c_h can be expressed in terms of R_h, S_h and their partial derivatives evaluated at $(\rho_h, \eta_{h,1})$: specifically,

$$\begin{aligned}
a_h &= -\sqrt{\rho_h \frac{R_h(\rho_h, \eta_{h,1})}{S_h(\eta_{h,1})}},\\
b_h &= \frac{\rho_h S_h(\eta_{h,1}) \frac{\partial R_h}{\partial y}(\rho_h, \eta_{h,1}) - \rho_h S_h'(\eta_{h,1}) R_h(\rho_h, \eta_{h,1})}{2 S_h(\eta_{h,1})^2},\\
c_h &= \frac{\rho_h^{3/2} N}{8\sqrt{R_h(\rho_h, \eta_{h,1}) S_h(\eta_{h,1})^7}},
\end{aligned}$$

where the numerator N is a polynomial in $R_h(\rho_h, \eta_{h,1})$, $S_h(\eta_{h,1})$ and their derivatives. By Lemma 3.12, R_h and S_h as well as their partial derivatives converge uniformly to R and S as well as their partial derivatives, respectively, with an error bound of $O(\lambda^{h/2})$. We also know that ρ_h and $\eta_{h,1}$ converge exponentially to ρ and η_1, respectively, see Lemma 3.9. This means that first replacing all occurrences of R_h and S_h by R and S, respectively, and then replacing all occurrences of ρ_h and $\eta_{h,1}$ by ρ and η_1, respectively, shows that $a_h = a + O(\lambda^{h/2})$, $b_h = b + O(\lambda^{h/2})$, and $c_h = c + O(\lambda^{h/2})$ where a, b, and c are the results of these replacements. Taking the limit for $h \to \infty$ in (30) shows that R and S and therefore a, b, and c play the same role with respect to F_∞ as R_h, S_h, a_h, b_h, and c_h play with respect to F_h, which implies that a, b, and c are indeed the constants from (4). □

Having dealt with the behaviour around the singularity, it remains to prove a uniform bound on $Y_{h,1}$ in a domain of the form (6) for fixed δ.

Lemma 3.15. *Let $\varepsilon > 0$ be such that all previous lemmata hold. There exist $\delta > 0$ and a positive integer h_0 such that $Y_{h,1}(x)$ has an analytic continuation to the domain*

$$\{x \in \mathbb{C} : |x| \le (1+\delta)|\rho_h|, |\mathrm{Arg}(x/\rho_h - 1)| > \pi/4\}$$

for all $h \ge h_0$, and has the uniform upper bound

$$|Y_{h,1}(x)| \le \tau - \rho + \frac{\varepsilon}{2} = \eta_1 + \frac{\varepsilon}{2}$$

for all $h \ge h_0$ and all x.

Proof. Let us define $r_h = \sup \mathcal{R}_h$, where

$$\mathcal{R}_h = \left\{ r : Y_{h,1} \text{ extends analytically to } D_r(0) \setminus D_{\delta_0}(\rho_h) \text{ and satisfies } |Y_{h,1}(x)| < \eta_1 + \frac{\varepsilon}{2} \text{ there} \right\},$$

with δ_0 as in the previous lemma. Note that trivially, $r_h \ge \rho$. If $\liminf_{h\to\infty} r_h > \rho$, we are done: in this case, there is some $\delta > 0$ such that $Y_{h,1}$ extends analytically to $D_{\rho(1+\delta)}(0) \setminus D_{\delta_0}(\rho_h)$ and satisfies $|Y_{h,1}(x)| < \eta_1 + \frac{\varepsilon}{2}$ there. As the previous lemma covers $D_{\delta_0}(\rho_h)$, this already completes the proof.

So let us assume that $\liminf_{h\to\infty} r_h = \rho$ and derive a contradiction. The assumption implies that there is an increasing sequence of positive integers h_j such that $\lim_{j\to\infty} r_{h_j} = \rho$. Without loss of generality, we may assume that $r_{h_j} \le \rho + \frac{\varepsilon}{2}$ for all j. Pick (for each sufficiently large j) a point x_{h_j} with $|x_{h_j}| = r_{h_j}$ and $|Y_{h_j,1}(x_{h_j})| = \eta_1 + \frac{\varepsilon}{2}$. If this were not possible, we could analytically continue $Y_{h_j,1}$ at every point x with $|x| = r_{h_j}$ and $x \notin D_{\delta_0}(\rho_{h_j})$ to a disk where $Y_{h_j,1}$ is still bounded by $\eta_1 + \frac{\varepsilon}{2}$. This analytic continuation is possible, since by Lemma 3.11 the pair $(\rho_{h_j}, \eta_{h_j,1})$ is the only solution to the simultaneous equations $F_{h_j}(x,y) = y$ and $\frac{\partial}{\partial y} F_{h_j}(x,y) = 1$ with $(x,y) \in \Xi_\varepsilon^{(1,2)}$, so the analytic implicit function theorem becomes applicable (compare e.g. the analytic continuation of q_h in

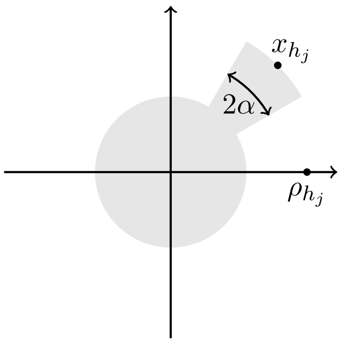

Figure 5. Illustration of the domain $x_{h_j}A$.

Lemma 3.8). By compactness, this would allow us to extend $Y_{h_j,1}$ to $D_r(0) \setminus D_{\delta_0}(\rho_{h_j})$ for some $r > r_{h_j}$ while still maintaining the inequality $|Y_{h_j,1}(x)| < \eta_1 + \frac{\varepsilon}{2}$, contradicting the choice of r_{h_j}.

Without loss of generality (choosing a subsequence if necessary), we can assume that x_{h_j} and $Y_{h_j,1}(x_{h_j})$ have limits x_∞ and y_∞, respectively. By construction, $|x_\infty| = \rho$ and $|y_\infty| = \eta_1 + \frac{\varepsilon}{2}$.

Since $x_{h_j} \notin D_{\delta_0}(\rho)$ for all j, $\operatorname{Arg} x_{h_j}$ is bounded away from 0. Thus we can find $\alpha > 0$ such that $|\operatorname{Arg} x_{h_j}| \geq 2\alpha$ for all j. Define the region A by

$$A = \left\{ z \in \mathbb{C} : |z| < \frac{1}{2} \text{ or } (|z| < 1 \text{ and } |\operatorname{Arg} z| < \alpha) \right\}.$$

Note that $x_{h_j}A$ avoids the part of the real axis that includes ρ_{h_j} (see Fig. 5), so the function $Y_{h_j,1}(x)$ is analytic in this region for all j by construction since $(x, Y_{h_j,1}(x)) \in \Xi_\varepsilon^{(1,2)}$ whenever $x \in x_{h_j}A$. So we have a sequence of functions $W_j(z) := Y_{h_j,1}(x_{h_j}z)$ that are all analytic on A and are uniformly bounded above by $\eta_1 + \frac{\varepsilon}{2}$ by our choice of x_{h_j}. By Montel's theorem, there is a subsequence of these functions (without loss of generality the sequence itself) that converges locally uniformly and thus to an analytic function W_∞ on A. This function needs to satisfy the following:

- $W_\infty(0) = 0$, since $W_j(0) = 0$ for all j,
- $W_\infty(z) = F_\infty(x_\infty z, W_\infty(z)) = x_\infty z(\Phi(x_\infty z + W_\infty(z)) - 1)$ for $z \in A$, since we have the uniform estimate

  $$W_j(z) = Y_{h_j,1}(x_{h_j}z) = F_{h_j}(x_{h_j}z, Y_{h_j,1}(x_{h_j}z)) = F_\infty(x_{h_j}z, Y_{h_j,1}(x_{h_j}z)) + O(\lambda^h).$$

  This is also equivalent to

  $$x_\infty z + W_\infty(z) = x_\infty z \Phi(x_\infty z + W_\infty(z)).$$

These two properties imply that $W_\infty(z) = Y(x_\infty z) - x_\infty z$, since Y is the unique function that is analytic at 0 and satisfies the implicit equation $Y(x) = x\Phi(Y(x))$. Implicit differentiation of $Y_{h_j,1}(x) = F_{h_j}(x, Y_{h_j,1}(x))$ for $x \in x_{h_j}A$ yields

$$Y'_{h_j,1}(x) = \frac{\frac{\partial F_{h_j}}{\partial x}(x, Y_{h_j,1}(x))}{1 - \frac{\partial F_{h_j}}{\partial y}(x, Y_{h_j,1}(x))} = \frac{\frac{\partial F_\infty}{\partial x}(x, Y_{h_j,1}(x)) + O(\lambda^{h_j})}{1 - \frac{\partial F_\infty}{\partial y}(x, Y_{h_j,1}(x)) + O(\lambda^{h_j})}. \tag{43}$$

Note that the numerator is uniformly bounded. Moreover, we recall again that the only solution to the simultaneous equations $F_\infty(x, y) = x(\Phi(y + x) - 1) = y$ and $\frac{\partial}{\partial y}F_\infty(x, y) = x\Phi'(y + x) = 1$ with $|x| \leq \rho + \varepsilon$ and $|x + y| \leq \tau + \varepsilon$ is $(x, y) = (\rho, \tau - \rho) = (\rho, \eta_1)$ by our assumptions on Φ.

By construction, there is a constant $\varepsilon_A > 0$ such that $|x-\rho| \geq \varepsilon_A$ whenever $x \in x_{h_j}A$ for some j. The map $(x, y) \mapsto \|(F_\infty(x, y) - y, \frac{\partial}{\partial y}F_\infty(x, y) - 1)\|$ is continuous on the compact set

$$\mathcal{K} := \{(x, y) : |x| \leq \rho + \varepsilon \text{ and } |x+y| \leq \tau + \varepsilon \text{ and } |x-\rho| \geq \varepsilon_A\}$$

and has no zero there (using the Euclidean norm on $\mathbb{C}^2$). Therefore, it attains a minimum $\delta_A > 0$ on $\mathcal{K}$.

Now for $x \in x_{h_j}A$, $|x| \leq \rho + \varepsilon$ holds by assumption, as does $|x + Y_{h_j,1}(x)| \leq \tau + \varepsilon$. Moreover, $|x-\rho| \geq \varepsilon_A$. Thus we can conclude that $(x, Y_{h_j,1}(x)) \in \mathcal{K}$ and therefore $\|(F_\infty(x, Y_{h_j,1}(x)) - Y_{h_j,1}(x), \frac{\partial F_\infty}{\partial y}(x, Y_{h_j,1}(x)) - 1)\| \geq \delta_A$ for all such x. Since

$$F_\infty(x, Y_{h_j,1}(x)) - Y_{h_j,1}(x) = F_{h_j}(x, Y_{h_j,1}(x)) - Y_{h_j,1}(x) + O(\lambda^{h_j}) = O(\lambda^{h_j}),$$

this means that $\left|1 - \frac{\partial F_\infty}{\partial y}(x, Y_{h_j,1}(x))\right| \geq \delta_A - O(\lambda^{h_j})$, so that the denominator in (43) is bounded below by a positive constant for sufficiently large j.

So we can conclude that $Y'_{h_j,1}(x)$ is uniformly bounded by a constant for $x \in x_{h_j}A$, implying that $W_j{}'(z)$ is uniformly bounded (for all $z \in A$ and all sufficiently large j) by a constant that is independent of j. Therefore, $W_j(z)$ is a uniformly equicontinuous sequence of functions on $\overline{A}$, the closure of A. By the Arzelà–Ascoli theorem, this implies that $W_j(z) \to W_\infty(z) = Y(x_\infty z) - x_\infty z$ holds even for all $z \in \overline{A}$, not only on A. In particular, $y_\infty = W_\infty(1) = Y(x_\infty) - x_\infty$. Here, we have $|x_\infty| \leq \rho$ and $|y_\infty| = \eta_1 + \frac{\varepsilon}{2}$ by assumption. However,

$$|Y(x) - x| \leq |Y(\rho) - \rho| = \eta_1$$

holds for all $|x| \leq \rho$ by the triangle inequality, so we finally reach a contradiction. □

We conclude this section with a summary of the results proven so far. The following proposition follows by combining the last two lemmata.

Proposition 3.16. *There exists a constant $\delta > 0$ such that $Y_{h,1}(x)$ can be continued analytically to the domain*

$$\{x \in \mathbb{C} : |x| \leq (1+\delta)|\rho_h|, |\operatorname{Arg}(x/\rho_h - 1)| > \pi/4\}$$

for every sufficiently large h. Moreover, $Y_{h,1}(x)$ is then uniformly bounded on this domain by a constant that is independent of h, and the following singular expansion holds near the singularity:

$$Y_{h,1}(x) = \eta_{h,1} + a_h\Big(1 - \frac{x}{\rho_h}\Big)^{1/2} + b_h\Big(1 - \frac{x}{\rho_h}\Big) + c_h\Big(1 - \frac{x}{\rho_h}\Big)^{3/2} + O\Big((\rho_h - x)^2\Big),$$

where the O-constant is independent of h and a_h, b_h, c_h converge at an exponential rate to a, b, c respectively as $h \to \infty$.

Remark 3.17. *Let $D = \gcd\{i \in \mathbb{N} \colon w_i \neq 0\}$ be the period of Φ. The purpose of this remark is to give indications how the results so far have to be adapted for the case $D > 1$.*

If $D > 1$, then for all trees of our simply generated family of trees, the number n of vertices will be congruent to 1 modulo D because all outdegrees are multiples of D. Trivially, the same is true for all trees with maximum protection number h.

By [10, Remark VI.17], both Y and $Y_{h,1}$ have D conjugate roots on its circle of convergence. Therefore, it is enough to study the positive root at the radius of convergence. Up to Theorem 3.10, no changes are required. In Lemma 3.11, there are exactly D solutions instead of exactly one solution to the simultaneous equations. Lemmata 3.12, 3.13, and 3.14 analyse the behaviour of $Y_{h,1}$ around the dominant positive singularity and remain valid without any change. In the proof of Lemma 3.15, we need to exclude balls around the conjugate roots. Proposition 3.16 must also be changed to exclude the conjugate roots.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

4. The exponential case: $w_1 \neq 0$

4.1 Asymptotics of the singularities

Proposition 3.16 that concluded the previous section shows that condition (2) of Theorem 2.1 is satisfied (with $\alpha = \frac{1}{2}$) by the generating functions $Y_{h,1}$ (and thus also $Y_{h,0}$, since $Y_{h,0}(x) = Y_{h,1}(x) + x$). It remains to study the behaviour of the singularity ρ_h of $Y_{h,0}$ and $Y_{h,1}$ to make the theorem applicable. As it turns out, condition (1) of Theorem 2.1 holds precisely if vertices of outdegree 1 are allowed in our simply generated family of trees. In terms of the weight generating function Φ, this can be expressed as $w_1 = \Phi'(0) \neq 0$. Starting with Lemma 4.3, we will assume that this holds. The case where vertices of outdegree 1 cannot occur (equivalently, $w_1 = \Phi'(0) = 0$) is covered in Section 5.

Let us define the auxiliary quantities $\eta_{h,k} := Y_{h,k}(\rho_h)$ for all $0 \le k \le h$. We know that these must exist and be finite for all sufficiently large h. Since the coefficients of $Y_{h,k}$ are nonincreasing in k in view of the combinatorial interpretation, we must have

$$\eta_{h,0} \ge \eta_{h,1} \ge \cdots \ge \eta_{h,h}. \tag{44}$$

Note also that the following system of equations holds:

$$\eta_{h,0} = \eta_{h,1} + \rho_h, \tag{45}$$

$$\eta_{h,k} = \rho_h \Phi(\eta_{h,k-1}) - \rho_h \Phi(\eta_{h,h}) \qquad \text{for } 1 \le k \le h, \tag{46}$$

in view of (8) and (7), respectively. Since $Y_{h,1}$ is singular at ρ_h by assumption, the Jacobian determinant of the system that determines $Y_{h,0}, Y_{h,1}, \ldots, Y_{h,h}$ needs to vanish (as there would otherwise be an analytic continuation by the analytic implicit function theorem). This determinant is given by

$$\begin{vmatrix} 1 & -1 & 0 \cdots & 0 & 0 \\ -\rho_h \Phi'(\eta_{h,0}) & 1 & 0 \cdots & 0 & \rho_h \Phi'(\eta_{h,h}) \\ 0 & -\rho_h \Phi'(\eta_{h,1}) & 1 \cdots & 0 & \rho_h \Phi'(\eta_{h,h}) \\ \vdots & \vdots & \vdots \ddots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & -\rho_h \Phi'(\eta_{h,h-1}) & 1 + \rho_h \Phi'(\eta_{h,h}) \end{vmatrix}.$$

Using column expansion with respect to the last column to obtain the determinant, we find that this simplifies to

$$\prod_{j=1}^{h} \left(\rho_h \Phi'(\eta_{h,j})\right) + \left(1 - \rho_h \Phi'(\eta_{h,0})\right)\left(1 + \sum_{k=2}^{h} \prod_{j=k}^{h} \left(\rho_h \Phi'(\eta_{h,j})\right)\right) = 0. \tag{47}$$

We will now use (45), (46), and (47) to determine an asymptotic formula for ρ_h. Throughout this section, B_i's will always be positive constants with $B_i < 1$ that depend on the specific family of simply generated trees, but nothing else.

Lemma 4.1. *There exist positive constants C and B_1 with $B_1 < 1$ such that $\eta_{h,k} \le C B_1^k$ for all sufficiently large h and all k with $0 \le k \le h$.*

Proof. Since we already know that $\eta_{h,1}$ converges to $\tau - \rho$ and that ρ_h converges to ρ, $\eta_{h,0}$ converges to τ by (45). By the monotonicity property (44), all $\eta_{h,k}$ must therefore be bounded by a single constant M for sufficiently large h. Since $\eta_{h,1}$ converges to $\tau - \rho$, we must have that $\rho_h \Phi'(\eta_{h,1})$ converges to $\rho \Phi'(\tau - \rho)$. Therefore, $\rho_h \Phi'(\eta_{h,1}) \le \rho \Phi'(\tau - \rho/2)$ for sufficiently large h. It follows that $\rho_h \Phi'(\eta_{h,1}) \le \rho \Phi'(\tau - \rho/2) < \rho \Phi'(\tau) = 1$. For all $1 \le j \le h$, we now have

$$\eta_{h,j} = \rho_h \Phi(\eta_{h,j-1}) - \rho_h \Phi(\eta_{h,h}) \le \rho_h \Phi'(\eta_{h,j-1})(\eta_{h,j-1} - \eta_{h,h}) \le \rho_h \Phi'(\eta_{h,1}) \eta_{h,j-1}.$$

Thus by induction

$$\eta_{h,k} \le \eta_{h,1}\big(\rho_h \Phi'(\eta_{h,1})\big)^{k-1} \le M\big(\rho\Phi'(\tau-\rho/2)\big)^{k-1}.$$

This proves the desired inequality for sufficiently large h and $1 \le k \le h$ with $B_1 = \rho\Phi'(\tau - \rho/2) < 1$, and we are done. □

With this bound, we will be able to refine the estimates for the system of equations, leading to better estimates for ρ_h and $\eta_{h,0}$. Recall from Lemma 3.9 that ρ_h and $\eta_{h,1}$ converge to their respective limits ρ and $\tau - \rho = \eta_1$ (at least) exponentially fast. Since $\eta_{h,0} = \eta_{h,1} + \rho_h$ by (45), this also applies to $\eta_{h,0}$. We show that an analogous statement also holds for $\eta_{h,k}$ with arbitrary k. In view of (46), it is natural to expect that $\eta_{h,k} \to \eta_k$, where η_k is defined recursively as follows: $\eta_0 = \tau$ and, for $k > 0$, $\eta_k = \rho\Phi(\eta_{k-1}) - \rho$, which also coincides with our earlier definition of $\eta_1 = \tau - \rho$. This is proven in the following lemma.

Lemma 4.2. *For a suitable constant $B_2 < 1$ and sufficiently large h, we have $\rho_h = \rho + O(B_2^h)$ and $\eta_{h,k} = \eta_k + O(B_2^h)$ for all k with $0 \le k \le h$, uniformly in k.*

Proof. For a suitable choice of B_2, the estimate for ρ_h has been established by Lemma 3.9, as has the estimate for $\eta_{h,k}$ in the cases where $k = 0$ and $k = 1$. Set $\delta_{h,k} = \eta_{h,k} - \eta_k$. Since $\eta_{h,h} \le CB_1^h$ by Lemma 4.1, we have $\Phi(\eta_{h,h}) = \Phi(0) + O(B_1^h) = 1 + O(B_1^h)$. Without loss of generality, suppose that $B_2 \ge B_1$. Then, using (46), we obtain

$$\begin{aligned}
\eta_{h,k} &= \rho_h\Phi(\eta_{h,k-1}) - \rho_h\Phi(\eta_{h,h}) \\
&= (\rho + O(B_2^h))\Phi(\eta_{k-1} + \delta_{h,k-1}) - (\rho + O(B_2^h))(1 + O(B_1^h)) \\
&= \rho(\Phi(\eta_{k-1}) + \Phi'(\xi_{h,k-1})\delta_{h,k-1}) - \rho + O(B_2^h) \\
&= \eta_k + \rho\Phi'(\xi_{h,k-1})\delta_{h,k-1} + O(B_2^h)
\end{aligned}$$

where $\xi_{h,k-1}$ is between η_{k-1} and $\eta_{h,k-1}$ (by the mean value theorem) and the O-constant is independent of k. Let M be this O-constant. We already know (compare the proof of Lemma 4.1) that $\eta_{h,k-1} \le \eta_{h,1} \le \tau - \rho/2$ for every $k \ge 2$ if h is sufficiently large. Likewise, it is easy to see that η_k is decreasing in k, hence $\eta_{k-1} \le \eta_1 = \tau - \rho$. Thus, $\xi_{h,k-1} \le \tau - \rho/2$ and $\rho\Phi'(\xi_{h,k-1}) \le \rho\Phi'(\tau - \rho/2) = B_1 < 1$. So we have, for every $k > 1$,

$$|\delta_{h,k}| = |\eta_{h,k} - \eta_k| \le B_1|\delta_{h,k-1}| + MB_2^h.$$

Iterating this inequality yields

$$|\delta_{h,k}| \le B_1^{k-1}|\delta_{h,1}| + (1 + B_1 + \cdots + B_1^{k-2})MB_2^h \le |\delta_{h,1}| + \frac{MB_2^h}{1 - B_1},$$

and the desired statement follows. □

From Lemma 4.1 and the fact that $\eta_{h,k} \to \eta_k$, we trivially obtain $\eta_k \le C \cdot B_1^k$, with the same constants B_1 and C as in Lemma 4.1. In fact, we can be more precise, and this is demonstrated in the lemma that follows. Since the expression $\rho\Phi'(0)$ occurs frequently in the following, we set $\zeta := \rho\Phi'(0)$. Recall that we assume $\Phi'(0) \ne 0$ until the end of this section.

Lemma 4.3. *The limit $\lambda_1 := \lim_{k\to\infty} \zeta^{-k}\eta_k$ exists. Moreover, we have*

$$\eta_k = \lambda_1\zeta^k(1 + O(B_1^k)),$$

with B_1 as in Lemma 4.1.

Proof. Recall that we defined the sequence $(\eta_k)_{k\ge 0}$ by $\eta_0 = \tau$ and $\eta_k = \rho\Phi(\eta_{k-1}) - \rho$ for $k \ge 1$. Using Taylor expansion, we obtain

$$\eta_k = \rho\Phi'(0)\eta_{k-1}(1 + O(\eta_{k-1})) = \zeta\eta_{k-1}(1 + O(\eta_{k-1})).$$

Since we already know that $\eta_{k-1} \leq C \cdot B_1^{k-1}$, this implies that

$$\eta_k = \zeta \eta_{k-1}(1 + O(B_1^k)).$$

Now it follows that the infinite product

$$\lambda_1 = \eta_0 \prod_{j \geq 1} \frac{\eta_j}{\zeta \eta_{j-1}} = \lim_{k \to \infty} \eta_0 \prod_{j=1}^{k} \frac{\eta_j}{\zeta \eta_{j-1}} = \lim_{k \to \infty} \zeta^{-k} \eta_k$$

converges. The error bound follows from noting that

$$\zeta^{-k} \eta_k = \lambda_1 \prod_{j \geq k+1} \frac{\zeta \eta_{j-1}}{\eta_j} = \lambda_1 \prod_{j \geq k+1} (1 + O(B_1^j)).$$

□

Next, we consider the expression in (47) and determine the asymptotic behaviour of its parts.

Lemma 4.4. *For large enough h and a fixed constant $B_3 < 1$, we have*

$$1 + \sum_{k=2}^{h} \prod_{j=k}^{h} (\rho_h \Phi'(\eta_{h,j})) = \frac{1}{1 - \zeta} + O(B_3^h)$$

and

$$\prod_{j=1}^{h} \rho_h \Phi'(\eta_{h,j}) = \lambda_2 \zeta^h (1 + O(B_3^h)),$$

where $\lambda_2 := \prod_{j \geq 1} \frac{\Phi'(\eta_j)}{\Phi'(0)}$.

Proof. Note that

$$\prod_{j=k}^{h} (\rho_h \Phi'(\eta_{h,j})) = \rho_h^{h-k+1} \prod_{j=k}^{h} \Phi'(\eta_{h,j}).$$

In view of Lemma 4.2, we have $\rho_h^{h-k+1} = \rho^{h-k+1}(1 + O(B_2^h))^{h-k+1} = \rho^{h-k+1}(1 + O(hB_2^h))$, uniformly in k. Moreover, Lemma 4.1 yields $\Phi'(\eta_{h,j}) = \Phi'(0) + O(\eta_{h,j}) = \Phi'(0) + O(B_1^j)$, uniformly in h. Thus

$$\prod_{j=k}^{h} \Phi'(\eta_{h,j}) = \Phi'(0)^{h-k+1} \prod_{j=k}^{h} (1 + O(B_1^j)) = \Phi'(0)^{h-k+1}(1 + O(B_1^k)).$$

Hence the expression simplifies to

$$1 + \sum_{k=2}^{h} \prod_{j=k}^{h} (\rho_h \Phi'(\eta_{h,j})) = 1 + (1 + O(hB_2^h)) \sum_{k=2}^{h} \zeta^{h-k+1}(1 + O(B_1^k)).$$

Since $\zeta < 1$ and $B_1 < 1$, we can simply evaluate the geometric series, and the expression further simplifies to

$$1 + \sum_{k=2}^{h} \zeta^{h-k+1} + O(B_3^h) = \frac{1 - \zeta^h}{1 - \zeta} + O(B_3^h) = \frac{1}{1 - \zeta} + O(B_3^h)$$

for an appropriately chosen $B_3 < 1$. This proves the first statement. For the second statement, we also use Lemma 4.2, along with the monotonicity of Φ' and the assumption that $\Phi'(0) \neq 0$, which implies that $\Phi'(\eta_j)$ is bounded away from 0. This yields

$$\prod_{j=1}^{h} \rho_h \Phi'(\eta_{h,j}) = \prod_{j=1}^{h} (\rho + O(B_2^h))(\Phi'(\eta_j) + O(B_2^h)) = \rho^h(1 + O(hB_2^h)) \prod_{j=1}^{h} \Phi'(\eta_j).$$

Since $\Phi'(\eta_j) = \Phi'(0) + O(\zeta^j)$ (by Lemma 4.3), the product that defines λ_2 converges. So we can rewrite the product term as

$$\prod_{j=1}^{h} \Phi'(\eta_j) = \Phi'(0)^h \prod_{j=1}^{h} \frac{\Phi'(\eta_j)}{\Phi'(0)} = \lambda_2 \Phi'(0)^h \prod_{j\geq h+1} \frac{\Phi'(0)}{\Phi'(\eta_j)},$$

and thus, using again the estimate $\Phi'(\eta_j) = \Phi'(0) + O(\zeta^j)$ on the remaining product,

$$\prod_{j=1}^{h} \rho_h \Phi'(\eta_{h,j}) = \lambda_2 \zeta^h (1 + O(hB_2^h))(1 + O(\zeta^h)).$$

This proves the desired formula for a suitable choice of B_3. □

Corollary 4.5. *For sufficiently large h, we have that*

$$\rho_h \Phi'(\eta_{h,0}) = 1 + \lambda_2 (1 - \zeta)\zeta^h (1 + O(B_3^h)), \tag{48}$$

where λ_2 and B_3 are as in Lemma 4.4.

Proof. Taking the asymptotic formulas from the statement of Lemma 4.4 and applying them to (47) we obtain the formula after solving for $\rho_h \Phi'(\eta_{h,0})$. □

In the proof of Lemma 4.2 we used the bound $\eta_{h,h} = O(B_1^h)$ (obtained from Lemma 4.1). In order to refine the process, we need a more precise estimate.

Lemma 4.6. *For sufficiently large h and a fixed constant $B_4 < 1$, we have that*

$$\eta_{h,h} = \lambda_1 (1 - \zeta)\zeta^h (1 + O(B_4^h)), \tag{49}$$

where λ_1 is as defined in Lemma 4.3.

Proof. Pick some $\alpha \in (0, 1)$ in such a way that $\zeta^\alpha > B_2$, with B_2 as in Lemma 4.2, and set $m = \lfloor \alpha h \rfloor$. From Lemma 4.3, we know that $\eta_m = \Theta(\zeta^{\alpha h})$. By Lemma 4.2, $\eta_{h,m} = \eta_m + O(B_2^h)$, so by our choice of α there is some $B_4 < 1$ such that $\eta_{h,m} = \eta_m (1 + O(B_4^h))$ for sufficiently large h.

Next, recall from (46) that

$$\eta_{h,k} = \rho_h \big(\Phi(\eta_{h,k-1}) - \Phi(\eta_{h,h})\big).$$

By the mean value theorem, there is some $\xi_{h,k} \in (\eta_{h,h}, \eta_{h,k-1})$ such that

$$\eta_{h,k} = \rho_h (\eta_{h,k-1} - \eta_{h,h}) \Phi'(\xi_{h,k}) = \rho_h (\eta_{h,k-1} - \eta_{h,h})(\Phi'(0) + O(\eta_{h,k-1})).$$

Assume now that $k \geq m$, so that $\eta_{h,k-1} = O(B_1^{\alpha h})$ by Lemma 4.1. Moreover, $\rho_h = \rho + O(B_2^h)$ by Lemma 4.2. So with $B = \max(B_2, B_1^\alpha)$, it follows that

$$\eta_{h,k} = \zeta (\eta_{h,k-1} - \eta_{h,h})(1 + O(B^h)),$$

uniformly for all $k \geq m$. Rewrite this as

$$\eta_{h,k-1} = \eta_{h,h} + \frac{\eta_{h,k}}{\zeta}(1 + O(B^h)).$$

Iterate this $h-m$ times to obtain

$$\eta_{h,m} = \sum_{j=0}^{h-m} \frac{\eta_{h,h}}{\zeta^j}(1+O(B^h))^j$$
$$= \eta_{h,h}\zeta^{-(h-m)}\frac{1-\zeta^{h-m+1}}{1-\zeta}(1+O(hB^h)).$$

Now recall that $\eta_{h,m} = \eta_m(1+O(B_4^h))$, and that $\eta_m = \lambda_1\zeta^m(1+O(B_1^{\alpha h}))$ by Lemma 4.3. Plugging all this in and solving for $\eta_{h,h}$, we obtain (49), provided that B_4 was also chosen to be greater than B and $\zeta^{1-\alpha}$. □

Now we can make use of this asymptotic formula for $\eta_{h,h}$ in order to obtain a refined estimate for $\eta_{h,0}$.

Proposition 4.7. *For a fixed constant $B_5 < 1$ and large enough h, we have that*

$$\eta_{h,0} = \tau + \frac{(1-\zeta)(\Phi(\tau)\lambda_2 - \Phi'(0)\lambda_1)}{\tau\Phi''(\tau)}\zeta^h + O((\zeta B_5)^h) \tag{50}$$

and

$$\rho_h = \rho + \frac{\lambda_1(1-\zeta)}{\Phi(\tau)}\zeta^{h+1} + O((\zeta B_5)^h), \tag{51}$$

where λ_1 and λ_2 are as in Lemmata 4.3 and 4.4 respectively.

Proof. From (45) and (46) with $k=1$, we have

$$\eta_{h,0} = \rho_h\big(\Phi(\eta_{h,0}) - \Phi(\eta_{h,h}) + 1\big). \tag{52}$$

By means of Taylor expansion and Lemma 4.6, we get

$$\eta_{h,0} = \rho_h\big(\Phi(\eta_{h,0}) - \Phi'(0)\eta_{h,h} + O(\eta_{h,h}^2)\big).$$

We multiply this by (48) and divide through by ρ_h to obtain

$$\eta_{h,0}\Phi'(\eta_{h,0}) = \big(\Phi(\eta_{h,0}) - \Phi'(0)\eta_{h,h} + O(\eta_{h,h}^2)\big)\big(1+\lambda_2(1-\zeta)\zeta^h(1+O(B_3^h))\big) \tag{53}$$

or, with $H(x) = x\Phi'(x) - \Phi(x)$,

$$H(\eta_{h,0}) = \big(-\Phi'(0)\eta_{h,h} + O(\eta_{h,h}^2)\big)(1+O(\zeta^h)) + \Phi(\eta_{h,0})\lambda_2(1-\zeta)\zeta^h(1+O(B_3^h)).$$

We plug in the asymptotic formula for $\eta_{h,h}$ from Lemma 4.6 and also note that $\Phi(\eta_{h,0}) = \Phi(\tau + O(B_2^h)) = \Phi(\tau) + O(B_2^h)$ by Lemma 4.2. This gives us

$$H(\eta_{h,0}) = (\Phi(\tau)\lambda_2 - \Phi'(0)\lambda_1)(1-\zeta)\zeta^h + O((\zeta B_5)^h), \tag{54}$$

where $B_5 = \max(\zeta, B_2, B_3, B_4)$. Now note that the function H is increasing (on the positive real numbers within the radius of convergence of Φ) with derivative $H'(x) = x\Phi''(x)$ and a unique zero at τ. So by inverting (54), we finally end up with

$$\eta_{h,0} = \tau + \frac{1}{H'(\tau)}(\Phi(\tau)\lambda_2 - \Phi'(0)\lambda_1)(1-\zeta)\zeta^h + O((\zeta B_5)^h),$$

completing the proof of the first formula. Now we return to (48), which gives us

$$\rho_h = \frac{1+\lambda_2(1-\zeta)\zeta^h(1+O(B_3^h))}{\Phi'(\eta_{h,0})} = \frac{1+\lambda_2(1-\zeta)\zeta^h(1+O(B_3^h))}{\Phi'(\tau) + \Phi''(\tau)(\eta_{h,0}-\tau) + O((\eta_{h,0}-\tau)^2)}.$$

Plugging in (50) and simplifying by means of the identities $\rho\Phi(\tau) = \tau$ and $\rho\Phi'(\tau) = 1$ now yields (51). □

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

4.2 Proof of Theorem 1.2

We are now finally ready to apply Theorems 2.1 and 2.2. The generating functions $Y_h(z) := Y_{h,0}(z) = Y_{h,1}(z) + z$ were defined precisely in such a way that $y_{h,n} = [z^n]Y_h(z)$ is the number of n-vertex trees for which the maximum protection number is less than or equal to h. Thus the random variable X_n in Theorem 2.1 becomes the maximum protection number of a random n-vertex tree. Condition (2) of Theorem 2.1 is satisfied in view of Proposition 3.16. Condition (1) holds by Proposition 4.7 with $\zeta = \rho\Phi'(0)$ and

$$\kappa = \frac{\lambda_1(1-\zeta)\zeta}{\rho\Phi(\tau)} = \frac{\lambda_1(1-\zeta)\zeta}{\tau}, \tag{55}$$

where λ_1 is as defined in Lemma 4.3 and we recall the definition of ζ as $\rho\Phi'(0)$. This already proves the first part of Theorem 1.2.

We can also apply Theorem 2.2: Note that the maximum protection number of a tree with size n is no greater than $n-1$, thus $y_{h,n} = y_n$ for $h \geq n-1$, and an appropriate choice of constant for Condition (1) in Theorem 2.2 would be $K = 1$. Conditions (2) and (3) are still covered by Proposition 3.16. Hence Theorem 2.2 applies, and the second part of Theorem 1.2 follows.

5. The double-exponential case: $w_1 = 0$

5.1 Asymptotics of the singularities

In Section 4.1, it was crucial in most of our asymptotic estimates that $w_1 = \Phi'(0) \neq 0$. In this section we assume that $w_1 = \Phi'(0) = 0$ and define r to be the smallest positive outdegree with nonzero weight:

$$r = \min\{i \in \mathbb{N} : i \geq 2 \text{ and } w_i \neq 0\} = \min\{i \in \mathbb{N} : i \geq 2 \text{ and } \Phi^{(i)}(0) \neq 0\}.$$

Our goal will be to determine the asymptotic behaviour of ρ_h in this case, based again on the system of equations that is given by (45), (46) and (47). Once again, B_i's will always denote positive constants with $B_i < 1$ (different from those in the previous section, but for simplicity we restart the count at B_1) that depend on the specific family of simply generated trees, but nothing else.

No part of the proof of Lemma 4.1 depends on $\Phi'(0) \neq 0$ and thus it also holds in the case which we are currently working in, so we already have an exponential bound on $\eta_{h,k}$. However, this bound is loose if $\Phi'(0) = 0$, and so we determine a tighter bound.

Lemma 5.1. *There exist positive constants C and B_1 with $B_1 < 1$ such that $\eta_{h,k} \leq CB_1^{r^k}$ for all sufficiently large h and all k with $0 \leq k \leq h$.*

Proof. From (46), we have that $\eta_{h,k} = \rho_h\Phi(\eta_{h,k-1}) - \rho_h\Phi(\eta_{h,h})$. Using the Taylor expansion about 0, this gives, for some $\xi_{h,k-1} \in (0, \eta_{h,k-1})$,

$$\begin{aligned}\eta_{h,k} &= \rho_h\left(\Phi(0) + \frac{\Phi^{(r)}(\xi_{h,k-1})}{r!}\eta_{h,k-1}^r\right) - \rho_h\Phi(\eta_{h,h}) \\ &\leq \rho_h\frac{\Phi^{(r)}(\xi_{h,k-1})}{r!}\eta_{h,k-1}^r \leq \rho_h\frac{\Phi^{(r)}(\eta_{h,1})}{r!}\eta_{h,k-1}^r.\end{aligned}$$

There is a constant M such that $\rho_h\frac{\Phi^{(r)}(\eta_{h,1})}{r!} \leq M$ for all sufficiently large h, since we already know that ρ_h and $\eta_{h,1}$ converge. So for sufficiently large h, we have $\eta_{h,k} \leq M\eta_{h,k-1}^r$ for all $k > 1$. Iterating this inequality yields

$$\eta_{h,k} \leq M^{\frac{r^{k-\ell}-1}{r-1}}\eta_{h,\ell}^{r^{k-\ell}}$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

for $0 \le \ell \le k$. In view of the exponential bound on $\eta_{h,\ell}$ provided by Lemma 4.1, we can choose ℓ so large that $M^{1/(r-1)}\eta_{h,\ell} \le \frac{1}{2}$ for all sufficiently large h. This proves the desired bound for $k \ge \ell$ with $B_1 = 2^{-r^{-\ell}}$ and a suitable choice of C (for $k < \ell$, it is implied by the exponential bound). □

Our next step is an analogue of Lemma 4.4.

Lemma 5.2. *For large enough h and the same constant $B_1 < 1$ as in the previous lemma, we have*

$$1 + \sum_{k=2}^{h} \prod_{j=k}^{h} (\rho_h \Phi'(\eta_{h,j})) = 1 + O(B_1^{r^h})$$

and

$$\prod_{j=1}^{h} \rho_h \Phi'(\eta_{h,j}) = O(B_1^{r^h}).$$

Proof. We already know that $\rho_h \Phi'(\eta_{h,1})$ converges to $\rho\Phi'(\tau - \rho) < 1$, so for sufficiently large h and some $q < 1$, we have $\rho_h \Phi'(\eta_{h,j}) \le \rho_h \Phi'(\eta_{h,1}) \le q$ for all $j \ge 1$. It follows that

$$\sum_{k=2}^{h} \prod_{j=k}^{h} (\rho_h \Phi'(\eta_{h,j})) \le \sum_{k=2}^{h} q^{h-k} \rho_h \Phi'(\eta_{h,h}) \le \frac{1}{1-q} \rho_h \Phi'(\eta_{h,h})$$

and

$$\prod_{j=1}^{h} \rho_h \Phi'(\eta_{h,j}) \le q^{h-1} \rho_h \Phi'(\eta_{h,h}).$$

Now both statements follow from the fact that $\Phi'(\eta_{h,h}) = \Phi'(0) + O(\eta_{h,h}) = O(\eta_{h,h})$ and the previous lemma. □

Taking the results from Lemma 5.2 and applying them to (47), we find that

$$\rho_h \Phi'(\eta_{h,0}) = 1 + O\big(B_1^{r^h}\big). \tag{56}$$

Additionally note that using Lemma 5.1 and Taylor expansion, we have that

$$\Phi(\eta_{h,h}) = 1 + O(B_1^{r^h}).$$

Now recall that (45) and (46) yield (see (52))

$$\eta_{h,0} = \rho_h \big(\Phi(\eta_{h,0}) - \Phi(\eta_{h,h}) + 1\big), \tag{57}$$

which now becomes

$$\eta_{h,0} = \rho_h \Phi(\eta_{h,0}) + O\big(B_1^{r^h}\big). \tag{58}$$

Taking advantage of the expressions in (56) and (58), we can now prove doubly exponential convergence of ρ_h and $\eta_{h,0}$ (using the approach of Proposition 4.7).

Lemma 5.3. *For large enough h, it holds that*

$$\rho_h = \rho + O\big(B_1^{r^h}\big) \qquad \text{and} \qquad \eta_{h,0} = \tau + O\big(B_1^{r^h}\big).$$

and thus also $\eta_{h,1} = \eta_{h,0} - \rho_h = \eta_1 + O(B_1^{r^h})$.

Proof. Multiplying (56) and (58) and dividing by ρ_h yields

$$\eta_{h,0} \Phi'(\eta_{h,0}) = \Phi(\eta_{h,0}) + O\big(B_1^{r^h}\big).$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

As in the proof of Proposition 4.7, we observe that the function $H(x) = x\Phi'(x) - \Phi(x)$ is increasing (on the positive real numbers within the radius of convergence of Φ) with derivative $H'(x) = x\Phi''(x)$ and a unique zero at τ. So it follows from this equation that $\eta_{h,0} = \tau + O\big(B_1^{r^h}\big)$. Using this estimate for $\eta_{h,0}$ in (56) it follows that $\rho_h = \rho + O\big(B_1^{r^h}\big)$. □

As in the previous section, we will approximate $\eta_{h,k}$ by η_k, defined recursively by $\eta_0 = \tau$ and $\eta_k = \rho(\Phi(\eta_{k-1}) - 1)$. As it turns out, this approximation is even more precise in the current case.

Lemma 5.4. *For a fixed constant $B_2 < 1$ and sufficiently large h, we have that*

$$\eta_{h,k} = \eta_k(1 + O(B_2^{r^h})),$$

uniformly for all $0 \le k \le h$.

Proof. Recall that, by (46), $\eta_{h,k} = \rho_h\Phi(\eta_{h,k-1}) - \rho_h\Phi(\eta_{h,h})$. By Taylor expansion, we find that

$$\eta_{h,k} = \rho_h\Phi(\eta_{h,k-1}) - \rho_h + O(\eta_{h,h}^r).$$

Since $\eta_{h,k} \ge \eta_{h,h}$, we have $\eta_{h,k} - O(\eta_{h,h}^r) = \eta_{h,k}(1 - O(\eta_{h,h}^{r-1}))$. Now we use the estimates $\eta_{h,h} = O(B_1^{r^h})$ from Lemma 5.1 and $\rho_h = \rho + O(B_1^{r^h})$ from Lemma 5.3 to obtain

$$\eta_{h,k} = \rho(\Phi(\eta_{h,k-1}) - 1)\big(1 + O(B_1^{r^h})\big).$$

We compare this to

$$\eta_k = \rho(\Phi(\eta_{k-1}) - 1).$$

Taking the logarithm in both these equations and subtracting yields

$$\log\frac{\eta_{h,k}}{\eta_k} = \log(\Phi(\eta_{h,k-1}) - 1) - \log\left(\Phi(\eta_{k-1}) - 1\right) + O(B_1^{r^h}). \tag{59}$$

For large enough h, we can assume that $\eta_{h,1} \le \tau$ and thus $\eta_{h,k} \le \tau$ for all $k \ge 1$. The auxiliary function

$$\Psi_1(u) = \log\big(\Phi(e^u) - 1\big)$$

is continuously differentiable on $(-\infty, \log(\tau)]$. Since $\lim_{u\to-\infty}\Psi_1'(u) = r$, as one easily verifies, $|\Psi_1'(u)|$ must be bounded by some constant K for all u in this interval, thus $|\Psi_1(u+v) - \Psi_1(u)| \le K|v|$ whenever $u, u+v \le \log(\tau)$. We apply this with $u + v = \log\eta_{h,k-1}$ and $u = \log\eta_{k-1}$ to obtain

$$\Big|\log\left(\Phi(\eta_{h,k-1}) - 1\right) - \log\left(\Phi(\eta_{k-1}) - 1\right)\Big| \le K\Big|\log\frac{\eta_{h,k-1}}{\eta_{k-1}}\Big|.$$

Plugging this into (59) yields

$$\Big|\log\frac{\eta_{h,k}}{\eta_k}\Big| \le K\Big|\log\frac{\eta_{h,k-1}}{\eta_{k-1}}\Big| + O(B_1^{r^h}). \tag{60}$$

We already know that $\big|\log\frac{\eta_{h,0}}{\eta_0}\big| = O(B_1^{r^h})$ and $\big|\log\frac{\eta_{h,1}}{\eta_1}\big| = O(B_1^{r^h})$ in view of Lemma 5.3. Iterating (60) gives us

$$\Big|\log\frac{\eta_{h,k}}{\eta_k}\Big| = O\big((1 + K + K^2 + \cdots + K^k)B_1^{r^h}\big),$$

which implies the statement for any $B_2 > B_1$. □

The next lemma parallels Lemma 4.3.

Lemma 5.5. *There exist positive constants λ_1 and $\mu < 1$ such that*

$$\eta_k = \lambda_1\mu^{r^k}\big(1 + O(B_1^{r^k})\big),$$

with the same constant B_1 as in Lemma 5.1.

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Proof. Note that Lemma 5.1 trivially implies that $\eta_k = O(B_1^{r^k})$. From the recursion

$$\eta_k = \rho(\Phi(\eta_{k-1}) - 1),$$

we obtain, by the properties of Φ,

$$\eta_k = \frac{\rho\Phi^{(r)}(0)\eta_{k-1}^r}{r!}(1 + O(\eta_{k-1})).$$

Set

$$\lambda_1 = \left(\frac{\rho\Phi^{(r)}(0)}{r!}\right)^{-1/(r-1)} = (\rho w_r)^{-1/(r-1)} \tag{61}$$

and divide both sides by λ_1 to obtain

$$\frac{\eta_k}{\lambda_1} = \left(\frac{\eta_{k-1}}{\lambda_1}\right)^r (1 + O(\eta_{k-1})).$$

Let us write $e^{\theta_{k-1}}$ for the final factor, where $\theta_{k-1} = O(\eta_{k-1})$. Taking the logarithm yields

$$\log\frac{\eta_k}{\lambda_1} = r\log\frac{\eta_{k-1}}{\lambda_1} + \theta_{k-1}.$$

We iterate this recursion k times to obtain

$$\begin{aligned}\log\frac{\eta_k}{\lambda_1} &= r^k\log\frac{\eta_0}{\lambda_1} + \sum_{j=0}^{k-1} r^{k-1-j}\theta_j \\ &= r^k\left(\log\frac{\eta_0}{\lambda_1} + \sum_{j=0}^{\infty} r^{-1-j}\theta_j\right) - \sum_{j=k}^{\infty} r^{k-1-j}\theta_j.\end{aligned}$$

The infinite series converge in view of the estimate $\theta_j = O(\eta_j) = O(B_1^{r^j})$ that we get from Lemma 5.1. Moreover, we have $\sum_{j=k}^{\infty} r^{k-1-j}\theta_j = O(B_1^{r^k})$ by the same bound. The result follows upon taking the exponential on both sides and multiplying by λ_1, setting

$$\mu := \exp\left(\log\frac{\eta_0}{\lambda_1} + \sum_{j=0}^{\infty} r^{-1-j}\theta_j\right) = \frac{\eta_0}{\lambda_1}\prod_{j=0}^{\infty} e^{\theta_j/r^{j+1}}. \tag{62}$$

Note that $\mu < 1$ because we already know that $\eta_k = O(B_1^{r^k})$. □

In order to further analyse the behaviour of the product $\prod_{j=1}^{h}(\rho_h\Phi'(\eta_{h,j}))$ in (47), we need one more short lemma.

Lemma 5.6. *For sufficiently large h, we have that*

$$\Phi'(\eta_{h,k}) = \Phi'(\eta_k)\big(1 + O(B_2^{r^h})\big),$$

uniformly for all $1 \le k \le h$, with the same constant B_2 as in Lemma 5.4.

Proof. Again, we can assume that h is so large that $\eta_{h,k} \le \tau$ for all $k \ge 1$. The auxiliary function $\Psi_2(u) = \log(\Phi'(e^u))$ is continuously differentiable on $(-\infty, \log\tau]$ and satisfies $\lim_{u\to-\infty}\Psi_2'(u) = r - 1$. Thus its derivative is also bounded, and the same argument as in Lemma 5.4 shows that

$$\left|\log\frac{\Phi'(\eta_{h,k})}{\Phi'(\eta_k)}\right| \le K\left|\log\frac{\eta_{h,k}}{\eta_k}\right|$$

for some positive constant K. Now the statement follows from Lemma 5.4. □

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

Lemma 5.7. *There exist positive constants λ_2, λ_3 and $B_3 < 1$ such that, for large enough h,*

$$\prod_{j=1}^{h}(\rho_h\Phi'(\eta_{h,j})) = \lambda_2\lambda_3^h\mu^{r^{h+1}}\big(1+O(B_3^{r^h})\big), \tag{63}$$

with μ as in Lemma 5.5.

Proof. First, observe that

$$\prod_{j=1}^{h}(\rho_h\Phi'(\eta_{h,j})) = \Big(\rho\big(1+O(B_1^{r^h})\big)\Big)^h\prod_{j=1}^{h}\Big(\Phi'(\eta_j)\big(1+O(B_2^{r^h})\big)\Big) = \rho^h\Big(\prod_{j=1}^{h}\Phi'(\eta_j)\Big)\big(1+O(hB_2^{r^h})\big)$$

in view of Lemmata 5.3 and 5.6 (recall that $B_2 > B_1$). Next, Taylor expansion combined with Lemma 5.5 gives us

$$\Phi'(\eta_k) = \frac{\Phi^{(r)}(0)}{(r-1)!}\eta_k^{r-1}(1+O(\eta_k)) = \frac{\Phi^{(r)}(0)\lambda_1^{r-1}}{(r-1)!}\mu^{(r-1)r^k}\big(1+O(B_1^{r^k})\big).$$

Set $\lambda_3 := \rho\frac{\Phi^{(r)}(0)\lambda_1^{r-1}}{(r-1)!} = rw_r\rho\lambda_1^{r-1}$, so that

$$\Phi'(\eta_k) = \frac{\lambda_3}{\rho}\mu^{(r-1)r^k}\big(1+O(B_1^{r^k})\big).$$

It follows that the infinite product

$$\Pi := \prod_{j=1}^{\infty}\frac{\rho\Phi'(\eta_j)}{\lambda_3\mu^{(r-1)r^j}}$$

converges, and that

$$\prod_{j=1}^{h}\frac{\rho\Phi'(\eta_j)}{\lambda_3\mu^{(r-1)r^j}} = \Pi\big(1+O(B_1^{r^h})\big).$$

Consequently,

$$\prod_{j=1}^{h}\Phi'(\eta_j) = \Pi\Big(\frac{\lambda_3}{\rho}\Big)^h\mu^{r^{h+1}-r}\big(1+O(B_1^{r^h})\big).$$

Putting everything together, the statement of the lemma follows with $\lambda_2 = \Pi\mu^{-r}$ and a suitable choice of $B_3 > B_2$. □

With this estimate for the product term in the determinant of the Jacobian (47), and the estimate for the sum term from Lemma 5.2, we can now obtain a better asymptotic formula for $\rho_h\Phi'(\eta_{h,0})$ than that which was obtained in (56). For large enough h, we have that

$$\rho_h\Phi'(\eta_{h,0}) = 1+\lambda_2\lambda_3^h\mu^{r^{h+1}}\big(1+O(B_3^{r^h})\big). \tag{64}$$

For the error term, recall that $B_1 < B_3$. Moreover, combining Lemmata 5.4 and 5.5 leads to

$$\eta_{h,h} = \lambda_1\mu^{r^h}\big(1+O(B_2^{r^h})\big)$$

https://doi.org/10.1017/S0963548324000099 Published online by Cambridge University Press

since B_2 was chosen to be greater than B_1, which we can apply to (57):

$$\begin{aligned}\eta_{h,0} &= \rho_h\big(\Phi(\eta_{h,0}) - \Phi(\eta_{h,h}) + 1\big)\\ &= \rho_h\Big(\Phi(\eta_{h,0}) - \frac{\Phi^{(r)}(0)}{r!}\eta_{h,h}^r + O(\eta_{h,h}^{r+1})\Big)\\ &= \rho_h\Big(\Phi(\eta_{h,0}) - w_r\lambda_1^r\mu^{r^{h+1}}\big(1+O(B_3^{r^h})\big)\Big) \end{aligned} \tag{65}$$

since B_3 was chosen to be greater than B_1 (and thus also μ) and B_2. As we did earlier to obtain Lemma 5.3, we multiply the two equations (64) and (65) and divide by ρ_h to find that

$$\eta_{h,0}\Phi'(\eta_{h,0}) = \Big(\Phi(\eta_{h,0}) - w_r\lambda_1^r\mu^{r^{h+1}}\big(1+O(B_3^{r^h})\big)\Big)\Big(1+\lambda_2\lambda_3^h\mu^{r^{h+1}}\big(1+O(B_3^{r^h})\big)\Big).$$

From this, the following result follows now in exactly the same way as Proposition 4.7 follows from (53).

Proposition 5.8. *For large enough h and a fixed constant $B_4 < 1$, we have that*

$$\eta_{h,0} = \tau + \frac{\Phi(\tau)\lambda_2\lambda_3^h - w_r\lambda_1^r}{\tau\Phi''(\tau)}\mu^{r^{h+1}} + O(\mu^{r^{h+1}}B_4^{r^h})$$

and

$$\rho_h = \rho\Big(1 + \frac{w_r\lambda_1^r}{\Phi(\tau)}\mu^{r^{h+1}} + O(\mu^{r^{h+1}}B_4^{r^h})\Big).$$

5.2 An adapted general scheme and the proof of Theorem 1.3

In this final section, we will first prove Theorems 1.4 and 1.5. Then, we will be able to put all pieces together and prove Theorem 1.3.

Proof of Theorem 1.4. We apply singularity analysis, and use the uniformity condition to obtain

$$y_{h,n} = \frac{A_h}{\Gamma(-\alpha)}n^{-\alpha-1}\rho_h^{-n}(1+o(1))$$

uniformly in h as $n\to\infty$ as well as

$$y_n = \frac{A}{\Gamma(-\alpha)}n^{-\alpha-1}\rho^{-n}(1+o(1)).$$

Since in addition $A_h \to A$ and $\rho_h = \rho(1+\kappa\zeta^{r^h} + o(\zeta^{r^h}))$, it holds that

$$\begin{aligned}\frac{y_{h,n}}{y_n} &= \Big(\frac{\rho_h}{\rho}\Big)^{-n}(1+o(1)) = \exp\big(-\kappa n\zeta^{r^h} + o(n\zeta^{r^h})\big)(1+o(1))\\ &= \exp\big(-\kappa n\zeta^{r^h}(1+o(1)) + o(1)\big).\end{aligned}$$

□

Proof of Theorem 1.5. Fix $\epsilon > 0$. If $h \ge m_n + \epsilon = \log_r\log_d(n) + \epsilon$, then Theorem 1.4 gives us

$$\mathbb{P}(X_n \le h) \ge \exp\big(-\kappa n^{1-r^{\epsilon}}(1+o(1)) + o(1)\big) = 1 - o(1),$$

thus $X_n \le h$ with high probability. If $\{m_n\} \le 1-\epsilon$, then this is the case for $h = \lceil m_n\rceil$, otherwise for $h = \lceil m_n\rceil + 1$. Similarly, if $h \le m_n - \epsilon = \log_r\log_d(n) - \epsilon$, then Theorem 1.4 gives us

$$\mathbb{P}(X_n \le h) \le \exp\big(-\kappa n^{1-r^{-\epsilon}}(1+o(1)) + o(1)\big) = o(1),$$

thus $X_n > h$ with high probability. If $\{m_n\} \ge \epsilon$, then this is the case for $h = \lfloor m_n\rfloor$, otherwise for $h = \lfloor m_n\rfloor - 1$. The statement now follows by combining the two parts. □

Remark 5.9. *As in Remark 3.17, we indicate the changes which are necessary for the case that the period D of Φ is greater than* 1.

Theorem 1.4 only depends on singularity analysis. It is well known (see [10, Remark VI.17]) that singularity analysis simply introduces a factor D in this situation, and as this factor D cancels because it occurs both in the asymptotic expansions of Y_h as well as Y, this theorem remains valid for $n \equiv 1$ (mod D).

Theorem 1.3 is now an immediate consequence of Theorems 1.4 and 1.5. In analogy to the proof of Theorem 1.2, the analytic conditions on the generating functions are provided by Proposition 3.16. The condition on the asymptotic behaviour of ρ_h is given by Proposition 5.8 (with $\zeta = \mu^r$). Thus the proof of Theorem 1.3 is complete.

Financial support

The research of C. Heuberger and S. J. Selkirk was funded in part by the Austrian Science Fund (FWF) [10.55776/P28466], Analytic Combinatorics: Digits, Automata and Trees and Austrian Science Fund (FWF) [10.55776/DOC78]. S. Wagner was supported by the Knut and Alice Wallenberg Foundation, grant KAW 2017.0112, and the Swedish research council (VR), grant 2022-04030. For open access purposes, the authors have applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

References

[1] Aizenberg, I. A and Yuzhakov, A. P. (1983) *Integral Representations and Residues in Multidimensional Complex Analysis*, Vol. 58 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI.
[2] Bóna, M. (2014) k-protected vertices in binary search trees. *Adv. Appl. Math.* **53** 1–11.
[3] Bóna, M. and Pittel, B. (2017) On a random search tree: asymptotic enumeration of vertices by distance from leaves. *Adv. Appl. Prob.* **49**(3) 850–876.
[4] Cheon, G.-S. and Shapiro, L. W. (2008) Protected points in ordered trees. *Appl. Math. Lett.* **21**(5) 516–520.
[5] Copenhaver, K. (2017) k-protected vertices in unlabeled rooted plane trees. *Graphs Comb.* **33**(2) 347–355.
[6] Devroye, L., Goh, M. K. and Zhao, R. Y. (2023) On the peel number and the leaf-height of Galton-Watson trees. *Comb. Probab. Comput.* **32**(1) 68–90.
[7] Devroye, L. and Janson, S. (2014) Protected nodes and fringe subtrees in some random trees. *Electron. Commun. Prob.* **19**(6) 10.
[8] Drmota, M. (2009) *Random Trees*. SpringerWienNewYork.
[9] Du, R. R. X. and Prodinger, H. (2012) Notes on protected nodes in digital search trees. *Appl. Math. Lett.* **25**(6) 1025–1028.
[10] Flajolet, P. and Sedgewick, R. (2009) *Analytic Combinatorics*. Cambridge University Press, Cambridge.
[11] Gaither, J., Homma, Y., Sellke, M. and Ward, M. D. On the number of 2-protected nodes in tries and suffix trees. In *23rd International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'12)* (N. Broutin and L. Devroye, eds), Montreal, Canada. Discrete Mathematics and Theoretical Computer Science, pp. 381–398.
[12] Gittenberger, B., Gołębiewski, Z., Larcher, I. and Sulkowska, M. (2023) Protection numbers in simply generated trees and Pólya trees. *Appl. Anal. Discrete Math.* **17** 1–24.
[13] Gołębiewski, Z. and Klimczak, M. (2019) Protection number of recursive trees. In *Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, Philadelphia PA. SIAM, pp. 45-53.
[14] Heuberger, C. and Prodinger, H. (2017) Protection number in plane trees. *Appl. Anal. Discrete Math.* **11** 314–326.
[15] Holmgren, C. and Janson, S. (2015) Asymptotic distribution of two-protected nodes in ternary search trees. *Electron. J. Prob.* **20** 1–20.
[16] Holmgren, C. and Janson, S. (2015) Limit laws for functions of fringe trees for binary search trees and random recursive trees. *Electron. J. Prob.* **20**(4) 51.
[17] Janson, S. (2012) Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.* **9** 103–252.
[18] Kim, H. and Stanley, R. P. (2016) A refined enumeration of hex trees and related polynomials. *Eur. J. Comb.* **54** 207–219.
[19] Mahmoud, H. M. and Ward, M. D. (2015) Asymptotic properties of protected nodes in random recursive trees. *J. Appl. Prob.* **52**(1) 290–297.
[20] Mahmoud, H. M. and Ward, M. D. (2012) Asymptotic distribution of two-protected nodes in random binary search trees. *Appl. Math. Lett.* **25**(12) 2218–2222.

[21] Mansour, T. (2011) Protected points in k-ary trees. *Appl. Math. Lett.* **24**(4) 478–480.
[22] Meir, A. and Moon, J. W. (1978) On the altitude of nodes in random trees. *Can. J. Math.* **30**(5) 997–1015.
[23] Prodinger, H. and Wagner, S. (2015) Bootstrapping and double-exponential limit laws. *Discrete Math. Theor. Comput. Sci.* **17**(1) 123–144.

Cite this article: Heuberger C, Selkirk SJ, and Wagner S (2024). The distribution of the maximum protection number in simply generated trees. *Combinatorics, Probability and Computing* **33**, 518–553. https://doi.org/10.1017/S0963548324000099