Node Profiles of Symmetric Digital Search Trees

Michael Drmota
Institute for Discrete Mathematics and Geometry
Vienna University of Technology
1040 Vienna
Austria

Michael Fuchs*
Department of Applied Mathematics
National Chiao Tung University
Hsinchu 300
Taiwan

Hsien-Kuei Hwang†
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

Ralph Neininger
Institute for Mathematics
Goethe University
60054 Frankfurt a.M.
Germany

May 30, 2019

Abstract

We give a detailed analysis in distribution of the profiles of random symmetric digital search trees, which are in close connection with the performance of the search complexity of random queries in such trees. While the expected profiles have been analyzed for several decades, the analysis of the variance turns out to be very difficult and challenging, and requires the combination of several different analytic techniques, as established the first time in this paper. We also prove by contraction method the asymptotic normality of the profiles in the range where the variance tends to infinity. As consequences of our results, we obtain a two-point concentration for the distributions of the height and the saturation level of random digital search trees.

AMS 2010 subject classifications. Primary 05A16,60C05, 68Q25; secondary 68P05, 60F05.

Key words. Digital search tree, level profile, two-point concentration, double-indexed recurrence, asymptotic transfer, Poissonization, Laplace transform, Mellin transform, contraction method, asymptotic normality.

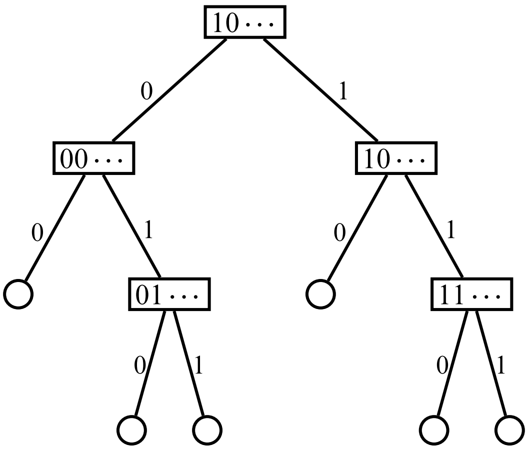
1 Introduction and Results

Digital trees are fundamental data structures for words or alphabets in computer algorithms whose analysis has attracted much attention over the last half century. One major such varieties is the *digital search tree* (DST for short), introduced by Coffman and Eve in 1970 [5] (see also [23] for more information). Such structures are closely related to the popular Lempel-Ziv compression scheme, and their asymptotic stochastic behaviors under random inputs are often more challenging than those for other digital tree families because of the natural occurrence of differential-functional equations instead of purely algebraic-functional equations.

*Partially supported by MOST under the grants MOST-104-2923-M-009-006-MY3 and MOST-105-2115-M-009-010-MY2.

†Partially supported by an Investigator Award from Academia Sinica under the Grant AS-IA-104-M03.

We begin with the definition of DST, which is the main object of study in this paper. In the simplest situation, it is built from digital data consisting of a sequence of records in the form of 0-1 strings. The first record is stored at the root of the tree. All other records are distributed to the left- or right-subtree according as their first bit being 0 or 1, respectively, and retain their relative order. The subtrees of the root are then built according to the same rules but by using the jth digit at level j in further directing the strings to their respective subtrees. The splitting process stops when the size of the subtree is either zero or one. Note that the resulting tree is a binary tree with internal nodes holding the records. External nodes, which represent places where future records can be inserted, are often added to the tree (in fact, two external nodes are automatically created in the algorithmic implementation for each new internal node); see Figure 1 for an example of a DST built from five records (internal nodes are represented by rectangles and external nodes by circles).



$$B_{5,0} = 0, \quad I_{5,0} = 1;$$

$$B_{5,1} = 0, \quad I_{5,1} = 2;$$

$$B_{5,2} = 2, \quad I_{5,2} = 2;$$

$$B_{5,3} = 4, \quad I_{5,3} = 0.$$

Figure 1: *A DST built from* 5 *records with its profiles.*

For the purpose of analysis, we assume that bits in the input strings are independent and identically distributed with a common Bernoulli(p) random variable with $0 < p < 1$. Throughout this paper, we fix $p = \frac{1}{2}$, namely, we consider only the symmetric case. This random model is called the *symmetric Bernoulli model* and the corresponding random tree is referred to as a *random symmetric DST*. It represents a simple model with reasonably good predictive power in general (for example, results holding in the Bernoulli model often have similar forms in more general Markov models; see [26]).

Under such a random model, we study in this paper the external and internal node profiles (referred to as the profiles for short) which are defined as follows: the external profile of a random symmetric DST of size n is a double-indexed sequence of random variables $B_{n,k}$ which counts the number of external nodes at (horizontal) level k; the internal profile $I_{n,k}$ is similarly defined (with external nodes replaced by internal nodes).

Profiles are *fine shape characteristics* encoding the level silhouette of the tree and they are closely connected to many other shape parameters such as height, width, total path length, saturation or fill-up level, and successful and unsuccessful search. In particular, we will discuss the unsuccessful search (or the depth), the height and the saturation level:

- Unsuccessful search U_n: the distance from the root to a randomly chosen external node with its distribution given by
$$\mathbb{P}(U_n = k) = \frac{\mathbb{E}(B_{n,k})}{n+1}. \tag{1}$$
- Height: the length of the longest path from the root to an external node, or $\max\{k : B_{n,k} > 0\}$;

- Saturation level: the last level from the root which is completely filled with internal nodes, or $\max\{k : I_{n,k} = 2^k\}$.

See for example [8, 16] and the references therein for more shape parameters in DSTs.

Historically, the external profile was among the very first shape parameters analyzed on DSTs due to the connection (1) to the unsuccessful search; see Knuth [23] and Konheim and Newman [24]. Yet our understanding of the profiles of symmetric DSTs has remained incomplete. Table 1 summarizes the current status for the profiles of tries, Patricia tries and DSTs, the latter two representing other major classes of digital trees.

Trees	$p = q$?	Mean	Variance	CLT
Tries	$0 < p < 1$	[32]	[32]	[32]
Patricia Tries	$p = \frac{1}{2}$	[28]	?	?
	$p \neq \frac{1}{2}$	[11, 28, 27]	[11, 27]	[27]
DSTs	$p = \frac{1}{2}$	[25, 10]	this paper	this paper
	$p \neq \frac{1}{2}$	[10]	[19]	?

Table 1: A summary of the analysis in distribution of profiles in the three major classes of digital trees under the Bernoulli model.

Briefly, in the case of random tries, the mean, the variance and the asymptotic normality of both profiles under the symmetric and asymmetric models are fully clarified in [32]. For Patricia tries, the expected profiles were studied in [28] for both symmetric and asymmetric models. Then the asymptotic variance and the asymptotic normality of the profiles, *inter alia*, under the asymmetric model are established in the recent papers [11, 27].

As regards symmetric DSTs, Louchard [25], following [23, 24], derived an explicit expression for the expected profiles; see also [8, 10, 29, 33]. Louchard also obtained an asymptotic approximation for the mean profiles in the most important range $k = \log_2 n + \mathcal{O}(1)$ (where most nodes lie), characterizing the asymptotic distribution of unsuccessful and successful search. These results were later extended in [8, 10, 22, 29]. We broaden the study in this paper to the variance of the profiles for which an arduous analysis is carried out. We also clarify the asymptotic normality of the profiles in the range where the variance becomes unbounded. Moreover, we will apply our results to the height and the saturation level. See also [6, 18, 26, 34] for other parameters and different types of results on profiles in DSTs .

We now state our results, focusing on the external profile. The corresponding results for the internal profile will be given in Section 5. First, we introduce the following function and sequence that are ubiquitous in the analysis of DSTs; see [23].

$$Q(z) = \prod_{\ell \geqslant 1} (1 - 2^{-\ell} z) \qquad \text{and} \qquad Q_n = \prod_{1 \leqslant \ell \leqslant n} (1 - 2^{-\ell}) = \frac{Q(1)}{Q(2^{-n})}.$$

Note that $\lim_{n\to\infty} Q_n$ exists and equals $Q(1) =: Q_\infty$.

In the next section, we will derive the following (known) result (see [8]) for the mean of the external profile.

Theorem 1. *The mean of the external profile satisfies*

$$\mathbb{E}(B_{n,k}) = 2^k F\big(2^{-k} n\big) + \mathcal{O}(1), \tag{2}$$

uniformly for $0 \leqslant k \leqslant n$*, where* $F(x)$ *is a positive function on* $\mathbb{R}^+$ *defined by*

$$F(x) = \sum_{j \geqslant 0} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_\infty} e^{-2^j x}. \tag{3}$$

The proof of (2) for small k, or more precisely, for k such that $2^{-k} n \to \infty$, will follow readily by simple elementary arguments, whereas for the remaining range complex-analytic tools will be used. More precisely, when $2^{-k} n \to \infty$ and $k \geqslant 1$, we will show that

$$\mathbb{E}(B_{n,k}) = \frac{2^k}{Q_k} \big(1 - 2^{-k}\big)^n \big(1 + \mathcal{O}\big(e^{-\frac{n}{2^k - 1}}\big)\big), \tag{4}$$

which is stronger than (2) if $2^k F\big(2^{-k} n\big) = \mathcal{O}(1)$.

Remark 1. Note that (4) indeed holds for all n and k but is more useful in the range when $2^{-k} n \to \infty$.

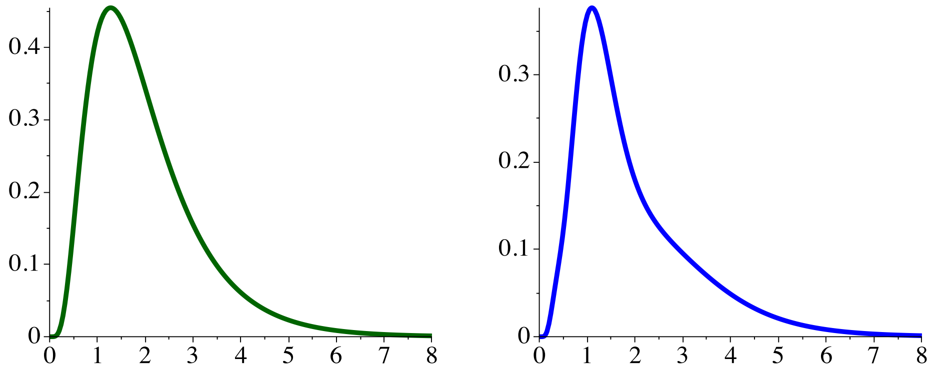


Figure 2: The functions F (left) and G (right).

On the other hand, the relation (2) is only a (useful) asymptotic approximation if the first term on the right-hand side is not bounded for large n. Thus, to understand when this holds, we derive more precise asymptotic behaviors of $F(x)$ for large and small x; see Figure 2 (left) for a graphical rendering of F.

Observe first that the series definition (3) of F extends to complex parameter z with $\Re(z) \geqslant 0$ and is itself an asymptotic expansion for large $|z|$:

$$F(z) = \frac{e^{-z}}{Q_\infty} + \mathcal{O}\big(e^{-2\Re(z)}\big), \qquad (\Re(z) \geqslant 0). \tag{5}$$

On the other hand, for small x with $X := \frac{1}{x \log 2}$ (see Proposition 1),

$$\begin{aligned} F(x) = \sqrt{\frac{\log 2}{2\pi}}\, X^{\frac{1}{2} + \frac{1}{\log 2}} \exp\bigg(&-\frac{(\log(X \log X))^2}{2 \log 2} - P(\log_2(X \log X))\bigg) \\ &\times \bigg(1 + \mathcal{O}\Big(\frac{(\log\log X)^2}{\log X}\Big)\bigg), \end{aligned} \tag{6}$$

with $P(t)$, $t \in \mathbb{R}$, a 1-periodic function whose Fourier series is given explicitly by

$$P(t) := \frac{\log 2}{12} + \frac{\pi^2}{6 \log 2} - \sum_{j \geqslant 1} \frac{\cos(2 j \pi t)}{j \sinh\big(\frac{2 j \pi^2}{\log 2}\big)}. \tag{7}$$

Note that the series in (7), representing the fluctuating part of $P(t)$, $t \in \mathbb{R}$, has a (peak-to-peak) amplitude less than 1.8×10^{-12}. The expansion (6) and (7) can be extended to complex z with $|z| \leqslant \varepsilon$ and $|\arg(z)| \leqslant \frac{\pi}{2} - \varepsilon$; see Proposition 1.

While it is well anticipated (from known results for tries and Patricia tries) that $\mathrm{Var}(B_{n,k})$ is asymptotically of the same form as (2) for $\mathbb{E}(B_{n,k})$ in most ranges of k of interest, the function involved is surprisingly very complicated, as shown in (9) below; see also Figure 2 (right).

Theorem 2. *The variance of the external profile satisfies*

$$\mathrm{Var}(B_{n,k}) = 2^k G\big(2^{-k} n\big) + \mathcal{O}(1), \tag{8}$$

uniformly for $0 \leqslant k \leqslant n$, *where* $G(x)$ *is a positive function on* $\mathbb{R}^+$ *defined by*

$$G(x) = \sum_{j,r \geqslant 0} \sum_{0 \leqslant h,\ell \leqslant j} \frac{(-1)^{r+h+\ell} 2^{-j-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}+2h+2\ell}}{Q_\infty Q_r Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \varphi\big(2^{r+j}, 2^h + 2^\ell; x\big), \tag{9}$$

with

$$\varphi(u, v; x) = e^{-ux} \int_0^x t e^{(u-v)t} \,\mathrm{d}t = \begin{cases} \dfrac{e^{-ux} + ((u-v)x - 1)\, e^{-vx}}{(u-v)^2}, & \text{if } u \neq v; \\ \frac{1}{2} x^2 e^{-ux}, & \text{if } u = v. \end{cases} \tag{10}$$

Remark 2. In the case when $2^{-k} n \to \infty$, we will in fact prove that

$$\mathbb{E}(B_{n,k}) \sim \mathrm{Var}(B_{n,k}).$$

Despite of its complicated form, the function G is very close to F in the following sense (see Section 3):

$$G(x) \sim \begin{cases} F(x), & \text{if } x \to \infty; \\ 2F(x), & \text{if } x \to 0; \end{cases} \tag{11}$$

see also [32] for the same type of results for symmetric tries, and Devroye [7] for a general bound for the profile variance. A more precise approximation when $x \to \infty$ is

$$G(x) = \frac{e^{-x}}{Q_\infty} + \mathcal{O}\big(x e^{-2x}\big),$$

where the second-order term differs from that of F; see (5).

The two theorems imply that the mean and the variance have very similar behaviors. In particular, they tend to infinity in the same range of k.

Corollary 1. *For large n and $0 \leqslant k \leqslant n$, $\mathbb{E}(B_{n,k}) \to \infty$ iff $\mathrm{Var}(B_{n,k}) \to \infty$.*

We now describe the range where the mean and the variance tend to infinity. Define two functions of n:

$$\begin{aligned} k_s &:= \log_2 n - \log_2 \log n + 1 + \frac{\log_2 \log n}{\log n}, \\ k_h &:= \log_2 n + \sqrt{2 \log_2 n} - \frac{1}{2} \log_2 \log_2 n + \frac{1}{\log 2} - \frac{3 \log\log n}{4\sqrt{2 (\log n)(\log 2)}}. \end{aligned} \tag{12}$$

Corollary 2. *The mean and the variance of $B_{n,k}$ tend to infinity iff there exists a positive sequence ω_n tending to infinity with n such that*

$$k_s + \frac{\omega_n}{\log n} \leqslant k \leqslant k_h - \frac{\omega_n}{\sqrt{\log n}}. \tag{13}$$

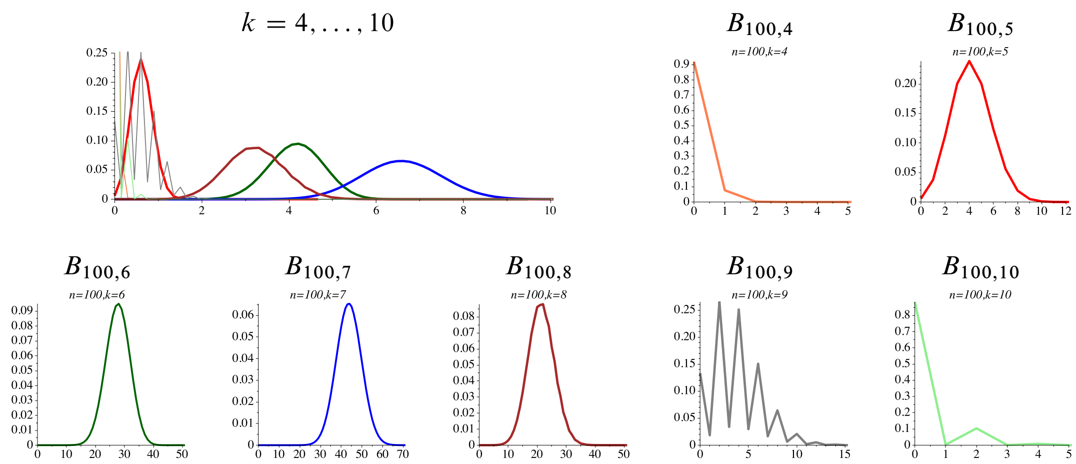


Figure 3: *Histograms of* $B_{100,k}$ *for* $k = 4, \dots, 10$ *with the same color used for each histogram: coral for* $k = 4$*, red for* $k = 5$*, dark green for* $k = 6$*, blue for* $k = 7$*, brown for* $k = 8$*, gray for* $k = 9$ *and light green for* $k = 10$*. Numerically,* $\mathbb{E}(H_{100}) = 8.98615\dots$

This range is very small (or almost all nodes are concentrated at these levels); see Figure 3. For convenience, we will refer to (13) as the *central range*, and it is exactly this range where the sequence of random variables $\{B_{n,k}\}_n$ follows a central limit theorem.

Theorem 3. *If* $\mathrm{Var}(B_{n,k}) \to \infty$*, then the external profile is asymptotically normally distributed:*

$$\frac{B_{n,k} - \mathbb{E}(B_{n,k})}{\sqrt{\mathrm{Var}(B_{n,k})}} \xrightarrow{d} \mathscr{N}(0,1),$$

where $\mathscr{N}(0,1)$ *denotes the standard normal random variable.*

Our proof of Theorem 3 relies on the contraction method, which has found fruitful applications to recursively defined random variables in the last three decades; see Neininger and Rüschendorf [31] and Section 4.

Results of a very similar nature for the internal profile are given in Section 5.

These new results for the internal and external profile have many consequences in view of their close connections to other shape parameters. We content ourselves here with an application to the height H_n of DSTs, which is related to $B_{n,k}$ by $H_n := \max\{k : B_{n,k} > 0\}$; see Section 6 for other consequences.

Theorem 4. *Define* k_H *as follows*

$$k_H = \left\lfloor \log_2 n + \sqrt{2\log_2 n} - \frac{1}{2}\log_2\log_2 n + \frac{1}{\log 2} \right\rfloor. \tag{14}$$

which is at the upper boundary of the central range (13). Then the distribution of H_n *is concentrated at the two points* k_H *and* $k_H + 1$*:*

$$\mathbb{P}(H_n = k_H \text{ or } H_n = k_H + 1) \longrightarrow 1, \qquad (n \to \infty). \tag{15}$$

The possibility that such a result might hold was mentioned in [1] for a closely related model; a heuristic derivation was given in [20]. See also [2, 3] for other two-point approximation results in probability theory.

It is interesting to compare (15) with known results for the height of tries and those for Patricia tries, which we summarize in Table 2; see Flajolet [12] for the height of symmetric tries, and Knessl and Szpankowski [21] for that of Patricia tries (with only non-rigorous proofs).

Trees	Expected height	Discrete concentration	References
Tries	$2\log_2 n + \mathcal{O}(1)$	no	[12]
Patricia tries	$\log_2 n + \sqrt{2\log_2 n} + \mathcal{O}(1)$	at 3 pts	[21] (non-rigorous)
DSTs	$\log_2 n + \sqrt{2\log_2 n}$ $-\frac{1}{2}\log_2\log_2 n + \mathcal{O}(1)$	at 2 pts	this paper

Table 2: A comparison of the height of random symmetric tries, Patricia tries and DSTs.

We describe briefly the methods and tools used in proving Theorems 1–3, which all start with the following distributional recurrence

$$B_{n,k} \stackrel{d}{=} B_{J_n,k-1} + B^*_{n-1-J_n,k-1}, \qquad (n,k \geqslant 1), \tag{16}$$

with the boundary conditions $B_{0,0} = 1$, $B_{0,k} = 0$ for $k \geqslant 1$, $B_{n,0} = 0$ for $n \geqslant 1$, where $J_n = \text{Binomial}\big(n-1, \frac{1}{2}\big)$, and $B^*_{n,k}$ is an independent copy of $B_{n,k}$.

To derive the asymptotic approximations for the mean (Theorem 1) and the variance (Theorem 2), we rely on the property, in view of (16), that all moments of $B_{n,k}$ satisfy recurrences of the following type

$$a_{n,k} = 2^{2-n} \sum_{0\leqslant j<n} \binom{n-1}{j} a_{j,k-1} + b_{n,k} \tag{17}$$

for some given sequence $b_{n,k}$. This recurrence looks standard but the complication here comes from the dependence of k on n. When k is small, more precisely, when $2^{-k}n \to \infty$, the tree shape at these levels has little variation and thus both mean and variance can be treated by simple elementary arguments. The hard ranges are when $2^{-k}n \asymp 1$ and $2^{-k}n \to 0$ for which our arguments are built upon the idea of *Poissonization* by defining the *Poisson generating functions*

$$\tilde{A}_k(z) = e^{-z}\sum_{n\geqslant 0} a_{n,k}\frac{z^n}{n!} \quad \text{and} \quad \tilde{B}_k(z) = e^{-z}\sum_{n\geqslant 0} b_{n,k}\frac{z^n}{n!}.$$

Then (17) is translated into the differential-functional equation

$$\tilde{A}_k(z) + \tilde{A}'_k(z) = 2\tilde{A}_{k-1}\big(\tfrac{1}{2}z\big) + \tilde{B}_k(z),$$

which amounts to describing the moments in the *Poisson model*. This equation will be solved via Laplace transform techniques, which lead to exact and asymptotic expressions whose asymptotic properties will be further examined via Mellin transform, saddle-point method and again Laplace transform. Finally, we will translate the results in the Poisson model to those in the Bernoulli model via de-Poissonization.

While these procedures are by now standard (see [15, 16]) and work well for the mean, the analysis of the variance is more subtle. Here, the most crucial step is to introduce a *Poissonized variance* in the Poisson model (see again [15, 16]) so as to provide an asymptotic equivalent to the variance after de-Poissonization. An appropriate adaptation in the current situation is to define the function

$$\tilde{V}_k(z) := \tilde{M}_{k,2}(z) - \tilde{M}_{k,1}(z)^2 - z\tilde{M}'_{k,1}(z)^2,$$

where $\tilde{M}_{k,2}(z)$ and $\tilde{M}_{k,1}(z)$ denote the Poisson generating functions of the second moment and the first moment of $B_{n,k}$, respectively. Then we show that $\tilde{V}_k(z)$ is well-approximated by $2^k G\big(2^{-k}z\big)$ for large

$|z|$, and that $\tilde{V}_k(n)$ is asymptotically equivalent to $\mathrm{Var}(B_{n,k})$. Once these are clarified, the next challenge is the asymptotic behaviors of $G(z)$, notably for small $|z|$, which turns out to be the most technical part of this paper (see Proposition 3) largely due to the complicated form of the Laplace transform of $G(z)$ (see (43) and (44)) and the uniformity for large parameters (see Lemma 8).

In addition to the asymptotics of the mean and the variance, we also prove the central limit theorem from the distributional recurrence of $B_{n,k}$ via the contraction method, which is built on recurrences and the corresponding asymptotic transfer. Finally, Theorem 4 and related properties will be proved in Section 6 by using the results from Sections 2–5 and the first and second moment method. The corresponding asymptotic estimates for the internal profiles will be given in Section 5.

An extended abstract of this paper (entitled *External Profile of Symmetric Digital Search Trees*) by the same authors has appeared in the *Proceedings of the Fourteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO17)*, and contains Theorems 1–4 and sketches of the proofs of the first two. The current paper provides the proofs and derives additionally the same types of asymptotic approximations to the internal profile and discusses some of their consequences.

2 Expected Values of the External Profile

In this section, we prove Theorem 1 for $\mathbb{E}(B_{n,k})$. As mentioned in the Introduction, most results given here are known. Nevertheless, we provide detailed proofs because the analysis of the variance will follow the same pattern.

We start from (16). Write $\mu_{n,k} = \mathbb{E}(B_{n,k})$. Then

$$\mu_{n,k} = 2^{2-n} \sum_{0 \leqslant j < n} \binom{n-1}{j} \mu_{j,k-1}, \qquad (n, k \geqslant 1),$$

with the boundary conditions $\mu_{0,0} = 1$, $\mu_{n,0} = 0$ for $n \geqslant 1$, and $\mu_{0,k} = 0$ for $k \geqslant 1$. We then consider the Poisson generating function

$$\tilde{M}_{k,1}(z) := e^{-z} \sum_{n \geqslant 0} \mu_{n,k} \frac{z^n}{n!}, \qquad (k \geqslant 0),$$

which satisfies the differential-functional equation

$$\tilde{M}_{k,1}(z) + \tilde{M}'_{k,1}(z) = 2\tilde{M}_{k-1,1}\big(\tfrac{1}{2}z\big), \qquad (k \geqslant 1), \tag{18}$$

with $\tilde{M}_{0,1}(z) = e^{-z}$.

We now solve this differential-functional equation using Laplace transform, which, by inverting and taking coefficients, leads to an exact expression for $\mu_{n,k}$.

2.1 Exact Expressions

To solve (18), we apply Laplace transform (subsequently denoted by $\mathscr{L}[\cdot; s]$) on both sides of (18), and obtain

$$\mathscr{L}[\tilde{M}_{k,1}(z); s] = \frac{4}{s+1} \mathscr{L}[\tilde{M}_{k-1,1}(z); 2s], \qquad (k \geqslant 1),$$

with $\mathscr{L}[\tilde{M}_{0,1}(z); s] = \frac{1}{s+1}$. A direct iteration then yields

$$\mathscr{L}[\tilde{M}_{k,1}(z); s] = \frac{4^k}{(s+1)(2s+1)\cdots(2^k s+1)},$$

for $k \geqslant 0$. By partial fraction expansion, we see that

$$\mathscr{L}[\tilde{M}_{k,1}(z); s] = 2^k \sum_{0 \leqslant j \leqslant k} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_{k-j}} \cdot \frac{1}{s + 2^{j-k}}, \tag{19}$$

which, by term-by-term inversion, gives

$$\tilde{M}_{k,1}(z) = 2^k \sum_{0 \leqslant j \leqslant k} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_{k-j}} e^{-2^{j-k} z}, \qquad (k \geqslant 0). \tag{20}$$

From this, we obtain the closed-form expression for the expected profile (first derived in [25])

$$\mu_{n,k} = 2^k \sum_{0 \leqslant j \leqslant k} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_{k-j}} \left(1 - 2^{j-k}\right)^n. \tag{21}$$

We now examine the asymptotic aspects.

2.2 Asymptotics of $\mu_{n,k}$

If $2^{-k} n \to \infty$, then an expansion for the mean can be derived by elementary arguments because the term in (21) with $j = 0$ is dominating. More precisely, we have

$$\mu_{n,k} = \frac{2^k}{Q_k} \left(1 - 2^{-k}\right)^n \left(1 + \mathcal{O}\left(\frac{\left(1 - 2^{1-k}\right)^n}{\left(1 - 2^{-k}\right)^n}\right)\right), \tag{22}$$

where the error term is bounded above by

$$\frac{\left(1 - 2^{1-k}\right)^n}{\left(1 - 2^{-k}\right)^n} = \exp\left(-n \sum_{\ell \geqslant 1} \frac{2^\ell - 1}{\ell 2^{k\ell}}\right) \leqslant \exp\left(-\frac{n}{2^k - 1}\right) \qquad (k \geqslant 1). \tag{23}$$

Substituting this into (22) proves the asymptotic estimate (4) for $\mu_{n,k}$ when $2^{-k} n \to \infty$ and $k \geqslant 0$.

If $2^{-k} n = \mathcal{O}(1)$, then no single term in (21) is dominating and $2^{-rk} n \to 0$ for $r \geqslant 2$, so we readily obtain, again by (21) (approximating $(1 - x)^n$ by e^{-xn} and by extending k to infinity), $\mu_{n,k} \sim 2^k F\left(2^{-k} n\right)$, but the asymptotics of F for small parameter remains unclear. We will use instead the Poissonization techniques (see [16, 17]) to derive the required asymptotic approximation; see Theorem 1.

We derive first a simple bound for $F(z)$ and its derivatives.

Lemma 1. *For $m \geqslant 0$ and $\Re(z) \geqslant 0$, the mth derivative of F satisfies the uniform bound*

$$\sup_{\Re(z) \geqslant 0} \left|F^{(m)}(z)\right| = \mathcal{O}\left(2^{\binom{m+1}{2}}\right). \tag{24}$$

Proof. By the definition (3)

$$F^{(m)}(z) = \sum_{j \geqslant 0} \frac{(-1)^{j+r} 2^{-\binom{j}{2}+jm}}{Q_j Q_\infty} e^{-2^j z} = \mathcal{O}\left(\sum_{j \geqslant 0} 2^{-\binom{j}{2}+jm}\right) = \mathcal{O}\left(2^{\binom{m+1}{2}}\right).$$

This proves the uniform bound (24). ∎

We then show that (20) can be brought into the following more useful form (both exact and asymptotic).

Lemma 2. *For $\Re(z) \geqslant 0$ and $k \geqslant 1$, the Poisson generating function $\tilde{M}_{k,1}(z)$ of the expected profile $\mu_{n,k}$ satisfies*

$$\tilde{M}_{k,1}(z) = 2^k \sum_{m\geqslant 0} \frac{2^{-\binom{m+1}{2}-km}}{Q_m} F^{(m)}\big(2^{-k}z\big),$$

where $F(z)$ is given in (3).

Proof. By Euler's identity (see [4, Corollary 2.2])

$$\sum_{j\geqslant 0} \frac{(-1)^j q^{\binom{j}{2}}}{(1-q)(1-q^2)\cdots(1-q^j)} z^j = \prod_{\ell\geqslant 0}\big(1-q^\ell z\big), \qquad (0<q<1),$$

we have

$$\frac{Q_\infty}{Q_{k-j}} = \prod_{\ell\geqslant 1}\big(1-2^{j-k-\ell}\big) = \sum_{m\geqslant 0} \frac{(-1)^m 2^{-\binom{m+1}{2}}}{Q_m} 2^{(j-k)m},$$

which is still valid for $j > k$ (in which case both sides are zero). Substituting the latter into (20) gives

$$\begin{aligned}
\tilde{M}_{k,1}(z) &= 2^k \sum_{0\leqslant j\leqslant k} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_{k-j}} e^{-2^{j-k}z} \\
&= 2^k \sum_{j\geqslant 0} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_\infty} \sum_{m\geqslant 0} \frac{(-1)^m 2^{-\binom{m+1}{2}+(j-k)m}}{Q_m} e^{-2^{j-k}z} \\
&= 2^k \sum_{m\geqslant 0} \frac{(-1)^m 2^{-\binom{m+1}{2}-km}}{Q_m Q_\infty} \sum_{j\geqslant 0} \frac{(-1)^j 2^{-\binom{j}{2}+jm}}{Q_j} e^{-2^{j-k}z},
\end{aligned}$$

where interchanging the sums is justified as in the proof of Lemma 1. This proves the lemma since the last series is equal to $(-1)^m Q_\infty F^{(m)}\big(2^{-k}z\big)$. ∎

From these two lemmas, we get

$$\tilde{M}_{k,1}(z) = 2^k F\big(2^{-k}z\big) + \mathcal{O}(1), \qquad (\Re(z) \geqslant 0), \tag{25}$$

which is the Poissonized version of (2).

The asymptotics of $\mu_{n,k}$ and that of $\tilde{M}_k(n)$ can be bridged by the analytic de-Poissonization techniques; see the survey paper [17]. For that purpose, it turns out that the use of JS-admissible functions, a notion introduced in [16], provides a more effective operational approach.

Throughout this paper, the generic symbols $\varepsilon, \varepsilon'$ always denote small positive quantities whose values are immaterial and not necessarily the same at each occurrence.

Definition 1 ([16]). *An entire function $\tilde{f}(z)$ is said to be JS-admissible, denoted by $\tilde{f}(z) \in \mathscr{JS}$, if the following two conditions hold for $|z| \geqslant 1$.*

(I) *There exists a constant $\alpha \in \mathbb{R}$ such that uniformly for $|\arg(z)| \leqslant \varepsilon$,*

$$\tilde{f}(z) = \mathcal{O}\big(|z|^\alpha\big).$$

(O) *Uniformly for* $\varepsilon \leqslant |\arg(z)| \leqslant \pi$,

$$f(z) := e^z \tilde{f}(z) = \mathcal{O}\big(e^{(1-\varepsilon')|z|}\big).$$

When $\tilde{f} \in \mathscr{JS}$, its coefficients can be expressed in terms of the Poisson-Charlier expansion (see [16])

$$n![z^n]e^z \tilde{f}(z) = \sum_{j \geqslant 0} \frac{\tilde{f}^{(j)}(n)}{j!} \tau_j(n), \tag{26}$$

which is not only an identity but also an asymptotic expansion, where the $\tau_j(n)$'s are essentially Charlier polynomials defined by

$$\tau_j(n) = n![t^n]e^t(t-n)^j = \sum_{0 \leqslant \ell \leqslant j} \binom{j}{\ell}(-n)^{j-\ell}\frac{n!}{(n-\ell)!}, \qquad (j = 0, 1, \dots).$$

In particular, $\tau_j(n)$ is a polynomial in n of degree $\lfloor \frac{1}{2} j \rfloor$; the expressions for $\tau_j(n)$, $0 \leqslant j \leqslant 5$, are given below.

$\tau_0(n)$	$\tau_1(n)$	$\tau_2(n)$	$\tau_3(n)$	$\tau_4(n)$	$\tau_5(n)$
1	0	$-n$	$2n$	$3n(n-2)$	$-4n(5n-6)$

For our purpose, we also need an additional uniformity property for JS-admissible functions as the level parameter k may also depend on n.

Lemma 3. *The functions $\tilde{M}_{k,1}(z)$ are uniformly JS-admissible, i.e, if $|z| \geqslant 1$, then for $|\arg(z)| \leqslant \varepsilon$*

$$\tilde{M}_{k,1}(z) = \mathcal{O}(|z|)$$

and for $\varepsilon \leqslant |\arg(z)| \leqslant \pi$

$$e^z \tilde{M}_{k,1}(z) = \mathcal{O}\big(e^{(1-\varepsilon')|z|}\big), \tag{27}$$

where all implied constants are absolute and hold uniformly for $k \geqslant 0$.

Proof. Let $M_{k,1}(z) := e^z \tilde{M}_{k,1}(z)$. We first rewrite (18) into the following form

$$M_{k,1}(z) = \int_0^z 2e^{\frac{1}{2}u} M_{k-1,1}\big(\tfrac{1}{2}u\big)\,\mathrm{d}u = 2z \int_0^1 e^{\frac{1}{2}tz} M_{k-1,1}\big(\tfrac{1}{2}tz\big)\,\mathrm{d}t. \tag{28}$$

Consider first the region $|\arg(z)| \geqslant \varepsilon$. The bound (27) holds trivially for $k = 0$ since $M_{0,1}(z) = 1$. Thus, we assume $k \geqslant 1$. By the trivial bound $\mu_{n,k} \leqslant 2n$, we get the *a priori* upper estimate $|M_k(z)| \leqslant 2|z|e^{|z|}$. Plugging this into (28) yields

$$\begin{aligned}
|M_{k,1}(z)| &\leqslant 2|z|^2 \int_0^1 te^{\frac{1}{2}t(\Re(z)+|z|)}\,\mathrm{d}t \\
&\leqslant 2|z|^2 \int_0^1 e^{\frac{1}{2}t(\cos\varepsilon+1)|z|}\,\mathrm{d}t \\
&\leqslant \frac{4|z|}{\cos\varepsilon + 1}\, e^{\frac{1}{2}(\cos\varepsilon+1)|z|}.
\end{aligned}$$

Since $\frac{1}{2}(\cos\varepsilon + 1) < 1$, this proves (27).

Now we consider the sector $|\arg(z)| \leqslant \varepsilon$. The required bound $\tilde{M}_{k,1}(z) = \mathcal{O}(|z|)$ will follow from (25) and the smallness of $F(z)/z$ to be proved in Proposition 1 below (see also Remark 3). It is also possible to give a direct proof although with a weaker estimate (sufficient for our de-Poissonization purposes). By (28)

$$\tilde{M}_{k,1}(z) = 2z \int_0^1 e^{-(1-t)z} \tilde{M}_{k-1,1}\left(\tfrac{1}{2}tz\right) \mathrm{d}t,$$

and induction on k, we deduce the (slightly worse) bound $\tilde{M}_{k,1}(z) = \mathcal{O}(|z|^{1+\varepsilon'})$. ∎

By the asymptotic expansion (26) and Lemma 3 (which also gives bounds on the derivatives of $\tilde{M}_{k,1}(z)$ by Ritt's theorem; see [16]), we can justify the "Poisson heuristic" for $\mu_{n,k}$ as follows.

Lemma 4. *For large n and $0 \leqslant k \leqslant n$*

$$\mu_{n,k} = \tilde{M}_{k,1}(n) + \mathcal{O}(1),$$

where the $\mathcal{O}$-term holds uniformly in k.

From this and (25), we obtain now Theorem 1 (except of the positivity of $F(x)$, which will be established in Proposition 2 below).

2.3 Asymptotics of $F(z)$

In this subsection, we derive an asymptotic expansion for $F(z)$ for small $|z|$ and prove the positivity of $F(x)$ on $\mathbb{R}^+$; the corresponding large-$|z|$ asymptotics is much easier; see (5).

Proposition 1. *For each integer $m \geqslant 0$, the mth derivative of F satisfies*

$$F^{(m)}(z) = \frac{\rho^{m+\frac{1}{2}+\frac{1}{\log 2}}}{\sqrt{2\pi \log_2 \rho}} \exp\Big(-\frac{(\log \rho)^2}{2\log 2} - P(\log_2 \rho)\Big)\left(1 + \mathcal{O}\big(|\log \rho|^{-1}\big)\right), \tag{29}$$

as $|z| \to 0$ in the sector $|\arg(z)| \leqslant \varepsilon$, where $P(t)$ is given in (7) and $\rho = \rho(z)$ solves the saddle-point equation

$$\frac{\rho}{\log \rho} = \frac{1}{z \log 2},$$

satisfying $|\rho| \to \infty$ as $|z| \to 0$.

Proof. By additivity, the Laplace transform of F has the form

$$\mathscr{L}[F(z); s] = \sum_{j \geqslant 0} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_\infty (s + 2^j)}, \qquad (\Re(s) > -1),$$

which equals the partial fraction expansion of the product

$$\mathscr{L}[F(z); s] = \prod_{j \geqslant 0} \frac{1}{1 + 2^{-j}s} = \frac{1}{Q(-2s)}. \tag{30}$$

Since we are interested in the asymptotics of $F(z)$ as $|z| \to 0$, which is reflected by the large-s asymptotics of $\mathscr{L}[F(z); s]$, we apply the Mellin transform techniques for that purpose; see Flajolet et al.'s survey paper [13] for more background tools and applications. In particular, taking logarithm on both sides of (30) and using the inverse Mellin transform gives

$$\log Q(-2s) = \sum_{j\geqslant 0} \log(1+2^{-j}s) = \frac{1}{2\pi i}\int_{-\frac12 - i\infty}^{-\frac12 + i\infty} \frac{\pi s^{-\omega}}{(1-2^{\omega})\omega \sin \pi\omega}\,\mathrm{d}\omega,$$

because the Mellin transform of $\log(1+s)$ equals

$$\int_0^\infty s^{\omega-1}\log(1+s)\,\mathrm{d}s = \frac{\pi}{\omega \sin \pi\omega}, \qquad (\Re(\omega)\in(-1,0)).$$

Now, by standard Mellin analysis (see [13]), we deduce that

$$\log Q(-2s) = \frac{(\log s)^2}{2\log 2} + \frac{\log s}{2} + P(\log_2 s) + \mathcal{O}\big(|s|^{-1}\big), \tag{31}$$

uniformly as $|s|\to\infty$ in the sector $|\arg s|\leqslant \pi-\varepsilon$.

Next, by the inverse Laplace transform, first for $z=r$ real, we have

$$F^{(m)}(r) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} \frac{s^m e^{rs}}{Q(-2s)}\,\mathrm{d}s. \tag{32}$$

It follows, by moving the line of integration to $\Re(s)=\rho$ and by substituting the asymptotic approximation (31), that

$$F^{(m)}(r) = \frac{1}{2\pi i}\int_{\rho-i\infty}^{\rho+i\infty} s^m \exp\left(rs - \frac{(\log s)^2}{2\log 2} - \frac{\log s}{2} - P(\log_2 s) + \mathcal{O}\big(|s|^{-1}\big)\right)\mathrm{d}s.$$

A standard application of the saddle-point method (see [14, Ch. VIII]) then yields (29) for real z with $z\to 0$.

When the imaginary part of z is not zero, we can still apply the same procedure but need to deform the integration contour in the representation (32) from the vertical line with real part 1 to the one where the portions from $1+i$ to $1+i\infty$ and $1-i\infty$ to $1-i$ are tilted slightly to the left; see the Appendix for details. ∎

Remark 3. As a consequence of the above proposition, we see that $F(z)$ is smaller than any polynomial of z as $|z|\to 0$ in the sector $|\arg(z)|\leqslant\varepsilon$.

Asymptotically, for large $X := \frac{1}{x\log 2}$, $x\in\mathbb{R}$,

$$\rho = X\left(\log X + \log\log X + \frac{\log\log X}{\log X} - \frac{(\log\log X)^2 - 2\log\log X}{2(\log X)^2} + \mathcal{O}\left(\frac{(\log\log X)^3}{(\log X)^3}\right)\right). \tag{33}$$

Substituting this into (29) gives the more explicit expression (6) (but with a worse error term).

Finally, we prove the positivity of F on the positive real line.

Proposition 2. *The function $F(x)$ is positive in $(0,\infty)$.*

Proof. Since $\mu_{n,k}\geqslant 0$, we see, by Theorem 1, that $F(x)\geqslant 0$ on $(0,\infty)$. Then, from (30),

$$(1+s)\mathscr{L}[F(x);s] = \prod_{j\geqslant 0}\frac{1}{1+2^{-j-1}s} = \mathscr{L}[F(x);\tfrac12 s].$$

The corresponding inverse Laplace transform yields the equation

$$F(x) + F'(x) = 2F(2x).$$

With this differential-functional equation, we prove the positivity of F by contradiction. Assume *a contrario* that $F(x)$ has a zero in $(0, \infty)$, say x_0. Then $F'(x_0) = 2F(2x_0) \geqslant 0$ and since $F'(x_0) > 0$ is not possible (for otherwise $F(x)$ would become negative in a neighborhood of x_0), we also have $F(2x_0) = 0$. Continuing this argument, we obtain arbitrarily large zeros. This is, however, impossible since we see from (5) that $F(x)$ is positive for all x large enough. ∎

3 The Variance of the External Profile

In this section, we prove Theorem 2 by the same approach used above for the mean, starting from the second moment $\nu_{n,k} := \mathbb{E}(B_{n,k}^2)$, which satisfies, by (16), the recurrence

$$\nu_{n,k} = 2^{2-n} \sum_{0 \leqslant j < n} \binom{n-1}{j} \nu_{j,k-1} + 2^{2-n} \sum_{0 \leqslant j < n} \binom{n-1}{j} \mu_{j,k-1} \mu_{n-1-j,k-1}, \qquad (n, k \geqslant 1),$$

with the boundary conditions $\nu_{0,0} = 1$, $\nu_{n,0} = 0$ for $n \geqslant 1$, and $\nu_{0,k} = 0$ for $k \geqslant 1$. Translating this recurrence into the corresponding Poisson generating functions

$$\tilde{M}_{k,2}(z) := e^{-z} \sum_{n \geqslant 0} \nu_{n,k} \frac{z^n}{n!}, \qquad (k \geqslant 0)$$

leads to the differential-functional equation

$$\tilde{M}_{k,2}(z) + \tilde{M}'_{k,2}(z) = 2\tilde{M}_{k-1,2}\left(\tfrac{1}{2}z\right) + 2\tilde{M}_{k-1,1}\left(\tfrac{1}{2}z\right)^2, \qquad (k \geqslant 1), \tag{34}$$

with $\tilde{M}_{0,2}(z) = e^{-z}$.

Since the variance is expected to be of the same order as the mean (notably when both tend to infinity), there is a cancellation between the dominant term in the asymptotic expansion for $\nu_{n,k}$ and that for $\mu_{n,k}^2$. Such a cancellation of dominant terms can be incorporated in the Poissonized variance (as in [15, 16]):

$$\tilde{V}_k(z) = \tilde{M}_{k,2}(z) - \tilde{M}_{k,1}(z)^2 - z\tilde{M}'_{k,1}(z)^2,$$

which itself also satisfies, after a straightforward calculation,

$$\tilde{V}_k(z) + \tilde{V}'_k(z) = 2\tilde{V}_{k-1}\left(\tfrac{1}{2}z\right) + z\tilde{M}''_{k,1}(z)^2, \qquad (k \geqslant 1), \tag{35}$$

with $\tilde{V}_0(z) = e^{-z} - (1+z)e^{-2z}$. In this form, the original inherent cancellation is nicely integrated into the same type of equation with an explicitly computable non-homogeneous function, and we need only to work out the asymptotics of $\tilde{V}_k(z)$, which will be proved to be asymptotically equivalent to the variance of $B_{n,k}$ in the major range of interest.

3.1 Exact Expressions

To justify the cancellation-free approach to computing the asymptotic variance, we still need more explicit expressions for $\tilde{M}_{k,2}(z)$ and $\tilde{V}_k(z)$. For that purpose, we apply Laplace transform on both sides of (34) and obtain

$$\mathscr{L}[\tilde{M}_{k,2}(z); s] = \frac{4}{s+1} \mathscr{L}[\tilde{M}_{k-1,2}(z); 2s] + \frac{4}{s+1} \mathscr{L}[\tilde{M}_{k-1,1}(z)^2; 2s], \qquad (k \geqslant 1),$$

with $\mathscr{L}[\tilde{M}_{0,2}(z);s] = \frac{1}{s+1}$. Iterating the recurrence gives

$$\mathscr{L}[\tilde{M}_{k,2}(z);s] = \frac{4^k}{(s+1)\cdots(2^k s+1)} + \sum_{0\leqslant j<k} \frac{4^{k-j}\mathscr{L}[\tilde{M}_{j,1}(z)^2; 2^{k-j}s]}{(s+1)\cdots(2^{k-j-1}s+1)}. \tag{36}$$

From (20), a manageable expression for the Laplace transform of $\tilde{M}_{k-j,1}(z)^2$ is given by

$$\mathscr{L}[\tilde{M}_{j,1}(z)^2;s] = 4^j \sum_{0\leqslant h,\ell\leqslant j} \frac{(-1)^{h+\ell}2^{-\binom{h}{2}-\binom{\ell}{2}}}{Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \cdot \frac{1}{s+2^{h-j}+2^{\ell-j}}.$$

This and the partial fraction expansion (19) yield

$$\begin{aligned}\sum_{0\leqslant j<k} & \frac{4^{k-j}\mathscr{L}[\tilde{M}_{j,1}(z)^2; 2^{k-j}s]}{(s+1)\cdots(2^{k-j-1}s+1)} \\ &= \sum_{(j,r,h,\ell)\in\mathscr{S}} \frac{2^{2j+1}(-1)^{r+h+\ell}2^{-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}}}{Q_r Q_{k-1-j-r} Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \cdot \frac{1}{(s+2^{r+1-k+j})(s+2^{h-k}+2^{\ell-k})},\end{aligned}$$

where

$$\mathscr{S} = \{(j,r,h,\ell)\ :\ 0\leqslant j\leqslant k-1, 0\leqslant r\leqslant k-1-j, 0\leqslant h,\ell\leqslant j\}.$$

From this and the expression

$$\frac{1}{(s+u)(s+v)} = \frac{1}{v-u}\left(\frac{1}{s+u} - \frac{1}{s+v}\right), \qquad (u\neq v),$$

we obtain, by term-by-term inversion,

$$\tilde{M}_{k,2}(z) = \tilde{M}_{k,1}(z) + \sum_{(j,r,h,\ell)\in\mathscr{S}} \frac{2^{2j+1}(-1)^{r+h+\ell}2^{-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}}}{Q_r Q_{k-1-j-r} Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \phi(2^{r+1-k+j}, 2^{h-k}+2^{\ell-k}; z),$$

where

$$\phi(u,v;z) = e^{-vz}\int_0^z e^{-(u-v)t}\,\mathrm{d}t = \begin{cases} \dfrac{e^{-uz}-e^{-vz}}{v-u}, & \text{if } u\neq v; \\ ze^{-uz}, & \text{if } u=v. \end{cases}$$

Taking the coefficients of z^n on both sides, we are led to the exact expression for $\nu_{n,k}$:

$$\nu_{n,k} = \mu_{n,k} + \sum_{(j,r,h,\ell)\in\mathscr{S}} \frac{2^{2j+1}(-1)^{r+h+\ell}2^{-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}}}{Q_r Q_{k-1-j-r} Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \delta(2^{r+1-k+j}, 2^{h-k}+2^{\ell-k}; n), \tag{37}$$

where

$$\begin{aligned}\delta(u,v;n) &= n\int_0^1 (1-u-(v-u)t)^{n-1}\,\mathrm{d}t \\ &= \begin{cases} \dfrac{(1-u)^n-(1-v)^n}{v-u}, & \text{if } u\neq v; \\ n(1-u)^{n-1}, & \text{if } u=v. \end{cases}\end{aligned}$$

Similarly, by (35) and the same procedure, we also have

$$\tilde{V}_k(z) = \sum_{(j,r,h,\ell)\in\mathscr{V}} \frac{2^{k-j}(-1)^{r+h+\ell}2^{-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}+2h+2\ell}}{Q_r Q_{k-j-r} Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \varphi\big(2^{r+j}, 2^h+2^\ell; 2^{-k}z\big), \tag{38}$$

where

$$\mathscr{V} = \{(j,r,h,\ell) \ : \ 0 \leqslant j \leqslant k, 0 \leqslant r \leqslant k-j, 0 \leqslant h,\ell \leqslant j\},$$

and $\varphi(u,v;z)$ is defined in (10). Note that the equality $2^{r+j} = 2^h + 2^\ell$ occurs if and only if (j,r,h,ℓ) belongs to the set

$$\{(j,r,h,\ell) \ : \ 1 \leqslant j \leqslant k, r=0, h=\ell=j-1 \quad \text{or} \quad 0 \leqslant j < k, r=1, h=\ell=j\},$$

and the corresponding terms in (38) are

$$\sum_{0<j\leqslant k} \frac{2^{-k-j}2^{-2\binom{j-1}{2}+4(j-1)}}{Q_{k-j}Q_1^2Q_{j-1}^2} \cdot \frac{z^2}{2}e^{-2^{j-k}z} - \sum_{0\leqslant j<k} \frac{2^{-k-j}2^{-2\binom{j}{2}+4j}}{Q_{k-j-1}Q_1Q_j^2} \cdot \frac{z^2}{2}e^{-2^{j+1-k}z},$$

which become zero since $Q_1 = \frac{1}{2}$. Hence, the equality part in the definition (10) of $\varphi(u,v;z)$ may be ignored.

3.2 Asymptotics of $\mathrm{Var}(B_{n,k})$

The range where $2^k n e^{-2^{-k}n} \to 0$ can be treated elementarily, as in the case of the mean. In this range of k (and even in the wider range where $2^{-k}n \to \infty$), we have

$$\mathrm{Var}(B_{n,k}) \sim \frac{2^k}{Q_k}(1-2^{-k})^n, \tag{39}$$

uniformly in k. To prove this, we use (37) and begin with the estimate

$$\delta(2^{r+1-k+j}, 2^{h-k}+2^{\ell-k}; n) = \mathcal{O}\big(n\big(1-2^{1-k}\big)^n\big),$$

where the implied constant is absolute in n and in k. Substituting this into (37) yields

$$\begin{aligned} \nu_{n,k} &= \mu_{n,k} + \mathcal{O}\Bigg(n\big(1-2^{1-k}\big)^n \sum_{(j,r,h,\ell)\in\mathscr{S}} 2^{2j}2^{-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}}\Bigg) \\ &= \mu_{n,k} + \mathcal{O}\big(n4^k\big(1-2^{1-k}\big)^n\big). \end{aligned}$$

By the asymptotic estimate (4) for $\mu_{n,k}$, we then have

$$\nu_{n,k} = \frac{2^k}{Q_k}\big(1-2^{-k}\big)^n\Bigg(1+\mathcal{O}\big(e^{-\frac{n}{2^k-1}}\big)+\mathcal{O}\bigg(n2^k\frac{\big(1-2^{1-k}\big)^n}{\big(1-2^{-k}\big)^n}\bigg)\Bigg).$$

From this estimate and (23), we see that

$$\nu_{n,k} \sim \frac{2^k}{Q_k}\big(1-2^{-k}\big)^n \sim \mu_{n,k},$$

because $2^k n e^{-2^{-k}n} \to 0$. Also, $\mu_{n,k} = o(1)$ in this range. Thus, we get $\mu_{n,k}^2 = o(\mu_{n,k})$ and then (39). Note that it is possible to extend slightly the range to $2^k e^{-2^{-k}n} \to 0$ because there is only one term containing the factor $n(1-2^{1-k})^n$ in the sum (37) (which is when $j=r=h=l=0$), and the contribution of all other terms is bounded above by $\mathcal{O}(4^k(1-2^{1-k})^n)$. Moreover, that (39) holds in the wider range $2^{-k}n \to \infty$ follows from a refinement of the expansion of Theorem 2, which can in turn be obtained by the analytic method below.

On the other hand, complex analytic tools apply in a wider range. In contrast to the mean, however, we do not prove an identity for $\tilde{V}_k(z)$ (compare with Lemma 2 and see Remark 4) but we directly prove an asymptotic result similar to the one for the mean in (25).

Lemma 5. *For $|\arg(z)| \leqslant \varepsilon$ and $k \geqslant 0$, $\tilde{V}_k(z)$ satisfies the expansion*

$$\tilde{V}_k(z) = 2^k G\big(2^{-k} z\big) + \mathcal{O}(1), \tag{40}$$

where $G(z)$ is defined in Theorem 2.

Proof. Similar to Lemma 2, we first consider Q_∞ / Q_{k-j-r} which satisfies the uniform bound

$$\frac{Q_\infty}{Q_{k-j-r}} = \prod_{\ell \geqslant 1}\big(1 - 2^{j+r-k-\ell}\big) = 1 + \mathcal{O}\big(2^{j+r-k}\big). \tag{41}$$

Here the product also makes sense for $j + r > k$ where it becomes zero and thus the bound also holds in this case. Substituting this into (38) gives

$$\begin{aligned}\tilde{V}_k(z) &= 2^k G\big(2^{-k} z\big) \\ &\quad + \sum_{j,r \geqslant 0} \sum_{0 \leqslant h,\ell \leqslant j} \frac{(-1)^{r+h+\ell} 2^{-j-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}+2h+2\ell}}{Q_\infty Q_r Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \mathcal{O}\big(2^{j+r} \varphi\big(2^{r+j}, 2^h + 2^\ell; 2^{-k} z\big)\big).\end{aligned}$$

To estimate the double sum, we split the summation range into two: (*i*) $h, \ell \leqslant \lfloor j/2 \rfloor$ and (*ii*) either $h > \lfloor j/2 \rfloor$ or $\ell > \lfloor j/2 \rfloor$, and denote the resulting sums $E_1(z)$ and $E_2(z)$, respectively. By the estimates

$$2^{j+r} \varphi\big(2^{r+j}, 2^h + 2^\ell; 2^{-k} z\big) = \begin{cases} \mathcal{O}(1), & \text{if } h, \ell \leqslant \lfloor j/2 \rfloor, \\ \mathcal{O}(2^{j+r}), & \text{if } h > \lfloor j/2 \rfloor \text{ or } \ell > \lfloor j/2 \rfloor, \end{cases}$$

where the implied constants are both absolut, we have

$$E_1(z) = \mathcal{O}\Bigg(\sum_{j,r \geqslant 0} \sum_{0 \leqslant h,\ell \leqslant j/2} 2^{-j-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}+2h+2\ell}\Bigg) = \mathcal{O}(1),$$

and

$$E_2(z) = \mathcal{O}\left(\sum_{j,r \geqslant 0} \sum_{\substack{0 \leqslant h,\ell \leqslant j \\ h > j/2 \text{ or } \ell > j/2}} 2^{-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}+r+2h+2\ell}\right) = \mathcal{O}\Bigg(\sum_{j \geqslant 1} j^{-1} 2^{-\frac{1}{8}j^2+\frac{5}{4}j}\Bigg) = \mathcal{O}(1).$$

This proves the claimed expansion. ∎

Remark 4. Comparing with Lemma 2, it would be natural to derive an identity for $\tilde{V}_k(z)$ in a way similar to that for $\tilde{M}_{k,1}(z)$ by replacing the first order asymptotics (41) by the full expansion

$$\frac{Q_\infty}{Q_{k-j-r}} = \prod_{\ell \geqslant 1}\big(1 - 2^{j+r-k-\ell}\big) = \sum_{m \geqslant 0} \frac{(-1)^m 2^{-\binom{m+1}{2}}}{Q_m} 2^{(j+r-k)m},$$

which is zero for $j + r > k$. However, doing so yields an expression that is no more absolutely convergent, as pointed out by one referee. Nevertheless, ignoring the convergence issue and carrying out all computations formally, one can expand $\tilde{V}_k(z)$ as

$$\tilde{V}_k(z) = 2^k \sum_{m \geqslant 0} \frac{2^{-\binom{m+1}{2}-mk}}{Q_m} H_m(2^{-k} z), \tag{42}$$

where $H_m(z)$ are suitable functions. Then, a formal calculation of the Laplace transform gives (after a lengthy computation)

$$\mathscr{L}[H_m(z);s] = s^m \sum_{j\geqslant 0} 4^{-j} \frac{\tilde{g}_j^*(2^{-j}s)}{Q(-2^{1-j}s)}, \tag{43}$$

where $\tilde{g}_j^*(s) = \mathscr{L}[z(\tilde{M}''_{j,1}(z))^2;s]$. Similarly,

$$\mathscr{L}\left[G^{(m)}(z);s\right] = s^m \sum_{j\geqslant 0} 4^{-j} \frac{\tilde{g}_j^*(2^{-j}s)}{Q(-2^{1-j}s)}, \tag{44}$$

where this relation indeed holds not just formally but also in a rigorous analytic sense since the series representation of $G^{(m)}(z)$ does converge absolutely for all $m \geqslant 0$ (this can be proved by bounding the derivatives of $\varphi(u,v;x)$). Thus, (43) and (44) suggest that

$$H_m(z) = G^{(m)}(z), \qquad (m \geqslant 0)$$

which in turn suggests that the following identity

$$\tilde{V}_k(z) = 2^k \sum_{m\geqslant 0} \frac{2^{-\binom{m+1}{2}-km}}{Q_m} G^{(m)}(2^{-k}z).$$

This identity, if true, would be the variance analogue of the identity in Lemma 2. This identity was claimed to hold at the end of Section 1 in the conference version of this paper. However, it is not rigorously proved. This shows the intricacy of the analysis for the variance.

Note that (40) gives the version of (8) under the Poisson model. From this we will deduce now Theorem 8 by the same approaches used to prove Theorem 1, namely, de-Poissonization techniques through the use of JS-admissible functions.

Lemma 6. *The functions $\tilde{M}_{k,2}(z)$ are uniformly JS-admissible. More precisely, if $|z| \geqslant 1$, then for $|\arg(z)| \leqslant \varepsilon$*

$$\tilde{M}_{k,2}(z) = \mathcal{O}(|z|^2),$$

and for $\varepsilon \leqslant |\arg(z)| \leqslant \pi$

$$e^z \tilde{M}_{k,2}(z) = \mathcal{O}(e^{(1-\varepsilon')|z|}),$$

where all implied constants in the $\mathcal{O}$-terms are absolute for $k \geqslant 0$.

Proof. We proved in [16, Prop. 2.4] that if $\tilde{g}(z)$ is JS-admissible, then $\tilde{f}(z)$ with

$$\tilde{f}(z) + \tilde{f}'(z) = 2\tilde{f}\left(\tfrac{1}{2}z\right) + \tilde{g}(z),$$

is also JS-admissible. Since $2\tilde{M}_{k-1,1}\left(\frac{1}{2}z\right)^2$ is uniformly JS-admissible (by Lemma 3), the same property holds for $\tilde{M}_{k,2}(z)$ by the same proof of [16, Prop. 2.4]. Alternatively, the bound for $\tilde{M}_{k,2}(z)$ can be derived from (40) and properties of $G(z)$ and $\tilde{M}_{k,1}(z)$. ∎

We are now ready to prove Theorem 2. Since $\tilde{M}_{k,2} \in \mathscr{JS}$, we have the expansion

$$v_{n,k} = \sum_{0\leqslant j\leqslant 3} \frac{\tilde{M}_{k,2}^{(j)}(n)}{j!} \tau_j(n) + \mathcal{O}(\tilde{M}_{k,2}(n)n^{-2}),$$

where we retain the terms from (26) with $j \geqslant 3$ so as to guarantee that the error term is $\mathcal{O}(1)$ (since $\tilde{M}_{k,2}(n) = \mathcal{O}(n^2)$). For $\mu_{n,k}$, we need an expansion with an error up to $\mathcal{O}(n^{-1})$:

$$\mu_{n,k} = \sum_{0 \leqslant j \leqslant 3} \frac{\tilde{M}_{k,1}^{(j)}(n)}{j!} \tau_j(n) + \mathcal{O}\big(\tilde{M}_{k,1}(n) n^{-2}\big),$$

so that $\mu_{n,k}^2$ is correct up to an error of order $\mathcal{O}(1)$. Then we obtain

$$\mathrm{Var}(B_{n,k}) = \nu_{n,k} - \mu_{n,k}^2 = \tilde{V}_k(n) + \mathcal{O}(1),$$

where we have used the relation $\tilde{M}_{k,2}(n) = \tilde{V}_k(n) + \tilde{M}_{k,1}(n)^2 + n\tilde{M}_{k,1}'(n)^2$. By (40), this proves the approximation in Theorem 2.

3.3 Asymptotics of $G(z)$

We now derive the asymptotic behaviors of G for small and large $|z|$, and prove that $G(x)$ is positive for $x \in (0, \infty)$. In particular, the asymptotic approximations of G will imply (11).

First, the asymptotics of $G(z)$ for large z follows directly from the defining series (9)

$$G(z) = \frac{e^{-z}}{Q_\infty} + \mathcal{O}\big(|z| e^{-2\Re(z)}\big),$$

for $\Re(z) > 0$. In contrast the small-$|z|$ asymptotics of G turns out to be very involved, which we now examine.

Proposition 3. *For each integer $m \geqslant 0$, $G^{(m)}$ satisfies the asymptotic estimate*

$$G^{(m)}(z) \sim 2F^{(m)}(z),$$

as $|z| \to 0$ in the sector $|\arg(z)| \leqslant \varepsilon$.

The proof of this proposition is long and technical and relies mostly on Laplace transform. Note that since the Laplace transform of $G^{(m)}$ is just s^m times the Laplace transform of G, it will be sufficient to consider only the case $m = 0$.

We start from the Laplace transform of $G(z)$, which by (43) and (44), is given by

$$\mathscr{L}[G(z); s] = \sum_{j \geqslant 0} R_j(s), \quad \text{where} \quad R_j(s) := \frac{\tilde{g}_j^*(2^{-j}s)}{4^j Q\big(-2^{1-j}s\big)}. \tag{45}$$

Here $\tilde{g}_j^*(s) = \mathscr{L}[z(\tilde{M}_{j,1}''(z))^2; s]$, which, by (20) and a straightforward computation, has the form

$$\tilde{g}_j^*(s) = 4^{-j} \sum_{0 \leqslant h,\ell \leqslant j} \frac{(-1)^{h+\ell} 2^{-\binom{h}{2}-\binom{\ell}{2}+2h+2\ell}}{Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \cdot \frac{1}{(s + 2^{h-j} + 2^{\ell-j})^2}.$$

Strangely, the dominating term in (45) is

$$R_2(s) \sim \frac{2}{Q(-2s)}$$

for large $|s|$, and the hard part of the analysis consists in showing that $\sum_{j \neq 2} R_j(s) = \mathcal{O}(|R_2(s)/s|)$.

Lemma 7. *For large $|s|$ in the half-plane $\Re(s) > 0$, R_0 and R_1 satisfy*

$$R_0(s) = \frac{s^{-2}}{Q(-2s)}(1 + \mathcal{O}(|s|^{-1})) \quad \textit{and} \quad R_1(s) = \frac{9\, s^{-1}}{Q(-2s)}(1 + \mathcal{O}(|s|^{-1})), \tag{46}$$

respectively, and R_j with fixed $j \geqslant 2$ satisfies

$$R_j(s) = \frac{(2j-3)!2^{\binom{j}{2}}}{((j-2)!)^2 Q(-2s)}\, s^{2-j}\big(1 + \mathcal{O}\big(|s|^{-1}\big)\big). \tag{47}$$

Proof. The estimates (46) for $j = 0$ and $j = 1$ follow from the closed-form expressions

$$\tilde{g}_0^*(s) = \frac{1}{(s+2)^2} \quad \text{and} \quad \tilde{g}_1^*(2^{-1}s) = \frac{4}{(s+2)^2} - \frac{32}{(s+3)^2} + \frac{64}{(s+4)^2},$$

and the functional relation $Q(-2s) = (1+s)Q(-s)$. Thus we assume now $j \geqslant 2$.

Since $\tilde{g}_j^*\big(2^{-j}s\big)$ is the Laplace transform of $4^j z \tilde{M}''_{j,1}(2^j z)^2$, we see that the large-$|s|$ behavior of the former is reflected from the small-$|z|$ behavior of the latter. Starting from (19) using Ritt's theorem for the asymptotics of the derivatives of an analytic function, we obtain successively the estimates in the following table.

$f(z)$	$\tilde{M}_{j,1}(z)$	$\tilde{M}''_{j,1}(z)$	$4^j z \tilde{M}''_{j,1}(2^j z)^2$
as $\lvert z\rvert \sim 0$	$\frac{z^j}{j!2^{j(j-3)/2}}$	$\frac{z^{j-2}}{(j-2)!2^{-j(j-3)/2}}$	$\frac{2^{j(j+1)}}{((j-2)!)^2}\, z^{2j-3}$
$\mathscr{L}[f(z); s]$	$\frac{4^j}{\prod_{0\leqslant \ell \leqslant j}(2^\ell s+1)}$	$\frac{4^j s^2}{\prod_{0\leqslant \ell \leqslant j}(2^\ell s+1)}$	$\tilde{g}_j^*\big(2^{-j}s\big)$
as $\lvert s\rvert \to \infty$	$2^{-j(j-3)/2}\, s^{-j-1}$	$2^{-j(j-3)/2}\, s^{-j+1}$	$\frac{(2j-3)!}{((j-2)!)^2}\, 2^{j(j+1)} s^{2-2j}$

where the entries in the second and the fourth rows give the asymptotics of f and its Laplace transform as $|z| \to 0$ and $|s| \to \infty$, respectively. All error terms are of the form $1 + \mathcal{O}(|z|)$ and $1 + \mathcal{O}(|s|^{-1})$, respectively,

From this table and the estimate

$$\begin{aligned} Q(-2^{1-j}s) &= \frac{Q(-2s)}{(1+s)(1+2^{-1}s)\cdots(1+2^{-(j-1)}s)} \\ &= s^{-j} Q(-2s) 2^{\binom{j}{2}}(1 + \mathcal{O}(|s|^{-1})), \end{aligned} \tag{48}$$

for large $|s|$, we obtain (47). ∎

We now enhance the asymptotic approximation (47) by incorporating the uniformity in j.

Lemma 8. *For $j \geqslant 2$, uniformly as $|s| \to \infty$ with $\Re(s) \geqslant \varepsilon$,*

$$\tilde{g}_j^*\big(2^{-j}s\big) = \begin{cases} \mathcal{O}\left(\frac{j}{\log_2 |s|}, 2^{-(\log_2 |s|)(\log_2 |s| - 5)}\right), & \textit{if } 1 \leqslant |s| \leqslant 2^{j+1}, \\ \mathcal{O}\left(2^{j^2+3j}|s|^{-2j+2}\right), & \textit{if } |s| \geqslant 2^{j+1}. \end{cases} \tag{49}$$

Proof. For notational convenience, we write the Laplace transform of h as h^* and the convolution as

$$(h_1^* \star h_2^*)(s) := \frac{1}{2\pi i}\int_{\frac{1}{2}s - i\infty}^{\frac{1}{2}s + i\infty} h_1^*(t) h_2^*(s-t)\,\mathrm{d}t.$$

Since $\tilde{g}_j^*(2^{-j}s) = \mathscr{L}[z\tilde{M}''_{j,1}(2^j z)^2; s]$, we see that, by the relation $\mathscr{L}[zh(z); s] = -h^*(s)'$,

$$\tilde{g}_j^*(2^{-j}s) = -(L_j \star L_j)'(s) = -(L_j \star L_j')(s),$$

where

$$L_j(s) := \frac{s^2}{\prod_{0\leqslant\ell\leqslant j}(1+2^{-\ell}s)}. \tag{50}$$

On the other hand, since

$$L_j'(s) = L_j(s)\left(\frac{2}{s} - \sum_{0\leqslant\ell\leqslant j}\frac{1}{s+2^\ell}\right),$$

we now derive an upper bound for each of the convolutions

$$L_j(s) \star \left(\frac{L_j(s)}{s}\right) \qquad \text{and} \qquad L_j(s) \star \left(\frac{L_j(s)}{s+2^\ell}\right), \quad 0 \leqslant \ell \leqslant j.$$

Note that if both $h_1^*(s), h_2^*(s) = \mathcal{O}(|s|^{-1})$ for large $|s|$, then we can replace $(h_1^* \star h_2^*)(s)$ (for $\Re(s) > 0$) by the integral

$$\frac{1}{2\pi i}\int_{\gamma(s)} h_1^*(t)h_2^*(s-t)\,\mathrm{d}t,$$

where $\gamma(s) := \left\{\frac{1}{2}s\,(1+iv) : -\infty < v < \infty\right\}$ is the symmetry line between 0 and s; such a choice implies $|t| = |s-t|$ for $t \in \gamma(s)$ and $|L_j(t)| = |L_j(s-t)|$. This simplifies our analysis.

By (50), it is also straightforward to show that for all s with $\Re(s) > 0$ and for all $t \in \gamma(s)$,

$$L_j(t) = \begin{cases} \mathcal{O}\left(|t|^3|s|^{-1}2^{-\frac{1}{2}(\log_2|t|)(\log_2|t|+1)}\right); & \text{for } 1 \leqslant |t| \leqslant 2^j, \\ \mathcal{O}\left(2^{j(j+1)/2}|t|^{-j+1}\right), & \text{for } |t| \geqslant 2^j. \end{cases} \tag{51}$$

Now, we are ready to prove (49). Assume first that $|s| \geqslant 2^{j+1}$. Then $|t| \geqslant 2^j$ for all $t \in \gamma(s)$. Furthermore, we have $|t+y| \geqslant |s|/2$ for all $t \in \gamma(s)$ and for all non-negative numbers y. Consequently, by (51),

$$\begin{aligned}\left|L_j(s) \star \left(\frac{L_j(s)}{s+y}\right)\right| &\leqslant \frac{2}{|s|}\int_{\gamma(s)} |L_j(t)|^2\,|\mathrm{d}t| \\ &= \mathcal{O}\left(\frac{2^{j(j+1)}}{|s|}\left(\frac{2}{|s|}\right)^{2j-2}|s|\int_0^\infty \frac{1}{(1+v^2)^{j-1}}\,\mathrm{d}v\right) \\ &= \mathcal{O}\left(\frac{2^{j^2+3j}}{j\,|s|^{2j-2}}\right),\end{aligned}$$

where we used the substitution $t = \frac{1}{2}s\,(1+iv)$. Since $\tilde{g}_j^*(2^{-j}s)$ can be written as the sum of $j+1$ terms of this form, (49) follows (in the case $|s| \geqslant 2^{j+1}$).

If $|s| \leqslant 2^{j+1}$, then we split the integral into two parts. The first one is by (51) bounded above by

$$\begin{aligned}
\frac{2}{|s|}\int_{t\in\gamma(s),\ |t|\leqslant 2^j} |L_j(t)|^2\,|\mathrm{d}t| &= \mathcal{O}\left(\frac{1}{|s|^3}\int_{t\in\gamma(s),\ |t|\leqslant 2^j} |t|^6 2^{-(\log_2|t|)(\log_2|t|+1)}\,|\mathrm{d}t|\right)\\
&= \mathcal{O}\left(\frac{1}{|s|^3}\int_{t\in\gamma(s),\ |t|\leqslant 2^j} |t|^6 2^{-(\log_2|t|)(\log_2|s|)}\,|\mathrm{d}t|\right)\\
&= \mathcal{O}\left(\frac{1}{|s|^3}\int_{t\in\gamma(s),\ |t|\leqslant 2^j} |t|^{6-\log_2|s|}\,|\mathrm{d}t|\right)\\
&= \mathcal{O}\left(|s|^4\left(\frac{|s|}{2}\right)^{-\log_2|s|}\int_0^\infty |1+iv|^{6-\log|s|}\,\mathrm{d}v\right)\\
&= \mathcal{O}\left(\frac{2^{-\log_2^2|s|+5\log_2|s|}}{\log|s|}\right),
\end{aligned}$$

where we used the inequality $|t| \geqslant \frac{1}{2}|s|$ and the substitution $t = \frac{1}{2}s\,(1+iv)$. Finally, the remaining integral is bounded above by

$$\begin{aligned}
\frac{2}{|s|}\int_{t\in\gamma(s),\ |t|\geqslant 2^j} |L_j(t)|^2\,|\mathrm{d}t| &= \mathcal{O}\left(2^{j(j+1)}\left(\frac{2}{|s|}\right)^{2j-2}\int_{\frac{2^j}{|s|}}^\infty \frac{1}{(1+v^2)^{j-1}}\,\mathrm{d}v\right)\\
&= \mathcal{O}\left(2^{j(j+1)}\left(\frac{2}{|s|}\right)^{2j-2}\frac{1}{j}\left(\frac{|s|}{2^j}\right)^{2j-5}\right)\\
&= \mathcal{O}\left(\frac{2^{-j^2+8j}}{j\,|s|^3}\right).
\end{aligned}$$

Since $|s| \leqslant 2^{j+1}$, we then have

$$\frac{2^{-j^2+8j}}{j\,|s|^3} = \mathcal{O}\left(\frac{2^{-\log_2^2|s|+5\log_2|s|}}{\log_2|s|}\right),$$

and accordingly

$$\tilde{g}_j^*(2^{-j}s) = \mathcal{O}\left(\frac{j}{\log_2|s|}\,2^{-\log_2^2|s|+5\log_2|s|}\right),$$

which completes the proof of the lemma. ∎

We now derive a precise asymptotics for $\mathscr{L}[G(z);s]$ for large $|s|$.

Lemma 9. *The Laplace transform of G satisfies*

$$\mathscr{L}[G(z);s] = \frac{2}{Q(-2s)}\bigl(1+\mathcal{O}(|s|^{-1})\bigr),$$

uniformly as $|s|\to\infty$ and $\Re(s)\geqslant\varepsilon$.

Proof. By Lemma 7, we have ($R_j(s)$ being defined in (45))

$$R_0(s)+R_1(s)+R_2(s) = \frac{2}{Q(-2s)}\bigl(1+\mathcal{O}(|s|^{-1})\bigr).$$

For the remaining terms, we examine the factor $Q(-2^{1-j}s)$. By (48), we have, uniformly for $|s|\to\infty$ with $\Re(s)>0$ and $|s|\geqslant 1$,

$$Q(-2^{1-j}s) = \begin{cases}\Omega(1), & \text{if } 1\leqslant|s|\leqslant 2^{j+1};\\ \Omega\left(|s|^{-j}|Q(-2s)|2^{\binom{j}{2}}\right), & \text{if } |s|\geqslant 2^{j+1}.\end{cases}$$

Now, it follows from Lemma 8 that

$$\sum_{3\leqslant j\leqslant \log_2|s|-1} R_j(s) = \mathcal{O}\left(\frac{1}{|Q(-2s)|}\sum_{3\leqslant j\leqslant \log_2|s|-1} 2^{\frac{1}{2}j(j+3)-(j-2)\log_2|s|}\right)$$
$$= \mathcal{O}\left(\frac{1}{|s\,Q(-2s)|}\right);$$

also by Lemma 8 and (31)

$$\sum_{j\geqslant \log_2|s|} R_j(s) = \mathcal{O}\left(\sum_{j\geqslant\log_2|s|}\frac{j}{\log_2|s|}2^{-(\log_2|s|)(\log_2|s|-5)-2j}\right)$$
$$= \mathcal{O}\big(2^{-(\log_2|s|)(\log_2|s|-3)}\big)$$
$$= \mathcal{O}\left(\frac{2^{-(\log_2|s|)(\log_2|s|-7)/2}}{|Q(-2s)|}\right)$$
$$= \mathcal{O}\left(\frac{1}{|s\,Q(-2s)|}\right),$$

This completes the proof of the lemma. ∎

Proof of Proposition 3. Proposition 3 now follows from Lemma 9 because $\frac{1}{Q(-2s)}$ is the Laplace transform of $F(z)$, details being similar to the proof of Theorem 1 (for the asymptotics of $F(z)$). ∎

Finally, we prove that $G(x)$ is a positive function.

Proposition 4. *The function $G(x)$ is positive on* $(0,\infty)$.

Proof. By (45) and the inverse Laplace transform, we see that

$$G(x) = \sum_{j\geqslant 0} 2^j \int_0^x (x-t)\tilde{M}''_{j,1}(2^j(x-t))^2 F(2^j t)\,\mathrm{d}t.$$

Since $F(x)$ is positive on $(0,\infty)$, we then deduce that $G(x)$ is also positive on $(0,\infty)$. ∎

4 Asymptotic Normality of $B_{n,k}$

We prove in this section Theorem 3, the central limit theorem for $B_{n,k}$, by the contraction method [31].

Recall that $\mu_{n,k} = \mathbb{E}(B_{n,k})$. Let $\sigma_{n,k} := \sqrt{\mathrm{Var}(B_{n,k})}$ for all $n,k\geqslant 0$. We consider the standardized random variables when $\sigma_{n,k} > 0$

$$X_{n,k} := \frac{B_{n,k}-\mu_{n,k}}{\sigma_{n,k}}, \tag{52}$$

and $X_{n,k} := 0$ otherwise. From (16), we obtain

$$X_{n,k} \overset{d}{=} \frac{\sigma_{J_n,k-1}}{\sigma_{n,k}} X_{J_n,k-1} + \frac{\sigma_{n-1-J_n,k-1}}{\sigma_{n,k}} X^*_{n-1-J_n,k-1} + b_{n,k}(J_n), \tag{53}$$

where all $X_{j,k}$, $X^*_{j,k}$ and $J_n = \mathrm{Binomial}\big(n-1,\frac{1}{2}\big)$ are independent, $X^*_{j,k}$ are distributed as $X_{j,k}$, and

$$b_{n,k}(j) := \frac{\mu_{j,k-1}+\mu_{n-1-j,k-1}-\mu_{n,k}}{\sigma_{n,k}}. \tag{54}$$

Our goal is to prove that $X_{n,k}$ converges in distribution to $\mathcal{N} := N(0,1)$ as $\sigma_{n,k} \to \infty$. This is achieved by showing that in the limit (53) becomes

$$X \stackrel{d}{=} \frac{1}{\sqrt{2}} X + \frac{1}{\sqrt{2}} X^*, \tag{55}$$

which is satisfied by the normal distribution. Thus, apart from the convergence of $X_{n,k}$ to X, we will show that the coefficients of the first two terms on the right hand side of (53) both converge to $\frac{1}{\sqrt{2}}$, and the third term there tends to 0 as $\sigma_{n,k} \to \infty$. In addition to these convergences, we derive (stronger) uniform bounds in the next two lemmas. In what follows, we use $\|X\|_3$ to denote the L_3 norm of a random variable X.

Lemma 10. *The L_3 norm of $b_{n,k}(J_n)$ is bounded above by*

$$\|b_{n,k}(J_n)\|_3 = \mathcal{O}\big(\sigma_{n,k}^{-1}\big). \tag{56}$$

Proof. Define the set $A_n = \big\{|J_n - \frac{1}{2}n| < n^{\frac{3}{4}}\big\}$ and denote by $\mathbf{1}_{A_n}$ the indicator function of A_n.

We first consider $b_{n,k}(J_n)$ on the complement A_n^c, where we have

$$\|b_{n,k}(J_n)\mathbf{1}_{A_n^c}\|_3 = \frac{\mathcal{O}(ne^{-2\sqrt{n}})}{\sigma_{n,k}} = \mathcal{O}\big(\sigma_{n,k}^{-1}\big).$$

Here we used the linear upper bound of $\mu_{n,k}$ (either by definition or by Theorem 1) and Chernoff's bound for binomial tails

$$\sum_{|j-\frac{1}{2}n|\geqslant n^{\frac{3}{4}}} 2^{1-n}\binom{n-1}{j} = \mathcal{O}\big(n^{-\frac{1}{4}}e^{-2\sqrt{n}}\big). \tag{57}$$

It remains to estimate $\|b_{n,k}(J_n)\mathbf{1}_{A_n}\|_3$. By Theorem 1 and Taylor series expansion,

$$\begin{aligned}
&\mu_{J_n,k-1} + \mu_{n-1-J_n,k-1} - \mu_{n,k} \\
&\qquad = 2^{k-1}F\big(2^{1-k}J_n\big) + 2^{k-1}F\big(2^{1-k}(n-1-J_n)\big) - 2^kF\big(2^{-k}n\big) + \mathcal{O}(1) \\
&\qquad = \mathcal{O}\bigg(1 + F'(2^{-k}n) + 4^{-k}\big(J_n - \tfrac{1}{2}(n-1)\big)^2 \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}}\big|F''\big(2^{1-k}x\big)\big|\bigg).
\end{aligned}$$

Thus

$$\|b_{n,k}(J_n)\mathbf{1}_{A_n}\|_3 = \sigma_{n,k}^{-1} \times \mathcal{O}\bigg(1 + 4^{-k}n \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}}\big|F''\big(2^{1-k}x\big)\big|\bigg).$$

Now by the asymptotic approximations (5) and (6) to $F(x)$, we see that

$$4^{-k}n \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}}\big|F''\big(2^{1-k}x\big)\big| = \begin{cases} \mathcal{O}\big(4^{-k}ne^{-(1-\varepsilon)2^{-k}n}\big), & \text{if } 2^{-k}n \to \infty; \\ \mathcal{O}\big(4^{-k}nF\big(2^{-k}n\big)\big), & \text{if } 2^{-k}n = \Theta(1); \\ \mathcal{O}\big(n^{-1}k^2F\big(2^{-k}n\big)\big), & \text{if } 2^{-k}n \to 0. \end{cases}$$

It follows that

$$4^{-k}n \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}}\big|F''\big(2^{1-k}x\big)\big| = \mathcal{O}(1),$$

and this completes the proof of (56). ∎

Lemma 11. *We have the uniform bound*

$$\left\| \frac{\sigma_{J_n,k-1}}{\sigma_{n,k}} - \frac{1}{\sqrt{2}} \right\|_3 = \mathcal{O}\big(\sigma_{n,k}^{-1}\big). \tag{58}$$

Proof. Note that (58) clearly holds if $\sigma_{n,k}$ is bounded since then, $\sigma_{J_n,k-1}$ is also bounded. We thus assume in the sequel that $\sigma_{n,k} \to \infty$ which holds for k with (13). Also, note that

$$\left\| \frac{\sigma_{J_n,k-1}}{\sigma_{n,k}} - \frac{1}{\sqrt{2}} \right\|_3 = \left\| \frac{\sigma_{J_n,k-1}^2 - \frac{1}{2}\sigma_{n,k}^2}{\sigma_{n,k}\left(\sigma_{J_n,k-1} + \frac{1}{\sqrt{2}}\sigma_{n,k}\right)} \right\|_3 = \mathcal{O}\left(\frac{\left\|\sigma_{J_n,k-1}^2 - \frac{1}{2}\sigma_{n,k}^2\right\|_3}{\sigma_{n,k}^2} \right).$$

To prove (58) from the last expression, we apply similar arguments as in the proof of Lemma 10 whose notations we again adopt.

First, on the set A_n^c the estimate (58) follows from the same exponential tail bounds for binomial distributions and $\sigma_{n,k}^2 = \mathcal{O}(n)$ (as in Lemma 10). On the other hand, on the set A_n we use Theorem 3 and the Taylor series expansion

$$\begin{aligned}\sigma_{J_n,k-1}^2 &= 2^{k-1}G\big(2^{1-k}J_n\big) + \mathcal{O}(1)\\ &= 2^{k-1}G\big(2^{-k}n\big) + \mathcal{O}\left(1 + G'(2^{-k}n) + \big(J_n - \tfrac{1}{2}n\big) \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}} \left|G'\big(2^{1-k}x\big)\right|\right).\end{aligned}$$

Thus

$$\left\|\sigma_{J_n,k-1}^2 - \tfrac{1}{2}\sigma_{n,k}^2\right\|_3 = \mathcal{O}\left(1 + \sqrt{n} \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}} \left|G'\big(2^{1-k}x\big)\right|\right).$$

Next, by the growth properties of $G(x)$ from Section 3.3, we see that

$$\sqrt{n} \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}} \left|G'\big(2^{1-k}x\big)\right| = \mathcal{O}\left(\sqrt{2^k G\big(2^{-k}n\big)}\right),$$

because

$$\sqrt{n} \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}} \left|G'\big(2^{1-k}x\big)\right| = \begin{cases} \mathcal{O}\left(2^{k/2}\sqrt{2^{-k/2}n}\, e^{-(1-\varepsilon)2^{-k}n}\right), & \text{if } 2^{-k}n \to \infty;\\ \mathcal{O}\left(\sqrt{n}G\big(2^{-k}n\big)\right), & \text{if } 2^{-k}n = \Theta(1);\\ \mathcal{O}\left(k\sqrt{2^k G\big(2^{-k}n\big)/n}\right), & \text{if } 2^{-k}n \to 0.\end{cases}$$

This proves (58). ∎

We now justify the convergence in distribution of (52) to standard normal through the use of the Zolotarev ζ_3-distance introduced in [35, 36]. Note that the fixed-point equation (55) appears frequently in the analysis of recursive algorithms and random trees; Section 5.3 of [31] gives about a dozen such examples. Most applications of the contraction method in the literature rely on the minimal L_2-metric. However, recurrences leading to the limit equation (55) are beyond the power of the minimal L_p-metrics as explained in detail in Section 2 of [31]. This deficiency of the minimal L_p-metrics was initially the main motivation to develop the Zolotarev ζ_s-distance in the context of the contraction method in [31]. That (55) on an appropriate subspace of probability measures endowed with the ζ_3-distance constitutes a contraction is also explained in details in Section 2 of [31]. This fact is the key for the present proof of the central limit law in Theorem 3.

We recall a few properties needed.

Zolotarev metric. The Zolotarev ζ_3-distance of two distributions $\mathcal{L}(Y)$, $\mathcal{L}(Z)$ on $\mathbb{R}$ is defined by

$$\zeta_3(Y,Z) := \zeta_3(\mathcal{L}(Y),\mathcal{L}(Z)) := \sup_{f\in\mathcal{F}_3} |\mathbb{E}(f(Y)-f(Z))|, \tag{59}$$

where

$$\mathcal{F}_3 := \{f \in C^2(\mathbb{R},\mathbb{R}) : |f^{(2)}(x) - f^{(2)}(y)| \leqslant |x-y|\},$$

denotes the space of twice continuously differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ such that the second derivative is Lipschitz continuous with Lipschitz constant 1. It is known that $\zeta_3(Y,Z) < \infty$ if $\mathbb{E}(Y) = \mathbb{E}(Z)$, $\mathbb{E}(Y^2) = \mathbb{E}(Z^2)$ and $\|Y\|_3, \|Z\|_3 < \infty$. The convergence in ζ_3 implies weak convergence on $\mathbb{R}$. We need that ζ_3 is $(3,+)$-ideal, namely,

$$\zeta_3(Y+W, Z+W) \leqslant \zeta_3(Y,Z) \tag{60}$$

and

$$\zeta_3(cY, cZ) = c^3\zeta_3(Y,Z), \tag{61}$$

for all W independent of (Y,Z) and all $c > 0$. Note that this implies that

$$\zeta_3\,(Y_1+Y_2, Z_1+Z_2) \leqslant \zeta_3(Y_1,Z_1) + \zeta_3(Y_2,Z_2), \tag{62}$$

for Y_1, Y_2 independent and Z_1, Z_2 independent such that the respective ζ_3 distances are finite.

The following bound is also used (see [9, Lemma 5.7] or [30, Lemma 2.1] for a slightly tighter bound):

$$\zeta_3(\mathcal{L}(Y),\mathcal{L}(Z)) \leqslant (\|Y\|_3^2 + \|Z\|_3^2)\|Y-Z\|_3, \tag{63}$$

where on the right-hand side the pair (Y,Z) forms any coupling of the distributions $\mathcal{L}(Y), \mathcal{L}(Z)$ on a joint probability space.

The ζ_3-distance is part of the family ζ_s, $s > 0$, of Zolotarev metrics.

Before we prove Theorem 3 with this metric, we need another crucial technical lemma.

Lemma 12 ($\mathcal{O}$- and o-transfer for (17))**.** *Let $\alpha > 1$ and consider the recurrence (17).*

(a) If $b_{n,k} = \mathcal{O}\left(2^{\alpha k}G\left(2^{-k}n\right)^{\alpha} + 1\right)$, then $a_{n,k} = \mathcal{O}\left(2^{\alpha k}G\left(2^{-k}n\right)^{\alpha}\right)$ for any sequence $k = k(n)$ such that $\sigma_{n,k} \to \infty$;

(b) If $b_{n,k} = o\left(2^{\alpha k}G\left(2^{-k}n\right)^{\alpha} + 1\right)$, then $a_{n,k} = o\left(2^{\alpha k}G\left(2^{-k}n\right)^{\alpha}\right)$ for any sequence $k = k(n)$ such that $\sigma_{n,k} \to \infty$.

See (13) for the range where $\sigma_{n,k} \to \infty$.

Proof. We start with part *(a)*, and prove by induction on k the uniform bound

$$|a_{n,k}| \leqslant C_0 2^{k\alpha}G\left(2^{-k}n\right)^{\alpha} + D_0 n^{1+\varepsilon}, \tag{64}$$

for some constants $C_0, D_0 > 0$ (specified below) and an arbitrarily small $\varepsilon > 0$. For $k = 0$ the bound (64) holds. Hence, we assume that (64) holds for all $k' < k$. To prove it for k, we substitute (64) into the recurrence (17), and obtain

$$|a_{n,k}| \leqslant 2^{2-n} \sum_{0\leqslant j<n} \binom{n-1}{j} \left(C_0 2^{(k-1)\alpha}G\left(2^{1-k}j\right)^{\alpha} + D_0 j^{1+\varepsilon}\right) + C_1 2^{k\alpha}G\left(2^{-k}n\right)^{\alpha} + D_1, \tag{65}$$

for some constants C_1 and D_1. As in the proof of Lemma 10, we split the above binomial sum into two parts according as $|j - \frac{1}{2}n| < n^{\frac{3}{4}}$ or $|j - \frac{1}{2}n| \geqslant n^{\frac{3}{4}}$. The sum over the latter range is, by Chernoff's bound (57), bounded above by

$$2^{2-n} \sum_{|j-\frac{1}{2}n|\geqslant n^{\frac{3}{4}}} \binom{n-1}{j} \left(C_0 2^{(k-1)\alpha} G\big(2^{1-k} j\big)^{\alpha} + D_0 j^{1+\varepsilon}\right) = \mathcal{O}\left(n^{\alpha} e^{-2\sqrt{n}}\right) = o(1),$$

because $\sigma_{n,k}^2$ is at most of linear growth. For the remaining sum with $|j - \frac{1}{2}n| < n^{\frac{3}{4}}$, we split it further into two parts, one for each summand inside the parentheses of the first term in (65). The second part is independent of k and can again be easily estimated by binomial concentration properties:

$$2D_0 \sum_{|j-\frac{1}{2}n|<n^{\frac{3}{4}}} 2^{1-n}\binom{n-1}{j} j^{1+\varepsilon} = 2^{-\varepsilon} D_0 n^{1+\varepsilon}(1+o(1)).$$

We are left with the sum

$$2^{1-\alpha} C_0 2^{k\alpha} \sum_{|j-\frac{1}{2}n|<n^{\frac{3}{4}}} 2^{1-n}\binom{n-1}{j} G\big(2^{1-k} j\big)^{\alpha}. \tag{66}$$

By Taylor expansion, we have

$$G\big(2^{1-k} j\big)^{\alpha} = G\big(2^{-k} n\big)^{\alpha} + \alpha G\big(2^{-k} n\big)^{\alpha-1} G'\big(2^{-k} n\big) \frac{2j-n}{2^k} + \mathcal{O}\left(E(n)\frac{(2j-n)^2}{4^k}\right)$$

with

$$E(n) := \max_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}} \left(\big|G\big(2^{1-k}x\big)\big|^{\alpha-2}\big|G'\big(2^{1-k}x\big)\big|^2 + \big|G\big(2^{1-k}x\big)\big|^{\alpha-1}\big|G''\big(2^{1-k}x\big)\big|\right).$$

Substituting this into (66), we deduce that

$$\begin{aligned} C_0 2^{1-\alpha} 2^{k\alpha} &\sum_{|x-\frac{1}{2}n|<n^{\frac{3}{4}}} 2^{1-n}\binom{n-1}{j} G\big(2^{1-k} j\big)^{\alpha} \\ &= C_0 2^{1-\alpha} 2^{k\alpha} G\big(2^{-k} n\big)^{\alpha}(1+o(1)) + o(1) + \mathcal{O}\left(4^{-k} 2^{k\alpha} n E(n)\right). \end{aligned} \tag{67}$$

Now, by the properties of $G(x)$ from Section 3.3,

$$4^{-k} n E(n) = \begin{cases} \mathcal{O}\Big(4^{-k} n e^{-\alpha 2^{-k} n + \mathcal{O}(2^{-k} n^{\frac{3}{4}})}\Big), & \text{if } 2^{-k} n \to \infty; \\ \mathcal{O}\big(4^{-k} n G\big(2^{-k} n\big)^{\alpha}\big), & \text{if } 2^{-k} n = \Theta(1); \\ \mathcal{O}\big(k^2 n^{-1} G\big(2^{-k} n\big)^{\alpha}\big), & \text{if } 2^{-k} n \to 0. \end{cases}$$

Thus, $\mathcal{O}(4^{-k} 2^{k\alpha} n E(n)) = o\big(1 + 2^{k\alpha} G\big(2^{-k} n\big)^{\alpha}\big)$, which shows that the last term in (67) can be dropped.

Collecting all estimates and substituting them into (65), we get

$$|a_{n,k}| \leqslant C_0 2^{1-\alpha} 2^{k\alpha} G\big(2^{-k} n\big)^{\alpha}(1+o(1)) + D_0 2^{-\varepsilon} n^{1+\varepsilon}(1+o(1)) + C_1 2^{k\alpha} G\big(2^{-k} n\big)^{\alpha} + D_1.$$

Since $2^{1-\alpha} < 1$ and $2^{-\varepsilon} < 1$, it is clear that one can choose C_0 and D_0 such that the latter is bounded by $C_0 2^{k\alpha} G\big(2^{-k} n\big)^{\alpha} + D_0 n^{1+\varepsilon}$. This proves the uniform bound (64).

Using the uniform bound (64), the assertion of part *(a)* can now be obtained by another induction on those $k = k(n)$ such that $\sigma_{n,k} \to \infty$, where we can use the same arguments as above since now D_1 in (65) can be dropped and thus the second term in the uniform bound (64) is not needed anymore.

Finally, part *(b)* follows *mutatis mutandis* the same method of proof. This proves the lemma. ∎

Proof of Theorem 3. Denote by

$$\mathbf{1}_{n,k} := \begin{cases} 1; & \text{if } \sigma_{n,k} > 0, \\ 0, & \text{otherwise.} \end{cases}$$

To bound the ζ_3-distance between $X_{n,k}$ and $\mathcal{N}(0,1)$ we define

$$\Xi_{n,k} := \frac{\sigma_{J_n,k-1}}{\sigma_{n,k}}\mathcal{N} + \frac{\sigma_{n-1-J_n,k-1}}{\sigma_{n,k}}\mathcal{N}^* + b_{n,k}(J_n), \tag{68}$$

if $\sigma_{n,k} > 0$ and $\Xi_{n,k} := 0$ otherwise, where $\mathcal{N}, \mathcal{N}^*$ and J_n are independent and $\mathcal{N}$ and $\mathcal{N}^*$ have the standard normal distribution. Conditioning on J_n and using (53), we obtain, for all $n, k \geqslant 0$,

$$\operatorname{Var}(\Xi_{n,k}) = \operatorname{Var}(X_{n,k}) = \operatorname{Var}(\mathbf{1}_{n,k}\mathcal{N}) = \mathbf{1}_{n,k}.$$

Since all the random variables in this display are centered and have finite third moments, we see that the ζ_3-distances between them are finite. Thus

$$\zeta_3(X_{n,k}, \mathbf{1}_{n,k}\mathcal{N}) \leqslant \zeta_3(X_{n,k}, \Xi_{n,k}) + \zeta_3(\Xi_{n,k}, \mathbf{1}_{n,k}\mathcal{N}).$$

We first bound the second term on the right-hand side. Note that Lemma 10 and 11 imply that

$$\|\Xi_{n,k}\|_3 = \mathcal{O}\big(\sigma_{n,k}^{-1} + 1\big).$$

Furthermore, we have the identity

$$\mathcal{N} \stackrel{d}{=} \tfrac{1}{\sqrt{2}}\mathcal{N} + \tfrac{1}{\sqrt{2}}\mathcal{N}^*, \tag{69}$$

where $\mathcal{N}$ and $\mathcal{N}^*$ are as above; compare with (55). Now, using the representations (68) and (69) and the inequality (63), we have

$$\begin{aligned}
\zeta_3(\Xi_{n,k}, \mathbf{1}_{n,k}\mathcal{N}) &= \zeta_3(\Xi_{n,k}, \mathcal{N}) \\
&= \mathcal{O}\left(\big(\sigma_{n,k}^{-2} + 1\big)\left\| \frac{\sigma_{J_n,k-1}}{\sigma_{n,k}}\mathcal{N} + \frac{\sigma_{n-1-J_n,k-1}}{\sigma_{n,k}}\mathcal{N}^* \right.\right. \\
&\qquad\qquad\left.\left. + b_{n,k}(J_n) - \left(\tfrac{1}{\sqrt{2}}\mathcal{N} + \tfrac{1}{\sqrt{2}}\mathcal{N}^*\right)\right\|_3\right) \\
&= \mathcal{O}\left(\big(\sigma_{n,k}^{-2} + 1\big)\left(\left\|\frac{\sigma_{J_n,k-1}}{\sigma_{n,k}} - \frac{1}{\sqrt{2}}\right\|_3 + \|b_{n,k}(J_n)\|_3\right)\right).
\end{aligned}$$

Consequently, by Lemma 10 and 11,

$$\zeta_3(\Xi_{n,k}, \mathbf{1}_{n,k}\mathcal{N}) = \mathcal{O}\big(\sigma_{n,k}^{-3} + \sigma_{n,k}^{-1}\big).$$

Thus,

$$\Delta(n,k) := \zeta_3(X_{n,k}, \mathbf{1}_{n,k}\mathcal{N}) \leqslant \zeta_3(X_{n,k}, \Xi_{n,k}) + \mathcal{O}\big(\sigma_{n,k}^{-3} + \sigma_{n,k}^{-1}\big).$$

We are left with the distance $\zeta_3(X_{n,k}, \Xi_{n,k})$. Using the definition (59) and conditioning on J_n, we see that

$$\begin{aligned}
\zeta_3(X_{n,k}, \Xi_{n,k}) \leqslant 2^{1-n} \sum_{0 \leqslant j < n} \binom{n-1}{j} \zeta_3\Bigg(&\frac{\sigma_{j,k-1}}{\sigma_{n,k}} X_{j,k-1} + \frac{\sigma_{n-1-j,k-1}}{\sigma_{n,k}} X^*_{n-1-j,k-1} + b_{n,k}(j), \\
&\frac{\sigma_{j,k-1}}{\sigma_{n,k}}\mathcal{N} + \frac{\sigma_{n-1-j,k-1}}{\sigma_{n,k}}\mathcal{N}^* + b_{n,k}(j)\Bigg).
\end{aligned}$$

We then majorize the latter expression by using the $+$-ideal properties of ζ_3 in (60), (61) and drop the $b_{n,k}(j)$ terms. This gives, by (62) and again the 3-ideal properties of ζ_3 in (60), (61),

$$\begin{aligned}\zeta_3(X_{n,k}, \Xi_{n,k}) &\leqslant 2^{1-n} \sum_{0\leqslant j<n} \binom{n-1}{j} \Bigg(\left(\frac{\sigma_{j,k-1}}{\sigma_{n,k}}\right)^3 \zeta_3(X_{j,k-1}, \mathbf{1}_{j,k-1}\mathcal{N}) \\ &\qquad + \left(\frac{\sigma_{n-1-j,k-1}}{\sigma_{n,k}}\right)^3 \zeta_3(X^*_{n-1-j,k-1}, \mathbf{1}_{n-1-j,k-1}\mathcal{N}^*)\Bigg) \\ &= 2^{-n} \sum_{0\leqslant j<n} \binom{n-1}{j} \left(\frac{\sigma_{j,k-1}}{\sigma_{n,k}}\right)^3 \Delta(j,k-1).\end{aligned}$$

Note that the corresponding summand is zero when $\mathbf{1}_{j,k-1} = 0$ or $\mathbf{1}_{n-1-j,k-1} = 0$. Collecting the estimates yields

$$\Delta(n,k) \leqslant 2^{-n} \sum_{0\leqslant j<n} \binom{n-1}{j} \left(\frac{\sigma_{j,k-1}}{\sigma_{n,k}}\right)^3 \Delta(j,k-1) + \mathcal{O}\big(\sigma_{n,k}^{-3} + \sigma_{n,k}^{-1}\big).$$

From this, we see that $\sigma_{n,k}^3 \Delta(n,k)$ is bounded above by a sequence $a_{n,k}$ satisfying (17) with $b_{n,k} = \mathcal{O}(\sigma_{n,k}^2 + 1)$. Thus

$$\sigma_{n,k}^3 \Delta(n,k) = o\left(\sigma_{n,k}^{2+\varepsilon}\right),$$

whenever $\sigma_{n,k} \to \infty$ by Lemma 12, part *(a)* (with $\alpha = 1+\varepsilon$ for any $\varepsilon > 0$). Thus $\Delta(n,k) \to 0$ as $\sigma_{n,k} \to \infty$, and this completes the proof of Theorem 3. ∎

5 Internal Profile

In this section, we present the results without proofs for the internal profile $I_{n,k}$ because all proofs used for the external profile extend to those for the internal profile, which satisfies the same form of recurrence as $B_{n,k}$, namely,

$$I_{n,k} \stackrel{d}{=} I_{J_n,k-1} + I^*_{n-1-J_n,k-1}, \qquad (n,k \geqslant 1),$$

where the notation is as in (16). The only differences lie in the boundary conditions: $I_{0,k} = 0$ for $k \geqslant 0$ and $I_{n,0} = 1$ for $n \geqslant 1$.

From this recurrence and the same method used in Section 2, we can derive the following (mostly known) result for the mean; see [10, 25].

Theorem 5. *The expected internal profile satisfies*

$$\mathbb{E}(I_{n,k}) = 2^k F_I\big(2^{-k}n\big) + \mathcal{O}(1),$$

uniformly for $0 \leqslant k < n$, *where* $F_I(x)$ *equals the antiderivative of* $F(x)$*:*

$$F_I(x) = 1 - \sum_{j\geqslant 0} \frac{(-1)^j 2^{-\binom{j+1}{2}}}{Q_j Q_\infty} e^{-2^j x}.$$

Moreover, $F_I(x)$ *is a positive function on* $\mathbb{R}^+$*; see Figure 4 for a plot.*

From the above expression, we see that as $x \to \infty$

$$F_I(x) = 1 - \frac{e^{-x}}{Q_\infty} + \mathcal{O}\left(e^{-2x}\right)$$

and as $x \to 0$, with the same method as in Section 2,

$$F_I(x) = \sqrt{\frac{\log 2}{2\pi}} \cdot \frac{X^{-\frac{1}{2}+\frac{1}{\log 2}}}{\log X} \exp\left(-\frac{\log(X \log X)^2}{2\log 2} - P\left(\log_2(X \log X)\right)\right) \times \left(1 + \mathcal{O}\left(\frac{(\log\log X)^2}{\log X}\right)\right),$$

where $X = \frac{1}{x \log 2}$ and the 1-periodic function $P(t)$ is defined in (7). The extra term $X^{-1}(\log X)^{-1}$, when compared with (6), comes from integration (or from the additional factor s^{-1} in the Laplace transform of $F_I(x)$ and the saddle-point approximation).

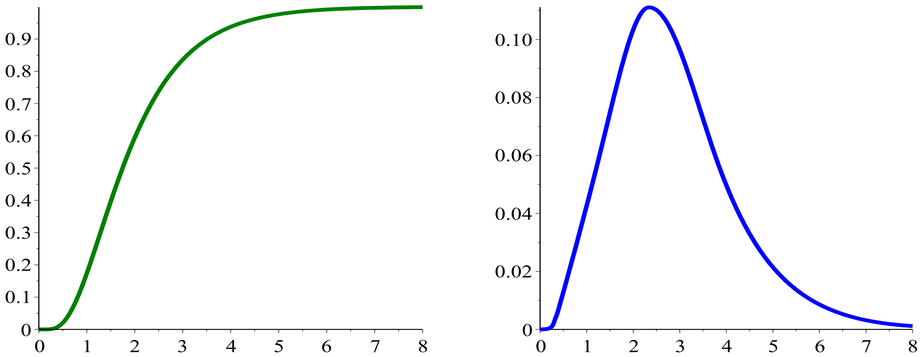


Figure 4: The functions F_I (left) and G_I (right).

Similarly, the same approach in Section 3 leads to the following asymptotic expansion for the variance.

Theorem 6. *The variance of the internal profile satisfies*

$$\mathrm{Var}(I_{n,k}) = 2^k G_I\left(2^{-k}n\right) + \mathcal{O}(1),$$

uniformly for $0 \leqslant k < n$*, where* $G_I(x)$ *is positive on* $\mathbb{R}^+$ *defined by*

$$G_I(x) = \sum_{j,r\geqslant 0}\sum_{0\leqslant h,\ell\leqslant j} \frac{(-1)^{r+h+\ell}2^{-j-\binom{r}{2}-\binom{h}{2}-\binom{\ell}{2}+h+\ell}}{Q_\infty Q_r Q_h Q_{j-h} Q_\ell Q_{j-\ell}} \varphi(2^{r+j}, 2^h + 2^l; x),$$

where φ *is defined in* (10).

Note that the only difference between $G(z)$ and $G_I(z)$ is that the exponent $2h + 2\ell$ in the series definition (9) of $G(z)$ is replaced by $h + \ell$; see Figure 4.

Proposition 5. *The function* $G_I(x)$ *satisfies*

$$G_I(x) \sim \begin{cases} \dfrac{e^{-x}}{Q_\infty}, & \textit{if } x \to \infty; \\ F_I(x), & \textit{if } x \to 0. \end{cases}$$

Corollary 3. *The variance of the internal profile tends to infinity iff there exists a positive sequence ω_n tending to infinity with n such that*

$$k_s + \frac{\omega_n}{\log n} \leqslant k \leqslant k_h - 1 - \frac{\omega_n}{\log n}, \tag{70}$$

where k_s and k_h are defined in (12).

Note that the only difference from the central range (13) for the external profile is the additional shift of -1 in the upper bound, a property implied from the fundamental relation

$$2I_{n,k} = I_{n,k+1} + B_{n,k+1}.$$

Finally, (70) is the range where the internal profile follows asymptotically a normal limit law.

Theorem 7. *If* $\mathrm{Var}(I_{n,k}) \to \infty$*, then $I_{n,k}$ is asymptotically normally distributed:*

$$\frac{I_{n,k} - \mathbb{E}(I_{n,k})}{\sqrt{\mathrm{Var}(I_{n,k})}} \xrightarrow{d} \mathscr{N}(0,1).$$

6 Applications

In this section, we apply our results on the profiles to establish the asymptotic two-point concentration of the height and the saturation level in random DSTs.

6.1 Height

We first prove Theorem 4 for the height of random DSTs. Recall that k_H is defined as follows:

$$k_H = \left\lfloor \log_2 n + \sqrt{2\log_2 n} - \frac{1}{2}\log_2\log_2 n + \frac{1}{\log 2} \right\rfloor.$$

To prove the asymptotic concentration of the height at k_H and k_H+1, we also need a finer approximation to the mean. By using more terms in the identity of Lemma 2 (together with Lemma 1 for the error term) and (26), we obtain

$$\mu_{n,k} = 2^k F(2^{-k}n) + F'(2^{-k}n) - 2^{-k-1}nF''(2^{-k}n) + \mathcal{O}\left(n^{-1} + 4^{-k}n\right). \tag{71}$$

Moreover, from Theorem 1 and Theorem 2 and the asymptotic growth of $G(z)$ and $H(z)$ from Section 2.3 and Section 3.3, respectively, we deduce that the variance of $B_{n,k}$ is asymptotically of the same order as the mean:

$$\sigma_{n,k}^2 = \Theta(\mu_{n,k}) \tag{72}$$

for k with $\mu_{n,k} \to \infty$; compare with Corollary 1.

Lemma 13. *For all k,*

$$1 - \sum_{\ell \geqslant 1} 2^{-\ell}\mu_{n,k+\ell} \leqslant \mathbb{P}(H_n \leqslant k) \tag{73}$$

and for all k with $\mu_{n,k+1} \to \infty$,

$$\mathbb{P}(H_n \leqslant k) = \mathcal{O}\left(\frac{1}{\mu_{n,k+1}}\right), \tag{74}$$

Proof. The proof relies on the first and the second moment method.

Noting that $I_{n,k} > 0$ if and only if $H_n > k$, we have, by the first moment method,

$$\mathbb{P}(H_n > k) = \mathbb{P}(I_{n,k} > 0) \leqslant \mathbb{E}(I_{n,k}).$$

On the other hand, in view of the relation $I_{n,k} = \sum_{\ell\geqslant 1} 2^{-\ell} B_{n,k+\ell}$, (73) follows from the inequality

$$\mathbb{P}(H_n > k) \leqslant \sum_{\ell\geqslant 1} 2^{-\ell}\mu_{n,k+\ell}.$$

For the upper bound, since $B_{n,k+1} > 0$ implies that $H_n > k$, we see that

$$\mathbb{P}(H_n > k) \geqslant \mathbb{P}(B_{n,k+1} > 0).$$

By the second moment method,

$$\mathbb{P}(H_n \leqslant k) \leqslant \mathbb{P}(B_{n,k+1} = 0) \leqslant \frac{\sigma_{n,k+1}^2}{\mu_{n,k+1}^2}.$$

From this (74) follows then from (72). ▮

Theorem 4 is then a consequence of the following two limit results.

Lemma 14. *The height of random DSTs satisfies*

$$\lim_{n\to\infty} \mathbb{P}(H_n \leqslant k_H + 1) = 1 \qquad \textit{and} \qquad \lim_{n\to\infty} \mathbb{P}(H_n \leqslant k_H - 1) = 0.$$

Proof. We first consider the expected value of the external profile around the level k_H and define

$$k_\ell := k_H + \ell = \left\lfloor \log_2 n + \sqrt{2\log_2 n} - \frac{1}{2}\log_2\log_2 n + \frac{1}{\log 2}\right\rfloor + \ell \qquad (\ell \in \mathbb{Z}).$$

Since $2^{-k_\ell} n \to 0$, we apply Proposition 1 for the asymptotics of $F(2^{-k_\ell}n)$ (and its derivatives). Note that the saddle-point equation in Proposition 1 has the form

$$\frac{\rho}{\log\rho} = \frac{2^{k_\ell}}{n\log 2},$$

or, with $R := \log\rho$,

$$\frac{R - \log R + \log\log 2}{\log 2} = \sqrt{2\log_2 n} - \frac{1}{2}\log_2\log_2 n + \frac{1}{\log 2} + \ell - \theta,$$

where $\theta = \theta_n$ denotes the fractional part of $\log_2 n + \sqrt{2\log_2 n} - \frac{1}{2}\log_2\log_2 n + 1/\log 2$. Now, by a direct bootstrapping argument, we obtain

$$R = \sqrt{2\log_2 n}\log 2 + 1 + \left(\ell + \frac{1}{2} - \theta\right)\log 2 + \frac{2\ell + 1 + \frac{2}{\log 2} - 2\theta}{2\sqrt{2\log_2 n}} + \mathcal{O}\left(\frac{1}{\log n}\right).$$

Substituting this into the asymptotics of $F(2^{-k_\ell}n)$ from Proposition 1, we have

$$k_\ell \log 2 + \log F(2^{-k_\ell}n) = -\sqrt{2\log_2 n}(\ell - 1 - \theta)\log 2 - \frac{3\log\log_2 n}{4} + \mathcal{O}(1).$$

On the other hand, from (71) and the estimates (by Proposition 1)

$$\frac{F'(2^{-k_\ell}n)}{F(2^{-k_\ell}n)} = \mathcal{O}(\rho) = \mathcal{O}(e^R)$$

and

$$\frac{F''(2^{-k_\ell}n)}{F(2^{-k_\ell}n)} = \mathcal{O}(\rho^2) = \mathcal{O}(e^{2R}),$$

we have

$$\mu_{n,k_\ell} = 2^{k_\ell} F(2^{-k_\ell}n)\left(1 + \mathcal{O}\big(n^{-1}e^{\mathcal{O}(\sqrt{\log n})}\big)\right).$$

Consequently, if $\ell \leqslant 0$, then

$$\mu_{n,k_\ell} \sim 2^{k_\ell} F(2^{-k_\ell}n) \geqslant \frac{e^{-\sqrt{2\log_2 n}(1+\theta)\log 2 + \mathcal{O}(1)}}{(\log_2 n)^{3/4}} \to \infty, \tag{75}$$

and if $\ell \geqslant 2$, then

$$\mu_{n,k_\ell} \sim 2^{k_\ell} F(2^{-k_\ell}n) \leqslant \frac{e^{-\sqrt{2\log_2 n}(1-\theta) + \mathcal{O}(1)}}{(\log_2 n)^{3/4}} \to 0. \tag{76}$$

Now, by Lemma 13, we show that

$$\sum_{\ell \geqslant 1} 2^{-\ell}\mu_{n,k_{\ell+1}} \to 0, \qquad \text{and} \qquad \mu_{n,k_0} \to \infty, \tag{77}$$

which will then prove the lemma.

Observe that the second limit of (77) follows directly from (75). We prove the first claim of (77), beginning with the inequality $\log(2X\log(2X)) \geqslant \log 2 + \log(X\log X)$ for $X \geqslant 2$, which in turn implies that

$$\log(2X\log(2X))^2 \geqslant \log(X\log X)^2 + 2(\log 2)\log(X\log X), \qquad (X \geqslant 2).$$

Thus, it follows from (6) that, for small x (with $X = 1/(x\log 2)$),

$$F(x/2) = \mathcal{O}\left(\frac{F(x)}{X\log X}\right) = \mathcal{O}\left(\frac{xF(x)}{\log(\frac{1}{x})}\right).$$

Hence, if $2^{-k}n \to 0$, then

$$\mu_{n,k+1} \leqslant C_2 \frac{2^{-k}n}{\log(2^k/n)}\mu_{n,k}, \tag{78}$$

for some constant $C_2 > 0$. From this, we have (for n sufficiently large)

$$\mu_{n,k_\ell} \leqslant \mu_{n,k_2}, \qquad (\ell \geqslant 2),$$

and thus

$$\sum_{\ell \geqslant 1} 2^{-\ell}\mu_{n,k_{\ell+1}} \leqslant \mu_{n,k_2} \to 0.$$

This proves (77) and the lemma. ∎

Remark 5. Observe that the only missing case in (75) and (76) is $\ell = 1$ for which we have

$$\mu_{n,k_1} \sim 2^{k_1} F(2^{-k_1} n) = \frac{e^{\theta\sqrt{2\log_2 n}\,\log 2 + \mathcal{O}(1)}}{(\log_2 n)^{3/4}}.$$

Thus, in this case, we have $\mu_{n,k_1} \to \infty$ if

$$\theta \geqslant \frac{3\log_2\log_2 n}{4\sqrt{2\log_2 n}}\left(1 + \frac{\omega_n}{\log_2\log_2 n}\right),$$

where ω_n is any sequence tending to infinity, and $\mu_{n,k_1} \to 0$ if

$$\theta \leqslant \frac{3\log_2\log_2 n}{4\sqrt{2\log_2 n}}\left(1 - \frac{\omega_n}{\log_2\log_2 n}\right),$$

where ω_n is as above. Finally, μ_{n,k_1} remains bounded if

$$\theta = \frac{3\log_2\log_2 n}{4\sqrt{2\log_2 n}}\left(1 \pm \frac{\mathcal{O}(1)}{\log_2\log_2 n}\right).$$

Thus we see that in almost all cases $\mu_{n,k_1} \to \infty$ and $\mu_{n,k_1+1} \to 0$, meaning that the height is in almost all cases asymptotic to $k_H + 1$; see also [20] where this was observed.

6.2 Saturation Level

Recall that the saturation level S_n of a binary tree with n internal nodes is defined as the maximal level with $I_{n,k} = 2^k$, that is, up to level S_n the binary tree is complete.

Define k_S as follows:

$$k_S = \lceil \log_2 n - \log_2 \log n \rceil$$

which is at the lower boundary of the central range (13).

Theorem 8. *The distribution of S_n is asymptotically concentrated on the two points $k_S - 1$ and k_S:*

$$\mathbb{P}(S_n = k_S - 1 \text{ or } S_n = k_S) \longrightarrow 1, \qquad (n \to \infty).$$

The proof of Theorem 8 is very similar to that of Theorem 4. The basic observation is that $S_n < k$ if and only if $\sum_{\ell \leqslant k} B_{n,\ell} > 0$. In particular, if $B_{n,k} > 0$ then $S_n < k$. Hence, as above, a direct application of the first and second moment method implies that

$$1 - \sum_{\ell \leqslant k} \mu_{n,\ell} \leqslant \mathbb{P}(S_n \geqslant k) \leqslant \frac{\sigma_{n,k}^2}{\mu_{n,k}^2}.$$

By using similar arguments as above, Theorem 8 then follows from the limit results:

$$\lim_{n\to\infty} \mathbb{P}(S_n \geqslant k_S - 1) = 1 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}(S_n \geqslant k_S + 1) = 0.$$

The only difference is that we now use the asymptotic expansion, for $2^{-k} n \to \infty$,

$$\mu_{n,k} \sim \sigma_{n,k}^2 \sim \frac{2^k}{Q_k}\left(1 - 2^{-k}\right)^n.$$

Acknowledgment

We thank the anonymous referees for a careful reading of the paper and for helpful comments and suggestions.

References

[1] D. Aldous and P. Shields (1988). A diffusion limit for a class of random-growing binary trees, *Probab. Theory Related Fields*, **79**, 509–542.

[2] C. W. Anderson (1970). Extreme value theory for a class of discrete distributions, with applications to some stochastic processes, *J. Appl. Probab.*, **7**, 99–113.

[3] C. W. Anderson, S. G. Coles and J. Hüsler (1997). Maxima of Poisson-like variables and related triangular arrays, *Ann. Appl. Probab.*, **7**, 953–971.

[4] G. E. Andrews (1998). *The Theory of Partitions*, Cambridge University Press.

[5] E. G. Coffman, Jr. and J. Eve (1970). File structures using hashing functions, *Commun. ACM*, **13**, 427–432.

[6] L. Devroye (1998). Universal limit laws for depths in random trees, *SIAM J. Comput.*, **28**, 409–432.

[7] L. Devroye (2005). Universal asymptotics for random tries and PATRICIA trees, *Algorithmica*, **42**, 11–29.

[8] M. Drmota (2009). *Random Trees: An Interplay between Combinatorics and Probability*, Springer-WienNewYork, Vienna.

[9] M. Drmota, S. Janson, and R. Neininger (2008). A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.* **18**, 288–333.

[10] M. Drmota and W. Szpankowski (2011). The expected profile of digital search trees, *J. Combin. Theory Ser. A*, **118**, 1939–1965.

[11] M. Drmota, A. Magner and W. Szpankowski (2018). Asymmetric Rényi problem, *Combin. Probab. Comput.*, accepted for publication.

[12] P. Flajolet (1983), On the performance evaluation of extendible hashing and trie searching, *Acta Info.* **20**, 345–369.

[13] P. Flajolet, X. Gourdon, and P. Dumas (1995). Mellin transforms and asymptotics: harmonic sums. *Theoret. Comput. Sci.*, **144**, 3–58.

[14] P. Flajolet and R. Sedgewick (2009). *Analytic Combinatorics*, Cambridge University Press.

[15] M. Fuchs, H.-K. Hwang and V. Zacharovas (2014). An analytic approach to the asymptotic variance of trie statistics and related structures, *Theor. Comput. Sci.*, **527**, 1–36.

[16] H.-K. Hwang, M. Fuchs and V. Zacharovas (2010). Asymptotic variance of random symmetric digital search trees, *Discrete Math. Theor. Comput. Sci. (Special issue in honor of Philippe Flajolet)*, **12**, 103–166.

[17] P. Jacquet and W. Szpankowski (1998). Analytical de-Poissonization and its applications, *Theoret. Comput. Sci.*, **201**, 1–62.

[18] P. Jacquet and W. Szpankowski (2015). *Analytic Pattern Matching: From DNA to Twitter*, Cambridge University Press, Cambridge.

[19] R. Kazemi and M. Vahidi-Asl (2011). The variance of the profile in digital search trees, *Discrete Math. Theor. Comput. Sci.*, **13**, 21–38.

[20] C. Knessl and W. Szpankowski (2000). Asymptotic behavior of the height in a digital search tree and the longest phrase of the Lempel-Ziv scheme, *SIAM J. Comput.*, **30**, 923–964.

[21] C. Knessl and W. Szpankowski (2002), Limit laws for the height in PATRICIA tries, *J. Algorithms* **44**, 63–97.

[22] C. Knessl and W. Szpankowski (2009). On the average profile of symmetric digital search trees, *Online J. Anal. Comb.*, **4**, 14 pp.

[23] D. E. Knuth (1998). *The Art of Computer Programming, Volume 3: Sorting and Searching*, Second Ed., Addison-Wesley, Reading, MA.

[24] A. G. Konheim and D. J. Newman (1973). A note on growing binary trees, *Discrete Math.*, **4**, 57–63.

[25] G. Louchard (1987). Exact and asymptotic distributions in digital and binary search trees, *RAIRO Inform. Thér. Appl.*, **21**, 479–495.

[26] G. Louchard, W. Szpankowski and J. Tang (1999). Average profile of the generalized digital search trees and the generalized Lempel-Ziv Algorithm, *SIAM J. Comp.*, **28**, 935–954.

[27] A. Magner and W. Szpankowski (2017). Profile of PATRICIA tries, *Algorithmica*, **80** (2018), 331–397.

[28] A. Magner, C. Knessl and W. Szpankowski (2014). Expected external profile of PATRICIA tries, *Proceedings of the Eleventh Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, 16–24.

[29] H. M. Mahmoud. *Evolution of Random Search Trees*, John Wiley & Sons Inc., New York, 1992.

[30] R. Neininger and L. Rüschendorf (2002). Rates of convergence for quicksort. *J. Algorithms*, **44**, 52–62.

[31] R. Neininger and L. Rüschendorf (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.* **14**, 378–418.

[32] G. Park, H.-K. Hwang, P. Nicodème, and W. Szpankowski (2009). Profile of tries, *SIAM J. Comput.*, **38**, 1821–1880.

[33] H. Prodinger (1995). Digital search trees and basic hypergeometric functions, *Bulletin of the EATCS*, **56**, 112–115.

[34] W. Szpankowski (2001). *Average Case Analysis of Algorithms on Sequences*, Wiley, New York.

[35] V. M. Zolotarev (1976). Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces, *Theory Probab. Appl.* **21**, 721–737; Erratum *ibid.* **22**, 881.

[36] V. M. Zolotarev (1977). Ideal metrics in the problem of approximating the distributions of sums of independent random variables, *Theory Probab. Appl.* **22**, 433–449.

Appendix: Proof of Proposition 1

We give a detailed proof of Proposition 1, which for convenience is re-stated here.

Proposition 1. *For each integer $m \geqslant 0$, the mth derivative of F satisfies*

$$F^{(m)}(z) = \frac{\rho^{m+\frac{1}{2}+\frac{1}{\log 2}}}{\sqrt{2\pi \log_2 \rho}} \exp\left(-\frac{(\log \rho)^2}{2\log 2} - P(\log_2 \rho)\right)\left(1 + \mathcal{O}\left(|\log \rho|^{-1}\right)\right), \tag{79}$$

as $|z| \to 0$ in the sector $|\arg(z)| \leqslant \varepsilon$, where $P(u)$ is given in (7) and ρ solves the equation

$$\frac{\rho}{\log \rho} = \frac{1}{z \log 2},$$

satisfying $|\rho| \to \infty$ as $|z| \to 0$.

Proof. Recall that

$$Q(s) := \prod_{j \geqslant 1} \left(1 - 2^{-j}s\right) \quad \text{and} \quad Q_n := \prod_{1 \leqslant j \leqslant n} \left(1 - 2^{-j}\right) = \frac{Q(1)}{Q(2^{-n})}.$$

Also

$$F(z) := \sum_{j \geqslant 0} \frac{(-1)^j 2^{-\binom{j}{2}}}{Q_j Q_\infty} e^{-2^j z}.$$

Since the Laplace transform $\mathscr{L}[F(z); s]$ of F is given by

$$\mathscr{L}[F(z); s] = \prod_{j \geqslant 0} \frac{1}{1 + 2^{-j}s} = \frac{1}{Q(-2s)} \qquad (\Re(s) > -1), \tag{80}$$

we have the Laplace inversion formula

$$F(z) = \frac{1}{2\pi i} \int_{1-i\infty}^{1+i\infty} \frac{e^{zs}}{Q(-2s)} \,\mathrm{d}s, \tag{81}$$

which is valid for $z = r$ where $r > 0$ is real. We are interested in the asymptotics of $F(r)$ as $r \to 0$, which is reflected by the large-s asymptotics of $\mathscr{L}[F(z); s]$. Our approach relies on the Mellin transform techniques and the saddle-point method; see the survey paper [13] for more background tools and applications on Mellin transform. In particular, taking logarithm on both sides of (80) (assuming that $1 + 2^{-j}s \neq 0$), we begin with the Mellin integral representation

$$\log Q(-2s) = \sum_{j \geqslant 0} \log(1 + 2^{-j}s) = \frac{1}{2\pi i} \int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \frac{\pi\, s^{-w}}{(1 - 2^w) w \sin \pi w} \,\mathrm{d}w,$$

because the Mellin transform of $\log(1+s)$ equals

$$\int_0^\infty s^{w-1} \log(1+s) \,\mathrm{d}s = \frac{\pi}{w \sin \pi w}, \qquad (\Re(w) \in (-1, 0)).$$

Note that if $w = u + iv$ and $s = |s|e^{ib}$ with u, v, b real and $|b| \leqslant \pi - \varepsilon$, then

$$\left|\frac{\pi s^{-w}}{(1 - 2^w) w \sin \pi w}\right| = \mathcal{O}\left(\frac{|s|^{-u} e^{-|v|(\pi - |b|)}}{|1 - 2^w||w|}\right),$$

provided that $|w|$ stays away from the zeros of the denominator. Thus by standard arguments, we deduce that (with $\beta := \frac{1}{2\log 2}$)

$$\log Q(-2s) = \beta(\log s)^2 + \frac{\log s}{2} + P(\log_2 s) + J(s), \tag{82}$$

when $|\arg(s)| \leqslant \pi - \varepsilon$, where the periodic function $P(u)$ has the Fourier series representation

$$P(u) = \frac{\log 2}{12} + \frac{\pi^2}{6\log 2} - \sum_{j\geqslant 1} \frac{\cos(2j\pi u)}{j \sinh \frac{2j\pi^2}{\log 2}}, \tag{83}$$

which also defines an analytic function as long as $|\Im(u)| \leqslant \pi - \varepsilon$; see Figure 5. Here the remainder $J(s)$ satisfies

$$\begin{aligned} J(s) &= \frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{\pi\, s^{-w}}{(1-2^w)w\sin\pi w}\,\mathrm{d}w \\ &= \frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \frac{\pi\, s^{w}}{(1-2^{-w})w\sin\pi w}\,\mathrm{d}w \\ &= -\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \frac{\pi\, (2s)^{w}}{(1-2^{w})w\sin\pi w}\,\mathrm{d}w \\ &= -\log Q\left(-\frac{1}{s}\right). \end{aligned}$$

We thus have the *identity*:

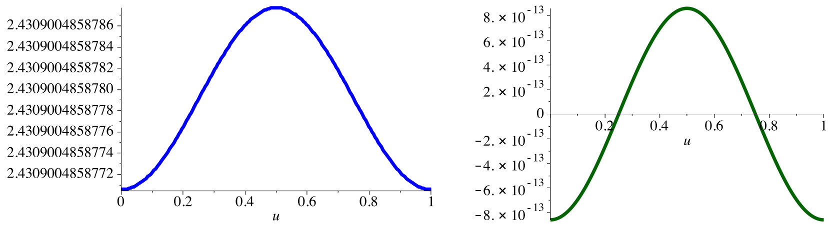


Figure 5: $P(u)$ in the unit interval (left) and the fluctuating part of $P(u)$ (right).

$$Q(-2s) = \frac{\sqrt{s}\, e^{\beta(\log s)^2 + P(\log_2 s)}}{Q(-\frac{1}{s})},$$

or

$$\boxed{\prod_{j\geqslant 0}\left(1+\frac{s}{2^j}\right) = \sqrt{s}\, e^{\beta(\log s)^2+P(\log_2 s)} \prod_{j\geqslant 1}\frac{1}{1+\frac{1}{2^j s}}},$$

which indeed holds, by analytic continuation, as long as $s \in \mathbb{C}\setminus(-\infty, 0]$. In particular, for large $|s|$ with $|\arg(s)| \leqslant \pi - \varepsilon$,

$$J(s) = -\frac{1}{s} + \frac{1}{6s^2} - \frac{1}{21s^3} + \frac{1}{60s^4} + \mathcal{O}\big(|s|^{-5}\big).$$

It follows, by substituting the asymptotic approximation (31), that

$$F(r) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} s^{-\frac{1}{2}} e^{rs-\beta(\log s)^2 - P(\log_2 s)} \Big(1+\mathcal{O}\big(|s|^{-1}\big)\Big)\,\mathrm{d}s.$$

Now, the asympotitcs of $F(r)$ as $r \to 0$ is obtained by a standard application of the saddle-point method. Therefore, we move the line of integration to $\Re(s) = \rho$, where $\rho > 0$ solves the saddle-point equation

$$\frac{\log \rho}{\rho} = \frac{r}{2\beta}.$$

Note that this does not change the value of the integral which is either clear from the domain of the Laplace transform of $f(z)$ or can also be seen directly since the integrand over the horizontal line segments of distance $T \gg 1$ from the positive real axis (and contained in a cone with $|\arg(s)| \leqslant \pi - \varepsilon$) is bounded above by

$$T^{-\frac{1}{2}} e^{r\Re(s) - \beta(\log T)^2},$$

implying that the integral along such lines is of order

$$T^{-\frac{1}{2}} \exp(-\beta(\log T)^2),$$

which decays to 0 as T tends to infinity. Thus, (with $s \mapsto \rho(1+it)$)

$$F(r) = \frac{\rho^{\frac{1}{2}} e^{\rho r}}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\rho t r - \beta(\log(\rho(1+it)))^2 - P(\log_2(\rho(1+it)))}}{\sqrt{1+it}} \Big(1 + \mathcal{O}\Big(\frac{1}{\rho|1+it|}\Big)\Big)\, \mathrm{d}t.$$

By a direct iterative argument, we obtain, with $R := \frac{2\beta}{r}$,

$$\rho = R\left(\log R + \log\log R + \frac{\log\log R}{\log R} - \frac{(\log\log R)^2 - 2\log\log R}{2(\log R)^2} + \mathcal{O}\left(\frac{|\log\log R|^3}{|\log R|^3}\right)\right).$$

Then we split the integral into two parts:

$$F(r) = \frac{\rho^{\frac{1}{2}} e^{\rho r}}{2\pi} \left(\int_{|t| \leqslant t_0} + \int_{|t| > t_0}\right) \frac{e^{i\rho t r - \beta(\log(\rho(1+it)))^2 - P(\log_2(\rho(1+it)))}}{\sqrt{1+it}} \Big(1 + \mathcal{O}\Big(\frac{1}{\rho|1+it|}\Big)\Big)\, \mathrm{d}t,$$

where $t_0 = (\log \rho)^{-\frac{2}{5}}$. Since

$$\Re((\log(\rho(1+it)))^2) = (\log \rho)^2 + (\log \rho)\log(1+t^2) + \tfrac{1}{4}\log(1+t^2)^2 - \arctan(t)^2$$

is a monotonic function of $|t|$ for fixed ρ, we have

$$\begin{aligned}
&\int_{|t| > t_0} \frac{e^{i\rho t r - \beta(\log(\rho(1+it)))^2 - P(\log_2(\rho(1+it)))}}{\sqrt{1+it}}\, \mathrm{d}t \\
&\quad = \mathcal{O}\left(e^{-\beta(\log \rho)^2} \int_{\log(1+t_0^2)}^{\infty} w^{-\frac{1}{2}} e^{-\beta(w \log \rho + \frac{1}{4}w^2) + w}\, \mathrm{d}w\right) \\
&\quad = \mathcal{O}\left(e^{-\beta(\log \rho)^2 - \varepsilon(\log \rho)^{\frac{1}{5}}}\right),
\end{aligned}$$

for some $\varepsilon > 0$. Now by the local expansions

$$\begin{aligned}
&i\rho t r - \beta(\log(\rho(1+it)))^2 \\
&\qquad = -\beta(\log \rho)^2 - \beta(\log \rho - 1)t^2 + \tfrac{1}{3}(2\log \rho - 3)it^3 + \mathcal{O}\big(t^4 \log \rho\big),
\end{aligned}$$

and

$$e^{-P(\log_2(\rho(1+it)))} = e^{-P(\log_2 \rho)} \left(1 - \frac{P'(\log_2 \rho)}{\log 2} it + \mathcal{O}(|t|^2)\right),$$

for $|t| \leqslant t_0$, we deduce that the integral with $|t| \leqslant t_0$ is asymptotic to

$$\begin{aligned} F(r) &= \frac{\rho^{\frac{1}{2}} e^{\rho r}}{2\pi} \int_{|t|\leqslant t_0} \frac{e^{i\rho t r - \beta(\log(\rho(1+it)))^2 - P(\log_2(\rho(1+it)))}}{\sqrt{1+it}} \Big(1 + \mathcal{O}\Big(\frac{1}{\rho|1+it|}\Big)\Big)\, \mathrm{d}t \\ &= \frac{\rho^{\frac{1}{2}} e^{\rho r - \beta(\log\rho)^2 - P(\log_2\rho)}}{2\sqrt{\pi\beta\log\rho}} \left(1 + \mathcal{O}\left((\log\rho)^{-1}\right)\right). \end{aligned}$$

Similarly, we also have

$$F^{(m)}(r) = \frac{\rho^{m+\frac{1}{2}} e^{\rho r - \beta(\log\rho)^2 - P(\log_2\rho)}}{2\sqrt{\pi\beta\log\rho}} \left(1 + \mathcal{O}\left(m^2(\log\rho)^{-1}\right)\right),$$

uniformly as $r \to 0$ and $m = o(\sqrt{\log\rho})$.

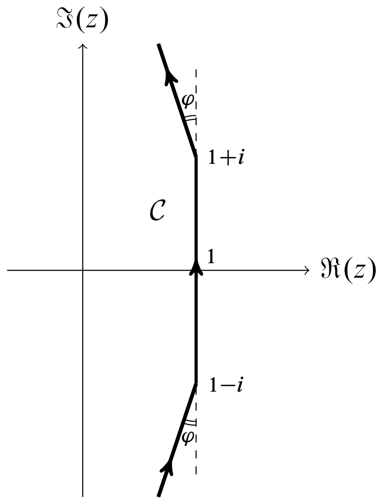


Figure 6: *The contour of integration in the integral representation of $F(z)$ when z is complex.*

We now look at the situation when $z = re^{i\theta}$ with $\theta \neq 0$ and $|\theta| \leqslant \varepsilon$. Here, (81) is no longer valid since the integral diverges. However, by the same idea of the Hankel contour used for extending the Gamma function, we can deform the original integration line into the following one:

$$\begin{aligned} \mathcal{C} := \{z = 1 - i + e^{-i(\frac{\pi}{2}+\varphi)}u \,:\, u \geqslant 0\} \\ \cup \{z = 1 + iu \,:\, -1 < u < 1\} \cup \{z = 1 + i + e^{i(\frac{\pi}{2}+\varphi)}u \,:\, u \geqslant 0\}, \end{aligned}$$

where $\varepsilon < \varphi$; see Figure 6. Then, we use (31) and make the substitution

$$\begin{aligned} F(z) &= \frac{1}{2\pi i} \int_{\mathcal{C}} s^{-\frac{1}{2}} e^{re^{i\theta}s - \beta(\log s)^2 - P(\log_2 s)} \left(1 + \mathcal{O}\big(|s|^{-1}\big)\right) \mathrm{d}s \\ &= \frac{e^{-\frac{1}{2}i\theta}}{2\pi i} \int_{e^{i\theta}\mathcal{C}} s^{-\frac{1}{2}} e^{rs - \beta(\log s - i\theta)^2 - P\left(\log_2(se^{-i\theta})\right)} \left(1 + \mathcal{O}\big(|s|^{-1}\big)\right) \mathrm{d}s, \end{aligned}$$

where $e^{i\theta}\mathcal{C}$ denotes the image of $\mathcal{C}$ under the mapping $s \mapsto e^{i\theta}s$. Note that the solution to the saddle-point equation

$$\frac{\log\rho(z)}{\rho(z)} = \frac{z}{2\beta} = \frac{e^{i\theta}}{R}$$

where $R := \frac{2\beta}{r}$, satisfies asymptotically for small r

$$\rho(z) = Re^{-i\theta}\biggl(\log R + \log\log R - i\theta + \frac{\log\log R - i\theta}{\log R} + \mathcal{O}\biggl(\frac{|\log\log R|^2}{|\log R|^2}\biggr)\biggr).$$

In particular ($\rho = \rho(|z|)$),

$$\rho(z) = \rho\, e^{-i\theta}\biggl(1 - \frac{i\theta}{\log R} + \frac{(\log\log R - 1)i\theta}{(\log R)^2} + \mathcal{O}\biggl(\frac{|\log\log R|^2}{|\log R|^2}\biggr)\biggr). \tag{84}$$

Since $|\theta| \leqslant \varepsilon$, we now deform the integration contour again into the vertical line $\Re(s) = \rho$ (which again does not change the value of the integral as can be seen by a similar argument as above) and proceed as before:

$$\begin{aligned} F(z) = \frac{e^{-\frac{1}{2}i\theta}}{2\pi i}\Biggl(\int_{\substack{s=\rho\cdot(1+it)\\ |t|\leqslant t_0}} + \int_{\substack{s=\rho\cdot(1+it)\\ |t|> t_0}}\Biggr) \\ s^{-\frac{1}{2}} e^{rs-\beta(\log s - i\theta)^2 - P(\log_2 s - i\theta)}\Bigl(1 + \mathcal{O}\bigl(|s|^{-1}\bigr)\Bigr)\,\mathrm{d}s. \end{aligned} \tag{85}$$

By the local expansion

$$\begin{aligned} r\rho i t - \beta(\log(\rho\cdot(1+it)) - i\theta)^2 &= -\beta(\log\rho - i\theta)^2 - 2\beta\theta t - \beta(\log\rho - 1 - i\theta)t^2 \\ &\quad + \frac{\beta}{3}(2\log\rho - 3 - 2i\theta)it^3 + \mathcal{O}\bigl((\log\rho)t^4\bigr), \end{aligned}$$

and the relations ($a \in \mathbb{R}, b > 0$)

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} t^m e^{-at-bt^2}\,\mathrm{d}t = \frac{e^{\frac{a^2}{4b}}}{2^{m+1}\sqrt{\pi b}}\sum_{0\leqslant\ell\leqslant\lfloor\frac{1}{2}m\rfloor}\frac{m!a^{m-2\ell}}{\ell!(m-2\ell)!b^{m-l}} \qquad (m = 0, 1, \ldots),$$

we deduce that the first integral on the RHS of (85) is asymptotic to

$$\begin{aligned} &\frac{\rho^{\frac{1}{2}}e^{-\frac{1}{2}i\theta - P\left(\log_2(\rho e^{-i\theta})\right) + r\rho - \beta(\log\rho - i\theta)^2 - \frac{\beta^2\theta^2}{\log\rho - 1 - i\theta}}}{2\sqrt{\pi\beta(\log\rho - 1 - i\theta)}}\Bigl(1 + \mathcal{O}\bigl((\log\rho)^{-1}\bigr)\Bigr) \\ &= \frac{\rho^{\frac{1}{2}}e^{-\frac{1}{2}i\theta - P\left(\log_2(\rho e^{-i\theta})\right) + r\rho - \beta(\log\rho - i\theta)^2}}{2\sqrt{\pi\beta\log\rho}}\Bigl(1 + \mathcal{O}\bigl((\log\rho)^{-1}\bigr)\Bigr). \end{aligned}$$

By (84), the right-hand side is asymptotic to

$$\frac{\rho(z)^{\frac{1}{2}}e^{-P(\log_2\rho(z)) + z\rho(z) - \beta(\log\rho(z))^2}}{2\sqrt{\pi\beta\log\rho(z)}}\Bigl(1 + \mathcal{O}\bigl(|\log\rho(z)|^{-1}\bigr)\Bigr).$$

It remains to prove the smallness of the other integral in (85), which is bounded above by

$$\begin{aligned} &\int_{\substack{s=\rho\cdot(1+it)\\ |t|>t_0}} s^{-\frac{1}{2}}e^{rs-\beta(\log s - i\theta)^2 - P(\log_2(se^{-i\theta}))}\,\mathrm{d}s \\ &\quad = \mathcal{O}\left(\rho^{\frac{1}{2}}e^{r\rho}\int_{t_0}^{\infty}(1+t^2)^{-\frac{1}{4}}e^{-\beta((\log\rho + \frac{1}{2}\log(1+t^2))^2 - (\theta - \arctan(t))^2)}\,\mathrm{d}t\right). \end{aligned}$$

The factor $(\theta - \arctan(t))^2$ being bounded for t in the range of integration, we obtain

$$\mathcal{O}\left(\rho^{\frac{1}{2}} e^{r\rho} \int_{t_0}^{\infty} (1+t^2)^{-\frac{1}{4}} e^{-\beta(\log\rho + \frac{1}{2}\log(1+t^2))^2}\, \mathrm{d}t\right) = \mathcal{O}\left(\rho^{\frac{1}{2}} e^{r\rho - \beta\log(\rho)^2 - \varepsilon(\log\rho)^{\frac{1}{5}}}\right),$$

which, by (84), is majorized by

$$\mathcal{O}\left(|\rho(z)|^{\frac{1}{2}} e^{\Re(z\rho(z) - \beta\log(\rho(z))^2) - \varepsilon|\log\rho(z)|^{\frac{1}{5}}}\right).$$

We thus obtain the approximation

$$F(z) = \frac{\rho(z)^{\frac{1}{2}} e^{z\rho(z) - \beta(\log\rho(z))^2 - P(\log_2\rho(z))}}{2\sqrt{\pi\beta\log\rho(z)}} \left(1 + \mathcal{O}\left(|\log\rho(z)|^{-1}\right)\right),$$

uniformly as $|z| \to 0$ in the sector $|\arg(z)| \leqslant \varepsilon$. The proof for the mth derivative of $F(z)$ is similar as above. ▮