Progress in Applied Mathematics
Vol. 1, No. 1, 2011, pp. 1-70
www.cscanada.org

ISSN 1925-251X [Print]
ISSN 1925-2528 [Online]
www.cscanada.net

Enumeration of General t-ary Trees and Universal Types

Zhilong ZHANG[1]
Charles Knessl[1,*]

Abstract: We consider t-ary trees characterized by their numbers of nodes and their total path length. When $t = 2$ these are called binary trees, and in such trees a parent node may have up to t child nodes. We give asymptotic expansions for the total number of trees with nodes and path length p, when n and p are large. We consider several different ranges of n and p. For $n \to \infty$ and $p = O(n^{3/2})$ we recover the Airy distribution for the path length in trees with many nodes, and also obtain higher order asymptotic results. For $p \to \infty$ and an appropriate range of n we obtain a limiting Gaussian distribution for the number of nodes in trees with large path lengths. The mean and variance are expressed in terms of the maximal root of the Airy function. Singular perturbation methods, such as asymptotic matching and WKB type expansions, are used throughout, and they are combined with more standard methods of analytic combinatorics, such as generating functions, singularity analysis, saddle point method, etc. The results are applicable to problems in information theory, that involve data compression schemes which parse long sequence into shorter phrases. Numerical studies show the accuracy of the various asymptotic approximations.

Key Words: Trees; Universal Types; Asymptotics; Path Length; Singular Perturbations

1. INTRODUCTION

1.1 Background and Motivation

Trees are the most important and fundamental data structures used in computer science. Mathematically, a tree is an acyclic connected graph where each node has zero or more children nodes, and all nodes except the root node have one parent node. A t-ary ($t \geq 2$) tree[1, 2] T is a finite set of n ($n \geq 0$) nodes with the following properties: (a) the set is empty, $T = \emptyset$; or (b) the set consists of a root, R, and the remaining nodes are partitioned into disjoint sets $T_1, \cdots, T_t$, each of which is a t-ary tree such that $T = \{R, T_1, \cdots, T_t\}$. The trees $T_1, \cdots, T_t$ are called subtrees of the root R. The bigger t is, the larger is the entropy of the t-ary tree. Nodes that do not have any children are called leaf nodes, which are also referred to as external nodes or terminal nodes. An internal node is any node of a tree that has child nodes. The depth of a node is the length of the path from the root to the node. The height of a tree is the length of the path from the root to the deepest node in the tree. We define $\mathcal{T}_n^*$ as the set of all possible t-ary trees with n nodes. The total path length p is defined as the sum, over all nodes, of the depths. We define the set $\mathcal{T}_p$ as the collection of t-ary trees that have path length p. We give an example of a 3-ary (or ternary) tree with 8 nodes in 1, to illustrate

[1]University of Illinois at Chicago, Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 851 South Morgan Street, Chicago, IL 60607-7045, USA. E-mail addresses: dragonzhsd@yahoo.com (Zhilong ZHANG); knessl@uic.edu (Charles Knessl).
*Corresponding author.
†Received 8 December 2010; accepted 5 January 2011.

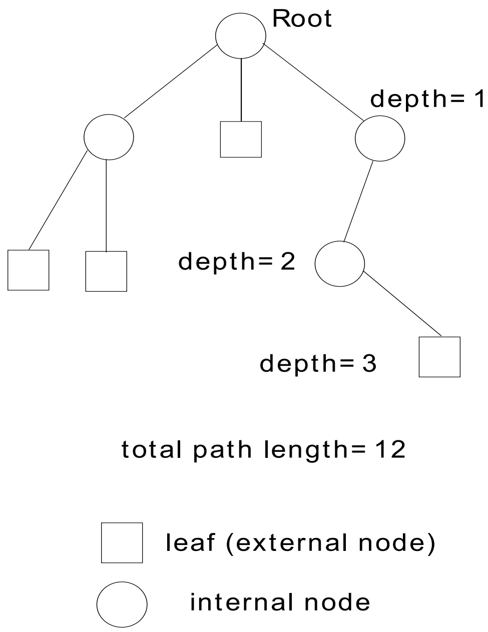


Figure 1: A ternary tree with 8 nodes and total path length 12

the basic concepts. In this example the total path length is $p = 12$.

In this paper, we discuss an application of t-ary unlabeled ordered trees (furthermore called t-ary trees) to information theory. In particular we will discuss the counting of Lempel-Ziv'78 (called LZ'78 hereafter) parsings[3, 4] and universal types[5, 6], which are explained below.

The LZ'78 is a dictionary-based compression scheme that maintains an explicit dictionary via variable-rate coding. The output codewords consist of an index referring to the longest matching dictionary entry and the first non-matching symbol. For example, let $v(j)$ denote the binary sequence of length $j2^j$ that lists all the 2^j binary words of length j, and let

$$v_1^{n(j)} = v(1)v(2)\cdots v(j), \quad n(j) = \sum_{i=1}^{j} i2^i = (j-1)2^{j+1} + 2.$$

It is easy to check that each $v(i)$ is parsed into its 2^i distinct i-tuples, and the maximal number of distinct phrases in parsing $v_1^{n(j)}$ is $2^{j+1} - 2$. For example, $v_1^{n(2)}$ is parsed as

$$v_1^{n(2)} = 0, 1, 00, 01, 10, 11.$$

In other words, the LZ'78 scheme partitions a word into phrases (blocks) of variable size, such that a new block is the shortest sub-word not seen before as a phrase. As a second example, consider the ternary alphabet $\{A, B, C\}$ and the word $ABAACAABBCCACB$. This word of length 14 will be parsed into the phrases $\{A, B, AA, C, AAB, BC, CA, CB\}$ by the LZ'78 scheme.

Tree structures have been extensively investigated for many years, and many interesting results have been published in the literature. Various questions concerning the statistics of randomly generated binary trees were studied in [1, 7–11]. The standard model assumes that all binary unlabeled ordered trees built on n nodes distribute uniformly. The total number of binary trees with n nodes is $|\mathcal{T}_n^*| = \binom{2n}{n}\frac{1}{n+1}$. Flajolet and Odlyzko[12] and Takacs[10] derived the average and the limiting distribution for the height, for $n \to \infty$. Louchard[13, 14] and Takacs[10, 11, 15] established the limiting distribution for the total path length, which can be expressed in terms of the Airy function[16, 17]. The Airy distribution arises in many topics, such as trees, discrete random walk, parking allocation, area under Brownian excursion, hashing tables, (see [10, 11, 13, 14, 18–21]). Properties of this 'Airy distribution' are discussed by Flajolet and Louchard[19] and Majumdar and Comtet[22].

Although profound and interesting results about the behavior of trees in the standard model were discovered, there are still many problems of practical importance that remain unsolved. Seroussi[5, 6] asked for the enumeration of binary trees with a given path length, when studying 'universal types' of sequences and distinct parsings of the Lempel-Ziv scheme. Seroussi observed that the number of possible parsings of sequences of length p corresponds to the cardinality of $\mathcal{T}_p$. Knessl and Szpankowski[23] studied the enumeration of binary trees ($t = 2$) and universal types. Here we generalize their results to t-ary trees. We shall first enumerate $\mathcal{T}_p$ (cf. also [24]), and then compute the limiting distribution of the number of nodes (phrases in the LZ'78 scheme) when a tree is chosen uniformly from the collection $\mathcal{T}_p$.

The method of types[25, 26], involves partitioning sequences of length p into classes according to type, or empirical distribution. It is a powerful technical tool in information theory. This method reduces the computations of rare event probabilities to a combinatorial analysis. Two sequences of the same length p 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 each letter of the alphabet. Seroussi[6] introduced *universal types* of individual sequences, and/or sequences generated by a stationary and ergodic source. Two sequences of the same length are said to be of the same universal type if and only if they produce the same set of phrases in the incremental parsing of the LZ'78 scheme. It was proved that such sequences have the same asymptotic empirical distribution[6]. But, every set of phrases defines uniquely a t-ary tree of path length p (cf. [27]), with the number of phrases corresponding to the number of nodes in the $\mathcal{T}_p$ model. For example, the strings $ABCAABBCCBBBAAC$ and $BCABBAACCAACBBB$ have the same set of phrases $\{A,\ B,\ C,\ AA,\ BB,\ CC,\ BBB,\ AAC\}$ and thus the corresponding ternary trees are the same. Hence, enumeration of $\mathcal{T}_p$ leads to counting universal types, or the different LZ'78 parsings of sequences of length p.

1.2 Mathematical Approach

We let $g(n, p)$ be the number of t-ary trees with n nodes and path length p. The total number trees with n nodes is

$$|\mathcal{T}_n^*| = \sum_{p=0}^{\infty} g(n, p) = \frac{1}{(t-1)n+1}\binom{tn}{n}, \tag{1}$$

which is called the generalized Catalan number[1]. But the total number of trees of path length p ($|\mathcal{T}_p| = \sum_{n=0}^{\infty} g(n, p)$) is unknown, and this is one of the main topics of this paper.

The generating function

$$G_n(w) = \sum_{p=0}^{\infty} w^p g(n, p), \tag{2}$$

satisfies the non-linear recurrence equation

$$G_{n+1}(w) = w^n \sum_{k_1+k_2+\cdots+k_t=n} \left[\prod_{i=1}^{t} G_{k_i}(w)\right], \tag{3}$$

and the double transform

$$G(z, w) = \sum_{n=0}^{\infty}\sum_{p=0}^{\infty} g(n, p) w^p z^n = \sum_{n=0}^{\infty} z^n G_n(w), \tag{4}$$

satisfies the functional equation

$$G(z,w) = 1 + zG^t(zw,w). \tag{5}$$

Note that the coefficient of $G(1,w)$ at w^p is $\sum_{n=0}^{\infty} g(n,p)$, which enumerates the number of t-ary trees with path length p. The functional equation (5) belongs to the class of quick-sort-like nonlinear functional equations (cf. [8, 20, 28–31]) that remains not fully analyzed, with some exceptions like the linear probing algorithm[19, 21].

In this paper we shall analyze $g(n,p)$, $G_n(w)$ and at times $G(z,w)$ asymptotically, for $n \to \infty$. The path length p is necessarily also large, and we shall study various asymptotic scales, such as $p = \binom{n}{2} - O(1)$, $p = O(n^2)$, $p = O(n^{3/2})$, $p = O(n^{4/3})$ and $p = n\log_t n + O(n)$. Our goal is to obtain a thorough understanding of the double sequence $g(n,p)$ when n and p are large. Then the distribution of the trees by path length alone, or number of nodes alone, will follow as special cases. These two "marginal" distributions are quite different, as the path length will follow an Airy distribution[10, 15], while the number of nodes will follow a Gaussian distribution, but with the maximal root of the Airy function $Ai(\cdot)$ appearing in the mean and variance[23].

Seroussi first conjectured and later proved[24] that for t-ary trees, as $p \to \infty$,

$$|\mathcal{T}_p| = t^{\mathcal{A}(p)}, \;\; \mathcal{A}(p) = h(t^{-1})\frac{tp}{\ln p}(1 + o(1)), \tag{6}$$

where $h(x) = -x\ln x - (1-x)\ln(1-x)$. He used information theory and combinatorial arguments. After reconciling notations, we shall show that our result agrees with (6) and we will refine it considerably.

Our approach to analyzing $g(n,p)$, (3) and (5) is analytic, as opposed to probabilistic or combinatorial. The non-linear recurrence equation (3) and the non-linear functional equation (5) appear too difficult to solve in closed form. Such equations cannot be analyzed by standard analytic tools such as transforms. Thus, we shall use methods of applied mathematics such as the **WKB** method and asymptotic matching. We will need to make certain assumptions about the forms of some of the asymptotic expansions, and about the asymptotic matching between different scales. The **WKB** method[9, 32] was named after the physicists Wentzel, Kramers and Brillouin. It assumes that the solution, say $G(\xi;n)$, to a differential equation, functional equation or recurrence has the asymptotic form

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

The equation satisfied by $G(\xi;n)$ implies that $\phi(\xi)$ and $S(\xi)$, $S^{(1)}(\xi)$, $\cdots$ satisfy certain simpler equations, that can often be explicitly solved. The asymptotic matching principle (cf. [32]) is a powerful tool developed over the past 50^+ years, and it allows us to relate expansions on different scales. For example, we shall find that $g(n,p)$ has very different expansions for $n \to \infty$ with $p = O(n^{3/2})$ and $p = O(n^{4/3})$. Asymptotic matching considers the behavior of $g(n,p)$ in an intermediate limit, which lies between the two scales and where $p/n^{3/2} \to 0$ but $p/n^{4/3} \to \infty$.

1.3 Organization

The organization of this paper is as follows. In section 2 we present our main results and briefly discuss them. In sections 3-7 we analyze $G_n(w)$, and then $g(n,p)$, for $n \to \infty$ and various ranges of the generating function variable w. In section 3 we consider $w > 1$ and $n \to \infty$. In section 4 we discuss the limit $w \downarrow 1$ and $n \to \infty$ with $w - 1 = O(n^{-1})$. We investigate the range with $w - 1 = O(n^{-3/2})$ in section 5. The scale $w - 1 = O(n^{-1})$ with $w \uparrow 1$ is studied in section 6, and in section 7 we consider $0 < w < 1$ and $n \to \infty$. In section 8 we discuss the asymptotic matching range between the limits in section 6 and section 7. This will be key to obtaining the exponential growth rate of the total number of trees of path length p, and to obtaining the mean and variance of the Gaussian distribution of the number of nodes in such trees. Finally, in section 9 we provide some numerical studies.

2. SUMMARY OF RESULTS FOR GENERAL *t*-ARY TREES

2.1 Generating Function, Moments and Distribution

Let $g(n, p)$ denote the number of t-ary trees with n nodes and total path length p. The recurrence relation satisfied by this function is

$$g(n+1, p) = \sum_{k_1+k_2+\cdots+k_t=n} \left\{ \sum_{s_1+s_2+\cdots+s_t=p-n} \left[\prod_{i=1}^{t} g(k_i, s_i) \right] \right\}, \quad n \geq 0,$$

with the boundary conditions

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

The generating function

$$G_n(w) = \sum_{p=0}^{\infty} g(n,p) w^p,$$

thus satisfies

$$G_{n+1}(w) = w^n \sum_{k_1+k_2+\cdots+k_t=n} \left[\prod_{i=1}^{t} G_{k_i}(w) \right], \quad n \geq 0, \quad G_0(w) = 1. \tag{7}$$

The double generation function

$$G(z,w) = \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} g(n,p) w^p z^n = \sum_{n=0}^{\infty} z^n G_n(w), \tag{8}$$

satisfies the functional equation

$$G(z,w) = 1 + zG^t(zw, w). \tag{9}$$

We shall mostly analyze (7), and compute approximations to $G_n(w)$ for $n \to \infty$ and various ranges of w. Then asymptotic results for $g(n, p)$ will follow by evaluating asymptotically the Cauchy integral

$$g(n,p) = \frac{1}{2\pi i} \int_C G_n(w) w^{-p-1} dw. \tag{10}$$

Here C is any closed contour about the origin in the w-plane. When $w = 1$, we let

$$a(z) \equiv G(z,1) = \sum_{n=0}^{\infty} G_n(1) z^n. \tag{11}$$

It is well known[1] that the total number of t-ary trees with n nodes is

$$\sum_{p=0}^{\infty} g(n,p) = \frac{1}{2\pi i} \int_C \frac{a(z)}{z^{n+1}} dz = G_n(1) = \frac{1}{(t-1)n+1} \binom{tn}{n}, \tag{12}$$

which is the generalized Catalan number.

When $t = 2$, we can easily solve (9) with $w = 1$, to obtain

$$a(z) \equiv G(z, 1; t = 2) = \frac{1}{2z}(1 - \sqrt{1 - 4z}).$$

When $t = 3$ with $w = 1$, we get two different expressions for $a(z)$:

$$a(z) \equiv G(z, 1; t = 3) = \frac{2\sqrt{3}}{3\sqrt{z}} \sin\left[\frac{1}{3}\arcsin\left(\frac{3\sqrt{3}}{2}\sqrt{z}\right)\right],$$

or

$$\begin{aligned} &G(z, 1; t = 3) = \\ &-\frac{1}{12z}\left[\left(\left(12\sqrt{3}\sqrt{27 - \frac{4}{z}} - 108\right)z^2\right]^{1/3} - \left[\left(12\sqrt{3}\sqrt{27 - \frac{4}{z}} - 108\right)z^2\right]^{-1/3}\right. \\ &- i\sqrt{3}\left\{\frac{1}{12z}\left[\left(12\sqrt{3}\sqrt{27 - \frac{4}{z}} - 108\right)z^2\right]^{1/3} - \left[\left(12\sqrt{3}\sqrt{27 - \frac{4}{z}} - 108\right)z^2\right]^{-1/3}\right\}. \end{aligned}$$

When $t > 3$, it is difficult to obtain even $G(z, 1)$ explicitly. However we can get a series expansion of $a(z)$, by using the Lagrange inversion theorem, since

$$z = \frac{G(z, 1) - 1}{G^t(z, 1)} = \frac{\bar{a}_{\geq 1}(z)}{(\bar{a}_{\geq 1}(z) + 1)^t}, \quad \bar{a}_{\geq 1}(z) = G(z, 1) - 1 = a(z) - 1.$$

Thus we obtain the generalized Catalan numbers,

$$[z^n]\bar{a}_{\geq 1}(z) = \frac{1}{n}[\bar{a}_{\geq 1}(z)^{n-1}](\bar{a}_{\geq 1}(z) + 1)^{tn} = \frac{1}{n}\binom{tn}{n-1} = \frac{1}{(t-1)n + 1}\binom{tn}{n}.$$

Here $[z^n]$ is an operator which extracts the coefficient of z^n in the expression that follows it.

We expand $G_n(w)$ and $G(z, w)$ about $w = 1$, and define $a_n,\ b_n,\ c_n$ as the coefficients in the expansion of $G_n(w)$,

$$G_n(w) = a_n + b_n(w - 1) + \frac{c_n}{2}(w - 1)^2 + O((w - 1)^3), \tag{13}$$

and $a(z),\ b(z),\ c(z)$ as the coefficients in the expansion of $G(z, w)$,

$$G(z, w) = a(z) + b(z)(w - 1) + \frac{c(z)}{2}(w - 1)^2 + O((w - 1)^3). \tag{14}$$

By differentiating (9) with respect to w and setting $w = 1$ we obtain

$$b(z) = G_w(z, 1) = \frac{tz^2 G^{t-1}(z, 1) G_z(z, 1)}{1 - tzG^{t-1}(z, 1)} = \frac{tz^2 a^{t-1}(z) a'(z)}{1 - tza^{t-1}(z)} = \frac{tz^2 a'^2(z)}{a(z)}, \tag{15}$$

and correspondingly

$$b_n = G_n'(1) = \sum_{p=1}^{\infty} pg(n, p),\ n \geq 0. \tag{16}$$

From (11) we obtain

$$a'(z) = \sum_{n=1}^{\infty} nG_n(1)z^{n-1} = \sum_{n=0}^{\infty}(n + 1)G_{n+1}(1)z^n. \tag{17}$$

Also, from the convolution of the generating function, we have

$$a^{t-1}(z) = \sum_{n=0}^{\infty} \frac{1}{n + 1}\binom{nt + t - 2}{n} z^n. \tag{18}$$

Using (11), (17) and (18) in (15), we obtain a recurrence relation for the b_n

$$b_{n+2} = t\left(\sum_{i=0}^{n}(n-i+1)a_{n-i+1}S_i + \sum_{i=0}^{n+1}S_i b_{n+1-i}\right), \tag{19}$$

where $S_i = \frac{1}{i+1}\binom{it+t-2}{i}$, $b_0 = 0$, $b_1 = 0$, and $a_n = G_n(1)$ as in (12). The average total path length is thus b_n/a_n. Similarly we can obtain higher order moments, and in particular we obtain $c(z)$ and c_n from

$$\begin{aligned} c(z) &= G_{ww}(z,1) \\ &= \frac{\left[t(t-1)za^{t-2}(z)(za'(z)+b(z))^2 + tz^2a^{t-1}(z)\left[za''(z)+2b'(z)\right]\right]}{1-tza^{t-1}(z)} \\ &= \frac{a'(z)}{a^2(z)}\Big[2t(t-1)z^2b(z)a'(z) + t(t-1)z^3(a'(z))^2 + t(t-1)zb^2(z) \\ &\quad + tz^3a(z)a''(z) + 2tz^2a(z)b'(z)\Big], \end{aligned} \tag{20}$$

and

$$c_n = G_n''(1) = \sum_{p=2}^{\infty} p(p-1)g(n,p),\ n \ge 0. \tag{21}$$

By expanding $a(z)$ about $z = z_0$, which is the singularity closest to $z = 0$, and using (9) with $w = 1$, we find that

$$\begin{aligned} &a(z) = y_0 + y_1(z_0-z)^{1/2} + y_2(z_0-z) + O((z_0-z)^{3/2}), \\ &z_0 = \frac{(t-1)^{t-1}}{t^t},\ \ y_0 = \frac{t}{t-1},\ \ y_1 = -\frac{\sqrt{2}t^{(t+1)/2}}{(t-1)^{(2+t)/2}},\ \ y_2 = \frac{2t^t(t+1)}{3(t-1)^{t+1}}. \end{aligned} \tag{22}$$

The above can be obtained by expanding (18) directly, but it is easier to use the relation $a(z) = 1 + za^t(z)$ and expand it about $z = z_0$. This shows that $y_0 = 1 + z_0y_0^t$ and then the higher order coefficients in (22), such as y_1 and y_2, can be similarly computed. By using (22) in (15) and (20) we then obtain the expansions of $b(z)$ and $c(z)$ about $z = z_0$ as

$$\begin{aligned} b(z) &= \frac{b_0}{z_0-z} + \frac{b_1}{\sqrt{z_0-z}} + O(1), \\ b_0 &= \frac{t}{2(t-1)^2}z_0,\ \ b_1 = \frac{-\sqrt{2}t(4t+1)}{6(t-1)^{5/2}}\sqrt{z_0}, \\ c(z) &= \frac{c_0}{(z_0-z)^{5/2}} + \frac{c_1}{(z_0-z)^2} + O((z_0-z)^{-3/2}), \\ c_0 &= \frac{5\sqrt{2}\,t^{3/2}}{8(t-1)^{5/2}}z_0^{5/2},\ \ c_1 = \frac{-(17t+5)t}{12(t-1)^3}z_0^2, \end{aligned} \tag{23}$$

where z_0 is defined in (22). Then by using singularity analysis we can infer the behavior of the coefficients in (13) for $n \to \infty$:

$$\begin{aligned}
a_n &= \frac{1}{z_0^n n^{3/2}}[m_0 + O(n^{-1})], \quad m_0 = \frac{\sqrt{t}}{\sqrt{2\pi}(t-1)^{3/2}},\\
b_n &= \frac{1}{z_0^n}[m_1 + \bar{m}_1 \frac{1}{\sqrt{n}} + O(n^{-1})],\\
m_1 &= \frac{t}{2(t-1)^2}, \quad \bar{m}_1 = \frac{-(4t+1)\sqrt{2t}}{6\sqrt{\pi}(t-1)^{5/2}};\\
c_n &= \frac{1}{z_0^n}[m_2 n^{3/2} + \bar{m}_2 n + O(\sqrt{n})],\\
m_2 &= \frac{5t\sqrt{t}}{3\sqrt{2\pi}(t-1)^{5/2}}, \quad \bar{m}_2 = \frac{-(17t+5)t}{12(t-1)^3}.
\end{aligned} \tag{24}$$

It can be easily shown that for each k

$$\frac{G_n^{(k+1)}(1)}{G_n^{(k)}(1)} = O(n^{3/2}),\ n \to \infty. \tag{25}$$

We define the distribution of the total path length L_n as

$$Pr\{L_n = p\} = \frac{g(n,p)}{\sum_{p=0}^{\infty} g(n,p)} = \frac{g(n,p)}{a_n}. \tag{26}$$

From (24) it follows that

$$\begin{aligned}
\mathbf{E}[L_n] &= \frac{\sqrt{t\pi}}{\sqrt{2(t-1)}} n^{3/2} + O(n),\\
\mathbf{Var}[L_n] &= \frac{t}{t-1}\left(\frac{5}{3} - \frac{\pi}{2}\right) n^3 + O(n^{5/2}).
\end{aligned} \tag{27}$$

A more complicated problem is to study the distribution of the number of nodes in trees for a given path length p, i.e., for the ensemble $\mathcal{T}_p$. We define the random variable N_p to be the number of nodes in a tree generated uniformly from $\mathcal{T}_p$. Its distribution is computed from $g(n,p)$ via

$$Pr\{N_p = n\} = \frac{g(n,p)}{\sum_{n=0}^{\infty} g(n,p)}. \tag{28}$$

We will compute $Pr\{N_p = n\}$ asymptotically, and we will also obtain the asymptotic structure of $g(n,p)$ for various ranges of n and p. The sums in (26) and (28) are finite, because $g(n,p)$ is non-zero only in the range

$$\sum_{J=2}^{n} \lfloor \log_t((t-1)J) \rfloor = p_{min}(n) \le p \le p_{max}(n) = \binom{n}{2}. \tag{29}$$

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

$$\begin{aligned}
n_{min}(p) &= min\{n : \binom{n}{2} \ge p\},\\
n_{max}(p) &= max\{n : \sum_{J=2}^{n} \lfloor \log_t((t-1)J) \rfloor \le p\}.
\end{aligned}$$

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

We now summarize the main results. The derivations are quite complicated and will be presented in later sections. Our first main result is for the cardinality of $\mathcal{T}_p$, as given below.

2.2 Result 1: Total Number of Trees

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

$$|\mathcal{T}_p| = \sum_{n=0}^{\infty} g(n,p) = \frac{1}{(\log_t p)\sqrt{p\pi}} \exp\left[\frac{p \ln\left(\frac{1}{z_0}\right)}{\log_t p}\left(1 - \frac{3}{2}A_0 \frac{\ln(t)}{a^{1/3}}(\ln p)^{-2/3} + \frac{M}{\ln p} + O((\ln p)^{-4/3})\right)\right], \quad (30)$$

where

$$M = (1 + A_1 \ln(t)) \ln\ln p - \ln\ln t + (k_2 - A_1 \ln a) \ln t, \quad (31)$$

$$A_0 = \frac{2}{3}\left(\frac{2t}{t-1}\right)^{1/3} |r_0|, \quad A_1 = \frac{1}{\ln t} - \frac{t+1}{9(t-1)},$$
$$r_0 = max\{z : Ai(z) = 0\} = -2.3381\cdots, \; a = \ln(t) \ln\left(\frac{1}{z_0}\right) = \ln(t) \ln\left(\frac{t^t}{(t-1)^{t-1}}\right). \quad (32)$$

The constant $k_2 = k_2(t)$ can be obtained by numerically solving a nonlinear integral equation. For example, when $t = 2$, $k_2 \approx 3.696$, and when $t = 3$, $k_2 \approx 2.727$. Here $Ai(\cdot)$ is the Airy function, defined as a solution of $f'' - zf = 0$ that decays as $z \to \infty$.

The exponential growth rate of the total number of trees of path length p is thus

$$\ln\left[\sum_n g(n,p)\right] \sim \frac{p \ln(\frac{1}{z_0})}{\log_t p} = \frac{p \ln(t) \ln\left(\frac{t^t}{(t-1)^{t-1}}\right)}{\ln p} = O(\frac{p}{\ln p}). \quad (33)$$

This agrees with the recent result of Seroussi[24], after we reconcile our notations. The correction term to (33) (cf. (30)) involves the maximal root of the Airy function and is of the order $O(p(\ln p)^{-5/3})$. We will indicate how to obtain further terms in the asymptotic series in (30), which involves powers of $(\ln p)^{-1/3}$, along with some $\ln(\ln p)$ factors.

Next we discuss the random variable N_p. This gives the probability that a t-ary tree with total path length p will have n nodes. We define the mean and variance by

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

2.3 Result 2: Mean and Variance

For the mean and variance of N_p we have the expansions

$$\mathcal{N}(p) = \frac{p}{\ln p} \ln t \left[1 - \frac{A_0 \ln t}{a^{1/3}(\ln p)^{2/3}} + \frac{M - A_1 \ln(t)}{\ln p} + O((\ln p)^{-4/3})\right], \quad (34)$$

and

$$\mathcal{V}(p) = \frac{p \ln^2(t) A_0}{3(\ln p)^{5/3} \ln\left(\frac{t^t}{(t-1)^{t-1}}\right) a^{1/3}} \left[1 - \frac{3A_1}{A_0}\left(\frac{a}{\ln p}\right)^{1/3} + O((\ln p)^{-2/3})\right]. \quad (35)$$

Furthermore, the limiting distribution of N_p is Gaussian, i.e.,

$$Pr\{N_p = n\} = \frac{g(n,p)}{\sum_{n=0}^{\infty} g(n,p)} \sim \frac{1}{\sqrt{2\pi\mathcal{V}(p)}} \exp\left[-\frac{(n-\mathcal{N}(p))^2}{2\mathcal{V}(p)}\right], \tag{36}$$

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

The most important scale for studying t-ary trees with a given number n of nodes, is $p = O(n^{3/2})$. Our results show that when studying trees with a given path length p the most important range of n is

$$n = \frac{p}{\log_t p} + O\left(p(\ln p)^{-5/3}\right),$$

or

$$p = n\log_t n + O\left(n(\ln n)^{1/3}\right).$$

This is relatively close to the upper limit $n_{max}(p)$ (or the lower limit $p_{min}(n)$) of the support of $g(n,p)$. Thus given a large p the number of nodes tends to be close to the maximum number possible, which means that the tree is close to being balanced.

Results 1 and 2 follow from our detailed analysis of $G_n(w)$ for various ranges of w, and of $g(n,p)$ for various ranges of n and p. Below we list these, first for $G_n(w)$.

2.4 Result 3: Asymptotics of the Generating Function

Consider t-ary trees with path length equal to p. Let $G_n(w)$ be its generating function, which satisfies equation (7). Then for $n \to \infty$ we have the following asymptotic expansions.

(a) Far Right Region : $n \to \infty,\ w > 1$

$$G_n(w) \sim w^{\binom{n}{2}} t^{n-1} G_*(w), \tag{37}$$

where $G_*(w)$ satisfies

$$G_*(w) = 1 + \frac{t-1}{2t}\frac{1}{w} + \frac{t-1}{t}\frac{1}{w^2} + O(w^{-3}), \quad w \to \infty, \tag{38}$$

$$G_*(w) \sim d_1\sqrt{w-1}\ \exp\left(\frac{d_0}{w-1}\right), \quad w \to 1^+, \tag{39}$$

$$\begin{aligned} d_0 &= (t-1)\int_0^{\ln(\frac{t}{t-1})} \frac{x}{e^x - 1}\,dx = -(t-1)\int_0^{\infty} \ln\left(1 - \frac{e^{-x}}{t}\right) dx, \\ d_1 &= \frac{t\sqrt{2}}{\sqrt{\pi}(t-1)} e^{d_0/2}. \end{aligned} \tag{40}$$

The numerical computation of $G_*(w)$, for $1 < |w| < \infty$, is discussed in sections 3 and 9.

(b) Right Region : $w = 1 + \beta/n,\ 0 < \beta < \infty$

$$G_n(w) \sim \sqrt{\frac{\beta}{n}}\,\hat{g}(\beta)\ \exp[n\Phi(\beta)], \tag{41}$$

where

$$\Phi(\beta) = \ln(t) + \frac{\beta}{2} + \phi(\beta), \quad \phi(\beta) \equiv -\frac{(t-1)}{\beta}\int_0^{\beta} \ln\left(1 - \frac{e^{-x}}{t}\right) dx,$$

$$\hat{g}(\beta) = \frac{t\sqrt{2t}}{\sqrt{\pi}(t-1)} e^{-\beta^2/4} e^{-\beta/2} \left(\frac{1-e^{-\beta}}{t-e^{-\beta}}\right)^{3/2} \exp\left[\frac{1}{2}\beta\phi(\beta) + \frac{t-1}{2}\beta\ln\left(1 - \frac{e^{-\beta}}{t}\right)\right].$$

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

$$
\begin{aligned}
G_n(w) &= \frac{1}{(t-1)n+1}\binom{tn}{n} + \frac{1}{z_0^n n^{3/2}}\left[C(a) + \frac{1}{\sqrt{n}}C^{(1)}(a) + O(n^{-1})\right] \\
&= \frac{1}{z_0^n n^{3/2}}\left[m_0 + C(a) + \frac{1}{\sqrt{n}}C^{(1)}(a) + O(n^{-1})\right],\ \ m_0 = \frac{\sqrt{t}}{\sqrt{2\pi}(t-1)^{3/2}}, \\
C(a) &= (-a)\bar{D}((-a)^{2/3}) = Y^{3/2}\bar{D}(Y), \quad Y = (-a)^{2/3}, \quad a < 0, \\
\bar{D}(Y) &= \frac{1}{2\pi i}\int_{Br} e^{sY}\left[\frac{1}{t-1}\left(\frac{2t}{t-1}\right)^{1/2}\sqrt{s} + \frac{1}{t-1}\left(\frac{2t}{t-1}\right)^{2/3}\frac{Ai'\left((\frac{t-1}{2t})^{1/3}s\right)}{Ai\left((\frac{t-1}{2t})^{1/3}s\right)}\right]ds.
\end{aligned}
\tag{42}
$$

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. Another expression for the leading term is, for the range of $a = -Y^{3/2} < 0$,

$$
\begin{aligned}
G_n(w) &\sim \frac{1}{z_0 n^{3/2}}(-a)\frac{d}{dY}\left[\frac{1}{2\pi i}\int_{Br} e^{sY}\frac{1}{s(t-1)}\left(\frac{2t}{t-1}\right)^{2/3}\frac{Ai'\left((\frac{t-1}{2t})^{1/3}s\right)}{Ai\left((\frac{t-1}{2t})^{1/3}s\right)}ds\right] \\
&= \frac{1}{z_0 n^{3/2}}\left(\frac{2t}{(t-1)^2}\right)(-a)\sum_{j=0}^{\infty}\exp\left(-|r_j|\left(\frac{2t}{t-1}\right)^{1/3}Y\right).
\end{aligned}
\tag{43}
$$

Here $z_0 = (t-1)^{t-1}t^{-t}$ is as in (22) and r_j are the roots of $Ai(z) = 0$, ordered as $0 > r_0 > r_1 > r_2 > \cdots$. The correction term $C^{(1)}(a)$ has the following integration representation for $a < 0$,

$$
\begin{aligned}
C^{(1)}(a) &= -a\bar{D}_1(Y) = \frac{Y^2}{2\pi i}\int_{Br} e^{sY}\mathcal{E}_*(s)ds, \\
\mathcal{E}_*(s) &= -\frac{(7t+1)}{6(t-1)^2}s + \frac{2t^2}{(t-1)^3}\left(\frac{h'(s)}{h(s)}\right)^2 - \frac{2t(t+1)}{3(t-1)^3h^2(s)}\int_s^{\infty}\frac{(h'(v))^3}{h(v)}dv \\
&= -\frac{(7t+1)}{6(t-1)^2}s + \frac{t(7t+1)}{3(t-1)^3}\left(\frac{h'(s)}{h(s)}\right)^2 + \frac{2t(t+1)}{3(t-1)^3}\left(\frac{h'(s)}{h(s)}\right)^2\ln[(h(s)] \\
&\quad - \frac{(t+1)}{3(t-1)^2}s\ln[h(s)] - \frac{(t+1)}{3(t-1)^2h^2(s)}\int_s^{\infty}h^2(v)\ln[h(v)]dv, \\
h(s) &= Ai\left((\frac{t-1}{2t})^{1/3}s\right).
\end{aligned}
\tag{44}
$$

For $a > 0$ (i.e., $w > 1$) we set $a = y^{3/2}$ with $y > 0$, and the leading term is

$$
\begin{aligned}
G_n(w) \sim &\frac{1}{z_0^n n^{3/2}}\left\{\frac{ta}{2(t-1)^2\pi^2}\left(\frac{t-1}{2t}\right)^{1/3}\int_0^{\infty}\frac{e^{\omega\tau y}}{h(\omega\tau)h(\omega^2\tau)}d\tau\right. \\
&\left. - \frac{2ta}{(t-1)^2\pi}\int_0^{\infty}\Re\left[e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega^2\tau y}\right]d\tau\right\},
\end{aligned}
\tag{45}
$$

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

(d) Left Region : $w = 1 - \gamma/n,\ 0 < \gamma < \infty$

$$
\begin{aligned}
G_n(w) &\sim \frac{1}{z_0^n n}\exp[v_0 n^{1/3}\gamma^{2/3} + v_1\gamma\ln(n)]F_0(\gamma), \\
F_0(\gamma) &= \frac{2t}{(t-1)^2}\gamma F_1(\gamma), \\
v_0 &= \left(\frac{2t}{t-1}\right)^{1/3}r_0 = -\left(\frac{2t}{t-1}\right)^{1/3}|r_0|, \quad v_1 = -\frac{(t+1)}{9(t-1)},
\end{aligned}
\tag{46}
$$

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

$$\frac{(e^\gamma-1)}{\gamma}F_1(\gamma)=\int_0^1 F_1(\gamma x_1)F_1(\gamma-\gamma x_1)e^{\nu_1\gamma H(\vec{x}(2))}dx_1+\sum_{i=3}^{t}\binom{t}{i}\frac{1}{t}\left(\frac{2}{t-1}\right)^{i-1}\gamma^{i-2}$$
$$\times\int_0^1\dots\int_0^{1-x_1-\dots-x_{(i-2)}}e^{\nu_1\gamma H(\vec{x}(i))}\left[\prod_{j=1}^{i}F_1(\gamma x_j)\right]dx_{(i-1)}\dots dx_1, \tag{47}$$
$$H(\vec{x}(i))=\sum_{j=1}^{i}x_j\ln(x_j),\quad \sum_{j=1}^{i}x_j=1.$$

For example, when $t=3$, $F_1(\gamma)$ satisfies

$$\frac{(e^\gamma-1)}{\gamma}F_1(\gamma)=\int_0^1 F_1(x_1\gamma)F_1((1-x_1)\gamma)e^{-2\gamma H(\vec{x}(2))/9}dx_1$$
$$+\frac{\gamma}{3}\int_0^1\int_0^{1-x_1}F_1(x_1\gamma)F_1(x_2\gamma)F_1((1-x_1-x_2)\gamma)e^{-2\gamma H(\vec{x}(3))/9}dx_2dx_1.$$

For $\gamma\to 0^+$, F_1 behaves as

$$F_1(\gamma)=1+\alpha_0\gamma\ln(\gamma)+\alpha_1\gamma+o(\gamma),$$

where

$$\alpha_0=-\frac{2(t+1)}{9(t-1)},\quad \alpha_1=\frac{(t-1)^2\kappa}{2t}+\frac{(t+1)}{3(t-1)}(1-\gamma_E),$$
$$\kappa\equiv\frac{2t(t+1)}{3(t-1)^3}\ln[h'(s_0)]+\frac{t(7t+1)}{3(t-1)^3}-\frac{(t+1)}{3(t-1)^2[h'(s_0)]^2}\int_{s_0}^{\infty}h^2(v)\ln[h(v)]dv \tag{48}$$
$$s_0=\left(\frac{2t}{t-1}\right)^{1/3}r_0,\quad \gamma_E=\text{Euler's constant}=0.5772\cdots.$$

For $\gamma\to\infty$, we obtain

$$F_1(\gamma)\sim\frac{(t-1)t^{1/(t-1)}}{2\sqrt{2\pi\ln(t)}}\frac{e^{k_2}}{\sqrt{\gamma}}\exp\left[\left(\frac{t+1}{9(t-1)}-\frac{1}{\ln(t)}\right)\gamma\ln(\gamma)\right], \tag{49}$$

where $k_2=k_2(t)$ can be obtained numerically (cf. sections 6 and 9).

(e) Far Left Region : $n\to\infty,\ 0<w<1$

$$G_n(w)\sim e^{n\log_t(n)\ln(w)}n^{\frac{\log_t(w)}{t-1}}(2\pi n)^{-1/2}w^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln(t)}}\exp\left(\frac{B_1^*(w,n)}{t-1}\right),$$
$$\times\exp\left\{\left(n+\frac{1}{t-1}\right)\left[g(w)+B_0^*(w,n)\right]\right\}\sqrt{-\log_t(w)-2B_1^*(w,n)-B_2^*(w,n)}. \tag{50}$$

Here the B_j^* have the Fourier series

$$B_0^*(w,n)=\sum_{\substack{k=-\infty\\k\neq0}}^{\infty}g_k(w)e^{2\pi ik\log_t(n)},$$
$$B_1^*(w,n)=\frac{2\pi i}{\ln(t)}\sum_{k=-\infty}^{\infty}k\,g_k(w)e^{2\pi ik\log_t(n)}, \tag{51}$$
$$B_2^*(w,n)=\frac{2\pi i}{\ln(t)}\sum_{k=-\infty}^{\infty}\left(\frac{2\pi i}{\ln(t)}k^2-k\right)g_k(w)e^{2\pi ik\log_t(n)}.$$

We cannot give the $g_k(w)$ explicitly, but can obtain $g(w)$ as $w \to 1^-$, as

$$\begin{aligned} g(w) &= \ln\left(\frac{1}{z_0}\right) + \left(\frac{2t}{t-1}\right)^{1/3} r_0(1-w)^{2/3} \\ &+ \left[\frac{1}{\ln(t)} - \frac{(t+1)}{9(t-1)}\right](w-1)\ln(1-w) + k_2(1-w) + o(1-w). \end{aligned} \tag{52}$$

The numerical computation of $g(w)$ is discussed in sections 7 and 9. The sum in B_0^* omits the $k = 0$ term which would correspond to $g(w)$, and the Fourier coefficients satisfy $g_k(w) = o(w-1)$ as $w \to 1^-$. Numerical studies show that the $g_k(\cdot)$ are very small unless w is very small. Thus we can neglect the B_j^* for $j = 0, 1, 2$, and we use the following approximation

$$\begin{aligned} G_n(w) &\approx w^{n\log_t(n)} e^{(n+\frac{1}{t-1})g(w)} n^{\frac{\log_t(w)}{t-1} - \frac{1}{2}} \\ &\times w^{\frac{t}{(t-1)^2} + \frac{1}{(t-1)\ln(t)}} \sqrt{\frac{-\log_t(w)}{2\pi}}, \end{aligned} \tag{53}$$

but $g(w)$ must still be found numerically for $w < 1$.

To get the main asymptotic results for the number of trees of a given path length, the important range is the asymptotic matching region between the left and the far left regions, corresponding $w \to 1^-$ but $n(1-w) = \gamma \to +\infty$. Because we have explicit analytic results for $g(w)$ as $w \to 1^-$ and $g_k(w) \to 0$ for $k \neq 0$, we were able to obtain the explicit expressions in Results 1 and 2. The central region, where $w - 1 = O(n^{-3/2})$, is the most important for the distribution of the path length in trees with n ($\to \infty$) nodes, and the leading term corresponds to an Airy distribution. Note that this Airy distribution persists even in the limit $t \to \infty$.

Next we give results for $g(n, p)$ for p and $n \to \infty$, and the main results are summarized as items (A)-(E). Going from (A) to (E) corresponds to increasing n or decreasing p.

2.5 Result 4: Expansion of the Numbers of Trees by Nodes and Path Length

Let $g(n, p)$ denote the number of t-ary trees built over n nodes with path length p. For $n, p \to \infty$, we have the following asymptotic approximations.

Region (A): $n \to \infty$ *with* $p = n(n-1)/2 - O(1)$

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

$$g(n, p) \sim \frac{t^{n-1}}{2\pi i} \int_C w^{L-1} G_*(w)\, dw. \tag{54}$$

Here C is a closed contour with $|w| > 1$ and $G_*(w)$ is the same as in (a) of Result 3.

Region (B): $n \to \infty$ *with* $p = O(n^2)$

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

$$\begin{aligned} g(n, p) &\sim \frac{t^n}{n^2} \frac{t\sqrt{t}}{\pi(t-1)} \beta_* e^{-\beta_*/2} \left(\frac{1 - e^{-\beta_*}}{t - e^{-\beta_*}}\right)^{3/2} \left[1 - 2\Lambda - \frac{(t-1)}{te^{\beta_*} - 1}\right]^{-1/2} \\ &\times \exp\left\{n\left[\beta_*(1 - 2\Lambda) - (t-1)\ln(1 - \frac{e^{-\beta_*}}{t})\right]\right\}, \end{aligned} \tag{55}$$

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

$$\beta_*^2\left(\frac{1}{2}-\Lambda\right)-(t-1)\beta_*\ln\left(1-\frac{e^{-\beta_*}}{t}\right)=(t-1)\int_{-\ln\left(1-\frac{e^{-\beta_*}}{t}\right)}^{\ln\left(\frac{t}{t-1}\right)}\frac{x}{e^x-1}\,dx. \tag{56}$$

Region (C): $n\to\infty$ *with* $p=O(n^{3/2})$

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

$$\begin{aligned} g(n,p)\sim &-\frac{1}{z_0^n n^3}\left(\frac{1}{3\Omega}\right)^{1/3}\sum_{j=0}^{\infty}\left\{\left[\frac{28t(\frac{2t}{t-1})^{2/3}r_j^2}{9(t-1)^2\Omega^3}+\frac{32t(\frac{2t}{t-1})^{5/3}r_j^5}{81(t-1)^2\Omega^5}\right]Ai\left(\frac{(\frac{2t}{t-1})^{2/3}r_j^2}{(3\Omega)^{4/3}}\right)\right.\\ &+\left(\frac{1}{3\Omega}\right)^{1/3}\left.\left[\frac{20t(\frac{2t}{t-1})^{1/3}r_j}{3(t-1)^2\Omega^2}+\frac{32t(\frac{2t}{t-1})^{4/3}r_j^4}{27(t-1)^2\Omega^4}\right]Ai'\left(\frac{(\frac{2t}{t-1})^{2/3}r_j^2}{(3\Omega)^{4/3}}\right)\right\}\\ &\times\exp\left(-\frac{4t|r_j|^3}{27(t-1)\Omega^2}\right). \end{aligned} \tag{57}$$

Again, here $r_j<0$ are the negative roots of the Airy function.

Region (D): $n\to\infty$ *with* $p=O(n^{4/3})$

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

$$\begin{aligned} g(n,p)&\sim\frac{1}{z_0^n n^{13/6}}n^{-\frac{t+1}{9(t-1)}\gamma_*}\left(\frac{64t^2\sqrt{2t}}{81(t-1)^{7/2}\sqrt{\pi}}\right)\frac{|r_0|^{9/2}}{\Theta^5}F_1(\gamma_*)\exp\left(-\frac{8tn^{1/3}|r_0|^3}{27(t-1)\Theta^2}\right),\\ \gamma_*&=\frac{16t}{27(t-1)}\frac{|r_0|^3}{\Theta^3}. \end{aligned} \tag{58}$$

Here $F_1(\cdot)$ is defined as the solution to (47).

Region (E): $n\to\infty$ *with* $p=n\log_t n+O(n)$

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

$$\begin{aligned} g(n,p)\approx&\frac{n^{\frac{\log_t(w_*)}{t-1}}}{2\pi n}\frac{w_*^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln(t)}}}{\sqrt{\alpha+w_*^2g''(w_*)}}e^{\frac{g(w_*)}{t-1}}\sqrt{-\log_t w_*}\\ &\times\exp\left[ng(w_*)-n\alpha\ln(w_*)\right]. \end{aligned} \tag{59}$$

Here $w_*=w_*(\alpha)$ is the solution of the equation $w_*g'(w_*)=\alpha$.

Dividing (57) by the expansion of a_n as $n\to\infty$ (cf. (24)) recovers the Airy distribution. In obtaining (59) we used (50) in (10) and neglected the non-constant terms in the Fourier series, which are numerically small. The integral was then evaluated by the saddle point method (cf. [9]). A refined approximation, which uses the non-constant terms, can be obtained by using (50) in (10). By using the correction term $C^{(1)}(a)$ in (42) and (44) in asymptotically inverting (10), we can obtain a correction term of order $O(n^{-1/2})$ to the Airy distribution in (57). Again, we only have partial information about the expansion for case (E), as we know $g(w)$ analytically only as $w\to 1$. Our approximation(s) to $g(n,p)$ in items (A), (D) and (E) involve the unknown functions $G_*(w), F_1(\gamma)$ and $g(w)$, whose numerical computation we discuss in section 9.

To get more insight into the structure of $g(n,p)$, we give formulas that apply in the asymptotic matching regions between the various scales. The results are summarized below, with the notation (AB) denoting the asymptotic matching region between the scales in (A) and (B), and so on. The result (AB) can be obtained by either expanding (54) as $L\to\infty$, or by expanding (55) as $\Lambda\to\left(\frac{1}{2}\right)^-$.

2.6 Result 5: Expansions in the Matching Regions

We summarize the expressions in the asymptotic matching regions.

Matching Region (AB) :

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

$$g(n,p) \sim \frac{t^n}{n^2\pi} \frac{\sqrt{2d_0}}{1-2\Lambda} \frac{1}{(t-1)} \exp\left[n\sqrt{2d_0(1-2\Lambda)} - \frac{1}{2}\sqrt{\frac{2d_0}{1-2\Lambda}}\right]. \tag{60}$$

Here d_0 is defined as in (40).

Matching Region (BC) :

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

$$\begin{aligned} g(n,p) &\sim \frac{1}{z_0^n n^2 \pi} \frac{9\sqrt{3}}{t\sqrt{(t-1)}} \Lambda^2 \exp\left(-\frac{3(t-1)}{2t} n\Lambda^2\right) \\ &= \frac{1}{z_0^n n^3 \pi} \frac{9\sqrt{3}}{t\sqrt{(t-1)}} \Omega^2 \exp\left(-\frac{3(t-1)}{2t}\Omega^2\right). \end{aligned} \tag{61}$$

Matching Region (CD) :

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

$$\begin{aligned} g(n,p) &\sim \frac{1}{z_0^n n^{13/6}} \frac{|r_0|^{9/2}}{\Theta^5} \frac{64t^2\sqrt{2t}}{81(t-1)^{7/2}\sqrt{\pi}} \exp\left[-\frac{8t}{27(t-1)} n^{1/3} \frac{|r_0|^3}{\Theta^2}\right] \\ &= \frac{1}{z_0^n n^3} \frac{|r_0|^{9/2}}{\Omega^5} \frac{64t^2\sqrt{2t}}{81(t-1)^{7/2}\sqrt{\pi}} \exp\left[-\frac{8t}{27(t-1)} \frac{|r_0|^3}{\Omega^2}\right]. \end{aligned} \tag{62}$$

Matching Region (DE) :

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

$$\begin{aligned} g(n,p) \sim \frac{1}{n^{13/6}} \frac{|r_0|^3}{\Theta^{7/2}} \frac{t^{\frac{1}{t-1}+2}}{\pi\sqrt{\ln(t)}} \frac{8\sqrt{3}}{27(t-1)^2} \exp\Bigg[& n\ln\left(\frac{1}{z_0}\right) - \frac{t+1}{9(t-1)}\gamma_* \ln(n) \\ &+ \left(\frac{t+1}{9(t-1)} - \frac{1}{\ln(t)}\right)\gamma_* \ln(\gamma_*) + k_2\gamma_* - \frac{1}{3}\left(\frac{2t}{t-1}\right)^{1/3} n^{1/3} |r_0|^{1/3} \gamma_*^{2/3}\Bigg]. \end{aligned} \tag{63}$$

Here γ_* is defined in (58), and we note that

$$\frac{1}{3}\left(\frac{2t}{t-1}\right)^{1/3} n^{1/3} |r_0|^{1/3} \gamma_*^{2/3} = \frac{8t}{27(t-1)} n^{1/3} \frac{|r_0|^3}{\Theta^2}.$$

In section 8 we will show that the asymptotic matching region (DE) leads to the Gaussian distribution in (36). In each of the above four matching regions, the results are completely explicit functions of n and p. The results in (BC) and (CD) give, respectively, the right and left tails of the Airy distribution in (57), which has support in the range $0 < \Omega < \infty$.

3. FAR RIGHT REGION

We consider (7) for a fixed $w > 1$ and $n \to \infty$. We let

$$G_n(w) = w^{\binom{n}{2}} t^{n-1} \bar{G}_n(w), \tag{64}$$

and find that (7) becomes

$$\begin{aligned}\bar{G}_{n+1}(w) &= \frac{1}{t^t} \sum_{i_1+i_2+\cdots+i_t=n} w^{\left(-\sum_{k,l\in[1,t],\,k\neq l} i_k i_l\right)} \prod_{j=1}^{t} \bar{G}_{i_j} \\ &= \frac{1}{t^t}\Bigg\{ t\bar{G}_n(w)\bar{G}_0^{t-1}(w) + t(t-1)w^{1-n}\bar{G}_{n-1}(w)\bar{G}_0^{t-2}(w) \\ &\quad + tw^{4-2n}\bar{G}_{n-2}(w)\left[(t-1)\bar{G}_2(w)\bar{G}_0^{t-2}(w) + \binom{t-1}{2}\frac{\bar{G}_1^2(w)\bar{G}_0^{t-3}(w)}{w}\right] + \cdots \Bigg\}.\end{aligned} \tag{65}$$

Using the initial values $G_0(w) = G_1(w) = 1$ and $G_2(w) = tw$, (64) leads to

$$\bar{G}_0(w) = t, \;\; \bar{G}_1(w) = 1, \;\; \bar{G}_2(w) = 1, \tag{66}$$

and (65) becomes

$$\bar{G}_{n+1}(w) = \bar{G}_n(w) + \frac{t-1}{t} w^{1-n}\bar{G}_{n-1}(w) + \frac{t-1}{t} w^{4-2n}\bar{G}_{n-2}(w) + O(w^{-3n}), \tag{67}$$

whose asymptotic solution is

$$\bar{G}_n(w) = G_*(w) + O(w^{-n}), \tag{68}$$

for some function $G_*(\cdot)$

For $w \to \infty$ and fixed $n \geq 4$, it can be established inductively from (7) that

$$G_n(w) = t^{n-1} w^{\binom{n}{2}} + \frac{(t-1)}{2} t^{n-2} w^{\binom{n}{2}-1} + t(t-1)t^{n-3} w^{\binom{n}{2}-2} + O\left(w^{\binom{n}{2}-3}\right). \tag{69}$$

Thus by comparing (69) to (64) with (68), we obtain

$$G_*(w) = 1 + \frac{(t-1)}{2t}\frac{1}{w} + \frac{(t-1)}{t}\frac{1}{w^2} + O(w^{-3}), \;\; w \to \infty. \tag{70}$$

In section 4 we use an asymptotic matching argument, to show that as $w \downarrow 1$ we have

$$G_*(w) \sim d_1 \sqrt{w-1}\, \exp\left(\frac{d_0}{w-1}\right), \quad w \to 1^+, \tag{71}$$

where the constants d_0 and d_1 are defined in (40). We have not been able to obtain $G_*(w)$ analytically except for its behaviors for $w \downarrow 1$ and $w \to \infty$. It can be obtained numerically by fixing $w > 1$ and iterating (65) until $\bar{G}_n(w)$ reaches some prescribed accuracy, for any given $t \geq 2$.

In Table 1, we give $G_*(w)$ for w in range [1.04, 5] for $t = 3$. The convergence of this procedure becomes very slow (and $G_*(w)$ becomes very large) when w exceeds 1 just slightly. This is consistent with our asymptotic analysis, that predicts another scale where $n \to \infty$ with $n(w-1)$ fixed. The numerical studies are consistent with our results in (39) and (40). As t increases from 2 to ∞, d_0 increases from $0.58224\cdots$ to 1, while d_1 decreases from $2.1350\cdots$ to 0.

We have assumed that w is *real* in our analysis thus far. However (64) is valid for complex w with $|w| > 1$ by the same arguments. We can use (65) to compute $G_*(w)$ for complex w with $|w| > 1$. In Figure 1 we plot $\Re[G_*(w)]$ and $\Im[G_*(w)]$ for $1.4 < |w| < 4$ and $t = 3$.

From (64) and (68) we can infer the behavior of $g(n, p)$ for p close to $p_{max} = \binom{n}{2}$. Setting

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

we have from (9)

$$g(n, p) \sim \frac{t^{n-1}}{2\pi i} \int_C w^{L-1} G_*(w)\, dw, \quad n \to \infty, \tag{72}$$

where C is a closed contour with $|w| > 1$. For $L = 0, 1, 2, \cdots$ we can use (70) to calculate (72), and we will obtain the asymptotic behavior of the integral as $L \to \infty$ in section 4.

Table 1: Numerical $G_*(w)$ for $t = 3$

w	$G_*(w)$
5	1.1014
4.5	1.1185
4	1.1424
3.5	1.1776
3	1.2346
2.5	1.3407
2	1.5981
1.8	1.8359
1.6	2.3357
1.4	3.8750
1.2	19.580
1.18	28.392
1.16	45.346
1.14	83.211
1.12	188.26
1.10	596.50
1.08	3419.3
1.06	64599
1.04	2.4500×10^7

4. RIGHT REGION

We consider the limit $w \downarrow 1$ and $n \to \infty$ with $w - 1 = O(n^{-1})$. We thus define β by

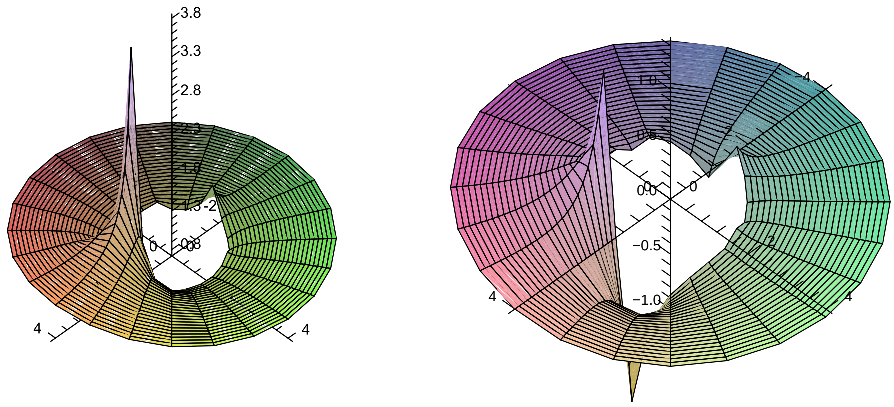
$$n(w - 1) = \beta = O(1), \quad \beta > 0, \tag{73}$$

and let

$$G_n(w) = f(\beta; n) = f(n(w - 1); n). \tag{74}$$

With (73) and (74), (7) becomes

$$f\left(\beta + \frac{\beta}{n}; n + 1\right) \sim t\left(1 + \frac{\beta}{n}\right)^n \sum_{m=0}^{n/t} \left[\prod_{k=1}^{t-1} G_{i_k}(1 + \frac{\beta}{n})\right] f\left(\beta(1 - \frac{m}{n}); n - m\right), \tag{75}$$



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

Figure 2: Plots of $\Re[G_*(w)]$ and $\Im[G_*(w)]$ for $1.4 < |w| < 4$ and $t = 3$

where

$$i_1 + i_2 + \cdots + i_{(t-1)} = m. \tag{76}$$

We make a few comments about the asymptotic relation in (75). For the scale $n \to \infty$ with $w > 1$, which we analyzed in the previous section, we found that only those terms in the multi-sum in (7) that had one $i_k = n$ and all other $i_k = 0$ were asymptotically important. On the β-scale we shall see the main contribution to the multi-sum will again come from the t corners of the lattice triangle $i_1 + i_2 + \cdots + i_t = n$, but now an infinite number of terms will contribute from the neighborhood of each corner point. For example for the corner with $i_1 = n - O(1)$, we will sum over the other indices $i_2,\ i_3,\ \cdots,\ i_t$ through all $O(1)$ values. Using the symmetry of the multi-sum, we considered only the corner where $i_t = n - m$ with $m = O(1)$, and then multiplied by the factor t, which is the first factor in the right side of (75). The exact value of the upper limit ($= n/t$) in the m-sum in (75) is asymptotically unimportant, as this limit will be extended to ∞. We also note that (75) would be exact if we omit the factor t and take the upper limit as $m = n$. But, our form of (75) is convenient for isolating the asymptotically dominant terms. For $i_k = O(1)$ (for $1 \le k \le t-1$) we wrote $G_{i_k}(w)$ as itself, while we set $i_t = n - m$ and used the β-scale form for $G_{i_t}(w)$ in (74). We can also view (75) as representing a linearization of the non-linear equation in (7).

For fixed m and $w \to 1$, the behavior of $G_m(w)$ follows from (13), since

$$G_m\left(1 + \frac{\beta}{n}\right) = a_m + b_m\frac{\beta}{n} + O(n^{-2}), \tag{77}$$

where $a_m = G_m(1)$ and b_m is as in (19). We shall make repeated use of the generating functions

$$a(z) = \sum_{m=0}^{\infty} a_m z^m, \quad b(z) \quad = \sum_{m=0}^{\infty} b_m z^m. \tag{78}$$

Also, we have

$$\left(1 + \frac{\beta}{n}\right)^n = e^{\beta}\left[1 - \frac{\beta^2}{2n} + O(n^{-2})\right], \quad n \to \infty. \tag{79}$$

For fixed $\beta > 0$ and $n \to \infty$, we assume that $f(\beta; n)$ has an expansion in the **WKB** form

$$f(\beta; n) = e^{n\Phi(\beta)} n^{-1/2}\left[g(\beta) + \frac{1}{n} g^{(1)}(\beta) + O(n^{-2})\right]. \tag{80}$$

The factor $n^{-1/2}$ is needed for asymptotically matching to another expansion that applies for $w-1=O(n^{-3/2})$, which is the "central region" discussed in section 5. In section 9 we present some numerical justification for the *ansatz* (80). The numerical calculations show that, however, the ratio of the correction term to the leading term in (80) is bigger than $O(n^{-1})$, perhaps of the order $O(n^{-1/2})$. Thus the series in (80) may be in powers of $n^{-1/2}$ instead of n^{-1}. However, this does not affect our computation of $\Phi(\beta)$ and $g(\beta)$.

We substitute (77) and (80) into (75) to find that the left hand side becomes

$$\begin{aligned}&\exp\left[(n+1)\Phi(\beta+\frac{\beta}{n})\right](n+1)^{-1/2}\left[g(\beta+\frac{\beta}{n})+\frac{1}{n+1}g^{(1)}(\beta+\frac{\beta}{n})+O(n^{-2})\right]\\ =&e^{n\Phi(\beta)}e^{\beta\Phi'(\beta)+\Phi(\beta)}n^{-1/2}\left[1+\frac{1}{n}\left(\beta\Phi'(\beta)+\frac{\beta^2}{2}\Phi''(\beta)\right)\right]\left[1-\frac{1}{2n}\right]\\ \times&\left[g(\beta)+\frac{1}{n}\left(g^{(1)}(\beta)+\beta g'(\beta)\right)+O(n^{-2})\right],\end{aligned} \tag{81}$$

and the sum in the right side becomes

$$\begin{aligned}&\sum_{m=0}^{n/t}\prod_{k=1}^{t-1}G_{i_k}(1+\frac{\beta}{n})\,f\left(\beta(1-\frac{m}{n});n-m\right)\\ =&\sum_{m=0}^{n/t}\prod_{k=1}^{t-1}\left[G_{i_k}(1)+\frac{\beta}{n}G'_{i_k}(1)+O(n^{-2})\right]\frac{1}{\sqrt{n-m}}\exp\left[(n-m)\Phi\left(\beta(1-\frac{m}{n})\right)\right]\\ &\times\left[g(\beta)+\frac{1}{n}\left(g^{(1)}(\beta)-m\beta g'(\beta)+O(n^{-2})\right)\right]\\ \sim&\ e^{n\Phi(\beta)}n^{-1/2}\sum_{m=0}^{\infty}e^{-m\beta\Phi'(\beta)-m\Phi(\beta)}[1+\frac{\beta^2m^2}{2n}\Phi''(\beta)][1+\frac{\beta m^2}{n}\Phi'(\beta)]\\ &\times\left[\sum_{m=0}^{\infty}\prod_{k=1}^{t-1}a_{i_k}+(t-1)\frac{\beta}{n}\sum_{m=0}^{\infty}b_{i_{(t-1)}}\prod_{k=1}^{t-2}a_{i_k}+O(n^{-2})\right]\\ &\times\left[g(\beta)+\frac{1}{n}\left(g^{(1)}(\beta)-m\beta g'(\beta)+O(n^{-2})\right)\right].\end{aligned} \tag{82}$$

Here we extended the limit on the m-sum to ∞, as this will only cause an exponentially small error, and thus not affect the terms in the series in (80). Diving (75) by $t(1+\beta/n)^n$ and letting $n\to\infty$ we obtain, after canceling some common factors, the limiting equation

$$\frac{1}{t}e^{-\beta}e^{\beta\Phi'+\Phi}=\sum_{m=0}^{\infty}e^{-m(\beta\Phi'+\Phi)}a_{i_1}a_{i_2}\cdots a_{i_{(t-1)}}. \tag{83}$$

In view of (76) the right side of (83) is just the $(t-1)^{st}$ power of the generating function $a(z)$, evaluated at $z=e^{-\beta\Phi'-\Phi}$. Thus (83) is a non-linear ODE for the function $\Phi(\beta)$, which is the exponential growth rate of $G_n(w)$ on the β-scale, in view of (80). We thus have

$$\frac{1}{t}e^{-\beta}=za^{t-1}(z)=1-\frac{1}{a(z)},\quad z=e^{-(\beta\Phi'+\Phi)},$$

or

$$a(z)=\frac{t}{t-e^{-\beta}}\,. \tag{84}$$

Using (84) to solve for $(\beta\Phi(\beta))'=-\ln z=\ln[te^{\beta}a^{t-1}(z)]$ leads to

$$(\beta\Phi(\beta))'=\ln(t^t)+t\beta-(t-1)\ln(te^{\beta}-1). \tag{85}$$

Integrating equation (85) gives

$$\begin{aligned}\Phi(\beta) &= \ln(t^t) + \frac{t\beta}{2} - \frac{(t-1)}{\beta}\int_0^\beta \ln(te^x - 1)dx \\ &= \ln(t) + \frac{\beta}{2} - \frac{(t-1)}{\beta}\int_0^\beta \ln\left(1 - \frac{e^{-x}}{t}\right)dx \\ &= \ln(t) + \frac{\beta}{2} + \phi(\beta),\end{aligned} \tag{86}$$

where

$$\begin{aligned}\phi(\beta) &\equiv -\frac{(t-1)}{\beta}\int_0^\beta \ln\left(1 - \frac{e^{-x}}{t}\right)dx \\ &= \frac{(t-1)}{\beta}\int_{-\ln(1-\frac{e^{-\beta}}{t})}^{\ln(\frac{t}{t-1})} \frac{u}{e^u - 1}\,du.\end{aligned} \tag{87}$$

We note that $\Phi(\beta) = c/\beta$ is a homogeneous solution to (85), however it must be discarded because asymptotic matching to the central region ($w - 1 = O(n^{-3/2})$) will force $\Phi(\beta)$ to be bounded as $\beta \to 0^+$. From (86) we then obtain $\Phi(0) = t\ln(t) - (t-1)\ln(t-1)$.

Now we determine $g(\beta)$ in (80). Using (79), (81) and (82) in (75), at order $e^{n\Phi}n^{-1/2}$ we obtain the linear equation

$$\begin{aligned}&\frac{e^{-\beta}}{t}e^{(\beta\Phi)'}\left[\beta g' + \left(\frac{\beta^2}{2}\Phi'' + \beta\Phi' + \frac{\beta^2}{2} - \frac{1}{2}\right)g\right] \\ &= -\beta g'\sum_{m=0}^{\infty} me^{-m(\beta\Phi)'}\left[\sum_{i_1+i_2+\cdots+i_{(t-1)}=m}\left(\prod_{k=1}^{t-1} a_{i_k}\right)\right] \\ &+ (t-1)\beta g\sum_{m=0}^{\infty} e^{-m(\beta\Phi)'}\left[\sum_{i_1+i_2+\cdots+i_{(t-1)}=m} b_{i_{(t-1)}}\left(\prod_{k=1}^{t-2} a_{i_k}\right)\right] \\ &+ g\sum_{m=0}^{\infty} e^{-m(\beta\Phi)'}\left(\frac{m}{2} + \frac{\beta^2 m^2}{2}\Phi'' + \beta m^2\Phi'\right)\left[\sum_{i_1+i_2+\cdots+i_{(t-1)}=m}\left(\prod_{k=1}^{t-1} a_{i_k}\right)\right].\end{aligned} \tag{88}$$

Note that $g^{(1)}$ drops out of (88), in view of (83). Since we determined Φ earlier, (88) is a simple first order linear ODE for $g(\beta)$.

To solve (88), we first define

$$S_j(\beta) = \sum_{m=0}^{\infty} m^j a_m e^{-m(\beta\Phi)'}, \quad j = 0, 1, 2, \tag{89}$$

and

$$T_0(\beta) = \sum_{m=0}^{\infty} b_m e^{-m(\beta\Phi)'}. \tag{90}$$

Again using

$$z = e^{-(\beta\Phi)'} = \frac{e^{-\beta}(t - e^{-\beta})^{t-1}}{t^t},$$

we obtain

$$S_0(\beta) = \sum_{m=0}^{\infty} a_m z^m = \frac{t}{t - e^{-\beta}}. \tag{91}$$

To evaluate $S_1(\beta)$, we differentiate (78) and multiply by z

$$S_1(\beta) = \sum_{m=0}^{\infty} m\, a_m\, z^m = z\, a'(z) = \frac{e^{-\beta}}{(t - e^{-\beta})(1 - e^{-\beta})}. \tag{92}$$

Differentiating $S_0(\beta)$ with respect to β we have

$$S_0'(\beta) = -\sum_{m=0}^{\infty} m\,(\beta\Phi)'' a_m\, e^{-m(\beta\Phi)'} = -(\beta\Phi)'' S_1(\beta), \tag{93}$$

and differentiating again yields

$$S_0''(\beta) = [(\beta\Phi)'']^2 S_2(\beta) - (\beta\Phi)''' S_1(\beta). \tag{94}$$

Using (91)-(94) and $\frac{\beta^2}{2}\Phi'' + \beta\Phi' = \frac{\beta}{2}(\beta\Phi)''$ in (88) leads to

$$\begin{aligned}
&\frac{e^{-\beta}}{t} e^{(\beta\Phi)'}\left[\beta g' + \left(\frac{\beta}{2}(\beta\Phi)' + \frac{\beta^2}{2} - \frac{1}{2}\right) g\right] \\
&= (t-1)\beta g S_0^{t-2} T_0 + \frac{(t-1)}{2} g S_0^{t-2} S_1 - (t-1)\beta g' S_0^{t-2} S_1 \\
&+ \frac{g\beta}{2}(\beta\Phi)''\left[(t-1)S_0^{t-2} S_2 + (t-1)(t-2) S_0^{t-3} S_1^2\right].
\end{aligned} \tag{95}$$

We set

$$g(\beta) = \sqrt{\beta}\,\hat{g}(\beta), \tag{96}$$

and substitute (96) into (95), with which (95) simplifies to

$$\begin{aligned}
&\left[S_0^{t-1} + (t-1) S_1 S_0^{t-2}\right]\frac{\hat{g}'(\beta)}{\hat{g}(\beta)} \\
=&(t-1) T_0 S_0^{t-2} + \frac{(\beta\Phi)''}{2}\left[(t-1) S_0^{t-2} S_2 + (t-1)(t-2) S_0^{t-3} S_1^2\right] \\
&- \frac{S_0^{t-1}}{2}\left[(\beta\Phi)'' + \beta\right].
\end{aligned} \tag{97}$$

From (91)-(94) we obtain

$$S_2(\beta) = \frac{e^{-\beta}(t - e^{-2\beta})}{t(1 - e^{-\beta})^3 (t - e^{-\beta})}, \tag{98}$$

and

$$(\beta\Phi)'' = -\frac{S_0'(\beta)}{S_1(\beta)} = \frac{t(1 - e^{-\beta})}{t - e^{-\beta}}. \tag{99}$$

We use the generating function of $b(z)$, which was previously expressed in terms of $a(z)$, to obtain

$$T_0(\beta) = \frac{e^{-2\beta}}{(t - e^{-\beta})(1 - e^{-\beta})^2}. \tag{100}$$

Solving (97) for $\hat{g}'(\beta)/\hat{g}(\beta)$ we obtain

$$\begin{aligned}
\frac{\hat{g}'(\beta)}{\hat{g}(\beta)} &= \frac{(2t\beta e^{\beta} + 4te^{\beta} - t\beta e^{2\beta} - te^{2\beta} - t\beta - 2e^{\beta} - 1)}{2(te^{\beta} - 1)(e^{\beta} - 1)} \\
&= -\frac{1}{2} - \frac{3e^{-\beta}}{2(1 - e^{-\beta})} - \frac{3e^{-\beta}}{2(t - e^{-\beta})} + \frac{t\beta e^{-\beta}}{2(t - e^{-\beta})} - \frac{t\beta}{2(1 - e^{-\beta})}.
\end{aligned} \tag{101}$$

Using MAPLE, (101) integrates to

$$\hat{g}(\beta) = (const.)e^{-\beta^2/4}e^{-\beta/2}\left(\frac{1-e^{-\beta}}{t-e^{-\beta}}\right)^{3/2} \times \exp\left\{\frac{(t-1)}{2}\beta\ln\left(1-\frac{e^{-\beta}}{t}\right)+\frac{1}{2}\left[\text{polylog}\left(2,\frac{e^{-\beta}}{t}\right)-t\,\text{dilog}\left(1-\frac{e^{-\beta}}{t}\right)\right]+M^*\right\}, \tag{102}$$

where $M^* = \frac{(t-1)}{2}\,\text{dilog}\left(\frac{t-1}{t}\right)$ and the dilog(x) and polylog(a,z) functions are defined by

$$\text{dilog}(x) = \int_1^x \frac{\ln(t)}{1-t}\,dx, \quad \text{polylog}(a,z) = \sum_{n=1}^{\infty}\frac{z^n}{n^a},$$
$$\text{dilog}(z) = \text{polylog}(2, 1-z).$$

We note that $\phi(\beta)$ in (87) may also be written as

$$\frac{\beta}{2}\phi(\beta) = M^* + \frac{1}{2}\left[\text{polylog}\left(2,\frac{e^{-\beta}}{t}\right)-t\,\text{dilog}\left(\frac{e^{-\beta}}{t}\right)\right]. \tag{103}$$

With (103), (102) becomes to

$$\hat{g}(\beta) = (const.)\,e^{-\beta^2/4}e^{-\beta/2}\left(\frac{1-e^{-\beta}}{t-e^{-\beta}}\right)^{3/2}\ \exp\left[\frac{(t-1)}{2}\beta\ln\left(1-\frac{e^{-\beta}}{t}\right)+\frac{\beta}{2}\phi(\beta)\right], \tag{104}$$

where $\phi(\beta)$ is defined in (87).

We assume that (80) asymptotically matches to the expansion for $w > 1$; we write this condition symbolically as

$$e^{n\Phi(\beta)}\sqrt{\frac{\beta}{n}}\hat{g}(\beta)\bigg|_{\beta\to\infty} \sim w^{\binom{n}{2}}t^{n-1}G_*(w)\bigg|_{w\to 1}. \tag{105}$$

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

$$\Phi(\beta) = \ln(t) + \frac{\beta}{2} - \frac{(t-1)}{\beta}\int_0^{\infty}\ln\left(1-\frac{e^{-x}}{t}\right)dx + O\left(\frac{e^{-\beta}}{\beta}\right),$$

and

$$\beta\phi(\beta) \to -(t-1)\int_0^{\infty}\ln\left(1-\frac{e^{-x}}{t}\right)dx \equiv d_0, \tag{106}$$

and thus

$$\hat{g}(\beta) \sim (const.)e^{-\beta^2/4}e^{-\beta/2}e^{d_0/2}t^{-3/2}, \ \ \beta \to \infty. \tag{107}$$

For $w \to 1$, we have

$$w^{\binom{n}{2}} = \left(1+\frac{\beta}{n}\right)^{\frac{n(n-1)}{2}} = e^{n\beta/2}e^{-\beta^2/4}e^{-\beta/2}\left[1+O\left(\frac{\beta^3}{n}\right)\right]. \tag{108}$$

By using (106)-(108) in (105) we find that the matching is possible provided that

$$G_*(w) \sim \frac{(const.)}{\sqrt{t}}\sqrt{w-1}\,e^{d_0/2}\exp\left(\frac{d_0}{w-1}\right), \ w \to 1^+. \tag{109}$$

This shows that $G_*(w)$ has an essential singularity at $w = 1$. We give numerical support for the behavior in (109) in section 9. In section 5 we will show that asymptotically matching the β-scale expansion (80), as

$\beta \to 0^+$, to the central region a-scale expansion (which corresponds to $n^{3/2}(w-1) = a = O(1)$), as $a \to +\infty$, determines the constant in (104) and (109) as

$$const. = \frac{t\sqrt{2t}}{\sqrt{\pi}(t-1)}. \tag{110}$$

With (109) and (110) we have established (41).

Note that if we used the following *ansatz*, which is slightly more general than (80),

$$f(\beta; n) \sim e^{n\Phi(\beta)} n^{\Psi(\beta)} g(\beta), \tag{111}$$

in (75), we would find that $\Phi(\beta)$ is as in (86), and 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 (104). Matching to the a-scale result would show that $\Psi_0 = -\frac{1}{2}$ and fix the multiplier constant as in (110). Using (111) and matching to the expansion for $w > 1$ would lead to

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

We use (80) to calculate $g(n, p)$ from the Cauchy integral (10). We write

$$w^{-p-1} = \left(1 + \frac{\beta}{n}\right)^{-p-1} = \exp\left[-\frac{p}{n}\beta + \frac{p}{2n^2}\beta^2 + O\left(\frac{p\beta^3}{n^3}\right)\right]. \tag{112}$$

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 $\beta = O(1)$, and (10) will have saddle point(s) where

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

or

$$\frac{1}{2} - \Lambda - \frac{t-1}{\beta}\ln\left(1 - \frac{e^{-\beta}}{t}\right) = \frac{t-1}{\beta^2}\int_{-\ln\left(1-\frac{e^{-\beta}}{t}\right)}^{\ln\left(\frac{t}{t-1}\right)} \frac{x}{e^x - 1}dx. \tag{113}$$

Equation (113) is a transcendental equation with a unique real solution $\beta = \beta_* = \beta_*(\Lambda)$, that satisfies

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

We only need to consider $\Lambda \in \left(0, \frac{1}{2}\right)$ in view of $p_{max}(n)$ in (29). The steepest descent directions at $\beta = \beta_*$ are $arg(\beta - \beta_*) = \pm\frac{\pi}{2}$, and (112) and (113) lead to

$$g(n,p) \sim \frac{\sqrt{\beta_*}}{n^2}\hat{g}(\beta_*)e^{\Lambda\beta_*^2/2}\frac{1}{\sqrt{2\pi|\phi''(\beta_*)|}}\exp\left[n\left(\ln t + \frac{\beta_*}{2} + \phi(\beta_*) - \Lambda\beta_*\right)\right]. \tag{114}$$

From (113) we have

$$\phi(\beta_*) = \left(\frac{1}{2} - \Lambda\right)\beta_* - (t-1)\ln\left(1 - \frac{e^{-\beta}}{t}\right), \tag{115}$$

and (87) leads to

$$|\phi''(\beta_*)| = \frac{1}{\beta_*}\left[1 - 2\Lambda - \frac{t-1}{te^{\beta_*} - 1}\right]. \tag{116}$$

Combining (104) and (110) with (115) and (116), we obtain from (114)

$$\begin{aligned} g(n,p) \sim \frac{t^{n+1}}{n^2}\frac{\sqrt{t}}{\pi(t-1)}\beta_* e^{-\beta_*/2}\left(\frac{1-e^{-\beta_*}}{t-e^{-\beta_*}}\right)^{3/2}\left[1-2\Lambda-\frac{t-1}{te^{\beta_*}-1}\right]^{-1/2} \\ \times \exp\left\{n\left[\beta_*(1-2\Lambda)-(t-1)\ln\left(1-\frac{e^{-\beta_*}}{t}\right)\right]\right\}, \end{aligned} \tag{117}$$

and thus we have established (55).

Now we discuss the asymptotic matching between (72), as $L \to \infty$, and (117), as $\Lambda \uparrow \frac{1}{2}$. We solve (113) asymptotically, for β_* large. We have

$$\phi(\beta) = \frac{d_0}{\beta} - \frac{t-1}{t}e^{-\beta} + O_R(e^{-2\beta}),\ \beta \to \infty, \tag{118}$$

with which (113) becomes

$$\Lambda - \frac{1}{2} = -\frac{d_0}{\beta^2} + \left(\frac{t-1}{t\beta} + \frac{t-1}{t\beta^2}\right)e^{-\beta} + O_R(e^{-2\beta}). \tag{119}$$

Here O_R means "roughly" of the order, and ignores factors algebraic in β. From (119) we find that

$$\beta_* \sim \sqrt{\frac{2d_0}{1-2\Lambda}}\left[1 - \frac{t-1}{2td_0}\left(1+\sqrt{\frac{2d_0}{1-2\Lambda}}\right)\exp\left(-\frac{2d_0}{1-2\Lambda}\right)\right],\ \ \Lambda \uparrow \frac{1}{2}. \tag{120}$$

Substituting (120) into (117), the right side becomes

$$\frac{\sqrt{2d_0}t^n}{\pi n^2(1-2\Lambda)(t-1)}\left[-\frac{1}{2}\sqrt{\frac{2d_0}{1-2\Lambda}} + n\sqrt{2d_0}\sqrt{1-2\Lambda}\right]. \tag{121}$$

We show that (121) matches with (72) as $L \to \infty$. We argue that there exists a saddle point in the range $w \sim 1$, and expand (72) for L large and estimate $G_*(w)$ by (71), which yields the integral

$$\frac{t^{n-1}}{2\pi i}\, d_1 \int_C w^{L-1}\sqrt{w-1}\exp\left(\frac{d_0}{w-1}\right)dw. \tag{122}$$

The integrand in (122) has a saddle point where

$$\frac{d}{dw}\left[L\ln w + \frac{d_0}{w-1}\right] = 0 \ \Rightarrow\ w = w_s \equiv 1 + \frac{d_0}{2L} + \sqrt{\frac{d_0}{L}}\sqrt{1+\frac{d_0}{4L}},$$

and then we obtain the following estimate for (122):

$$\frac{t^{n-1}d_1}{\sqrt{2\pi}}\sqrt{w_s-1}\left[\frac{2d_0w_s^2}{(w_s-1)^3} - L\right]^{-1/2}\exp\left[L\ln w_s + \frac{d_0}{w_s-1}\right]. \tag{123}$$

We can simplify (123) for $L \to \infty$ by using

$$w_s = 1 + \sqrt{\frac{d_0}{L}} + \frac{d_0}{2L} + O(L^{-3/2}), \tag{124}$$

and note that

$$L = \binom{n}{2} - p = \frac{n^2}{2} - p - \frac{n}{2} = n^2\left(\frac{1}{2} - \Lambda\right) - \frac{n}{2}. \tag{125}$$

Using (124) in (123) leads to

$$\frac{t^n}{L}\frac{d_1\sqrt{d_0}}{2t\sqrt{\pi}}\exp\left[2\sqrt{Ld_0} - \frac{d_0}{2}\right] = \frac{t^n}{L}\frac{\sqrt{d_0}}{\sqrt{2}\pi(t-1)}\exp(2\sqrt{Ld_0}). \tag{126}$$

If we use (125) in (126) and expand for $n \to \infty$, $\Lambda \to \frac{1}{2}$ with $n(\frac{1}{2} - \Lambda) \to \infty$, we obtain exactly (121). This confirms the asymptotic matching between the Λ-scale and L-scale results.

5. CENTRAL REGION

In this section we analyze (7) for $w-1=O(n^{-3/2})$ and then obtain an expansion for $g(n,p)$ that is valid for $p=O(n^{3/2})$. We define a by

$$w-1=\frac{a}{n^{3/2}},\ -\infty<a<\infty. \tag{127}$$

We will separately consider the cases, $a<0$ and $a>0$, as needed.

We set

$$\begin{aligned} G_n(w) &= \frac{1}{z_0^n}\left[a_n z_0^n+\frac{1}{n^{3/2}}\bar{C}_n(a)\right] \\ &= \frac{1}{(t-1)n+1}\binom{nt}{n}+\frac{1}{z_0^n n^{3/2}}\bar{C}_n(n^{3/2}(w-1)), \end{aligned} \tag{128}$$

and (12) leads to

$$\bar{C}_n(0)=0. \tag{129}$$

Since a_n satisfies the recurrence

$$a_{n+1}=\sum_{k_1+k_2+\cdots+k_t=n}\left(\prod_{i=1}^{t}a_{k_i}\right),\quad n\geq 0,\ \ a_0=1, \tag{130}$$

we use (128) in (7) and obtain

$$\begin{aligned} &\frac{1}{z_0(n+1)^{3/2}}\bar{C}_{n+1}\left(\left(1+\frac{1}{n}\right)^{3/2}a\right) \\ =&a_{n+1}z_0^n\left[\left(1+\frac{a}{n^{3/2}}\right)^n-1\right] \\ &+t\left(1+\frac{a}{n^{3/2}}\right)^n\sum_{l=0}^{n}\frac{S_{l,1}z_0^l}{k_1{}^{3/2}}\bar{C}_{k_1}\left(k_1{}^{3/2}a\right) \\ &+\binom{t}{2}\left(1+\frac{a}{n^{3/2}}\right)^n\sum_{l=0}^{n}\frac{S_{l,2}z_0^l}{(k_1k_2)^{3/2}}\bar{C}_{k_1}\left(k_1{}^{3/2}a\right)\bar{C}_{k_2}\left(k_2{}^{3/2}a\right) \\ &+\cdots \\ &+\binom{t}{t}\left(1+\frac{a}{n^{3/2}}\right)^n\sum_{k_1+k_2+\cdots+k_t=n}\left[\prod_{i=1}^{t}\frac{1}{k_i{}^{3/2}}\bar{C}_{k_i}\left(k_i{}^{3/2}a\right)\right], \end{aligned} \tag{131}$$

where

$$S_{l,i}=\sum_{k_{i+1}+\cdots+k_t=l}\left(\prod_{j=i+1}^{t}a_{k_j}\right),\ \ l=n-(k_1+\cdots+k_i).$$

We write $G_n(w)$ as a Taylor series around $w=1$, setting

$$G_n(w)=\sum_{j=0}^{\infty}\frac{M_{j,n}}{j!}(w-1)^j. \tag{132}$$

From (12) we have $M_{0,n}=a_n,\ M_{1,n}=b_n$, and $M_{2,n}=c_n$. Dividing (7) by w^n, differentiating N times with respect to w, and setting $w = 1$, we obtain

$$\begin{aligned}&\sum_{i=0}^{N}\binom{N}{i}(-1)^{N-i}\frac{(n+N-i-1)!}{(n-1)!}M_{i,n+1}\\&=\sum_{l_1+l_2+\cdots+l_t=n}\left[\sum_{j_1+j_2+\cdots+j_t=N}\binom{N}{j_1,\,j_2,\,\cdots,\,j_t}\prod_{k=1}^{t}M_{j_k,l_k}\right].\end{aligned}\tag{133}$$

For $n\to\infty$ we write

$$M_{i,n}=z_0^{-n}\tilde{M}_{i,n}=z_0^{-n}n^{\frac{3}{2}(i-1)}\left[m_i+\frac{1}{\sqrt{n}}\bar{m}_i+O(n^{-1})\right].\tag{134}$$

We will obtain asymptotic approximations to $\bar{C}_n(a)$, which is equivalent to obtaining expansions for $M_{j,n}$ for each j. For the leading term, we shall analyze (131). For the correction term, we shall analyze the functional equation (9) for the double transform.

5.1 Analysis of the Basic Recurrence

We consider (7), which becomes (131) on the a-scale. Using (24) we obtain

$$a_{n+1}z_0^n\left[\left(1+\frac{a}{n^{3/2}}\right)^n-1\right]=\frac{m_0}{z_0}\frac{a}{n^2}+\frac{m_0}{2z_0}\frac{a^2}{n^{5/2}}+O(n^{-3}).\tag{135}$$

We expand $\bar{C}_n(a)$ in (128) as

$$\bar{C}_n(a)=C^{(0)}(a)+\frac{1}{\sqrt{n}}C^{(1)}(a)+\frac{1}{n}C^{(2)}(a)+O(n^{-3/2}),\tag{136}$$

and then estimate the various terms in (131). In view of (129) we see that $C^{(j)}(0)=0$ for all $j\geq 0$.

Using the generating function $a(z)$ defined in (11) and the estimate in (24), we find that

$$\begin{aligned}&\sum_{l=0}^{n}a_l z_0^l=\frac{t}{t-1}-\frac{2\,m_0}{\sqrt{n}}+O(n^{-3/2}),\\&\sum_{l=0}^{n}S_{l,1}z_0^l=\left(\frac{t}{t-1}\right)^{t-1}-(t-1)\left(\frac{t}{t-1}\right)^{t-2}\frac{2\,m_0}{\sqrt{n}}+O(n^{-3/2}).\end{aligned}\tag{137}$$

From (136) and the Euler-MacLaurin formula we have

$$\begin{aligned}&\sum_{l=0}^{n}\frac{S_{l,2}z_0^l}{(k_1k_2)^{3/2}}\bar{C}_{k_1}\left({k_1}^{3/2}a\right)\bar{C}_{k_2}\left({k_2}^{3/2}a\right)\\&\sim\left(\frac{t}{t-1}\right)^{t-2}\frac{1}{n^2}\int_0^1\frac{C^{(0)}(x^{3/2}a)C^{(0)}((1-x)^{3/2}a)}{(x(1-x))^{3/2}}dx.\end{aligned}\tag{138}$$

Here we used

$$\sum_{l=0}^{n}S_{l,2}z_0^l\sim\left(\frac{t}{t-1}\right)^{t-2},\ \ n\to\infty.$$

We write the first sum in the right side of (131) as

$$\begin{aligned}&\sum_{l=0}^{n}\frac{S_{l,1}z_0^l}{k_1{}^{3/2}}\bar{C}_{k_1}\left(k_1{}^{3/2}a\right)\\&=\sum_{l=0}^{n}\frac{S_{l,1}z_0^l}{n^{3/2}}\left[\left(1-\frac{l}{n}\right)^{-3/2}C^{(0)}\left(\left(1-\frac{l}{n}\right)^{3/2}a\right)-C^{(0)}(a)\right]\\&+\sum_{l=0}^{n}\frac{S_{l,1}z_0^l}{n^2}\left[\left(1-\frac{l}{n}\right)^{-2}C^{(1)}\left(\left(1-\frac{l}{n}\right)^{3/2}a\right)-C^{(1)}(a)\right]\\&+\sum_{l=0}^{n}S_{l,1}z_0^l\left[\frac{C^{(0)}(a)}{n^{3/2}}+\frac{C^{(1)}(a)}{n^2}\right]+O(n^{-5/2}).\end{aligned}\tag{139}$$

From (18) for $l\to\infty$ we have

$$S_{l,1}=\frac{\sqrt{t-1}}{t\sqrt{2\pi}z_0}\frac{1}{l^{3/2}}+O(l^{-5/2}).\tag{140}$$

Using (140) and the Euler-Maclaurin formula we obtain

$$\begin{aligned}&\sum_{l=0}^{n}\frac{S_{l,1}z_0^l}{n^{3/2}}\left[\left(1-\frac{l}{n}\right)^{-3/2}C^{(0)}\left(\left(1-\frac{l}{n}\right)^{3/2}a\right)-C^{(0)}(a)\right]\\&\sim\frac{\sqrt{t-1}}{t\sqrt{2\pi}z_0}\frac{1}{n^2}\int_0^1\frac{1}{x^{3/2}}\left[\frac{C^{(0)}((1-x)^{3/2}a)}{(1-x)^{3/2}}-C^{(0)}(a)\right]dx+O(n^{-5/2}).\end{aligned}\tag{141}$$

Thus we have

$$\begin{aligned}&\sum_{l=0}^{n}\frac{S_{l,1}z_0^l}{k_1{}^{3/2}}\bar{C}_{k_1}\left(k_1{}^{3/2}a\right)\\&=\frac{\sqrt{t-1}}{t\sqrt{2\pi}z_0}\frac{1}{n^2}\int_0^1\frac{1}{x^{3/2}}\left[\frac{C^{(0)}((1-x)^{3/2}a)}{(1-x)^{3/2}}-C^{(0)}(a)\right]dx\\&+\left(\frac{t}{t-1}\right)^{t-1}\frac{C^{(0)}(a)}{n^{3/2}}+\left(\frac{t}{t-1}\right)^{t-1}\left[C^{(1)}(a)-\frac{2(t-1)^2m_0}{t}C^{(0)}(a)\right]\frac{1}{n^2}+O(n^{-5/2}).\end{aligned}\tag{142}$$

Multiplying (131) 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 (135) leads to

$$\begin{aligned}&\frac{1}{z_0(n+1)^{3/2}}\left(1+\frac{a}{n^{3/2}}\right)^{-n}\bar{C}_{n+1}\left(\left(1+\frac{1}{n}\right)^{3/2}a\right)-a_{n+1}z_0^n\left[1-\left(1+\frac{a}{n^{3/2}}\right)^{-n}\right]\\&=\frac{1}{z_0n^{3/2}}\left[1+O(n^{-1})\right]\left[1-\frac{a}{\sqrt{n}}+O(n^{-1})\right]\\&\times\left[C^{(0)}(a)+\frac{1}{\sqrt{n}}C^{(1)}(a)+O(n^{-1})\right]-\frac{m_0}{z_0n^{3/2}}\frac{a}{\sqrt{n}}+O(n^{-5/2})\\&=\frac{1}{z_0n^{3/2}}C^{(0)}(a)+\frac{1}{z_0n^2}\left[C^{(1)}(a)-am_0-aC^{(0)}(a)\right]+O(n^{-5/2}).\end{aligned}\tag{143}$$

We note that cubic and higher order terms in $\bar{C}_{k_i}(a)$ in the right side of (131) are at most of the order $O(n^{-5/2})$. Comparing (143) with the linear and quadratic terms in the right side of (131) in term of $\bar{C}_{k_i}(a)$, and using (138), (141) and (142), we obtain at $O(n^{-2})$ the limiting integral equation

$$\begin{aligned} -am_0 - aC^{(0)}(a) = &-\frac{2m_0(t-1)^2}{t}C^{(0)}(a) \\ &+ \frac{(t-1)^2}{2t}\int_0^1 \frac{C^{(0)}(x^{3/2}a)C^{(0)}((1-x)^{3/2}a)}{(x(1-x))^{3/2}}dx \\ &+ \frac{\sqrt{t-1}}{\sqrt{2t\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{C^{(0)}((1-x)^{3/2}a)}{(1-x)^{3/2}} - C^{(0)}(a)\right]dx. \end{aligned} \tag{144}$$

Here m_0 is the zeroth moment, which we obtained in (24) as

$$m_0 = \frac{\sqrt{t}}{\sqrt{2\pi}(t-1)^{3/2}}.$$

Thus, the leading term in (136) satisfies the non-linear integral equation (144). Note that at $O(n^{-5/2})$, we would ultimately obtain a linear integral equation for the correction term $C^{(1)}(a)$ in (136). However, it is easier to obtain $C^{(1)}(a)$ by analyzing the double transform equation (9), which we will present in the next subsection.

We introduce

$$D(y) = \sum_{j=0}^{\infty} \frac{u_j}{(j+1)!}y^{\frac{3}{2}j}, \quad y > 0, \tag{145}$$

and set

$$u_j = m_{j+1}, \tag{146}$$

and

$$\bar{D}(Y) = \sum_{j=1}^{\infty} Y^{\frac{3}{2}(j-1)}\frac{m_j}{j!}(-1)^j = \sum_{i=0}^{\infty} Y^{\frac{3}{2}i}\frac{u_i}{(i+1)!}(-1)^{i+1}, \quad -a = Y^{3/2} > 0. \tag{147}$$

We also note that $\bar{D}(0) = -u_0 = -m_1$ and

$$\sum_{k=0}^{\infty} \frac{m_k}{k!}a^k = m_0 + \sum_{k=1}^{\infty} \frac{m_k}{k!}(-a)^{k-1}(-a)(-1)^k = m_0 - a\bar{D}(Y). \tag{148}$$

To analyze (144) for $a < 0$ we first note that, in view of (147), $C^{(0)}$ and $\bar{D}$ are related by

$$C^{(0)}(a) = (-a)\bar{D}((-a)^{2/3}) = Y^{3/2}\bar{D}(Y). \tag{149}$$

Then (144) becomes, for $Y > 0$,

$$\begin{aligned} a^2\bar{D}(Y) - a\,m_0 = &\frac{2\,m_0(t-1)^2}{t}a\bar{D}(Y) \\ &+ \frac{(t-1)^2}{2t}a^2\int_0^1 \bar{D}(Yx)\bar{D}(Y-Yx)dx \\ &- \frac{\sqrt{t-1}}{\sqrt{2t\pi}}a\int_0^1 \frac{\bar{D}(Y-Yx)-\bar{D}(Y)}{x^{3/2}}dx. \end{aligned} \tag{150}$$

Integrating by parts and using $\bar{D}(0) = -m_1$ we obtain

$$\int_0^1 \frac{\bar{D}(Y-Yx)-\bar{D}(Y)}{x^{3/2}}dx = 2\bar{D}(Y) + 2m_1 - 2Y\int_0^1 x^{-1/2}\bar{D}'(Y-Yx)dx. \tag{151}$$

Using (151) in (150), and multiplying by Y/a^2 we obtain

$$Y\bar{D}(Y) = -\frac{\sqrt{2(t-1)}}{\sqrt{t\pi}}\int_0^Y \frac{\bar{D}'(v)}{\sqrt{Y-v}}dv + \frac{(t-1)^2}{2t}\int_0^Y \bar{D}(v)\bar{D}(Y-v)dv. \tag{152}$$

Similarly, for $a > 0$ we have

$$C^{(0)}(a) = aD(a^{2/3}) = y^{3/2}D(y), \tag{153}$$

and then (144) yields, for $y > 0$,

$$-yD(y) = -\frac{\sqrt{2(t-1)}}{\sqrt{t\pi}}\int_0^y \frac{D'(v)}{\sqrt{y-v}}dv + \frac{(t-1)^2}{2t}\int_0^y D(v)D(y-v)dv. \tag{154}$$

In view of (132) and (134) we obtain the leading order approximation to $G_n(w)$ (for $n \to \infty$ with a fixed $a > 0$) as

$$\begin{aligned} G_n(w) &\sim \sum_{j=0}^{\infty} z_0^{-n} n^{\frac{3}{2}(j-1)} \frac{m_j}{j!}(w-1)^j \\ &= \frac{1}{z_0^n n^{3/2}} \sum_{j=0}^{\infty} \frac{m_j}{j!} a^j \\ &= \frac{1}{z_0^n n^{3/2}} \left[m_0 + a\sum_{j=0}^{\infty} \frac{m_{j+1}}{(j+1)!} a^j \right] \\ &= \frac{1}{z_0^n n^{3/2}} \left[m_0 + aD(a^{2/3}) \right], \quad a > 0. \end{aligned} \tag{155}$$

and for $n \to \infty$ with a fixed $a < 0$ we have

$$G_n(w) \sim \frac{1}{z_0^n n^{3/2}} \left[m_0 - a\bar{D}((-a)^{2/3}) \right], \quad -a = Y^{3/2} > 0. \tag{156}$$

We note that (152) differs from (154) only by the sign on the left-hand side. However (152) is susceptible to be solved by a Laplace transform while (154) is not.

Setting

$$D_*(s) = \mathcal{L}\{\bar{D}(Y)\} \equiv \int_0^\infty e^{-sY}\bar{D}(Y)dY, \tag{157}$$

where $\mathcal{L}$ is the Laplace transform operator, we then have

$$\mathcal{L}\{Y\bar{D}(Y)\} = -D_*'(s), \quad \bar{D}(0) = -m_1 = -\frac{t}{2(t-1)^2},$$

$$\mathcal{L}\left\{\bar{D}'(Y) * \frac{1}{\sqrt{Y}}\right\} = \left[sD_*(s) - \bar{D}(0)\right]\sqrt{\frac{\pi}{s}}.$$

Here the * denotes convolution. Thus (152) transforms to

$$-D_*'(s) = \frac{(t-1)^2}{2t}D_*^2(s) - \frac{\sqrt{2(t-1)}}{\sqrt{t}}\sqrt{s}D_*(s) - \frac{\sqrt{t}}{\sqrt{2}(t-1)^{3/2}}\frac{1}{\sqrt{s}}. \tag{158}$$

This Riccati equation can be solved by setting

$$D_*(s) = \frac{\sqrt{2t}}{(t-1)^{3/2}}\sqrt{s} + U(s), \qquad U(s) = \frac{2t}{(t-1)^2}\frac{F'(s)}{F(s)}, \tag{159}$$

which yields the Airy equation

$$F''(s) = \frac{(t-1)s}{2t} F(s). \tag{160}$$

The solution that has acceptable behavior as $s \to +\infty$ is

$$F(s) = (const.) Ai\left(\left(\frac{t-1}{2t}\right)^{1/3} s\right), \tag{161}$$

where $Ai(\cdot)$ is the Airy function. Using (161) in (159) and inverting the Laplace transform, we have

$$\bar{D}(Y) = \frac{1}{2\pi i} \int_{Br} e^{sY} \left[\frac{\sqrt{2t}}{(t-1)^{3/2}} \sqrt{s} + \frac{1}{t-1} \left(\frac{2t}{t-1}\right)^{2/3} \frac{Ai'\left(\left(\frac{t-1}{2t}\right)^{1/3} s\right)}{Ai\left(\left(\frac{t-1}{2t}\right)^{1/3} s\right)} \right] ds, \quad Y = (-a)^{2/3}, \tag{162}$$

which is valid only for $a < 0$. Here Br is any vertical contour on which $\Re(s) > r_0 = max\{z : Ai(z) = 0\}$. We can also write (162) as

$$\begin{aligned} \bar{D}(Y) &= \frac{1}{2\pi i} \frac{d}{dY} \left\{ \int_{Br} e^{sY} \left[\frac{\sqrt{2t}}{(t-1)^{3/2} \sqrt{s}} + \frac{1}{s(t-1)} \left(\frac{2t}{t-1}\right)^{2/3} \frac{Ai'\left(\left(\frac{t-1}{2t}\right)^{1/3} s\right)}{Ai\left(\left(\frac{t-1}{2t}\right)^{1/3} s\right)} \right] ds \right\} \\ &= \frac{d}{dY} \left[\frac{\sqrt{2t}}{(t-1)^{3/2} \sqrt{\pi Y}} + \frac{1}{(t-1)} \left(\frac{2t}{t-1}\right)^{2/3} \sum_{j=0}^{\infty} \frac{1}{r_j} \exp\left(\left(\frac{2t}{t-1}\right)^{1/3} r_j Y\right) \right] \\ &= \frac{m_0}{a} + \frac{2t}{(t-1)^2} \sum_{j=0}^{\infty} \exp\left(\left(\frac{2t}{t-1}\right)^{1/3} r_j Y\right). \end{aligned} \tag{163}$$

Here $0 > r_0 > r_1 > \cdots$ and r_j are the roots of the Airy function $Ai(\cdot)$. We evaluated the integral by the residue theorem, closing the Br contour in the left half-plane.

Using (163) in (156) we have the leading order approximation

$$G_n(w) \sim \frac{1}{z_0^n n^{3/2}} (-a) \frac{2t}{(t-1)^2} \sum_{j=0}^{\infty} \exp\left(-\left(\frac{2t}{t-1}\right)^{1/3} |r_j| Y\right), \quad Y = (-a)^{2/3} > 0. \tag{164}$$

We show that (164) is consistent with the fact that $G_n(1) = a_n \sim m_0/(z_0^n n^{3/2})$ as $n \to \infty$. It is well known that

$$Ai(z) \sim \frac{1}{\sqrt{\pi}} (-z)^{-1/4} \sin\left(\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}\right), \quad z \to -\infty, \tag{165}$$

and hence we can approximate the roots r_j by

$$|r_j| \sim \left(\frac{3j\pi}{2}\right)^{2/3}, \quad j \to \infty. \tag{166}$$

For $Y \to 0^+$ we can estimate the sum in (164) by the Euler-MacLaurin formula, hence

$$\begin{aligned} &\sum_{j=0}^{\infty} \exp\left(-\left(\frac{2t}{t-1}\right)^{1/3} |r_j| Y\right) \\ &\sim \int_0^{\infty} \exp\left[-\left(\frac{2t}{t-1}\right)^{1/3} \left(\frac{3x\pi}{2}\right)^{2/3} Y\right] dx \\ &= \frac{\sqrt{t-1}}{Y^{3/2} \pi \sqrt{2t}} \int_0^{\infty} e^{-u} \sqrt{u}\, du = \frac{-\sqrt{t-1}}{2a\sqrt{2t\pi}} = \frac{(t-1)^2 m_0}{2t(-a)}. \end{aligned}$$

We next derive an expression valid for $a > 0$, and show that, for $a \to \infty$, the a-scale expansion asymptotically matches to the β-scale, for $\beta \to 0^+$. By using the integral equation (154) and a **WKB**-type expansion we can obtain the behavior of the central region approximation as $a = y^{3/2} \to +\infty$. We assume that

$$D(y) \sim K(y)e^{\Psi(y)}, \quad y \to \infty, \tag{167}$$

where $\Psi(y) \gg \log[K(y)]$. We also expect that $\Psi'(y) > 0$ and $\Psi'(y) \to \infty$ as $y \to \infty$. Using (167) to estimate the various terms in (154) we have

$$\begin{aligned}
&\int_0^y \frac{1}{\sqrt{v}} D'(y-v)dv \sim \int_0^y \frac{1}{\sqrt{v}} [K'e^{\Psi} + K\Psi' e^{\Psi}](y-v)dv \\
&\sim \int_0^y e^{\Psi(y)} \frac{e^{-v\Psi'(y)}}{\sqrt{v}} \left[\left(1 + \frac{1}{2}v^2\Psi''(y)\right) (\Psi'(y) - v\Psi''(y)) (K(y) - vK'(y)) + K'(y) \right] dv \\
&\sim \sqrt{\Psi'(y)\pi} K(y) e^{\Psi(y)} + K(y)e^{\Psi(y)} \int_0^\infty \frac{e^{-v\Psi'(y)}}{\sqrt{v}} \left[\frac{K'}{K} - v\Psi'' - v\Psi' \frac{K'}{K} + \frac{1}{2}v^2 \Psi''\Psi' \right] dv \\
&= K(y)e^{\Psi(y)} \sqrt{\pi} \left[\sqrt{\Psi'} + \frac{K'}{K}\frac{1}{\sqrt{\Psi'}} - \frac{1}{2(\Psi')^{3/2}} \left(\Psi'' + \Psi' \frac{K'}{K} \right) + \frac{3}{8} \frac{\Psi''}{(\Psi')^{3/2}} \right].
\end{aligned} \tag{168}$$

We estimate the non-linear term as

$$\begin{aligned}
&\int_0^y D(v)D(y-v)dv = 2\int_0^{y/2} D(v)D(y-v)dv \\
&\sim 2\int_0^\infty D(v)e^{\Psi(y)}e^{-v\Psi'(y)}K(y)dv \\
&\sim 2D(0)K(y)e^{\Psi(y)} \frac{1}{\Psi'(y)}.
\end{aligned} \tag{169}$$

Recalling that $D(0) = u_0 = m_1 = t/[2(t-1)^2]$ and using (168) and (169) in (154) we get the leading order estimate

$$yK(y)e^{\Psi(y)} \sim \frac{\sqrt{2(t-1)}}{\sqrt{t}} \sqrt{\Psi'(y)} K(y) e^{\Psi(y)},$$

and hence

$$\Psi(y) = \frac{t}{6(t-1)} y^3 = \frac{t}{6(t-1)} a^2. \tag{170}$$

At the next order, using (168) and (169) in (154), after some simplification we obtain

$$0 = -\frac{\sqrt{t}}{\sqrt{2(t-1)}} \frac{1}{2\Psi'} + \frac{K'}{K}\frac{1}{\sqrt{\Psi'}} - \frac{1}{2(\Psi')^{3/2}} \left(\Psi'' + \Psi' \frac{K'}{K} \right) + \frac{3}{8}\frac{\Psi''}{(\Psi')^{3/2}}, \tag{171}$$

and thus

$$\frac{K'}{K} = \frac{\sqrt{t}}{\sqrt{2(t-1)}} \frac{1}{\Psi'} + \frac{1}{4}\frac{\Psi''}{\Psi'} = \frac{1}{y} + \frac{1}{2y} = \frac{3}{2y}. \tag{172}$$

By integrating (172) we have $K(y) = (const.')y^{3/2} = (const.')a$, so we have formally established that

$$\bar{C}_n(a) \sim (const.')a^2 \exp\left[\frac{t}{6(t-1)} a^2 \right], \quad a \to +\infty, \tag{173}$$

for some constant $const.'$. We find that $\bar{C}_n(a) \sim aD(a^{2/3})$ dominates the first term in the right hand side of (128) as $a \to +\infty$.

The asymptotic matching condition for the a- and β-scales is, in view of (128), (74) and (80),

$$\frac{1}{z_0^n n^{3/2}}\bar{C}_n(a)\bigg|_{a\to\infty} \sim e^{n\Phi(\beta)}n^{-1/2}g(\beta)\bigg|_{\beta\to 0^+}. \tag{174}$$

From (86), (96) and (104) we obtain, for $\beta \to 0^+$,

$$\Phi(\beta) = \ln\left(\frac{t^t}{(t-1)^{t-1}}\right) + \frac{t}{6(t-1)}\beta^2 + O(\beta^3), \quad g(\beta) \sim \frac{const.}{(t-1)^{3/2}}\beta^2. \tag{175}$$

Because $\beta = a/\sqrt{n}$ we find that the matching is indeed possible if

$$const. = (t-1)^{3/2} const.' \tag{176}$$

where $const.$ is the constant in (104). Our formal study suggests that the non-linear integral equation (158) may be estimated by a linear one for $y \to \infty$. The non-linear term does not affect the exponential growth rate $\Psi(y) = ty^3/[6(t-1)]$, however it does affect the algebraic factor $K(y) \propto y^{3/2}$.

To determine the constant in (176), we continue (162) into the range $a > 0$. Since $Y = (-a)^{2/3}$, $\arg(Y) = \pm\frac{2\pi}{3}$ for $a > 0$. It is convenient to define

$$h(s) = Ai\left(\left(\frac{t-1}{2t}\right)^{1/3} s\right). \tag{177}$$

By deforming the Bromwich contour in (162) to a piecewise linear one 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 paths, we obtain

$$\bar{D}(Y) - \frac{m_0}{a} = \frac{2t}{(t-1)^2}\frac{1}{2\pi i}\int_0^\infty \left[\frac{h'(\omega\tau)}{h(\omega\tau)}\omega e^{\omega\tau Y} - \frac{h'(\omega^2\tau)}{h(\omega^2\tau)}\omega^2 e^{\omega^2\tau Y}\right]d\tau, \quad \omega = e^{2\pi i/3}. \tag{178}$$

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

$$G_n(w) \sim \frac{2t}{(t-1)^2 z_0^n \pi n^{3/2}}(-a)\int_0^\infty \Re\left[e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}\omega e^{\omega\tau Y}\right]d\tau. \tag{179}$$

To evaluate the integral in (178) we shall use the Wronskian identity (cf. [16, 32])

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

which in term of $h(\cdot)$ leads to

$$e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)} + e^{-\pi i/6}\frac{h'(\omega^2\tau)}{h(\omega^2\tau)} = -\left(\frac{t-1}{2t}\right)^{1/3}\frac{1}{2\pi h(\omega\tau)h(\omega^2\tau)}. \tag{180}$$

The integral

$$I_1 = \int_0^\infty e^{\pi i/6}\frac{h'(\omega\tau)}{h(\omega\tau)}e^{\omega\tau Y}d\tau, \quad Y > 0, \tag{181}$$

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

$$I_2 = \int_0^\infty e^{-\pi i/6}\frac{h'(\omega^2\tau)}{h(\omega^2\tau)}e^{\omega^2\tau Y}d\tau, \quad Y > 0, \tag{182}$$

may be continued into the range $\arg(Y) \in \left(-\frac{5\pi}{6}, \frac{\pi}{6}\right)$ (in this range $\arg(\omega^2 Y) \in \left(-\frac{3\pi}{2}, -\frac{\pi}{2}\right)$). Using (180) we write $I_1 = I_3 + I_4$ where

$$\begin{aligned} I_3 &= -\left(\frac{t-1}{2t}\right)^{1/3} \frac{1}{2\pi} \int_0^\infty \frac{e^{\omega\tau Y}}{h(\omega\tau)h(\omega^2\tau)} d\tau, \\ I_4 &= -\int_0^\infty e^{-\pi i/6} \frac{h'(\omega^2\tau)}{h(\omega^2\tau)} e^{\omega\tau Y} d\tau, \end{aligned} \tag{183}$$

It is well known that (cf. [16]) as $z \to \infty$

$$Ai(z) \sim \frac{1}{2\sqrt{\pi}} z^{-1/4} \exp\left(-\frac{2}{3} z^{3/2}\right), \quad |\arg(z)| < \pi, \tag{184}$$

so that both $h(\omega\tau)$ and $h(\omega^2\tau)$ grow faster than exponentially as $\tau \to +\infty$, and hence I_3 is an entire function of Y. The integral I_4 in (183) is an analytic function defined in the range $arg(Y) \in \left(-\frac{\pi}{6}, \frac{5\pi}{6}\right)$ or $\arg(Y) \in \left(-\frac{13\pi}{6}, -\frac{7\pi}{6}\right)$. We let $\arg(Y) = -\frac{2\pi}{3}$ and set

$$Y = \omega^2 y = \omega^2 a^{2/3}, \tag{185}$$

with $a > 0$. Thus we have continued $I_1 + I_2$ into the range $a > 0$, using $I_1 + I_2 = I_2 + I_3 + I_4$. This continues the right side of (178) to $a > 0$. We next rotate τ in I_4 by $\omega^{-1} = e^{-2\pi i/3}$ and use (185). After some simplification, this yields

$$\begin{aligned} &\frac{2t}{(t-1)^2\pi}(-a) \int_0^\infty \Re\left[e^{\pi i/6} \frac{h'(\omega\tau)}{h(\omega\tau)} \omega e^{\omega\tau Y}\right] d\tau \\ =&\frac{ta}{2(t-1)^2\pi^2} \left(\frac{t-1}{2t}\right)^{1/3} \int_0^\infty \frac{e^{\tau y}}{h(\omega\tau)h(\omega^2\tau)} d\tau \\ &- \frac{2ta}{(t-1)^2\pi} \int_0^\infty \Re\left[e^{\pi i/6} \frac{h'(\omega\tau)}{h(\omega\tau)} e^{\omega^2\tau y}\right] d\tau. \end{aligned} \tag{186}$$

We expand the right hand side of (186) as $a = y^{3/2} \to +\infty$. We approximate the second integral by Watson's lemma, and find that it is $O(a/y) = O(\sqrt{y})$. We evaluate the first integral by Laplace's method. For $\tau \to \infty$ we use (184) to estimate the integrand, thus

$$h(\omega\tau)h(\omega^2\tau) = |h(\omega\tau)|^2 \sim \frac{1}{4\pi} \left(\frac{2t}{t-1}\right)^{1/6} \frac{1}{\sqrt{\tau}} \exp\left[\frac{4}{3}\left(\frac{t-1}{2t}\right)^{1/2} \tau^{3/2}\right],$$

Here we used the reflection principle, since $Ai(z)$ is real for real z. We thus obtain

$$\begin{aligned} &\frac{t}{2(t-1)^2\pi^2} \left(\frac{t-1}{2t}\right)^{1/3} \int_0^\infty \frac{e^{\tau y}}{h(\omega\tau)h(\omega^2\tau)} d\tau \\ \sim&\frac{\sqrt{2t}}{(t-1)^{3/2}\pi} \int_0^\infty e^{\tau y} \sqrt{\tau} \exp\left[-\frac{4}{3}\left(\frac{t-1}{2t}\right)^{1/2} \tau^{3/2}\right] d\tau \\ \sim&\frac{t}{(t-1)^2\pi} y \int_{-\infty}^\infty \exp\left[\frac{t}{6(t-1)} y^3 - \frac{t-1}{2ty}\left(\tau - \frac{t}{2(t-1)} y^2\right)^2\right] d\tau \\ =&\frac{\sqrt{2}t^{3/2}}{\sqrt{\pi}(t-1)^{5/2}} y^{3/2} \exp\left(\frac{t}{6(t-1)} y^3\right). \end{aligned} \tag{187}$$

The major contribution to the integral in (187) is from the point $\tau = \frac{t}{2(t-1)} y^2$. Since (187) dominates the second integral in the right side of (186), we see that the right hand side of (187) gives the expansion of

$D(y)$ for $y \to \infty$. This computation confirms the result obtained by the **WKB** *ansatz* (167), and determines the constant as

$$const. = \frac{\sqrt{2}t^{3/2}}{\sqrt{\pi}(t-1)}. \tag{188}$$

To calculate the correction term $C^{(1)}(a)$ in (136), we analyze the double transform equation (9) in the next subsection.

5.2 Double Transform

We re-consider the central region by using the functional equation (9). We introduce the scaling

$$w = 1 + \frac{a}{n^{3/2}}, \quad z = z_0\left(1 + \frac{\xi}{n}\right), \tag{189}$$

with

$$G(z, w) = \hat{G}(\xi, a) = \hat{G}\left(\left(\frac{z}{z_0} - 1\right)n, (w-1)n^{3/2}\right). \tag{190}$$

From (189) we have

$$\frac{zw}{z_0} - 1 = \frac{\xi}{n} + \left(1 + \frac{\xi}{n}\right)\frac{a}{n^{3/2}}, \tag{191}$$

and in terms of (ξ, a), (9) becomes

$$\hat{G}(\xi, a) - 1 = z_0\left(1 + \frac{\xi}{n}\right)\left[\hat{G}\left(\xi + \frac{a}{\sqrt{n}} + \frac{a\xi}{n^{3/2}}, a\right)\right]^t. \tag{192}$$

We define $G_1(\xi, a)$ by

$$\hat{G}(\xi, a) = a(z_0)\left[1 + \frac{1}{\sqrt{n}}G_1(\xi, a)\right] = \frac{t}{t-1}\left[1 + \frac{1}{\sqrt{n}}G_1(\xi, a)\right], \tag{193}$$

and then expand $G_1 = G_1(\xi, a; n)$ for $n \to \infty$ in the form

$$G_1(\xi, a; n) = G^{(0)}(\xi, a) + \frac{1}{\sqrt{n}}G^{(1)}(\xi, a) + \frac{1}{n}G^{(2)}(\xi, a) + O(n^{-3/2}). \tag{194}$$

Using (193) and (194) in (192) we obtain at the first two orders ($O(n^{-1})$ and $O(n^{-3/2})$) the equations

$$aG^{(0)}_{\xi} + \frac{\xi}{t} + \frac{t-1}{2}\left[G^{(0)}\right]^2 = 0, \tag{195}$$

and

$$aG^{(1)}_{\xi} + (t-1)G^{(0)}G^{(1)} + (t-1)aG^{(0)}G^{(0)}_{\xi} + \xi G^{(0)} + \frac{a^2}{2}G^{(0)}_{\xi\xi} + \frac{(t-1)(t-2)}{6}\left[G^{(0)}\right]^3 = 0. \tag{196}$$

To solve (195) we set

$$G^{(0)}(\xi, a) = \frac{2a}{t-1}\frac{H_\xi}{H}, \tag{197}$$

to obtain the Airy equation $\frac{2t}{t-1}a^2H_{\xi\xi} + \xi H = 0$ and thus

$$H(\xi, a) = Ai\left(\left(\frac{t-1}{2ta^2}\right)^{1/3}(-\xi)\right), \quad \xi < 0. \tag{198}$$

We use (198) in (197), note that

$$z^{-n-1}dz = z_0^{-n-1}\left(1+\frac{\xi}{n}\right)^{-n-1}\frac{z_0}{n}d\xi = \frac{1}{z_0^n n}e^{-\xi}\left[1+O\left(\frac{1}{n}\right)\right]d\xi, \tag{199}$$

and invert asymptotically the transform over z in (8). The leading term for $G_n(w)$ thus becomes

$$\begin{aligned} G_n(w) &= \frac{1}{2\pi i}\int_C \frac{G(z,w)}{z^{n+1}}dz \\ &\sim \frac{1}{z_0^n n}\frac{1}{2\pi i}\int_{Br^-}\frac{t}{t-1}e^{-\xi}\left[1+\frac{1}{\sqrt{n}}\frac{-2a}{(t-1)^{2/3}(2ta^2)^{1/3}}\frac{Ai'\left(-\left(\frac{t-1}{2t}\right)^{1/3}\xi|a|^{-2/3}\right)}{Ai\left(-\left(\frac{t-1}{2t}\right)^{1/3}\xi|a|^{-2/3}\right)}\right]d\xi. \end{aligned} \tag{200}$$

Here $\Re(\xi) < 0$ on Br^-. Setting $\xi = -(-a)^{2/3}s = -Ys$ for $a < 0$ and $Y > 0$, and interpreting the $O(1)$ term in the brackets in (200) as a distribution, via

$$\frac{1}{2\pi i}\int_{Br} e^{-\xi}d\xi = \frac{1}{2\pi i}\int_{Br} e^{Ys}Y ds = Y\delta(Y) = 0,$$

(200) becomes

$$\begin{aligned} G_n(w) &\sim \frac{1}{z_0^n n^{3/2}}\frac{-(2t)^{2/3}a}{2\pi i(t-1)^{5/3}}\int_{Br}\frac{Ai'\left(\left(\frac{t-1}{2t}\right)^{1/3}s\right)}{Ai\left(\left(\frac{t-1}{2t}\right)^{1/3}s\right)}e^{Ys}ds \\ &= \frac{1}{z_0^n n^{3/2}}\frac{-(2t)^{2/3}a}{2\pi i(t-1)^{5/3}}\frac{d}{dY}\left(\int_{Br}\frac{Ai'\left(\left(\frac{t-1}{2t}\right)^{1/3}s\right)}{Ai\left(\left(\frac{t-1}{2t}\right)^{1/3}s\right)}\frac{e^{Ys}}{s}ds\right). \end{aligned} \tag{201}$$

The first integral in (201) must be also interpreted as distribution. We thus have regained the leading term for the scale $w = 1 + O(n^{-3/2})$ and $w < 1$, as (201) is equivalent to (164).

We next obtain the correction term $G^{(1)}(\xi, a)$. Differentiating (195) and rearranging leads to

$$taG^{(0)}_{\xi\xi} + 1 + t(t-1)G^{(0)}G^{(0)}_{\xi} = 0. \tag{202}$$

We rewrite the linear equation (196) as

$$aG^{(1)}_{\xi} + (t-1)G^{(0)}G^{(1)} = -\xi G^{(0)} - \frac{a^2}{2}G^{(0)}_{\xi\xi} - (t-1)aG^{(0)}G^{(0)}_{\xi} - \frac{(t-1)(t-2)}{6}\left[G^{(0)}\right]^3. \tag{203}$$

Using (197) and (202) in (203), and multiplying by H^2, we obtain

$$\frac{d}{d\xi}\left[H^2G^{(1)}\right] = -\frac{(t+1)}{t(t-1)}\xi HH_{\xi} + \frac{1}{2t}H^2 + \frac{2(t+1)}{3(t-1)^2}a^2\frac{H_{\xi}^3}{H}, \tag{204}$$

where

$$H(\xi) = h(s) = Ai\left(\left(\frac{t-1}{2t}\right)^{1/3}s\right),\ \ \xi = -(-a)^{2/3}s. \tag{205}$$

For $\xi \to -\infty$, (184) and (205) yield

$$h(s) = H(\xi) \sim \exp\left[-\frac{2}{3}(-\xi)^{3/2}\left(\frac{t-1}{2ta^2}\right)^{1/2}\right], \tag{206}$$

and hence

$$\frac{H_\xi}{H} \sim \frac{\sqrt{t-1}\sqrt{-\xi}}{\sqrt{2}t(-a)}, \ \ \xi \to -\infty. \tag{207}$$

Thus the right side of (204) is asymptotic to

$$\begin{aligned}&(-\xi)^{3/2}H^2(\xi)\left[-\frac{1}{a}\left(\frac{t+1}{t\sqrt{2t(t-1)}}+\frac{t+1}{3t\sqrt{2t(t-1)}}\right)\right],\\ =&(-\xi)^{3/2}H^2(\xi)\left(\frac{-1}{a}\right)\frac{4(t+1)}{3t\sqrt{2t(t-1)}}.\end{aligned} \tag{208}$$

Due to the algebraic factor $(-\xi)^{3/2}$ in (208), the solution to (204) would grow like $(-\xi)^{5/2}$. To avoid this growth we set

$$G^{(1)} = a_*\xi + \hat{G}^{(1)}, \tag{209}$$

with which (204) becomes

$$\frac{d}{d\xi}\left[H^2\hat{G}^{(1)}\right] = -a_*(H^2+2\xi HH_\xi) - \frac{(t+1)}{t(t-1)}\xi HH_\xi + \frac{1}{2t}H^2 + \frac{2(t+1)}{3(t-1)^2}\,a^2\frac{H_\xi^3}{H}. \tag{210}$$

By choosing $a_* = \frac{-2(t+1)}{3t(t-1)}$ we can avoid the growth as $\xi \to -\infty$, and then (210) becomes

$$\frac{d}{d\xi}\left[H^2\hat{G}^{(1)}\right] = \frac{7t+1}{6t(t-1)}H^2 + \frac{(t+1)}{3t(t-1)}\xi HH_\xi + \frac{2(t+1)}{3(t-1)^2}\,a^2\frac{H_\xi^3}{H}. \tag{211}$$

As $\xi \to -\infty$, the right side of (211) will be $O(H^2)$, rather than $O((-\xi)^{3/2}H^2)$.

In term of s, (211) becomes

$$-\frac{1}{Y}\frac{d}{ds}\left[h^2(s)\hat{G}^{(1)}\right] = \frac{7t+1}{6t(t-1)}h^2(s) + \frac{(t+1)}{3t(t-1)}sh(s)h'(s) - \frac{2(t+1)}{3(t-1)^2}\frac{a^2}{Y^3}\frac{(h'(s))^3}{h(s)}. \tag{212}$$

To solve (212), we let

$$\hat{G}^{(1)} = \frac{t-1}{t}Y\mathcal{E}_*(s), \tag{213}$$

with which (212) becomes

$$\frac{d}{ds}\left[h^2(s)\mathcal{E}_*(s)\right] = -\frac{(7t+1)}{6(t-1)^2}h^2(s) - \frac{(t+1)}{3(t-1)^2}sh(s)h'(s) + \frac{2t(t+1)}{3(t-1)^3}\frac{(h'(s))^3}{h(s)}. \tag{214}$$

Before solving (214), we first establish a relationship between $C^{(1)}$ and $G^{(1)}$ in (194). From (128) and (136) we have

$$G_n(w) - G_n(1) \sim \frac{1}{z_0^n n^{3/2}}\left[C^{(0)}(a) + \frac{1}{\sqrt{n}}C^{(1)}(a) + O(n^{-1})\right], \tag{215}$$

and inverting (8) using (194), (209) and (213) leads to

$$\begin{aligned}G_n(w) - G_n(1) &= \frac{1}{2\pi i}\int_C \frac{G(z,w)-G(z,1)}{z^{n+1}}dz\\ &\sim \frac{1}{z_0^n n}\frac{1}{2\pi i}\int_{Br^-}\left(\frac{t}{t-1}\right)e^{-\xi}\left[\frac{1}{\sqrt{n}}G^{(0)}(\xi,a) + \frac{1}{n}\hat{G}^{(1)}(\xi,a) + O(n^{-3/2})\right]d\xi\\ &= \frac{1}{z_0^n n^{3/2}}\frac{1}{2\pi i}\int_{Br^-} e^{Ys}\left[\left(\frac{t}{t-1}\right)YG^{(0)}(-Ys,a) + \frac{1}{\sqrt{n}}Y^2\mathcal{E}_*(s) + O(n^{-1})\right]ds.\end{aligned} \tag{216}$$

Comparing (215) with (216), we conclude that

$$C^{(1)}(a) = \frac{1}{2\pi i}\int_{Br^-} e^{Ys}Y^2\mathcal{E}_*(s)ds. \tag{217}$$

We define $\bar{D}_1(Y)$ from

$$C^{(1)}(a) = (-a)\bar{D}_1((-a)^{2/3}) = Y^{3/2}\bar{D}_1(Y), \quad a < 0, \tag{218}$$

and note that (132), (134) and (136) give

$$\bar{D}_1(Y) = \sum_{L=1}^{\infty}\frac{\bar{m}_L}{L!}Y^{\frac{3}{2}(L-1)}(-1)^L, \tag{219}$$

where $\bar{m}_L$ is defined in (134). Using (217) and (218), we can express $\mathcal{E}_*(s)$ in term of $\bar{D}_1(Y)$ as

$$\mathcal{E}_*(s) = \int_0^{\infty} e^{-sY}\frac{\bar{D}_1(Y)}{\sqrt{Y}}dY. \tag{220}$$

We return to (214), and write its general solution as

$$\mathcal{E}_*(s) = \frac{c_*}{h^2(s)} + \frac{1}{h^2(s)}\int_s^{\infty}\left[\frac{(7t+1)}{6(t-1)^2}h^2(v) + \frac{(t+1)}{3(t-1)^2}sh(v)h'(v) - \frac{2t(t+1)}{3(t-1)^3}\frac{(h'(v))^3}{h(v)}\right]dv. \tag{221}$$

However, $h(s)$ decays faster than exponentially as $s \to \infty$ so that $h^{-2}(s)$ has unacceptable growth. Thus we must set $c_* = 0$ in order for $\mathcal{E}_*(s)$ to be a proper Laplace transform. Using the fact that $h(s)$ satisfies (160) and integrating by parts yields

$$\int^s vh(v)h'(v)dv = \frac{2t}{t-1}\int^s h'(v)h''(v)dv = \frac{t}{t-1}(h'(s))^2,$$
$$\int^s h^2(v)dv = sh^2(s) - 2\int^s vh(v)h'(v)dv = sh^2(s) - \frac{t}{t-1}(h'(s))^2,$$

and thus (221) simplifies to

$$\mathcal{E}_*(s) = -\frac{(7t+1)}{6(t-1)^2}s + \frac{2t^2}{(t-1)^3}\left(\frac{h'(s)}{h(s)}\right)^2 - \frac{2t(t+1)}{3(t-1)^3h^2(s)}\int_s^{\infty}\frac{(h'(v))^3}{h(v)}dv. \tag{222}$$

We have thus obtained (44) and hence (42), since for $a < 0$

$$\begin{aligned}
G_n(w) &= \sum_{j=0}^{\infty} M_{j,n}\frac{(w-1)^j}{j!} \\
&= \frac{1}{z_0^n n^{3/2}}\left[\sum_{j=0}^{\infty} m_j\frac{a^j}{j!} + \frac{1}{\sqrt{n}}\sum_{j=0}^{\infty}\bar{m}_j\frac{a^j}{j!} + O(n^{-1})\right] \\
&= \frac{1}{z_0^n n^{3/2}}\left[m_0 + \sum_{j=1}^{\infty} m_j\frac{a^j}{j!} + \frac{1}{\sqrt{n}}\sum_{j=1}^{\infty}\bar{m}_j\frac{a^j}{j!} + O(n^{-1})\right] \\
&= \frac{1}{z_0^n n^{3/2}}\left[\frac{\sqrt{t}}{\sqrt{2\pi}(t-1)^{3/2}} - a\bar{D}((-a)^{2/3}) + \frac{1}{\sqrt{n}}(-a)\bar{D}_1((-a)^{2/3}) + O(n^{-1})\right] \\
&= \frac{1}{z_0^n n^{3/2}}\left[\frac{\sqrt{t}}{\sqrt{2\pi}(t-1)^{3/2}} + Y^{3/2}\bar{D}(Y) + \frac{1}{\sqrt{n}}Y^{3/2}\bar{D}_1(Y) + O(n^{-1})\right].
\end{aligned} \tag{223}$$

To asymptotically match the a-scale result to that valid for $w < 1$ and $w - 1 = O(n^{-1})$, we shall need the behavior of $\bar{D}_1(Y)$ as $Y = (-a)^{2/3} \to \infty$. This will be determined by the singularity of $\mathcal{E}_*(s)$ with the largest real part. The Laplace transform of the leading term $\bar{D}(Y)$ has simple poles at $(\frac{2t}{t-1})^{1/3} r_j$ where r_j are the negative roots of the Airy function and the pole at $(\frac{2t}{t-1})^{1/3} r_0$ determines the asymptotic behavior as $Y \to \infty$ $(a \to -\infty)$. The dominant singularity of $\mathcal{E}_*(s)$ is also at $(\frac{2t}{t-1})^{1/3} r_0$. However, (222) has a more complicated structure, as there is a double pole combined with a logarithmic branch point at $s = s_0 \equiv (\frac{2t}{t-1})^{1/3} r_0$.

We expand (222) near the dominant singularity at s_0. Integrating by parts we have

$$\begin{aligned}\int_s^\infty \frac{(h'(v))^3}{h(v)} dv &= \int_s^\infty (h'(v))^2 d[\ln(h(v))] \\ &= -(h'(s))^2[\ln(h(s))] - \int_s^\infty 2h''(v)h'(v)[\ln(h(v))]dv \\ &= -(h'(s))^2[\ln(h(s))] - \frac{t}{t-1}\int_s^\infty vh'(v)h(v)[\ln(h(v))]dv,\end{aligned} \tag{224}$$

and

$$\begin{aligned}&\int_s^\infty vh'(v)h(v)[\ln(h(v))]dv = \int_s^\infty vh(v)[\ln(h(v))]dh(v) \\ &= -sh^2(s)\ln[h(s)] - \int_s^\infty \left\{h^2(v)\ln[h(v)] + vh(v)h'(v)\ln[h(v)] + vh(v)h'(v)\right\} dv,\end{aligned} \tag{225}$$

which may be re-arranged to

$$2\int_s^\infty vh'(v)h(v)\ln[h(v)]dv = -sh^2(s)\ln[h(s)] - \int_s^\infty h^2(v)\ln[h(v)]dv + \frac{2t}{t-1}(h'(s))^2. \tag{226}$$

Using (224)-(226) in (222) we obtain the following alternative form for $\mathcal{E}_*(s)$

$$\begin{aligned}\mathcal{E}_*(s) = &-\frac{(7t+1)}{6(t-1)^2}s + \frac{t(7t+1)}{3(t-1)^3}\left(\frac{h'(s)}{h(s)}\right)^2 + \frac{2t(t+1)}{3(t-1)^3}\left(\frac{h'(s)}{h(s)}\right)^2 \ln[(h(s)] \\ &- \frac{(t+1)}{3(t-1)^2}s\ln[h(s)] - \frac{(t+1)}{3(t-1)^2h^2(s)}\int_s^\infty h^2(v)\ln[h(v)]dv,\end{aligned} \tag{227}$$

which is convenient for studying the limit $s \to s_0$. Next we let $\tau = s - s_0$ and expand (227) about $\tau = 0$. Noting that $h(s_0) = h''(s_0) = 0$ and

$$\begin{aligned}\int_s^\infty h^2(v)[\ln h(v)]dv &= \int_{s_0}^\infty h^2(v)[\ln h(v)]dv - \int_{s_0}^s h^2(v)[\ln h(v)]dv \\ &= \int_{s_0}^\infty h^2(v)[\ln h(v)]dv + O(\tau^3\ln\tau),\end{aligned} \tag{228}$$

(227) leads to

$$\begin{aligned}\mathcal{E}_*(s) = &\frac{2t(t+1)}{3(t-1)^3\tau^2}\ln[h'(s_0)\tau] + \frac{t(7t+1)}{3(t-1)^3\tau^2} \\ &- \frac{(t+1)}{3(t-1)^2[h'(s_0)]^2\tau^2}\int_{s_0}^\infty h^2(v)\ln[h(v)]dv + O(\ln\tau), \quad \tau \to 0.\end{aligned} \tag{229}$$

Here

$$h'(s_0) = \left(\frac{t-1}{2t}\right)^{1/3} Ai'\left(\left(\frac{t-1}{2t}\right)^{1/3} s_0\right) = \left(\frac{t-1}{2t}\right)^{1/3} Ai'(r_0). \tag{230}$$

We define $\kappa = \kappa(t)$ by

$$\kappa \equiv \frac{2t(t+1)}{3(t-1)^3}\ln[h'(s_0)] + \frac{t(7t+1)}{3(t-1)^3} - \frac{(t+1)}{3(t-1)^2[h'(s_0)]^2}\int_{s_0}^{\infty} h^2(v)\ln[h(v)]dv, \tag{231}$$

and thus we obtain

$$\begin{aligned}\bar{D}_1(Y) &= \frac{\sqrt{Y}}{2\pi i}\int_{Br} e^{sY}\mathcal{E}_*(s)ds \\ &= \sqrt{Y}\exp\left(-|r_0|\left(\frac{2t}{t-1}\right)^{1/3} Y\right)\frac{1}{2\pi i}\int_{Br_+} e^{sY}\left[\frac{2t(t+1)}{3(t-1)^3}\frac{\ln\tau}{\tau^2} + \frac{\kappa}{\tau^2} + O(\ln\tau)\right]d\tau,\end{aligned} \tag{232}$$

where $\Re(\tau) > 0$ on Br_+. Using $e^{sY} = e^{s_0 Y}e^{\tau Y}$ and explicitly evaluating the integral in (232) we have

$$\begin{aligned}\bar{D}_1(Y) &= Y^{3/2}\exp\left(-|r_0|\left(\frac{2t}{t-1}\right)^{1/3} Y\right) \\ &\times\left[-\frac{2t(t+1)}{3(t-1)^3}\ln Y + \kappa + \frac{2t(t+1)}{3(t-1)^3}(1-\gamma_E) + O\left(\frac{\ln Y}{Y}\right)\right], \quad Y\to\infty,\end{aligned} \tag{233}$$

where γ_E is the Euler constant.

Combining (233) and (164) we have shown that for $Y \to \infty$ the two term approximation on the a-scale behaves as

$$\begin{aligned}G_n(w) &= \frac{1}{z_0^n n^{3/2}}\left[\frac{\sqrt{t}}{\sqrt{2\pi}(t-1)^{3/2}} + Y^{3/2}\bar{D}(Y) + \frac{1}{\sqrt{n}}Y^{3/2}\bar{D}_1(Y) + O(n^{-1})\right] \\ &\sim \frac{1}{z_0^n n^{3/2}}\exp\left[-|r_0|\left(\frac{2t}{t-1}\right)^{1/3}(-a)^{2/3}\right] \\ &\times\left\{\frac{2t}{(t-1)^2}(-a) + \frac{a^2}{\sqrt{n}}\left[-\frac{4t(t+1)}{9(t-1)^3}\ln(-a) + \kappa + \frac{2t(t+1)}{3(t-1)^3}(1-\gamma_E)\right]\right\}.\end{aligned} \tag{234}$$

This expression applies in the limit $w \to 1$ but $a = n^{3/2}(w-1) \to -\infty$ and will be used in section 6.

5.3 Transform Inversion

We now invert the transform over w using (10) and obtain an approximation to $g(n,p)$. We assume that

$$\Omega \equiv pn^{-3/2} = O(1), \quad 0 < \Omega < \infty, \tag{235}$$

and with (127) we obtain

$$w^{-p-1}dw = \left(1 + \frac{a}{n^{3/2}}\right)^{-1-\Omega n^{3/2}} n^{-3/2}\, da = \frac{e^{-a\Omega}}{n^{3/2}}[1 + O(n^{-3/2})]da. \tag{236}$$

Thus the scale $w - 1 = O(n^{-3/2})$ in the transform space translates to (235) in the (n,p) space, and (236) and (234) lead to

$$\begin{aligned}g(n,p) &= \frac{1}{z_0^n n^3}\frac{1}{2\pi i}\int_{Br_a}\left[m_0 + C^{(0)}(a) + \frac{1}{\sqrt{n}}C^{(1)}(a) + O(n^{-1})\right]e^{-a\Omega}da \\ &\sim \frac{1}{z_0^n n^3}\frac{1}{2\pi i}\int_{Br_a}\left[m_0 - a\bar{D}(Y)\right]e^{-a\Omega}da \\ &= \frac{1}{z_0^n n^3}(-a)\frac{1}{(2\pi i)^2}\int_{Br_a}\left[\frac{d}{dY}\int_{Br}\frac{2t}{(t-1)^2}\frac{h'(s)}{sh(s)}ds\right]e^{-a\Omega}da.\end{aligned} \tag{237}$$

Here we take $\Re(a) < 0$ on the vertical contour Br_a, since we used (163) for $\bar{D}(Y)$.

We note that as $-a$ goes from $-\infty\, i$ to $+\infty\, i$, $Y = (-a)^{2/3}$ goes from $\infty e^{-\pi i/3}$ to $\infty e^{\pi i/3}$. To evaluate the double integral in (237), we let

$$Y = Z^2,$$

and the contour Br_a is mapped to C_*, which goes from $\infty e^{-\pi i/6}$ to $\infty e^{\pi i/6}$ in the Z-plane. Thus (237) becomes

$$\begin{aligned} g(n,p) &\sim \frac{1}{z_0^n n^3}\frac{1}{(2\pi i)^2}\int_{C_*} e^{\Omega Z^3}\frac{3t}{(t-1)^2}Z^4\left[\frac{d}{dZ}\int_{Br}\frac{h'(s)}{sh(s)}e^{sZ^2}ds\right]dZ \\ &= \frac{1}{z_0^n n^3}\frac{-1}{(2\pi i)^2}\int_{\infty\omega^2}^{\infty\omega} e^{-\Omega Z^3}\frac{6t}{(t-1)^2}Z^5\left[\int_{Br}\frac{h'(s)}{h(s)}e^{sZ^2}ds\right]dZ, \end{aligned} \tag{238}$$

where we changed Z to $-Z$ and used $\omega = e^{2\pi i/3}$ in the last integral. By integrating by parts and after re-arrangement, we obtain

$$g(n,p) \sim \frac{1}{z_0^n n^3}\frac{-1}{(2\pi i)^2}\int_{\infty\omega^2}^{\infty\omega}\int_{Br} e^{-\Omega Z^3}e^{sZ^2}\frac{h'(s)}{h(s)}\left[\frac{20t}{3(t-1)^2}\frac{Zs}{\Omega^2}+\frac{8t}{9(t-1)^2}\frac{s^2(1+2sZ^2)}{\Omega^3}\right]dsdZ. \tag{239}$$

Furthermore we set

$$Z = \frac{s}{3\Omega} + W, \quad W = \frac{\xi}{(3\Omega)^{1/3}}, \tag{240}$$

and recall the integral representation of the Airy function

$$Ai(U) = \frac{1}{2\pi i}\int_{\infty\omega^2}^{\infty\omega} e^{\xi U}e^{-\xi^3/3}d\xi. \tag{241}$$

Differentiating (241) with respect to U corresponds to multiplying the integrand by ξ. Thus (239) yields

$$\begin{aligned} g(n,p) \sim &-\frac{1}{z_0^n n^3}\frac{1}{2\pi i}\int_{Br}\frac{h'(s)}{h(s)}\exp\left[\frac{2s^3}{27\Omega^2}\right]\frac{1}{(3\Omega)^{1/3}}\left\{\frac{20t}{3(t-1)^2}\frac{s}{\Omega^2(3\Omega)^{1/3}}Ai'(U(s))\right. \\ &+\left[\frac{20t}{3(t-1)^2 3\Omega^3}+\frac{8t}{9(t-1)^2\Omega^3}\right]s^2 Ai(U(s)) + \frac{16t}{9(t-1)^2}\frac{s^3}{\Omega^3(3\Omega)^{2/3}}Ai''(U(s)) \\ &\left.+\frac{32t}{9(t-1)^2}\frac{s^4}{\Omega^4(3\Omega)^{1/3}}Ai'(U(s)) + \frac{16t}{81(t-1)^2}\frac{s^5}{\Omega^5}Ai(U(s))\right\}ds, \end{aligned} \tag{242}$$

where we evaluated the integral over Z and set

$$U(s) = \frac{s^2}{(3\Omega)^{4/3}}. \tag{243}$$

Using the fact that $Ai''(U(s)) = U(s)Ai(U(s))$ to simplify (242) and noting that integrand has singularities at $s = (\frac{2t}{t-1})^{1/3}r_j$, we evaluate the integral using the residue theorem and thus establish (57). It is also possible to get an $O(n^{-1/2})$ correction to (242) or (57), by using (222) and (232) to compute $C^{(1)}(a)$ in (237), and then evaluate the integral in a manner similar to that above.

Finally we discuss the asymptotic behavior of the right side of (57) (or (242)) for $\Omega \to 0$ and $\Omega \to \infty$. For $\Omega \to 0$ the dominant term in the sum in (57) is $j = 0$, which is exponentially larger than the other terms, and thus we obtain

$$\begin{aligned} &\frac{1}{z_0^n n^3}\frac{1}{(3\Omega)^{1/3}}\exp\left(\frac{-4t|r_0|^3}{27(t-1)\Omega^2}\right)\frac{64t}{81(t-1)^2}\left(\frac{2t}{t-1}\right)^{5/3}\frac{|r_0|^5}{\Omega^5}Ai\left(\left(\frac{2t}{t-1}\right)^{2/3}\frac{r_0^2}{(3\Omega)^{4/3}}\right) \\ &\sim \frac{1}{z_0^n n^3}\exp\left(\frac{-8t|r_0|^3}{27(t-1)\Omega^2}\right)\frac{64\sqrt{2}t^{5/2}}{81(t-1)^{7/2}}\frac{|r_0|^{9/2}}{\sqrt{\pi}\Omega^5}. \end{aligned} \tag{244}$$

This gives an approximation to $g(n, p)$ that applies for $n^{4/3} \ll p \ll n^{3/2}$. In (244) we used the asymptotic expansion (184) of the Airy function. Note that (244) arises from two parts of the $j = 0$ term(s) in (57), the part proportional to $r_0^5 Ai(\cdot)$ and that proportional to $r_0^4 Ai'(\cdot)$. In section 6 we will use (244) to discuss the asymptotic matching between the scales $p = O(n^{3/2})$ and $p = O(n^{4/3})$. The limit $\Omega \to \infty$ corresponds to the left tail of the Airy distribution, and corresponds to the expression in (62), which applies for $n^{3/2} \ll p \ll n^2$.

The limit $\Omega \to \infty$ is difficult to obtain from the sum in (57), due to a lot of cancelation. However, we have previously established the asymptotic matching between the expansions for $w = 1+O(n^{-3/2})$ ($a-$scale) and $w = 1 + O(n^{-1})$ ($\beta-$ scale). This implies the matching in (n, p) space of the expansions for $p = O(n^{3/2})$ (Ω – scale), (where (242) applies), and for $p = O(n^2)$ (Λ – scale). Thus the behavior of (242) or (57) as $\Omega \to \infty$ is the same as that of (117) as $\Lambda = p/n^2 = \Omega/\sqrt{n} \to 0$. The latter is easily obtained from (113) and (117), as we show below.

As $\Lambda \to 0$, we have $\beta_* \to 0$ and from (87) we get

$$\phi(\beta) = -(t-1)\ln\frac{t-1}{t} - \frac{\beta}{2} + \frac{t\beta^2}{6(t-1)} + O(\beta^3), \ \ \beta \to 0. \tag{245}$$

From (113) we have

$$\beta_* = \beta_*(\Lambda) \sim \frac{3(t-1)}{t}\Lambda, \ \ \Lambda \to 0, \tag{246}$$

so that

$$1 - 2\Lambda - \frac{t-1}{te^{\beta_*} - 1} \sim \Lambda, \tag{247}$$

and

$$\beta_*(1-2\Lambda) - (t-1)\ln\left(1 - \frac{e^{-\beta_*}}{t}\right) \sim (t-1)\ln\left(\frac{t}{t-1}\right) - \frac{3(t-1)}{2t}\Lambda^2. \tag{248}$$

Using (246)-(248) in the right side of (117) yields

$$\begin{aligned}
&\frac{t^{n+1}}{n^2\pi}\frac{\sqrt{t}}{(t-1)^3}\frac{\beta_*^{5/2}}{\sqrt{\Lambda}}\exp\left\{n\left[(t-1)\ln\left(\frac{t}{t-1}\right) - \frac{3(t-1)}{2t}\Lambda^2\right]\right\} \\
&\sim \frac{t^n}{n^2\pi}\frac{3^{5/2}}{t\sqrt{t-1}}\Lambda^2\left(\frac{t}{t-1}\right)^{n(t-1)}\exp\left(-\frac{3(t-1)}{2t}n\Lambda^2\right) \\
&= \frac{1}{z_0^n n^3\pi}\frac{9\sqrt{3}}{t\sqrt{t-1}}\Omega^2\exp\left(-\frac{3(t-1)}{2t}\Omega^2\right), \ \ \Omega = \frac{p}{n^{3/2}}.
\end{aligned} \tag{249}$$

This is the asymptotic behavior of (242) or (57) as $\Omega \to \infty$. We have completed the analysis of the scales $w - 1 = O(n^{-3/2})$ and $p = O(n^{3/2})$.

6. LEFT REGION

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

$$w = 1 - \frac{\gamma}{n}, \ \gamma > 0. \tag{250}$$

Furthermore, we set

$$G_n(w) = \frac{1}{z_0^n n}e^{v_0 n^{1/3}\gamma^{2/3}}e^{v_1\gamma\ln n}F(\gamma; n). \tag{251}$$

Here v_0 and v_1 are constants that will be determined soon. The form in (251) is suggested by the behavior of the a-scale result as $a \to -\infty$ (cf. (134) with $-a = \gamma\sqrt{n}$).

6.1 Analysis of the Basic Recurrence

Using (251) in (7) yields

$$\begin{aligned}
&\frac{1}{z_0^{n+1}(n+1)}\exp\left[v_0(n+1)^{1/3}\left[\left(1+\frac{1}{n}\right)\gamma\right]^{2/3}+v_1\gamma\ln(n+1)\right]F\left(\gamma\left(1+\frac{1}{n}\right);n+1\right)\\
&=\binom{t}{1}\left(1-\frac{\gamma}{n}\right)^n\sum_{(i_1,i_2,\cdots,i_t)\in\mathcal{S}_1}\frac{1}{z_0^{i_1}i_1}\exp\left(v_0 i_1^{1/3}\left(\gamma\frac{i_1}{n}\right)^{2/3}+v_1\gamma\ln i_1\right)F\left(\frac{i_1}{n}\gamma;i_1\right)\\
&\times\left[\prod_{k=2}^{t}G_{i_k}\left(1-\frac{\gamma}{n}\right)\right]\\
&+\binom{t}{2}\left(1-\frac{\gamma}{n}\right)^n\sum_{(i_1,i_2,\cdots,i_t)\in\mathcal{S}_2}\prod_{j=1}^{2}\frac{1}{z_0^{i_j}i_j}\exp\left(v_0 i_j^{1/3}\left(\gamma\frac{i_j}{n}\right)^{2/3}+v_1\gamma\ln i_j\right)F\left(\frac{i_j}{n}\gamma;i_j\right)\\
&\times\left[\prod_{k=3}^{t}G_{i_k}\left(1-\frac{\gamma}{n}\right)\right]\\
&+\cdots\\
&+\binom{t}{t-1}\left(1-\frac{\gamma}{n}\right)^n\sum_{(i_1,i_2,\cdots,i_t)\in\mathcal{S}_{t-1}}\left[\prod_{k=1}^{t-1}\frac{1}{z_0^{i_k}i_k}\exp\left(v_0 i_k^{1/3}\left(\gamma\frac{i_k}{n}\right)^{2/3}+v_1\gamma\ln i_k\right)F\left(\frac{i_k}{n}\gamma;i_k\right)\right]\\
&\times G_{i_t}\left(1-\frac{\gamma}{n}\right)\\
&+\left(1-\frac{\gamma}{n}\right)^n\sum_{(i_1,i_2,\cdots,i_t)\in\mathcal{S}_t}\left[\prod_{k=1}^{t}\frac{1}{z_0^{i_k}i_k}\exp\left(v_0 i_k^{1/3}\left(\gamma\frac{i_k}{n}\right)^{2/3}+v_1\gamma\ln i_k\right)F\left(\frac{i_k}{n}\gamma;i_k\right)\right],
\end{aligned}\tag{252}$$

where

$$\mathcal{S}_J=\{i_1,i_2,\cdots,i_J=O(n);i_{J+1},\cdots,i_t=O(1)\},\quad J=1,\ 2,\ \cdots,\ t;$$

and

$$\mathcal{S}_1\cup\mathcal{S}_2\cup\cdots\cup\mathcal{S}_t=\{(i_1,i_2,\cdots,i_t):i_1+i_2+\cdots+i_t=n\}.$$

We need to say a few words about the form of (252). Note that in (252) we wrote the product $G_{i_1}G_{i_2}\cdots G_{i_t}$ using the form (251) for the first few i_k (with $1\le k\le J$) and used $G_{i_k}(w)=G_{i_k}(1-\gamma/n)$ for the remaining i_k (with $J+1\le k\le t$). We proceed to give a simple geometric interpretation of (252). For the binary case $t=2$ we found in [23] that in the single sum over $i_1+i_2=n$ in (7), two ranges are equally important asymptotically. These correspond to (1) $i_1=O(1)$ and $i_2=n-O(1)$ or $i_2=O(1)$ and $i_1=n-O(1)$, and (2) i_1,i_2 both $O(n)$. The first range corresponds to the endpoints of the line segment $i_1+i_2=n$ ($i_1,i_2\ge 0$) and the second range to the interior of the segment. In the first range we can use (252) to approximate one of the two factors in $G_{i_1}(w)G_{i_2}(w)$ (that which has $i_k=O(n)$) but not the other. If both $i_k=O(n)$ then we use (252) for both factors. Now consider $t=3$ where the double sum in (252) corresponds to summing over points in the lattice triangle $i_1+i_2+i_3=n$. We view this triangle as having corners at $(n,0,0)$, $(0,n,0)$, and $(0,0,n)$, three edges where, say, $i_3=0$ and $i_1+i_2=n$, and an interior where i_1,i_2,i_3 are all $O(n)$. Near the corner $(n,0,0)$, the first factor in $G_{i_1}(w)G_{i_2}(w)G_{i_3}(w)$ may be approximated by using (252), but not the other two. By symmetry the analysis will be the same at all 3 corners. Similarly, near an edge two of the three factors can be approximated using (252). In the interior we use (252) for all three factors. Then we find that these three geometric regions all contribute equally as $n\to\infty$. The corner contributions will be proportional to $F(\gamma)$, the edge contributions will involve a convolution integral of F with itself, and the interior contribution will be a double convolution integral, with the integrand involving $F(\gamma x)F(\gamma y)F(\gamma(1-x-y))$. The latter will correspond to a cubic non-linearity in the limiting integral equation. The equation (252) corresponds to generalizing the above argument to arbitrary t, where we view the $(t-1)$ fold sum in (7) (where $i_1+i_2+\cdots+i_t=n$) as being over a $t-1$ dimensional lattice hyper-triangle. This hyper-triangle has t corners, $\binom{t}{2}$ edges, etc., which correspond to the binomial coefficients in (252). We will now show that the corners, edges, $\cdots$, interior all contribute equally for $n\to\infty$, and this will lead to a limiting non-linear integral equation for $F(\gamma;n)$. Note that the last sum in (252) corresponds to the interior of the hyper-triangle, and the terms proportional to $\binom{t}{J}$ will have $i_1,i_2,\cdots,i_J=O(n)$ and $i_{J+1},\cdots,i_t=O(1)$.

Later, in subsection 6.2, we will analyze the scale (250) using the double transform equation (9), and this will involve less combinatorial analysis.

We note that

$$\sum_{k=1}^{t} i_k^{1/3}\left(\frac{i_k}{n}\right)^{2/3}=n^{1/3},$$

and define $H(\cdot)$ by

$$\sum_{k=1}^{t}\frac{i_k}{n}\ln(i_k)=\ln n+H(\vec{x}(t)),$$

where

$$H(\vec{x}(m))=\sum_{j=1}^{m}x_j\ln(x_j),\qquad \sum_{j=1}^{m}x_j=1. \tag{253}$$

We assume that

$$F(\gamma;n)\to F_0(\gamma),\qquad n\to\infty. \tag{254}$$

Then the first sum in (252) is asymptotic to

$$\begin{aligned}&\binom{t}{1}\left(1-\frac{\gamma}{n}\right)^n\sum_{(i_1,i_2,\cdots,i_t)\in\mathcal{S}_1}\frac{1}{z_0^{i_1}i_1}\exp\left(v_0 i_1^{1/3}\left(\gamma\frac{i_1}{n}\right)^{2/3}+v_1\gamma\ln i_1\right)F\left(\frac{i_1}{n}\gamma;i_1\right)\\&\times\left[\prod_{k=2}^{t}G_{i_k}\left(1-\frac{\gamma}{n}\right)\right]\\&\sim\frac{1}{z_0^n n}\exp(v_0n^{1/3}\gamma^{2/3})n^{v_1\gamma}F_0(\gamma)\sum_{i_2=0}^{\infty}\cdots\sum_{i_t=0}^{\infty}\prod_{k=2}^{t}\left[G_{i_k}(1)z_0^{i_k}\right]\\&=\frac{1}{z_0^n n}\exp(v_0n^{1/3}\gamma^{2/3})n^{v_1\gamma}\left(\frac{t}{t-1}\right)^{t-1}F_0(\gamma).\end{aligned} \tag{255}$$

We estimate the second sum in (252) by the Euler-Maclaurin formula to obtain for $t=2$

$$\frac{1}{z_0^n n}\exp(v_0n^{1/3}\gamma^{2/3})n^{v_1\gamma}\int_0^1 e^{v_1\gamma H(\vec{x}(2))}\frac{F_0(\gamma x_1)F_0(\gamma-\gamma x_1)}{x_1(1-x_1)}dx_1,$$

and for $t\geq 3$

$$\frac{1}{z_0^n n}\exp(v_0n^{1/3}\gamma^{2/3})n^{v_1\gamma}\int_0^1\ldots\int_0^{1-x_1-\ldots-x_{(t-2)}}e^{v_1\gamma H(\vec{x}(t))}\left[\prod_{j=1}^{t}\frac{F_0(\gamma x_j)}{x_j}\right]dx_{(t-1)}\ldots dx_1. \tag{256}$$

The other sums in (252) can be similarly approximated by the Euler-Maclaurin formula. Using (255) and (256) in (252) we get the limiting equation

$$\begin{aligned}\frac{1}{z_0}(e^\gamma - 1)F_0(\gamma) = &\binom{t}{2}\left(\frac{t}{t-1}\right)^{t-2}\int_0^1 e^{v_1\gamma H(\vec{x}(2))}\frac{F_0(\gamma x_1)F_0(\gamma-\gamma x_1)}{x_1(1-x_1)}dx_1 \\ &+\sum_{i=3}^{t}\binom{t}{i}\left(\frac{t}{t-1}\right)^{t-i} \\ &\times\int_0^1 \dots\int_0^{1-x_1-\dots-x_{(i-2)}} e^{v_1\gamma H(\vec{x}(i))}\left[\prod_{j=1}^{i}\frac{F_0(\gamma x_j)}{x_j}\right]dx_{(i-1)}\dots dx_1.\end{aligned} \tag{257}$$

Here $H(\vec{x}(\cdot))$ is as in (253), we used $(1-\frac{\gamma}{n})^n \sim e^{-\gamma}$, and

$$\sum_{\mathcal{S}_J}\prod_{k=J+1}^{t}\left[G_{i_k}(1-\frac{\gamma}{n})z_0^{i_k}\right] \sim \sum_{i_{J+1}=0}^{\infty}\cdots\sum_{i_t=0}^{\infty}\prod_{k=J+1}^{t}\left[G_{i_k}(1)z_0^{i_k}\right] = \left(\frac{t}{t-1}\right)^{t-J}.$$

Equation (257) 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[28–31]. Note that for the integral to converge, we must have $F_0(0) = 0$, but this follows also from asymptotic matching to the a-scale. Setting

$$F_0(\gamma) = \frac{2t}{(t-1)^2}\gamma F_1(\gamma), \tag{258}$$

(257) simplifies to

$$\begin{aligned}\frac{e^\gamma - 1}{\gamma}F_1(\gamma) = &\int_0^1 F_1(\gamma x_1)F_1(\gamma-\gamma x_1)e^{v_1\gamma H(\vec{x}(2))}dx_1 + \sum_{i=3}^{t}\binom{t}{i}\frac{1}{t}\left(\frac{2}{t-1}\right)^{i-1}\gamma^{i-2} \\ &\times\int_0^1 \dots\int_0^{1-x_1-\dots-x_{(i-2)}} e^{v_1\gamma H(\vec{x}(i))}\left[\prod_{j=1}^{i}F_1(\gamma x_j)\right]dx_{(i-1)}\dots dx_1.\end{aligned} \tag{259}$$

Setting $\gamma = 0$ in (259) we find that

$$F_1(0) = 1.$$

We study the behavior of (259) as $\gamma \to 0$. We expand $F_1(\cdot)$ as

$$F_1(\gamma) = 1 + \alpha_0\gamma\ln\gamma + \alpha_1\gamma + o(\gamma). \tag{260}$$

Using (260) in (259) and expanding for γ small we obtain at $O(\gamma)$

$$\alpha_1 + \frac{1}{2} = (\alpha_0 + v_0)\int_0^1 H(\vec{x}(2))dx_1 + \alpha_1 + \frac{(t-2)}{3(t-1)} = -\frac{1}{2}(\alpha_0 + v_1) + \alpha_1 + \frac{(t-2)}{3(t-1)}, \tag{261}$$

so that

$$\alpha_0 + v_1 = -\frac{(t+1)}{3(t-1)}, \tag{262}$$

and α_1 remains arbitrary. This shows that the solution of (259) is not unique, but becomes unique if α_1 is specified.

To uniquely determine $F_1(\gamma)$ we use asymptotic matching between the a and γ scales. For $\gamma \to 0^+$ we obtain, from (251), (254), (258) and (260),

$$\frac{1}{nz_0^n}\frac{2t}{(t-1)^2}\gamma[1 + \alpha_0\gamma\ln\gamma + \alpha_1\gamma + o(\gamma)](1 + v_1\gamma\ln n)\exp(v_0 n^{1/3}\gamma^{2/3}). \tag{263}$$

The expression in (263) should agree with the a-scale approximation as $a \to -\infty$, which we obtained in (234). Noting that $\gamma = -a/\sqrt{n}$ and comparing (263) to (234) we obtain

$$v_0 = \left(\frac{2t}{t-1}\right)^{1/3} r_0 = -\left(\frac{2t}{t-1}\right)^{1/3} |r_0|, \tag{264}$$

and

$$\alpha_1\gamma + \alpha_0\gamma\ln\gamma + v_1\gamma\ln n = \gamma\left[-\frac{2(t+1)}{9(t-1)}\ln(\gamma\sqrt{n}) + \frac{(t-1)^2}{2t}\kappa + \frac{(t+1)}{3(t-1)}(\gamma-\gamma_E)\right], \tag{265}$$

where γ_E is the Euler constant. From (265) we must have

$$v_1 = -\frac{(t+1)}{9(t-1)}, \quad \alpha_0 = -\frac{2(t+1)}{9(t-1)}, \tag{266}$$

and

$$\alpha_1 = \frac{(t-1)^2}{2t}\kappa + \frac{(t+1)}{3(t-1)}(1-\gamma_E). \tag{267}$$

Note that (266) is consistent with (262). With (267) the solution to (250) is unique and may be computed, e.g., in the form of the series in (260).

To analyze (259) further (with $v_1 = -(t+1)/[9(t-1)]$) we let

$$F_1(\gamma) = \exp(-v_1\gamma\ln\gamma)\,F_2(\gamma), \tag{268}$$

and obtain

$$\begin{aligned}(e^\gamma-1)F_2(\gamma) = &\int_0^\gamma F_2(u_1)F_2(\gamma-u_1)du_1 + \sum_{i=3}^{t}\binom{t}{i}\frac{1}{t}\left(\frac{2}{t-1}\right)^{i-1}\\ &\times\int_0^\gamma\int_0^{\gamma-u_1}\cdots\int_0^{\gamma-u_1-\cdots-u_{(i-2)}}\left[\prod_{j=1}^{i}F_2(u_j)\right]du_{(i-1)}\ldots du_1,\end{aligned} \tag{269}$$

where $u_1+\cdots+u_i=\gamma$ for each i in the sum. Introducing the Laplace transform

$$F(\theta) = \int_0^\infty e^{-\gamma\theta}F_2(\gamma)d\gamma, \tag{270}$$

we obtain from (269)

$$\begin{aligned}F(\theta-1) &= \sum_{i=1}^{t}\binom{t}{i}\frac{1}{t}\left(\frac{2}{t-1}\right)^{i-1}F^i(\theta)\\ &= \frac{t-1}{2t}\left\{\left(1+\frac{2}{t-1}F(\theta)\right)^t - 1\right\}.\end{aligned} \tag{271}$$

We will show later that (271) also follows as a limiting form of the double transform equation (9).

Then (260) gives the behavior of $F_2(\gamma)$ as $\gamma\to 0$

$$F_2(\gamma) = 1 - \frac{(t+1)}{3(t-1)}\gamma\ln\gamma + \alpha_1\gamma + o(\gamma), \quad \gamma\to 0^+,$$

and thus we have

$$F(\theta) = \frac{1}{\theta} + \frac{(t+1)}{3(t-1)}\frac{\ln\theta}{\theta^2} + \left[\alpha_1 + \frac{(t+1)(\gamma_E-1)}{3(t-1)}\right]\frac{1}{\theta^2} + o(\theta^{-2}), \quad \theta\to+\infty. \tag{272}$$

Using (271) we can refine (272) to

$$F(\theta) = \frac{1}{\theta} + \left[\frac{(t+1)}{3(t-1)}\ln\theta + \alpha_*\right]\frac{1}{\theta^2} + \frac{G(\theta)}{\theta^3} + O_R(\theta^{-4}), \tag{273}$$

where

$$\alpha_* = \alpha_1 + \frac{(t+1)(\gamma_E - 1)}{3(t-1)} = \frac{(t-1)^2}{2t}\kappa,$$

and

$$G(\theta) = \left[\frac{(t+1)}{3(t-1)}\ln\theta + \alpha_*\right]^2 - \frac{(t+1)^2}{9(t-1)^2}\ln\theta - \frac{t+1}{3(t-1)}\alpha_* + \frac{(t+1)^2}{18(t-1)^2}. \tag{274}$$

In section 9 we shall show that using some of the higher-order terms in expansion (272) will make it more efficient to numerically compute $F(\theta)$, and hence $F_2(\gamma)$.

Although we cannot solve (269) or (271) explicitly, we can guess the behavior of $F_2(\gamma)$ as $\gamma \to +\infty$, which is needed for asymptotic matching purposes. This corresponds to $F(\theta)$ as $\theta \to -\infty$. Let us assume that $F_2(\gamma)$ in (269) behaves as

$$F_2(\gamma) \sim e^{k_1\gamma\ln\gamma} e^{k_2\gamma}\gamma^{k_3}k_4, \quad \gamma \to \infty. \tag{275}$$

Using (275) in (269) we find that the last term ($i = t$) in the sum dominates. Evaluating this $(t-1)$ fold integral by Laplace's method, we obtain

$$\begin{aligned}
&\frac{e^\gamma}{\gamma} e^{k_1\gamma\ln\gamma} e^{k_2\gamma}\gamma^{k_3}k_4 \\
&\sim \frac{1}{t}\left(\frac{2}{t-1}\right)^{t-1}\gamma^{t-2}\int_0^1 \cdots \int_0^{1-x_1-\cdots-x_{(t-2)}} \left[\prod_{j=1}^{t} e^{k_1\gamma x_j\ln(\gamma x_j)} e^{k_2\gamma x_j}(\gamma x_j)^{k_3}k_4\right] dx_{(t-1)}\ldots dx_1, \\
&\sim \frac{1}{t}\left(\frac{2}{t-1}\right)^{t-1}\gamma^{t-2}(\gamma)^{tk_3}\left(\frac{1}{t}\right)^{tk_3} k_4^t e^{k_2\gamma} e^{k_1\gamma\ln\gamma} e^{k_1\gamma H((\vec{\frac{1}{t}})(t))} \\
&\times \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left[-k_1\gamma\frac{1}{2}\sum_{1\le i,j\le t}\left(x_i - \frac{1}{t}\right)\left(x_j - \frac{1}{t}\right)\left|\frac{\partial^2 H}{\partial x_i\partial x_j}\left(\left(\vec{\frac{1}{t}}\right)(t)\right)\right|\right] dx_{(t-1)}\ldots dx_1 \\
&= \frac{1}{t}\left(\frac{2}{t-1}\right)^{t-1}\gamma^{t-2}(\gamma)^{tk_3}\left(\frac{1}{t}\right)^{tk_3} k_4^t e^{k_2\gamma} e^{k_1\gamma\ln\gamma} e^{-k_1\gamma\ln t}\frac{2^{(t-1)/2}}{t^{t/2}}\left(\frac{\pi}{\gamma|k_1|}\right)^{(t-1)/2}.
\end{aligned} \tag{276}$$

Here $H((\vec{\frac{1}{t}})(t))$ means that H is evaluated at $x_1 = x_2 = \cdots = x_t = t^{-1}$, which is where H in (253) is maximal. In obtaining (276) we evaluated the last integral using

$$\begin{aligned}
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left[-k_1\gamma\frac{1}{2}\sum_{1\le i,j\le t}\left(x_i - \frac{1}{t}\right)\left(x_j - \frac{1}{t}\right)\left|\frac{\partial^2 H}{\partial x_i\partial x_j}\left(\left(\vec{\frac{1}{t}}\right)(t)\right)\right|\right] dx_{(t-1)}\ldots dx_1 \\
&= \frac{2^{(t-1)/2}}{t^{t/2}}\left(\frac{\pi}{\gamma|k_1|}\right)^{(t-1)/2}.
\end{aligned}$$

Comparing the left and right sides of (276) we conclude that

$$k_1 = -\frac{1}{\ln t},\ \ k_3 = -\frac{1}{2},\ \ k_4 = t^{\frac{1}{(t-1)}}\frac{(t-1)}{2\sqrt{2\pi\ln(t)}}. \tag{277}$$

The constant k_2 remain free, though it is uniquely determined once we know α_1 in (260). In section 9 we will compute k_2 numerically. To summarize, in view of (268), we have obtained

$$F_1(\gamma) \sim t^{\frac{1}{(t-1)}}\frac{(t-1)}{2\sqrt{2\pi\ln(t)}}\frac{e^{k_2\gamma}}{\sqrt{\gamma}}\exp\left[\left(\frac{(t+1)}{9(t-1)} - \frac{1}{\ln t}\right)\gamma\ln\gamma\right],\ \ \gamma \to \infty. \tag{278}$$

Using (278) and (258) in (251) gives the behavior of the γ-scale result for $\gamma \to \infty$, and we will use this in section 7 to asymptotically match to another approximation, that is valid for $0 < w < 1$ and $n \to \infty$.

We next analyze $F(\theta)$ for $\theta \to -\infty$, which will correspond to $F_2(\gamma)$ for $\gamma + \infty$. We let $\tau = -\theta \to +\infty$ and find that (271) is consistent with an asymptotic expansion of the form

$$F(\theta) \sim \frac{(t-1)}{2} t^{\frac{1}{t-1}} A^{t^{\tau}} \left[1 + \sum_{L=1}^{\infty} b_L A^{-Lt^{\tau}} \right], \quad \tau = -\theta \to -\infty. \tag{279}$$

Using (279) to asymptotically invert the transform in (270) yields

$$\begin{aligned}
F_2(\gamma) &= \frac{1}{2\pi i} \int_{Br} e^{\gamma\theta} F(\theta) d\theta \\
&\sim \frac{(t-1)}{2} t^{\frac{1}{t-1}} \frac{1}{2\pi i} \int_{Br} e^{\gamma\theta} A^{t^{-\theta}} d\theta \\
&= \frac{(t-1)}{2} t^{\frac{1}{t-1}} \frac{1}{2\pi i} \int_C \frac{1}{\ln t} \exp\left[-\frac{\gamma}{\ln t} \ln u \right] e^{u \ln A} \frac{du}{u} \\
&= \frac{(t-1)}{2} t^{\frac{1}{t-1}} \frac{1}{2\pi i} \int_C \frac{1}{\ln t} u^{-\frac{\gamma}{\ln t}-1} e^{u\ln A} du \\
&= \frac{(t-1)}{2} t^{\frac{1}{t-1}} \frac{1}{\ln t} \frac{(\ln A)^{\frac{\gamma}{\ln t}}}{\Gamma\left(1 + \frac{\gamma}{\ln t}\right)} \\
&\sim \frac{(t-1)}{2} t^{\frac{1}{t-1}} \frac{1}{\sqrt{2\pi\gamma}\sqrt{\ln t}} \exp\left[\frac{\gamma}{\ln t} (1 + \ln\ln A) \right] \exp\left[-\frac{\gamma}{\ln t} \ln\left(\frac{\gamma}{\ln t} \right) \right], \quad \gamma \to \infty.
\end{aligned} \tag{280}$$

Here the contour C goes from $-\infty + i0^-$ to $-\infty + i0^+$, encircling the branch cut along the negative real axis in the u-plane. Also, we used the asymptotic behavior of the Gamma function

$$\Gamma(x) = \frac{\sqrt{2\pi}}{\sqrt{x}} x^x e^{-x}, \quad x \to \infty.$$

By comparing (275) with (277) to (280) we find that k_2 and A are related by

$$k_2 = \frac{\ln(\ln t) + \ln(\ln A) + 1}{\ln t}. \tag{281}$$

In section 9 we shall numerically obtain A for $t = 3$ and then use (281) to obtain the corresponding value of k_2. We note that (279) implies that

$$t^{\theta} \left[\ln[F(\theta)] - \ln\left(\frac{(t-1)}{2} t^{\frac{1}{t-1}} \right) \right] = \ln A - (const.) t^{-\tau} A^{-t^{\tau}} (1 + o(1)), \quad \tau = -\theta \to \infty, \tag{282}$$

so that the left hand side of (282) should converge to the constant $\ln A$ super-exponentially fast as $\theta \to -\infty$.

Now we use the result on the γ-scale to obtain $g(n, p)$ for $p = O(n^{4/3})$. We scale

$$p = \Theta n^{4/3}, \quad \Theta = O(1), \quad \Theta > 0, \tag{283}$$

and use (250), so that

$$w^{-p} = \exp\left[\frac{\gamma p}{n} + O\left(\frac{p}{n^2} \right) \right] \sim \exp(\gamma n^{1/3} \Theta). \tag{284}$$

With (251), (254), (258) and (284), we obtain from (10)

$$\begin{aligned}
\frac{1}{2\pi i} \int_C G_n(w) w^{-p-1} dw \sim \frac{1}{z_0^n n 2\pi i} \int_{Br} \frac{2t}{(t-1)^2} \gamma F_1(\gamma) n^{-\frac{(t+1)\gamma}{9(t-1)}} \\
\times \exp\left[n^{1/3} \left(\gamma\Theta - \left(\frac{2t}{t-1} \right)^{1/3} |r_0| \gamma^{2/3} \right) \right] \frac{d\gamma}{n}.
\end{aligned} \tag{285}$$

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, where

$$\frac{d}{d\gamma}\left(\gamma\Theta - \left(\frac{2t}{t-1}\right)^{1/3}|r_0|\gamma^{2/3}\right) = 0 \Rightarrow \gamma = \gamma_* \equiv \frac{16t}{27(t-1)}\frac{|r_0|^3}{\Theta^3}. \tag{286}$$

Then we use the standard Laplace method to estimate (285) and obtain

$$g(n,p) \sim \frac{1}{z_0^n n^{13/6}}\frac{|r_0|^{9/2}}{\Theta^5}\frac{64t^{5/2}\sqrt{2}}{81(t-1)^{7/2}\sqrt{\pi}}n^{-\frac{(t+1)\gamma_*}{9(t-1)}}F_1(\gamma_*)\exp\left[-\frac{8t}{27(t-1)}n^{1/3}\frac{|r_0|^3}{\Theta^2}\right], \tag{287}$$

which is valid for $\Theta = pn^{-4/3} = O(1)$ and as $n \to \infty$. Here $F_1(\cdot)$ must obtained by numerically solving (259) or (269). Thus we have derived (58).

Since $F_1(0) = 1$, expanding (287) as $\Theta \to \infty$ corresponds to simply neglecting the factors $n^{-(t+1)\gamma_*/[9(t-1)]}$ and $F_1(\gamma_*)$. Thus (287) asymptotically matches to the result valid for $p = O(n^{3/2})$, in view of (242) and the fact that $\Omega = \Theta n^{-1/6}$. We can also deduce the behavior of the right side of (287) as $\Theta \to 0$, which corresponds to $p = o(n^{4/3})$ and $\gamma_* \to \infty$. Using (278) and (286) in (287) we obtain

$$\begin{aligned}
&\frac{1}{z_0^n n^{13/6}}\frac{|r_0|^{9/2}}{\Theta^5}\frac{64t^{5/2}\sqrt{2}}{81(t-1)^{7/2}\pi}\frac{t^{\frac{1}{(t-1)}}(t-1)}{2\sqrt{2\ln t}}\frac{3\sqrt{3(t-1)}}{4\sqrt{t}}\frac{\Theta^{3/2}}{|r_0|^{3/2}}\exp\left[-\frac{8t}{27(t-1)}n^{1/3}\frac{|r_0|^3}{\Theta^2}\right]\\
&\times\exp\left[-\frac{(t+1)}{9(t-1)}\gamma_*\ln n + \left(\frac{(t+1)}{9(t-1)} - \frac{1}{\ln t}\right)\gamma_*\ln\gamma_* + k_2\gamma_*\right]\\
=&\frac{1}{z_0^n n^{13/6}}\frac{|r_0|^3}{\Theta^{7/2}}\frac{1}{\pi\sqrt{\ln t}}t^{\frac{1}{t-1}+2}\frac{8\sqrt{3}}{27(t-1)^2}\\
&\times\exp\left[-\frac{1}{3}\left(\frac{2t}{t-1}\right)^{1/3}|r_0|n^{1/3}\gamma_*^{2/3} - \frac{(t+1)}{9(t-1)}\gamma_*\ln n + \left(\frac{(t+1)}{9(t-1)} - \frac{1}{\ln t}\right)\gamma_*\ln\gamma_* + k_2\gamma_*\right].
\end{aligned} \tag{288}$$

This yields an approximation to $g(n,p)$ valid for $p \ll n^{4/3}$, and is explicit except for the constant $k_2 = k_2(t)$. We will use (288) in sections 7 and 8, for asymptotic matching purposes. In fact, the expression in (288) will be key to seeing where $g(n,p)$ is maximal as a function of n for a fixed large p.

6.2 Analysis of the Functional Equation

We briefly re-analyze the left region by using the functional equation (9). We use the scaling

$$w = 1 - \frac{\gamma}{n},\ \ z = z_0\left(1 - \frac{v_0\gamma^{2/3}}{n^{2/3}} - \frac{v_1\gamma\ln n}{n} + v_1\gamma\frac{\ln\gamma}{n} - \frac{s\gamma}{n}\right), \tag{289}$$

with

$$G(z,w) = \hat{G}(s,\gamma) = \hat{G}\left(\left(1 - \frac{v_0\gamma^{2/3}}{n^{2/3}} - \frac{v_1\gamma\ln n}{n} + v_1\gamma\frac{\ln\gamma}{n} - \frac{z}{z_0}\right)\frac{n}{\gamma},\ (1-w)n\right). \tag{290}$$

Also from (9) we have

$$G(z,w) - G(z,1) = z[G^t(zw,w) - G^t(z,1)]. \tag{291}$$

From (289) we have

$$zw = z_0\left(1 - \frac{v_0\gamma^{2/3}}{n^{2/3}} - \frac{v_1\gamma\ln n}{n} + \frac{v_1\gamma\ln\gamma}{n} - \frac{(s+1)\gamma}{n} + O(n^{-5/3})\right), \tag{292}$$

and (9) becomes

$$\hat{G}(s,\gamma) = 1 + z_0\left(1 - \frac{v_0\gamma^{2/3}}{n^{2/3}} - \frac{v_1\gamma\ln n}{n} + v_1\gamma\frac{\ln\gamma}{n} - \frac{s\gamma}{n}\right) \times\left[\hat{G}\left(s+1+O(n^{-2/3}),\gamma\right)\right]^t. \tag{293}$$

Thus from (8) we obtain

$$\begin{aligned} G_n(w) &= \frac{1}{2\pi i}\int_C \frac{G(z,w)}{z^{n+1}}dz \\ &\sim \frac{1}{2\pi i}\int_C \hat{G}(s,\gamma)\frac{1}{z_0^n}e^{v_0 n^{1/3}\gamma^{2/3}}e^{v_1\gamma\ln n}e^{-v_1\gamma\ln\gamma}e^{s\gamma}\frac{\gamma ds}{n} \\ &= \frac{1}{z_0^n n}e^{v_0 n^{1/3}\gamma^{2/3}}e^{v_1\gamma\ln n}\frac{1}{2\pi i}\int_C \hat{G}(s,\gamma)e^{-v_1\gamma\ln\gamma}e^{s\gamma}\gamma ds. \end{aligned} \tag{294}$$

Comparing (294) with (251), we have

$$\frac{1}{2\pi i}\int_C \hat{G}(s,\gamma)e^{-v_1\gamma\ln\gamma}e^{s\gamma}\gamma ds \sim F(\gamma;n). \tag{295}$$

Setting

$$G(z,w) = \hat{G}(s,\gamma) \sim \frac{t}{t-1}\left[1 + B(s,\gamma)\right]. \tag{296}$$

Using (296) in (293), for $n\to\infty$ we obtain

$$\begin{aligned} \frac{t}{t-1}\left[1+B(s,\gamma)\right] &= 1 + \frac{(t-1)^{t-1}}{t^t}\left(\frac{t}{t-1}\right)^t\left[1+B(s+1,\gamma)\right]^t \\ &= 1 + \frac{1}{t-1}\left[1+B(s+1,\gamma)\right]^t. \end{aligned} \tag{297}$$

For $s\to\infty$, the solution to (297) has the form

$$\frac{1}{s} + \frac{\ln s}{s^2} + \frac{f(\gamma)}{s^2} + o\left(\frac{1}{s^2}\right),$$

where $f(\cdot)$ is an arbitrary function. By asymptotically matching to the a-scale we find that $f(\gamma)$ is a constant, and then $B(s,\gamma)$ will be independent of γ, so we write $B(s,\gamma) = B(s)$. Thus we have

$$B(s) = \frac{1}{t}\left\{\left[1+B(s+1)\right]^t - 1\right\}. \tag{298}$$

Setting $s = \theta - 1$, (298) becomes

$$B(\theta-1) = \frac{1}{t}\left\{\left[1+B(\theta)\right]^t - 1\right\}. \tag{299}$$

Rescaling $B(\theta)$ as

$$B(\theta) = \frac{2}{t-1}\tilde{B}(\theta),$$

(299) then leads to

$$\tilde{B}(\theta-1) = \sum_{i=1}^{t}\binom{t}{i}\frac{1}{t}\left(\frac{2}{t-1}\right)^{i-1}\tilde{B}^i(\theta). \tag{300}$$

Using (296) in (295), and recalling (258) and (268), we obtain, for $n\to\infty$,

$$\frac{1}{2\pi i}\int_C \frac{t}{t-1}\left[1+B(s)\right]e^{-v_1\gamma\ln\gamma}e^{s\gamma}\gamma ds \sim \frac{2t}{(t-1)^2}\gamma e^{-v_1\gamma\ln\gamma}F_2(\gamma). \tag{301}$$

Here again we interpreted

$$\frac{1}{2\pi i}\int_C e^{s\gamma}\gamma ds = \gamma\delta(\gamma) = 0, \tag{302}$$

as a distribution. By comparing (271) to (300) (or the inverse of (270) to (301)) we see that $\tilde{B}(\theta) = F(\theta)$. We have thus shown that (7) and (9) lead to equivalent limiting equations, that apply apply on the γ-scale.

7. FAR LEFT REGION

We consider (7) for $0 < w < 1$ and $n \to \infty$, and (10) for $n \to \infty$ with $p - n\log_t n = O(n)$. Note that $p_{min}(n)$ in (29) is contained in the scale $p - n\log_t n = O(n)$.

First we assume that $G_n(w)$ has an expansion of the form

$$G_n(w) = e^{-n\ln(n)f(w)}e^{ng(w)}n^{h(w)}q(w)[1 + o(1)], \tag{303}$$

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

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

with (303) the left side of (7) becomes

$$e^{-n\ln(n)f(w)}e^{ng(w)}n^{-f(w)+h(w)}q(w)e^{g(w)-f(w)}[1 + o(1)]. \tag{304}$$

To evaluate the right side we treat the sum as an implicit Laplace-type integral. The major contribution comes from the central region where each i_k is $O(n)$. Then, using (303), the sum in (7) becomes asymptotic to

$$e^{n\ln w}\sum_{i_1+\cdots+i_t=n} e^{-f(w)[i_1\ln i_1+\cdots+i_t\ln i_t]}e^{ng(w)}[i_1 i_2\cdots i_t]^{h(w)}q^t(w). \tag{305}$$

We use the Euler-Maclaurin formula to approximate the above sum and obtain

$$e^{n\ln w}n^{2h(w)+1}q^2(w)e^{ng(w)}e^{-n\ln(n)f(w)}\int_0^1 [x_1(1-x_1)]^{h(w)}\, e^{-nf(w)H(\vec{x}(2))}dx_1, \quad t = 2,$$

and

$$\begin{aligned} &e^{n\ln w}n^{th(w)+t-1}q^t(w)e^{ng(w)}e^{-n\ln(n)f(w)} \\ &\times\int_0^1\cdots\int_0^{1-x_1-\cdots-x_{t-2}}(x_1\cdots x_t)^{h(w)}e^{-nf(w)H(\vec{x}(t))}dx_{t-1}\cdots dx_1, \quad t\ge 3. \end{aligned} \tag{306}$$

Here $H(\vec{x}(t)) = \sum_{j=1}^t x_j\log x_j$ is as in (253). For now we assume that $f(w) > 0$ and estimate (306) by Laplace's method for $n \to \infty$, with the major contribution coming from $\vec{x} = (x_1,\cdots,x_t) = (\frac{1}{t},\cdots,\frac{1}{t}) \equiv \frac{\vec{1}}{t}$.

Using

$$H(\frac{\vec{1}}{t}(t)) = -\ln t,\ \ H'(\frac{\vec{1}}{t}(t)) = 0,\ \ H''(\frac{\vec{1}}{t}(t)) = 2t, \tag{307}$$

(306) becomes asymptotic to

$$e^{ng(w)}e^{-n\ln(n)f(w)}n^{th(w)+t-1}q^t(w)\frac{2^{(t-1)/2}}{t^{t/2}}\left(\frac{\pi}{nf(w)}\right)^{\frac{(t-1)}{2}}\left(\frac{1}{t}\right)^{th(w)}\exp\left[n(f(w)\ln t + \ln(w))\right]. \tag{308}$$

Comparing (303) to (308) we conclude that

$$f(w) = -\frac{\ln w}{\ln t} = -\log_t w > 0, \tag{309}$$

$$h(w) = \frac{\log_t w}{(t-1)} - \frac{1}{2}, \tag{310}$$

and

$$q(w) = w^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln t}} \sqrt{-\log_t(w)}\, \frac{1}{\sqrt{2\pi}} e^{\frac{g(w)}{t-1}}. \tag{311}$$

But, the function $g(w)$ remains undetermined. Summarizing the result so far, with the *ansatz* (303) we obtained

$$G_n(w) \sim w^{n\log_t n} e^{\left(n+\frac{1}{t-1}\right)g(w)} n^{\frac{\log_t w}{t-1}} w^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln t}} \frac{\sqrt{-\log_t(w)}}{\sqrt{2n\pi}}[1+o(1)]. \tag{312}$$

The numerical studies in section 9 show that (312) is approximately correct for $w > 0$, however there are oscillations that become numerically significant when w becomes small. Thus we re-consider (7) with the following more general *ansatz*

$$G_n(w) = e^{-n\ln(n)f(w)} e^{nB(w,n)} n^{h(w)} Q(w,n)[1+o(1)]. \tag{313}$$

We allow $B(w,n)$ and $Q(w,n)$ to depend weakly on n, in such a way that $B(w,n+1) \sim B(w,n)$ and $Q(w,n+1) \sim Q(w,n)$ for $n \to \infty$. Repeating the calculation that we did with the previous *ansatz* (303), now we find that as $n \to \infty$ the left side of (7) becomes

$$e^{-n\ln(n)f(w)} e^{(n+1)B(w,n+1)} n^{h(w)-f(w)} Q(w,n+1) e^{-f(w)}. \tag{314}$$

The right side of (7) is asymptotic to

$$\begin{aligned} e^{n\ln(w)} \sum_{i_1+\cdots+i_t=n} & [i_1 i_2 \cdots i_t]^{h(w)} [Q(w,i_1)\times\cdots\times Q(w,i_t)] \exp\left[i_1 B(w,i_1)+\cdots+i_t B(w,i_t)\right] \\ & \times \exp\left\{-f(w)[i_1\ln i_1+\cdots+i_t\ln i_t]\right\}. \end{aligned} \tag{315}$$

We shall again use Laplace's method to evaluate (315). We first use a multi-variable Taylor expansion to expand about $\vec{i} = (i_1,\cdots,i_t) = (\frac{n}{t},\cdots,\frac{n}{t})$ and obtain

$$\begin{aligned} \sum_{k=1}^{t} i_k B(w,i_k) &= nB\left(w,\frac{n}{t}\right) \\ &+ \left[\frac{n}{t}B_{2,2}\left(w,\frac{n}{t}\right) + 2B_{2,1}\left(w,\frac{n}{t}\right)\right] \sum_{1\le l,m\le t-1} \left(i_l-\frac{n}{t}\right)\left(i_m-\frac{n}{t}\right)+\cdots. \end{aligned} \tag{316}$$

Here $B_{2,1}(\cdot,\cdot)$ and $B_{2,2}(\cdot,\cdot)$ means that the first and second derivative of the second variable respectively.

Using (316) in (315) leads to, for $n \to \infty$,

$$\begin{aligned} & e^{n\ln w} e^{-n\ln(n)f(w)} e^{n(\ln t)f(w)} n^{t\,h(w)+t-1} \left(\frac{1}{t}\right)^{t\,h(w)} e^{nB\left(w,\frac{n}{t}\right)} Q^t\left(w,\frac{n}{t}\right) \\ & \times \frac{1}{\sqrt{2}}\left(\frac{2}{t}\right)^{t/2} \left[\frac{\pi}{n\left(f(w)-\frac{2n}{t}B_{2,1}\left(w,\frac{n}{t}\right)-\frac{n^2}{t^2}B_{2,2}\left(w,\frac{n}{t}\right)\right)}\right]^{(t-1)/2}. \end{aligned} \tag{317}$$

Comparing (314) to (317) we regain (309) and (310). Since we assumed that $B(w,n)$ varies slowly with n, we use

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

and from (314) and (317) we also conclude that

$$B(w,n) = B\left(w,\frac{n}{t}\right). \tag{318}$$

The most general solution of (318) is the Fourier series

$$B(w,n) = g(w) + \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} g_k(w) e^{2\pi i(\log_t n)k} \equiv g(w) + B_0^*(w,n), \tag{319}$$

where the sum represents an arbitrary, zero-mean periodic function of $\log_t(n)$, with period one. We see that (319) is consistent with our assumption of the slow variation of $B(w,n)$. However, we cannot determine explicitly the Fourier coefficient $g_k(w)$ using only the recurrence (7). This would seem to require an exact solution to (7), which is not feasible. However, we will be able to obtain some analytic information about the Fourier coefficients $g(w)$ and $g_k(w)$, and will study $g(w)$ numerically in section 9.

With (319) we write

$$\frac{n}{t} B_{2,1}\left(w, \frac{n}{t}\right) = nB_n(w,n) = B_1^*(w,n),$$
$$\frac{n^2}{t^2} B_{2,2}\left(w, \frac{n}{t}\right) = n^2 B_{nn}(w,n) = B_2^*(w,n)$$

where

$$\begin{aligned} B_1^*(w,n) &= \frac{2\pi i}{\ln(t)} \sum_{k=-\infty}^{\infty} k\, g_k(w) e^{2\pi i k \log_t(n)}, \\ B_2^*(w,n) &= \frac{2\pi i}{\ln(t)} \sum_{k=-\infty}^{\infty} \left(\frac{2\pi i}{\ln(t)} k^2 - k\right) g_k(w) e^{2\pi i k \log_t(n)}\,. \end{aligned} \tag{320}$$

Comparing the factors in (314) and (317) that are $O(1)$ in n, we obtain

$$Q^t\left(w, \frac{n}{t}\right) = Q(w,n) w^{\frac{t}{t-1}+\frac{1}{\ln t}} e^{g(w)} e^{B_0^*(w,n)+B_1^*(w,n)} \left[\frac{f(w) - 2B_1^*(w,n) - B_2^*(w,n)}{2\pi}\right]^{\frac{(t-1)}{2}}. \tag{321}$$

A particular solution to (321) is

$$Q_p(w,n) = w^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln t}} \frac{1}{\sqrt{2\pi}} e^{\frac{1}{t-1}[g(w)+B_0^*(w,n)+B_1^*(w,n)]} \sqrt{f(w) - 2B_1^*(w,n) - B_2^*(w,n)}. \tag{322}$$

Note that if $B_j^*(w,n) = 0$ and $Q(w,n) = q(w)$, (322) agrees with (311).

Setting

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

we see that $\tilde{Q}$ satisfies

$$\tilde{Q}^t\left(w, \frac{n}{t}\right) = \tilde{Q}(w,n). \tag{323}$$

Setting

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

we find that

$$t\tilde{G}\left(w, \frac{n}{t}\right) = \tilde{G}(w,n). \tag{324}$$

The most general solution to (326) is

$$\tilde{G}(w,n) = n \times [\text{periodic function of } \log_t n, \text{ of period } 1]\,. \tag{325}$$

But then $\tilde{G}(w,n)$ can be incorporated into the factor $\exp[nB(w,n)]$ in (313). Thus using (309), (310), (319) and (322) in (313), we have established (50) and (51).

Next we examine the asymptotic matching between the results for $0 < w < 1$ and $w = 1 - O(n^{-1})$. For $\gamma \to \infty$ we expand (251) to obtain

$$\begin{aligned}&\frac{1}{z_0^n n}\exp\left[\left(\frac{2t}{t-1}\right)^{1/3} r_0 n^{1/3}\gamma^{2/3} - \frac{(t+1)}{9(t-1)}\gamma\ln n\right]\frac{t}{(t-1)}\frac{t^{1/(t-1)}\sqrt{\gamma}}{\sqrt{2\pi\ln t}}e^{k_2\gamma}\\&\times\exp\left[\left(\frac{(t+1)}{9(t-1)} - \frac{1}{\ln t}\right)\gamma\ln\gamma\right].\end{aligned} \tag{326}$$

Here we also used (254), (258) and (278). By asymptotic matching, (326) should be consistent with the expansion of (313) as $w \to 1$. Since (326) has 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 (326) must match to the right side of (312), which ignored the oscillations, as the latter is expanded for $w \to 1$.

Noting that

$$w^{n\log_t n} = \left(1 - \frac{\gamma}{n}\right)^{n\log_t n} = \exp\left[-\frac{\gamma}{\ln t}\ln n + O\left(\frac{\ln n}{n}\right)\right],$$

and, for $1 - w$ sufficiently small,

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

(312) becomes

$$\frac{1}{\sqrt{2\pi\ln t}}\sqrt{\frac{1-w}{n}}\exp\left(-\frac{\gamma}{\ln t}\ln n\right)\exp\left[\left(n + \frac{1}{t-1}\right)g(1 - \frac{\gamma}{n})\right]. \tag{327}$$

Thus the matching is possible provided that as $w \uparrow 1$, $g(w)$ has the expansion

$$\begin{aligned}g(w) = {}&\ln\left(\frac{1}{z_0}\right) + \left(\frac{2t}{t-1}\right)^{1/3} r_0\, n^{1/3}(1-w)^{2/3} + \left(\frac{1}{\ln t} - \frac{(t+1)}{9(t-1)}\right)(w-1)\ln(1-w)\\&+ k_2(1-w) + o(1-w), \quad w \uparrow 1.\end{aligned} \tag{328}$$

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

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

$$G_n(w) = C_n w^{p_{min}(n)} + \cdots + t^{n-1}w^{\binom{n}{2}}, \tag{329}$$

where

$$\begin{aligned}p_{min}(n) = \sum_{J=1}^{n}\lfloor\log_t((t-1)J)\rfloor &= \left(n + \frac{1}{t-1}\right)\lfloor\log_t((t-1)n+1)\rfloor - \frac{t\left(t^{\lfloor\log_t((t-1)n+1)\rfloor} - 1\right)}{(t-1)^2}\\&= \left(n + \frac{1}{t-1}\right)\lfloor\log_t((t-1)n)\rfloor - \frac{t\left(t^{\lfloor\log_t((t-1)n)\rfloor} - 1\right)}{(t-1)^2}.\end{aligned} \tag{330}$$

Using (329) in (7) we can obtain a recurrence relation for the C_n. For $w \to 0$ we find that the largest terms in the multi-sum in the right side of (7) are these very close to the centroid of the hyper-triangle $i_1 + i_2 + \cdots + i_t = n$, where each i_k is approximately n/t. But, we can give a simple combinatorial argument that determines C_n explicitly.

We recall that C_n is the number of t-ary trees that have minimum path length possible, for this particular number of nodes n. Suppose first that n is of the form

$$n = \frac{t^{h+1}-1}{t-1} = 1 + t + t^2 + \cdots + t^h,$$

for some integer h. Then it is possible for the tree to be completely balanced, and this clearly leads to the shortest path possible. Furthermore, there is precisely one such balanced tree, hence

$$C_n = 1, \quad n = \frac{t^{h+1} - 1}{t - 1}. \tag{331}$$

Note that for $t = 3$ this corresponds to the subsequence $n \in \{1, 4, 13, 40, \cdots\}$, and h is precisely the height of the balanced tree.

Now consider values of n in the range

$$\frac{t^{h+1} - 1}{t - 1} < n < \frac{t^{h+2} - 1}{t - 1}.$$

It is no longer possible to configure the nodes into a balanced tree. Letting

$$h = h(n) = \lfloor \log_t (n(t-1) + 1) \rfloor,$$

the shortest path length is attained by putting $(t^{h+1} - 1)/(t - 1)$ into a balanced tree of height h, and the remaining nodes at depth $h + 1$. There a total of t^{h+1} possible positions for these additional nodes, and every different choice of positions leads to a different tree. Hence we have

$$C_n = \binom{t^{h+1}}{n - \frac{t^{h+1}-1}{t-1}}, \quad h = \lfloor \log_t (n(t-1) + 1) \rfloor, \tag{332}$$

and this even contains (331), in the special case of a balanced tree, where the lower argument of the binomial coefficient becomes zeros.

We can approximate C_n in (332) for $n \to \infty$. by Stirling's formula, and we denote this approximation by C_n^{asy}. For now we observe that only that C_n grows roughly exponentially with n, with some oscillations due to appearance of the floor function in the definition of h. More precisely, we will have

$$\ln\left(C_n^{asy}\right) = n\, f_0(\Omega_*) + (\ln n) f_1(\Omega_*) + f_2(\Omega_*) + o(1), \tag{333}$$

where $\Omega_* = \{\log_t (n(t-1) + 1)\}$ with $\{\cdot\}$ denoting the fractional part. The functions $f_k(\Omega_*)$ can be explicitly computed from (332) and Stirling's formula. Thus for n fixed we have shown that

$$G_n(w) \sim C_n w^{p_{min}(n)}, \quad w \to 0, \tag{334}$$

where C_n can be obtained from (332).

We next examine the asymptotic matching of (334) as $n \to \infty$, with (317) for $w \to 0$. Using the Fourier series

$$\{x\} = \frac{1}{2} - \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \frac{e^{2\pi i k x}}{2\pi i k} = \frac{1}{2} - \sum_{k=1}^{\infty} \frac{\sin(2\pi k x)}{\pi k}$$

and

$$t^{-\{x\}} = \frac{t-1}{t} \sum_{-\infty}^{\infty} \frac{1}{\ln t + 2k\pi i} e^{2\pi i k x},$$

we can represent $p_{min}(n)$ as

$$\begin{aligned} p_{min}(n) = n\log_t n + n\left[-\frac{1}{2} - \frac{1}{\ln t} + \frac{\ln(t-1)}{\ln t}\right] \\ + n\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{\ln t}{2\pi ik(\ln t + 2\pi ik)} e^{2\pi ik\log_t[(t-1)n]} \\ + \frac{1}{t-1}\log_t n + \frac{t+1}{2(t-1)^2} + \frac{\ln(t-1)}{(t-1)\ln t} \\ + \frac{1}{t-1}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{e^{2\pi ik\log_t[(t-1)n]}}{2\pi ik}. \end{aligned} \tag{335}$$

This form gives immediately the asymptotic behavior of $p_{min}(n)$ as $n \to \infty$. If (317) were to contain (334) as a special case (at least for $n \to \infty$), it would need to behave as $C_n^{asy}\exp(p_{min}(n)\ln w)$ for $w \to 0$, where C_n^{asy} denotes the asymptotic behavior of C_n as $n \to \infty$. By comparing (335) to (50) we find that the largest factors, i.e., $\exp(n\log_t n\ln w) = w^{n\log_t n}$, agree automatically, and the $\exp(O(n))$ factors match if

$$\begin{aligned} g(w) \sim \left(-\frac{1}{2} - \frac{1}{\ln t} + \frac{\ln(t-1)}{\ln t}\right)\ln w, \quad w \to 0, \\ g_k(w) \sim \frac{\ln t}{2\pi ik(\ln t + 2\pi ik)} e^{2\pi ik\log_t(t-1)}\ln w, \quad w \to 0. \end{aligned} \tag{336}$$

We comment that the second, $O(1)$, terms in the expansions of the $g_k(w)$ as $w \to 0$ could be obtained by taking into account the effects of C_n^{asy}, and using the expression in (333). With (336) we then have

$$e^{\frac{g(w)}{t-1}} w^{\frac{t}{(t-1)^2} + \frac{1}{(t-1)\ln t}} \sim w^{\frac{t+1}{2(t-1)^2} + \frac{\ln(t-1)}{(t-1)\ln t}}, \quad w \to 0,$$

and the term $\frac{\log_t n}{(t-1)}$ in (335) agrees with the factor $n^{\frac{\log_t w}{t-1}} = w^{\frac{\log_t n}{t-1}}$ in (50). Moreover,

$$\begin{aligned} \exp\left[\frac{1}{t-1}\Big(B_0^*(w,n) + B_1^*(w,n)\Big)\right] &= \exp\left\{\frac{1}{t-1}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty}\left(1 + \frac{2\pi ik}{\ln t}\right) g_k(w) e^{2\pi ik\log_t(n)}\right\} \\ &\sim \exp\left\{\frac{\ln w}{t-1}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{e^{2\pi ik\log_t[(t-1)n]}}{2\pi ik}\right\}, \quad w \to 0 \end{aligned}$$

which corresponds to the last term in (335). However, (50) still has the factor

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

But to evaluate the above as $w \to 0$ would require the second terms in the expansion of the $g_k(w)$ as $w \to 0$. Note that $2B_1^*(w,n) + B_2^*(w,n) \sim -\log_t w$, in view of (336).

We thus conclude that the expansion for $w < 1$, when expanded for $w \to 0$, may not be able to match to (333) as $n \to \infty$. This suggests that yet another scale may be needed to completely understand the asymptotic behavior of $G_n(w)$. This scale would have $n \to \infty$ and $w \to 0$ simultaneously. However, it is not important for studying the asymptotic behavior for $\sum_n g(n,p)$, as this requires $w \to 1$, as we show in section 8.

Finally, we use the form (312) to get an approximation to $g(n, p)$. We set

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

so that

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

and we obtain, for α fixed,

$$g(n, p) \approx \frac{1}{2\pi i}\int_C \frac{1}{w} e^{\frac{g(w)}{t-1}} n^{\frac{\log_t w}{t-1}} \sqrt{\frac{-\log_t w}{2\pi n}}\, w^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln t}} \exp\left[n(g(w) - \alpha \ln w)\right] dw. \tag{337}$$

Note that this neglects the Fourier coefficients $g_k(w)$ in (319), which tend to be numerically quite small. For $n \to \infty$ we use the saddle point method to estimate (337). There is a saddle point in the range $w = O(1)$ if we are able to solve the equation $g'(w) = \alpha/w$. Our numerical studies in section 9 suggest that this is indeed possible. The standard saddle point approximation then yields

$$\begin{aligned} g(n, p) \approx\ & \frac{n^{\frac{\log_t(w_*)}{t-1}}}{2\pi n} \frac{w_*^{\frac{t}{(t-1)^2}+\frac{1}{(t-1)\ln(t)}}}{\sqrt{\alpha + w_*^2 g''(w_*)}} e^{\frac{g(w_*)}{t-1}} \sqrt{-\log_t w_*} \\ & \times \exp\left[ng(w_*) - n\alpha \ln(w_*)\right]. \end{aligned} \tag{338}$$

Here $w_* = w_*(\alpha)$ must be obtained by numerically solving $w_* g'(w_*) = \alpha$. In section 8 we obtain w_* analytically for $\alpha \to \infty$. Thus we have established (59).

8. THE MATCHING BETWEEN THE LEFT AND FAR LEFT REGIONS

In section 7 we showed that the expansions for $w = 1 - O(n^{-1})$ ($p = O(n^{4/3})$) and $0 < w < 1$ ($p = n\log_t n + O(n)$) match in some intermediate limit, where $\gamma = n(1 - w) \to \infty$ and $w \to 1$. In this section we will examine in more detail this matching region, as this is the key to understanding the distribution of the number of nodes in trees with a given large total path length p (cf. (28)). For a fixed p, we move from right to left as we increase n. The result in (287) shows that for a fixed $\Theta = pn^{-4/3}$, $g(n, p)$ still grows with n, due to the dominant exponential factor of z_0^{-n} (we recall that z_0 is as in (22) and $0 < z_0 < 1$). However, for $p = n\log_t n + O(n)$, (59) shows that for $p/n - \log_t n = O(1)$, $g(n, p)$ respectively grows (decays) with n as $g(w_*) - \alpha \log w_* > 0$ (< 0). But, numerical studies (cf. section 9) show that the quantity $g(w_*) - \alpha \ln(w_*)$ is negative unless $w_* \to 1$. Thus, to find the limiting distribution of the number of nodes we need the maximum of $g(n, p)$ over n, and this occurs exactly in the matching region, where this function will be shown to reach a Gaussian peak.

As $(p - n\log_t n)/n = \alpha \to \infty$ we have $w_* \to 1$ and the non-constant Fourier coefficients in (50) (i.e., $g_k(w)$, $k \neq 0$) vanish as $w \to 1$. Thus we use (53) and (59). For $\alpha \to \infty$, we can solve the equation

$$wg'(w) = \alpha, \tag{339}$$

for the saddle point $w = w_*(\alpha)$ asymptotically, using the following relations, which follow from the asymptotic formula for $g(w)$ as $w \to 1$,

$$\begin{aligned} g(w) = \ln\left(\frac{1}{z_0}\right) - \left(\frac{2t}{t-1}\right)^{1/3} |r_0|(1-w)^{2/3} + \left[\frac{1}{\ln t} - \frac{(t+1)}{9(t-1)}\right](w-1)\ln(1-w) \\ - k_2(w-1) + o(w-1), \end{aligned} \tag{340}$$

$$g'(w) = \frac{2}{3}\left(\frac{2t}{t-1}\right)^{1/3}|r_0|(1-w)^{-1/3} + \left[\frac{1}{\ln t} - \frac{(t+1)}{9(t-1)}\right]\ln(1-w) + \frac{1}{\ln t} - \frac{(t+1)}{9(t-1)} - k_2 + o(1), \tag{341}$$

$$g''(w) \sim \frac{2}{9}\left(\frac{2t}{t-1}\right)^{1/3}|r_0|(1-w)^{-4/3}, \quad w \to 1^-. \tag{342}$$

In the calculations that follow it is useful to define $A_0,\ A_1,\ A_2$ by

$$A_0 = \frac{2}{3}\left(\frac{2t}{t-1}\right)^{1/3}|r_0|,\ A_1 = \frac{1}{\ln t} - \frac{(t+1)}{9(t-1)},\ A_2 = A_1 - k_2 = \frac{1}{\ln t} - \frac{(t+1)}{9(t-1)} - k_2, \tag{343}$$

and Δ by

$$w_* = 1 - \Delta. \tag{344}$$

With (343) and (344) we rewrite (339) as

$$\alpha = \left[A_0\Delta^{-1/3} + A_1\ln\Delta + A_2 + o(1)\right](1-\Delta),\ \ \Delta \to 0, \tag{345}$$

which can be rearranged to obtain

$$\begin{aligned}\frac{\Delta}{A_0^3} &= [\alpha - A_1\ln\Delta - A_2 + \alpha\Delta + o(1)]^{-3} \\ &= \alpha^{-3}\left[1 + \frac{3}{\alpha}(A_1\ln\Delta + A_2) + o(\alpha^{-1})\right].\end{aligned} \tag{346}$$

Thus we have

$$\Delta \sim \frac{{A_0}^3}{\alpha^3},\ \ \alpha \to \infty, \tag{347}$$

and (347) can be refined to the expansion

$$\Delta = \left(\frac{A_0}{\alpha}\right)^3\left[1 + \frac{\delta_1}{\alpha} + \frac{\delta_2}{\alpha^2} + \cdots\right],\ \ \alpha \to \infty, \tag{348}$$

where the δ_j may depend weakly on α, as $\ln(\alpha)$, and δ_1 is given by

$$\delta_1 = 3A_2 - 9A_1\ln\left(\frac{\alpha}{A_0}\right). \tag{349}$$

We next calculate $g(w_*) - \alpha\ln(w_*)$ as $w_* \to 1$ and $\alpha \to \infty$. Note that this is the exponential growth rate in n of the approximation in (59). With (344), (348) and (349) we get

$$\begin{aligned}g(w_*) - \alpha\ln(w_*) &= g(1-\Delta) - \alpha\ln(1-\Delta) \\ &= \ln\left(\frac{1}{z_0}\right) + \frac{A_1}{2}\Delta\ln\Delta + \left(\frac{3}{2}A_2 + k_2\right)\Delta - \frac{\alpha\Delta}{2} + O(\alpha\Delta^2) \\ &= \ln\left(\frac{1}{z_0}\right) - \frac{A_0^3}{2\alpha^2} + \frac{A_0^3}{\alpha^3}\left[3A_1\ln\left(\frac{\alpha}{A_0}\right) + k_2\right] + O_R(\alpha^{-4}).\end{aligned} \tag{350}$$

Here we used (345) in the alternate form $\alpha\Delta \sim A_0\Delta^{2/3} + A_1\Delta\ln\Delta + A_2\Delta$.

To locate the maximum of $ng(w_*) - n\alpha\ln(w_*)$ over n, we need to solve asymptotically

$$\frac{\partial}{\partial n}[ng(w_*) - n\alpha\ln(w_*)] = g(w_*) + \frac{\ln n + 1}{\ln t}\ln(w_*) = 0. \tag{351}$$

We note that here w_* depends on n through α, and from (339) we have $g'(w_*) = \alpha/w_*$. Also, from $n\alpha = p - n\log_t n$ we get $\partial(n\alpha)/\partial n = -\log_t n - 1/\ln t$. We write the solution to (351) as $n = \tilde{n} = \tilde{n}(p)$ and set

$$F(n;p) = ng(w_*) - n\alpha \ln w_*, \tag{352}$$

so that

$$F_n(n;p) = g(w_*) + \frac{\ln n + 1}{\ln t}\ln(w_*), \tag{353}$$

and

$$\begin{aligned} F_{nn}(n;p) &= g'(w_*)\frac{\partial w_*}{\partial n} + \frac{\ln w_*}{n\ln t} + \frac{(\ln n + 1)}{w_* \ln t}\frac{\partial w_*}{\partial n} \\ &= \left(\alpha + \frac{\ln n + 1}{\ln t}\right)\frac{1}{w_*}\frac{\partial w_*}{\partial n} + \frac{\ln w_*}{n\ln t}. \end{aligned} \tag{354}$$

We define

$$\Psi_0(p) = F(\tilde{n}(p);p), \tag{355}$$

and

$$\mathcal{V}_0(p) = -1/F_{nn}(\tilde{n}(p);p). \tag{356}$$

We will show that $\Psi_0(p)$ provides an asymptotic approximation to the exponential growth rate of the total number of trees of path length p, while $\tilde{n}(p)$ and $\mathcal{V}_0(p)$ yield asymptotically the mean and variance of the Gaussian distribution of the number of nodes in such trees.

For $w_* \to 1$ we use (340) and (348) in (351), and this leads to

$$\begin{aligned} &\ln\left(\frac{1}{z_0}\right) - \frac{\ln n + 1}{\ln t}\frac{A_0^3}{\alpha^3}\left[1 + \frac{3}{\alpha}\left(A_2 - 3A_1\ln\left(\frac{A_0}{\alpha}\right)\right)\right] \\ &- \frac{3}{2}\frac{A_0^3}{\alpha^2}\left[1 + \frac{2}{\alpha}\left(A_2 - 3A_1\ln\left(\frac{\alpha}{A_0}\right)\right)\right] + k_2\frac{A_0^3}{\alpha^3} - 3A_1\frac{A_0^3}{\alpha^3}\ln\left(\frac{A_0}{\alpha}\right) + o(\alpha^{-3}) = 0. \end{aligned}$$

To leading order, $\ln(1/z_0)$ must be balanced by the term proportional to $\alpha^{-3}\ln n$, and we thus have

$$\alpha \sim \frac{A_0}{(\ln t)^{1/3}}\frac{(\ln n)^{1/3}}{(-\ln z_0)^{1/3}}, \quad n = \tilde{n}(p). \tag{357}$$

This shows that for a fixed p, the maximum of $g(n,p)$ occurs in the range $p = n\log_t n + O[n(\ln n)^{1/3}]$. We compare this to $p_{min} = n\log_t n + O(n)$ (cf. (29)). From (347) and (357) we conclude that

$$\tilde{\Delta} \sim \frac{-\ln t \ln z_0}{\ln n} \sim 1 - \tilde{w} \quad n = \tilde{n}(p) \to \infty. \tag{358}$$

where $\tilde{w}$, $\tilde{\Delta}$ and $\tilde{\alpha}$ are defined by

$$\begin{aligned} &\tilde{w} = \tilde{w}(p) = w_*(\tilde{n}(p),p), \quad \tilde{\Delta} = \tilde{\Delta}(p) = 1 - w_*(\tilde{n}(p),p), \\ &\tilde{\alpha} = \frac{p - \tilde{n}(p)\log_t[\tilde{n}(p)]}{\tilde{n}(p)}, \end{aligned} \tag{359}$$

and $w_* = w_*(n,p)$ satisfies (339). Using (59) we obtain

$$\begin{aligned} \sum_n g(n,p) \sim\ & \frac{1}{2\pi\tilde{n}}\frac{(\tilde{w})^{\frac{t}{(t-1)^2} + \frac{1}{(t-1)\ln t}}}{\sqrt{\tilde{\alpha} + \tilde{w}^2 g''(\tilde{w})}} n^{\frac{1}{t-1}\log_t \tilde{w}} e^{\frac{1}{t-1}g(\tilde{w})}\sqrt{-\log_t \tilde{w}} \\ &\times \sum_{n=-\infty}^{\infty}\exp\left[ng(\tilde{w}) - n\alpha\ln\tilde{w} + \frac{1}{2}F_{nn}(\tilde{n}(p);p)[n - \tilde{n}(p)]^2\right], \end{aligned} \tag{360}$$

where we expanded $g(n, p)$ about $n = \tilde{n}(p)$, and thus estimated the sum $\sum_n g(n, p)$ by Laplace's method. Now, we have

$$n^{\log_t \tilde{w}} = \exp[\log_t n \ln(1 - \tilde{\Delta})] \sim \exp[-\ln(z_0^{-1})] = z_0,$$
$$e^{g(\tilde{w})} \sim \exp[\ln(z_0^{-1})] = \frac{1}{z_0},$$

and

$$\tilde{\alpha} + \tilde{w}^2 g''(\tilde{w}) \sim \tilde{\alpha} + \frac{1}{3}A_0(\tilde{\Delta})^{-4/3} \sim \frac{1}{3}A_0\left(\frac{\ln \tilde{n}}{-\ln t\ \ln z_0}\right)^{4/3}.$$

Thus (360) yields

$$\sum_n g(n, p) \sim \frac{1}{2\pi\tilde{n}}\left(\frac{1}{\ln \tilde{n}}\right)^{7/6}\sqrt{\frac{3}{A_0}}\,(\ln t)^{2/3}(-\ln z_0)^{7/6} \times \sqrt{\frac{2\pi}{-F_{nn}(\tilde{n}(p); p)}}\exp[F(\tilde{n}(p); p)]. \tag{361}$$

In view of (351) and (352) we get

$$F(\tilde{n}(p); p) = -\left(p + \frac{\tilde{n}}{\ln t}\right)\ln \tilde{w} = \left(p + \frac{\tilde{n}}{\ln t}\right)[\tilde{\Delta} + O(\tilde{\Delta}^2)]. \tag{362}$$

From (357), (358), (361) and (362) we obtain the growth rate

$$\ln\left[\sum_n g(n, p)\right] \sim \frac{p}{\ln p}(-\ln t\ \ln z_0) = \frac{p}{\ln p}\ln t\ \ln\left(\frac{t^t}{(t-1)^{t-1}}\right), \quad p \to \infty. \tag{363}$$

To refine (363), we consider (339) and (351) as a simultaneous system to determine $\tilde{w}$ and $\tilde{n}$ as functions of p. Expanding (351) around $w_* = 1$ and setting

$$S = S(p) = \ln \tilde{n}, \tag{364}$$

we obtain

$$\ln\left(\frac{1}{z_0}\right) - \frac{S}{\ln t}\Delta - \frac{S}{2\ln t}\Delta^2 - \frac{3}{2}A_0\Delta^{2/3} - A_1\Delta\ln\Delta + k_2\Delta - \frac{1}{\ln t}\Delta + O_R(S^{-2}) = 0, \tag{365}$$

where we also used (340). As $S \to \infty$ it follows from (365) that Δ has the expansion

$$\Delta = \frac{a}{S} + \frac{b}{S^{4/3}} + \frac{c}{S^{5/3}} + \frac{d}{S^2} + O_R(S^{-7/3}), \tag{366}$$

where

$$a = \ln t\ \ln\left(\frac{1}{z_0}\right), \quad b = 0, \tag{367}$$

$$c = -\frac{3}{2}(\ln t)A_0 a^{2/3} = -\frac{3}{2}(\ln t)^{5/3}A_0\left[\ln\left(\frac{1}{z_0}\right)\right]^{2/3}, \tag{368}$$

$$d = \ln t\left[\left(k_2 - \frac{1}{\ln t}\right)a - \frac{a^2}{2\ln t} + A_1 a\ln\left(\frac{S}{a}\right)\right], \tag{369}$$

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

We use (341) and (345), and re-write (339) as

$$\frac{p}{\tilde{n}} - \log_t \tilde{n} = (1 - \Delta)\left[A_0\Delta^{-1/3} + A_1\ln\Delta + A_2 + o(1)\right]. \tag{370}$$

We define n_* and Q by

$$\tilde{n}(p) = pn_*(p), \quad Q = \ln p, \tag{371}$$

with which (370) leads to

$$\begin{aligned}\frac{1}{n_*} - \frac{S}{\ln t} = A_0\left(\frac{S}{a}\right)^{1/3} - A_0\frac{c}{3a^{4/3}}\frac{1}{S^{1/3}} - A_0\frac{d}{3a^{4/3}}\frac{1}{S^{2/3}} \\ + A_1\left[-\ln\left(\frac{S}{a}\right) + \frac{c}{a}\frac{1}{S^{2/3}}\right] + A_2 - A_0\frac{a^{2/3}}{S^{2/3}} + O_R(S^{-1}).\end{aligned} \tag{372}$$

Here we also used (366) to expand Δ in (370). Noting that

$$S = Q + \ln n_* \sim Q,$$

and

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

we re-write (372) as

$$\frac{1}{n_*} = \frac{Q}{\ln t} + \frac{\ln n_*}{\ln t} + \frac{A_0}{a^{1/3}}Q^{1/3} + A_2 + A_1\ln a - A_1\ln Q + O_R(Q^{-1/3}). \tag{373}$$

Thus we obtain

$$n_* = \frac{\ln t}{Q} + \frac{v}{Q^{5/3}} + \frac{v'}{Q^2} + O_R(Q^{-7/3}), \tag{374}$$

where

$$v = -\frac{A_0}{a^{1/3}}(\ln t)^2,$$

and

$$v' = \ln t\,[\ln Q - \ln(\ln t)] + [A_1\ln Q - (A_2 + A_1\ln a)]\,(\ln t)^2.$$

Thus, as $p \to \infty$ (with $Q = \ln p$) we have

$$\begin{aligned}\tilde{n}(p) = \frac{p\ln t}{Q}\left[1 - \frac{A_0}{a^{1/3}}\frac{\ln t}{Q^{2/3}} + \frac{Z}{Q} + O_R(Q^{-4/3})\right], \\ Z = (A_1\ln t + 1)\ln Q - \ln\ln t - (A_2 + A_1\ln a)\ln t.\end{aligned} \tag{375}$$

Recalling that $n = \tilde{n}(p)$ corresponds to the maximum of $g(n, p)$ for a fixed large p, (376) is also the expansion of the asymptotic mean $\mathcal{N}(p)$, so we have derived (34). Also we note that

$$n_{max}(p) - \tilde{n}(p) \sim A_0\frac{(\ln t)^2}{a^{1/3}}\frac{p}{(\ln p)^{5/3}}, \quad p \to \infty,$$

where $n_{max}(p)$ is the inverse of $p_{max}(n)$ in (29). This gives an estimate of how the average number of nodes in a tree of path length $p \to \infty$ differs from the maximum number possible.

Next we refine (363) and derive an approximation to the variance in (356). From (355), (362) and (366) we obtain

$$\begin{aligned}\Psi_0(p) &= \left(p + \frac{\tilde{n}}{\ln t}\right)\left[\frac{a}{S} + \frac{c}{S^{5/3}} + \left(d + \frac{a^2}{2}\right)\frac{1}{S^2} + O_R(S^{-7/3})\right] \\ &= \left(p + \frac{p}{\ln p} + O\left(\frac{p}{Q^{5/3}}\right)\right)\left\{\frac{a}{Q} + \frac{c}{Q^{5/3}}\right. \\ &\quad \left. + \left[d + \frac{a^2}{2} + a(\ln Q - \ln(\ln t))\right]\frac{1}{Q^2} + O_R(S^{-7/3})\right\} \\ &= \frac{p}{Q}\left[a + \frac{c}{Q^{2/3}} + \left(a + \frac{a^2}{2} + d + a\ln Q - a\ln(\ln t)\right)\frac{1}{Q} + O_R(S^{-4/3})\right],\end{aligned} \tag{376}$$

where $a = -\ln t\ \ln z_0$ is as in (367). From (369) we find that

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

This establishes the result in (30) for $\ln\left[\sum_n g(n,p)\right]$.

To get the variance, we use (354) and (356). From (348) we obtain

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

and then (354) leads to

$$-nF_{nn} = \left[\frac{\ln n}{\ln t} + \alpha + \frac{1}{\ln t}\right]^2 \frac{3A_0^3}{\alpha^4}\left\{1 + \frac{1}{\alpha}\left[4A_2 + 3A_1 - 12A_1 \ln\left(\frac{\alpha}{A_0}\right)\right]\right\} + O\left(\frac{1}{\ln n}\right). \tag{377}$$

We set $n = \tilde{n}(p)$ and use

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

Then from (357) and (376) we obtain

$$\tilde{\alpha} \sim \frac{A_0}{(-\ln t \ln z_0)^{1/3}}\left[\ln p + \ln\ln t - \ln Q + O(Q^{-2/3})\right]^{1/3} \sim \frac{A_0}{(-\ln t \ln z_0)^{1/3}} Q^{1/3}.$$

Using the above in (377) we get

$$F_{nn}(\tilde{n}(p);p) \sim -\frac{1}{\tilde{n}(p)}\frac{Q^2}{(\ln t)^2}\frac{3}{A_0}[\ln t(\ln z_0)]^{4/3} Q^{-4/3} \sim -\frac{Q^{5/3}}{p}\frac{3\ln(z_0^{-1})a^{1/3}}{A_0(\ln t)^2}. \tag{378}$$

Thus we have obtained the leading term in the variance in (35). We will next use a somewhat different method to derive the correction term.

Using (378) and (376) in (361) yields (30). In the limit $\alpha \to \infty$ and $pn^{-4/3} \to 0$ we have derived the approximation

$$\begin{aligned} g(n,p) \sim\ & \frac{3}{\pi}\frac{t^{[\,t/(t-1)-1/6]}}{2^{5/3}(t-1)^{5/6}\sqrt{\ln t|r_0|}}\frac{1}{n}\left(\frac{A_0}{\alpha}\right)^{7/2}\exp\left[-\frac{1}{t-1}\frac{A_0^3}{\alpha^3}\frac{\ln n}{\ln t}\right] \\ & \times \frac{1}{z_0^n}\exp\left[-\frac{A_0^3 n}{2\alpha^2} + n\frac{A_0^3 n}{\alpha^3}(k_2 + 3A_1\ln\alpha - 3A_1\ln A_0)\right]. \end{aligned} \tag{379}$$

We obtained (379) by expanding (59) for $w_* \to 1$, and (379) matches to (288), as can be seen by replacing α in (379) by $p/n - \log_t n \sim p/n$. In this limit we note that

$$\gamma_* = \frac{16t}{27(t-1)}\frac{|r_0|^3}{\Theta^3} = \frac{16t}{27(t-1)}\frac{|r_0|^3 n^4}{p^3} \sim \frac{A_0^3}{\alpha^3}n.$$

Expanding (379) about $n = \tilde{n}(p)$ (or $\alpha = \tilde{\alpha}$) yields the Gaussian form in (36).

In view of (379), we define

$$\begin{aligned} H(\alpha) &= -\frac{A_0^3}{2\alpha^2} + \frac{A_0^3 n}{\alpha^3}(k_2 + 3A_1\ln\alpha - 3A_1\ln A_0), \\ \Phi(n,p) &\equiv -n\ln z_0 + nH(\alpha). \end{aligned} \tag{380}$$

Thus $\partial\Phi/\partial n = 0$ is equivalent to

$$-\ln z_0 + H(\alpha) - \frac{1}{\ln t}H'(\alpha) - \frac{p}{n}H'(\alpha) = 0. \tag{381}$$

We write the solution to (381) as

$$n = \hat{n}(p), \quad \hat{\alpha}(p) = \frac{p}{\hat{n}(p)} - \log_t[\hat{n}(p)],$$

where p is fixed. Since $\alpha = p/n - \log_t n$, we have

$$\frac{\partial^2\Phi}{\partial n^2} = -\frac{1}{n\ln t}H'(\alpha) + \frac{1}{n}\left(\frac{p}{n} + \frac{1}{\ln t}\right)^2 H''(\alpha), \tag{382}$$

and we define $\hat{\Phi}$ and $\hat{\mathcal{V}}(p)$ by

$$\hat{\Phi} = \Phi(\hat{n}(p), p), \quad \hat{\mathcal{V}}(p) = -\frac{1}{\Phi_{nn}(\hat{n}(p), p)}. \tag{383}$$

Solving for $\hat{n}$ asymptotically as $p \to \infty$ regains the expansion of the mean in (376). However using the more implicit form (381) has some advantage numerically (cf. section 9). Given p we can solve (381) numerically for $\hat{n}$ or $\hat{\alpha}$, and then compute $\hat{\Phi}$ from (380), and $\hat{\mathcal{V}}$ from (382) and (383). We can then use these numerical values of $(\hat{\Phi}, \hat{n}, \hat{\mathcal{V}})$ as approximations to the growth rate, mean and variance, and these are asymptotically equivalent to the results in (30)-(35). For example, solving (381) for $p \to \infty$ we obtain

$$\hat{\alpha} = A_0\left(\frac{Q}{a}\right)^{1/3} + A_2 + A_1\ln a - A_1\ln Q + O_R(Q^{-1/3}), \tag{384}$$

and we have

$$H''(\hat{\alpha}) = -\frac{3A_0^3}{\hat{\alpha}^4} + \frac{A_0^3}{\hat{\alpha}^5}\left[36A_1\ln\left(\frac{\hat{\alpha}}{A_0}\right) - 12A_2 - 9A_1\right]. \tag{385}$$

Using (384) in (385) yields

$$H''(\hat{\alpha}) = -\frac{3}{A_0}\left(\frac{a}{Q}\right)^{4/3} - \frac{9A_1}{A_0^2}\left(\frac{a}{Q}\right)^{5/3} + O_R(Q^{-2}). \tag{386}$$

Then we have

$$\frac{\partial^2\Phi}{\partial n^2}(\hat{n}(p), p) = \frac{1}{\hat{n}}\left(\frac{p}{\hat{n}} + \frac{1}{\ln t}\right)^2 H''(\hat{\alpha}) + O(\frac{1}{\hat{n}\ln\hat{n}}). \tag{387}$$

Using (386) in (387) we obtain the two term approximation to the variance in (35). Note that the leading terms in the expansion of the growth rate and the mean do not involve the root r_0 of the Airy function, but the leading term for the variance does involve r_0. The correction terms for the growth rate and mean are smaller than the leading terms by factor of $Q^{-2/3} = (\ln p)^{-2/3}$, while the correction term to the variance is smaller than its leading term by a factor of only in $O(Q^{-1/3}) = O((\ln p)^{-1/3})$. We have thus derived the various formulas in Results 1 and 2 of section 2.

9. NUMERICAL STUDIES

We provide numerical results that determine some unknown constants or functions, and these also give support to the various assumptions we made. Throughout this section we set $t = 3$, as the numerical trends are essentially independent of t.

For $w > 1$, we provided in Table 1, numerical values of $G_*(w)$ ($\sim G_n(w)t^{1-n}w^{-\binom{n}{2}}, n \to \infty$) for various w, and plotted this function in (cf. section 2). We test the accuracy of the asymptotic relation in (39). Let $\bar{G}_n(w)$ be the exact solution to (65). Our result shows that $\bar{G}_n(w) \to G_*(w)$ as $n \to \infty$ for $w > 1$, and that

$$(w-1)\ln[G_*(w)] \to d_0 = (t-1)\int_0^{\ln\left(\frac{t}{t-1}\right)} \frac{x}{e^x-1}dx = 0.732426\cdots, \tag{388}$$

as $w \downarrow 1$. In Table 2, we compute $(w-1)\ln[\bar{G}_\infty(w)]$ and $(w-1)\{\ln[\bar{G}_\infty(w)] - \frac{1}{2}\ln(w-1)\}$ for various $w > 1$. For each fixed w, $\bar{G}_\infty(w)$ is calculated by iterating (65) for n large enough until $\bar{G}_n(w)$ converges to a constant to 3 digits. Both of the tabulated functions should converge to d_0 as $w \downarrow 1$, with the latter one converging more rapidly, since it includes information from the algebraic factor $\sqrt{w-1}$ in (39). Table 2 shows that both functions are indeed approaching d_0 as $w \to 1$, with the second function having a faster rate of convergence. In Table 3 we calculate

$$D_1(w) \equiv \exp\left(-\frac{d_0}{w-1}\right)\frac{\bar{G}_\infty(w)}{\sqrt{w-1}}, \tag{389}$$

for various $w > 1$, to try to confirm the constant d_1 in (39). Our analysis predicts that $D_1(w) \to d_1 = 1.72613\cdots$ as $w \downarrow 1$. It became very difficult to compute $\bar{G}_\infty(w)$ for $w \leq 1.04$, since for $w \gtrsim 1$ the convergence of $\bar{G}_n$ to $\bar{G}_\infty$ 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 data trend is consistent with convergence to the theoretical value, but w would have to made much closer to 1 for us to make a more definite conclusion.

Table 2: Numerical d_0

w	$(w-1)\ln(\bar{G}_\infty)$	$(w-1)\{\ln[\bar{G}_\infty(w)] - \frac{1}{2}\ln(w-1)\}$
2	0.4688	0.4688
1.8	0.4860	0.5753
1.6	0.5090	0.6622
1.4	0.5418	0.7251
1.2	0.5949	0.7558
1.18	0.6023	0.7566
1.16	0.6103	0.7569
1.14	0.6190	0.7566
1.12	0.6285	0.7557
1.10	0.6391	0.7542
1.08	0.6510	0.7520
1.06	0.6645	0.7490
1.04	0.6805	0.7449
1.02	0.696	0.735
1	0.7324	0.7324

We next study the β-scale, where $w = 1 + \beta/n = 1 + O(n^{-1})$ with $\beta > 0$. Then (41) gives the asymptotic result. In Table 4 and Table 5, we compare

$$\frac{1}{n}\ln\left[G_n\left(1+\frac{\beta}{n}\right)\right], \tag{390}$$

to $\Phi(\beta) = \ln t + \frac{\beta}{2} + \phi(\beta)$ (cf. (41)) for $\beta = 0.25,\ 0.5,\ 1,\ 2$ and 4, and for various n. Our **WKB** expansion predicts that (390) should approach $\Phi(\beta)$ as $n \to \infty$. The data in the tables clearly demonstrates this convergence. Also, the data are consistent with an $O(n^{-1}\ln n)$ error term, which is indicated by (41). The smaller the β the slower the convergence (cf. Table 4), which is consistent with our analysis that once β becomes $O(n^{-1/2})$, the expansion becomes invalid and we have to use the a-scale result.

Table 3: Numerical $D_1(w)$

w	$D_1(w)$
5	0.458554
4.5	0.484995
4	0.516684
3.5	0.555658
3	0.605303
2.5	0.671797
2	0.768269
1.8	0.821631
1.6	0.889573
1.4	0.981769
1.2	1.12419
1.18	1.14386
1.16	1.16520
1.14	1.18857
1.12	1.21446
1.10	1.24360
1.08	1.27714
1.06	1.31707
1.04	1.36747
1	1.72614

Table 4: Numerical exponential rate $\beta = 0.25, 0.5, 1$

(a) $\beta = 0.25$

n	$\frac{1}{n}\ln(G_n)$	$\Phi(\beta)$
10	1.4913	1.9234
20	1.6756	
30	1.7473	
40	1.7861	
50	1.8106	
75	1.8450	
100	1.8631	
125	1.8743	
150	1.8819	
200	1.8917	
250	1.8977	
500	1.9100	
750	1.9142	
1000	1.9164	

(b) $\beta = 0.5$

n	$\frac{1}{n}\ln(G_n)$	$\Phi(\beta)$
10	1.5659	1.9597
20	1.7416	
30	1.8075	
40	1.8424	
50	1.8640	
75	1.8938	
100	1.9091	
125	1.9185	
150	1.9284	
200	1.9329	
250	1.9379	
500	1.9482	
750	1.9518	
1000	1.9536	

(c) $\beta = 1$

n	$\frac{1}{n}\ln(G_n)$	$\Phi(\beta)$
10	1.7169	2.0778
20	1.8842	
30	1.9439	
40	1.9747	
50	1.9935	
75	2.0194	
100	2.0328	
125	2.0410	
150	2.0466	
200	2.0537	
250	2.0581	
500	2.0673	
750	2.0706	
1000	2.0723	

In Table 6, we compare

$$\sqrt{n}e^{-n\Phi(\beta)}G_n\left(1+\frac{\beta}{n}\right), \tag{391}$$

to $\sqrt{\beta}\hat{g}(\beta)$, for $\beta = 0.5,\ 1$, and 2, and for various n. Our analysis predicts that the limit of (391) should be $\sqrt{\beta}\hat{g}(\beta)$. The data again demonstrate the convergence suggested by the **WKB** *ansatz* (80). However, the convergence is much slower than that predicted by the $O(n^{-1})$ error term in (80). The numerical results

Table 5: Numerical exponential rate $\beta = 2, 4$

(a) $\beta = 2$

n	$\frac{1}{n}\ln(G_n)$	$\Phi(\beta)$
10	2.0235	2.4192
20	2.2067	
30	2.2717	
40	2.3054	
50	2.3262	
75	2.3548	
100	2.3696	
125	2.3787	
150	2.3849	
200	2.3928	
250	2.3977	
500	2.4078	
750	2.4113	
1000	2.4132	

(b) $\beta = 4$

n	$\frac{1}{n}\ln(G_n)$	$\Phi(\beta)$
10	2.6267	3.2787
20	2.9181	
30	3.0273	
40	3.0850	
50	3.1208	
75	3.1701	
100	3.1956	
125	3.2113	
150	3.2219	
200	3.2353	
250	3.2435	
500	3.2604	
750	3.2662	
1000	3.2692	

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 expansion in (80). To resolve this issue more conclusively, we would have to study higher-order asymptotic matchings between the a-scale and the β-scale. This would require, among other things, continuing (44) into the range $a > 0$ and evaluating the result for $a \to +\infty$.

Table 6: Numerical results for $\sqrt{n}G_n\exp(-n\Phi)$ for $\beta = 0.5, 1, 2$

(a) $\beta = 0.5$

n	$\sqrt{n}e^{-n\Phi}G_n$	$\sqrt{\beta}\hat{g}$
10	0.06161	0.07439
20	0.05697	
30	0.05699	
40	0.05792	
50	0.05902	
75	0.06148	
100	0.06327	
125	0.06453	
150	0.06544	
200	0.06664	
250	0.06741	
500	0.06919	
750	0.06996	
1000	0.07042	

(b) $\beta = 1$

n	$\sqrt{n}e^{-n\Phi}G_n$	$\sqrt{\beta}\hat{g}$
10	0.08564	0.12849
20	0.09311	
30	0.09853	
40	0.10208	
50	0.10450	
75	0.10815	
100	0.11027	
125	0.11172	
150	0.11280	
200	0.11434	
250	0.11541	
500	0.11821	
750	0.11952	
1000	0.12037	

(c) $\beta = 2$

n	$\sqrt{n}e^{-n\Phi}G_n$	$\sqrt{\beta}\hat{g}$
10	0.06045	0.08266
20	0.06384	
30	0.06556	
40	0.06671	
50	0.06757	
75	0.06907	
100	0.07008	
125	0.07084	
150	0.07143	
200	0.07233	
250	0.07298	
500	0.07481	
750	0.07574	
1000	0.07633	

Next we consider the a-scale result. Recall that obtaining the right tail of the Airy distribution involved showing that

$$D(y) = D(a^{2/3}) \sim \frac{\sqrt{2}\,t^{3/2}}{\sqrt{\pi}(t-1)^{5/2}}\,a\exp\left[\frac{t}{6(t-1)}a^2\right],\ a \to +\infty. \tag{392}$$

In Table 7 we give $D(a^{2/3})a^{-1}e^{-\frac{t}{6(t-1)}a^2}$ for various values of $a \geq 1$. The table clearly shows that the function is converging to the constant $3\sqrt{3}/(4\pi) = 0.73290\cdots$. To compute the exact values of $D(y)$ we used (145), (146) and (154). Then the u_j in (146) were computed by solving the recurrence

$$V_{n+1} = \frac{\sqrt{t}}{\sqrt{2(t-1)}}\left[\left(\frac{3}{2}n+1\right)V_n + \frac{(t-1)^2}{2t}\sum_{i=0}^{n} V_i V_{n-i}\right], \quad v_0 = \frac{t}{2(t-1)^2}, \tag{393}$$

and using

$$u_n = m_{n+1} = \frac{(n+1)!}{\Gamma(\frac{3}{2}n+1)}V_n.$$

Table 7: a-scale numerical constant

a	$D(a^{2/3})a^{-1}e^{-a^2/4}$
1	0.7651650024
2	0.6619976925
3	0.6722963306
4	0.6940454045
5	0.7091235006
6	0.7174633676
7	0.7220450344
8	0.7247911444
9	0.7265872121
10	0.7278366898
11	0.7287445990
12	0.7294264389
13	0.7299521449
14	0.7303663257
15	0.7306986147
20	0.7316718710
25	0.7321178292
30	0.7323589279
50	0.7327082900
100	0.7328557843
∞	0.7329037680

Next we discuss the expansion on the γ-scale, where $w = 1 - O(n^{-1})$. We need to solve (269) or (271). We make use of the asymptotic behavior of $F(\theta)$ in (273) as $\theta \to \infty$, and use the following numerical scheme: (1) Fix a large N, (2) let $\tilde{F}(L)$ satisfy the recurrence

$$\tilde{F}(L-1) = \tilde{F}(L) + \tilde{F}^2(L) + \frac{1}{3}\tilde{F}^3(L), \quad L = N, N-1, \cdots, -M+1, \tag{394}$$

subject to the terminal condition

$$\tilde{F}(N) = \frac{1}{N} + \frac{\frac{2}{3}\ln N + \alpha_*}{N^2} + \frac{(\frac{2}{3}\ln N + \alpha_*)^2 - \frac{4}{9}\ln N - \frac{2}{3}\alpha_* + \frac{2}{9}}{N^3},$$

(3) iterate (394) backward until $L = -M+1$, and (4) use the approximation

$$F(-M) \approx \tilde{F}(M).$$

For M positive and large, (279) implies that for $t = 3$

$$3^{-M}\left[\ln[F(-M)] - \frac{1}{2}\ln 3\right] = \ln A - (const.)3^{-M}A^{-3^M} + \cdots, \quad M \to \infty. \tag{395}$$

The left side of (395) should converge to the constant $\ln A$ very rapidly as $M \to \infty$, if $A > 1$. Our numerical results show that $M = 12$ is sufficient to get $\ln A$ to several decimal places. We take $M = 12$, solve (394) for various N and provide $3^{-M}\left[\ln[F(-M)] - \frac{1}{2}\ln 3\right]$, which is an estimate to $\ln A$. This shows that for $t = 3$ (cf. Table 8)

$$\ln A \approx 6.696, \tag{396}$$

and then (281) yields

$$k_2 \approx 2.727. \tag{397}$$

We recall that k_2 corresponds to the exponential growth rate of $F_1(\gamma)$ in (278), and thus that of the approximation in (251) on the γ-scale, or (46)-(49).

Table 8: Numerical ln(A)

N	$3^{-M}\left\{\ln[F(-M)] - \frac{1}{2}\ln 3\right\}$
1000	6.6948
2000	6.6958
3000	6.6961
5000	6.6962
10000	6.6963

Table 9: Numerical $g(w)$

w	$g_{num}^{(n)}(w)$
0.999999	1.9081
0.99999	1.9070
0.9999	1.9023
0.999	1.8832
0.99	1.8127
0.9	1.6314
0.8	1.5735
0.7	1.5638
0.6	1.5867
0.5	1.6402
0.4	1.7297
0.3	1.8697
0.2	2.0964
0.1	2.53
0.01	4.15-4.20
0.001	5.83-6.08
0.0001	7.52-8.03
0.00001	9.21-10.01

We next consider $n \to \infty$ with $0 < w < 1$, where approximation (53) applies (or its refined form (50), that contains oscillations). Neglecting the oscillations and defining

$$g_{num}^{(n)}(w) \equiv \frac{1}{n+\frac{1}{2}} \ln\left[\frac{G_n(w)w^{-n\log_3 n}\sqrt{2n\pi\ln 3}}{n^{\frac{\ln w}{2\ln 3}} w^{\frac{3}{4}+\frac{1}{2\ln 3}}\sqrt{-\ln w}}\right], \tag{398}$$

Table 10: Numerical k_2

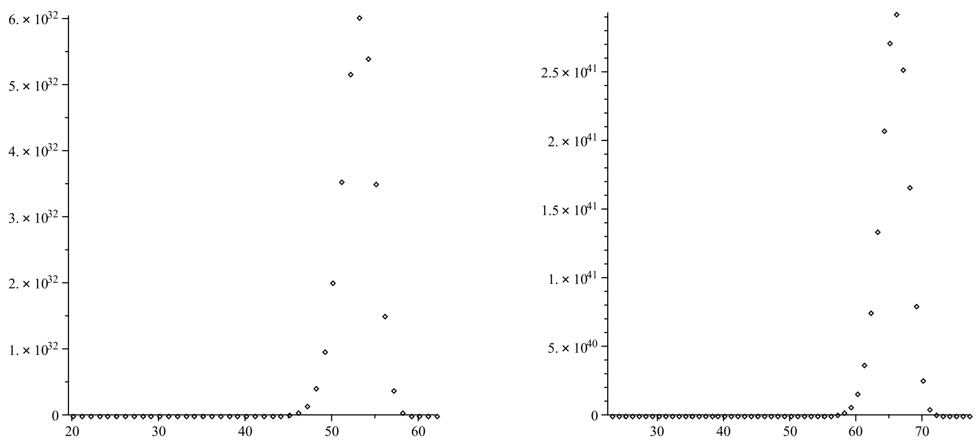
w	$g_{num}^{(n)}(w)$	$CONST$
0.8	1.57353	2.9789
0.9	1.63140	2.8994
0.925	1.65929	2.8773
0.95	1.69757	2.8529
0.975	1.75525	2.8228
0.99	1.81266	2.7960

we should have $g_{num}^{(n)}(w) \to g(w)$ as $n \to \infty$ each fixed $0 < w < 1$. We also should have $n^{-1} \ln \left[G_n(w) w^{-n \log_3 n}\right] \to g(w)$, but (398) converges faster. In Table 9 we list data for various w and n large. These data show that for n large enough, $g_{num}^{(n)}(w)$ is constant to several decimal places, with the exception of very small values of w. When $w \leq 0.01$ the function oscillates with n over a certain range. The oscillations corresponding to $g_k(w)$ in (51) indeed exist, but they are numerically small,unless w itself is very small. For practical purposes we can use (53) and (59) as the approximations to $G_n(w)$ and $g(n, p)$, though $g(w)$ still needs to be determined numerically, e.g., using Table 9. We also observe that $g(w)$ is reaching the theoretical value $g(1) = \ln\left(\frac{27}{4}\right) = \cdots$ as $w \to 1$.

We study the matching region between the scales $0 < w < 1$ and $w = 1 - O(n^{-1})$ in Table 10. We have shown that $g(w)$ behaves as in (340) for $w \uparrow 1$. Let us define

$$\begin{aligned} CONST &= CONST(w; n) \\ &\equiv \frac{1}{1-w} \left\{ g_{num}^{(n)}(w) - \left[\ln\left(\frac{27}{4}\right) + 3^{1/3} r_0 (1-w)^{2/3} + \left(\frac{1}{\ln 3} - \frac{2}{9}\right)(w-1)\ln(1-w) \right] \right\}. \end{aligned} \tag{399}$$

We can choose n sufficiently large to make (399) independent of n to several decimal places. As $w \to 1$, $CONST$ should converge to the constant k_2. The convergence becomes very slow with n when w is very close to one. However the data show that $CONST$ is decreasing and converging to the theoretical value in (397). Thus we have provided numerical evidence both for the value of k_2 and the asymptotic form in (340).



(a) $p = 190$ with n in the range [20..62]

(b) $p = 250$ with n in the range [23..77]

Figure 3: Plots of $g(n, p)$

Finally, we provide numerical studies for our results on the growth rate of the total number of trees of a given path length, and the distribution of the trees by their number of nodes. In Figure 3, we plot the exact $g(n, p)$ for $p = 190$ and $p = 250$ with $n \in [n_{min}(p), n_{max}(p)] = [20, 62]$ and $[23, 77]$ respectively. The graphs possibly resemble a Gaussian near their peak, but these values of p are too small to conclude this definitely.

ACKNOWLEDGEMENTS

This work was supported in part by NSA Grant H 98230-08-1-0102.

REFERENCES

[1] Knuth, D. E. (1997). *The art of computer programming, volume 1: Fundamental algorithms* (3rd Edition). Reading, MA: Addison-Wesley.

[2] Preiss, B. R. (1999). *Data structures and algorithms with object-oriented design patterns in Java*. New York: John Wiley & Sons.

[3] Ziv, J., & Lempel, A. (1977). A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory, 23*(3), 337-343.

[4] Ziv, J., & Lempel, A. (1978). Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory, 24*(5), 530-536.

[5] Seroussi, G. (2004). *Universal types and simulation of individual sequences, volume 2976 of LATIN 2004: theoretical informatics*. Heidelberg, Berlin: Springer.

[6] Seroussi, G. (2006). On universal types. *IEEE Transactions on Information Theory, 52*(1), 171-189.

[7] Flajolet, P., & Sedgewick, R. (2009). *Analytic combinatorics*. New Jersey: Cambridge University Press.

[8] Knuth, D. E. (2000). *Selected papers on the analysis of algorithms*. Cambridge: Cambridge Univeristy Press.

[9] Szpankowski, W. (2001). *Average case analysis of algorithm on sequences*. New York: John Wiley & Sons.

[10] Takács, L. (1991). Conditional limit theorem for branching processes. *J. Applied Mathematics and Stochastic Analysis, 4*(4), 263-292.

[11] Takács, L. (1993). The asymptotic distribution of the total heights of random rooted trees. *Acta Sci. Math. (Szegad), 57*(1-4), 613-625.

[12] Flajolet, P., & Odlyzko, A. (1982). The average height of binary trees and other simple trees. *Journal of Computer and System Sciences, 25*(2), 171-213.

[13] Louchard, G. (1984). The Brownian excursion area: A numerical analysis. *International Journal of Computers and Mathematics with Applications, 10*(6), 413-417.

[14] Louchard, G. (1984). Kac's formula, Levy's local time and Brownian excursion. *J. Appl. Prob., 21*(3), 479-499.

[15] Takács, L. (1991). A Bernoulli excursion and its various applications. *J. Appl. Probab., 23*(3), 557-585.

[16] Abramowitz, M., & Stegun, I. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York: Dover.

[17] Andrews, G., Askey, R., & Roy, R. *Special functions*. Cambridge: Cambridge University Press, 1999.

[18] Csörgö, M., Shi, Z., & Yor, M. (1999). Some asymptotic properties of the local time of the uniform

empirical process. *Bernoulli, 5*(6), 1035-1058.

[19] Flajolet, P., & Louchard, G. (2001). Analytic variation on the Airy distribution. *Algorithmica, 31*(3), 361-377.

[20] Flajolet, P., Poblete, P., & Viola, A. (1998). On the analysis of linear probing hashing. *Algorithmica, 22*(4), 490-515.

[21] Knuth, D. E. (1998). Linear probing and graphs. *Algorithmica, 22*(4), 561-568.

[22] Majumdar, S. N., & Comtet, A. (2005). Airy distribution function: From the area under a Brownian excursion to the maximal height of fluctuating interfaces. *Journal of Statistical Physics, 119*(3-4), 777-826.

[23] Knessl, C., & Szpankowski, W. (2005). Enumeration of binary trees and universal types. *Discrete Mathematics & Theoretical Computer Science, 7*(1), 313-400.

[24] Seroussi, G. (2006). On the number of *t*-ary trees with a given path length. *Algorithmica, 46*(3-4), 557-565.

[25] Csiszár, I. (1998). The method of types. *IEEE Transactions on Information Theory, 44*(6), 2505-2523.

[26] Csiszár, I., & Körner, J. (1981). *Information theory: Coding theorems for discrete memoryless systems*. New York: Academic Press.

[27] Jacquet, P., & Szpankowski, W. (1995). Asymptotic behavior of the Lempel-Ziv parsing scheme and digital search trees. *Theoretical Computer Science, 144*(1-2), 161-197.

[28] Hennequin, P. (1989). Combinatorial analysis of quicksort algorithm. *Theoretical Informatics and Applications, 23*(3), 317-333.

[29] Knessl, C., & Szpankowski, W. (1999). Quicksort algorithm again revisted. *Discrete Mathematics & Theoretical Computer Science, 3*(2), 43-64.

[30] Rëgnier, M. (1989). A limiting distribution for quicksort. *Theoretical Informatics and Applications, 23*(3), 335-343.

[31] Rösler, U. (1991). A limit theorem for quicksort. *Theoretical Informatics and Applications, 25*(1), 85-100.

[32] Bender, C., & Orszag, S. (1999). *Advanced mathematical methods for scientists and engineers*. New York: Springer-Verlag.