*Discrete Mathematics and Theoretical Computer Science* DMTCS vol. **7**, 2005, 313–400

# *Enumeration of Binary Trees and Universal Types*

Charles Knessl[1†] and Wojciech Szpankowski[2‡]

[1]*Dept. Mathematics, Statistics & Computer Science, University of Illinois at Chicago, Chicago, Illinois 60607-7045, USA. Email:* `knessl@uic.edu`
[2]*Department of Computer Science, Purdue University, West Lafayette, IN, USA. Email:* `spa@cs.purdue.edu`

*received Aug 6, 2004*, *accepted Dec 22, 2005*.

Binary unlabeled ordered trees (further called binary trees) were studied at least since Euler, who enumerated them. The number of such trees with $n$ nodes is now known as the Catalan number. Over the years various interesting questions about the statistics of such trees were investigated (e.g., height and path length distributions for a randomly selected tree). Binary trees find an abundance of applications in computer science. However, recently Seroussi posed a new and interesting problem motivated by information theory considerations: how many binary trees of a *given path length* (sum of depths) are there? This question arose in the study of *universal types* of sequences. Two sequences of length $p$ have the same universal type if they generate the same set of phrases in the incremental parsing of the Lempel-Ziv'78 scheme since one proves that such sequences converge to the same empirical distribution. It turns out that the number of distinct types of sequences of length $p$ corresponds to the number of binary (unlabeled and ordered) trees, $T_p$, of given path length $p$ (and also the number of distinct Lempel-Ziv'78 parsings of length $p$ sequences). We first show that the number of binary trees with given path length $p$ is asymptotically equal to $T_p \sim 2^{2p/(\log_2 p)(1+O(\log^{-2/3} p))}$. Then we establish various limiting distributions for the number of nodes (number of phrases in the Lempel-Ziv'78 scheme) when a tree is selected randomly among all trees of given path length $p$. Throughout, we use methods of analytic algorithmics such as generating functions and complex asymptotics, as well as methods of applied mathematics such as the WKB method and matched asymptotics.

**Keywords:** Binary trees, types, Lempel-Ziv'78, path length

†The work was supported by NSF Grant DMS 05-03745 and NSA Grant MDA 904-03-1-0036.
‡The work was supported by NSF Grants CCR-0208709 and DMS 05-03745, NIH Grant R01 GM068959-01, and NSA Grant MDA 904-03-1-0036.

# Contents

**1 Introduction** **314**

**2 Summary of Results** **317**

**3 Far Right Region** **327**

**4 Right Region** **330**

**5 Central Region** **337**
5.1 Moment equations . . . 338
5.2 Analysis of the basic recurrence . . . 353
5.3 Transform inversion . . . 357

**6 Left Region** **360**

**7 Far Left Region** **365**

**8 The Matching Region Between the Left and Far Left Scales** **371**

**9 Numerical Studies** **379**

**A APPENDIX** **396**

# 1 Introduction

Trees are the most important nonlinear structures that arise in computer science. Applications are in abundance (cf. [15, 17]); in this paper we discuss a novel application of binary unlabeled ordered trees (further called binary trees) in information theory (e.g., counting Lempel-Ziv'78 parsings and universal types). Tree structures have been the object of extensive mathematical investigations for many years, and many interesting facts have been discovered. Enumeration of binary trees, which are of principal importance to computer science, has been known already by Euler. Nowadays, the number of such trees built on $n$ nodes is called the Catalan number.

Since Euler and Cayley, various interesting questions concerning statistics of randomly generated binary trees were investigated (cf. [9, 15, 17, 24, 26, 27]). In the standard model, one selects uniformly a tree among all binary unlabeled ordered trees built on $n$ nodes, $\mathcal{T}_n^*$ (where $|\mathcal{T}_n^*| = \binom{2n}{n}\frac{1}{n+1}$ =Catalan number). For example, Flajolet and Odlyzko [6] and Takacs [26] established the average and the limiting distribution for the height (longest path), while Louchard [18, 19] and Takacs [25, 26, 27] derive the limiting distribution for the path length (sum of all paths from the root to all nodes). As we indicate below, these limiting distributions are expressible in terms of the Airy's function (cf. [1, 2]).

While deep and interesting results concerning the behavior of binary trees in the standard model were uncovered, there are still many important unsolved problems of practical importance. Recently, Seroussi [22], when studying universal types for sequences and distinct parsings of the Lempel-Ziv scheme, asked for the enumeration of binary trees with a *given path length*. Let $\mathcal{T}_p$ be the set of binary trees of given path length $p$. Seroussi observed that the cardinality of $\mathcal{T}_p$ corresponds to the number of possible parsings of sequences of length $p$ in the Lempel-Ziv'78 scheme, and the number of universal types (that we discuss below). We shall first enumerate $\mathcal{T}_p$ (cf. also Seroussi [23]), and then compute the limiting distribution of the number of nodes (phrases in the LZ'78 scheme) when a tree is selected uniformly among $\mathcal{T}_p$. To the best of our knowledge these problems were never addressed before, with the exception of [22]. We show below that they are much harder than the corresponding problems in the more standard $\mathcal{T}_n^*$ model.

As mentioned above, the problem of enumerating binary trees of a given path length arose in Seroussi's research on universal types. The method of types [4] is a powerful technique in information theory, large deviations, and analysis of algorithms. It reduces calculations of the probability of rare events to a combinatorial analysis. Two sequences (over a finite alphabet) are of the same *type* if they have the same empirical distribution. For memoryless sources, the type is measured by the relative frequency of symbol occurrences, while for Markov sources one needs to count the number of pairs of symbols. It turns out (cf. [12]) that the number of sequences of a given Markovian type can be counted by enumerating Eulerian paths in a multigraph. Recently, Seroussi [22] introduced *universal types* (for individual sequences and/or for sequences generated by a stationary and ergodic source). Two sequences of the same length $p$ are said to be of the same universal type if they generate the same set of phrases in the incremental parsing of the Lempel-Ziv'78 scheme. It is proved that such sequences have the same asymptotic empirical distribution. But, every set of phrases defines uniquely a binary tree of path length $p$ [11, 22] (with the number of phrases corresponding to the number of nodes in the $\mathcal{T}_p$ model). For example, strings $10101100$ and $01001011$ have the same set of phrases $\{1, 0, 10, 11, 00\}$ and therefore the corresponding binary trees are the same. Thus, enumeration of $\mathcal{T}_p$ leads to counting universal types and different LZ'78 parsings of sequences of length $p$.

Let us now summarize our main results. It is easy to see that the generating function $B(z, w) = \sum_{n,p\geq 0} b(n, p)z^n w^p$ of the number $b(n, p)$ of binary trees with $n$ nodes and path length $p$ satisfies the following functional equation [15]

$$B(z, w) = 1 + zB^2(zw, w). \tag{1.1}$$

Observe that this equation is asymmetric with respect to $z$ and $w$. When enumerating trees in $\mathcal{T}_n^*$, we set $w = 1$ to get the well known algebraic equation $B(z, 1) = 1 + zB^2(z, 1)$ that can be explicitly solved as $B(z, 1) = \left(1 - \sqrt{1 - 4z}\right)/(2z)$ leading to the Catalan number. A randomly (uniformly) selected tree from $\mathcal{T}_n^*$ has path length $L_n$ that is asymptotically distributed as Airy's distribution [25, 26], that is,

$$\Pr\{L_n/\sqrt{2n^3} \leq x\} \to W(x)$$

where $W(x)$ is the Airy distribution function defined by its moments [7]. The Airy distribution arises in surprisingly many contexts, such as parking allocations, hashing tables, trees, discrete random walks, area under Brownian bridge, etc. [7, 8, 16, 18, 19, 25, 26, 27].

Setting $z = 1$ in (1.1) we arrive at

$$B(1, w) = 1 + B^2(w, w)$$

which is not *algebraically* solvable. Observe that the coefficient of $B(1,w)$ at $w^p$ of $\sum_{n=0}^{\infty} b(n,p)$ enumerates binary trees with *given* path length $p$. We denote the set of all such binary trees as $\mathcal{T}_p$ and study in this paper its combinatorial and statistical properties. We shall show (and also give explicitly the error term, which involves the Airy function) that

$$T_p := |\mathcal{T}_p| = \frac{1}{(\log_2 p)\sqrt{\pi p}} 2^{\frac{2p}{\log_2 p}\left(1+c_1 \log^{-2/3} p + c_2 \log^{-1} p + O(\log^{-4/3} p)\right)}$$

for large $p$, where $c_1$ and $c_2$ are explicitly computable constants. Seroussi first conjectured the form of the leading term in the exponent of the above asymptotic result, proved an upper bound of that form, and has recently obtained [23] a proof for the matching lower bound, using information-theoretic and combinatorial arguments for $t$-ary trees.

In this paper we further analyze the random variable $N_p$ representing the number of nodes in a randomly selected tree from the assembly $\mathcal{T}_p$. We show that $(N_p - \mathbf{E}[N_p])/\sqrt{\mathbf{Var}\,[N_p]}$ is asymptotically normal. Finally, after deriving various asymptotic expansions for $B_n(w) = \sum_{p\geq 0} w^p b(n,p)$, we analyze the number of trees $b(n,p)$ with $n$ nodes and path length $p$ for various ranges of $n$ and $p$. We also obtain $b(n,p)$ in the asymptotic matching regions for various scales. In passing, we point out that $T_p = |\mathcal{T}_p|$ corresponds to the number of distinct universal types in Seroussi's sense and the number of distinct parsings of binary sequences of length $p$, while $b(n,p)$ enumerates the number of Lempel-Ziv'78 parsings with $n$ phrases. Observe that then $N_p$ represents the number of phrases in LZ'78 parsing of a sequence of length $p$ in the $\mathcal{T}_p$ model.

The functional equation (1.1) falls into the class of quicksort-like nonlinear functional equations (cf. [10, 8, 17, 13, 20, 21]) that is still not fully analyzed (with some exceptions like the linear probing algorithm [8, 16]). Nonlinear functional equations of type (1.1) are not particularly suitable for analytic tools which work fine for linear functional equations (cf. [9, 24]). Therefore, we turn to methods of applied mathematics such as matched asymptotics and the WKB method [3]. These make certain assumptions about the forms of some asymptotic expansions and their asymptotic matching. When stating our main results (see Result 3 in section 2), we discuss the assumptions in more detail. The methods we use are analytic methods that are especially suitable for problems that cannot be solved exactly by transform methods (cf. [13, 14]).

The WKB method [3, 24] was named after the physicists Wentzel, Kramers and Brillouin. It *assumes* that the solution, $B(\xi;n)$, to a recurrence, functional equation or differential equation has the following asymptotic form

$$B(\xi;n) \sim e^{n\phi(\xi)} \left[ A(\xi) + \frac{1}{n} A^{(1)}(\xi) + \frac{1}{n^2} A^{(2)}(\xi) + \cdots \right], \quad n \to \infty$$

where $\phi(\xi)$ and $A(\xi), A^{(1)}(\xi), \ldots$ are unknown functions. These functions must be determined from the equation itself, often in conjunction with another tool known as the *asymptotic matching* principle (cf. [3]).

The outline for the paper is as follows: In Section 2 we present our main results and their consequences. In particular, we enumerate the number of binary trees of given path length and count the number of nodes in a randomly selected tree in the $\mathcal{T}_p$ model. Derivations are given in Sections 3–8. Numerical studies are discussed in Section 9.

## 2 Summary of Results

We let $b(n,p)$ denote the number of binary trees with $n$ nodes and path length $p$. This function satisfies the recurrence relation

$$b(n,p) = \sum_{k+\ell=n-1} \sum_{r+s+n-1=p} b(k,r)b(\ell,s), \quad n \geq 1$$

with the boundary conditions

$$b(0,0) = 1; \quad b(0,p) = 0, \quad p \geq 1.$$

The generating function

$$B_n(w) = \sum_{p=0}^{\infty} b(n,p)w^p$$

becomes

$$B_{n+1}(w) = w^n \sum_{\ell=0}^{n} B_\ell(w)B_{n-\ell}(w), \quad n \geq 0 \tag{2.1}$$

with $B_0(w) = 1$. Furthermore, the double generating function

$$B(z,w) = \sum_{n=0}^{\infty}\sum_{p=0}^{\infty} b(n,p)w^p z^n = \sum_{n=0}^{\infty} z^n B_n(w)$$

satisfies the functional equation

$$B(z,w) = 1 + zB^2(zw,w). \tag{2.2}$$

We shall mostly analyze (2.1), and then obtain asymptotic results for $b(n,p)$ by expanding the Cauchy integral (cf. [24])

$$b(n,p) = \frac{1}{2\pi i}\int_C B_n(w)w^{-p-1}dw. \tag{2.3}$$

Here $C$ is any closed loop about the origin in the $w$-plane.

We can solve (2.2) when $w = 1$, noting that $B(0,1) = 1$, to obtain

$$B(z,1) \equiv a(z) = \frac{1}{2z}\left[1 - \sqrt{1-4z}\right]$$

and thus

$$\sum_{p=0}^{\infty} b(n,p) = \frac{1}{2\pi i}\int_C \frac{a(z)}{z^{n+1}}dz = B_n(1) = \frac{1}{n+1}\binom{2n}{n} \tag{2.4}$$

is the Catalan number. This gives the total number of trees with $n$ nodes, regardless of the total path length. By expanding (2.2) about $w = 1$, with

$$B_n(w) = a_n + b_n(w-1) + \frac{1}{2}c_n(w-1)^2 + O((w-1)^3) \tag{2.5}$$

$$B(z,w) = a(z) + b(z)(w-1) + \frac{1}{2}c(z)(w-1)^2 + O((w-1)^3) \tag{2.6}$$

we are led to

$$b(z) = B_w(z,1) = \frac{2z^2 B(z,1) B_z(z,1)}{1 - 2zB(z,1)} = \frac{2z^2 a(z) a'(z)}{1 - 2za(z)}$$

and thus

$$b_n = B_n'(1) \equiv \sum_{p=1}^{\infty} p b(n,p) = 4^n - \frac{3n+1}{n+1}\binom{2n}{n}, \quad n \geq 0. \tag{2.7}$$

This gives the average total path length. Higher-order moments can be obtained in a similar manner. In particular, we obtain

$$\begin{aligned} c_n &= B_n''(1) = \sum_{p=2}^{\infty} p(p-1) b(n,p) \\ &= -4^n\left(\frac{13}{2}n + 4\right) + \binom{2n}{n}\left[\frac{10}{3}n^2 + \frac{44}{3}n + 2 + \frac{2}{n+1}\right], \quad n \geq 0. \end{aligned} \tag{2.8}$$

Asymptotically, for $n \to \infty$, we obtain from (2.4), (2.9) and (2.10) via Stirling's formula

$$\begin{aligned} a_n &= \frac{4^n}{\sqrt{\pi} n^{3/2}}[1 + O(n^{-1})] \\ b_n &= 4^n\left[1 - \frac{3}{\sqrt{\pi n}} + O(n^{-1})\right] \\ c_n &= 4^n\left[\frac{10}{3\sqrt{\pi}} n^{3/2} - \frac{13}{2} n + O(\sqrt{n})\right]. \end{aligned} \tag{2.9}$$

The constants $b_n$ and $c_n$ are related to the first two moments of the Airy distribution (cf. [7] and [25]). We can easily show that for each $j$

$$\frac{B_n^{(j+1)}(1)}{B_n^{(j)}(1)} = O(n^{3/2}), \quad n \to \infty. \tag{2.10}$$

It is known [18, 19, 25, 26] that the distribution of the total path length $L_n$, that is,

$$\Pr\{L_n = p\} = \frac{b(n,p)}{\displaystyle\sum_{p=0}^{\infty} b(n,p)} \tag{2.11}$$

follows an Airy distribution as $n \to \infty$, and most of the mass occurs in the range $p = O(n^{3/2})$. In particular,

$$\begin{aligned} \mathbf{E}[L_n] &= \sqrt{\pi} n^{3/2} + O(n), \\ \mathbf{Var}\,[L_n] &= \left(\frac{10}{3} - \pi\right) n^3 + O(n^{5/2}). \end{aligned}$$

More precisely,

$$\Pr\{L_n/\sqrt{2n^3} \leq x\} \xrightarrow{d} \Pr\{\mathcal{A} \leq x\}$$

where $\xrightarrow{d}$ denotes convergence in distribution and $\mathcal{A}$ is the random variable possessing the Airy distribution. It is characterized by moments [7]

$$\mathbf{E}[\mathcal{A}^r] = \frac{-\Gamma(-1/2)}{\Gamma((3r-1)/2)}\Omega_r$$

where $\Omega_r$ are determined by the recurrence

$$2\Omega_r = (3r-4)r\Omega_{r-1} + \sum_{j=1}^{r-1}\binom{r}{j}\Omega_j\Omega_{r-j}, \quad r \geq 1$$

with $\Omega_0 = -1$. Following Flajolet and Louchard [7] we observe that

$$\sum_{r\geq 0}\Omega_r\frac{w^r}{r!} = -\frac{\Phi_{2/3}(w)}{\Phi_{-1/3}(w)}, \qquad \Phi_\nu = {}_2F_0\left(\frac{1}{2}+\nu, \frac{1}{2}-\nu; \frac{3}{2}w\right)$$

where

$${}_2F_0(a,b;z) = 1 + \frac{a\cdot b}{1!}z + \frac{a(a+1)\cdot b(b+1)}{2!}z^2 + \cdots$$

is the generalized hypergeometric function (cf. [1]), and the above is a formal power series (that actually diverges).

A more difficult problem studied in this paper is to investigate the distribution of the number of nodes in trees of a fixed path length $p$, that is, for the ensemble $\mathcal{T}_p$. Let $N_p$ be the number of nodes for a tree uniformly generated from $\mathcal{T}_p$. It is a random variable distributed as

$$\Pr\{N_p = n\} = \frac{b(n,p)}{\sum_{n=0}^{\infty} b(n,p)}. \tag{2.12}$$

We shall compute this distribution asymptotically, and also obtain the asymptotic structure of $b(n,p)$ for various ranges of $n$ and $p$. We note that the sums in (2.11) and (2.12) are actually finite, since $b(n,p)$ is only non-zero in the range

$$\sum_{J=2}^{n}\lfloor\log_2 J\rfloor = p_{\min}(n) \leq p \leq p_{\max}(n) = \binom{n}{2}. \tag{2.13}$$

Here $p_{\min}$ and $p_{\max}$ are the minimal and maximal total path lengths possible in a tree with $n$ nodes. If we view the problem as having $p$ fixed and varying $n$, then $b(n,p)$ is non-zero in the range $n \in [n_{\min}(p), n_{\max}(p)]$ where

$$n_{\min}(p) = \min\{n : \binom{n}{2} \geq p\}$$

and

$$n_{\max}(p) = \max\left\{n : \sum_{J=2}^{n}\lfloor\log_2 J\rfloor \leq p\right\}.$$

Asymptotically, for $n \to \infty$, $[p_{\min}, p_{\max}] \sim \left[n \log_2 n, \frac{n^2}{2}\right]$ and, for $p \to \infty$, $[n_{\min}, n_{\max}] \sim \left[\sqrt{2p}, \frac{p}{\log_2 p}\right]$.

We now summarize our main results. Our derivations are quite complicated, and are delayed until Sec. 3– 8. In passing, we should add that we use ideas of applied mathematics, such as linearization and asymptotic matching. We shall make certain assumptions about the forms of the asymptotic expansions, as well as the asymptotic matching between the various scales. In particular, we shall use the WKB method discussed in the introduction.

We next formulate our main result concerning the cardinality of $\mathcal{T}_p$.

**Result 1** *The total number of trees of path length p is, for $p \to \infty$*

$$|\mathcal{T}_p| = \sum_{n=0}^{\infty} b(n,p) = \frac{1}{(\log_2 p)\sqrt{\pi p}} \exp\left[\frac{2p\log 2}{\log_2 p}\left(1 - \frac{3}{2}A_0 \frac{\log 2}{a^{1/3}} Q^{-2/3} + M(Q)Q^{-1} + O(Q^{-4/3})\right)\right] \tag{2.14}$$

*where $Q = \log p$ and*

$$M(Q) \quad = \quad (\log Q)(1 + A_1 \log 2) - \log\log 2 + (k_2 - A_1 \log a)\log 2, \tag{2.15}$$

$$A_0 = \frac{2}{3}4^{1/3}|r_0| = 2.4743\ldots, \quad A_1 = \frac{1}{\log 2} - \frac{1}{3} = 1.1093\ldots, \tag{2.16}$$

$$a = 2(\log 2)^2 = .96090\ldots, \quad r_0 = \max\{z : \mathrm{Ai}(z) = 0\} = -2.3381\ldots,$$

*Here $k_2 \approx 3.696$ is obtained by numerically solving a nonlinear integral equation, and $\mathrm{Ai}(\cdot)$ is the Airy function [2], defined as a solution of the differential equation $f'' - zf = 0$ that decays as $z \to \infty$.*

It follows that the exponential growth rate of the total number of trees of path length $p$ is

$$\log\left[\sum_n b(n,p)\right] \sim \frac{p}{\log p} 2(\log 2)^2 \tag{2.17}$$

with the correction terms involving the least negative root of the Airy function. This result was also recently obtained by Seroussi [23]. We will indicate how to formally obtain further terms in the asymptotic series in (2.14).

Next, we discuss the random variable $N_p$. Let

$$\mathcal{N}(p) := \mathbf{E}[N_p] = \frac{\sum_{n=0}^{\infty} n b(n,p)}{\sum_{n=0}^{\infty} b(n,p)}, \qquad \mathcal{V}(p) := \mathbf{Var}\,[N_p] = \frac{\sum_{n=0}^{\infty} (n - \mathcal{N}(p))^2 b(n,p)}{\sum_{n=0}^{\infty} b(n,p)}.$$

**Result 2** *With the notation as above, we find that*

$$\mathcal{N}(p) = \frac{p}{Q}\log 2\left[1-\frac{\log 2}{a^{1/3}}\frac{A_0}{Q^{2/3}}+\frac{M(Q)-A_1\log 2}{Q}+O(Q^{-4/3})\right], \tag{2.18}$$

$$\mathcal{V}(p) = \frac{p}{Q^{5/3}}\frac{(\log 2)A_0}{6a^{1/3}}\left[1-\frac{3A_1}{A_0}\left(\frac{a}{Q}\right)^{1/3}+O(Q^{-2/3})\right], \tag{2.19}$$

*where $A_0$, $A_1$ are defined in (2.16) and $M(Q)$ is given by (2.15). Furthermore, the limiting distribution of $N_p$ is Gaussian, that is,*

$$\Pr\{N_p=n\}=\frac{b(n,p)}{\sum_{n=0}^{\infty}b(n,p)}\sim\frac{1}{\sqrt{2\pi\mathcal{V}(p)}}\ \exp\left[-\frac{(n-\mathcal{N}(p))^2}{2\mathcal{V}(p)}\right], \tag{2.20}$$

*for $p\to\infty$ and $n-\mathcal{N}(p)=O(\mathcal{V}^{1/2}(p))=O(\sqrt{p}(\log p)^{-5/6})$.*

We note that while the most important scale for (2.11) is $p=O(n^{3/2})$, that for (2.12) is

$$p=n\log_2 n+O[n(\log n)^{1/3}],$$

which is close to the lower limit $p_{\min}(n)$ (or upper limit $n_{\max}(p)$) of the support of $b(n,p)$.

The above results are derived through the following main technical result. It gives detailed asymptotic results for the solution $B_n(w)$ to (2.1) as $n\to\infty$, for various ranges of $w$.

This main result makes certain assumptions about the forms of various asymptotic expansions and the matching between them. Specifically, the result in item (a) below assumes that the function in (3.1) has the limiting behavior in (3.5). The result in (b) is based upon the WKB expansion in (4.9), while (2.20) assumes the asymptotic matching between the ranges in (a) and (b). To obtain the result in (c) we assumed that the function in (5.2) has an expansion of the form (5.101). The asymptotic matching assumption between ranges (b) and (c) determines a multiplicative constant in (2.21) The result in part (d) assumes the limit in (6.5), and (2.29) (respecively, (2.30)) assumes asymptotic matching between ranges (d) and (e) (respectively, (e)). The expansion in region (e) assumes the WKB form in (7.1).

**Result 3** *Consider binary trees with path length equal to $p$. Let $B_n(w)$ be its generating function satisfying (2.1). Then for $n\to\infty$ we have the following asymptotic expansions.*

*(a)* **far right region***: $n\to\infty$, $w\gg 1$*

$$B_n(w)\sim w^{\binom{n}{2}}2^{n-1}B_*(w) \tag{2.21}$$

*where $B_*(w)$ satisfies*

$$B_*(w)=1+\frac{1}{4w}+\frac{1}{2w^2}+O(w^{-3}),\quad w\to\infty; \tag{2.22}$$

$$B_*(w)\sim d_1\sqrt{w-1}\exp\left(\frac{d_0}{w-1}\right),\quad w\to 1^+, \tag{2.23}$$

$$d_0 = \int_0^{\log 2} \frac{\xi}{e^{\xi}-1} d\xi = .58224\ldots, \qquad d_1 = \frac{4}{\sqrt{2\pi}} e^{d_0/2} = 2.1350\ldots. \tag{2.24}$$

*The numerical calculation of $B_*(w)$ is discussed in Sections 3 and 9.*

*(b)* **right region***:* $w = 1 + \beta/n$, $0 < \beta < \infty$

$$B_n(w) \sim \sqrt{\frac{\beta}{n}} \hat{g}(\beta) \exp[n\Phi(\beta)], \tag{2.25}$$

*where*

$$\begin{aligned} \Phi(\beta) &= \log 2 + \frac{\beta}{2} + \frac{1}{\beta} \int_{-\log\left(1-\frac{1}{2}e^{-\beta}\right)}^{\log 2} \frac{\xi}{e^{\xi}-1} d\xi \equiv \log 2 + \frac{\beta}{2} + \phi(\beta), \\ \hat{g}(\beta) &= \frac{4}{\sqrt{\pi}} e^{-\beta^2/4} e^{-\beta/2} \left(\frac{1-e^{-\beta}}{2-e^{-\beta}}\right)^{3/2} \exp\left[\frac{1}{2}\beta\phi(\beta) + \frac{1}{2}\beta \log\left(1 - \frac{1}{2}e^{-\beta}\right)\right]. \end{aligned}$$

*(c)* **central region***:* $w = 1 + a/n^{3/2}$, $-\infty < a < \infty$

$$B_n(w) = \frac{1}{n+1}\binom{2n}{n} + \frac{4^n}{n^{3/2}} \left[C(a) + \frac{1}{\sqrt{n}} C^{(1)}(a) + O(n^{-1})\right], \tag{2.26}$$

*where*

$$\begin{aligned} C(a) &= (-a)\bar{D}((-a)^{2/3}) = Y^{3/2}\bar{D}(Y), \;\; Y = (-a)^{2/3}, \;\; a < 0, \\ \bar{D}(Y) &= \frac{1}{2\pi i} \int_{Br} e^{sY} \left[2\sqrt{s} + 4^{2/3} \frac{\mathrm{Ai}'(4^{-1/3}s)}{\mathrm{Ai}(4^{-1/3}s)}\right] ds. \end{aligned}$$

*Here $Br$ is a vertical contour on which $\Re(s) > 0$, and $\sqrt{s}$ is analytic for $\Re(s) > 0$ and positive for $s$ real and positive. An alternate expression for the leading term is, for $a = -Y^{3/2} < 0$,*

$$\begin{aligned} B_n(w) &\sim \frac{4^n}{n^{3/2}}(-a)\frac{d}{dY}\left[\frac{1}{2\pi i}\int_{Br} \frac{4^{2/3}}{s} \frac{\mathrm{Ai}'(4^{-1/3}s)}{\mathrm{Ai}(4^{-1/3}s)} e^{sY} ds\right] \\ &= \frac{4^{n+1}}{n^{3/2}}(-a) \sum_{j=0}^{\infty} \exp(-|r_j| 4^{1/3} Y) \end{aligned} \tag{2.27}$$

*where $0 > r_0 > r_1 > r_2 > \ldots$ and $r_j$ are the roots of $\mathrm{Ai}(z) = 0$. The correction term has the integral representation, for $a < 0$,*

$$C^{(1)}(a) = -a\bar{D}_1(Y) = \frac{Y^2}{2\pi i} \int_{Br} e^{sY} \mathcal{E}_*(s) ds, \tag{2.28}$$

$$\begin{aligned} \mathcal{E}_*(s) &= -\frac{5}{2}s + 8\left(\frac{h'(s)}{h(s)}\right)^2 - \frac{4}{h^2(s)} \int_s^{\infty} \frac{(h'(v))^3}{h(v)} dv \\ &= -\frac{5}{2}s + 10\left(\frac{h'(s)}{h(s)}\right)^2 + 4\left(\frac{h'(s)}{h(s)}\right)^2 \log[h(s)] - s\log[h(s)] \\ &- \frac{1}{h^2(s)} \int_s^{\infty} h^2(v) \log[h(v)] dv, \qquad h(s) = \mathrm{Ai}(4^{-1/3}s). \end{aligned}$$

*For $a > 0$ we let $a = y^{3/2}$ with $y > 0$ and the leading term is*

$$B_n(w) \sim \frac{4^n}{n^{3/2}} \left( \frac{a}{\pi^2 4^{1/3}} \int_0^\infty \frac{e^{\tau y}}{h(\omega\tau)h(\omega^2\tau)} d\tau - \frac{4a}{\pi} \int_0^\infty \Re\left[ e^{\pi i/6} \frac{h'(\omega\tau)}{h(\omega\tau)} e^{\omega^2 \tau y} \right] d\tau \right) \quad (2.29)$$

*where $\omega = \exp(2\pi i/3)$.*

*(d)* **left region***: $w = 1 - \gamma/n$, $0 < \gamma < \infty$*

$$\begin{aligned} B_n(w) &\sim \frac{4^n}{n} \exp[\nu_0 n^{1/3} \gamma^{2/3} + \nu_1 \gamma \log n] F_0(\gamma), \\ F_0(\gamma) &= 4\gamma F_1(\gamma), \\ \nu_0 &= 4^{1/3} r_0 = -4^{1/3}|r_0|, \quad \nu_1 = -\frac{1}{3}, \end{aligned} \quad (2.30)$$

*where $F_1(\cdot)$ satisfies the non-linear integral equation*

$$\begin{aligned} \frac{e^\gamma - 1}{\gamma} F_1(\gamma) &= \int_0^1 F_1(\gamma x) F_1(\gamma - \gamma x) e^{-\gamma H(x)/3} dx, \\ H(x) &= x \log x + (1-x) \log(1-x) \end{aligned} \quad (2.31)$$

*and behaves, for $\gamma \to 0^+$, as*

$$F_1(\gamma) = 1 - \frac{2}{3} \gamma \log \gamma + \alpha_1 \gamma + O(\gamma)$$

*where*

$$\begin{aligned} \alpha_1 &= \frac{7}{2} - \gamma_E + \log[h'(s_0)] - \frac{1}{4[h'(s_0)]^2} \int_{s_0}^\infty h^2(v) \log[h(v)] dv = 2.9622\ldots, \\ s_0 &= 4^{1/3} r_0, \quad \gamma_E = \text{Euler's constant} = .57721\ldots. \end{aligned} \quad (2.32)$$

*For $\gamma \to \infty$ we have*

$$F_1(\gamma) \sim \frac{1}{\sqrt{2\pi \log 2}} \frac{e^{k_2 \gamma}}{\sqrt{\gamma}} \exp\left[ \left( \frac{1}{3} - \frac{1}{\log 2} \right) \gamma \log \gamma \right] \quad (2.33)$$

*where $k_2 \approx 3.696$ is found numerically (cf. Sections 6 and 9).*

*(e)* **far left region***: $n \to \infty$, $0 < w \ll 1$*

$$\begin{aligned} B_n(w) &\sim e^{n(\log_2 n) \log w} e^{n[g(w) + B_0^*(w,n)]} n^{\log_2 w} (2\pi n)^{-1/2} \\ &\times e^{g(w)} w^{2 + \frac{1}{\log 2}} e^{B_0^*(w,n) + B_1^*(w,n)} \sqrt{-\log_2 w - B_1^*(w,n) - \frac{1}{4} B_2^*(w,n)}. \end{aligned} \quad (2.34)$$

*Here*

$$
\begin{aligned}
B_0^*(w,n) &= \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} g_k(w) e^{2\pi i(\log_2 n)k} \\
B_1^*(w,n) &= \frac{2\pi i}{\log 2} \sum_{k=-\infty}^{\infty} k g_k(w) e^{2\pi i(\log_2 n)k} \\
B_2^*(w,n) &= \frac{2\pi i}{\log 2} \sum_{k=-\infty}^{\infty} \left[\frac{2\pi i}{\log 2} k^2 - k\right] g_k(w) e^{2\pi i(\log_2 n)k}
\end{aligned}
\tag{2.35}
$$

*and $g(w)$ has the asymptotic expansion*

$$
\begin{aligned}
g(w) = \log 4 + 4^{1/3} r_0 (1-w)^{2/3} + \left(\frac{1}{\log 2} - \frac{1}{3}\right)(w-1)\log(1-w) \\
-k_2(w-1) + o(w-1), \quad w \to 1^-.
\end{aligned}
\tag{2.36}
$$

The numerical calculation of $g(w)$ is discussed in Sections 7 and 9. The sum in $B_0^*$ omits the term $k=0$ and the non-constant Fourier coefficients $g_k(w)$ satisfy $g_k(w) = o(w-1)$ as $w \to 1^-$. Numerical studies show that unless $w$ is very small, the $g_k(\cdot)$ are very small and we can use the approximation

$$
B_n(w) \approx w^{n\log_2 n} e^{(n+1)g(w)} n^{\log_2 w} n^{-1/2} w^{2+\frac{1}{\log 2}} \sqrt{\frac{-\log_2 w}{2\pi}}. \tag{2.37}
$$

This corresponds to neglecting $B_j^*$ in (2.34) for $j = 0,1,2$.

We comment that our analysis suggests that yet another scale exists, which has $n \to \infty$ and $w \to 0$ simultaneously, and where a different expansion for $B_n(w)$ is needed. We have not been able to analyze this scale, but it is not needed to obtain the asymptotic results for the number of trees of a given total path length. For this the important range is the asymptotic matching region between the left and far left regions, corresponding to $w \to 1^-$, but $n(1-w) = \gamma \to +\infty$. Since we have explicit analytic results for $g(w)$ as $w \to 1^-$, and $g_k(w) \to 0$ for $k \neq 0$, we can use the above results to obtain the explicit expressions in (2.14) - (2.20). To obtain the distribution of the path length in trees with $n$ ($\to \infty$) nodes, the central region (c) is the most important, and the leading term corresponds to (the transform of) the Airy distribution.

We next give results for $b(n,p)$ for $n$ and $p \to \infty$, and summarize the main results as items (A)-(E) below. Going from (A) to (E) corresponds to increasing $n$ or decreasing $p$.

**Result 4** *Consider binary trees built over $n$ nodes with path length $p$. Let $b(n,p)$ denote the number of such trees. Then we have the following for $p, n \to \infty$.*

*(A)*

$$
n \to \infty, \quad p = \binom{n}{2} - L, \quad L = O(1), \quad L \geq 0
$$

$$
b(n,p) \sim 2^{n-1} \frac{1}{2\pi i} \int_C w^{L-1} B_*(w)\, dw \tag{2.38}
$$

*where $C$ is a closed loop with $|w| > 1$ and $B_*(w)$ is as in (a) in Result 3.*

*(B)*

$$p, n \to \infty \text{ with } \Lambda = p/n^2 \in \left(0, \frac{1}{2}\right)$$

$$b(n,p) \sim \frac{\sqrt{2}}{\pi}\beta_* e^{-\beta_*/2}\left(\frac{1-e^{-\beta_*}}{2-e^{-\beta_*}}\right)^{3/2}\left[1-2\Lambda-\frac{1}{2e^{\beta_*}-1}\right]^{-1/2} \\ \times\frac{2^{n+1}}{n^2}\exp\left(n\left[\beta_*(1-2\Lambda)-\log\left(1-\frac{1}{2}e^{-\beta_*}\right)\right]\right) \tag{2.39}$$

*where* $\beta_* \equiv \beta_*(\Lambda)$ *is defined implicitly by*

$$\beta_*^2\left(\frac{1}{2}-\Lambda\right)+\beta_*\log\left(1-\frac{1}{2}e^{-\beta_*}\right) = \int_{-\log\left(1-\frac{1}{2}e^{-\beta_*}\right)}^{\log 2}\frac{\xi}{e^{\xi}-1}d\xi.$$

*(C)*

$$p, n \to \infty \textit{ with } \Omega = p/n^{3/2} \in (0,\infty)$$

$$b(n,p) \sim -\frac{4^n}{n^3}\left(\frac{1}{3\Omega}\right)^{1/3}\sum_{j=0}^{\infty}\left\{\left[\frac{56}{9}\frac{4^{2/3}r_j^2}{\Omega^3}+\frac{64}{81}\frac{4^{5/3}r_j^5}{\Omega^5}\right]\mathrm{Ai}\left(\frac{r_j^2 4^{2/3}}{3^{4/3}\Omega^{4/3}}\right)\right. \\ \left.+\left(\frac{1}{3\Omega}\right)^{1/3}\left[\frac{40}{3}\frac{4^{1/3}r_j}{\Omega^2}+\frac{64}{27}\frac{4^{4/3}r_j^4}{\Omega^4}\right]\mathrm{Ai}'\left(\frac{r_j^2 4^{2/3}}{3^{4/3}\Omega^{4/3}}\right)\right\}\exp\left(-\frac{8|r_j|^3}{27\Omega^2}\right). \tag{2.40}$$

*Here* $r_j < 0$ *are the roots of the Airy function.*

*(D)*

$$p, n \to \infty \text{ with } \Theta = p/n^{4/3} \in (0,\infty)$$

$$\begin{aligned} b(n,p) &\sim \frac{4^n}{n^{13/6}}n^{-\gamma_*/3}\frac{2^9}{3^4\sqrt{\pi}}\frac{|r_0|^{9/2}}{\Theta^5}F_1(\gamma_*)\exp\left[-\frac{16n^{1/3}|r_0|^3}{27\Theta^2}\right], \\ \gamma_* &= \frac{32}{27}\frac{|r_0|^3}{\Theta^3} \end{aligned} \tag{2.41}$$

*where* $F_1(\cdot)$ *satisfies (2.31) (cf. item (d) in Result 3).*

*(E)*

$$p, n \to \infty \textit{ with } p = n\log_2 n + \alpha n,\ \alpha = O(1)$$

$$b(n,p) \approx \frac{n^{\log_2 w_*}}{2\pi n}\frac{w_*^{2+\frac{1}{\log 2}}}{\sqrt{\alpha + w_*^2 g''(w_*)}}e^{g(w_*)}\sqrt{-\log_2 w_*} \\ \times\exp\left[ng(w_*)-n\alpha\log w_*\right] \tag{2.42}$$

*where* $w_* = w_*(\alpha)$ *is the solution to* $w_* g'(w_*) = \alpha$*.*

From item (C) the Airy distribution can be recovered by dividing by the expansion of $a_n$ in (2.6). Then (2.37) becomes of the form $n^{-3/2}$ times the Airy density.

In obtaining (2.42) we used (2.34) in (2.3) and neglected the non-constant terms in the Fourier series, which are numerically small. Then we evaluated the integral by the saddle point method (cf. [24]). A refined approximation that uses also the non-constant terms in the Fourier series' in (2.35), can be obtained by using (2.34) in (2.3). We can also obtain an $O(n^{-1/2})$ correction term to the Airy distribution in (2.40) by using the correction term (i.e., $C^{(1)}(a)$) in (2.26) in asymptotically inverting (2.3). Our approximation(s) to $b(n,p)$ involve the unknown functions $B_*(w)$, $F_1(\gamma)$ and $g(w)$, whose numerical calculation we discuss in Section 9.

In view of the complexity of the results in items (A)-(E), we can get more insight into the structure of $b(n,p)$ by giving formulas that apply in the asymptotic matching regions between the various scales. We summarize these below, with the notation (AB) denoting the asymptotic matching region between the scales in (A) and (B) , and so on. Note that the (AB) result can be obtained by either expanding (2.38) as $L \to \infty$ (using (2.23)), or by expanding (2.39) as $\Lambda \to \left(\frac{1}{2}\right)^-$.

**Result 5** *The following matching asymptotics hold:*

**(AB)**

$$n, p \to \infty;\ L = \binom{n}{2} - p \to \infty,\ \Lambda = p/n^2 \to \frac{1}{2}$$

$$b(n,p) \sim \frac{2^n}{\pi n^2}\left(\frac{2d_0}{1-2\Lambda}\right)^{1/2}\frac{1}{\sqrt{1-2\Lambda}}\exp\left(n\sqrt{2d_0(1-2\Lambda)} - \frac{1}{2}\sqrt{\frac{2d_0}{1-2\Lambda}}\right) \tag{2.43}$$

*where $d_0$ is given in (2.24).*

**(BC)**

$$n, p \to \infty;\ \Lambda = p/n^2 \to 0,\ \Omega = p/n^{3/2} \to \infty$$

$$b(n,p) \sim \frac{4^n}{n^2}\frac{9\sqrt{3}}{2\pi}\Lambda^2 \exp\left(-\frac{3}{4}n\Lambda^2\right) = \frac{4^n}{n^3}\frac{9\sqrt{3}}{2\pi}\Omega^2 \exp\left(-\frac{3}{4}\Omega^2\right). \tag{2.44}$$

**(CD)**

$$n, p \to \infty;\ \Omega = p/n^{3/2} \to 0,\ \Theta = p/n^{4/3} \to \infty$$

$$\begin{aligned} b(n,p) &\sim \frac{4^n}{n^{13/6}}\frac{|r_0|^{9/2}}{\Theta^5}\frac{2^9}{3^4\sqrt{\pi}} \exp\left[-\frac{16n^{1/3}|r_0|^3}{27\Theta^2}\right] \\ &= \frac{4^n}{n^3}\frac{|r_0|^{9/2}}{\Omega^5}\frac{2^9}{3^4\sqrt{\pi}} \exp\left[-\frac{16|r_0|^5}{27\Omega^2}\right]. \end{aligned} \tag{2.45}$$

**(DE)**

$$n, p \to \infty;\ \Theta = p/n^{4/3} \to 0,\ \alpha = p/n - \log_2 n \to \infty$$

$$b(n,p) \sim \frac{1}{n^{13/6}} \frac{|r_0|^3}{\Theta^{7/2}} \frac{1}{\pi\sqrt{\log 2}} \frac{64}{9\sqrt{3}} \ \exp\Big\{ n\log 4 - \frac{\gamma_*}{3}\log n + \left(\frac{1}{3} - \frac{1}{\log 2}\right)\gamma_* \log\gamma_* + k_2\gamma_* - \frac{16}{27}\frac{n^{1/3}|r_0|^3}{\Theta^2} \Big\} \tag{2.46}$$

*where $\gamma_*$ is given below (2.41).*

We will show in Section 8 that the asymptotic matching region (DE) leads to the Gaussian distribution in (2.20). Note that in each of the four matching regions, our results are completely explicit functions of $n$ and $p$. The result in (BC) (resp., (CD)) gives the right (resp., left) tail of the Airy distribution in (2.40). The right tail has not been characterized this precisely in previous studies [5] (cf. also [7, 18, 19, 25, 26, 27]).

# 3 Far Right Region

We consider (2.1) for a fixed $w > 1$ and $n \to \infty$. We set

$$B_n(w) = w^{\binom{n}{2}} 2^{n-1} \bar{B}_n(w) \tag{3.1}$$

to find that (2.1) becomes

$$\begin{aligned} \bar{B}_{n+1}(w) &= \frac{1}{4}\sum_{\ell=0}^{n} w^{-\ell(n-\ell)} \bar{B}_\ell(w)\bar{B}_{n-\ell}(w) \\ &= \frac{1}{4}[2\bar{B}_0(w)\bar{B}_n(w) + 2w^{1-n}\bar{B}_1(w)\bar{B}_{n-1}(w) \\ &\qquad + 2w^{4-2n}\bar{B}_2(w)\bar{B}_{n-2}(w) + \cdots]. \end{aligned} \tag{3.2}$$

But $B_0(w) = B_1(w) = 1$ and $B_2(w) = 2w$ so that (3.1) yields

$$\bar{B}_0(w) = 2,\ \bar{B}_1(w) = 1,\ \bar{B}_2(w) = 1 \tag{3.3}$$

and (3.2) becomes

$$\bar{B}_{n+1}(w) = \bar{B}_n(w) = \frac{1}{2}w^{1-n}\bar{B}_{n-1}(w) + \frac{1}{4}w^{4-2n}\bar{B}_{n-2}(w) + O(w^{-3n}) \tag{3.4}$$

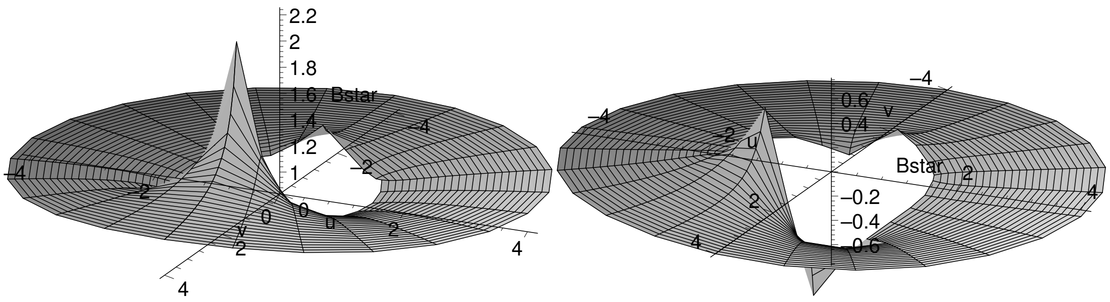
whose asymptotic solution is

$$\bar{B}_n(w) = B_*(w) + O(w^{-n}) \tag{3.5}$$

for some function $B_*(\cdot)$.

From (2.1) we obtain the first few $B_n(w)$ as

$$\begin{aligned} B_1(w) = 1,\ B_2(w) = 2w,\ B_3(w) &= w^2 + 4w^3 \\ B_4(w) &= 4w^4 + 2w^5 + 8w^6 \\ B_5(w) &= 6w^6 + 8w^7 + 8w^8 + 4w^9 + 16w^{10} \end{aligned} \tag{3.6}$$




(a) $\Re[B_*(w)]$ (b) $\Im[B_*(w)]$

**Fig. 1:** Plots of $\Re[B_*(w)]$ and $\Im[B_*(w)]$ for $1.5 < |w| < 4$.

and it is easy to inductively establish that for $w \to \infty$ and fixed $n$

$$B_n(w) = 2^{n-1}w^{\binom{n}{2}} + 2^{n-3}w^{\binom{n}{2}-1} + 2^{n-2}w^{\binom{n}{2}-2} + O\left(w^{\binom{n}{2}-3}\right), \tag{3.7}$$

for $n \geq 4$. By comparing (3.7) to (3.1) with (3.5) we conclude that

$$B_*(w) = 1 + \frac{1}{4w} + \frac{1}{2w^2} + O(w^{-3}),\ w \to \infty. \tag{3.8}$$

In section 4 we will argue, by asymptotic matching, that as $w \downarrow 1$ we have

$$B_*(w) \sim d_1\sqrt{w-1}\ \exp\left(\frac{d_0}{w-1}\right),\ w \to 1^+ \tag{3.9}$$

where the constants $d_0$ and $d_1$ are given in (2.24). We have not been able to determine $B_*(w)$ analytically except for its behaviors as $w \to \infty$ and $w \downarrow 1$. It can easily be obtained numerically by fixing $w > 1$ and iterating (3.2) until $\bar{B}_n(w)$ settles to a limit to some prescribed accuracy.

In Table 1, we give $B_*(w)$ for $w$ in the range $[1.04, 5]$. We note that the convergence of this procedure becomes very slow (and $B_*(w)$ becomes very large) when $w$ exceeds 1 only slightly. This is certainly consistent with the asymptotic analysis, that predicts another scale where $n \to \infty$ with $n(w-1)$ fixed.

Our analysis thus far assumed that $w$ is *real*. However, the same arguments show that (3.1) remains valid for complex $w$ with $|w| > 1$. We can also use (3.2) to calculate $B_*(w)$ for complex $w$ with $|w| > 1$. In Figure 1 we plot $\Re[B_*(w)]$ and $\Im[B_*(w)]$ for $1.5 < |w| < 4$. These surfaces are viewed from the $+u$ perspective, where $w = u + iv$.

With (3.1) and (3.5) we can infer the behavior of $b(n,p)$ for $p$ close to $p_{\max}(n) = \binom{n}{2}$. Setting

$$L = \binom{n}{2} - p = O(1)$$

we obtain from (2.3)

$$b(n,p) \sim \frac{2^{n-1}}{2\pi i}\int_C w^{L-1}B_*(w)dw,\ n \to \infty, \tag{3.10}$$

**Tab. 1:**

| $w$ | $B_*(w)$ |
|---|---|
| 5 | 1.0757 |
| 4.5 | 1.0884 |
| 4 | 1.1060 |
| 3.5 | 1.1319 |
| 3 | 1.1735 |
| 2.5 | 1.2500 |
| 2 | 1.4306 |
| 1.8 | 1.5922 |
| 1.6 | 1.9184 |
| 1.4 | 2.8430 |
| 1.2 | 10.088 |
| 1.18 | 13.502 |
| 1.16 | 19.502 |
| 1.14 | 31.426 |
| 1.12 | 59.739 |
| 1.10 | 148.17 |
| 1.08 | 587.3 |
| 1.06 | 5981 |
| 1.04 | $6.55 \times 10^5$ |

where $C$ is a closed contour on which $|w| > 1$. We can use (3.8) to compute (3.10) for $L = 0, 1, 2, \ldots,$ and in Section 4 we will obtain the asymptotic behavior of the integral as $L \to \infty$.

# 4 Right Region

We consider the limit $w \downarrow 1$ and $n \to \infty$ with

$$n(w-1) = \beta = O(1),\ \beta > 0. \tag{4.1}$$

We let

$$B_n(w) = f(\beta; n) = f(n(w-1); n) \tag{4.2}$$

with which (2.1) becomes

$$f\left(\beta + \frac{\beta}{n}; n+1\right) = 2\left(1+\frac{\beta}{n}\right)^n \sum_{\ell=0}^{n/2} B_\ell\left(1+\frac{\beta}{n}\right) f\left(\beta\left(1-\frac{\ell}{n}\right); n-\ell\right). \tag{4.3}$$

We used the symmetry of the sum in (2.1) to truncate the upper limit on the sum in (4.3) at $\ell = n/2$. To be more precise, we should distinguish the cases where $n$ is odd or even, but it will become apparent that for $n \to \infty$ the asymptotically dominant terms in (4.3) are those with $\ell = O(1)$, and the terms near $\ell = n/2$ are exponentially smaller than the dominant ones. We also note that in (4.3) we re-wrote only the second factor in the convolution sum in (2.1) using the scaling in (4.1) and (4.2).

The behavior of $B_\ell(w)$ for fixed $\ell$ and $w \to 1$ follows from (2.5), and we have

$$B_\ell\left(1+\frac{\beta}{n}\right) = a_\ell + b_\ell \frac{\beta}{n} + O(n^{-2}) \tag{4.4}$$

where

$$a_\ell = \frac{1}{\ell+1}\binom{2\ell}{\ell}, \quad b_\ell = 4^\ell - \frac{3\ell+1}{\ell+1}\binom{2\ell}{\ell} \tag{4.5}$$

and these have the generating functions

$$\sum_{\ell=0}^{\infty} a_\ell z^\ell = \frac{1}{2z}\left[1-\sqrt{1-4z}\right] \tag{4.6}$$

$$\sum_{\ell=0}^{\infty} b_\ell z^\ell = \frac{1}{z}\left[1-\sqrt{1-4z}\right] + \frac{1}{1-4z} - \frac{3}{\sqrt{1-4z}}. \tag{4.7}$$

We also have

$$\left(1+\frac{\beta}{n}\right)^n = e^\beta\left[1-\frac{\beta^2}{2n} + O(n^{-2})\right],\ n \to \infty. \tag{4.8}$$

We assume, for fixed $\beta > 0$ and $n \to \infty$, that $f(\beta; n)$ has an expansion of the WKB form

$$f(\beta; n) = e^{n\Phi(\beta)} n^{-1/2}\left[g(\beta) + \frac{1}{n} g^{(1)}(\beta) + O(n^{-2})\right]. \tag{4.9}$$

The factor $n^{-1/2}$ must be included in order to asymptotically match to another expansion that applies for $w-1=O(n^{-3/2})$. This is the "central region" analyzed in section 5. In section 9 we provide numerical justification for the *ansatz* (4.9). The numerical studies suggest, however, that the ratio of the correction term to the leading term in (4.9) is larger than $O(n^{-1})$, perhaps $O(n^{-1/2})$. Thus the series in (4.9) may actually be in powers of $n^{-1/2}$ rather than $n^{-1}$. However, this will not affect our calculation of $\Phi(\beta)$ and $g(\beta)$.

We use (4.4) and (4.9) in (4.3) and extend the limit on the sum from $n/2$ to $\infty$. We note that

$$\begin{aligned}
&\exp\left[(n+1)\Phi\left(\beta+\frac{\beta}{n}\right)\right](n+1)^{-1/2}\left[g\left(\beta+\frac{\beta}{n}\right)+\frac{1}{n+1}g^{(1)}\left(\beta+\frac{\beta}{n}\right)+O(n^{-2})\right]\\
= \quad & e^{n\Phi(\beta)}e^{\beta\Phi'(\beta)+\Phi(\beta)}n^{-1/2}\left[1+\frac{1}{n}\left(\beta\Phi'(\beta)+\frac{1}{2}\beta^2\Phi''(\beta)\right)\right]\left[1-\frac{1}{2n}\right]\\
\times \quad & \left[g(\beta)+\frac{1}{n}(g^{(1)}(\beta)+\beta g'(\beta))+O(n^{-2})\right] \qquad (4.10)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{\ell=0}^{n/2} B_\ell\left(1+\frac{\beta}{n}\right) f\left(\beta\left(1-\frac{\ell}{n}\right);n-\ell\right)\\
= \quad & \sum_{\ell=0}^{n/2}\left[B_\ell(1)+\frac{\beta}{n}B_\ell'(1)+O(n^{-2})\right]\frac{1}{\sqrt{n-\ell}}\ \exp\left[(n-\ell)\Phi\left(\beta-\frac{\beta\ell}{n}\right)\right]\\
\times \quad & \left[g(\beta)+\frac{1}{n}(g^{(1)}(\beta)-\ell\beta g'(\beta))+O(n^{-2})\right]\\
= \quad & e^{n\Phi(\beta)}n^{-1/2}\sum_{\ell=0}^{\infty} e^{-\beta\ell\Phi'(\beta)}\left[a_\ell+\frac{\beta}{n}b_\ell\right]\left[1+\frac{\ell}{2n}\right]e^{-\ell\Phi(\beta)}\\
\times \quad & \left[1+\frac{\beta^2\ell^2}{2n}\Phi''(\beta)\right]\left[1+\frac{\beta\ell^2}{n}\Phi'(\beta)\right]\left[g(\beta)+\frac{1}{n}(g^{(1)}(\beta)-\ell\beta g'(\beta))+O(n^{-2})\right]. \qquad (4.11)
\end{aligned}$$

Dividing (4.3) by $2(1+\beta/n)^n$ and letting $n\to\infty$ we obtain, after cancelling the common factor $n^{-1/2}\exp[n\Phi(\beta)]$, the limiting equation

$$\frac{1}{2}e^{-\beta}e^{\beta\Phi'+\Phi}=\sum_{\ell=0}^{\infty}a_\ell e^{-\ell(\beta\Phi'+\Phi)}. \qquad (4.12)$$

The above is a non-linear ODE for the function $\Phi(\beta)$, which is in view of (4.9) the exponential growth rate of $B_n(w)$ on the $\beta$-scale. Using (4.6) to evaluate the sum in (4.12) leads to, after some simplification,

$$\frac{1}{2}e^{-\beta}=\frac{1}{2}\left[1-\sqrt{1-4e^{-(\beta\Phi)'}}\right] \qquad (4.13)$$

or

$$(\beta\Phi)'=\log 4+2\beta-\log(2e^{\beta}-1). \qquad (4.14)$$

Equation (4.14) integrates to

$$\begin{aligned}\Phi(\beta) &= \log 4 + \beta - \frac{1}{\beta}\int_0^\beta \log(2e^v - 1)dv \\ &= \log 2 + \frac{\beta}{2} - \frac{1}{\beta}\int_0^\beta \log\left(1 - \frac{1}{2}e^{-v}\right)dv \\ &= \log 2 + \frac{\beta}{2} + \phi(\beta) \qquad (4.15)\end{aligned}$$

where

$$\phi(\beta) = -\frac{1}{\beta}\int_0^\beta \log\left(1 - \frac{1}{2}e^{-v}\right)dv = \frac{1}{\beta}\int_{-\log\left(1-\frac{1}{2}e^{-\beta}\right)}^{\log 2} \frac{\zeta}{e^\zeta - 1}d\zeta. \qquad (4.16)$$

We note that $\Phi(\beta) = c/\beta$ is a homogeneous solution to (4.14), but such solutions must be discarded since asymptotic matching to the central region $(w - 1 = O(n^{-3/2}))$ will force $\Phi(\beta)$ to be bounded as $\beta \to 0^+$ (in fact $\Phi(0) = \log 4$).

We next determine $g(\beta)$ in (4.9). Using (4.8), (4.10), and (4.11) in (4.3), we obtain, at order $e^{n\Phi}n^{-1/2} \cdot n^{-1}$, the linear equation

$$\begin{aligned}&\frac{1}{2}e^{-\beta}e^{(\beta\Phi)'}\left[\beta g' + \left(\frac{\beta^2}{2}\Phi'' + \beta\Phi' + \frac{\beta^2}{2} - \frac{1}{2}\right)g\right] = -\beta g'\sum_{\ell=0}^\infty \ell a_\ell e^{-\ell(\beta\Phi)'} + \beta g\sum_{\ell=0}^\infty b_\ell e^{-\ell(\beta\Phi)'} \\ &+g\sum_{\ell=0}^\infty a_\ell\left(\frac{\ell}{2} + \frac{\beta^2\ell^2}{2}\Phi'' + \ell^2\beta\Phi'\right)e^{-\ell(\beta\Phi)'}. \qquad (4.17)\end{aligned}$$

Note that in view of (4.12) $g^{(1)}$ drops out of (4.17). The latter is a simple first order linear ODE since we have determined $\Phi$.

To integrate (4.17) we define the sums

$$S_j(\beta) = \sum_{\ell=0}^\infty \ell^j a_\ell e^{-\ell(\beta\Phi)'};\ j = 0, 1, 2 \qquad (4.18)$$

$$T_0(\beta) = \sum_{\ell=0}^\infty b_\ell e^{-\ell(\beta\Phi)'}. \qquad (4.19)$$

Setting

$$z = e^{-(\beta\Phi)'} = \frac{1}{4}e^{-\beta}(2 - e^{-\beta})$$

we have

$$\sum_{\ell=0}^\infty z^\ell a_\ell = S_0(\beta) = \frac{2}{2 - e^{-\beta}}. \qquad (4.20)$$

By differentiating (4.6) we have

$$\sum_{\ell=0}^{\infty} \ell z^{\ell} a_{\ell} = S_1(\beta) = \frac{z}{\sqrt{1-4z}}\left(\frac{1}{1+\sqrt{1-4z}}\right)^2 = \frac{1}{4}\left(\frac{1}{1-e^{-\beta}} - \frac{2}{2-e^{-\beta}}\right). \tag{4.21}$$

Also, by differentiating $S_0$ with respect to $\beta$ we obtain

$$S_0'(\beta) = -\sum_{\ell=0}^{\infty} \ell(\beta\Phi)'' a_{\ell} e^{-\ell(\beta\Phi)'} = -(\beta\Phi)'' S_1(\beta) \tag{4.22}$$

and then

$$S_0''(\beta) = [(\beta\Phi)'']^2 S_2(\beta) - (\beta\Phi)''' S_1(\beta). \tag{4.23}$$

Using (4.20) - (4.23) in (4.17) and noting that $\frac{\beta^2}{2}\Phi'' + \beta\Phi' = \frac{\beta}{2}(\beta\Phi)''$ leads to

$$\frac{1}{2}e^{-\beta}e^{(\beta\Phi)'}\left[\beta g' + \left(\frac{\beta}{2}(\beta\Phi)'' + \frac{\beta^2}{2} - \frac{1}{2}\right)g\right] = \beta g T_0 + \frac{S_0'}{(\beta\Phi)''}\beta g' - \frac{S_0'}{2(\beta\Phi)''}g + \frac{\beta}{2}\frac{1}{(\beta\Phi)''}[S_0'' + (\beta\Phi)''' S_1]g. \tag{4.24}$$

But, (4.12) yields $\frac{1}{2}e^{-\beta}e^{(\beta\Phi)'} = S_0$ and then we set

$$g(\beta) = \sqrt{\beta}\hat{g}(\beta) \tag{4.25}$$

with which (4.24) becomes

$$\left[S_0 - \frac{S_0'}{(\beta\Phi)''}\right]\frac{\hat{g}'(\beta)}{\hat{g}(\beta)} = T_0 - \frac{S_0}{2}[\beta + (\beta\Phi)''] + \frac{S_0'' + S_1(\beta\Phi)'''}{2(\beta\Phi)''}. \tag{4.26}$$

From (4.14) and (4.20) we obtain

$$S_0' = -\frac{2e^{-\beta}}{(2-e^{-\beta})^2} = -(\beta\Phi)''' \tag{4.27}$$

and

$$S_0 - \frac{S_0'}{(\beta\Phi)''} = \frac{1}{1-e^{-\beta}}. \tag{4.28}$$

Also, (4.7) yields

$$T_0(\beta) = \frac{4}{2-e^{-\beta}} + \frac{1}{(1-e^{-\beta})^2} - \frac{3}{1-e^{-\beta}}. \tag{4.29}$$

Combining (4.27) - (4.29) in (4.26) and multiplying the equation by $1 - e^{-\beta}$ leads to

$$\begin{aligned} \frac{\hat{g}'(\beta)}{\hat{g}(\beta)} &= -3 + \frac{1}{1-e^{-\beta}} + \frac{4(1-e^{-\beta})}{2-e^{-\beta}} - \frac{1-e^{-\beta}}{2-e^{-\beta}}\left(\beta + 2 - \frac{2}{2-e^{-\beta}}\right) \\ &+ \frac{1}{2}(1-e^{-\beta})e^{-\beta}\left[\frac{1}{(1-e^{-\beta})^2} - \frac{2}{(2-e^{-\beta})^2}\right] \end{aligned} \tag{4.30}$$

and this integrates to

$$\hat{g}(\beta) = (\text{const.})\, e^{-\beta^2/4} e^{-\beta/2}\left(\frac{1-e^{-\beta}}{2-e^{-\beta}}\right)^{3/2} \exp\left[\frac{\beta}{2}\log\left(1 - \frac{1}{2}e^{-\beta}\right) + \frac{\beta}{2}\phi(\beta)\right] \tag{4.31}$$

where const. is an arbitrary constant and $\phi(\beta)$ is defined by (4.15).

We next assume that (4.9) asymptotically matches to the expansion for $w > 1$; we write this condition symbolically as

$$e^{n\Phi(\beta)}\sqrt{\frac{\beta}{n}}\hat{g}(\beta)\bigg|_{\beta\to\infty} \sim w^{\binom{n}{2}} 2^{n-1} B_*(w)\bigg|_{w\to 1}. \tag{4.32}$$

The matching condition applies on some intermediate scale where $\beta = n(w-1) \to \infty$ but $w \to 1$. For $\beta \to \infty$ we have

$$\begin{aligned} \Phi(\beta) &= \frac{\beta}{2} + \log 2 - \frac{1}{\beta}\int_0^\infty \log\left(1 - \frac{1}{2}e^{-v}\right) dv + O\left(\frac{e^{-\beta}}{\beta}\right), \\ \beta\phi(\beta) &\to -\int_0^\infty \log\left(1 - \frac{1}{2}e^{-v}\right) dv \equiv d_0 = .58224\ldots \end{aligned} \tag{4.33}$$

and thus

$$\hat{g}(\beta) \sim (\text{const.})\, e^{-\beta^2/4} e^{-\beta/2} e^{d_0/2} 2^{-3/2},\ \beta \to \infty. \tag{4.34}$$

For $w \to 1$, we have

$$w^{\binom{n}{2}} = \left(1 + \frac{\beta}{n}\right)^{\frac{n(n-1)}{2}} = e^{n\beta/2} e^{-\beta^2/4} e^{-\beta/2}\left[1 + O\left(\frac{\beta^3}{n}\right)\right]. \tag{4.35}$$

Using (4.33) - (4.35) in (4.32) we see that the matching is possible provided that

$$B_*(w) \sim \frac{(\text{const.})}{\sqrt{2}}\sqrt{w-1}\, e^{d_0/2} \exp\left(\frac{d_0}{w-1}\right),\ w \to 1^+. \tag{4.36}$$

We also give numerical evidence for this behavior in Section 9. We will show in Section 5 that asymptotically matching the $\beta$-scale expansion (4.9), as $\beta \to 0^+$, to the central region expansion (that applies for $n^{3/2}(w-1) = a = O(1)$), as $a \to +\infty$, determines the constant in (4.31) and (4.36) as

$$\text{const.} = \frac{4}{\sqrt{\pi}}. \tag{4.37}$$

With (4.36) and (4.37) we obtain (2.25).

We comment that using the following *ansatz*

$$f(\beta; n) \sim e^{n\Phi(\beta)} n^{\Psi(\beta)} g(\beta) \tag{4.38}$$

in (4.3), which is slightly more general than (4.9), we would find that $\Phi(\beta)$ is as in (4.15), that

$$\Psi(\beta) = \Psi_0 \text{ is a constant,}$$

and then

$$g(\beta) = \beta^{-\Psi_0} \hat{g}(\beta)$$

where $\hat{g}$ is as in (4.31). Then matching to the $a$-scale result will show that $\Psi_0 = -\frac{1}{2}$ and fix the multiplicative constant as in (4.37). Using (4.38) and matching to the expansion for $w > 1$ would yield

$$B_*(w) \sim d_1 (w-1)^{-\Psi_0} \ \exp\left(\frac{d_0}{w-1}\right), \ w \to 1.$$

Finally, we use (4.9) to calculate $b(n,p)$ from the Cauchy formula (2.3). We write

$$w^{-p-1} = \left(1 + \frac{\beta}{n}\right)^{-p-1} = \ \exp\left[-\frac{p}{n}\beta + \frac{p}{2n^2}\beta^2 + O\left(\frac{p\beta^3}{n^3}\right)\right]. \tag{4.39}$$

Thus if we let $p, n \to \infty$ in such a way that $p/n^2$ is fixed, then $e^{n\Phi(\beta)}$ and $e^{-p\beta/n}$ grow at the same linear rate in $n$, for $\beta = O(1)$. Then (2.3) will have saddle point(s) where

$$\Phi'(\beta) = \frac{1}{2} + \phi'(\beta) = \Lambda \equiv \frac{p}{n^2},$$

or

$$\frac{1}{2} - \Lambda - \frac{1}{\beta}\log\left(1 - \frac{e^{-\beta}}{2}\right) = \frac{1}{\beta^2}\int_{-\log\left(1-\frac{1}{2}e^{-\beta}\right)}^{\log 2} \frac{\zeta}{e^{\zeta}-1} d\zeta. \tag{4.40}$$

This transcendental equation has a unique real solution $\beta = \beta_* = \beta_*(\Lambda)$ that satisfies

$$\beta_* \to \infty \text{ as } \Lambda \uparrow \frac{1}{2}, \ \beta_* \to 0^+ \text{ as } \Lambda \to 0^+.$$

In view of $p_{\max}(n)$ in (2.13) we need only consider $\Lambda \in \left(0, \frac{1}{2}\right)$. The steepest descent directions at $\beta = \beta_*$ are $\arg(\beta - \beta_*) = \pm\frac{\pi}{2}$ and then (4.39) and (4.40) lead to the estimate

$$b(n,p) \sim \frac{\sqrt{\beta_*}}{n^2}\hat{g}(\beta_*) e^{\Lambda\beta_*^2/2} \frac{1}{\sqrt{2\pi\phi''(\beta_*)}} \ \exp\left[n\left(\log 2 + \frac{\beta_*}{2} + \phi(\beta_*) - \Lambda\beta_*\right)\right]. \tag{4.41}$$

In view of (4.40) we have

$$\phi(\beta_*) = \left(\frac{1}{2} - \Lambda\right)\beta_* - \log\left(1 - \frac{1}{2}e^{-\beta_*}\right) \tag{4.42}$$

and (4.15) yields

$$\phi''(\beta_*) = \frac{1}{\beta_*}\left[1 - 2\Lambda - \frac{1}{2e^{\beta_*} - 1}\right]. \tag{4.43}$$

Combining (4.31) and (4.37) with (4.42) and (4.43), (4.41) becomes

$$\begin{aligned} b(n,p) \sim \frac{2^{n+1}}{n^2}\frac{\sqrt{2}}{\pi}\beta_* e^{-\beta_*/2}\left(\frac{1-e^{-\beta_*}}{2-e^{-\beta_*}}\right)^{3/2}\left[1 - 2\Lambda - \frac{1}{2e^{\beta_*}-1}\right]^{-1/2} \\ \times\ \exp\left\{n\left[\beta_*(1-2\Lambda) - \log\left(1 - \frac{1}{2}e^{-\beta_*}\right)\right]\right\}, \end{aligned} \tag{4.44}$$

which establishes (2.39).

Finally we discuss the asymptotic matching between (4.44), as $\Lambda \uparrow \frac{1}{2}$, and (3.10), as $L \to \infty$. We can solve (4.40) for $\beta_*$ asymptotically, as $\beta_*$ becomes large. We have

$$\phi(\beta) = \frac{d_0}{\beta} - \frac{1}{2\beta}e^{-\beta} + O_R(e^{-2\beta}),\ \beta \to \infty \tag{4.45}$$

and thus (4.40) becomes

$$\Lambda - \frac{1}{2} = -\frac{d_0}{\beta^2} + \left(\frac{1}{2\beta} + \frac{1}{2\beta^2}\right)e^{-\beta} + O_R(e^{-2\beta}). \tag{4.46}$$

Here again $O_R$ means roughly of the order, and neglects factors algebraic in $\beta$. Inverting the relation in (4.46) we find that

$$\beta_* \sim \sqrt{\frac{2d_0}{1-2\Lambda}}\left[1 - \frac{1}{4d_0}\left(\sqrt{\frac{2d_0}{1-2\Lambda}} + 1\right)\ \exp\left(-\sqrt{\frac{2d_0}{1-2\Lambda}}\right)\right],\ \Lambda \uparrow \frac{1}{2}. \tag{4.47}$$

Using (4.47) in (4.44) the right side becomes

$$\frac{\sqrt{2d_0}2^n}{\pi n^2(1-2\Lambda)}\ \exp\left[-\frac{1}{2}\sqrt{\frac{2d_0}{1-2\Lambda}} + n\sqrt{2d_0}\sqrt{1-2\Lambda}\right]. \tag{4.48}$$

We show that (4.48) agrees with (3.10) as $L \to \infty$. To expand (3.10) for $L$ large we argue that there is a saddle in the range $w \sim 1$ and approximate $B_*(w)$ by (3.9), which leads to the integral

$$\frac{2^{n-1}}{2\pi i}d_1\int_C w^{L-1}\sqrt{w-1}\ \exp\left(\frac{d_0}{w-1}\right)dw. \tag{4.49}$$

This has a saddle where

$$\frac{d}{dw}\left[L\log w + \frac{d_0}{w-1}\right] = 0 \Rightarrow$$

$$w = w_S \equiv 1 + \frac{d_0}{2L} + \sqrt{\frac{d_0}{L}}\sqrt{1 + \frac{d_0}{4L}}$$

and then the standard estimate for (4.49) is

$$\frac{2^{n-1}d_1}{\sqrt{2\pi}}\sqrt{w_S-1}\left[\frac{2d_0w_S^2}{(w_S-1)^3}-L\right]^{-1/2}\ \exp\left[L\log w_S+\frac{d_0}{w_S-1}\right]. \tag{4.50}$$

For $L\to\infty$ we can further simplify (4.50) using

$$w_S=1+\sqrt{\frac{d_0}{L}}+\frac{d_0}{2L}+O(L^{-3/2}) \tag{4.51}$$

and then we note that

$$L=\binom{n}{2}-p=\frac{n^2}{2}-p-\frac{n}{2}=n^2\left(\frac{1}{2}-\Lambda\right)-\frac{n}{2}. \tag{4.52}$$

With (4.51), (4.50) simplifies to

$$\frac{2^n}{L}\frac{d_1\sqrt{d_0}}{4\sqrt{\pi}}\ \exp\left(2\sqrt{Ld_0}-\frac{d_0}{2}\right)=\frac{2^n}{L}\frac{\sqrt{d_0}}{\sqrt{2\pi}}\ \exp(2\sqrt{Ld_0}). \tag{4.53}$$

But, if we use (4.52) in (4.53) and expand for $\Lambda\to\frac{1}{2}$, $n\to\infty$ with $n\left(\frac{1}{2}-\Lambda\right)\to\infty$, we obtain precisely (4.48). This verifies the asymptotic matching between the $L$-scale and $\Lambda$-scale results.

# 5 Central Region

In this section we analyze (2.1) for $w-1=O(n^{-3/2})$ and then obtain an expansion for $b(n,p)$ that applies for $p=O(n^{3/2})$. We define $a$ by

$$w-1=\frac{a}{n^{3/2}},\ -\infty<a<\infty. \tag{5.1}$$

At times we will need to separately consider the cases $a<0$ and $a>0$.

We set

$$\begin{aligned} B_n(w) &= 4^n\left[a_n4^{-n}+\frac{1}{n^{3/2}}\bar{C}_n(a)\right] \\ &= \frac{1}{n+1}\binom{2n}{n}+\frac{4^n}{n^{3/2}}\bar{C}_n(n^{3/2}(w-1)) \end{aligned} \tag{5.2}$$

and note that (2.4) yields

$$\bar{C}_n(0)=0. \tag{5.3}$$

Since $a_n$ satisfies the recurrence

$$a_{n+1}=\sum_{\ell=0}^{n}a_\ell a_{n-\ell},\ a_0=1, \tag{5.4}$$

using (5.2) in (2.1) leads to

$$\begin{aligned}
&\frac{4}{(n+1)^{3/2}}\bar{C}_{n+1}\left(\left(1+\frac{1}{n}\right)^{3/2}a\right) = a_{n+1}4^{-n}\left[\left(1+\frac{a}{n^{3/2}}\right)^{n}-1\right] \\
&+2\left(1+\frac{a}{n^{3/2}}\right)^{n}\sum_{\ell=0}^{n}\frac{a_\ell 4^{-\ell}}{(n-\ell)^{3/2}}\bar{C}_{n-\ell}\left(\left(1-\frac{\ell}{n}\right)^{3/2}a\right) \\
&+\left(1+\frac{a}{n^{3/2}}\right)^{n}\sum_{\ell=0}^{n}\left[\frac{1}{\ell(n-\ell)}\right]^{3/2}\bar{C}_\ell\left(\left(\frac{\ell}{n}\right)^{3/2}a\right)\bar{C}_{n-\ell}\left(\left(1-\frac{\ell}{n}\right)^{3/2}a\right). \qquad (5.5)
\end{aligned}$$

We can also write $B_n(w)$ as a Taylor series about $w=1$, setting

$$B_n(w) = \sum_{j=0}^{\infty}\frac{M_{j,n}}{j!}(w-1)^j. \qquad (5.6)$$

In view of (2.5) we have $M_{0,n} = a_n$, $M_{1,n} = b_n$, and $M_{2,n} = c_n$. By multiplying (2.1) by $w^{-n}$, differentiating $N$ times with respect to $w$ and setting $w=1$, we obtain

$$\sum_{i=0}^{N}\binom{N}{i}(-1)^{N-i}\frac{(n+N-i-1)!}{(n-1)!}M_{i,n+1} = \sum_{\ell=0}^{n}\sum_{j=0}^{N}\binom{N}{j}M_{j,\ell}M_{N-j,n-\ell}. \qquad (5.7)$$

We will obtain asymptotic approximations to $\bar{C}_n(a)$, or, equivalently, $M_{j,n}$. In subsection 5.1 we shall analyze (5.7), while in subsection 5.2 we analyze (5.5). Then in subsection 5.3 we will use the results to obtain the approximation to $b(n,p)$. We can also get analogous results by analyzing the functional equation (2.2) for the double transform; we discuss this further in Appendix A.

## 5.1 Moment equations

We consider (5.7). For $n\to\infty$ we write

$$M_{i,n} \equiv 4^n\tilde{M}_{i,n} = 4^n n^{\frac{3}{2}(i-1)}\left[m_i + \frac{1}{\sqrt{n}}\bar{m}_i + O(n^{-1})\right] \qquad (5.8)$$

and (2.9) shows that

$$m_0 = \frac{1}{\sqrt{\pi}},\ \bar{m}_0 = 0,\ m_1 = 1,\ \bar{m}_1 = -\frac{3}{\sqrt{\pi}},\ m_2 = \frac{10}{3\sqrt{\pi}},\ \bar{m}_2 = -\frac{13}{2}. \qquad (5.9)$$

Isolating the terms with $i=N$, $i=N-1$, and $j=0$, $j=N$ in (5.7) and using (5.8) we obtain

$$\begin{aligned}
\tilde{M}_{N,n+1} - nN\tilde{M}_{N-1,n+1} \quad &+ \quad \sum_{i=0}^{N-2}\binom{N}{i}(-1)^{N-i}\frac{(n+N-i-1)!}{(n-1)!}\tilde{M}_{i,n+1} \qquad (5.10)\\
&= \quad \frac{1}{2}\sum_{\ell=0}^{n}\tilde{M}_{0,\ell}\tilde{M}_{N,n-\ell} + \frac{1}{4}\sum_{j=1}^{N-1}\binom{N}{j}\sum_{\ell=0}^{n}\tilde{M}_{j,\ell}\tilde{M}_{N-j,n-\ell}.
\end{aligned}$$

Using (5.8) in the double sum in (5.10) we obtain

$$\begin{aligned}\sum_{j=1}^{N-1}\binom{N}{j}\sum_{\ell=0}^{n}\tilde{M}_{j,\ell}\tilde{M}_{N-j,n-\ell} &\sim \sum_{j=1}^{N-1}\binom{N}{j}\sum_{\ell=0}^{n} n^{\frac{3}{2}(j-1)}(n-\ell)^{\frac{3}{2}(N-j-1)}m_j m_{N-j} \\ &\sim \sum_{j=1}^{N-1} m_j m_{N-j}\left[\int_0^1 x^{\frac{3}{2}(j-1)}(1-x)^{\frac{3}{2}(N-j-1)}dx\right] n^{\frac{3}{2}N}n^{-2}.\end{aligned} \tag{5.11}$$

Here we approximated the inner sum by an integral via the Euler-MacLaurin formula. By definition (5.8), we have $\tilde{M}_{0,\ell} = a_\ell 4^{-\ell}$ and we write

$$\sum_{\ell=0}^{n} a_\ell 4^{-\ell}\tilde{M}_{N,n-\ell} = \sum_{\ell=0}^{n} a_\ell 4^{-\ell}\left[\tilde{M}_{N,n-\ell} - \tilde{M}_{N,n}\right] + \tilde{M}_{N,n}\sum_{\ell=0}^{n} a_\ell 4^{-\ell}. \tag{5.12}$$

Using the generating function $a(z)$ below (2.3) and the estimate in (2.9) we find that

$$\sum_{\ell=0}^{n} a_\ell 4^{-\ell} = 2 - \frac{2}{\sqrt{\pi n}} + O(n^{-3/2}). \tag{5.13}$$

Again by (5.8) and the Euler-MacLaurin formula we obtain

$$\sum_{\ell=0}^{n} a_\ell 4^{-\ell}\left[\tilde{M}_{N,n-\ell} - \tilde{M}_{N,n}\right] \sim n^{\frac{3}{2}N}n^{-2}m_N\frac{1}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[(1-x)^{\frac{3}{2}(N-1)} - 1\right]dx. \tag{5.14}$$

We also have

$$\begin{aligned}\tilde{M}_{N,n+1} - nN\tilde{M}_{N-1,n+1} &= (n+1)^{\frac{3}{2}(N-1)}\left[m_N + \frac{1}{\sqrt{n+1}}\bar{m}_N + O(n^{-1})\right] \\ &- nN(n+1)^{\frac{3}{2}(N-2)}\left[m_{N-1} + \frac{1}{\sqrt{n+1}}\bar{m}_{N-1} + O(n^{-1})\right] \\ &= m_N n^{\frac{3}{2}N}n^{-\frac{3}{2}} - Nm_{N-1}n^{\frac{3}{2}N}n^{-2} + O\left(n^{\frac{3}{2}N}n^{-\frac{5}{2}}\right).\end{aligned} \tag{5.15}$$

Using (5.11) - (5.15) in (5.10), dividing by $n^{\frac{3}{2}N}n^{-2}$, and letting $n \to \infty$ we obtain the limiting equation

$$\begin{aligned}-Nm_{N-1} &= -\frac{1}{\sqrt{\pi}}m_N + m_N\frac{1}{2\sqrt{\pi}}\int_0^1 \frac{(1-x)^{\frac{3}{2}(N-1)} - 1}{x^{3/2}}dx \\ &+ \frac{1}{4}\sum_{j=1}^{N-1}\binom{N}{j}m_j m_{N-j}\int_0^1 x^{\frac{3}{2}(j-1)}(1-x)^{\frac{3}{2}(N-j-1)}dx,\end{aligned} \tag{5.16}$$

which applies for $N \geq 1$. Setting

$$u_N = m_{N+1} \tag{5.17}$$

and replacing $N$ by $N+2$ in (5.16) we obtain

$$\begin{aligned} -(N+2)u_N &= -\frac{3}{2\sqrt{\pi}}(N+1)u_{N+1}\int_0^1 \frac{y^{\frac{3}{2}N+\frac{1}{2}}}{\sqrt{1-y}}dy \\ &+ \frac{1}{4}\sum_{i=0}^{N}\frac{(N+2)!}{(i+1)!(N+1-i)!}u_i u_{N-i}\int_0^1 x^{\frac{3}{2}i}(1-x)^{\frac{3}{2}(N-i)}dx. \end{aligned} \qquad (5.18)$$

Here we also used

$$\int_0^1 \frac{(1-x)^{\frac{3}{2}(N-1)}-1}{x^{3/2}}dx = 2-3\int_0^1 \frac{N-1}{\sqrt{x}}(1-x)^{\frac{1}{2}(3N-5)}dx, \qquad (5.19)$$

which follow by integration by parts.

Equation (5.18) can be recast by defining

$$D(y) = \sum_{j=0}^{\infty}\frac{u_j}{(j+1)!}y^{\frac{3}{2}j},\ y>0. \qquad (5.20)$$

Multiplying (5.18) by $y^{\frac{3}{2}N}/(N+2)!$ and summing over $n$, we obtain

$$-D(y) = -\frac{1}{y\sqrt{\pi}}\int_0^y \frac{D'(v)}{\sqrt{y-v}}dv + \frac{1}{4y}\int_0^y D(v)D(y-v)dv. \qquad (5.21)$$

In view of (5.6) we have the corresponding leading order approximation to $B_n(w)$ (for $n\to\infty$ with $a$ fixed) as

$$\begin{aligned} B_n(w) &\sim \sum_{j=0}^{\infty} 4^n n^{\frac{3}{2}(j-1)}\frac{m_j}{j!}(w-1)^j \\ &= \frac{4^n}{n^{3/2}}\sum_{j=0}^{\infty}\frac{m_j}{j!}a^j \\ &= \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} + a\sum_{j=0}^{\infty}\frac{m_{j+1}}{(j+1)!}a^j\right] \\ &= \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} + aD(a^{2/3})\right],\ a>0. \end{aligned} \qquad (5.22)$$

We can simplify (5.18) by setting

$$u_N = \frac{(N+1)!}{\Gamma\left(\frac{3}{2}N+1\right)}V_N, \quad V_0 = 1 \qquad (5.23)$$

and using the identities

$$\begin{aligned} \int_0^1 x^{\frac{3}{2}i}(1-x)^{\frac{3}{2}(N-i)}dx &= B\left(\frac{3}{2}i+1, \frac{3}{2}(N-i)+1\right) = \frac{\Gamma\left(\frac{3}{2}i+1\right)\Gamma\left(\frac{3}{2}(N-i)+1\right)}{\Gamma\left(\frac{3}{2}N+2\right)} \\ \int_0^1 \frac{1}{\sqrt{1-x}}x^{\frac{3}{2}N+\frac{1}{2}}dx &= B\left(\frac{3}{2}N+\frac{3}{2}, \frac{1}{2}\right) = \sqrt{\pi}\frac{\Gamma\left(\frac{3}{2}N+\frac{3}{2}\right)}{\Gamma\left(\frac{3}{2}N+2\right)}, \end{aligned} \qquad (5.24)$$

where $B(\cdot,\cdot)$ and $\Gamma(\cdot)$ are the Beta and Gamma functions, respectively. Using (5.23) and (5.24), (5.18) becomes

$$V_{N+1} = \left(\frac{3}{2}N + 1\right) V_N + \frac{1}{4}\sum_{i=0}^{N} V_i V_{N-i},\ V_0 = 1. \tag{5.25}$$

Since the $V_i$ are all positive, (5.25) shows that $V_{N+1} \geq \left(\frac{3}{2}N + 1\right) V_N$ and and thus

$$V_N \geq \left(\frac{3}{2}\right)^N \frac{\Gamma\left(N + \frac{2}{3}\right)}{\Gamma\left(\frac{2}{3}\right)},\ N \geq 0. \tag{5.26}$$

This estimate shows that the $V_N$ grow faster than exponentially, and thus (5.25) cannot be easily solved by generating functions, despite the non-linearity having the form of a convolution sum.

We can also rewrite (5.18) by introducing $\bar{D}(Y)$

$$\bar{D}(Y) = \sum_{j=1}^{\infty} Y^{\frac{3}{2}(j-1)} \frac{m_j}{j!}(-1)^j = \sum_{i=0}^{\infty} Y^{\frac{3}{2}i} \frac{u_i}{(i+1)!}(-1)^{i+1} \tag{5.27}$$

and then similarly to (5.22) we find that

$$B_n(w) \sim \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} - a\bar{D}((-a)^{2/3})\right], \quad -a = Y^{3/2} > 0. \tag{5.28}$$

This gives a leading order approximation to $B_n(w)$ for $n \to \infty$ and a fixed $a < 0$. We also note that

$$\sum_{L=0}^{\infty} \frac{m_L}{L!} a^L = m_0 + \sum_{L=1}^{\infty} \frac{m_L}{L!}(-a)^{L-1}(-a)(-1)^L = \frac{1}{\sqrt{\pi}} - a\bar{D}(Y). \tag{5.29}$$

Using (5.27) in (5.18) leads to the non-linear integral equation

$$Y\bar{D}(Y) = -\frac{1}{\sqrt{\pi}}\int_0^Y \frac{\bar{D}'(v)}{\sqrt{Y-v}}dv + \frac{1}{4}\int_0^Y \bar{D}(v)\bar{D}(Y-v)dv. \tag{5.30}$$

Note that this differs from (5.21) only by the sign of the left-hand side. However, (5.30) is susceptible to solution by a Laplace transform, whereas (5.21) is not.

Introducing

$$D_*(s) = \mathcal{L}\{\bar{D}(Y)\} = \int_0^{\infty} e^{-sY}\bar{D}(Y)dY, \tag{5.31}$$

where $\mathcal{L}$ is the Laplace transform operator, and noting that

$$\begin{aligned} \mathcal{L}\{Y\bar{D}(Y)\} &= -D_*'(s),\ \ \bar{D}(0) = -m_1 = -1, \\ \mathcal{L}\left\{\bar{D}'(Y) * \frac{1}{\sqrt{Y}}\right\} &= [sD_*(s) - \bar{D}(0)]\sqrt{\frac{\pi}{s}} \end{aligned}$$

(where $*$ denotes convolution) we obtain

$$-D_*'(s) = -\sqrt{s}D_*(s) + \frac{1}{4}D_*^2(s) - \frac{1}{\sqrt{s}}. \tag{5.32}$$

This Riccati equation is easily solved by setting

$$D_* = 2\sqrt{s} + U(s), \;\; U(s) = 4\frac{F'(s)}{F(s)}, \tag{5.33}$$

which leads to the Airy equation

$$F''(s) = \frac{s}{4}F(s). \tag{5.34}$$

The only solution with acceptable behavior as $s \to +\infty$ is given by

$$F(s) = \text{ (const.) Ai}(4^{-1/3}s) \tag{5.35}$$

where $\text{Ai}(\cdot)$ is the Airy function. Using (5.35) in (5.33) and inverting the Laplace transform, we obtain the integral representation

$$\bar{D}(Y) = \frac{1}{2\pi i}\int_{Br} e^{sY}\left[2\sqrt{s} + 4^{2/3}\frac{\text{Ai}'(4^{-1/3}s)}{\text{Ai}(4^{-1/3}s)}\right] ds, \quad Y = (-a)^{2/3} \tag{5.36}$$

which applies only for $a < 0$. Here $Br$ is any vertical contour in which $\Re(s) > r_0 = max\{z : \text{Ai}(z) = 0\}$. We can also write (5.36) as

$$\begin{aligned}
\bar{D}(Y) &= \frac{1}{2\pi i}\frac{d}{dY}\left\{\int_{Br} e^{sY}\left[\frac{2}{\sqrt{s}} + \frac{4^{2/3}}{s}\frac{\text{Ai}'(4^{-1/3}s)}{\text{Ai}(4^{-1/3}s)}\right] ds\right\} \\
&= \frac{d}{dY}\left[\frac{2}{\sqrt{\pi Y}} + 4^{2/3}\sum_{j=0}^{\infty}\frac{1}{r_j}e^{4^{1/3}r_jY}\right] \\
&= \frac{1}{\sqrt{\pi}a} + 4\sum_{j=0}^{\infty} e^{4^{1/3}r_jY}.
\end{aligned} \tag{5.37}$$

Here $0 > r_0 > r_1 > r_2 > \ldots$ where $r_j$ are the roots of the Airy function $\text{Ai}(\cdot)$, and we evaluated the integral as a residue series, by closing the $Br$ contour in the left half-plane.

By using (5.37) in (5.28) we obtain the leading order approximation

$$B_n(w) \sim \frac{4^{n+1}}{n^{3/2}}(-a)\sum_{j=0}^{\infty} e^{-4^{1/3}|r_j|Y}, \;\; Y = (-a)^{2/3} > 0. \tag{5.38}$$

We note that (5.38) is consistent with the fact that $B_n(1) = a_n \sim 4^n/(n^{3/2}\sqrt{\pi})$ as $n \to \infty$. It is well known that as $z \to -\infty$

$$\text{Ai}(z) \sim \frac{1}{\sqrt{\pi}}(-z)^{-1/4}\sin\left(\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}\right) \tag{5.39}$$

and thus the roots $r_j$ satisfy

$$|r_j| \sim \left(\frac{3j\pi}{2}\right)^{2/3}, \; j \to \infty. \tag{5.40}$$

Thus, as $Y \to 0^+$ ($w \to 1^-$), we can approximate the sum in (5.38) by the Euler-MacLaurin formula with

$$\sum_{j=0}^{\infty} e^{-4^{1/3}|r_j|Y} \sim \frac{1}{Y^{3/2}} \int_0^\infty \exp[-(3x\pi)^{2/3}]dx = \frac{-1}{4\sqrt{\pi}a}.$$

We next derive a representation valid for $a > 0$, and show that as $a \to \infty$ the $a$-scale expansion asymptotically matches to the $\beta$-scale result, as $\beta \to 0^+$. We can get a rough idea of the behavior of the central region approximation as $a = y^{3/2} \to +\infty$ by using the integral equation (5.21) and a WKB-type expansion. Let us assume that

$$D(y) \sim K(y)e^{\Psi(y)},\ y \to \infty \tag{5.41}$$

where $\Psi(y) \gg \log[K(y)]$. We use (5.41) to approximate the various terms in (5.21). Expecting that $\Psi'(y) > 0$ and $\Psi'(y) \to \infty$ as $y \to \infty$, we have

$$\begin{aligned}
&\int_0^y \frac{1}{\sqrt{v}} D'(y-v)dv \sim \int_0^y \frac{1}{\sqrt{v}}(K'e^\Psi + K\Psi' e^\Psi)(y-v)dv \\
&\sim \int_0^\infty e^{\Psi(y)} \frac{e^{-v\Psi'(y)}}{\sqrt{v}} \left[ \left(1 + \frac{1}{2}v^2\Psi''(y)\right)(\Psi'(y) - v\Psi''(y))(K(y) - vK'(y)) + K'(y)\right] dv \\
&\sim \sqrt{\Psi'(y)\pi} K(y)e^{\Psi(y)} + K(y)e^{\Psi(y)} \int_0^\infty \frac{e^{-v\Psi'}}{\sqrt{v}} \left[\frac{K'}{K} - v\Psi'' - v\Psi'\frac{K'}{K} + \frac{1}{2}v^2\Psi''\Psi'\right] dv \\
&= K(y)e^{\Psi(y)}\sqrt{\pi}\left[\sqrt{\Psi'} + \frac{K'}{K}\frac{1}{\sqrt{\Psi'}} - \frac{1}{2(\Psi')^{3/2}}\left(\Psi'' + \Psi'\frac{K'}{K}\right) + \frac{3}{8}\frac{\Psi''}{(\Psi')^{3/2}}\right].
\end{aligned} \tag{5.42}$$

The non-linear term in (5.21) we approximate as

$$\begin{aligned}
\int_0^y D(v)D(y-v)dv &= 2\int_0^{y/2} D(v)D(y-v)dv \\
&\sim 2\int_0^\infty D(v)e^{\Psi(y)}e^{-v\Psi'(y)}K(y)dv \\
&\sim 2D(0)K(y)e^{\Psi'(y)}\frac{1}{\Psi'(y)}.
\end{aligned} \tag{5.43}$$

Recalling that $D(0) = u_0 = m_1 = 1$, we use (5.42) and (5.43) in (5.21). After multiplying by $-y$ we obtain, to leading order,

$$yK(y)e^{\Psi(y)} \sim \sqrt{\Psi'(y)}K(y)e^{\Psi(y)}$$

so that

$$\Psi(y) = \frac{y^3}{3} = \frac{a^2}{3}. \tag{5.44}$$

At the next order, after cancelling the common factors $Ke^\Psi$, we are led to

$$0 = -\frac{1}{2\Psi'} + \frac{K'}{K}\frac{1}{\sqrt{\Psi'}} - \frac{1}{2(\Psi')^{3/2}}\left[\Psi'' + \Psi'\frac{K'}{K}\right] + \frac{3}{8}\frac{\Psi''}{(\Psi')^{3/2}} \tag{5.45}$$

and thus

$$\frac{K'}{K} = \frac{1}{4}\frac{\Psi''}{\Psi'} + \frac{1}{\sqrt{\Psi'}} = \frac{1}{2y} + \frac{1}{y} = \frac{3}{2y}$$

so that $K(y) = (\text{const.})' y^{3/2} = (\text{const.})' a$. We thus have formally obtained

$$\bar{C}_n(a) \sim (\text{const.})' a^2 e^{a^2/3}, \; a \to +\infty \tag{5.46}$$

for some constant (const.)$'$. We note that $\bar{C}_n(a) \sim aD(a^{2/3})$ dominates the first term in the right side of (5.2). In view of (5.2), (4.2) and (4.9) the asymptotic matching of the $a$- and $\beta$-scales requires that

$$\frac{4^n}{n^{3/2}}\bar{C}_n(a)\bigg|_{a\to\infty} \sim e^{n\Phi(\beta)} n^{-1/2} g(\beta)\bigg|_{\beta\to 0^+}. \tag{5.47}$$

From (4.15), (4.25), and (4.31) we have

$$\Phi(\beta) = \log 4 + \frac{\beta^2}{3} + O(\beta^3), \; g(\beta) \sim (\text{const.})\, \beta^2. \tag{5.48}$$

Since $\beta = a/\sqrt{n}$ we see that the matching is indeed possible provided that

$$\text{const.}' = \text{const.} \tag{5.49}$$

where the latter constant is the one that arose in Section 4. We note that our formal analysis suggests that the non-linear integral equation (5.21) may be approximated by a linear one for $y \to \infty$. The non-linear term does not affect the exponential growth rate $\Psi = y^3/3$, but it does affect the algebraic growth rate $K \propto y^{3/2}$.

To determine the remaining constant, we continue (5.37) into the range $a > 0$. Since $Y = (-a)^{2/3}$, this corresponds to $\arg(Y) = \pm\frac{2\pi}{3}$. We define

$$h(s) = \mathrm{Ai}(4^{-1/3}s). \tag{5.50}$$

By deforming the Bromwich contour in (5.37) to a piecewise linear contour that goes from $s = e^{-2\pi i/3}\infty$ to $s = 0$ and then from $s = 0$ to $s = e^{2\pi i/3}\infty = \omega\infty$, and parameterizing the two pieces, we are led to

$$\bar{D}(Y) - \frac{1}{\sqrt{\pi}a} = \frac{4}{2\pi i}\int_0^\infty \left[\omega\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega\tau Y} - \omega^2\frac{h'(\omega^2\tau)}{h(\omega^2\tau)}e^{\omega^2\tau Y}\right] d\tau. \tag{5.51}$$

Since $\Re(\omega)$, $\Re(\omega^2) < 0$, these integrals converge for $Y > 0$. We can also write the approximation to $B_n(w)$ for $w = 1 + O(n^{-3/2})$ as

$$B_n(w) \sim \frac{4^{n+1}}{\pi n^{3/2}}(-a)\int_0^\infty \Re\left[e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega\tau Y}\right] d\tau. \tag{5.52}$$

We shall use the Wronskian identity (cf. [1, 3])

$$\mathrm{Ai}(\omega z)\omega^2\mathrm{Ai}'(\omega^2 z) - \omega\mathrm{Ai}'(\omega z)\mathrm{Ai}(\omega^2 z) = \frac{i}{2\pi}$$

which in terms of $h(\cdot)$ becomes

$$e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}+e^{-\pi i/6}\frac{h'(\omega^2\tau)}{h(\omega^2\tau)}=-\frac{1}{2\pi 4^{1/3}}\frac{1}{h(\omega\tau)h(\omega^2\tau)}. \tag{5.53}$$

The integral

$$I_1=\int_0^\infty e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega\tau Y}d\tau,\ Y>0$$

may be continued analytically into the range $\arg(Y)\in\left(-\frac{\pi}{6},\frac{5\pi}{6}\right)$, while

$$I_2=\int_0^\infty e^{-\pi i/6}\frac{h'(\omega^2\tau)}{h(\omega^2\tau)}e^{\omega^2\tau Y}d\tau,\ Y>0$$

may be similarly continued into the range $\arg(Y)\in\left(-\frac{5\pi}{6},\frac{\pi}{6}\right)$ $\left(\text{in this range } \arg(\omega^2 Y)\in\left(-\frac{3\pi}{2},-\frac{\pi}{2}\right)\right)$. Using (5.53) we write $I_1=I_3+I_4$ where

$$\begin{aligned} I_3 &= -\frac{1}{2\pi 4^{1/3}}\int_0^\infty \frac{e^{\omega\tau Y}}{h(\omega\tau)h(\omega^2\tau)}d\tau && (5.54)\\ I_4 &= -\int_0^\infty e^{-\pi i/6}\frac{h'(\omega^2\tau)}{h(\omega^2\tau)}e^{\omega\tau Y}d\tau. && (5.55)\end{aligned}$$

Since $\mathrm{Ai}(\cdot)$ satisfies (cf. [1])

$$\mathrm{Ai}(z)\sim\frac{1}{2\sqrt{\pi}}z^{-1/4}\ \exp\left(-\frac{2}{3}z^{3/2}\right),\ |\arg z|<\pi \tag{5.56}$$

we see that both $h(\omega\tau)$ and $h(\omega^2\tau)$ grow faster than exponentially as $\tau\to+\infty$, and thus $I_3$ defines an entire function of $Y$. The integral $I_4$ may be viewed as defining an analytic function in the range $\arg(Y)\in\left(-\frac{\pi}{6},\frac{5\pi}{6}\right)$ or $\arg(Y)\in\left(-\frac{13\pi}{6},-\frac{7\pi}{6}\right)$. Now we let $\arg(Y)=-\frac{2\pi}{3}$ and set

$$Y=\omega^2 y=\omega^2 a^{2/3} \tag{5.57}$$

where $a$ is real and positive. We have thus shown that $I_1+I_2=I_2+I_3+I_4$. This continues the right side of (5.52) to $a>0$ and shows that

$$\begin{aligned}\frac{4}{\pi}(-a)\int_0^\infty \Re\left[e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega\tau Y}\right]d\tau &= \frac{a}{\pi^2 4^{1/3}}\int_0^\infty\frac{e^{\tau y}}{h(\omega\tau)h(\omega^2\tau)}d\tau && (5.58)\\ &- \frac{4a}{\pi}\int_0^\infty \Re\left[e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega^2\tau y}\right]d\tau.\end{aligned}$$

Next we estimate the right side as $a=y^{3/2}\to+\infty$. The second integral can be expanded by Watson's lemma, and is $O(a/y)=O(\sqrt{y})$. The first integral we evaluate by Laplace's method, first using (5.56) to approximate the integrand for $\tau\to+\infty$, with

$$h(\omega\tau)h(\omega^2\tau)=|h(\omega\tau)|^2\sim\frac{(4^{-1/3}\tau)^{-1/2}}{4\pi}\ \exp\left[\frac{4}{3}(4^{-1/3}\tau)^{3/2}\right].$$

Here we also used the reflection principle, since $\mathrm{Ai}(z)$ is real for real $z$. We thus have

$$\begin{aligned}\frac{4^{2/3}}{4\pi^2}\int_0^\infty \frac{e^{\tau y}}{|h(\omega\tau)|^2}d\tau &\sim \frac{4^{2/3}}{\pi}\int_0^\infty e^{\tau y}e^{-\frac{2}{3}\tau^{3/2}}2^{-1/3}\sqrt{\tau}d\tau \\ &\sim \frac{2}{\pi}\sqrt{y^2}\int_{-\infty}^\infty \exp\left[\frac{y^3}{3}-\frac{1}{4y}(\tau-y^2)^2\right]d\tau \\ &= \frac{4}{\sqrt{\pi}}y^{3/2}\exp\left(\frac{y^3}{3}\right). \qquad (5.59)\end{aligned}$$

Here we used the Laplace method to find that the major contribution to the integral came from $\tau = y^2$. Since this is asymptotically large, it is permissible to first approximate the integrand for $\tau$ large. We note that the right side of (5.59) is also the expansion of $D(y)$ as $y \to +\infty$. Thus, this calculation verifies that obtained formally by the WKB *ansatz* (5.41), and also determines the constant as

$$\text{const.} = \frac{4}{\sqrt{\pi}}. \qquad (5.60)$$

The leading term on the $a$-scale for $a > 0$ is thus given by $4^n/n^{3/2}$ times the right side of (5.58), and this is precisely the result in (2.29).

We next calculate the correction term $\bar{m}_i$ in (5.8). We shall show that this corresponds to $C^{(1)}(a)$ if $\bar{C}_n(a)$ has the expansion $\bar{C}_n(a) = C^{(0)}(a) + n^{-1/2}C^{(1)}(a) + O(n^{-1})$. We return to (5.10) and use (5.8). For $n \to \infty$ we have

$$\begin{aligned}&\sum_{i=0}^N \binom{N}{i}(-1)^{N-i}\frac{(n+N-i-1)!}{(n-1)!}\tilde{M}_{i,n+1} = \tilde{M}_{N,n+1} - nN\tilde{M}_{N-1,n+1} \\ &+\binom{N}{2}n(n+1)\tilde{M}_{N-2,n+1} + O\left(n^{\frac{3}{2}N-3}\right) \\ &= n^{\frac{3}{2}(N-1)}\left[m_N + \frac{1}{\sqrt{n}}(\bar{m}_N - Nm_{N-1})\right. \\ &\left.+\frac{1}{n}\left(\tilde{m}_N + \frac{3}{2}(N-1)m_N - N\bar{m}_{N-1} + \binom{N}{2}m_{N-2}\right) + O(n^{-3/2})\right] \qquad (5.61)\end{aligned}$$

where $\tilde{m}_N$ is the third term in the expansion in (5.8), i.e.,

$$\tilde{M}_{N,n} = n^{\frac{3}{2}(N-1)}\left[m_N + \frac{1}{\sqrt{n}}\bar{m}_N + \frac{1}{n}\tilde{m}_N + O(n^{-3/2})\right].$$

We next estimate more precisely (5.12):

$$\frac{1}{2}\sum_{\ell=0}^n a_\ell[\tilde{M}_{N,n-\ell} - \tilde{M}_{N,n}] + \frac{1}{2}\tilde{M}_{N,n}\sum_{\ell=0}^n a_\ell 4^{-\ell}. \qquad (5.62)$$

From (5.13) and (5.8) we obtain

$$\frac{1}{2}\tilde{M}_{N,n}\sum_{\ell=0}^n a_\ell 4^{-\ell} = n^{\frac{3}{2}(N-1)}\left[m_N + \frac{1}{\sqrt{n}}\bar{m}_N + \frac{1}{n}\tilde{m}_N - \frac{1}{\sqrt{\pi n}}m_N + O(n^{-3/2})\right].$$

To obtain the first term in (5.62) we use an intermediate limit, writing to sum as

$$\begin{aligned}
&\sum_{\ell=0}^{n} a_\ell 4^{-\ell}[\tilde{M}_{N,n-\ell} - \tilde{M}_{N,n}] \qquad (5.63)\\
&= \sum_{\ell=0}^{n} a_\ell 4^{-\ell} n^{\frac{3}{2}(N-1)} \left\{ \left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N-1)} \left[m_N + \frac{\bar{m}_N}{\sqrt{n-\ell}}\right] - m_N - \frac{\bar{m}_N}{\sqrt{n}} + O(n^{-1}) \right\}.
\end{aligned}$$

Now, by Euler-MacLaurin

$$\begin{aligned}
\sum_{\ell=0}^{n} a_\ell 4^{-\ell} n^{\frac{3}{2}(N-1)} \frac{\bar{m}_N}{\sqrt{n}} \left[\left(1-\frac{\ell}{n}\right)^{\frac{3}{2}N-2} - 1\right] \quad &\sim \quad \frac{1}{\sqrt{\pi}} n^{\frac{3}{2}N-\frac{5}{2}} \bar{m}_N \qquad (5.64)\\
&\times \quad \int_0^1 \frac{1}{x^{3/2}} \left[(1-x)^{\frac{3}{2}N-2} - 1\right] dx
\end{aligned}$$

where we used $a_\ell 4^{-\ell} \sim 1/[\sqrt{\pi}\ell^{3/2}]$. The remaining terms in (5.63) must be estimated more precisely, as we need both the $O(n^{\frac{3}{2}N-2})$ leading term (as in (5.14)) and the $O(n^{\frac{3}{2}N-\frac{5}{2}})$ correction. Breaking up the sum into two pieces we obtain

$$\begin{aligned}
\sum_{\ell=0}^{n} a_\ell 4^{-\ell} \left[\left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N-1)} - 1\right] \quad &= \quad \sum_{\ell=0}^{L-1} a_\ell 4^{-\ell} \left[\left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N-1)} - 1\right] \qquad (5.65)\\
&+ \quad \sum_{\ell=L}^{n} a_\ell 4^{-\ell} \left[\left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N-1)} - 1\right] \equiv S_1 + S_2.
\end{aligned}$$

Here $L$ satisfies the asymptotic bounds $L \to \infty$ but $L/n \to 0$ as $n \to \infty$. We use a binomial approximation in the first sum and approximate $a_\ell$ for $\ell \to \infty$ in the second. Thus

$$\begin{aligned}
S_1 \quad &= \quad \sum_{\ell=0}^{L-1} a_\ell 4^{-\ell} \left[-\frac{3}{2}(N-1)\frac{\ell}{n} + O\left(\frac{\ell^2}{n^2}\right)\right]\\
&= \quad -\frac{3}{2}\frac{N-1}{n} \sum_{\ell=0}^{L-1} \ell a_\ell 4^{-\ell} + O\left(\frac{L^{3/2}}{n^2}\right)\\
&= \quad -\frac{3}{2}\frac{N-1}{n} \sum_{\ell=0}^{L} \ell a_\ell 4^{-\ell} + O\left(\frac{1}{n\sqrt{L}}, \frac{L^{3/2}}{n^2}\right). \qquad (5.66)
\end{aligned}$$

Using (2.4) we have

$$\begin{aligned}\sum_{\ell=0}^{L} \ell a_\ell 4^{-\ell} &= \sum_{\ell=0}^{L} (\ell+1)4^{-\ell} a_\ell - \sum_{\ell=0}^{L} 4^{-\ell} a_\ell \\ &= \sum_{\ell=0}^{L} \binom{2\ell}{\ell} 4^{-\ell} - 2 + o_L(1) \\ &= \frac{2(L+1)}{4^{L+1}} \binom{2L+2}{L+1} - 2 + o_L(1) \\ &= \frac{2\sqrt{L}}{\sqrt{\pi}} - 2 + O_L\left(\frac{1}{\sqrt{L}}\right). \end{aligned} \tag{5.67}$$

To evaluate $S_2$ we use Euler-MacLaurin to find that

$$\begin{aligned} S_2 &= \sum_{\ell=L}^{n} \left[\frac{1}{\sqrt{\pi}}\ell^{-3/2} + O(\ell^{-5/2})\right]\left[\left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N+1)} - 1\right] \\ &= \frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{n}} \int_{L/n}^{1} \frac{1}{x^{3/2}} \left[(1-x)^{\frac{3}{2}(N-1)} - 1\right] dx + O\left(\frac{1}{n\sqrt{L}}\right) \\ &= \frac{1}{\sqrt{\pi}\sqrt{n}} \int_0^1 \frac{1}{x^{3/2}} \left[(1-x)^{\frac{3}{2}(N-1)} - 1\right] dx + \frac{3}{\sqrt{\pi}}(N-1)\frac{\sqrt{L}}{n} + O\left(\frac{1}{n\sqrt{L}}\right). \end{aligned} \tag{5.68}$$

Combining (5.66) - (5.68) and letting $L \to \infty$ in such a way that $L^{3/2}/n \to 0$ we have shown that

$$\frac{1}{2}(S_1+S_2) = \frac{1}{2\sqrt{\pi n}} \int_0^1 \frac{1}{x^{3/2}} \left[(1-x)^{\frac{3}{2}(N-1)} - 1\right] dx + \frac{3}{2}\frac{N-1}{n} + O(n^{-3/2}). \tag{5.69}$$

Next, we expand the last sum in (5.10) as

$$\begin{aligned} &\sum_{j=1}^{N-1} \binom{N}{j} \sum_{\ell=0}^{n} \tilde{M}_{j,\ell} \tilde{M}_{N-j,n-\ell} = \sum_{j=1}^{N-1} \binom{N}{j} n^{\frac{3}{2}N-3} \left[ m_j m_{N-j} \sum_{\ell=0}^{n} \left(\frac{\ell}{n}\right)^{\frac{3}{2}(j-1)} \left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N-j-1)} \right. \\ &\left. + \frac{2}{\sqrt{n}} m_j \bar{m}_{N-j} \sum_{\ell=0}^{n} \left(\frac{\ell}{n}\right)^{\frac{3}{2}(j-1)} \left(1-\frac{\ell}{n}\right)^{\frac{3}{2}(N-j)-2} + O(n^{-1}) \right] \\ &= n^{\frac{3}{2}N-2} \sum_{j=1}^{N-1} m_j m_{N-j} \binom{N}{j} \int_0^1 x^{\frac{3}{2}(j-1)} (1-x)^{\frac{3}{2}(N-j-1)} dx \\ &+ n^{\frac{3}{2}N-\frac{5}{2}} \sum_{j=1}^{N-1} 2 m_j \bar{m}_{N-j} \binom{N}{j} \int_0^1 x^{\frac{3}{2}(j-1)} (1-x)^{\frac{3}{2}(N-j)-2} dx + O(n^{\frac{3}{2}N-3}). \end{aligned} \tag{5.70}$$

Combining (5.61), (5.62) - (5.64), (5.69) and (5.70) we see that at $O\left(n^{\frac{3}{2}N-2}\right)$ we regain (5.16). At

$O\left(n^{\frac{3}{2}N-\frac{5}{2}}\right)$ we obtain, multiplying by 4,

$$\begin{aligned} 2N(N-1)m_{N-2} - 4N\bar{m}_{N-1} &= -\frac{2}{\sqrt{\pi}}(3N-4)\bar{m}_N \int_0^1 \frac{(1-x)^{\frac{3}{2}N-3}}{\sqrt{x}}dx \\ &+ 2\sum_{j=1}^{N-1}\binom{N}{j} m_j \bar{m}_{N-j} \int_0^1 x^{\frac{3}{2}(j-1)}(1-x)^{\frac{3}{2}(N-j)-2}dx. \end{aligned} \tag{5.71}$$

This holds for $N \geq 2$. Here we also rewrote the integral in (5.64) using integration by parts. Note that when $N = 2$ (5.71) becomes

$$4m_0 - 8\bar{m}_1 = -\frac{8}{\sqrt{\pi}}\bar{m}_2 + 8m_1\bar{m}_1$$

and this is consistent with (5.9).

We solve (5.71). Note that once $m_N$ is known this is a linear equation for $\bar{m}_N$. We set

$$\bar{D}_1(Y) = \sum_{L=1}^{\infty} \frac{\bar{m}_L}{L!}(-1)^L Y^{\frac{3}{2}(L-1)} \tag{5.72}$$

and recall the definition of $\bar{D}(Y)$ in (5.27). Since $m_0 = 1/\sqrt{\pi}$, we have

$$\begin{aligned} \sum_{N=2}^{\infty} 2N(N-1)m_{N-2}\frac{(-1)^N}{N!}Y^{\frac{3}{2}(N-1)} &= 2\sum_{L=0}^{\infty}\frac{m_L}{L!}(-1)^L Y^{\frac{3}{2}(L+1)} \\ &= 2\left[\frac{Y^{3/2}}{\sqrt{\pi}} + Y^3\bar{D}(Y)\right]. \end{aligned} \tag{5.73}$$

Also,

$$-4\sum_{N=2}^{\infty} N\bar{m}_{N-1}\frac{(-1)^N}{N!}Y^{\frac{3}{2}(N-1)} = 4\sum_{L=1}^{\infty}\frac{\bar{m}_L}{L!}(-1)^L Y^{\frac{3}{2}L} = 4Y^{3/2}\bar{D}_1(Y), \tag{5.74}$$

$$\begin{aligned} &\sum_{N=2}^{\infty}(-1)^N 2Y^{\frac{3}{2}(N-1)}\sum_{j=1}^{N-1}\frac{m_j\bar{m}_{N-j}}{j!(N-j)!}\int_0^1 x^{\frac{3}{2}(j-1)}(1-x)^{\frac{3}{2}(N-j-1)}\frac{dx}{\sqrt{1-x}} \\ &\qquad = 2Y^{3/2}\int_0^1 \bar{D}(Yx)\frac{\bar{D}_1(Y-Yx)}{\sqrt{1-x}}dx, \end{aligned} \tag{5.75}$$

and

$$
\begin{aligned}
&-\frac{2}{\sqrt{\pi}}\sum_{N=2}^{\infty}(3N-4)\frac{(-1)^N}{N!}Y^{\frac{3}{2}(N-1)}\bar{m}_N\int_0^1\frac{(1-x)^{\frac{3}{2}N-3}}{\sqrt{x}}dx\\
&=-\frac{2}{\sqrt{\pi}}\int_0^1\frac{1}{\sqrt{x(1-x)}}\frac{1}{1-x}\sum_{N=2}^{\infty}\frac{3(N-1)-1}{N!}(-1)^N\bar{m}_N(Y-Yx)^{\frac{3}{2}(N-1)}dx\\
&=-\frac{2}{\sqrt{\pi}}\int_0^1\frac{1}{\sqrt{x}(1-x)^{3/2}}[2(Y-Yx)\bar{D}_1(Y-Yx)-\bar{D}_1(Y-Yx)-\bar{m}_1]dx\\
&=-\frac{4}{\sqrt{\pi}}\int_0^1\frac{Y}{\sqrt{x(1-x)}}\bar{D}_1'(Y-Yx)dx+\frac{2}{\sqrt{\pi}}\int_0^1\frac{\bar{D}_1(Y-Yx)-\bar{D}_1(0)}{\sqrt{x}(1-x)^{3/2}}dx\\
&=-\frac{4Y}{\sqrt{\pi}}\int_0^Y\frac{\bar{D}_1'(Y-v)}{\sqrt{v}\sqrt{Y-v}}dv+\frac{2Y}{\sqrt{\pi}}\int_0^Y\frac{\bar{D}_1(v)-\bar{D}_1(0)}{\sqrt{Y-v}\,v^{3/2}}dv.
\end{aligned}
\tag{5.76}
$$

Combining (5.72) - (5.76) in (5.71) the equation transforms to (after dividing by $Y$)

$$
\begin{aligned}
&\frac{2Y^{1/2}}{\sqrt{\pi}}+2Y^2\bar{D}(Y)+4Y^{1/2}\bar{D}_1(Y)\\
&=2\int_0^Y\frac{\bar{D}_1(v)}{\sqrt{v}}\bar{D}(Y-v)dv-\frac{4}{\sqrt{\pi}}\int_0^Y\frac{1}{\sqrt{Y-v}}\frac{d}{dv}\left[\frac{\bar{D}_1(v)-\bar{D}_1(0)}{\sqrt{v}}\right]dv.
\end{aligned}
\tag{5.77}
$$

To solve (5.77) we define

$$
\mathcal{E}_*(s)=\int_0^\infty e^{-sY}\frac{\bar{D}_1(Y)}{\sqrt{Y}}dY \tag{5.78}
$$

and note that

$$
\begin{aligned}
\int_0^\infty e^{-sY}\int_0^Y\frac{1}{\sqrt{Y-v}}\frac{d}{dv}\left[\frac{\bar{D}_1(v)-\bar{D}_1(0)}{\sqrt{v}}\right]dv\,dY &= \sqrt{\frac{\pi}{s}}\int_0^\infty e^{-sY}\frac{d}{dY}\left[\frac{\bar{D}_1(Y)-\bar{D}_1(0)}{\sqrt{Y}}\right]dY\\
&= \sqrt{\pi s}\int_0^\infty e^{-sY}\left[\frac{\bar{D}_1(Y)-\bar{D}_1(0)}{\sqrt{Y}}\right]dY\\
&= \sqrt{\pi s}\,\mathcal{E}_*(s)-\pi\bar{D}_1(0).
\end{aligned}
\tag{5.79}
$$

Since $\bar{D}_1(0)=-\bar{m}_1=3/\sqrt{\pi}$, the Laplace transform of (5.77) becomes

$$
\frac{1}{s^{3/2}}+2D_*''(s)-4\mathcal{E}_*'(s)=[2D_*(s)-4\sqrt{s}]\mathcal{E}_*(s)+12. \tag{5.80}
$$

Recalling that $h(s)=\mathrm{Ai}(4^{-1/3}s)$ we write (5.33) as

$$
D_*(s)=2\sqrt{s}+4\frac{h'(s)}{h(s)} \tag{5.81}
$$

whose derivatives yield

$$
D_*'(s)=\frac{1}{\sqrt{s}}+s-4\left(\frac{h'(s)}{h(s)}\right)^2 \tag{5.82}
$$

and

$$D''_*(s) = -\frac{1}{2s^{3/2}} + 1 - 2s\frac{h'(s)}{h(s)} + 8\left(\frac{h'(s)}{h(s)}\right)^3, \tag{5.83}$$

where we used the fact that $h''(s) = \frac{1}{4}sh(s)$. Using (5.81) - (5.83) in (5.80) we obtain, after some rearrangement,

$$\mathcal{E}'_*(s) + 2\frac{h'(s)}{h(s)}\mathcal{E}_*(s) = -\frac{5}{2} - s\frac{h'(s)}{h(s)} + 4\left(\frac{h'(s)}{h(s)}\right)^3 \tag{5.84}$$

or

$$\frac{d}{ds}[h^2(s)\mathcal{E}_*(s)] = -\frac{5}{2}h^2(s) - sh(s)h'(s) + 4\frac{(h'(s))^3}{h(s)}. \tag{5.85}$$

The general solution of (5.85) may be written as

$$\mathcal{E}_*(s) = \frac{c_*}{h^2(s)} + \frac{1}{h^2(s)}\int_s^\infty \left[\frac{5}{2}h^2(v) + vh(v)h'(v) - 4\frac{(h'(v)^3}{h(v)}\right] dv. \tag{5.86}$$

But, $h(s)$ decays exponentially as $s \to \infty$ and so $h^{-2}$ grows. Thus, if $\mathcal{E}_*(s)$ is to be a proper Laplace transform, we must set $c_* = 0$. Use of integration by parts and the fact that $h(s)$ satisfies (5.34) leads to

$$\begin{aligned} \int^s vh(v)h'(v)dv = 4\int^s h'(v)h''(v)dv &= 2(h'(s))^2 \\ \int^s h^2(v)dv &= sh^2(s) - 2\int^s vh(v)h'(v)dv = sh^2(s) - 4(h'(s))^2 \end{aligned}$$

and thus (5.86) simplifies to

$$\mathcal{E}_*(s) = -\frac{5}{2}s + 8\left(\frac{h'(s)}{h(s)}\right)^2 - \frac{4}{h^2(s)}\int_s^\infty \frac{(h'(v))^3}{h(v)}dv. \tag{5.87}$$

We have thus established (2.26), since for $a < 0$

$$\begin{aligned} B_n(w) &= \sum_{j=0}^\infty M_{j,n}\frac{(w-1)^j}{j!} \\ &= \frac{4^n}{n^{3/2}}\left[\sum_{j=0}^\infty m_j\frac{a^j}{j!} + \frac{1}{\sqrt{n}}\sum_{j=0}^\infty \bar{m}_j\frac{a^j}{j!} + O(n^{-1})\right] \\ &= \frac{4^n}{n^{3/2}}\left[m_0 + \sum_{j=1}^\infty m_j\frac{a^j}{j!} + \frac{1}{\sqrt{n}}\sum_{j=1}^\infty \bar{m}_j\frac{a^j}{j!} + O(n^{-1})\right] \\ &= \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} - a\bar{D}((-a)^{2/3}) + \frac{1}{\sqrt{n}}(-a)\bar{D}_1(Y) + O(n^{-1})\right] \\ &= \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} + Y^{3/2}\bar{D}(Y) + \frac{1}{\sqrt{n}}Y^{3/2}\bar{D}_1(Y) + O(n^{-1})\right]. \end{aligned} \tag{5.88}$$

In order to asymptotically match the $a$-scale result to that valid for $w-1=O(n^{-1})$ with $w<1$, we will need the behavior of $\bar{D}_1(Y)$ as $Y=(-a)^{2/3}\to\infty$. Thus we must locate the singularity of $\mathcal{E}_*(s)$ with the largest real part, or an appropriate saddle point. The Laplace transform of the leading term $\bar{D}(Y)$ had simple poles at $4^{1/3}r_j$ where $r_j$ were the negative roots of the Airy function, with the pole at $4^{1/3}r_0$ determining the asymptotic behavior as $Y\to\infty$ ($a\to-\infty$) (cf. (5.37)). The dominant singularity in $\mathcal{E}_*(s)$ is also at $4^{1/3}, r_0$. but (5.87) shows that the structure is more complicated, as there is a double pole combined with a logarithmic branch point at $s=s_0\equiv 4^{1/3}r_0$.

We expand (5.87) near the dominant singularity at $s_0$. Integrating by parts we obtain

$$\begin{aligned}\int_s^\infty \frac{(h'(v))^3}{h(v)}dv &= \int_s^\infty (h'(v))^2 d[\log h(v)] \qquad (5.89)\\ &= -(h'(s))^2[\log h(s)] - \frac{1}{2}\int_s^\infty vh'(v)h(v)[\log h(v)]dv\end{aligned}$$

and

$$\int_s^\infty vh(v)h'(v)[\log h(v)]dv = \int_s^\infty vh(v)[\log h(v)]dh(v) \qquad (5.90)$$
$$= \quad -sh^2(s)[\log h(s)] - \int_s^\infty \{h^2(v)[\log h(v)] + vh(v)h'(v)\log h(v) + vh(v)h'(v)\}dv$$

which may be rewritten as

$$\begin{aligned}2\int_s^\infty vh(v)h'(v)[\log h(v)]dv = &- sh^2(s)[\log h(s)] + 2(h'(s))^2\\ &- \int_s^\infty h^2(v)[\log h(v)]dv. \qquad (5.91)\end{aligned}$$

Using (5.89) - (5.91) in (5.87) we obtain

$$\begin{aligned}\mathcal{E}_*(s) = &- \frac{5}{2}s + 10\left(\frac{h'(s)}{h(s)}\right)^2 + 4\left(\frac{h'(s)}{h(s)}\right)^2\log h(s)\\ &- s\log h(s) - \frac{1}{h^2(s)}\int_s^\infty h^2(v)[\log h(v)]dv. \qquad (5.92)\end{aligned}$$

We let $\tau=s-s_0$ and expand (5.92) about $\tau=0$. We have $h(s_0)=h''(s_0)=0$ and

$$\begin{aligned}\int_s^\infty h^2(v)[\log h(v)]dv &= \int_{s_0}^\infty h^2(v)[\log h(v)]dv - \int_{s_0}^s h^2(v)[\log h(v)]dv\\ &= \int_{s_0}^\infty h^2(v)[\log h(v)]dv + O(\tau^3\log\tau),\end{aligned}$$

and then (5.92) yields

$$\mathcal{E}_*(s) = \frac{4}{\tau^2}\log(h'(s_0)\tau) + \frac{10}{\tau^2} - \frac{\int_{s_0}^\infty h^2(v)[\log h(v)]dv}{(h'(s_0))^2\tau^2} + O(\log\tau),\ \tau\to 0. \qquad (5.93)$$

Here

$$h'(s_0) = 4^{-1/3}\mathrm{Ai}'(4^{-1/3}s_0) = 4^{-1/3}\mathrm{Ai}'(r_0). \tag{5.94}$$

Defining

$$\kappa \equiv 10 + 4\log(h'(s_0)) - \frac{1}{(h'(s_0))^2}\int_{s_0}^{\infty} h^2(v)[\log h(v)]dv \tag{5.95}$$

we have

$$\begin{aligned}\bar{D}_1(Y) &= \frac{\sqrt{Y}}{2\pi i}\int_{Br} e^{sY}\mathcal{E}_*(s)ds \\ &= \sqrt{Y}\,\exp(-|r_0|4^{1/3}Y)\frac{1}{2\pi i}\int_{Br_+} e^{\tau Y}\left[\frac{4\log\tau}{\tau^2} + \frac{\kappa}{\tau^2} + O(\log\tau)\right]d\tau \end{aligned}\tag{5.96}$$

where $\Re(\tau) > 0$ on $Br_+$. Evaluating explicitly the integrals in (5.96) shows that

$$\bar{D}_1(Y) = Y^{3/2}\,\exp(-|r_0|4^{1/3}Y)\left[-4\log Y + \kappa + 4 - 4\gamma_E + O\left(\frac{\log Y}{Y}\right)\right],\ Y \to \infty \tag{5.97}$$

where $\gamma_E$ is the Euler constant.

Combining (5.97) with (5.38) we have shown that as $Y \to \infty$ the two term approximation on the $a$-scale behaves as

$$B_n(w) = \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} - a\bar{D}(Y) - \frac{a}{\sqrt{n}}\bar{D}_1(Y) + O(n^{-1})\right] \tag{5.98}$$

$$\sim \frac{4^n}{n^{3/2}}\,\exp\left[-4^{1/3}|r_0|(-a)^{2/3}\right]\left\{-4a + \frac{a^2}{\sqrt{n}}\left(-\frac{8}{3}\log(-a) + \kappa + 4 - 4\gamma_E\right)\right\}. \tag{5.99}$$

This expression will be used in Section 6.

To summarize, we have used the moment equations (5.7) to get the two-term approximation (5.98) to $B_n(w)$, valid for $a = n^{3/2}(w-1) = O(1)$ and $a < 0$. For $a > 0$ we must use (5.51) and (5.58) with $y = a^{2/3}$.

## 5.2 Analysis of the basic recurrence

We consider (2.1), which on the $a$-scale becomes (5.5). Using (2.9) we obtain

$$a_{n+1}4^{-n}\left[\left(1 + \frac{a}{n^{3/2}}\right)^n - 1\right] = \frac{4}{\sqrt{\pi}}\frac{a}{n^2} + \frac{2}{\sqrt{\pi}}\frac{a^2}{n^{5/2}} + O(n^{-3}). \tag{5.100}$$

We expand $\bar{C}_n(a)$ in (5.2) as

$$\bar{C}_n(a) = C^{(0)}(a) + \frac{1}{\sqrt{n}}C^{(1)}(a) + O(n^{-1}) \tag{5.101}$$

and then estimate the various terms in (5.5). We note that $C^{(j)}(0) = 0$ for all $j \geq 0$, in view of (5.3).

By (5.101) and the Euler-MacLaurin formula we have

$$
\begin{aligned}
&\sum_{\ell=0}^{n} \left[\frac{1}{\ell(n-\ell)}\right]^{3/2} \bar{C}_\ell \left(\left(\frac{\ell}{n}\right)^{3/2} a\right) \bar{C}_{n-\ell}\left(\left(1-\frac{\ell}{n}\right)^{3/2} a\right) \\
&= \frac{1}{n^2}\int_0^1 \frac{C^{(0)}(x^{3/2}a)C^{(0)}((1-x)^{3/2}a)}{[x(1-x)]^{3/2}} dx \\
&+ \frac{2}{n^{5/2}}\int_0^1 \frac{C^{(0)}(x^{3/2}a)C^{(1)}((1-x)^{3/2}a)}{x^{3/2}(1-x)^2} dx + O(n^{-3}). \qquad (5.102)
\end{aligned}
$$

Note that the integrals converge since $C^{(j)}(0) = 0$, and that the boundary effects near $x = 0$ and $x = 1$ do not appear until the third ($O(n^{-3})$) term in the expansion.

Next we write the first sum in the right side of (5.5) as

$$
\begin{aligned}
&\sum_{\ell=0}^{n} a_\ell 4^{-\ell} (n-\ell)^{-3/2} \bar{C}_{n-\ell}\left(\left(1-\frac{\ell}{n}\right)^{3/2} a\right) \\
&= \sum_{\ell=0}^{n} a_\ell 4^{-\ell} \frac{1}{n^{3/2}} \left[\left(1-\frac{\ell}{n}\right)^{-3/2} C^{(0)}\left(\left(1-\frac{\ell}{n}\right)^{3/2} a\right) - C^{(0)}(a)\right] \\
&+ \sum_{\ell=0}^{n} a_\ell 4^{-\ell} \frac{1}{n^2} \left[\left(1-\frac{\ell}{n}\right)^{2} C^{(1)}\left(\left(1-\frac{\ell}{n}\right)^{3/2} a\right) - C^{(1)}(a)\right] \\
&+ \frac{C^{(0)}(a)}{n^{3/2}} \sum_{\ell=0}^{n} a_\ell 4^{-\ell} + \frac{C^{(1)}(a)}{n^2} \sum_{\ell=0}^{n} a_\ell 4^{-\ell} + \frac{C^{(2)}(a)}{n^{5/2}} \sum_{\ell=0}^{n} a_\ell 4^{-\ell} + O(n^{-3}). \qquad (5.103)
\end{aligned}
$$

The last three terms in (5.103) may be estimated using (5.13). By Euler-Maclaurin we have

$$
\begin{aligned}
&\sum_{\ell=0}^{n} a_\ell 4^{-\ell} \left[\left(1-\frac{\ell}{n}\right)^{-2} C^{(1)}\left(\left(1-\frac{\ell}{n}\right)^{3/2} a\right) - C^{(1)}(a)\right] \\
&\sim \frac{1}{\sqrt{n}}\frac{1}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}} \left[\frac{C^{(1)}((1-x)^{3/2}a)}{(1-x)^2} - C^{(1)}(a)\right] dx \qquad (5.104)
\end{aligned}
$$

where we approximated $a_\ell$ for $\ell \to \infty$. An argument completely analogous to (5.65) shows that

$$
\begin{aligned}
&\sum_{\ell=0}^{n} a_\ell 4^{-\ell} \left[ \left(1-\frac{\ell}{n}\right)^{-3/2} C^{(0)} \left( \left(1-\frac{\ell}{n}\right)^{3/2} a \right) C^{(0)}(a) \right] \qquad (5.105)\\
&= \sum_{\ell=0}^{L} a_\ell 4^{-\ell} \left\{ \frac{3}{2}\frac{\ell}{n}[C^{(0)}(a) - aC^{(0)\prime}(a)] + O\left(\frac{\ell^2}{n^2}\right) \right\}\\
&+ \sum_{\ell=L+1}^{n} \frac{1}{\sqrt{\pi}}[\ell^{-3/2} + O(\ell^{-5/2})] \left[ \left(1-\frac{\ell}{n}\right)^{-3/2} C^{(0)} \left( \left(1-\frac{\ell}{n}\right)^{3/2} a \right) - C^{(0)}(a) \right]\\
&= [C^{(0)}(a) - aC^{(0)\prime}(a)] \frac{3}{2n} \sum_{\ell=0}^{L} \ell a_\ell 4^{-\ell}\\
&+ \frac{1}{\sqrt{\pi n}} \int_{L/n}^{1} x^{-3/2} [(1-x)^{-3/2} C^{(0)}((1-x)^{3/2} a) - C^{(0)}(a)] dx + O\left( \frac{L^{3/2}}{n^2}, \frac{1}{n\sqrt{L}} \right)\\
&\sim \frac{1}{\sqrt{\pi n}} \int_{0}^{1} x^{-3/2} [(1-x)^{-3/2} C^{(0)}((1-x)^{3/2} a) - C^{(0)}(a)] dx\\
&+ [C^{(0)}(a) - aC^{(0)\prime}(a)] \frac{3}{2n} \left[ \sum_{\ell=0}^{L} \ell a_\ell 4^{-\ell} - 2\sqrt{\frac{L}{\pi}} \right]\\
&\sim \frac{1}{\sqrt{\pi n}} \int_{0}^{1} x^{-3/2} [(1-x)^{-3/2} C^{(0)}((1-x)^{3/2} a) - C^{(0)}(a)] dx - \frac{3}{n}[C^{(0)}(a) - aC^{(0)\prime}(a)].
\end{aligned}
$$

Here we also used (5.67). Multiplying (5.5) by

$$
w^{-n} = \left(1 + \frac{a}{n^{3/2}}\right)^{-n} = 1 - \frac{a}{\sqrt{n}} + \frac{a^2}{2n} + O(n^{-3/2})
$$

and using (5.100) yields

$$
\begin{aligned}
&\frac{4}{(n+1)^{3/2}} \left(1 + \frac{a}{n^{3/2}}\right)^{-n} \bar{C}_{n+1} \left( \left(1+\frac{1}{n}\right)^{3/2} a \right) - a_{n+1} 4^{-n} \left[ 1 - \left(1 + \frac{a}{n^{3/2}}\right)^{-n} \right]\\
&= \frac{4}{n^{3/2}} \left[ 1 - \frac{3}{2n} + O(n^{-2}) \right] \left[ 1 - \frac{a}{\sqrt{n}} + \frac{a^2}{2n} + O(n^{-3/2}) \right] \qquad (5.106)\\
&\times \left[ C^{(0)}(a) + \frac{1}{\sqrt{n}} C^{(1)}(a) + \frac{1}{n} \left( C^{(2)}(a) + \frac{3}{2} a C^{(0)\prime}(a) \right) \right]\\
&- \frac{4}{\sqrt{\pi} n^{3/2}} \left[ \frac{a}{\sqrt{n}} - \frac{a^2}{2n} + O(n^{-3/2}) \right]\\
&= \frac{4}{n^{3/2}} C^{(0)}(a) + \frac{4}{n^2} \left[ -aC^{(0)}(a) - \frac{a}{\sqrt{\pi}} + C^{(1)}(a) \right]\\
&+ \frac{4}{n^{5/2}} \left[ -aC^{(1)}(a) + \frac{a^2}{2\sqrt{\pi}} + \frac{a^2}{2} C^{(0)}(a) + \frac{3}{2} a C^{(0)\prime}(a) - \frac{3}{2} C^{(0)}(a) + C^{(2)}(a) \right] + O(n^{-3}).
\end{aligned}
$$

Comparing (5.106) with (5.102) + 2 × (5.103), using (5.104), (5.105) and (5.13), we find at $O(n^{-2})$

$$-4aC^{(0)}(a) - \frac{4}{\sqrt{\pi}}a = -\frac{4}{\sqrt{\pi}}C^{(0)}(a) + \frac{2}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{C^{(0)}((1-x)^{3/2}a)}{(1-x)^{3/2}} - C^{(0)}(a)\right]dx + \int_0^1 \frac{C^{(0)}(x^{3/2}a)C^{(0)}((1-x)^{3/2}a)}{x^{3/2}(1-x)^{3/2}}dx \tag{5.107}$$

and at $O(n^{-5/2})$

$$-4aC^{(1)}(a) + \frac{2}{\sqrt{\pi}}a^2 + 2a^2C^{(0)}(a) = -\frac{4}{\sqrt{\pi}}C^{(1)}(a) + \frac{2}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{C^{(1)}((1-x)^{3/2}a)}{(1-x)^2} - C^{(1)}(a)\right]dx = 2\int_0^1 \frac{C^{(0)}(x^{3/2}a)C^{(1)}((1-x)^{3/2}a)}{x^{3/2}(1-x)^2}dx. \tag{5.108}$$

Thus, the leading term in (5.101) satisfies the non-linear integral equation (5.107) and the correction term satisfies the linear equation (5.108).

We can analyze these equations for $a < 0$ by setting

$$C^{(0)}(a) = (-a)\bar{D}((-a)^{2/3}) = Y^{3/2}\bar{D}(Y). \tag{5.109}$$

Then (5.107) becomes

$$4a^2\bar{D}(Y) - \frac{4}{\sqrt{\pi}}a = \frac{4}{\sqrt{\pi}}a\bar{D}(Y) + a^2\int_0^1 \bar{D}(Yx)\bar{D}(Y-Yx)dx - \frac{2a}{\sqrt{\pi}}\int_0^1 \frac{\bar{D}(Y-Yx) - \bar{D}(Y)}{x^{3/2}}dx. \tag{5.110}$$

Integrating by parts and using $\bar{D}(0) = -1$ we obtain

$$\int_0^1 x^{-3/2}[\bar{D}(Y-Yx) - \bar{D}(Y)]dx = 2 + 2\bar{D}(Y) - 2Y\int_0^1 x^{-1/2}\bar{D}'(Y-Yx)dx. \tag{5.111}$$

Using (5.111) in (5.110), dividing by $4a^2/Y$ and noting that $Y^{3/2} = -a$, we see that (5.110) is the same as (5.30). Similarly, we set

$$C^{(1)}(a) = (-a)\bar{D}_1((a)^{2/3}) = Y^{3/2}\bar{D}_1(Y) \tag{5.112}$$

in (5.108). Using also (5.109) this yields

$$4a^2\bar{D}_1(Y) + \frac{2}{\sqrt{\pi}}a^2 - \frac{4}{\sqrt{\pi}}a\bar{D}_1(Y) - 2a^3\bar{D}(Y) = -\frac{2a}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{\bar{D}_1(Y-Yx)}{\sqrt{1-x}} - \bar{D}_1(Y)\right]dx + 2a^2\int_0^1 \frac{\bar{D}(Yx)\bar{D}_1(Y-Yx)}{\sqrt{1-x}}dx. \tag{5.113}$$

Dividing (5.113) by $a^2/\sqrt{Y} = Y^{5/2}$ leads to

$$4\sqrt{Y}\bar{D}_1(Y) + \frac{2}{\sqrt{\pi}}\sqrt{Y} + \frac{4}{\sqrt{\pi Y}}\bar{D}_1(Y) + 2Y^2\bar{D}(Y) \tag{5.114}$$
$$= 2\int_0^Y \frac{\bar{D}_1(v)}{\sqrt{v}}\bar{D}(Y-v)dv + \frac{2}{\sqrt{\pi Y}}\int_0^Y \frac{1}{v^{3/2}}\left[\bar{D}_1(Y-v)\sqrt{\frac{Y}{Y-v}} - \bar{D}_1(Y)\right]dv.$$

But the last term in (5.114) is

$$\frac{2}{\sqrt{\pi Y}}\int_0^Y \frac{1}{v^{3/2}}\left[\bar{D}_1(Y-v)\sqrt{\frac{Y}{Y-v}} - \bar{D}_1(0)\sqrt{\frac{Y}{Y-v}} - \bar{D}_1(Y) + \bar{D}_1(0)\left(\sqrt{\frac{Y}{Y-v}}-1\right) + \bar{D}_1(0)\right]dv$$
$$= \frac{4}{\sqrt{\pi Y}}\bar{D}_1(0) - \frac{4}{\sqrt{\pi Y}}\int_0^Y \left\{\bar{D}_1(0) - \bar{D}_1(Y) + \sqrt{\frac{Y}{Y-v}}[\bar{D}_1(Y-v) - \bar{D}_1(0)]\right\}d(v^{-1/2})$$
$$= \frac{4}{\sqrt{\pi Y}}\bar{D}_1(Y) + \frac{4}{\sqrt{\pi Y}}\int_0^Y \frac{1}{\sqrt{v}}\frac{d}{dv}\left[\frac{\bar{D}_1(Y-v) - \bar{D}_1(0)}{\sqrt{Y-v}}\right]dv \tag{5.115}$$

where we integrated by parts and used

$$\int_0^1 \left(\frac{1}{\sqrt{1-x}} - 1\right)\frac{dx}{x^{3/2}} = 2.$$

Using (5.115) in (5.114) we regain equation (5.77). We have thus shown the equivalence of the moment method and the perturbation expansion of this subsection. Note also that (5.107) and (5.108) apply also for $a \geq 0$.

## 5.3 Transform inversion

We now use the results of subsections 5.1 and 5.2 to obtain an approximation to $b(n,p)$.

We use (2.3) and evaluate the Cauchy integral asymptotically. We assume that

$$\Omega \equiv pn^{-3/2} = O(1),\ 0 < \Omega < \infty \tag{5.116}$$

and with (5.1) we have

$$w^{-p-1}dw = \left(1 + \frac{a}{n^{3/2}}\right)^{-1-\Omega n^{3/2}}\frac{da}{n^{3/2}} = \frac{e^{-a\Omega}}{n^{3/2}}[1 + O(n^{-3/2})]da. \tag{5.117}$$

Thus the scale $w - 1 = O(n^{-3/2})$ in the transform space translates to (5.116) in the $(n,p)$ space, and (5.117) and (5.98) lead to

$$\begin{aligned} b(n,p) &= \frac{4^n}{n^3}\frac{1}{2\pi i}\int_{Br_a}\left[\frac{1}{\sqrt{\pi}} + C^{(0)}(a) + \frac{1}{\sqrt{n}}C^{(1)}(a) + O(n^{-1})\right]e^{-a\Omega}da \\ &\sim \frac{4^n}{n^3}\frac{1}{2\pi i}\int_{Br_a}\left[\frac{1}{\sqrt{\pi}} - a\bar{D}(Y)\right]e^{-a\Omega}da \\ &= \frac{4^n}{n^3}(-a)\frac{1}{(2\pi i)^2}\int_{Br_a}\left[\frac{d}{dY}\int_{Br}\frac{4}{s}\frac{h'(s)}{h(s)}e^{sY}ds\right]e^{-a\Omega}da. \end{aligned} \tag{5.118}$$

Here we take $\Re(a) < 0$ on the vertical contour $Br_a$, since we used the expression (5.37) for $\bar{D}(Y)$.

To evaluate the double integral in (5.118) we note that as $-a$ goes from $-\infty i$ to $+\infty i$, $Y = (-a)^{2/3}$ goes from $\infty e^{-\pi i/3}$ to $\infty e^{\pi i/3}$. We let

$$Y = Z^2$$

and the contour $Br_a$ is mapped to $C_1$, that goes from $\infty e^{-\pi i/6}$ to $\infty e^{\pi i/6}$. Then (5.118) becomes

$$\begin{aligned} b(n,p) &\sim \frac{4^n}{n^3}\frac{1}{(2\pi i)^2}\int_{C_1} e^{\Omega Z^3}\frac{3}{2}Z^4\left[\frac{d}{dZ}\int_{Br}\frac{4}{s}\frac{h'(s)}{h(s)}e^{sZ^2}ds\right]dZ \\ &= \frac{4^n}{n^3}\frac{-1}{(2\pi i)^2}\int_{\infty\omega^2}^{\infty\omega} e^{-\Omega Z^3}12Z^5\left[\int_{Br}\frac{h'(s)}{h(s)}e^{sZ^2}ds\right]dZ, \end{aligned} \tag{5.119}$$

where the last integral was obtained by changing $Z$ to $-Z$, and $\omega = e^{2\pi i/3}$. We next write $d(e^{-\Omega Z^3}) = -3\Omega Z^2 e^{-\Omega Z^3}$ and do some integrating by parts, noting that

$$Z^4 e^{-\Omega Z^3} = \frac{1}{9\Omega^2}\frac{\partial^2}{\partial Z^2}(e^{-\Omega Z^3}) + \frac{2Z}{3\Omega}e^{-\Omega Z^3}.$$

Then (5.119) becomes

$$\begin{aligned} b(n,p) &\sim -\frac{4^n}{n^3}\frac{1}{(2\pi i)^2}\left\{\int_{\infty\omega^2}^{\infty\omega} e^{-\Omega Z^3}\left[\int_{Br}\frac{h'(s)}{h(s)}(Ze^{sZ^2})_{ZZ}\frac{1}{9\Omega^2}ds\right.\right. \qquad (5.120) \\ &+ \left.\left.\int_{Br}\frac{h'(s)}{h(s)}e^{sZ^2}\frac{8Z^2}{\Omega}ds\right]dZ\right\} \\ &= -\frac{4^n}{n^3}\frac{1}{(2\pi i)^2}\int_{\infty\omega^2}^{\infty\omega}\int_{Br} e^{-\Omega Z^3}e^{sZ^2}\frac{h'(s)}{h(s)}\left[\frac{40}{3}\frac{Zs}{\Omega^2} + \frac{16}{9}\frac{s^2(1+2sZ^2)}{\Omega^3}\right]ds\,dZ. \end{aligned}$$

We furthermore set

$$Z = \frac{s}{3\Omega} + W,\ W = \frac{\zeta}{(3\Omega)^{1/3}} \tag{5.121}$$

and note that the Airy function has the integral representation

$$\frac{1}{2\pi i}\int_{\infty\omega^2}^{\infty\omega} e^{\zeta U}e^{-\zeta^3/3}d\zeta = \mathrm{Ai}(U). \tag{5.122}$$

Differentiating (5.122) with respect to $U$ corresponds to multiplying the integrand by $\zeta$. Hence (5.120) becomes

$$\begin{aligned} b(n,p) &\sim -\frac{4^n}{n^3}\frac{1}{2\pi i}\int_{Br}\frac{h'(s)}{h(s)}\ \exp\left(\frac{2s^3}{27\Omega^2}\right)\frac{1}{(3\Omega)^{1/3}}\left\{\frac{40}{3\Omega^2}\frac{s}{(3\Omega)^{1/3}}\mathrm{Ai}'(U(s))\right. \\ &+ \left(\frac{40}{9\Omega^3} + \frac{16}{9\Omega^3}\right)s^2\mathrm{Ai}(U(s)) + \frac{32}{9\Omega^3}\frac{s^3}{(3\Omega)^{2/3}}\mathrm{Ai}''(U(s)) + \frac{64}{27\Omega^4}\frac{s^4}{(3\Omega)^{1/3}}\mathrm{Ai}'(U(s)) \\ &+ \left.\frac{32}{81\Omega^5}s^5\mathrm{Ai}(U(s))\right\}ds \end{aligned} \tag{5.123}$$

where

$$U(s) = \frac{s^2}{(3\Omega)^{4/3}}. \tag{5.124}$$

Using $\mathrm{Ai}''(U(s)) = U(s)\mathrm{Ai}(U(s))$ to simplify (5.123) and noting that the integrand has singularities at $s = 4^{1/3}r_j$, we evaluate the integral as a residue series and the result is precisely the series in (2.40). It is also possible to obtain and $O(n^{-1/2})$ correction to (5.123) or (2.40), by using (5.87) and (5.96) to compute $C^{(1)}(a)$ in (5.118) and evaluating the integral in a manner similar to that above.

Finally we discuss the expansion of the right side of (2.40) (or (5.123)) for $\Omega \to 0$ and $\Omega \to \infty$. For $\Omega \to 0$ the term in the sum with $j = 0$ is exponentially larger than those with $j \geq 1$, and thus we get

$$\begin{aligned} &\frac{4^n}{n^3}\frac{1}{(3\Omega)^{1/3}} \exp\left(-\frac{8}{27}\frac{|r_0|^3}{\Omega^2}\right)\frac{64}{81}\frac{|r_0|^5}{\Omega^5}4^{5/3}\mathrm{Ai}\left(\frac{r_0^2 4^{2/3}}{3^{4/3}\Omega^{4/3}}\right) \\ &\sim \frac{4^n}{n^3} \exp\left(-\frac{16}{27}\frac{|r_0|^3}{\Omega^2}\right)\frac{2^9}{3^4}\frac{|r_0|^{9/2}}{\sqrt{\pi}\Omega^5}. \end{aligned} \tag{5.125}$$

Here we also used the asymptotic expansion (5.56) of the Airy function. We will use expression (5.125) in Section 6, when we discuss the asymptotic matching of the results for $p = O(n^{3/2})$ and $p = O(n^{4/3})$. The limit $\Omega \to 0$ corresponds to the left tail of the Airy distribution.

The right tail corresponds to $\Omega \to \infty$ and this limit is hard to evaluate from the sum in (2.40), as there is a lot of cancelation. However, we have already shown the asymptotic matching between the expansions for $w = 1 + O(n^{-1})$ ($\beta$-scale) and $w = 1 + O(n^{-3/2})$ ($a$-scale). This matching implies also the matching in $(n, p)$ space of the expansion for $p = O(n^2)$ ($\Lambda$-scale) and for $p = O(n^{3/2})$, where (5.123) applies. Thus the behavior of (5.123) as $\Omega \to \infty$ is the same as that of (4.44) as $\Lambda = p/n^2 = \Omega/\sqrt{n} \to 0$. The latter is easily obtained from (4.40) and (4.44).

As $\Lambda \to 0$, we have $\beta_* \to 0$ and from (4.16) we obtain

$$\phi(\beta) = \log 2 - \frac{1}{2}\beta + \frac{1}{3}\beta^2 + O(\beta^3),\ \beta \to 0. \tag{5.126}$$

It follows from (4.40) $\left(\text{since } \phi'(\beta) + \frac{1}{2} \sim \frac{2}{3}\beta\right)$ that

$$\beta_* = \beta_*(\Lambda) \sim \frac{3}{2}\Lambda,\ \Lambda \to 0. \tag{5.127}$$

Then

$$1 - 2\Lambda - \frac{1}{2e^{\beta_*} - 1} \sim \Lambda \tag{5.128}$$

and

$$\beta_*(1 - 2\Lambda) + \log 2 - \log(2 - e^{-\beta_*}) \sim \log 2 - \frac{3}{4}\Lambda^2. \tag{5.129}$$

Using (5.127) - (5.129) in (4.44) the right side becomes

$$\begin{aligned} \frac{2^{n+1}\sqrt{2}}{n^2\pi}\frac{\beta_*^{5/2}}{\sqrt{\Lambda}} \exp\left[n\log 2 - \frac{3}{4}n\Lambda^2\right] &\sim \frac{9\sqrt{3}}{2\pi}\frac{4^n}{n^2}\Lambda^2 \exp\left(-\frac{3}{4}n\Lambda^2\right) \\ &= \frac{9\sqrt{3}}{2\pi}\frac{4^n}{n^3}\Omega^2 \exp\left(-\frac{3}{4}\Omega^2\right),\ \Omega = \frac{p}{n^{3/2}}. \end{aligned} \tag{5.130}$$

This is the behavior of (5.123) or (2.40) as $\Omega \to \infty$. We have thus completed analysis of the scales $w - 1 = O(n^{-3/2})$ and $p = O(n^{3/2})$.

# 6 Left Region

We consider the scale $w - 1 = O(n^{-1})$ with $w < 1$, defining $\gamma$ by

$$w = 1 - \frac{\gamma}{n}, \ \gamma > 0. \tag{6.1}$$

We return to (2.1), and break up the sum over $[0, n]$ into three pieces: $[0, L] \cup [L+1, n-L-1] \cup [n-L-1, n]$. Furthermore, we set

$$B_n(w) = \frac{4^n}{n} e^{\nu_0 n^{1/3} \gamma^{2/3}} e^{\nu_1 \gamma \log n} F(\gamma; n). \tag{6.2}$$

Here $\nu_0$ and $\nu_1$ are constants that will be determined shortly. The form in (6.2) is suggested by the behavior of the $a$-scale result as $a \to -\infty$ (cf. (5.8) with $-a = \gamma\sqrt{n}$). Using (6.2) in (2.1) leads to

$$\begin{aligned}
&\frac{4^{n+1}}{n+1} \exp\left[(n+1)^{1/3}\nu_0 \left[\left(1+\frac{1}{n}\right)\gamma\right]^{2/3} + \nu_1 \gamma \log(n+1)\right] F\left(\gamma\left(1+\frac{1}{n}\right); n+1\right) \\
&= 2\left(1-\frac{\gamma}{n}\right)^n \sum_{\ell=0}^{L} B_\ell \left(1-\frac{\gamma}{n}\right) \frac{4^{n-\ell}}{n-\ell} \exp\left\{\nu_0 (n-\ell)^{1/3} \left[\gamma\left(1-\frac{\ell}{n}\right)\right]^{2/3}\right\} \\
&\times \exp\left[\nu_1 \gamma \left(1-\frac{\ell}{n}\right)\log(n-\ell)\right] F\left(\left(1-\frac{\ell}{n}\right)\gamma; n-\ell\right) \\
&+ \left(1-\frac{\gamma}{n}\right)^n \sum_{\ell=L+1}^{n-L-1} \frac{4^n}{\ell(n-\ell)} \exp\left\{\nu_0 \ell^{1/3}\left(\gamma\frac{\ell}{n}\right)^{2/3} + \nu_0 (n-\ell)^{1/3}\left[\gamma\left(1-\frac{\ell}{n}\right)\right]^{2/3}\right\} \\
&\times \exp\left\{\nu_1 \gamma \left[\frac{\ell}{n}\log \ell + \left(1-\frac{\ell}{n}\right)\log(n-\ell)\right]\right\} F\left(\frac{\ell}{n}\gamma; \ell\right) F\left(\left(1-\frac{\ell}{n}\right)\gamma; n-\ell\right).
\end{aligned} \tag{6.3}$$

Here we used the symmetry of the sum in (2.1). Also, in the second sum in (6.3) we used (6.2) to approximate both $B_\ell(w)$ and $B_{n-\ell}(w)$, while in the first sum we only used (6.2) to approximate the latter. This careful treatment of "edge effects" is necessary to derive the limiting equation.

We note that

$$\ell^{1/3}\left(\frac{\ell}{n}\right)^{2/3} + (n-\ell)^{1/3}\left(1-\frac{\ell}{n}\right)^{2/3} = n^{1/3}$$

and

$$\frac{\ell}{n}\log \ell + \left(1-\frac{\ell}{n}\right)\log(n-\ell) = \log n + H\left(\frac{\ell}{n}\right)$$

where

$$H(x) = x \log x + (1-x)\log(1-x). \tag{6.4}$$

Now we assume that

$$F(\gamma;n) \to F_0(\gamma),\ n \to \infty. \tag{6.5}$$

Then choosing $L$ so that $L \to \infty$ and $L/n \to 0$ as $n \to \infty$, the first sum in (6.3) is asymptotic to

$$\begin{aligned}
&\sum_{\ell=0}^{L} B_\ell \left(1 - \frac{\gamma}{n}\right) \frac{4^{n-\ell}}{n-\ell}\ \exp(\nu_0 n^{1/3}\gamma^{2/3}) n^{\nu_1\gamma} F_0(\gamma) \\
&\sim \frac{4^n}{n}\ \exp(\nu_0 n^{1/3}\gamma^{2/3}) n^{\nu_1\gamma} F_0(\gamma) \sum_{\ell=0}^{\infty} B_\ell(1) 4^{-\ell} \\
&= \frac{4^n}{n}\ \exp(\nu_0 n^{1/3}\gamma^{2/3}) n^{\nu_1\gamma} 2F_0(\gamma).
\end{aligned} \tag{6.6}$$

The second sum in (6.3) we approximate by the Euler-MacLaurin formula to obtain:

$$\frac{4^n}{n}\ \exp(\nu_0 n^{1/3}\gamma^{2/3}) n^{\nu_1\gamma} \int_0^1 e^{\nu_1\gamma H(x)} \frac{F_0(\gamma x)F_0(\gamma - \gamma x)}{x(1-x)} dx. \tag{6.7}$$

Using (6.6) and (6.7) in (6.3) we obtain the limiting equation

$$4(e^\gamma - 1)F_0(\gamma) = \int_0^1 \frac{e^{\nu_1\gamma H(x)}}{x(1-x)} F_0(\gamma x) F_0(\gamma - \gamma x) dx. \tag{6.8}$$

Here we also used $(1-\gamma/n)^n \sim e^{-\gamma}$.

This is a non-linear integral equation that is somewhat similar to one that arises in the study of the limiting distribution of the number of comparisons in the Quicksort algorithm [10, 13, 20, 21]. Note also that we must have $F_0(0) = 0$ for the integral to converge, but this will be shown to follow by asymptotic matching to the $a$-scale. Setting

$$F_0(\gamma) = 4\gamma F_1(\gamma), \tag{6.9}$$

equation (6.8) simplifies to

$$\frac{e^\gamma - 1}{\gamma} F_1(\gamma) = \int_0^1 F_1(\gamma x) F_1(\gamma - \gamma x) e^{\nu_1\gamma H(x)} dx. \tag{6.10}$$

Setting $\gamma = 0$ in (6.10) we conclude that

$$F_1(0) = 1. \tag{6.11}$$

We examine more closely the behavior of (6.10) as $\gamma \to 0$. We expand $F_1(\cdot)$ as

$$F_1(\gamma) = 1 + \alpha_0 \gamma \log \gamma + \alpha_1 \gamma + o(\gamma).$$

Using (6.11) in (6.10) and expanding for $\gamma$ small we obtain at $O(\gamma)$

$$\begin{aligned}
\alpha_1 + \frac{1}{2} = (\alpha_0 + \nu_1)\left[\int_0^1 H(x)dx\right] + \alpha_1 \quad &= \quad 2(\alpha_0 + \nu_1)\left[\int_0^1 x \log x\ dx\right] + \alpha_1 \\
&= \quad -\frac{1}{2}(\alpha_0 + \nu_1) + \alpha_1
\end{aligned} \tag{6.12}$$

so that

$$\alpha_0 + \nu_1 = -1 \tag{6.13}$$

and $\alpha_1$ is arbitrary. We can compute the higher-order terms in (6.11) in terms of $\alpha_1$, but the latter cannot be determined from (6.10) alone. This shows that the solution of (6.10) is not unique.

To uniquely determine $F_1(\gamma)$ we use asymptotic matching between the $a$ and $\gamma$ scales. For $\gamma \to 0^+$ we obtain, from (6.2), (6.5), (6.9) and (6.11),

$$\frac{4^{n+1}}{n}\gamma[1 + \alpha_0\gamma\log\gamma + \alpha_1\gamma + o(\gamma)](1 + \nu_1\gamma\log n)\ \exp(\nu_0 n^{1/3}\gamma^{2/3}). \tag{6.14}$$

The above should agree with the $a$-scale approximation as $a \to -\infty$, which we obtained in (5.99). Noting that $\gamma = -a/\sqrt{n}$ and comparing (5.99) to (6.14) we conclude that

$$\nu_0 = 4^{1/3} r_0 = -4^{1/3}|r_0| \tag{6.15}$$

and

$$\nu_1\gamma\log n + \alpha_0\gamma\log\gamma + \alpha_1\gamma = \gamma\left[-\frac{2}{3}\log(\gamma\sqrt{n}) + \frac{\kappa}{4} + 1 - \gamma_E\right]. \tag{6.16}$$

where $\gamma_E$ is the Euler constant. From (6.16) we must have

$$\nu_1 = -\frac{1}{3},\ \alpha_0 = -\frac{2}{3} \tag{6.17}$$

and

$$\alpha_1 = \frac{\kappa}{4} + 1 - \gamma_E = 2.9692\ldots. \tag{6.18}$$

We note that (6.17) is consistent with (6.13). With (6.18) the solution to (6.10) is unique and may be computed, e.g., in the form of the series (6.11).

To further analyze (6.10) (with $\nu_1 = -1/3$) we let

$$F_1(\gamma) = \exp\left(\frac{1}{3}\gamma\log\gamma\right)F_2(\gamma) \tag{6.19}$$

and obtain

$$(e^{\gamma} - 1)F_2(\gamma) = \int_0^{\gamma} F_2(v)F_2(\gamma - v)dv. \tag{6.20}$$

Introducing the Laplace transform

$$F(\theta) = \int_0^{\infty} e^{-\gamma\theta}F_2(\gamma)d\gamma \tag{6.21}$$

we obtain from (6.20) the discrete logistic equation (in reverse time):

$$F(\theta - 1) = F(\theta) + F^2(\theta). \tag{6.22}$$

The expansion (6.11) translates to

$$F_2(\gamma) = 1 - \gamma \log \gamma + \alpha_1 \gamma + o(\gamma), \quad \gamma \to 0^+$$

and hence

$$F(\theta) = \frac{1}{\theta} + \frac{\log \theta}{\theta^2} + \frac{\alpha_1 + \gamma_E - 1}{\theta^2} + o(\theta^{-2}), \quad \theta \to +\infty. \tag{6.23}$$

We can refine (6.23) to (using (6.22))

$$F(\theta) = \frac{1}{\theta} + \frac{\log \theta + \alpha_*}{\theta^2} + \frac{G(\theta)}{\theta^3} + O_R(\theta^{-4}) \tag{6.24}$$

where $\alpha_* = \alpha_1 + \gamma_E - 1 = 2.5464\ldots$ and

$$G(\theta) = (\log \theta + \alpha_*)^2 - \log \theta - \alpha_* + \frac{3}{2}. \tag{6.25}$$

We shall see in Section 9 that knowing some of the higher-order terms in expansion (6.23) will allow for more efficient numerical calculation of $F(\theta)$, and hence $F_2(\gamma)$.

While we cannot solve (6.20) or (6.22) explicitly, we can guess the behavior of $F_2(\gamma)$ as $\gamma \to +\infty$, which will be needed for asymptotic matching purposes. This will correspond to knowing $F(\theta)$ as $\theta \to -\infty$. Let us assume that $F_2(\gamma)$ in (6.20) behaves as

$$F_2(\gamma) \sim e^{k_1 \gamma \log \gamma} e^{k_2 \gamma} \gamma^{k_3} k_4, \quad \gamma \to \infty. \tag{6.26}$$

We use (6.26) in (6.20) and evaluate the integral by Laplace's method, to yield

$$\begin{aligned} \frac{e^\gamma}{\gamma} e^{k_1 \gamma \log \gamma} e^{k_2 \gamma} \gamma^{k_3} k_4 \quad &\sim \quad \int_0^1 \gamma^{2k_3} x^{k_3} (1-x)^{k_3} k_4^2 e^{k_2 \gamma} e^{k_1 [\gamma \log \gamma + \gamma H(x)]} dx \\ &\sim \quad \gamma^{2k_3} \left(\frac{1}{4}\right)^{k_3} k_4^2 e^{k_2 \gamma} e^{k_1 \gamma \log \gamma} e^{k_1 \gamma H\left(\frac{1}{2}\right)} \int_{-\infty}^{\infty} e^{2\gamma k_1 \left(x - \frac{1}{2}\right)^2} dx \\ &= \quad \gamma^{2k_3} \left(\frac{1}{4}\right)^{k_3} k_4^2 e^{k_2 \gamma} e^{k_1 \gamma \log \gamma} e^{-k_1 \gamma \log 2} \sqrt{\frac{\pi}{-2\gamma k_1}}. \end{aligned} \tag{6.27}$$

We conclude that

$$k_1 = -\frac{1}{\log 2}, \quad k_3 = -\frac{1}{2}, \quad k_4 = \frac{1}{\sqrt{2\pi \log 2}}. \tag{6.28}$$

The constant $k_2$ remains arbitrary, though it is uniquely determined once we know $\alpha_1$ in (6.11). The numerical studies in Section 9 will show that $k_2 \approx 3.696$. To summarize, we have obtained

$$F_1(\gamma) \sim \frac{1}{\sqrt{2\pi \log 2}} \frac{e^{k_2 \gamma}}{\sqrt{\gamma}} \ \exp\left[\left(\frac{1}{3} - \frac{1}{\log 2}\right) \gamma \log \gamma\right] \quad \gamma \to \infty. \tag{6.29}$$

Using (6.29) and (6.9) in (6.2) yields the behavior of the $\gamma$-scale result for $\gamma \to \infty$, and this will be used in Section 7 to asymptotically match to an approximation valid for $0 < w < 1$ and $n \to \infty$.

We also note that for $\theta \equiv -\tau \to -\infty$, (6.22) is consistent with an asymptotic expansion of the form

$$F(\theta) \sim A^{2^\tau}\left[1 + \sum_{L=1}^{\infty} b_L A^{-L2^\tau}\right], \qquad \tau = -\theta \to \infty. \tag{6.30}$$

Using (6.30) in (6.22) leads to the recurrence

$$b_{L-1} + \sum_{i=0}^{L} b_i b_{L-i} = \begin{cases} 0, & L \text{ odd} \\ b_{L/2}, & L \text{ even} \end{cases} \tag{6.31}$$

so that $b_1 = -1/2$, $b_2 = -1/8$, $b_3 = 0$, etc. Using (6.30) to asymptotically invert the transform in (6.21) leads to

$$\begin{aligned} F_2(\gamma) &= \frac{1}{2\pi i}\int_{Br} e^{\gamma\theta} F(\theta) d\theta \\ &\sim \frac{1}{2\pi i}\int_{Br} e^{\gamma\theta} A^{2^{-\theta}} d\theta \\ &= \frac{1}{2\pi i}\int_{C} \frac{1}{\log 2} e^{-\gamma \log_2 u} e^{u \log A}\frac{du}{u} \\ &= \frac{1}{\log 2}\frac{(\log A)^{\gamma/\log 2}}{\Gamma(1+\gamma/\log 2)} \\ &\sim \frac{1}{\sqrt{2\pi\gamma}}\frac{1}{\sqrt{\log 2}}\ \exp\left[-\frac{\gamma}{\log 2}\log\left(\frac{\gamma}{\log 2}\right)\right]\ \exp\left[\frac{\gamma}{\log 2}(1+\log\log A)\right],\ \gamma \to \infty. \end{aligned} \tag{6.32}$$

Here the contour $C$ goes from $-\infty + i0^-$ to $-\infty + i0^+$, encircling the branch cut along the negative real axis in the $u$-plane. By comparing (6.26) (with (6.28)) to (6.32) we conclude that

$$k_2 = \frac{\log\log 2}{\log 2} + \frac{1 + \log\log A}{\log 2}. \tag{6.33}$$

In Section 9 we shall numerically obtain $A$ and then (6.33) will be used to obtain $k_2$. We note that (6.30) implies that

$$2^\theta \log[F(\theta)] = \log A - \frac{1}{2}2^{-\tau}A^{-2^\tau}(1 + o(1)), \quad \tau = -\theta \to \infty \tag{6.34}$$

so that the left side should approach a constant value super-exponentially fast as $\theta \to -\infty$.

Finally, we invert (2.3) to get $b(n,p)$ using the result on the $\gamma$-scale. We scale

$$p = \Theta n^{4/3},\ \Theta = O(1), \quad \Theta > 0 \tag{6.35}$$

and use (6.1) so that

$$w^{-p} = \ \exp\left[\frac{\gamma p}{n} + O\left(\frac{p}{n^2}\right)\right] \sim\ \exp(\gamma n^{1/3}\Theta). \tag{6.36}$$

With (6.2), (6.5), (6.9) and (6.36), (2.3) becomes

$$\frac{1}{2\pi i}\int_C B_n(w) w^{-p-1} dw \sim \frac{4^{n+1}}{n}\frac{1}{2\pi i}\int_{Br} 4\gamma n^{-\gamma/3} F_1(\gamma)\ \exp[n^{1/3}(\gamma\Theta - |r_0|4^{1/3}\gamma^{2/3})]\frac{d\gamma}{n}. \tag{6.37}$$

For $n \to \infty$ we evaluate the last integral by the saddle point method, noting that there is a saddle point along the real axis, at

$$\frac{d}{d\gamma}(\gamma\Theta - |r_0|4^{1/3}\gamma^{2/3}) = 0 \Rightarrow \gamma = \gamma_* \equiv \frac{32}{27}\frac{|r_0|^3}{\Theta^3}. \tag{6.38}$$

Then the standard Laplace estimate of (6.37) yields

$$b(n,p) \sim \frac{4^n}{n^{13/6}}\frac{|r_0|^{9/2}}{\Theta^5}\frac{512}{81}\frac{n^{-\gamma_*/3}}{\sqrt{\pi}}F_1(\gamma_*)\ \exp\left[-\frac{16}{27}n^{1/3}\frac{|r_0|^3}{\Theta^2}\right] \tag{6.39}$$

which applies for $\Theta = pn^{-4/3}$ fixed as $n \to \infty$. Here $F_1(\cdot)$ must be obtained by numerically solving (6.10) or (6.20). We have thus established (2.41).

Since $F_1(0) = 1$, expanding (6.39) as $\Theta \to \infty$ corresponds simply to dropping the factors $n^{-\gamma_*/3}$ and $F_1(\gamma_*)$. Thus (6.39) asymptotically matches to the result valid for $p = O(n^{3/2})$, in view of (5.125) and the fact that $\Omega = \Theta n^{-1/6}$. We can also infer the behavior of the right side of (6.39) as $\Theta \to 0$, which corresponds to $\gamma_* \to \infty$ and $p = o(n^{3/2})$. Using (6.29) we obtain

$$\begin{aligned}
&\frac{4^n}{n^{13/6}}\frac{512}{81\pi}\frac{1}{\sqrt{2\log 2}}\frac{|r_0|^{9/2}}{\Theta^5}\frac{\Theta^{3/2}}{|r_0|^{3/2}}\frac{3\sqrt{3}}{4\sqrt{2}}\ \exp\left[-\frac{16}{27}n^{1/3}\frac{|r_0|^3}{\Theta^2}\right] \\
&\times\ \exp\left[-\frac{1}{3}\gamma_*\log n + \left(\frac{1}{3} - \frac{1}{\log 2}\right)\gamma_*\log\gamma_* + k_2\gamma_*\right] \\
&= \frac{1}{n^{13/6}}\frac{|r_0|^3}{\Theta^{7/2}}\frac{64}{9\sqrt{3}}\frac{1}{\pi\sqrt{\log 2}}\ \exp\left[n\log 4 - 4^{1/3}\frac{9}{16}|r_0|n^{1/3}\gamma_*^{2/3}\right. \\
&\left. - \frac{1}{3}\gamma_*\log n + \left(\frac{1}{3} - \frac{1}{\log 2}\right)\gamma_*\log\gamma_* + k_2\gamma_*\right].
\end{aligned} \tag{6.40}$$

We will use (6.40) for asymptotic matching purposes in Sections 7 and 8.

# 7 Far Left Region

We consider (2.1) for $0 < w < 1$ and $n \to \infty$, and (2.3) for $n \to \infty$ with $p - n\log_2 n = O(n)$. Note that $p - n\log_2 n = O(n)$ contains $p_{\min}(n)$ in (2.13) (see also (7.29)).

We assume first that $B_n(w)$ has an expansion of the form

$$B_n(w) = e^{-n\log n f(w)}e^{ng(w)}n^{h(w)}q(w)[1 + o(1)], \tag{7.1}$$

for $0 < w < 1$ and $n \to \infty$. If $f(w) > 0$ this means that $B_n(w)$ will decay faster than exponentially as $n \to \infty$. Since

$$(n+1)\log(n+1) = n\log n + \log n + 1 + O(n^{-1})$$

we see that with (7.1) the left side of (2.1) becomes

$$e^{-n\log n f(w)}e^{ng(w)}n^{-f(w)+h(w)}q(w)e^{g(w)-f(w)}[1 + o(1)]. \tag{7.2}$$

To evaluate the right side we treat the sum as an implicit Laplace-type integral. Using (7.1) the sum becomes asymptotic to

$$e^{n\log w}\sum_{\ell=0}^{n}\ell^{h(w)}(n-\ell)^{h(w)}q^2(w)e^{ng(w)}\exp\{-[\ell\log\ell+(n-\ell)\log(n-\ell)]f(w)\}. \tag{7.3}$$

We approximate the sum via Euler-MacLaurin to get

$$e^{n\log w}n^{2h(w)+1}q^2(w)e^{ng(w)}e^{-n\log nf(w)}\int_0^1 x^{h(w)}(1-x)^{h(w)}e^{-nf(w)H(x)}dx \tag{7.4}$$

where $H(x)=x\log x+(1-x)\log(1-x)$ is as in (6.4). Assuming for now that $f(w)>0$ we can evaluate (7.4) as a Laplace integral for $n\to\infty$, with the major contribution coming from $x=\frac{1}{2}+O(n^{-1/2})$. Using

$$H\left(\frac{1}{2}\right)=-\log 2,\ H'\left(\frac{1}{2}\right)=0,\ H''\left(\frac{1}{2}\right)=4, \tag{7.5}$$

(7.4) becomes asymptotic to

$$e^{ng(w)}e^{-n\log nf(w)}q^2(w)n^{2h(w)+1}\sqrt{\frac{\pi}{2f(w)n}}\left(\frac{1}{2}\right)^{2h(w)}\exp[n(f(w)\log 2+\log w)]. \tag{7.6}$$

Comparing (7.1) to (7.6) we conclude that

$$f(w)=-\frac{\log w}{\log 2}=-\log_2 w>0, \tag{7.7}$$

$$h(w)=\log_2 w-\frac{1}{2}, \tag{7.8}$$

and

$$q(w)=w^{2+\frac{1}{\log 2}}\sqrt{-\log_2 w}\frac{1}{\sqrt{2\pi}}e^{g(w)}. \tag{7.9}$$

However, the function $g(w)$ remains undetermined. To summarize, assuming the *ansatz* (7.1) we have obtained

$$B_n(w)=w^{n\log_2 n}e^{(n+1)g(w)}n^{\log_2 w}\sqrt{\frac{-\log_2 w}{2\pi n}}w^{2+\frac{1}{\log 2}}[1+o(1)]. \tag{7.10}$$

The numerical studies in Section 9 suggest that (7.10) is approximately correct for $w>0$, but there are oscillations present that become numerically significant when $w$ becomes small. We thus re-examine (2.1) with the more general *ansatz*

$$B_n(w)=e^{-n\log nf(w)}e^{nB(w,n)}n^{h(w)}Q(w,n)[1+o(1)]. \tag{7.11}$$

We allow $B(w,n)$ and $Q(w,n)$ to depend *weakly* upon $n$, in such a way that $B(w,n+1) \sim B(w,n)$ and $Q(w,n+1) \sim Q(w,n)$ as $n \to \infty$. Repeating the calculation we did with (7.1) we now find that for $n \to \infty$ the left side of (2.1) is

$$e^{-n\log n f(w)} e^{(n+1)B(w,n+1)} n^{h(w)-f(w)} Q(w,n+1) e^{-f(w)}. \tag{7.12}$$

The sum in the right side of (2.1) becomes

$$\begin{aligned}&e^{n\log w}\sum_{\ell=0}^{n} \ell^{h(w)}(n-\ell)^{h(w)} Q(w,\ell)Q(w,n-\ell)e^{\ell B(w,\ell)+(n-\ell)B(w,n-\ell)}\\ &\times \exp\{-[\ell\log\ell + (n-\ell)\log(n-\ell)]f(w)\}.\end{aligned} \tag{7.13}$$

Again evaluating this by the Laplace method, we use

$$\begin{aligned}\ell B(w,\ell) \quad &+ \quad (n-\ell)B(w,n-\ell)\\ &= \quad nB\left(w,\frac{n}{2}\right) + \frac{1}{n}\left(\ell-\frac{n}{2}\right)^2\left[\frac{n^2}{2}B_{nn}\left(w,\frac{n}{2}\right) + 2nB_n\left(w,\frac{n}{2}\right)\right] + \ldots,\end{aligned} \tag{7.14}$$

which follows upon expanding in Taylor series about $\ell = n/2$. Then (7.13) becomes, for $n \to \infty$,

$$\begin{aligned}&e^{n\log w}e^{-n\log n f(w)}e^{n(\log 2)f(w)}n^{1+2h(w)}\left(\frac{1}{2}\right)^{2h(w)}\\ &\times e^{nB\left(w,\frac{n}{2}\right)}Q^2\left(w,\frac{n}{2}\right)\sqrt{\frac{\pi}{2n}}\left[f(w)-nB_n\left(w,\frac{n}{2}\right)-\frac{n^2}{4}B_{nn}\left(w,\frac{n}{2}\right)\right]^{-1/2}.\end{aligned}$$

Comparing (7.12) to (7.14) we again obtain (7.7). Then using

$$(n+1)B(w,n+1) = nB(w,n) + B(w,n) + nB_n(w,n) + \cdots$$

we find that $B(w,n)$ satisfies

$$B(w,n) = B\left(w,\frac{n}{2}\right). \tag{7.15}$$

The most general solution of (7.15) is

$$B(w,n) = g(w) + \sum_{\substack{k=-\infty\\ k\neq 0}}^{\infty} g_k(w)e^{2\pi i(\log_2 n)k}, \tag{7.16}$$

where the sum represents an arbitrary, zero-mean periodic function of $\log_2 n$, with period one. Note that (7.15) and (7.16) are indeed consistent with our assumption of slow variation. Unfortunately, we cannot determine explicitly the Fourier coefficients $g_k(w)$ using purely the recurrence (2.1). With (7.16) we write

$$\begin{aligned}nB_n\left(w,\frac{n}{2}\right) \quad &= \quad nB_n(w,n) = B_1^*(w,n),\\ n^2B_{nn}\left(w,\frac{n}{2}\right) \quad &= \quad n^2B_{nn}(w,n) = B_2^*(w,n)\end{aligned}$$

where

$$\begin{aligned} B_1^*(w,n) &= \frac{2\pi i}{\log 2}\sum_{k=-\infty}^{\infty} k g_k(w) e^{2\pi i(\log_2 n)k} \\ B_2^*(w,n) &= \frac{2\pi i}{\log 2}\sum_{k=-\infty}^{\infty}\left[\frac{2\pi i}{\log 2}k^2 - k\right] g_k(w) e^{2\pi i(\log_2 n)k}. \end{aligned} \tag{7.17}$$

Returning to (7.12) and (7.14) we regain (7.8) and instead of (7.9) we obtain the equation

$$\begin{aligned} Q^2\left(w,\frac{n}{2}\right) &= Q(w,n) w^{2+\frac{1}{\log 2}}\sqrt{2\pi} e^{g(w)} \\ &\times \; e^{B_0^*(w,n)+B_1^*(w,n)}\sqrt{-\log_2 w - B_1^*(w,n) - \frac{1}{4}B_2^*(w,n)}. \end{aligned} \tag{7.18}$$

Here we wrote (7.16) as $B(w,n) = g(w) + B_0^*(w,n)$.

A particular solution to (7.18) is

$$Q_p(w,n) = w^{2+\frac{1}{\log 2}}\frac{1}{\sqrt{2\pi}} e^{g+B_0^*+B_1^*}\sqrt{-\log_2 w - B_1^* - \frac{1}{4}B_2^*} \tag{7.19}$$

and setting

$$Q(w,n) = Q_p(w,n)\tilde{Q}(w,n)$$

we find that $\tilde{Q}$ satisfies

$$\tilde{Q}^2\left(w,\frac{n}{2}\right) = \tilde{Q}(w,n). \tag{7.20}$$

Setting

$$\tilde{G}(w,n) = \log[\tilde{Q}(w,n)]$$

we find that

$$2\tilde{G}\left(w,\frac{n}{2}\right) = \tilde{G}(w,n) \tag{7.21}$$

whose most general solution is

$$\tilde{G}(w,n) = n \times \; [\text{periodic function of } \log_2 n]. \tag{7.22}$$

However, then $\tilde{Q}(w,n)$ can be incorporated into the factor $\exp[nB(w,n)]$ in (7.11). Thus in (7.11) we can use (7.7), (7.8), (7.16) and (7.19). We have thus established (2.34) and (2.35).

We next examine the asymptotic matching between the results for $0 < w < 1$ and $w = 1 - O(n^{-1})$. For $\gamma \to \infty$ we expand (6.2) to get

$$\frac{4}{n}^{n+1} \exp[4^{1/3} r_0 n^{1/3}\gamma^{2/3} - \frac{1}{3}\gamma\log n]\frac{e^{k_2\gamma}}{\sqrt{2\pi\log 2}}\sqrt{\gamma}\exp\left[\left(\frac{1}{3} - \frac{1}{\log 2}\right)\gamma\log\gamma\right] \tag{7.23}$$

where we also used (6.5), (6.9) and (6.29). By asymptotic matching, (7.23) should agree with the expansion of (7.11) as $w \to 1$. Since (7.23) contains no oscillatory terms, we conclude that $nB_0^*(w,n) \to 0$ as $w \to 1$ and thus $g_k(w) = o(1-w)$ as $w \to 1$, for each $k$. Then (7.23) must match to the right side of (7.10) (which neglected the oscillations), as the latter is expanded for $w \to 1$. Noting that

$$w^{n \log_2 n} = \left(1 - \frac{\gamma}{n}\right)^{n \log_2 n} \sim \exp\left[-\frac{\gamma}{\log 2} \log n\right]$$

and

$$n^{\log_2 w} \sim 1,$$

(7.10) becomes

$$\frac{1}{\sqrt{2\pi \log 2}} \sqrt{\frac{1-w}{n}} \exp\left(-\frac{\gamma}{\log 2} \log n\right) e^{(n+1)g(1-\gamma/n)}.$$

Thus, the matching is possible provided that

$$\begin{aligned} g(w) &= \log 4 + 4^{1/3} r_0 (1-w)^{2/3} + \left(\frac{1}{\log 2} - \frac{1}{3}\right)(w-1)\log(1-w) \\ &+ k_2(1-w) + o(1-w), \qquad w \uparrow 1. \end{aligned} \tag{7.24}$$

We have thus used asymptotic matching to infer the behavior of $g(w)$ as $w \to 1$, and this will play a major role in Section 8.

Finally we discuss briefly the limit $w \to 0$ with $n$ fixed. From the discussion in Section 3, each $B_n(w)$ is a polynomial of the form

$$B_n(w) = C_n w^{p_{\min}(n)} + \cdots + 2^{n-1} w^{\binom{n}{2}} \tag{7.25}$$

where

$$p_{\min}(n) = \sum_{J=2}^{n} \lfloor \log_2 J \rfloor = (n+1)\lfloor \log_2 n \rfloor - 2(2^{\lfloor \log_2 n \rfloor} - 1).$$

By using (7.25) in (2.1) we find that the dominant terms for $w \to 0$ come from near the midpoint of the sum $\ell \approx n/2$. More precisely, for $n$ even we let $n = 2N$ and obtain

$$C_{2N+1} = \sum_{j \in S(N)} C_{N+j} C_{N-j} \tag{7.26}$$

where

$$S(N) = \{\ell : p_{\min}(N+\ell) + p_{\min}(N-\ell) = 2p_{\min}(N)\}.$$

For $n$ odd we set $n = 2N+1$ and get

$$C_{2N+2} = \sum_{j \in T(N)} C_{N+j} C_{N+1-j} \tag{7.27}$$

where

$$T(N) = \{\ell : p_{\min}(N+\ell) + p_{\min}(N+1-\ell) = p_{\min}(N) + p_{\min}(N+1)\}.$$

Note that $0 \in S(N)$ for all $N$, and $0, 1 \in T(N)$. If $\ell \in S(N)$ then $-\ell \in S(N)$, and if $\ell \in T(N)$ then $1-\ell \in T(N)$. For example, we have $S(1) = \{0\}$, $S(2) = \{-1, 0, 1\}$, $S(3) = \{0\}$, $T(0) = \{0, 1\}$, $T(1) = \{0, 1\}$, $T(2) = \{0, 1\}$. Thus for $n$ fixed we have

$$B_n(w) \sim C_n w^{p_{\min}(n)}, \; w \to 0. \tag{7.28}$$

We next show that (7.28) cannot be a special case of (7.14), as the latter is expanded for $w \to 0$. Using the Fourier series'

$$\begin{aligned}
\{x\} &= \frac{1}{2} - \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{e^{2\pi ikx}}{2\pi ik} = \frac{1}{2} - \sum_{k=1}^{\infty} \frac{\sin(2\pi kx)}{\pi k}, \\
2^{-\{x\}} &= \frac{1}{2} \sum_{k=-\infty}^{\infty} \frac{1}{\log 2 + 2k\pi i} e^{2\pi ikx},
\end{aligned}$$

where the first sum omits the term $k = 0$, we can represent $p_{\min}(n)$ as

$$\begin{aligned}
p_{\min}(n) &= n \log_2 n + n \left[ -\frac{1}{2} - \frac{1}{\log 2} \right] \\
&= n \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{\log 2}{2\pi ik(\log 2 + 2\pi ik)} e^{2\pi ik \log_2 n} \\
&= \log_2 n + \frac{3}{2} + \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{e^{2\pi ik \log_2 n}}{2\pi ik}.
\end{aligned} \tag{7.29}$$

This form yields the asymptotic behavior as $n \to \infty$. Now, if (7.14) were to contain (7.28) as a special case (at least for $n \to \infty$), it would need to behave as $C_n^{asy}\ \exp(p_{\min}(n) \log w)$ for $w \to 0$, where $C_n^{asy}$ is understood as the asymptotic behavior of $C_n$ as $n \to \infty$. By comparing (7.29) to (2.34) we see that the largest factors, i.e., $\exp(n \log_2 n \log w)$, match automatically, and the factors that are $\exp(O(n))$ agree if

$$\begin{aligned}
g(w) &\sim \left( -\frac{1}{2} - \frac{1}{\log 2} \right) \log w, \; w \to 0, \\
g_k(w) &\sim \frac{\log 2}{2\pi ik(\log 2 + 2\pi ik)} \log w, \; w \to 0.
\end{aligned} \tag{7.30}$$

With (7.30) we then have

$$e^{g(w)} w^{2+\frac{1}{\log 2}} \sim w^{3/2}, \quad w \to 0,$$

and the term $\log_2 n$ in $p_{\min}(n)$ corresponds to the factor $n^{\log_2 w} = w^{\log_2 n}$ in (2.34). Furthermore

$$\begin{aligned}\exp[B_0^*(w,n) + B_1^*(w,n)] &= \exp\left\{\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty}\left(1+\frac{2\pi i k}{\log 2}\right) g_k(k) e^{2\pi i k \log_2 n}\right\} \\ &\sim \exp\left\{\log w \left[\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{e^{2\pi i k \log_2 n}}{2\pi i k}\right]\right\}, \qquad w \to 0\end{aligned}$$

which corresponds to the last sum in (7.29). However, (2.34) still has the factor

$$\frac{1}{\sqrt{2\pi n}}\sqrt{-\log_2 w - B_1^*(w,n) - \frac{1}{4}B_2^*(w,n)}$$

which for $w \to 0$ would look like $\sqrt{-\log_2 w}$ times a function of $n$, and such a factor can have no analog in (7.28), regardless of the form of $C_n$ for $n \to \infty$.

This shows that yet another scale is needed in order to completely understand the asymptotic behavior of $B_n(w)$. This scale would have $w \to 0$ and $n \to \infty$ simultaneously. We have not been able to identify this scale precisely nor to analyze it. It is, however, not important for understanding the asymptotic for $\sum_n b(n,p)$, as this requires $w \to 1$, as shown in Section 8.

Finally, we use the form (7.10) to obtain an approximation to $b(n,p)$. We set

$$p = n\log_2 n + \alpha n,\ \alpha = O(1)$$

so that

$$w^{n\log_2 n} w^{-p-1} = w^{-\alpha n - 1}$$

and we have, for $\alpha$ fixed,

$$b(n,p) \approx \frac{1}{2\pi i}\int_C \frac{1}{w} e^{g(w)} n^{\log_2 w} \sqrt{-\frac{\log_2 w}{2\pi n}}\, w^{2+\frac{1}{\log 2}} \exp[n(g(w) - \alpha \log w)]dw. \tag{7.31}$$

For $n \to \infty$ we can use the saddle point method. There is a saddle in the range $w = O(1)$ if we can solve the equation $g'(w) = \alpha/w$. The numerical studies of $g(w)$ in Section 9 suggest this is indeed possible, and then the standard saddle point approximation leads to (2.42).

# 8 The Matching Region Between the Left and Far Left Scales

In Section 7 we showed that the expansions for $0 < w < 1$ ($p = n\log_2 n + O(n)$) and $w = 1 - O(n^{-1})$ ($p = O(n^{4/3})$) can match in some intermediate limit. Here we examine in more detail the solution in this matching region, as this is the key to understanding the distribution of the number of nodes in trees of a given (large) total path length $p$ (cf. (2.12)). For a fixed $p$, as we increase $n$ we move from right to left. The result in (6.39) shows that for a fixed $\Theta = pn^{-4/3}$, $b(n,p)$ still grows with $n$, due to the dominant exponential factor of $4^n$. However, for $p = n\log_2 n + O(n)$, (2.42) shows that for fixed $p/n - \log_2 n$, $b(n,p)$ grows (decays) with $n$ according as $g(w_*) - \alpha \log w_* > 0$ ($< 0$). However, numerical studies (cf. Section 9) show that this quantity is negative unless $w_* \to 1$. Thus to find the limiting distribution of the number of nodes we need the maximum of $b(n,p)$ over $n$, and this occurs precisely in the matching region, where this function will achieve a Gaussian peak.

As $(p - n\log_2 n)/n = \alpha \to \infty$ we have $w_* \to 1$ and the non-constant Fourier coefficients in (2.34) (i.e., $g_k(w)$, $k \neq 0$) vanish as $w \to 1$. Thus we use (2.37) and (2.42). For $\alpha \to \infty$ we can solve the equation

$$wg'(w) = \alpha \tag{8.1}$$

for the saddle $w = w_*(\alpha)$ asymptotically, using the relations

$$g(w) = \log 4 - 4^{1/3}|r_0|(1-w)^{2/3} + \left(\frac{1}{\log 2} - \frac{1}{3}\right)(w-1)\log(1-w) - k_2(w-1) + o(w-1), \tag{8.2}$$

$$g'(w) = 4^{1/3}\frac{2}{3}|r_0|\left(\frac{1}{1-w}\right)^{1/3}\left(\frac{1}{\log 2} - \frac{1}{3}\right)\log(1-w) + -\frac{1}{3} + \frac{1}{\log 2} - k_2 + o(1), \tag{8.3}$$

and

$$g''(w) \sim \frac{2}{9}4^{1/3}|r_0|(1-w)^{-4/3}, \quad w \to 1^-. \tag{8.4}$$

To facilitate the calculation, we set

$$A_0 = \frac{2}{3}4^{1/3}|r_0|,\ A_1 = \frac{1}{\log 2} - \frac{1}{3},\ A_2 = -\frac{1}{3} + \log 2 - k_2 = A_1 - k_2 \tag{8.5}$$

and

$$w_* = 1 - \Delta. \tag{8.6}$$

We rewrite (8.1) as

$$\alpha = [A_0\Delta^{-1/3} + A_1\log\Delta + A_2 + o(1)](1-\Delta), \quad \Delta \to 0 \tag{8.7}$$

which can be rearranged to give

$$\begin{aligned}
\frac{\Delta}{A_0^3} &= [\alpha - A_1\log\Delta - A_2 + \alpha\Delta + o(1)]^{-3} \\
&= \alpha^{-3}\left[1 - \frac{A_1\log\Delta + A_2}{\alpha} + \Delta + o\left(\frac{1}{\alpha}\right)\right]^{-3} \\
&= \alpha^{-3}\left[1 + \frac{3}{\alpha}(A_1\log\Delta + A_2) + o(\alpha^{-1})\right].
\end{aligned} \tag{8.8}$$

It follows that

$$\Delta \sim \frac{A_0^3}{\alpha^3},\ \alpha \to \infty \tag{8.9}$$

and this can be refined to the expansion

$$\Delta = \left(\frac{A_0}{\alpha}\right)^3 \left[1 + \frac{\delta_1}{\alpha} + \frac{\delta_2}{\alpha^2} + \cdots\right], \ \alpha \to \infty \tag{8.10}$$

where

$$\delta_1 = 3A_2 - 9A_1 \log\left(\frac{\alpha}{A_0}\right). \tag{8.11}$$

We next evaluate $g(w_*) - \alpha \log w_*$ as $\alpha \to \infty$ and $w_* \to 1$. We note that this is the exponential growth rate, in $n$, of (2.42). Using (8.6), (8.10) and (8.11) we obtain

$$\begin{aligned} g(w_*) - \alpha \log w_* &= g(1-\Delta) - \alpha \log(1-\Delta) \\ &= \log 4 + \frac{1}{2} A_1 \Delta \log \Delta + \left(\frac{3}{2} A_2 + k_2\right) \Delta - \frac{1}{2} \alpha \Delta + O(\alpha \Delta^2) \\ &= \log 4 - \frac{A_0^3}{2} \frac{1}{\alpha^2} + \frac{A_0^3}{\alpha^3} \left[3A_1 \log\left(\frac{\alpha}{A_0}\right) + k_2\right] + O_R(\alpha^{-4}). \end{aligned} \tag{8.12}$$

Here we also used (8.7) in the form $\alpha\Delta \sim A_0 \Delta^{2/3} + A_1 \Delta \log \Delta + A_2 \Delta$.

To find the maximum of $ng(w_*) - n\alpha \log w_*$ we need to asymptotically solve

$$\frac{\partial}{\partial n}[ng(w_*) - n\alpha \log w_*] = g(w_*) + \frac{\log n + 1}{\log 2} \log w_* = 0. \tag{8.13}$$

Here we note that $w_*$ depends on $n$ through $\alpha$, but $g'(w_*) = \alpha / w_*$ by (8.1). Also, $n\alpha = p - n \log_2 n$ so that $\partial(n\alpha)/\partial n = -\log_2 n - 1/\log 2$. Let us write the solution of (8.13) as $n = \tilde{n} = \tilde{n}(p)$. Setting

$$F(n; p) = ng(w_*) - n\alpha \log w_* \tag{8.14}$$

we have

$$F_n(n; p) = g(w_*) + \frac{\log n + 1}{\log 2} \log w_* \tag{8.15}$$

and

$$\begin{aligned} F_{nn}(n; p) &= g'(w_*) \frac{\partial w_*}{\partial n} + \frac{\log w_*}{n \log 2} + \frac{\log n + 1}{w_* \log 2} \frac{\partial w_*}{\partial n} \\ &= \left(\alpha + \frac{\log n + 1}{\log 2}\right) \frac{1}{w_*} \frac{\partial w_*}{\partial n} + \frac{1}{n \log 2} \log w_*. \end{aligned} \tag{8.16}$$

We also define

$$\Psi_0(p) = F(\tilde{n}(p); p) \tag{8.17}$$

and

$$\mathcal{V}_0(p) = -1/F_{nn}(\tilde{n}(p); p). \tag{8.18}$$

We will see that $\Psi_0(p)$, $\tilde{n}(p)$ and $\mathcal{V}_0(p)$ will provide asymptotic approximations to the exponential growth rate of the total number of trees of path length $p$, and the mean and variance of the Gaussian distribution of the number of nodes in such trees.

We use (8.2) and (8.10) in (8.13), which as $w_* \to 1$ yields

$$\log 4 - \frac{\log n + 1}{\log 2}\left[\frac{A_0^3}{\alpha^3} + \frac{3A_0^3}{\alpha^4}\left(A_2 - 3A_1 \log\left(\frac{\alpha}{A_0}\right)\right)\right]$$
$$-\frac{3}{2}A_0\frac{A_0^2}{\alpha^2}\left[1 + \frac{2}{\alpha}\left(A_2 - 3A_1\log\left(\frac{\alpha}{A_0}\right)\right)\right] + k_2\frac{A_0^3}{\alpha^3} - A_1\frac{A_0^3}{\alpha^3}3\log\left(\frac{A_0}{\alpha}\right) + o(\alpha^{-3}) = 0.$$

To leading order, $\log 4$ must be balanced by the term proportional to $\alpha^{-3}\log n$, and we have

$$\alpha \sim \frac{A_0}{(\log 2)^{1/3}}\frac{(\log n)^{1/3}}{(\log 4)^{1/3}} = \frac{A_0(\log n)^{1/3}}{2^{1/3}(\log 2)^{2/3}}, \quad n = \tilde{n}(p). \tag{8.19}$$

This shows that for a fixed $p$, the maximum of $b(n,p)$ occurs in the range $p = n\log_2 n + O[n(\log n)^{1/3}]$. We compare this to $p_{\min}(n) = n\log_2 n + O(n)$ (cf. (2.13)). From (8.9) and (8.19) we conclude that

$$\tilde{\Delta} \sim \frac{2(\log 2)^2}{\log n} \sim 1 - \tilde{w}, \quad n = \tilde{n}(p) \to \infty. \tag{8.20}$$

Here we write $w_* = w_*(n,p)$ via (8.1) and let

$$\begin{aligned} \tilde{w} &= \tilde{w}(p) = w_*(\tilde{n}(p),p),\ \tilde{\Delta} = \tilde{\Delta}(p) = 1 - w_*(\tilde{n}(p),p), \\ \tilde{\alpha} &= \frac{p - \tilde{n}(p)\log_2[\tilde{n}(p)]}{\tilde{n}(p)}. \end{aligned} \tag{8.21}$$

Using (2.42) we have

$$\begin{aligned} \sum_n b(n,p) &\sim \frac{1}{2\pi\tilde{n}}\frac{(\tilde{w})^{2+\frac{1}{\log 2}}}{\sqrt{\tilde{\alpha} + \tilde{w}^2 g''(\tilde{w})}} n^{\log_2 \tilde{w}} e^{g(\tilde{w})}\sqrt{-\log_2 \tilde{w}} \\ &\times \sum_{n=-\infty}^{\infty} \exp\left[ng(\tilde{w}) - n\alpha\log\tilde{w} + \frac{1}{2}F_{nn}(\tilde{n}(p);p)(n - \tilde{n}(p))^2\right] \end{aligned} \tag{8.22}$$

where we approximated the sum by the Laplace method. Now,

$$\begin{aligned} n^{\log_2 \tilde{w}} &= \exp[\log_2 n \log(1 - \tilde{\Delta})] \sim \exp[-2\log 2] = \frac{1}{4}, \\ e^{g(\tilde{w})} &\sim \exp[\log 4] = 4 \end{aligned}$$

and

$$\tilde{\alpha} + \tilde{w}^2 g''(\tilde{w}) \sim \tilde{\alpha} + \frac{1}{3}A_0(\tilde{\Delta})^{-4/3} \sim \frac{1}{3}A_0 2^{-4/3}(\log 2)^{-8/3}(\log\tilde{n})^{4/3}.$$

Thus (8.22) becomes

$$\begin{aligned}\sum_n b(n,p) \quad &\sim \quad \frac{1}{2\pi\tilde{n}}\left(\frac{1}{\log\tilde{n}}\right)^{7/6}\sqrt{\frac{3}{A_0}}2^{7/6}(\log 2)^{11/6} \\ &\times \quad \sqrt{\frac{2\pi}{-F_{nn}(\tilde{n}(p);p)}}\;\exp[F(\tilde{n}(p);p)]. \qquad (8.23)\end{aligned}$$

In view of (8.13) and (8.14) we have

$$F(\tilde{n}(p);p)) = -\left(p+\frac{\tilde{n}}{\log 2}\right)\log\tilde{w} = \left(p+\frac{\tilde{n}}{\log 2}\right)[\tilde{\Delta}+O(\tilde{\Delta}^2)]. \qquad (8.24)$$

From (8.23), (8.24), (8.19) and (8.20) we obtain the growth rate

$$\log\left[\sum_n b(n,p)\right] \sim \frac{p}{\log p}(2\log^2 2), \quad p\to\infty. \qquad (8.25)$$

To refine (8.25) and make (8.23) more explicit, we consider (8.1) and (8.13) as a simultaneous system to determine $\tilde{w}$ and $\tilde{n}$ (or, equivalently, $\tilde{\alpha}$) as functions of $p$. By expanding (8.13) about $w_* = 1$ and setting

$$S = S(p) = \log\tilde{n} \qquad (8.26)$$

we obtain

$$0 = \log 4 - \frac{S}{\log 2}\Delta - \frac{S}{2\log 2}\Delta^2 - \frac{3}{2}A_0\Delta^{2/3} - A_1\Delta\log\Delta + k_2\Delta - \frac{1}{\log 2}\Delta + O_R(S^{-2}) \qquad (8.27)$$

where we also used (8.2). From (8.27) it follows that as $S\to\infty$, $\Delta$ has the expansion

$$\Delta = \frac{a}{S} + \frac{b}{S^{4/3}} + \frac{c}{S^{5/3}} + \frac{d}{S^2} + O_R(S^{-7/3}) \qquad (8.28)$$

where

$$\begin{aligned} a &= 2(\log 2)^2,\ b = 0, & (8.29)\\ c &= -\frac{3}{2}(\log 2)A_0 a^{2/3} = -\frac{3}{2}2^{2/3}A_0(\log 2)^{7/3}, & (8.30)\\ d &= \log 2\left[\left(k_2 - \frac{1}{\log 2}\right)a - \frac{a^2}{2\log 2} + A_1 a\log\left(\frac{S}{a}\right)\right] & (8.31)\end{aligned}$$

and we note that $d$ depends weakly on $S$ (as $\log S$).

Next we use (8.3) and rewrite (8.1) as

$$\frac{p}{\tilde{n}} - \log_2\tilde{n} = (1-\Delta)[A_0\Delta^{-1/3} + A_1\log\Delta + A_2 + o(1)]. \qquad (8.32)$$

We set

$$\tilde{n}(p) = pn_*(p), \quad Q = \log p \tag{8.33}$$

with which (8.32) becomes

$$\begin{aligned}\frac{1}{n_*} - \frac{S}{\log 2} &= A_0 \left(\frac{S}{a}\right)^{1/3} - A_0 \frac{c}{3a^{4/3}} \frac{1}{S^{1/3}} - A_0 \frac{d}{3a^{4/3}} \frac{1}{S^{2/3}} \\ &+ A_1 \left[ -\log\left(\frac{S}{a}\right) + \frac{c}{a} \frac{1}{S^{2/3}} \right] + A_2 - A_0 \frac{a^{2/3}}{S^{2/3}} + O_R(S^{-1}) \end{aligned} \tag{8.34}$$

where we expanded $\Delta$ in (8.32) using (8.28). Noting that

$$S = Q + \log n_* \sim Q$$

and

$$S^{1/3} \sim Q^{1/3} + \frac{1}{3} Q^{-2/3} \log n_*$$

we rewrite (8.34) as

$$\frac{1}{n_*} = \frac{Q}{\log 2} + \frac{\log n_*}{\log 2} + \frac{A_0}{a^{1/3}} Q^{1/3} + A_2 + A_1 \log a - A_1 \log Q + O_R(Q^{-1/3}). \tag{8.35}$$

It follows that $n_*$ has the expansion

$$n_* = \frac{\log 2}{Q} + \frac{\nu}{Q^{5/3}} + \frac{\nu'}{Q^2} + O_R(Q^{-7/3}) \tag{8.36}$$

where

$$\nu = -\frac{A_0}{a^{1/3}} (\log 2)^2$$

and

$$\nu' = \log 2[\log Q - \log\log 2] + A_1 \log Q (\log 2)^2 - (A_2 + A_1 \log a)(\log 2)^2.$$

Hence, as $p \to \infty$ (with $Q = \log p$) we have

$$\begin{aligned}\tilde{n}(p) &= \frac{p}{Q} \log 2 \left[ 1 - \frac{A_0}{a^{1/3}} \frac{\log 2}{Q^{2/3}} + \frac{Z}{Q} + O_R(Q^{-4/3}) \right], \\ Z &= (A_1 \log 2 + 1) \log Q - \log\log 2 - (A_2 + A_1 \log a) \log 2. \end{aligned} \tag{8.37}$$

Recalling that $n = \tilde{n}(p)$ corresponds to the maximum of $b(n,p)$ for a fixed large $p$, (8.37) is also the expansion of $\mathcal{N}(p)$, so we have established (2.18). We also note that

$$n_{\max}(p) - \tilde{n}(p) \sim A_0 \frac{(\log 2)^2}{a^{1/3}} \frac{p}{(\log p)^{5/3}} = o(p),\ p \to \infty.$$

Here $n_{\max}(p)$ is the inverse of $p_{\min}(n)$ in (2.13).

Next we refine (8.25) and also obtain the variance in (8.18). From (8.17), (8.24) and (8.28) we have

$$\begin{aligned}
\Psi_0(p) &= \left(p + \frac{\tilde{n}}{\log 2}\right)\left[\frac{a}{S} + \frac{c}{S^{5/3}} + \left(d + \frac{a^2}{2}\right)\frac{1}{S^2} + O_R(S^{-7/3})\right] \\
&= \left[p + \frac{p}{\log p} + O\left(\frac{p}{Q^{5/3}}\right)\right]\left\{\frac{a}{Q} + \frac{c}{Q^{5/3}}\right. \\
&\quad \left. + \left[d + \frac{a^2}{2} + a(\log Q - \log\log 2)\right]\frac{1}{Q^2} + O_R(Q^{-7/3})\right\} \\
&= \frac{p}{Q}\left[a + \frac{c}{Q^{2/3}} + \left(a + \frac{a^2}{2} + d - a\log\log 2 + a\log Q\right)\frac{1}{Q} + O_R(Q^{4/3})\right]. \qquad (8.38)
\end{aligned}$$

From (8.31) we see that

$$d + a + \frac{a^2}{2} = a\log 2\left[k_2 + A_1\log\left(\frac{Q}{a}\right)\right] + o(1),\ Q \to \infty$$

so that (8.38) leads to (2.14) with (2.15).

To obtain the variance, we use (8.16) and (8.18). From (8.10) we obtain

$$\frac{\partial w_*}{\partial n} \sim -\frac{3}{\alpha^4}\frac{A_0^3}{n}\left(\frac{p}{n} + \frac{1}{\log 2}\right)\left[1 + \frac{4A_2}{\alpha} + \frac{3A_1}{\alpha} - \frac{12A_1}{\alpha}\log\left(\frac{\alpha}{A_0}\right)\right]$$

and then (8.16) becomes

$$-nF_{nn} = \left[\frac{\log n}{\log 2} + \alpha + \frac{1}{\log 2}\right]^2\frac{3A_0^3}{\alpha^4}\left\{1 + \frac{1}{2}\left[4A_2 + 3A_1 - 12A_1\log\left(\frac{\alpha}{A_0}\right)\right]\right\} + O\left(\frac{1}{\log n}\right). \qquad (8.39)$$

Now let $n = \tilde{n}(p)$ and use

$$\left[\frac{\log n}{\log 2} + \alpha + \frac{1}{\log 2}\right]^2 = \left[\frac{Q}{\log 2} + O(Q^{1/3})\right]^2 = \frac{Q^2}{(\log 2)^2} + O(Q^{4/3}).$$

From (8.19) and (8.37) we obtain

$$\tilde{\alpha} \sim \frac{A_0}{2^{1/3}(\log 2)^{2/3}}[\log p + \log\log 2 - \log Q + O(Q^{-2/3})]^{1/3} \sim \frac{A_0}{2^{1/3}(\log 2)^{2/3}}Q^{1/3}.$$

Using the above in (8.39) we obtain

$$F_{nn}(\tilde{n}(p); p) \sim -\frac{1}{\tilde{n}(p)}\frac{Q^2}{(\log 2)^2}\frac{3}{A_0}2^{4/3}(\log 2)^{8/3}Q^{-4/3} \sim -\frac{Q^{5/3}}{p}\frac{6a^{1/3}}{A_0\log 2}. \qquad (8.40)$$

We have thus obtained the leading term in the variance in (2.19). We will obtain the correction term using a somewhat different method (cf. (8.46) and (8.47)).

Using (8.40) and (8.37) in (8.23) leads to (2.14). To summarize, in the limit $\alpha \to \infty$ and $pn^{-4/3} \to 0$ we have obtained the approximation

$$\begin{aligned} b(n,p) \quad &\sim \quad \frac{3}{\pi} \frac{2^{1/6}}{\sqrt{\log 2} |r_0|^{1/2}} \frac{1}{n} \left(\frac{A_0}{\alpha}\right)^{7/2} \exp\left[-\frac{A_0^3}{\alpha^3} \frac{\log n}{\log 2}\right] \\ &\times \quad 4^n \exp\left[-\frac{A_0^3 n}{2\alpha^2} + \frac{A_0^3 n}{\alpha^3}(k_2 + 3A_1 \log \alpha - 3A_1 \log A_0)\right]. \end{aligned} \tag{8.41}$$

This follows by expanding (2.42) for $w_* \to 1$, and also matches to (6.40), as can be seen by replacing $\alpha$ in (8.41) by $p/n - \log_2 n \sim p/n$. We note that in this limit

$$\gamma_* = \frac{32}{27} \frac{|r_0|^3}{\Theta^3} = \frac{32}{27} \frac{|r_0|^3 n^4}{p^3} \sim \frac{A_0^3}{\alpha^3} n.$$

Expanding (8.41) about $n = \tilde{n}(p)$ (or $\alpha = \tilde{\alpha}$) leads to the Gaussian form in (2.20).

For numerical purposes it sometimes proves useful to base our approximations on (8.41). To this end we define

$$\begin{aligned} H(\alpha) \quad &= \quad -\frac{A_0^3}{2\alpha^2} + \frac{A_0^3}{\alpha^3}(k_2 + 3A_1 \log \alpha - 3A_1 \log A_0) \\ \Phi(n,p) \quad &\equiv \quad n \log 4 + nH(\alpha). \end{aligned} \tag{8.42}$$

Then $\partial \Phi / \partial n = 0$ is equivalent to solving

$$\log 4 + H(\alpha) - \frac{1}{\log 2} H'(\alpha) - \frac{p}{n} H'(\alpha) = 0. \tag{8.43}$$

We denote the solution to (8.43) by

$$n = \hat{n}(p), \quad \hat{\alpha}(p) = \frac{p}{\hat{n}(p)} - \log_2[\hat{n}(p)]$$

where we are viewing the problem for a fixed $p$. We also have, since $\alpha = p/n - \log_2 n$,

$$\frac{\partial^2 \Phi}{\partial n^2} = -\frac{1}{n \log 2} H'(\alpha) + \frac{1}{n} \left(\frac{p}{n} + \frac{1}{\log 2}\right)^2 H''(\alpha)$$

and we get $\hat{\Phi} = \Phi(\hat{n}(p), p)$ and

$$\hat{\mathcal{V}}(p) = -\frac{1}{\Phi_{nn}(\hat{n}(p), p)}.$$

Solving for $\hat{n}$ asymptotically as $p \to \infty$ regains the expansion of the mean in (8.37), though using the more implicit form (8.43) may be advantageous numerically (cf. Section 9). Thus we can use $(\hat{\Phi}, \hat{n}, \hat{\mathcal{V}})$ as the approximations to the growth rate, mean and variance, and these are asymptotically equivalent to the results in (2.14) - (2.19). For example, solving (8.43) for $p \to \infty$ we have

$$\hat{\alpha} = A_0 \left(\frac{Q}{a}\right)^{1/3} + A_2 + A_1 \log a - A_1 \log Q + O_R(Q^{-1/3}) \tag{8.44}$$

and we note that

$$H''(\hat{\alpha}) = -\frac{3A_0^3}{\hat{\alpha}^4} + \frac{A_0^3}{\hat{\alpha}^5}\left[36A_1 \log\left(\frac{\hat{\alpha}}{A_0}\right) - 12A_2 - 9A_1\right]. \tag{8.45}$$

Using (8.44) in (8.45) leads to

$$H''(\hat{\alpha}) = -\frac{3}{A_0}\left(\frac{a}{Q}\right)^{4/3} - \frac{9A_1}{A_0^2}\left(\frac{a}{Q}\right)^{5/3} + O_R(Q^{-2}). \tag{8.46}$$

Thus,

$$\frac{\partial^2 \Phi}{\partial n^2}(\hat{n}(p), p) = \frac{1}{\hat{n}}\left(\frac{p}{\hat{n}} + \frac{1}{\log 2}\right)^2 H''(\hat{\alpha}) + O\left(\frac{1}{\hat{n}\log\hat{n}}\right). \tag{8.47}$$

Using (8.46) in (8.47) leads to the two term approximation to the variance in (2.19). Note that while the leading terms in the expansion of the growth rate ($\log\sum_n b(n,p)$) and the mean do not involve the root $r_0$ of the Airy function, the leading term for the variance does involve this root. Also, the correction terms for the growth rate and mean are smaller than the leading terms by factors of $Q^{-2/3} = (\log p)^{-2/3}$, but the correction term to the variance is smaller by only an $O(Q^{-1/3}) = O((\log p)^{-1/3})$ factor. This completes the proof of Result 2.

# 9 Numerical Studies

We give numerical results that determine some unknown constants/functions, and also provide back-up to the various assumptions we made.

For $w > 1$ we gave in Table 1, numerical values of $B_*(w)$ ($\sim B_n(w)2^{1-n}w^{-\binom{n}{2}}$, $n \to \infty$) for various $w$, and sketched this function in Figure 1 (cf. Section 2). We test the accuracy of the asymptotic relation in (2.23). Let $\bar{B}_n(w)$ be the exact solution to (3.2). Then our analysis predicts that $\bar{B}_n(w) \to B_*(w)$ as $n \to \infty$ for each $w > 1$, and then

$$(w-1)\log[B_*(w)] \to d_0 = \int_0^{\log 2} \frac{\zeta}{e^\zeta - 1} d\zeta = .58224\ldots \tag{9.1}$$

as $w \downarrow 1$. In Table 2, we compute $(w-1)\log[\bar{B}_\infty(w)]$ and also $(w-1)\left[\log[\bar{B}_\infty(w)] - \frac{1}{2}\log(w-1)\right]$ for various $w > 1$. Here, for each fixed $w$, $\bar{B}_\infty(w)$ is computed by iterating (3.2) for $n$ sufficiently large until $\bar{B}_n(w)$ settles to a constant to 3 digits. Both of the tabulated functions should approach $d_0$ as $w \downarrow 1$, with the latter function approaching more rapidly, since it includes information from the algebraic factor $\sqrt{w-1}$ in (2.23). We see that both functions are indeed approaching $d_0$ as $w \to 1$, with the second function having a faster rate of approach. Next we attempt to verify the constant $d_1$ in the asymptotic relation (2.23), by computing

$$D_1(w) \equiv \exp\left(-\frac{d_0}{w-1}\right)\frac{\bar{B}_\infty(w)}{\sqrt{w-1}}$$

in Table 3, for various $w > 1$. Our analysis predicts that $D_1(w) \to d_1 = 2.1350\ldots$ as $w \downarrow 1$. It became very difficult to compute $\bar{B}_\infty(w)$ for $w \le 1.04$, since for $w \gtrsim 1$ the convergence of $\bar{B}_n$ to $\bar{B}_\infty$ is very slow. We see from Table 3 that $D_1(w)$ still changes appreciably as $w$ goes from 1.08 to 1.06 to 1.04. The

**Tab. 2:**

| $w$ | $(w-1)\log(\bar{B}_\infty)$ | $(w-1)\left[\log(\bar{B}_\infty) - \frac{1}{2}\log(w-1)\right]$ |
|---|---|---|
| 2 | .358 | .358 |
| 1.8 | .372 | .461 |
| 1.6 | .390 | .544 |
| 1.4 | .417 | .601 |
| 1.2 | .462 | .623 |
| 1.18 | .468 | .622 |
| 1.16 | .475 | .621 |
| 1.14 | .482 | .620 |
| 1.12 | .490 | .618 |
| 1.10 | .499 | .614 |
| 1.08 | .510 | .611 |
| 1.06 | .521 | .606 |
| 1.04 | .535 | .600 |
| 1.02 | .552 | .592 |
| 1 | .582 | .582 |

**Tab. 3:**

| $w$ | $D_1(w)$ |
|---|---|
| 5 | .46499 |
| 4.5 | .49262 |
| 4 | .52591 |
| 3.5 | .56716 |
| 3 | .62021 |
| 2.5 | .69229 |
| 2 | .79923 |
| 1.8 | .85978 |
| 1.6 | .93848 |
| 1.4 | 1.0485 |
| 1.2 | 1.2274 |
| 1.18 | 1.2531 |
| 1.16 | 1.2812 |
| 1.14 | 1.3123 |
| 1.12 | 1.3473 |
| 1.10 | 1.3872 |
| 1.08 | 1.4338 |
| 1.06 | 1.490 |
| 1.04 | 1.56 |
| 1 | 2.1350 |

data are not inconsistent with the theoretical value, but $w$ would have to be made much closer to one for us to reach a more definite conclusion.

We next consider the $\beta$-scale, where $w = 1 + \beta/n = 1 + O(n^{-1})$ with $\beta > 0$. Here the asymptotic result is given by (2.25). In Tables 4 and 5, we compare

$$\frac{1}{n}\log\left[B_n\left(1+\frac{\beta}{n}\right)\right] \tag{9.2}$$

to $\Phi(\beta) = \log 2 + \frac{\beta}{2} + \phi(\beta)$ (cf. (2.25)) for $\beta = .25, .5, 1, 2$ and 4, respectively, and for various $n$. Our WKB expansion predicts that (9.2) should approach $\Phi(\beta)$ as $n \to \infty$. The tables clearly demonstrate this convergence. The data are also consistent with an $O(n^{-1}\log n)$ error term, which is predicted by (2.25). The convergence is slowest when $\beta = .25$ (cf. Table 4), which is to be expected since our analysis predicts that once $\beta$ becomes $O(n^{-1/2})$, the expansion becomes invalid and we must use the $a$-scale result.

In Table 6, we compare

$$\sqrt{n}e^{-n\Phi(\beta)}B_n\left(1+\frac{\beta}{n}\right) \tag{9.3}$$

to $\sqrt{\beta}\hat{g}(\beta)$, which should be the limit of (9.3) as $n \to \infty$. Table 6 has $\beta = .5, 1$, and 2, respectively. Again the tables show the convergence predicted by the WKB *ansatz* (4.9). However, the convergence is much slower than one would expect with the $O(n^{-1})$ error term in (4.9). The numerical data suggest that the error term may be $O(n^{-1/2})$ (i.e., a term $n^{-1/2}g^{(1/2)}(\beta)$ should be included in the series in (4.9)). To resolve this issue more conclusively, we would have to examine higher-order asymptotic matching between the $a$-scale and $\beta$-scale. This would necessitate, among other things, continuing (2.28) into the range $a > 0$ and the evaluating the result for $a \to +\infty$.

We next consider the $a$-scale result. We recall that obtaining the right tail of the Airy distribution involved showing that

$$D(y) = D(a^{2/3}) \sim \frac{4}{\sqrt{\pi}}ae^{a^2/3},\ a \to +\infty. \tag{9.4}$$

In Table 7 we give $D(a^{2/3})a^{-1}e^{-a^2/3}$ for various values of $a \geq 1$. This clearly shows that the function is approaching the constant $4/\sqrt{\pi} = 2.2567\ldots$ . To compute the exact values of $D(y)$ we used the recurrence (5.25), (5.23) and (5.20).

Now consider the expansion on the $\gamma$-scale, where $w = 1 - O(n^{-1})$. Here we must solve (6.20) or, equivalently, (6.22). Given the asymptotic behavior of $F(\theta)$ in (6.24) as $\theta \to \infty$, we used the following numerical scheme: (i) Fix a large $N$, (ii) let $\tilde{F}(J)$ satisfy the recurrence

$$\tilde{F}(J-1) = \tilde{F}(J) + \tilde{F}^2(J),\ J = N, N-1, \ldots, -M+1, \tag{9.5}$$

subject to the "terminal" condition

$$\tilde{F}(N) = \frac{1}{N} + \frac{\log N + \alpha_*}{N^2} + \frac{(\log N + \alpha_*)^2 - \log N - \alpha_* + \frac{3}{2}}{N^3},$$

(iii) use the approximation

$$F(-M) \approx \tilde{F}(-M).$$

**Tab. 4:**

| $\beta = .25$ | | | $\beta = .5$ | | | $\beta = 1$ | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $\frac{1}{n}\log(B_n)$ | $\Phi(\beta)$ | $n$ | $\frac{1}{n}\log(B_n)$ | $\Phi(\beta)$ | $n$ | $\frac{1}{n}\log(B_n)$ | $\Phi(\beta)$ |
| 10 | 1.0517 | 1.4039 | 10 | 1.1310 | 1.4475 | 10 | 1.2910 | 1.5822 |
| 20 | 1.1984 | | 20 | 1.2702 | | 20 | 1.4245 | |
| 30 | 1.2569 | | 30 | 1.3232 | | 30 | 1.4724 | |
| 40 | 1.2889 | | 40 | 1.3514 | | 40 | 1.4972 | |
| 50 | 1.3092 | | 50 | 1.3689 | | 50 | 1.5125 | |
| 75 | 1.3378 | | 75 | 1.3931 | | 75 | 1.5335 | |
| 100 | 1.3530 | | 100 | 1.4055 | | 100 | 1.5445 | |
| 125 | 1.3624 | | 125 | 1.4132 | | 125 | 1.5513 | |
| 150 | 1.3688 | | 150 | 1.4184 | | 150 | 1.5559 | |
| 200 | 1.3769 | | 200 | 1.4251 | | 200 | 1.5618 | |
| 250 | 1.3820 | | 250 | 1.4292 | | 250 | 1.5655 | |
| 500 | 1.3923 | | 500 | 1.4377 | | 500 | 1.5732 | |
| 750 | 1.3959 | | 750 | 1.4408 | | 750 | 1.5759 | |
| 1000 | 1.3978 | | 1000 | 1.4423 | | 1000 | 1.5774 | |
| 2000 | 1.4006 | | 2000 | 1.4448 | | 2000 | 1.5796 | |
| 3000 | 1.4017 | | 3000 | 1.4456 | | 3000 | 1.5804 | |
| 5000 | 1.4025 | | 5000 | 1.4463 | | 5000 | 1.5811 | |
| 10,000 | 1.4031 | | 10,000 | 1.4469 | | 10,000 | 1.5816 | |

**Tab. 5:**

| $\beta = 2$ | | | $\beta = 4$ | | |
|---|---|---|---|---|---|
| $n$ | $\frac{1}{n}\log(B_n)$ | $\Phi(\beta)$ | $n$ | $\frac{1}{n}\log(B_n)$ | $\Phi(\beta)$ |
| 10 | 1.6123 | 1.9498 | 10 | 2.2336 | 2.8364 |
| 20 | 1.7670 | | 20 | 2.5013 | |
| 30 | 1.8223 | | 30 | 2.6024 | |
| 40 | 1.8512 | | 40 | 2.6559 | |
| 50 | 1.8690 | | 50 | 2.6892 | |
| 75 | 1.8936 | | 75 | 2.7351 | |
| 100 | 1.9064 | | 100 | 2.7588 | |
| 125 | 1.9143 | | 125 | 2.7734 | |
| 150 | 1.9197 | | 150 | 2.7833 | |
| 200 | 1.9266 | | 200 | 2.7959 | |
| 250 | 1.9308 | | 250 | 2.8035 | |
| 500 | 1.9397 | | 500 | 2.8193 | |
| 750 | 1.9428 | | 750 | 2.8247 | |
| 1000 | 1.9444 | | 1000 | 2.8275 | |
| 2000 | 1.9469 | | 2000 | 2.8318 | |
| 3000 | 1.9478 | | 3000 | 2.8333 | |
| 5000 | 1.9486 | | 5000 | 2.8344 | |
| 10,000 | 1.9491 | | 10,000 | 2.8354 | |

**Tab. 6:**

| $\beta = .5$ | | | $\beta = 1$ | | | $\beta = 2$ | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $\sqrt{n}e^{-n\Phi}B_n$ | $\sqrt{\beta}\hat{g}$ | $n$ | $\sqrt{n}e^{-n\Phi}B_n$ | $\sqrt{\beta}\hat{g}$ | $n$ | $\sqrt{n}e^{-n\Phi}B_n$ | $\sqrt{\beta}\hat{g}$ |
| 10 | .13346 | .18155 | 10 | .17190 | .28196 | 10 | .10819 | .16437 |
| 20 | .12886 | | 20 | .19081 | | 20 | .11552 | |
| 30 | .13149 | | 30 | .20327 | | 30 | .11961 | |
| 40 | .13512 | | 40 | .21132 | | 40 | .12245 | |
| 50 | .13859 | | 50 | .21688 | | 50 | .12461 | |
| 75 | .14540 | | 75 | .22557 | | 75 | .12844 | |
| 100 | .14995 | | 100 | .23085 | | 100 | .13105 | |
| 125 | .15307 | | 125 | .23456 | | 125 | .13300 | |
| 150 | .15530 | | 150 | .23737 | | 150 | .13454 | |
| 200 | .15833 | | 200 | .24144 | | 200 | .13687 | |
| 250 | .16033 | | 250 | .24433 | | 250 | .13858 | |
| 500 | .16525 | | 500 | .25202 | | 500 | .14338 | |
| 750 | .16749 | | 750 | .25574 | | 750 | .14583 | |
| 1000 | .16885 | | 1000 | .25812 | | 1000 | .14740 | |
| 2000 | .17158 | | 2000 | .26297 | | 2000 | .15072 | |
| 3000 | .17299 | | 3000 | .26509 | | 3000 | .15256 | |
| 4000 | .17369 | | 4000 | .26684 | | 4000 | .15343 | |
| 5000 | .17426 | | 5000 | .26791 | | 5000 | .15419 | |
| 10,000 | .17578 | | 10,000 | .27078 | | 10,000 | .15623 | |

**Tab. 7:**

| $a$ | $D(a^{2/3})a^{-1}e^{-a^2/3}$ |
|---|---|
| 1 | 2.242 |
| 2 | 2.035 |
| 3 | 2.102 |
| 4 | 2.168 |
| 5 | 2.204 |
| 6 | 2.222 |
| 7 | 2.232 |
| 8 | 2.238 |
| 9 | 2.242 |
| 10 | 2.245 |
| 11 | 2.247 |
| 12 | 2.248 |
| 13 | 2.249 |
| 14 | 2.250 |
| 15 | 2.251 |
| 20 | 2.253 |
| 25 | 2.254 |
| 30 | 2.255 |
| 50 | 2.256 |
| 100 | 2.256 |
| $\infty$ | 2.2567... |

**Tab. 8:**

| $N$ | $2^{-5}\tilde{F}(-5)$ |
|---|---|
| 1000 | 6.8809 |
| 2000 | 6.8807 |
| 3000 | 6.8804 |
| 5000 | 6.8800 |
| 10,000 | 6.8794 |

For $M$ large and positive, (6.30) shows that

$$2^{-M}\log[F(-M)] = \log A - \frac{1}{2}2^{-M}A^{-2^M} + \cdots, \ M \to \infty. \tag{9.6}$$

Thus the left side of (9.5) should approach the constant $\log A$ very rapidly (as a double exponential) as $M \to \infty$, if $A > 1$. We have found that taking $M = 5$ is sufficient to get $\log A$ to several decimal places. In Table 8 we take $M = 5$, solve (9.5) for various $N$ and give $2^{-5}\tilde{F}(-5)$, which is an approximation to $\log A$ that should become exact as $N \to \infty$. This shows that

$$\log A \approx 6.880 \tag{9.7}$$

and then (6.33) yields

$$k_2 \approx 3.696. \tag{9.8}$$

The latter constant corresponds to the exponential growth rate of $F_1(\gamma)$ in (6.29), and thus that of the approximation on the $\gamma$-scale in (6.2) or (2.30) - (2.33).

Next we consider $n \to \infty$ with $0 < w < 1$, where the approximation (2.37) applies (or its refinement (2.34), that allows for oscillations). If we were to ignore the oscillations and define

$$g_{\text{num}}^{(n)}(w) \equiv \frac{1}{n+1}\log\left[\frac{\sqrt{2\pi n\log 2}B_n(w)w^{-n\log_2 n}}{n^{\log_2 w}\sqrt{-\log w}w^{2+\frac{1}{\log 2}}}\right], \tag{9.9}$$

then we should have $g_{\text{num}}^{(n)}(w) \to g(w)$ as $n \to \infty$ for each fixed $0 < w < 1$. Of course we should also have $n^{-1}\log[B_n(w)w^{-n\log_2 n}] \to g(w)$, but including the other factors in the right side of (9.9) should help with the convergence, as shown by the asymptotic analysis of Section 7. In Table 9, we plot $g_{\text{num}}^{(n)}(w)$ large $n$, for several $w < 1$. These data show that for $n$ sufficiently large, $g_{\text{num}}^{(n)}(w)$ is constant to several decimal places, except when $w \leq .001$, where this function oscillates (with $n$) over a certain range. This range is indicated in Table 9 for $w = .001$, .0001 and .00001. Our study suggests that while the oscillations corresponding to $g_k(w)$ in (2.35) are indeed present, they are numerically very small, unless $w$ itself is very small. We also recall that the analysis in Section 7 predicted that $g_k(w)$ has a logarithmic singularity, for each $k$, as $w \to 0$ (cf. (7.30)). For practical purposes we can use (2.37) and (2.42) as the approximations to $B_n(w)$ and $b(n,p)$, though $g(w)$ must still be determined numerically, using, e.g., Table 9. We also note from the table that $g(w)$ is approaching, as $w \to 1$, the theoretical value $g(1) = \log 4 = 1.3862\ldots$ .

In Table 10 we examine the matching region between the scales $w = 1 - O(n^{-1})$ and $0 < w < 1$. By asymptotic matching we showed that $g(w)$ behaves as (8.2) for $w \uparrow 1$. Let us define CONST =

$$\mathrm{CONST}(w;n)$$

$$\equiv \frac{1}{1-w}\left\{ g_{\mathrm{num}}^{(n)}(w) - \left[\log 4 + 4^{1/3} r_0 (1-w)^{2/3} + \left(\frac{1}{\log 2} - \frac{1}{3}\right)(w-1)\log(1-w)\right]\right\}. \quad (9.10)$$

For $w$ not small, we can choose $n$ sufficiently large to make (9.10) independent of $n$ to several decimal places. Then as $w \to 1$ our analysis shows that (9.9) should approach the constant $k_2$. The convergence becomes very slow with $n$ when $w$ is only slightly less than one, but the data shows that CONST is decreasing to some value $< 3.792$, and this is consistent with our theoretical value in (9.8). This provides numerical evidence both for the form (8.2), as well as the numerical value of $k_2$.

Finally we do some numerical studies of our main results, which consist of the growth rate of the total number of trees of a given path length, and the distribution of these trees as characterized by their numbers of nodes. In Figure 2, we plot the exact $b(n,p)$ for $p = 190$ and $p = 250$ with $n \in [n_{\min}(p), n_{\max}(p)] = [20, 49]$ and $[23, 61]$. While this graph possibly resembles a Gaussian near the peak, the value of $p$ is too small for us to be certain. In Table 11 we compare the exact values of $\log[\sum_n b(n,p)]$ to the one-term, two-term and three-term asymptotic approximations to the exponent in (2.14). The one-term approximation is $ap/Q = 2(\log 2)^2 p/\log p$, the two-term includes the $O(Q^{-2/3})$ correction, and the three-term includes also the $O_R(Q^{-1})$ correction. Table 11 shows that the one-term approximation underestimates the true value by about 20%, the two-term approximation is very poor, while the three-term approximation is quite accurate. In Table 12, we give the exact mean total path length $\mathcal{N}(p)$ (cf. the one- two- and three-term asymptotic approximations in (2.18). We again see that the two-term approximation is actually worse than the one-term, and the three-term result is better than the one-term. In Table 13, we give the exact variance $\mathcal{V}(p)$ along with the one- and two-term approximations in (2.19). We see that the one-term approximation is fairly accurate (10% errors), while the two-term underestimates significantly the true value.

To explain these numerical trends we note that the ratios of coefficients of the second term to the first term are respectively given by:

growth rate: $R_0 \equiv \frac{3}{2} A_0 \frac{\log 2}{a^{1/3}} = 2^{1/3}(\log 2)^{1/3}|r_0| = 2.6070\ldots$

mean: $R_1 \equiv \frac{2}{3} 2^{1/3}(\log 2)^{1/3}|r_0| = 1.7380\ldots$

variance: $R_2 \equiv 3\frac{A_1}{A_0} a^{1/3} = \frac{9}{2}\frac{1}{|r_0|}\left(1 - \frac{\log 2}{3}\right)\left(\frac{1}{2\log 2}\right)^{1/3} = 1.3272\ldots$ .

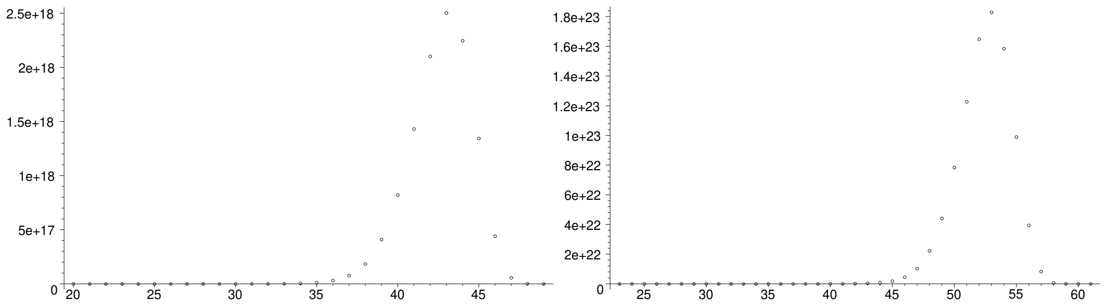
All of these numbers exceed unity. To compensate for the ratio being large we need to have $Q^{-2/3}$ (growth rate, mean) or $Q^{-1/3}$ (variance) sufficiently small. However, when $p = 250, Q^{1/3} = (\log p)^{1/3} = 1.767\ldots$ and $Q^{2/3} = 3.123\ldots$; increasing $p$ to 500 gives $Q^{1/3} = 1.838\ldots$ and $Q^{2/3} = 3.380\ldots$ . Thus to make the second term in the growth rate numerically smaller than the first, we need $p \approx 67.3$. Indeed, Table 11 shows that the two-term approximation is actually negative for $p = 20 - 50$. To make the second

**Tab. 9:**

| $w$ | $g_{\mathrm{num}}^{(n)}(w)$ |
|---|---|
| .999999 | 1.3853 |
| .99999 | 1.3841 |
| .9999 | 1.3791 |
| .999 | 1.3605 |
| .99 | 1.3042 |
| .9 | 1.26 |
| .8 | 1.37 |
| .7 | 1.55 |
| .6 | 1.77 |
| .5 | 2.06 |
| .4 | 2.44 |
| .3 | 2.94 |
| .2 | 3.66 |
| .1 | 4.94 |
| .01 | 9.33 |
| .001 | 13.7 - 13.8 |
| .0001 | 18.1 - 18.4 |
| .00001 | 22.5 - 23.0 |

**Tab. 10:**

| $w$ | $g_{\text{num}}^{(n)}(w)$ | CONST |
|---|---|---|
| .8 | 1.379 | 4.525 |
| .9 | 1.269 | 4.273 |
| .925 | 1.256 | 4.200 |
| .95 | 1.254 | 4.116 |
| .975 | 1.2715 | 4.010 |
| .99 | 1.3042 | 3.915 |
| .999 | 1.3606 | 3.792 |




(a) $p = 190$ with $n$ in the range $[20..49]$ (b) $p = 250$ with $n$ in the range $[23..61]$

**Fig. 2:** Plots of $b(n, p)$.

**Tab. 11:**

| $p$ | $\log[\sum_n b(n,p)]$ | 1-term | 2-term | 3-term |
|---|---|---|---|---|
| 20 | 6.4425 | 6.4151 | $< 0$ | 8.860 |
| 30 | 9.2943 | 8.4755 | $< 0$ | 11.47 |
| 40 | 11.702 | 10.419 | $< 0$ | 13.92 |
| 50 | 14.301 | 12.281 | $< 0$ | 16.25 |
| 75 | 20.145 | 16.692 | .27965 | 21.72 |
| 100 | 25.659 | 20.865 | 1.2131 | 26.86 |
| 125 | 30.940 | 24.876 | 2.1739 | 31.77 |
| 150 | 36.037 | 28.765 | 3.1545 | 36.50 |
| 175 | 40.989 | 32.558 | 4.1501 | 41.11 |
| 200 | 45.822 | 36.272 | 5.1575 | 45.61 |
| 225 | 50.553 | 39.918 | 6.1743 | 50.00 |
| 250 | 55.198 | 43.507 | 7.1988 | 54.33 |

**Tab. 12:**

| $p$ | $\mathcal{N}(p)$ | 1-term | 2-term | 3-term |
|---|---|---|---|---|
| 20 | 9.019 | 4.627 | .7572 | 7.138 |
| 30 | 11.719 | 6.113 | 1.415 | 9.244 |
| 40 | 14.050 | 7.516 | 2.044 | 11.21 |
| 50 | 16.395 | 8.859 | 2.657 | 13.08 |
| 75 | 21.704 | 12.04 | 4.148 | 17.47 |
| 100 | 26.624 | 15.05 | 5.600 | 21.58 |
| 125 | 31.287 | 17.94 | 7.027 | 25.51 |
| 150 | 35.766 | 20.75 | 8.433 | 29.30 |
| 175 | 40.093 | 23.48 | 9.824 | 32.99 |
| 200 | 44.300 | 26.16 | 11.20 | 36.58 |
| 225 | 48.406 | 28.79 | 12.56 | 40.10 |
| 250 | 52.426 | 31.38 | 13.92 | 43.55 |

term one half as large as the leading term, we would need $Q^{2/3} = 2R_0$ or $p = \exp[(2R_0)^{3/2}] \approx 148,000$. This is well outside the range of Table 11, where the maximum value of $p$ is 250. Thus, the fact that the two-term approximations tend to actually be worse than the one-term ones is not surprising, in view of the relatively large values of $R_j$ and the fact that $p$ must be extremely large before $(\log p)^{-1/3}$ or $(\log p)^{-2/3}$ is sufficiently small. The numerical results are not particularly bad, considering that the asymptotic series involves powers of $Q^{-1/3}$.

We have also tested the accuracy of the approximations $(\hat{\Phi}, \hat{n}, \hat{\mathcal{V}})$, defined by (8.43) and below. However, we found these to be worse than the (more explicit) results in (2.15), (2.18) and (2.19). We would guess that having an explicit expression for $g(w)$ (rather than only the local behavior in (8.2)) would lead to more accurate approximations. However, refining (8.2) (i.e., computing explicitly the $O(1-w)$ error term) would seem to require a lot of work, as we would need to compute higher-order terms in the expansions on the $a$- and $\gamma$-scales, and use higher-order asymptotic matching. The $\gamma$-scale would involve numerically solving certain integral equations.

Another idea to approximating $\log\left(\sum_n b(n,p)\right)$, $\mathcal{N}$ and $\mathcal{V}$ would be to define $(\hat{w}_a, \hat{n}_a)$ as the solution to the system

$$\begin{aligned} wg_a'(w) = \alpha = \frac{p}{n} - \log_2 n, \\ g_a(w) + \frac{\log n + 1}{\log 2} \log w = 0 \end{aligned} \tag{9.11}$$

where

$$g_a(w) \equiv \log 4 + 4^{1/3} r_0 (1-w)^{2/3} + \left(\frac{1}{\log 2} - \frac{1}{3}\right)(w-1)\log(1-w) - k_2(w-1) \tag{9.12}$$

is the truncation of the asymptotic series (8.2), up to the point that we explicitly computed it. For a given (large) $p$, we solve (9.11) (with (9.12)) for $(w, n)$ numerically. Then our approximations to the growth rate, mean and variance are

$$\hat{\Phi}_a = \hat{\Phi}_a(p) = \hat{n}_a[g_a(\hat{w}_a) - \hat{\alpha}_a \log(\hat{w}_a)],$$

$\hat{n}_a$ and

$$\begin{aligned} \mathcal{V}_a(p) &= = -\left[\frac{\log(\hat{w}_a)}{\hat{n}_a \log 2} + \left(\hat{\alpha}_a + \frac{\log(\hat{n}_a)+1}{\log 2}\right)\frac{1}{\hat{w}_a}\frac{\partial w}{\partial n}\right]^{-1} \\ &= \hat{n}_a\left[-\frac{\log(\hat{w}_a)}{\log 2} + \left(\frac{p}{\hat{n}_a} + \frac{1}{\log 2}\right)^2 \frac{1}{\hat{w}_a}\frac{1}{\hat{w}_a g_a''(\hat{w}_a) + g_a'(\hat{w}_a)}\right]^{-1}. \end{aligned}$$

Of course, asymptotically these approximations have the same order of accuracy as the explicit results in (2.15), (2.18) and (2.19), as both are based on the approximation (9.13) for $g(w)$. However, numerically $(\hat{\Phi}_a, \hat{n}_a, \mathcal{V}_a)$ lead to somewhat different results. In Table 14, we give these approximations for various $p$ in the range [20,250]. These should be compared to the exact values in Tables 11- 13. In doing the numerics we first replaced $n$ by

$$n = \exp\left(g_a(w)\frac{\log 2}{-\log w} - 1\right) \tag{9.13}$$

**Tab. 13:**

| $p$ | $\mathcal{V}(p)$ | 1-term | 2-term |
|---|---|---|---|
| 20 | .5792 | .9306 | .07377 |
| 30 | .9125 | 1.129 | .1326 |
| 40 | 1.230 | 1.315 | .1854 |
| 50 | 1.347 | 1.491 | .2350 |
| 75 | 1.902 | 1.897 | .3508 |
| 100 | 2.334 | 2.272 | .4595 |
| 125 | 2.719 | 2.625 | .5636 |
| 150 | 3.092 | 2.961 | .6644 |
| 175 | 3.448 | 3.284 | .7625 |
| 200 | 3.789 | 3.597 | .8586 |
| 225 | 4.118 | 3.901 | .9529 |
| 250 | 4.437 | 4.198 | 1.045 |

**Tab. 14:**

| $p$ | $\hat{w}_a$ | $\hat{\Phi}_a$ | $\hat{n}_a$ | $\mathcal{V}_a$ |
|---|---|---|---|---|
| 20 | .7536 | 8.649 | 7.331 | .1377 |
| 30 | .7719 | 11.34 | 9.596 | .2245 |
| 40 | .7837 | 13.85 | 11.69 | .3061 |
| 50 | .7923 | 16.23 | 13.68 | .3842 |
| 75 | .8065 | 21.81 | 18.33 | .5689 |
| 100 | .8156 | 27.03 | 22.67 | .7433 |
| 125 | .8223 | 32.02 | 26.81 | .9106 |
| 150 | .8274 | 36.83 | 30.80 | 1.072 |
| 175 | .8315 | 41.50 | 34.67 | 1.230 |
| 200 | .8350 | 46.06 | 38.44 | 1.383 |
| 225 | .8379 | 50.52 | 42.13 | 1.534 |
| 250 | .8405 | 54.90 | 45.74 | 1.683 |

in the first equation in (9.11), and solved numerically for $\hat{w}_a = \hat{w}_a(p)$. Then $\hat{n}_a$ is easily obtained from (9.13). The data in Table 14 shows that the approximations for $\Psi(p)$ and $\mathcal{N}(p)$ are roughly comparable to the three-term approximations in Tables 11 and 12. The approximation to $\mathcal{V}(p)$ is better than the two-term approximation in Table 13, but worse than the one-term. In Table 14 we also give the value of $\hat{w}_a$. We note that this quantity converges to one quite slowly as $p \to \infty$, and the use of (9.13) is only justified for $w \to 1$.

# A APPENDIX

Here we briefly re-examine the central region by using the functional equation (2.2). We introduce the scaling

$$w = 1 + \frac{a}{n^{3/2}},\ z = \frac{1}{4}\left(1 + \frac{\zeta}{n}\right) \tag{A.1}$$

with

$$B(z, w) = \bar{B}(\zeta, a) = \bar{B}((4z-1)n, (w-1)n^{3/2}). \tag{A.2}$$

From (A.1) we have

$$4zw - 1 = \frac{\zeta}{n} + \left(1 + \frac{\zeta}{n}\right)\frac{a}{n^{3/2}}$$

and (2.2) becomes

$$\bar{B}(\zeta, a) - 1 = \frac{1}{4}\left(1 + \frac{\zeta}{n}\right)\left[\bar{B}\left(\zeta + \frac{a}{\sqrt{n}} + \frac{a\zeta}{n^{3/2}}, a\right)\right]^2. \tag{A.3}$$

Upon setting

$$\bar{B}(\zeta, a) = 2\left[1 + \frac{1}{\sqrt{n}}B_1(\zeta, a)\right]$$

we obtain from (A.3)

$$\begin{aligned}\frac{2}{\sqrt{n}}B_1(\zeta, a) = \frac{2}{\sqrt{n}}B_1\left(\zeta + \frac{a}{\sqrt{n}} + \frac{a\zeta}{n^{3/2}}, a\right) + \frac{\zeta}{n} + \frac{2\zeta}{n^{3/2}}B_1\left(\zeta + \frac{a}{\sqrt{n}} + \frac{a\zeta}{n^{3/2}}, a\right)\\ + \frac{1}{n}B_1^2\left(\zeta + \frac{a}{\sqrt{n}} + \frac{a\zeta}{n^{3/2}}, a\right) + \frac{\zeta}{n^2}B_1^2\left(\zeta + \frac{a}{\sqrt{n}} + \frac{a\zeta}{n^{3/2}}, a\right).\end{aligned} \tag{A.4}$$

Expanding $B_1 = B_1(\zeta, a; n)$ as

$$B_1 = B^{(0)}(\zeta, a) + \frac{1}{\sqrt{n}}B^{(1)}(\zeta, a) + O(n^{-1})$$

we obtain from (A.4) at the first two orders ($O(n^{-1})$ and $O(n^{-3/2})$) the equations

$$0 = 2aB_\zeta^{(0)} + \zeta + [B^{(0)}]^2 \tag{A.5}$$

and

$$0 = 2aB_\zeta^{(1)} + a^2B_{\zeta\zeta}^{(0)} + 2\zeta B^{(0)} + 2B^{(0)}B^{(1)} + 2aB^{(0)}B_\zeta^{(0)}. \tag{A.6}$$

To solve (A.5) we set

$$B^{(0)}(\zeta, a) = 2a\frac{H_\zeta}{H} \tag{A.7}$$

to obtain the Airy equation $4a^2 H_{\zeta\zeta} + \zeta H = 0$ and thus

$$H(\zeta, a) = \mathrm{Ai}((4a^2)^{-1/3}(-\zeta)),\ \zeta < 0. \tag{A.8}$$

We use (A.8) in (A.7) and note that

$$z^{-n-1}dz = \frac{1}{4n}4^{n+1}\left(1+\frac{\zeta}{n}\right)^{-n-1} d\zeta = \frac{4^n}{n}e^{-\zeta}\left[1+O\left(\frac{1}{n}\right)\right]d\zeta. \tag{A.9}$$

We have

$$\begin{aligned} B_n(w) &= \frac{1}{2\pi i}\int_C \frac{B(z,w)}{z^{n+1}}dz \\ &\sim \frac{4^n}{n}\frac{1}{2\pi i}\int_{Br^-} 2e^{-\zeta}\left[1+\frac{1}{\sqrt{n}}\frac{-2a}{(2a^2)^{1/3}}\frac{\mathrm{Ai}'(-4^{-1/3}\zeta|a|^{-2/3})}{\mathrm{Ai}(-4^{-1/3}\zeta|a|^{-2/3})}\right]d\zeta. \end{aligned} \tag{A.10}$$

Here $\Re(\zeta) < 0$ on $Br^-$. Scaling $\zeta = -(-a)^{2/3}s = -Ys$ for $a < 0$ and $Y > 0$, and interpreting the leading term as

$$\frac{1}{2\pi i}\int_{Br} e^{-\zeta}d\zeta = \frac{1}{2\pi i}\int_{Br} e^{Ys}Y\,ds = Y\delta(Y) = 0$$

via distributions, (A.10) becomes

$$\begin{aligned} B_n(w) &\sim \frac{4^n}{n^{3/2}}\frac{-4^{2/3}a}{2\pi i}\int_{Br}\frac{\mathrm{Ai}'(4^{-1/3}s)}{\mathrm{Ai}(4^{-1/3}s)}e^{Ys}ds \\ &= \frac{4^n}{n^{3/2}}\frac{-4^{2/3}a}{2\pi i}\frac{d}{dY}\left(\int_{Br}\frac{\mathrm{Ai}'(4^{-1/3}s)}{\mathrm{Ai}(4^{-1/3}s)}\frac{e^{Ys}}{s}ds\right) \end{aligned} \tag{A.11}$$

where again the first integral in (A.11) must be interpreted using distributions. We have thus regained the leading term for the scale $w = 1 + O(n^{-3/2})$ and $w < 1$. However, the present derivation is not as clean as those in Section 5, where all the functions that arose may be interpreted in the classical sense.

We next obtain the correction term $B^{(1)}(\zeta, a)$. Differentiating (A.5) yields

$$0 = 2aB^{(0)}_{\zeta\zeta} + 1 + 2B^{(0)}B^{(0)}_{\zeta}. \tag{A.12}$$

We rewrite the linear equation (A.6) as

$$\begin{aligned} aB^{(1)}_{\zeta} + B^{(0)}B^{(1)} &= a\left[B^{(1)}_{\zeta} + 2\frac{H_\zeta}{H}B^{(1)}\right] \\ &= -\zeta B^{(0)} - \frac{a^2}{2}B^{(0)}_{\zeta\zeta} - aB^{(0)}B^{(0)}_{\zeta}. \end{aligned} \tag{A.13}$$

By multiplying (A.13) by $H^2$ and using (A.7) and (A.12), we obtain

$$\frac{d}{d\zeta}[H^2 B^{(1)}] = -\frac{3}{2}\zeta H H_\zeta + \frac{1}{4}H^2 + 2a^2\frac{H_\zeta^3}{H}. \tag{A.14}$$

Now set
$$H(\zeta) = h(s) = \mathrm{Ai}(4^{-1/3}s),\ \zeta = -(-a)^{2/3}s.$$
For $\zeta \to -\infty$ we have
$$h(s) = H(\zeta) \approx \exp\left[-\frac{2}{3}(-\zeta)^{3/2}(4a^2)^{-1/2}\right]$$
which follows from (A.8) and the exponential decay rate of the Airy function $\mathrm{Ai}(\cdot)$. It follows that
$$\frac{H_\zeta}{H} \sim \frac{\sqrt{-\zeta}}{2(-a)}, \quad \zeta \to -\infty$$
and the right side of (A.14) will be asymptotic to
$$(-\zeta)^{3/2}H(\zeta)^2\left(-\frac{3}{4a} - \frac{1}{4a}\right) = (-\zeta)^{3/2}H(\zeta)^2\left(-\frac{1}{a}\right).$$
To avoid this growth we set
$$B^{(1)} = a_*\zeta + \bar{B}^{(1)} \tag{A.15}$$
with which (A.14) becomes
$$\frac{d}{d\zeta}[H^2\bar{B}^{(1)}] = -a_*(H^2 + 2\zeta HH_\zeta) - \frac{3}{2}\zeta HH_\zeta + \frac{1}{4}H^2 + 2a^2\frac{H_\zeta^3}{H}. \tag{A.16}$$
By choosing $a_* = -1$ we can avoid growth in $\bar{B}^{(1)}$ as $\zeta \to -\infty$, and (A.16) then becomes
$$\frac{d}{d\zeta}[H^2\bar{B}^{(1)}] = \frac{5}{4}H^2 + \frac{1}{2}\zeta HH_\zeta + 2a^2\frac{H_\zeta^3}{H}. \tag{A.17}$$
By using $H(\zeta) = h(s)$ and
$$B^{(0)} = 2a\frac{H_\zeta}{H} = 2a\frac{h'(s)}{h(s)}\frac{1}{(-Y)}$$
we see that the approximation to $B_n(w)$ becomes
$$\begin{aligned} B_n(w) &= \frac{4^n}{n}\frac{1}{2\pi i}\int_{Br^-} 2e^{-\zeta}\left[1 + \frac{B^{(0)}(\zeta,a)}{\sqrt{n}} + \frac{B^{(1)}(\zeta,a)}{n} + \cdots\right]d\zeta \qquad \text{(A.18)} \\ &= \frac{4^n}{n^{3/2}}\left\{\frac{-4a}{2\pi i}\frac{d}{dY}\int_{Br_+}\frac{h'(s)}{h(s)}e^{Ys}ds + \frac{2}{2\pi i}\frac{Y}{\sqrt{n}}\int_{Br_+}B^{(1)}(-Ys,a)e^{Ys}ds + O(n^{-1})\right\}. \end{aligned}$$
Also, in terms of $s$, (A.17) is
$$-\frac{1}{Y}\frac{d}{ds}[h^2(s)\bar{B}^{(1)}] = \frac{5}{4}h^2(s) + \frac{1}{2}sh(s)h'(s) - \frac{2a^2}{Y^3}\frac{(h'(s))^3}{h(s)}. \tag{A.19}$$
The solution to (A.19) that has acceptable behavior as $s \to \infty$ is given by
$$\bar{B}^{(1)} = \frac{Y}{2}\mathcal{E}_*(s) \tag{A.20}$$

where $\mathcal{E}_*(s)$ is given by (5.87). Noting that $Y^2 = -a\sqrt{Y}$, we have thus regained the correction term $C^{(1)}(a)$ in (2.26) and (2.28). Here we must again interpret

$$\frac{1}{2\pi i}\int_{Br} \zeta e^{-\zeta} d\zeta = \frac{Y^2}{2\pi i}\int_{Br} s e^{Ys} ds = Y^2\delta'(Y) = 0$$

as a distribution.

We have thus shown that a perturbation analysis of (2.2) yields results essentially equivalent to those obtained by the methods of Section 5. However, here we had to interpret certain integrals as distributions; this issue never arose when we analyzed (2.1) or the moment equations (5.7).

## Acknowledgements

We thank Dr. Gadiel Seroussi (HPL, Palo Alto) for motivating the problem, and sharing with us his early results concerning the number of trees of a given path length.

## References

[1] M. Abramowitz, and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1964.

[2] G. Andrews, R. Askey and R. Roy, *Special Functions*, Cambridge University Press, 1999.

[3] C. Bender and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, Mc-Graw Hill, 1978.

[4] I. Csiszár, and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[5] M. Csőrgő, Z. Shi, and M. Yor, Some Asymptotic Properties of the Local Time of the Uniform Empirical Processes, *Bernoulli*, 5, 1035-1058, 1999.

[6] P. Flajolet, and A. Odlyzko, The Average Height of Binary Trees and Other Simple Trees, *J. Computer and System Sciences*, 25, 171–213, 1982.

[7] P. Flajolet and G. Louchard, Analytic Variations on the Airy Distribution, *Algorithmica*, 361-377, 2001.

[8] P. Flajolet, P. Poblete, and A. Viola, On the Analysis of Linear Probing Hashing, *Algorithmica*, 22, 490–515, 1998.

[9] P. Flajolet and R. Sedgewick, *Analytical Combinatorics*, in preparation; see also INRIA TR-1888 1993, TR-2026 1993 and TR-2376 1994.

[10] P. Hennequin, Combinatorial Analysis of Quicksort Algorithm, *Theoretical Informatics and Applications*, 23, 317–333, 1989.

[11] P. Jacquet, and W. Szpankowski, Asymptotic Behavior of the Lempel-Ziv Parsing Scheme and Digital Search Trees, *Theoretical Computer Science*, 144, 161–197, 1995.

[12] P. Jacquet and W. Szpankowski, Precise Worst Case Minimax Redundancy for Markov Sources *IEEE Trans. Information Theory*, 50, 2004.

[13] C. Knessl and W. Szpankowski, Quicksort algorithm again revisited, *Discrete Math. Theor. Comput. Sci.*, 3 43–64, 1999.

[14] C. Knessl and W. Szpankowski, Asymptotic Behavior of the Height in a Digital Search Tree and the Longest Phrase of the Lempel-Ziv Scheme, *SIAM J. Computing*, 30, 923-964, 2000.

[15] D. E. Knuth, *The Art of Computer Programming. Fundamental Algorithms,* Vol. 1, Third Edition, Addison-Wesley, Reading, MA, 1997.

[16] D. E. Knuth, Linear Probing and Graphs, *Algorithmica*, 22, 561–568, 1998.

[17] D. E. Knuth, *Selected Papers on the Analysis of Algorithms*, Cambridge University Press, Cambridge, 2000.

[18] G. Louchard, Kac's Formula, Lévy Local Time and Brownian Excursion, *J. Appl. Probab.*, 21, 479-499, 1984.

[19] G. Louchard, The Brownian Excursion Area: A Numerical Analysis, *Comp. & Maths. with Appls.*, 10, 413-417, 1984.

[20] M. Régnier, A Limiting Distribution for Quicksort, *Theoretical Informatics and Applications*, 23, 335–343, 1989.

[21] U. Rösler, A Limit Theorem for Quicksort, *Theoretical Informatics and Applications*, 25, 85–100, 1991.

[22] G. Seroussi, On Universal Types, *IEEE Trans. Information Theory*, 52, 2006.

[23] G. Seroussi, On the Number of $t$-ary Trees with a Given Pathlength, *Algorithmica*, 2006.

[24] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, John Wiley & Sons, New York, 2001.

[25] L. Takács, A Bernoulli Excursion and its Various Applications, *J. Appl. Probab.*, 23, 557-585, 1991.

[26] L. Takács, Conditional Limit Theorems for Branching Processes, *J. Applied Mathematics and Stochastic Analysis*, 4, 263-292, 1991.

[27] L. Takács, The Asymptotic Distribution of the Total Heights of Random Rooted Trees, *Acta Sci. Math. (Szegad)*, 57, 613-625, 1993.