Asymptotics of the Moments of Extreme-value related distribution functions

Guy Louchard* Helmut Prodinger†

June 22, 2005

Abstract

The asymptotic cost of many algorithms and combinatorial structures is related to the extreme-value Gumbel distribution $\exp(-\exp(-x))$. The following list is not exhaustive: Trie, Digital Search Tree, Leader Election, Adaptive Sampling, Counting Algorithms, trees related to the Register Function, Composition of Integers, some structures represented by Markov chains (Column-Convex Polyominoes, Carlitz Compositions), Runs and number of distinct values of some multiplicity in sequences of geometrically distributed random variables. Sometimes we can start from an exact (discrete) probability distribution, sometimes from an asymptotic analysis of the discrete objects (e.g., urn models) before establishing the relationship with the Gumbel distribution function. Also some Markov chains are either exactly and directly given by the structure itself, or as a limiting Markov process.

The main motivation of the paper is to compute the asymptotic distribution and the moments of the random variables in question. The moments are usually given by a dominant part and a small fluctuating part. We use Laplace and Mellin transforms and singularity analysis and aim for a unified treatment in all cases. Furthermore, our goal is a purely mechanical computation of dominant and fluctuating components, with the help of a computer algebra system. We provide each time the first three moments, but the treatment is (almost) completely automatic. We need some real analysis for the approximations and apart from that only easy complex analysis; simple poles and a few special functions.

1 Introduction.

The asymptotic cost of many algorithms is related to the extreme-value Gumbel distribution function (DF): $\exp(-\exp(-x))$. Also several combinatorial structures lead to the same behaviour.

A first aim of this paper is to come up with a classification of these objects. First of all, we can distinguish four classes depending on the *origin* of the objects. In the first class, the algorithms, we find the classical Trie, the Digital search tree, the Leader Election, the Adaptive Sampling, several Counting Algorithms (LogLog counting, Approximate Counting, Probabilistic Counting). In the second class, the combinatorial structures, we have the trees related to the registers needed in arithmetic expressions evaluation, the composition of integers, the number of distinct values of some multiplicity in sequences of geometrically distributed random variables (RV). A third class is given by some structures represented by Markov chains: the column-convex polyominoes, the Carlitz compositions, the runs of geometrically distributed RV, the monotone runs of uniformly distributed integer RV. A last class is related to singularity analysis techniques leading also to a Gumbel distribution function. As an example, we analyze the $M/G/1$ queue.

A second way of classifying our objects is to separate the class where we start from an exact (discrete) probability distribution from the class where we first make an asymptotic analysis of the

*Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium, email: louchard@ulb.ac.be

†University of the Witwatersrand, The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, P.O. Wits, 2050 Johannesburg, South Africa, e-mail: hproding@sun.ac.za

discrete objects (for instance with urn models) before establishing the relationship with the Gumbel DF. Also some Markov chains are either exactly and directly given by the structure itself, or as a limiting Markov process.

The second and most important motivation of the paper is to compute the asymptotic distribution and the moments of our related RV. The moments are usually given by a dominant part and a small fluctuating part. There we use Laplace and Mellin transforms and singularity analysis. We wanted to obtain a unified treatment of all our cases.

Our aim is to derive a purely mechanical computation of dominant and fluctuating components, with the help of computer algebra system (we use Maple here). As an example, we provide each time the first three moments, (the third moment is very rarely computed in the literature) but the treatment is completely automatic (with some human guidance of course).

A last but not least motivation is to simplify the analytic treatment by using only easy complex analysis: only simple poles are needed, we do not use alternating series (so we do not need Rice's formula). A small number of analytic functions are the only tools we need. This should be compared with the complicated techniques sometimes used in previous papers.

The paper is organized as follows: in Section 2, we recall the definitions, notations and known properties related to the dominant part of the moments. Section 3 is devoted to a detailed analysis of the fluctuating part of the moments. Section 4 deals with the direct limit case, leading from the discrete distribution to the Gumbel DF or a convergent series of such. The crucial rate of convergence is given in each case (with classical analysis techniques). In Section 5, we compute the asymptotic form of the moments, in a mechanized way. Section 6 is devoted to the Markov chain approach: the rate of convergence and the asymptotics are computed. In Section 7, we give an example of a preliminary singularity analysis: the $M/G/1$ queue. Section 8 concludes the paper. Due to length constraints, some algoritms are detailed in the complete version of this paper ([40]): Digital Search Tree, Leader Election, Approximate Counting, trees related to the Register Function. The number of distinct values of some multiplicity in sequences of geometrically distributed RV is analyzed in [41].

If we compare the approach in this paper with other ones that appeared previously, then we can notice the following. Traditionally, one would stay with exact enumerations as long as possible, and only at a late stage move to asymptotics. Doing this, one would, in terms of asymptotics, carry many unimportant contributions around, which makes the computations quite heavy, especially when it comes to higher moments. Here, however, approximations are carried out as early as possible, and this allows for streamlined (and often automatic) computations of the higher moments.

Let us finally point out some recent related works: in [19], S. Janson studies, by other methods, the characteristic function and moments of the integer-valued random variable $\lfloor X+\alpha \rfloor$, where X is a countinuous random variable. In [52] and [45], P. Robert and H. Mohamed use Poisson processes to analyze some algoritms with an underlying tree structure, without resorting to complex analysis.

2 Definitions, notations and known Properties.

Let us first give the notations we will use throughout the paper.

$$\begin{aligned}
L &:= \ln 2,\\
\log &:= \log_2,\\
\varepsilon &:= \text{ small real } > 0,\\
\tilde{\alpha} &:= \alpha/L,\\
Q(j) &:= \prod_{k=1}^{j}(1-2^{-k}), \quad Q(0)=1,\\
Q &:= Q(\infty) = 0.288788095087\ldots,\\
Q^{-}(j) &:= 1/Q(j),\\
R(j) &:= \frac{(-1)^{j+1}}{2^{j(j+1)/2}}Q^{-}(j),\ R(0)=-1.
\end{aligned}$$

These functions appear in many analyses of algorithms (see, for instance, Flajolet [7], Flajolet and Sedgewick [13], Louchard [32]).

The following facts will be frequently used:

$$\begin{aligned}
|Q - Q(i)| &= \mathcal{O}(2^{-i}), \\
Q(i) &\geq Q, \\
\ln(1-u) &> -u - u^2, \ u < 0.68, \\
(1-u)^n &< e^{-nu}, u \in]0,1[, \\
(1-u)^n &= e^{-nu}\left[1 - nu^2/2 + \mathcal{O}(nu^3)\right], \ u \in]0,1[, \\
(1-u)^n &= \left[1 - nu + n(n-1)/2u^2 + \mathcal{O}(nu^3)\right], \ u \in]0,1/n[, \\
\binom{n}{i} u^i (1-u)^{n-i} &= (nu)^i e^{-nu}/i![1 + \mathcal{O}(1/n) + \mathcal{O}(u) + \mathcal{O}(nu^2)], u \in]0,1[, i \text{ fixed },
\end{aligned}$$

this is the Poisson approximation.

For all integer-valued RV X_n analyzed in this paper, we set

$$p(j) := \Pr(X_n = j), \quad P(j) := \Pr(X_n \leq j).$$

Setting $\eta = j - \log n$, we will first compute f and F such that

$$p(j) \sim f(\eta), \quad P(j) \sim F(\eta), \quad n \to \infty,$$

and, of course,

$$f(\eta) = F(\eta) - F(\eta - 1).$$

Asymptotically, the distribution will be a periodic function of the fractional part of $\log n$. The distribution $P(j)$ does not converge in the weak sense, it does however converge along subsequences n_m for which the fractional part of $\log n_m$ is constant.

Next, we must show that

$$\mathbb{E}\left(X_n^k\right) = \sum_j j^k p(j) \sim \sum_j (\eta + \log n)^k [F(\eta) - F(\eta - 1)], \tag{1}$$

and/or

$$\mathbb{E}\left(\left[2^{X_n}/n\right]^d\right) = \sum_j (2^j/n)^d p(j) \sim \sum_j 2^{\eta d}[F(\eta) - F(\eta - 1)], \tag{2}$$

by computing a suitable rate of convergence. This is related to an uniform integrability condition (see Loève [30, Section 11.4].)

Finally we will use the following result from Hitczenko and Louchard [14] related to the dominant part of the moments (the '$\tilde{}$' sign is related to the moments of the integer-valued RV X_n).

Lemma 2.1 *Let a (integer-valued) RV X_n be such that $\Pr(X_n - \log n \leq \eta) \sim F(\eta)$, where $F(\eta)$ is the DF of a continuous RV Z with mean m_1, second moment m_2, variance σ^2 and centered moments μ_k. Assume that $F(\eta)$ is either an extreme-value DF or a convergent series of such and that (1) is satisfied. Let*

$$\varphi(\alpha) = \mathbb{E}(e^{\alpha Z}) = 1 + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} m_k = e^{\alpha m_1} \lambda(\alpha),$$

say, with

$$\lambda(\alpha) = 1 + \frac{\alpha^2}{2}\sigma^2 + \sum_{k=3}^{\infty} \frac{\alpha^k}{k!} \mu_k.$$

Let w, κ's (with or without subscripts) denote periodic functions of $\log n$, with period 1 and with usually small mean and amplitude. Actually, these functions depend on the fractional part of $\log n$: $\{\log n\}$.

Then the corresponding moments of X_n are given by

$$\begin{aligned}
\mathbb{E}(X_n - \log n) &\sim \int_{-\infty}^{+\infty} x[F(x) - F(x-1)]dx + w_1 \\
&= \tilde{m_1} + w_1, \quad \textit{with} \quad \tilde{m_1} = m_1 + 1/2, \\
\textit{var}(X_n) &\sim \mathbb{E}(X_n - (\log n + \tilde{m_1} + w_1))^2 \\
&\sim \int_{-\infty}^{+\infty} x^2[F(x) - F(x-1)]dx - \tilde{m_1}^2 + \kappa_2 \\
&= m_2 + m_1 + 1/3 - \tilde{m_1}^2 + \kappa_2 = \tilde{\sigma}^2 + \kappa_2, \quad \textit{with} \quad \tilde{\sigma}^2 = \sigma^2 + 1/12.
\end{aligned}$$

More generally, the centered moments of X_n are asymptotically given by $\tilde{\mu}_i + \kappa_i$, where

$$\Theta(\alpha) := 1 + \sum_{k=2}^{\infty} \frac{\alpha^k}{k!}\tilde{\mu}_k = \frac{2}{\alpha} sh(\alpha/2)\lambda(\alpha).$$

The neglected part is of order $1/n^\beta$ with $0 < \beta < 1$.

For instance, we derive

$$\begin{aligned}
\tilde{\mu}_2 &= \tilde{\sigma}^2 = \mu_2 + 1/12, \\
\tilde{\mu}_3 &= \mu_3, \\
\tilde{\mu}_4 &= \mu_4 + \sigma^2/2 + 1/80, \\
\tilde{\mu}_5 &= \mu_5 + 5/6\mu_3.
\end{aligned}$$

The moments of $X_n - \log n$ are asymptotically given by $\tilde{m}_i + w_i$, where the generating function of $\tilde{m}_i$ is given by

$$\phi(\alpha) := \int_{-\infty}^{\infty} e^{\alpha\eta} f(\eta) d\eta = 1 + \sum_{i=1}^{\infty} \frac{\alpha^i}{i!}\tilde{m}_i = \varphi(\alpha)\frac{e^\alpha - 1}{\alpha}. \tag{3}$$

The convergence domain for α will be studied in Section 5. This leads to

$$\begin{aligned}
\tilde{m}_1 &= m_1 + 1/2, \\
\tilde{m}_2 &= m_2 + m_1 + 1/3, \\
\tilde{m}_3 &= m_3 + 3m_2/2 + m_1 + 1/4;
\end{aligned}$$

w_i and κ_i will be analyzed in the next section.

Note that

$$\Theta(\alpha) = \phi(\alpha)e^{-\alpha\tilde{m}_1}.$$

This leads to

$$\begin{aligned}
\tilde{\mu}_2 &= \tilde{m}_2 - \tilde{m}_1^2 \\
\tilde{\mu}_3 &= \tilde{m}_3 + 2\tilde{m}_1^3 - 3\tilde{m}_2\tilde{m}_1.
\end{aligned} \tag{4}$$

Also

$$\varphi(\alpha) = \int_{-\infty}^{+\infty} e^{\alpha\eta} F'(\eta) d\eta = -\alpha \int_{-\infty}^{+\infty} e^{\alpha\eta} F(\eta) d\eta,$$

with suitable boundary conditions (which are satisfied in all examples we present).

3 The fluctuating components.

3.1 The moments of $X_n - \log n$.

To analyze the periodic component w_i to be added to the moments $\tilde{m}_i$ we return to the proof of the Lemma 2.1 given in Hitczenko and Louchard [14]. For instance, under (1),

$$\mathbb{E}(X_n - \log n) \sim E^{(1)}(n) = \sum_{j=1}^{\infty}[F(j - \log n) - F(j - \log n - 1)][j - \log n]. \tag{5}$$

Set $y = 2^{-x}$ and $G(y) = F(x)$. Equation (5) becomes a *harmonic sum*

$$E^{(1)}(n) := \sum_j [G(n/2^j) - G(n/2^{j+1})][-\log(n/2^j)],$$

the Mellin transform of which is (for a good reference on Mellin transforms, see Flajolet et al. [9] or Szpankowski [53])

$$\frac{2^s}{1-2^s}\Upsilon_1^*(s), \tag{6}$$

and

$$\begin{aligned}\Upsilon_1^*(s) &= \int_0^\infty y^{s-1}[G(y) - G(y/2)][-\log y]dy \\ &= \int 2^{-sx}[F(x) - F(x-1)]xLdx.\end{aligned}$$

Then

$$\Upsilon_1^*(s) = L\,\phi'(\alpha)\big|_{\alpha=-Ls}. \tag{7}$$

The fundamental strip of (6) is usually of the form $s \in \langle -C_1, 0\rangle$, $C_1 > 0$. This will be detailed in Section 5.

Set also

$$\Upsilon_0^*(s) = L\,\phi(\alpha)\big|_{\alpha=-Ls},\ \Upsilon_0^*(0) = L.$$

We assume now that all poles of $\frac{2^s}{1-2^s}\Upsilon_1^*(s)$ are simple, which will be the case in all our examples, and given by $s = 0$, $s = \chi_l$, with $\chi_l := 2l\pi i/L$, and, in some cases (DST for example), $s = C_3$, $C_3 > 0$, but we will check that the residue of $\frac{2^s}{1-2^s}\Upsilon_1^*(s)$ at C_3 is 0.

Using

$$E^{(1)}(n) = \frac{1}{2\pi i}\int_{C_2-i\infty}^{C_2+i\infty}\frac{2^s}{1-2^s}\Upsilon_1^*(s)n^{-s}ds, \quad -C_1 < C_2 < 0,$$

the asymptotic expression of $E^{(1)}(n)$ is obtained by moving the line of integration to the right, for instance to the line $\Re = C_4 > 0$, taking residues into account. This gives

$$E^{(1)}(n) = -\operatorname{Res}\left[\frac{2^s}{1-2^s}\Upsilon_1^*(s)n^{-s}\right]\Bigg|_{s=0} - \sum_{l\neq 0}\operatorname{Res}\left[\frac{2^s}{1-2^s}\Upsilon_1^*(s)n^{-s}\right]\Bigg|_{s=\chi_l} + \mathcal{O}(n^{-C_4}), \tag{8}$$

under a "slow increase property" (see below).

The residue at $s = 0$ gives of course

$$\tilde{m}_1 = \Upsilon_1^*(0)/L = \phi'(0).$$

The other residues lead to

$$w_1 = \sum_{l\neq 0}\Upsilon_1^*(\chi_l)e^{-2l\pi i\log n}/L. \tag{9}$$

More generally,

$$\mathbb{E}(X_n - \log n)^k \sim \tilde{m}_k + w_k,$$

with

$$w_k = \sum_{l \neq 0} \Upsilon_k^*(\chi_l) e^{-2l\pi i \log n}/L, \tag{10}$$

and

$$\Upsilon_k^*(s) = L\,\phi^{(k)}(\alpha)\Big|_{\alpha=-Ls}.$$

The residues analysis is similar to the previous one.

To compute the periodic component κ_i to be added to the centered moments $\tilde{\mu}_i$, we first set

$$\mathfrak{m}_1 := \tilde{m}_1 + w_1.$$

The variance of $X_n - \log n$ is asymptotically given by

$$\mathbb{E}[(X_n - \log n - \mathfrak{m}_1)^2] \sim \tilde{m}_2 + w_2 - \mathfrak{m}_1^2 = \tilde{\mu}_2 + \kappa_2.$$

The third centered moment is asymptotically given by

$$\mathbb{E}[(X_n - \log n - \mathfrak{m}_1)^3] \sim \tilde{m}_3 + w_3 - 3(\tilde{m}_2 + w_2)\mathfrak{m}_1 + 3\mathfrak{m}_1^3 - \mathfrak{m}_1^3 = \tilde{\mu}_3 + \kappa_3.$$

More generally, we start from

$$\phi(\alpha) := 1 + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} \tilde{m}_k = \varphi(\alpha) \frac{e^\alpha - 1}{\alpha}.$$

We replace $\tilde{m}_k$ by $\tilde{m}_k + w_k$, leading to

$$\phi_p(\alpha) = \phi(\alpha) + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} w_k.$$

But it is easy to check that

$$\sum_{l \neq 0} \phi(-L\chi_l) e^{-2l\pi i \log n} = 0,$$

so we obtain

$$\begin{aligned}
\phi_p(\alpha) &= \phi(\alpha) + \sum_{k=0}^{\infty} \sum_{l \neq 0} \phi^{(k)}(\alpha)\Big|_{\alpha=-L\chi_l} e^{-2l\pi i \log n} \frac{\alpha^k}{k!} \\
&= \phi(\alpha) + \sum_{l \neq 0} \phi(\alpha - L\chi_l) e^{-2l\pi i \log n} \\
&= \sum_{l} \phi(\alpha - L\chi_l) e^{-2l\pi i \log n}.
\end{aligned} \tag{11}$$

Finally, we compute

$$\Theta_p(\alpha) = \phi_p(\alpha) e^{-\alpha \mathfrak{m}_1} = 1 + \sum_{k=2}^{\infty} \frac{\alpha^k}{k!} (\tilde{\mu}_k + \kappa_k) = \Theta(\alpha) + \sum_{k=2}^{\infty} \frac{\alpha^k}{k!} \kappa_k, \tag{12}$$

leading to the (exponential) generating function (GF) of κ_k. This leads to

$$\begin{aligned}
\kappa_2 &= w_2 - w_1^2 - 2\tilde{m}_1 w_1, \\
\kappa_3 &= 6\tilde{m}_1^2 w_1 + 6\tilde{m}_1 w_1^2 + 2w_1^3 - 3\tilde{m}_2 w_1 - 3\tilde{m}_1 w_2 - 3w_1 w_2 + w_3.
\end{aligned} \tag{13}$$

All algebraic manipulations of this paper will be mechanically performed by Maple. We will give explicit expressions for $\tilde{\mu}_3$ and κ_3.

It will appear that $\Upsilon^*_k(s)$ are analytic functions (in some domain), depending on classical functions such as Γ, ζ, $\psi(k,s)$ (the $(k+1)$-gamma function) and an entire function $N(s)$ which is the analytic continuation of

$$\sum_{j\geq 1}(-1)^{\nu(j)}/j^s,$$

where $\nu(j)$ denotes the number of ones in the binary representation of j. We will also encounter the entire functions

$$\begin{aligned} H_1(\alpha) &:= \sum_{k\geq 0} e^{\alpha(k+1)}R(k), \\ H_2(\alpha) &:= \sum_{k\geq 0} e^{\alpha k}2^k R(k). \end{aligned}$$

S. Janson (private communication) mentioned that, with $\theta(z) := \prod_{k\geq 1}(1-z/2^k)$, we have

$$\begin{aligned} H_1(\alpha) &= -e^{\alpha}\theta(e^{\alpha}), \\ H_2(\alpha) &= -\theta(2e^{\alpha}) = (e^{\alpha}-1)\theta(e^{\alpha}) = (e^{-\alpha}-1)H_1(\alpha). \end{aligned}$$

The justification of (8) is classically done by considering the rectangle (for details, see Section .5.1)

$$[C_2-(k+1/2)\chi_l,\ C_2+(k+1/2)\chi_l,\ C_4+(k+1/2)\chi_l,\ C_4-(k+1/2)\chi_l].$$

But we know that $\Gamma(s)$ decreases exponentially in the direction $i\infty$:

$$|\Gamma(\sigma+it)| \sim \sqrt{2\pi}|t|^{\sigma-1/2}e^{-\pi|t|/2}.$$

Also, we have a "slow increase property" for all functions we encounter: let $s=\sigma+it$,

$$|\zeta(s)| = \mathcal{O}(|t|^{1-\sigma}),\ \sigma<1,$$

and a uniform version exists,

$$N(s) = \mathcal{O}(1+|s|^4),\ \sigma>-1.$$

(See Flajolet and Martin [10]).

For $H_.(-Ls)$, we use the "sum splitting technique" as described in Knuth [29, p. 131], and used in Flajolet [7]: let $\sigma>-1$ and $\rho(x)$ be an increasing function.

For $k<\rho(|s|)$, the contribution is bounded by

$$\rho(|s|)\sup_{0\leq j\leq\rho(|s|)}\left(\frac{1}{2^{\sigma j+j^2}}\right) = \mathcal{O}\left(1+\rho(|s|)2^{\rho(|s|)}\right),$$

for $k\geq\rho(|s|)$, the contribution is bounded by

$$\sum_{\rho(|s|)}^{\infty}\frac{2^k}{2^{k^2}} = \mathcal{O}\left(\frac{2^{\rho(|s|)}}{2^{2\rho(|s|)}}\right).$$

Choosing $\rho(x)=\log x$ insures the slow increase property.

It is not always evident that the limiting function $F(\eta)$ is indeed a distribution function. But here we can use the following lemma from Janson [19]. Assume that Y_n is a sequence of integer-valued random variables that is monotone. If, for every sequence of integers k_n,

$$\Pr(Y_n\leq k_n) = F(k_n-a_n)+o(1), n\to\infty,$$

for a right-countinuous function F with $\lim_{x\to-\infty}F(x)=0, \lim_{x\to\infty}F(x)=1$ and a sequence a_n such that $a_n\to\infty$ and $a_{n+1}-a_n\to 0$ as $n\to\infty$, then F is a distribution function.

These conditions will be satisfied in our examples.

3.2 The moments of $2^{X_n}/n$.

Let us analyze the moments $\mathbb{E}[(2^{X_n}/n)^d]$. This amounts to computing

$$\sum_j [G(n/2^j) - G(n/2^{j+1})](2^j/n)^d,$$

the Mellin transform of which is

$$\frac{2^s}{1-2^s}\Upsilon^*_{1,d}(s), \tag{14}$$

with (we put an extra 1 index in the expressions of this Section to make the difference with the moments of $X_n - \log n$)

$$\begin{aligned}
\Upsilon^*_{1,d}(s) &= \int_0^\infty y^{s-1}[G(y) - G(y/2)]y^{-d}dy \\
&= \int 2^{-sx}2^{dx}[F(x) - F(x-1)]Ldx \\
&= L\ \phi(\alpha)|_{\alpha=L(d-s)}\,.
\end{aligned}$$

So (14) leads to

$$\mathbb{E}[(2^{X_n}/n)^d] \sim \tilde{m}_{1,d} + w_{1,d}, \tag{15}$$

with

$$\begin{aligned}
\tilde{m}_{1,d} &= \phi(Ld), \\
w_{1,d} &= \sum_{l\neq 0} \Upsilon^*_{1,d}(\chi_l)e^{-2l\pi i \log n}/L.
\end{aligned} \tag{16}$$

Similarly, we derive

$$\begin{aligned}
\phi_1(\alpha) &= 1 + \sum_{k=1}^\infty \frac{\alpha^k}{k!}\tilde{m}_{1,k} = 1 + \sum_{k=1}^\infty \frac{\alpha^k}{k!}\phi(Lk), \\
\phi_{p1}(\alpha) &= 1 + \sum_{k=1}^\infty \frac{\alpha^k}{k!}[\tilde{m}_{1,k} + w_{1,k}] \\
&= 1 + \sum_{k=1}^\infty \frac{\alpha^k}{k!}\left[\phi(Lk) + \sum_{l\neq 0} \phi(\alpha)|_{\alpha=L(k-\chi_l)}\, e^{-2l\pi i \log n}\right] \\
&= \sum_0^\infty \frac{\alpha^k}{k!}\left[\phi(Lk) + \sum_{l\neq 0} \phi(\alpha)|_{\alpha=L(k-\chi_l)}\, e^{-2l\pi i \log n}\right], \\
\mathfrak{m}_{1,1} &= \tilde{m}_{1,1} + w_{1,1} = \phi(L) + \sum_{l\neq 0} \phi(\alpha)|_{\alpha=L(1-\chi_l)}\, e^{-2l\pi i \log n}, \\
\Theta_1(\alpha) &= \phi_1(\alpha)e^{-\alpha\tilde{m}_{1,1}} = 1 + \sum_{k=2}^\infty \frac{\alpha^k}{k!}\tilde{\mu}_{1,k}, & (17) \\
\Theta_{p1}(\alpha) &= \phi_{p1}(\alpha)e^{-\alpha\mathfrak{m}_{1,1}} = \Theta_1(\alpha) + \sum_{k=2}^\infty \frac{\alpha^k}{k!}\kappa_{1,k}. & (18)
\end{aligned}$$

4 Direct limits. The rate of convergence.

We will not give details on the description of algorithms, in order to keep the paper's length within reasonable limits. All algorithms are completely explained in the references we provide. We deal here with the direct limit case, leading from the discrete distribution to the Gumbel DF or a convergent series of such. We will compute $f(\eta)$ and $F(\eta)$. The crucial rate of convergence is given in each case (with classical analysis techniques).

4.1 The depth of the Trie.

From Pittel [46], Jacquet and Régnier [18], Louchard [31], we have

$$P(j) = (1 - 1/2^j)^n. \tag{19}$$

Let $\eta = j - \log n$, this leads to

$$F(\eta) = \exp(-2^{-\eta}). \tag{20}$$

First we must bound $|P(j) - F(\eta)|$. Let $1/2 < \beta < 1$. We will consider three ranges

- for $j < \beta \log n$, $P(j)$ is small. Indeed

$$P(j) \leq \exp(-n^{1-\beta}).$$

- for $\beta \log n \leq j < 2 \log n$, we have

$$|P(j) - F(\eta)| = F(\eta)\mathcal{O}(n/n^{2\beta}) = \mathcal{O}(1/n^{2\beta-1}).$$

- for $j = 2\log n + x$, $x \geq 0$, we have

$$|P(j) - F(\eta)| = \mathcal{O}(n/(n^4 2^{2x})) = \mathcal{O}(2^{-2x}/n^3).$$

To justify (1), we must bound the difference between the moments of X_n and the moments based on $F(\eta)$:

$$\begin{aligned}
&\Big| \sum_j j^k \left[[P(j) - P(j-1)] - [F(\eta) - F(\eta - 1)] \right] \Big| \\
&\leq \Bigg[\mathcal{O}((\beta \log n)^{k+1} \exp(-n^{1-\beta})) \\
&\quad + (2\log n)^{k+1}\mathcal{O}(1/n^{2\beta-1}) + \mathcal{O}\Bigg(\sum_{x \geq 0} (2 \log n + x)^k 2^{-2x}/n^3 \Bigg) \Bigg] = \mathcal{O}(1/n^{2\beta-1-\varepsilon}).
\end{aligned}$$

4.2 The search cost in Digital Search Tree.

From Louchard [32], we know that

$$p(j) = \frac{2^{j-1}}{n} \sum_{k=1}^{j-1} \frac{R(j-1-k)}{Q(k-1)} \left[(1 - 1/2^k)^{n-1} - 1 \right].$$

This can also be written as

$$p(j) = \frac{2^{j-1}}{n} \sum_{i=0}^{j-2} \frac{R(i)}{Q(j-i-2)} \left[(1 - 1/2^{j-1-i})^{n-1} - 1 \right]. \tag{21}$$

Let $\eta = j - \log n$, this leads to

$$f(\eta) = 2^{\eta-1} \sum_{i=0}^{\infty} R(i) \left[\exp(-2^{\eta-i-1}) - 1 \right] /Q. \tag{22}$$

We also obtain

$$P(j) = \frac{1}{n} \left[(2^j - 1) + \sum_{k=1}^{j-1} 2^k \frac{R(j-1-k)}{Q(k-1)} (1 - 1/2^k)^{n-1} \right],$$

$$F(\eta) = 2^{\eta-1} \sum_{i=0}^{\infty} R(i)/2^i \left[\exp(-2^{-(\eta-i-1)}) - 1 \right] /Q. \tag{23}$$

First we must bound $|p(j) - f(\eta)|$. Let $0 < \beta < 1$. We will consider three ranges.

- for $j < \beta \log n$, $p(j)$ is small. Indeed

$$|p(j)| \leq \frac{1}{Q^2} \frac{2^{j-1}}{n} 4 = \mathcal{O}(1/n^{1-\beta}).$$

- for $\beta \log n \leq j < 2 \log n$ we use again the "sum splitting technique." Set $r = 2\sqrt{\log n}$.

  1. truncating the sum in (21) to r leads to an error E_1:

  $$E_1 \leq \frac{n^2}{Q^2 n} \frac{1}{2^{r^2/2}} = \mathcal{O}(1/n).$$

  2. replacing $Q(j-i-2)$ in the truncated sum by Q leads to an error E_2:

  $$E_2 \leq \frac{r 2^{j-1}}{Qn} 2^{-j+r+2} 2 = \mathcal{O}(1/n^{1-\varepsilon}).$$

  3. replacing $(1 - 1/2^{j-1-i})^{n-1}$ in the truncated sum by $\exp(-2^{-(\eta-i-1)})$ leads to an error E_3:

  $$E_3 \leq \frac{r 2^{j-1}}{Q^2 n} \sup_{i \leq r} \mathcal{O}\Big(\frac{n 2^{2i}}{2^{2j}}\Big) = \mathcal{O}\Big(\frac{r 2^{2r}}{2^j}\Big) = \mathcal{O}(1/n^{\beta-\varepsilon}).$$

  4. completing the sum in (22) leads to an error E_4:

  $$E_4 \leq \frac{n^2}{Q^2 n} \frac{1}{2^{r^2/2}} = \mathcal{O}(1/n).$$

  So, finally,

  $$|p(j) - f(\eta)| = \mathcal{O}(1/n^{\beta-\varepsilon}).$$

- for $j = 2\log n + x$, $x \geq 0$, we set $r = 2\sqrt{\log n} + 2\sqrt{x}$. So $1/2^{r^2/2} \leq 2^{-2x}/n^2$. We proceed now as in the second range

  1. $$E_1 \leq \frac{n^2 2^x}{Q^2 n} \frac{2^{-2x}}{n^2} = \mathcal{O}(2^{-x}/n).$$

  2. $$E_2 \leq \frac{r 2^j}{Qn} 2^{-j+r+2} n 2^r / 2^j = \mathcal{O}(2^{-(1-\varepsilon)x}/n^{2-\varepsilon}).$$

  3. $$E_3 \leq \frac{r 2^{j-1}}{Q^2 n} \sup_{i \leq r} \mathcal{O}\Big(\frac{n 2^{2i}}{2^{2j}}\Big) = \mathcal{O}\Big(\frac{r 2^{2r}}{n^2 2^x}\Big) = \mathcal{O}(2^{-(1-\varepsilon)x}/n^{2-\varepsilon}).$$

  4. $$E_4 \leq \frac{n^2 2^x}{Q^2 n} \frac{2^{-2x}}{n^2} = \mathcal{O}(2^{-x}/n).$$

  So, finally,

  $$|p(j) - f(\eta)| = \mathcal{O}(2^{-x}/n).$$

Now we must bound $|\sum_j j^k [p(j) - f(\eta)]|$. This is done like in the Trie case.

Several generalizations can be found in the literature: the m-ary Digital search tree, the b-bucket Digital Search tree. See, for instance, Louchard [33], Louchard and Szpankowski [42], Louchard Szpankowski and Tang [43], Hubalek [15], Hubalek et al. [16].

4.3 The duration of the Leader Election Algorithm.

From Prodinger [48], Fill, Mahmoud and Szpankowski [5], we have

$$P(j) = \frac{n}{2^j}\sum_{i=1}^{2^j}(1-i/2^j)^{n-1}. \tag{24}$$

Let $\eta = j - \log n$, this leads to

$$F(\eta) = 2^{-\eta}\sum_{i=1}^{\infty}\exp(-i2^{-\eta}) = 2^{-\eta}\exp(-2^{-\eta})/[1-\exp(-2^{-\eta})]. \tag{25}$$

First we must bound $|P(j) - F(\eta)|$. Let $11/12 < \beta < 1$. We will consider three ranges

- for $j < \beta \log n$, $P(j)$ is small. Indeed

$$P(j) \le \frac{n}{2^j}\sum_{i=1}^{2^j}\exp(-(n-1)i/2^j) \le n\frac{1-e^{-n}}{1-e^{-n/2^j}}e^{-n/2^j} = \mathcal{O}\left(n\exp(-n^{1-\beta})\right).$$

- for $\beta \log n \le j < \log n$ we set $r = n^{1/4}$.

  1. truncating the sum in (24) leads to an error E_1:

$$E_1 \le \frac{n}{2^j}\frac{e^{-nr/2^j} - e^{-n}}{1-e^{-n/2^j}}e^{-n/2^j} = \mathcal{O}(n^{1-\beta})\exp(-n^{1/4}).$$

  2. replacing $(1-i/2^j)^{n-1}$ in the truncated sum by $\exp(-i2^{-\eta})$ leads to an error E_3:

$$E_3 \le \frac{n}{2^j}\sum_{i=1}^{r} n(i/2^j)^2 = \mathcal{O}\Big(\frac{n^2r^3}{2^{3j}}\Big) = \mathcal{O}(1/n^{3\beta-11/4}).$$

  3. completing the sum in (25) leads to an error E_4:

$$E_4 \le \frac{n}{2^j}\frac{e^{-nr/2^j}}{1-e^{-n/2^j}}e^{-n/2^j} = \mathcal{O}(n^{1-\beta})\exp(-n^{1/4}).$$

  So, finally,

$$|P(j) - F(\eta)| = \mathcal{O}(1/n^{3\beta-11/4}).$$

- for $j = \log n + \eta$, $\eta \ge 0$, we set $\xi := 2^{-\eta} < 1$. We obtain

$$\begin{aligned} F(\eta) &= \bar{F}(\xi) = \xi e^{-\xi}/(1-e^{-\xi}), \\ P(j) &\sim \bar{P}(\xi) = \xi\sum_{i=1}^{n/\xi} e^{-i\xi}(1-i^2\xi^2/(2n)), \end{aligned}$$

  which can be explicitly computed. Moreover,

$$\begin{aligned} p(j) &= P(j) - P(j-1) \sim \bar{P}(\xi) - \bar{P}(2\xi), \\ f(\eta) &= F(\eta) - F(\eta-1) = \bar{F}(\xi) - \bar{F}(2\xi), \end{aligned}$$

  and Maple gives us

$$p(j) - f(\eta) \sim e^{-n}n\xi/4 - e^{-n}n\xi^2/8.$$

Now we must bound $|\sum_j j^k[[P(j)-P(j-1)] - [F(\eta) - F(\eta-1)]]|$. Again this is done like in the Trie case.

4.4 The Adaptive sampling algorithm.

R denotes the last (left) bucket size, with value r. From Flajolet [8] and Louchard [34], we have

$$
\begin{aligned}
p(r,j) &= \binom{n}{r}\left(\frac{1}{2^j}\right)^r\left(1-\frac{1}{2^j}\right)^{n-r}\left[1-\sum_{k=0}^{b-r}\binom{n-r}{k}\left(\frac{1}{2^j}\right)^k\left(1-\frac{1}{2^j}\right)^{n-r-k}\right],\\
p(j) &= \sum_{r=0}^{b} p(r,j).
\end{aligned}
$$

We are interested in the random variables: $W := X_n - \log n$ and $Z := r2^{X_n}/n$. So we set $\eta = j - \log n$, this leads to

$$
f(r,\eta) = \exp(-2^{-\eta})\frac{2^{-r\eta}}{r!}\left[1-\exp(-2^{-\eta})\sum_{k=0}^{b-r}\frac{2^{-k\eta}}{k!}\right].
$$

Also we derive

$$
P(r,j) = \binom{n}{r}\sum_{u=1}^{j}\left(\frac{1}{2^u}\right)^r\left(1-\frac{1}{2^u}\right)^{n-r}\left[1-\sum_{k=0}^{b-r}\binom{n-i}{k}\left(\frac{1}{2^u}\right)^k\left(1-\frac{1}{2^u}\right)^{n-r-k}\right]
$$

and

$$
F(r,\eta) = \sum_{l=0}^{\infty}\exp(-2^{-(\eta-l)})\frac{2^{-r(\eta-l)}}{r!}\left[1-\exp(-2^{-(\eta-l)})\sum_{k=0}^{b-r}\frac{2^{-k(\eta-l)}}{k!}\right]. \tag{26}
$$

First we must bound $|p(j)-f(\eta)|$. Of course, we can treat each index k, r separately. We will consider the first term in $p(r,j)$, denoted by $p_1(r,j)$. The other terms are analyzed similarly.

4.4.1

To analyze the moments of X_n, we will consider three ranges. Let $1/2 < \beta < 1$.

- for $j < \beta\log n$, $p_1(r,j)$ is small. Indeed
$$p_1(r,j) \le n^r\exp(-n^{1-\beta}).$$
- for $\beta\log n \le j < 2\log n$ we have
$$|p_1(r,j)-f_1(r,\eta)| = f_1(r,\eta)\mathcal{O}[1/n+1/n^\beta+n/n^{2\beta}] = f_1(r,\eta)[\mathcal{O}(1/n^\beta)+\mathcal{O}(1/n^{2\beta-1})].$$
- for $j = 2\log n + x$, $x \ge 0$, we have
$$|p_1(r,j)-f_1(r,\eta)| = \mathcal{O}\left[(n/(n^2 2^x))^r[1/n+1/n^2+n/n^4]\right] = \mathcal{O}(1/(n^{r+1}2^{rx})).$$

Now we must bound $|\sum_j j^k[p_1(r,j)-f_1(r,\eta)]|$. This is done like in the Trie case.

4.4.2

To analyze the moments of Z i.e. $\mathbb{E}\left[(r2^{X_n}/n)^d\right]$, we must bound

$$
S_1 := \sum_j (2^j/n)^d|p_1(r,j)-f_1(r,\eta)|.
$$

Assume

$$
d < r, \quad \frac{r-d+1}{r-d+2} < \beta < 1.
$$

- for $j < \beta \log n$, we have
$$S_1 < 2\beta \log n (n^\beta/n)^d n^r \exp(-n^{1-\beta}).$$
- for $\beta \log n \leq j < \log n$, we have
$$\begin{aligned} S_1 &< \sum_{j=\beta\log n}^{\log n} \left(\frac{n}{2^j}\right)^r (2^j/n)^d [1/n + \mathcal{O}(1/n^\beta) + \mathcal{O}(1/n^{2\beta-1})] \\ &= \mathcal{O}\Bigg(\sum_{j=\beta\log n}^{\log n} 2^{j(d-r)}/n^{d-r}[\mathcal{O}(1/n^\beta) + \mathcal{O}(1/n^{2\beta-1})]\Bigg) \\ &= \mathcal{O}(\log n \; n^{(1-\beta)(r-d)}[\mathcal{O}(1/n^\beta) + \mathcal{O}(1/n^{2\beta-1})]) = \mathcal{O}(1/n^C),\ C > 0. \end{aligned}$$
- for $j = \log n + \eta$, $\eta \geq 0$, we have
$$S_1 < \mathcal{O}\Bigg(\sum_{\eta \geq 0} 2^{d\eta}/2^{r\eta}[1/n + 1/n^2 + n/n^4]\Bigg) = \mathcal{O}(1/n) \quad \text{if } d < r.$$

Actually, it is shown in Louchard [34], due to some cancellations, that we must only have $d \leq b$.

4.5 The Approximate Counting algorithm.

From Flajolet [7, Proposition 1], we have
$$p(j) = \sum_{k=0}^{j-1} 2^k \frac{-R(k)}{Q(j-1-k)}(1 - 1/2^{j-k})^n.$$
Let $\eta = j - \log n$, this leads to
$$f(\eta) = \sum_{k=0}^{\infty} 2^k \frac{-R(k)}{Q} \exp(-2^{-\eta+k}).$$
(Compare also [26, 47, 49].) It has been pointed out in [32] that it has some similarities with the Digital Search Tree distribution (see Section 4.2). Actually, as noticed by S. Janson (private communication), the distribution for approximate counting is the same as for unsuccessful search in Digital Search Tree (not only asymptotically).

The rate of convergence problem is completely solved in Flajolet [7]. Also, we obtain
$$\begin{aligned} P(j) &= -\sum_{k=0}^{j-1} \sum_{u=k+1}^{j} 2^k \frac{R(k)}{Q(u-1-k)}(1 - 1/2^{u-k})^n, \\ F(\eta) &= -\sum_{k=0}^{\infty} 2^k \frac{R(k)}{Q} \sum_{i=k}^{\infty} \exp(-2^{-\eta+i}). \end{aligned} \tag{27}$$
The algorithm can be generalized by changing the base. The analysis is quite similar.

4.6 The LogLog counting algorithm.

This is one of the most delicate cases. From Durand and Flajolet [4] we must analyze the moments $\mathbb{E}(\Delta^d)$ of $\Delta := m/n2^J$, where $J := \sum_1^m J_i$ and J_i has distribution (19) with n now given by ν_i : the number of keys in bucket i. ν_i is $Bin[\tilde{n}, \tilde{n}(1-1/m))]$ with $\tilde{n} := n/m$ and m is a design parameter. Actually, $\{\nu_1 \dots \nu_m\}$ is given by a multinomial distribution. We know that we can construct a "box"
$$[\tilde{n} - \tilde{n}^\theta, \tilde{n} + \tilde{n}^\theta]^m, 1/2 < \theta < 1, \tag{28}$$
such that, by large deviation analysis, the probability that $\{\nu_1 \dots \nu_m\}$ is outside this box is bounded by $\exp(-C\tilde{n}^{2\theta-1})$. We rewrite Δ as
$$\Delta = 2^{\sum X_i/m}, \text{ with } X_i := J_i - \log_2(\tilde{n}).$$

4.6.1

Let us first assume that ν_i is exactly given by its mean $\tilde{n}$. Dropping the i-index from X_i and the $\tilde{}$ sign from $\tilde{n}$ for convenience, we first compute $\mathbb{E}(2^{\tau X}) = \mathbb{E}((2^J/n)^\tau)$. We proceed as in the Adaptive sampling case. We must bound

$$S_1 := \sum_j (2^j/n)^\tau |P(j) - F(\eta)|,$$

where $P(j)$, $F(\eta)$ are given by (19), (20). Actually, we will see that we must have $\tau < 1$. Let $(1-\tau)/(2-\tau) < \beta < 1$.

- for $j < \beta \log n$, we have
$$S_1 < 2\beta \log n (n^\beta/n)^\tau \exp(-n^{1-\beta}).$$
- for $\beta \log n \leq j < \log n$, we have
$$S_1 < \sum_{j=\beta\log n}^{\log n} \frac{n}{2^{2j}}(2^j/n)^\tau = \sum_{j=\beta\log n}^{\log n} 2^{j(\tau-2)}/n^{\tau-1} = \mathcal{O}\left(\log n \; n^{1-\tau-\beta(2-\tau)}\right).$$
- for $j = \log n + \eta$, $\eta \geq 0$, we have
$$S_1 < \sum_{\eta\geq 0} 2^{\tau\eta} 2^{-2\eta}/n = \mathcal{O}(1/n) \text{ if } \tau < 2.$$

4.6.2

Now we assume that we are inside the box (28) and we must bound

$$S_2 := \sum_j (2^j/n)^\tau \left| P(j) - (1-1/2^j)^{n+\xi n^\theta} \right|,$$

with $|\xi| < 1$. Let $1/2 < \beta < 1$, $1/2 < \theta < \beta$, $\theta < \tau + \beta(1-\tau)$.

- for $j < \beta \log n$, we have
$$\begin{aligned} S_2 &< \beta \log n (n^\beta/n)^\tau [\exp(-((n+\xi n^\theta)/n^\beta) + \exp(-n^{1-\beta})] \\ &= \mathcal{O}((2\beta \log n)(n^\beta/n)^\tau \exp(-n^{1-\beta})) \text{ if } \theta < \beta. \end{aligned}$$
- for $\beta \log n \leq j < \log n$, we have
$$(1-1/2^j)^n - (1-1/2^j)^{n+\xi n^\theta} = (1-1/2^j)^n (1-(1-1/2^j)^{\xi n^\theta}) = \mathcal{O}(n^\theta/2^j). \qquad (29)$$

  This leads to
$$S_2 = \mathcal{O}\Bigg(\sum_{j=\beta\log n}^{\log n} \frac{n^\theta}{2^j}(2^j/n)^\tau\Bigg) = \mathcal{O}\Bigg(\sum_{j=\beta\log n}^{\log n} 2^{j(\tau-1)}/n^{\tau-\theta}\Bigg) = \mathcal{O}\left(\log n/n^{\tau-\theta+\beta(1-\tau)}\right).$$
- for $j = \log n + \eta$, $\eta \geq 0$, (29) leads to $\mathcal{O}(n^\theta/(n2^\eta))$ and
$$S_2 = \mathcal{O}\Bigg(\sum_{\eta\geq 0} n^\theta/(n2^\eta) 2^{\tau\eta}\Bigg) = \mathcal{O}(1/n^{1-\theta}) \text{ if } \tau < 1.$$

4.6.3

Now we consider the case $\nu > n + n^{\theta}$. We will show that

$$S_3 := \sum_j p(j)(2^j/n)^{\tau} \exp(-Cn^{2\theta-1}), \text{ with } p(j) \text{ based on } \nu,$$

is small.

- for $j < \beta \log n$, we have
$$S_3 < \beta \log n (n^{\beta}/n)^{\tau} \exp(-Cn^{2\theta-1}).$$
- for $\beta \log n \le j < \log n$, we have
$$S_3 < \sum_{j=\beta \log n}^{\log n} (2^j/n)^{\tau} \exp(-Cn^{2\theta-1}) = \mathcal{O}\left(\log n \exp(-Cn^{2\theta-1})\right).$$
- for $j = \log n + \eta$, $\eta \ge 0$, we notice that $1 < \nu/n < m$. We set $\rho := \log_2 m$. We have
$$\begin{aligned} S_3 &< \sum_0^{\rho} 2^{\tau\eta} \exp(-Cn^{2\theta-1}) + \sum_{\eta\ge\rho} [(1-1/(n2^{\eta}))^{\nu} - (1-1/(2n2^{\eta}))^{\nu}]\, 2^{\tau\eta} \exp(-Cn^{2\theta-1}) \\ &= \mathcal{O}\left(2^{\tau\rho} \exp(-Cn^{2\theta-1})\right) \\ &+ \sum_{\eta\ge\rho} \left[e^{-\nu/(n2^{\eta})}(1-\nu/(2n^2 2^{2\eta})) - e^{-\nu/(2n2^{\eta})}(1-\nu/(2.4n^2 2^{2\eta}))\right] 2^{\tau\eta} \exp(-Cn^{2\theta-1}) \\ &= \mathcal{O}(m^{\tau} \exp(-Cn^{2\theta-1})) + \sum_{\eta\ge\rho} e^{-\nu/(2n2^{\eta})}[\nu/(2n2^{\eta}) + \mathcal{O}(\nu/n^2 2^{2\eta})] 2^{\tau\eta} \exp(-Cn^{2\theta-1}). \end{aligned}$$

The sum converges if $\tau < 1$.

4.6.4

Now we consider the case $0 < \nu < n - n^{\theta}$. We have here $0 < \nu/n < 1$. The analysis proceeds like in the previous subsection. We omit the details.

As $\Delta = 2^{\sum X_i/m}$, we see that we must now have $d < m$.

4.7 The number of Registers in arithmetic expressions evaluation.

From Flajolet [6], Flajolet and Prodinger [11], and Louchard [31], we have

$$\begin{aligned} P(j) &= 1 + 2\sum_{k=1}^{\infty} (1 - 2k^2 4^{j+1}/n) e^{-k^2 4^{j+1}/n}, \\ p(j) &= 2 \sum_{k \text{ odd } >0}^{\infty} (2k^2 4^j/n - 1) e^{-k^2 4^j/n}. \end{aligned}$$

Let $\eta = j - \log_4 n$, this leads to

$$\begin{aligned} F(\eta) &= 1 + 2\sum_{k=1}^{\infty} (1 - 8k^2 4^{\eta}) e^{-4k^2 4^{\eta}}, \\ f(\eta) &= 2 \sum_{k \text{ odd } >0}^{\infty} (2k^2 4^{\eta} - 1) e^{-k^2 4^{\eta}}. \end{aligned} \tag{30}$$

(Compare also [12, 22].)

We have no rate of convergence problem here: the convergence of moments is already proved in Flajolet [6].

4.8 Product forms: Probabilistic Counting and first empty part in Compositions of integers.

The Probabilistic Counting is analyzed in Flajolet and Martin [10] (compare also [28]). Compositions of integers are analyzed in Hitczenko and Louchard [14], where it is shown that the first empty part value is asymptotically distributed as the first empty urn in an urn model where we throw n GEOM(1/2) RV. This is exactly equivalent to the Probabilistic Counting. Moreover, it is shown in [14] that the urns are asymptotically independent with probability $(1-1/2^i)^n$ for the urn i to be empty. So, setting (we omit n in $A_.(j)$ for convenience)

$$A_1(j) := \prod_{i=1}^{j}[1-(1-1/2^i)^n] \le 1,$$

we obtain here

$$P(j) = 1 - A_1(j). \tag{31}$$

In Flajolet and Martin [10], they use

$$P^*(j) = 1 - \sum_{i=0}^{2^j}(-1)^{\nu(i)}(1-i/2^j)^n,$$

where again $\nu(j)$ denotes the number of ones in the binary representation of j.

$P(j)$ and $P^*(j)$ are asymptotically equivalent (see Section 5.9 for details), but we will start here from (31) to illustrate the manipulation of product forms. We have

$$\begin{aligned}
\eta &= j - \log n,\\
A_2(j) &:= \prod_{k=0}^{j-1}\left[1-\exp(-e^{-L(\eta-k)})\right] \le 1,\\
p(j) &= (1-1/2^j)^n A_1(j-1),\\
A_3(\eta) &:= \prod_{k=0}^{\infty}[1-\exp(-e^{-L(\eta-k)}] \le 1,\\
F(\eta) &:= 1 - A_3(\eta),\\
f(\eta) &:= \exp(-e^{-L\eta})A_3(\eta-1).
\end{aligned} \tag{32}$$

4.8.1

Let us first bound $\Delta_1 := |A_1(j) - A_2(j)|$. We have

$$A_1(j) = \prod_{k=0}^{j-1}\left[1-\exp\big(-e^{-L(\eta-k)}\big)\left[1-\frac{n}{2^{1+2(j-k)}}+\mathcal{O}\left(\frac{n}{2^{3(j-k)}}\right)\right]\right].$$

Set

$$\varphi(x) := \exp(-e^{-Lx}).$$

We derive

$$\begin{aligned}
A_1(j) &= \prod_{k=0}^{j-1}\left[[1-\varphi(\eta-k)]\left[1+\frac{\varphi(\eta-k)}{2(1-\varphi(\eta-k))}\left[e^{-2L(\eta-k)}/n+\mathcal{O}(e^{-3L(\eta-k)}/n^2)\right]\right]\right]\\
&= A_2(j)A_4(j),
\end{aligned}$$

with

$$A_4(j) := \prod_{k=0}^{j-1}\left[1+\frac{\varphi(\eta-k)}{2(1-\varphi(\eta-k))}\left[e^{-2L(\eta-k)}/n+\mathcal{O}(e^{-3L(\eta-k)}/n^2)\right]\right].$$

We have the following equivalents:

$$
\begin{aligned}
\eta - k < 0 \quad &: \quad \frac{\varphi(\eta-k)}{1-\varphi(\eta-k)} e^{-2L(\eta-k)} < \frac{\varphi(\eta-k)}{\varphi(0)} e^{-2L(\eta-k)}, \\
\eta - k \geq 0 \quad &: \quad \varphi(\eta-k) = 1 - e^{-L(\eta-k)} + \mathcal{O}(e^{-2L(\eta-k)}), \\
&\quad\ \ 1-\varphi(\eta-k) = e^{-L(\eta-k)} + \mathcal{O}(e^{-2L(\eta-k)}), \\
&\quad\ \ \frac{\varphi(\eta-k)}{1-\varphi(\eta-k)} e^{-2L(\eta-k)} = \varphi(\eta-k)\left[e^{-L(\eta-k)} + \mathcal{O}(e^{-2L(\eta-k)})\right].
\end{aligned}
$$

All these functions are bounded. Now

$$
\begin{aligned}
\ln(A_4(j)) \quad = \quad & \mathcal{O}\left[\sum_{k=0}^{j-1}\left[\varphi(\eta-k)e^{-L(\eta-k)}/n + \varphi(\eta-k)e^{-2L(\eta-k)}/n\right.\right. \\
& \left.\left. +\ \varphi^2(\eta-k)e^{-2L(\eta-k)}/n^2 + \varphi^2(\eta-k)e^{-4L(\eta-k)}/n^2\right]\right] = \mathcal{O}(1/n),\ \text{uniformly on } \eta.
\end{aligned}
$$

So

$$A_4(j) = 1 + \mathcal{O}(1/n),\ \text{uniformly on } \eta,$$

and

$$\Delta_1 = A_2(j)\mathcal{O}(1/n).$$

4.8.2

Let us now bound $\Delta_2 := |A_2(j) - A_3(\eta)|$. We have

$$A_3(\eta) - A_2(j) = A_2(j)[1 - A_5(j)],$$

with

$$A_5(j) := \prod_{k=j}^{\infty}[1 - \exp(-e^{-L(\eta-k)})].$$

We obtain, independently of j,

$$A_5 := \prod_{l=0}^{\infty}[1 - \exp(-ne^{Ll})].$$

But, as n is large,

$$\ln(A_5) \geq -\sum_0^{\infty}\exp(-ne^{Ll}) + \mathcal{O}(\sum_0^{\infty}\exp(-2ne^{Ll})) \sim -\int_0^{\infty} e^{-nu}\frac{du}{Lu} \sim -\frac{e^{-n}}{Ln}.$$

So

$$A_5 \geq 1 - \frac{e^{-n}}{Ln},$$

and

$$\Delta_2 \leq A_2(j)\frac{e^{-n}}{Ln},$$

uniformly on j.

4.8.3

Now we bound $\Delta_3 := |P(j) - F(\eta)|, \eta \geq 0$. We have

$$
\begin{aligned}
\Delta_3 \quad &= \quad A_2(j)\mathcal{O}(1/n) = \left(1 - \exp(e^{-L\eta})\right)\mathcal{O}(1/n) \\
&= \quad \left[1 - \left(1 - e^{-L\eta} + \mathcal{O}(e^{-2L\eta})\right)\right]\mathcal{O}(1/n) = e^{-L\eta}\left[1 + \mathcal{O}(e^{-L\eta})\right]\mathcal{O}(1/n).
\end{aligned}
$$

4.8.4

Now we can proceed as in Section 4.6. We will analyze $\mathbb{E}((2^{X_n}/n)^\tau)$ (the analysis of the RV X_n is similar). Let us first assume that ν_i is exactly given by its mean $\tilde{n}$. We must bound

$$S_{1,1} := \sum_j (2^j/n)^\tau |p(j) - f(\eta)|,$$
$$S_{1,2} := \sum_j (2^j/n)^\tau |P(j) - F(\eta)|.$$

Actually, we will see that we must have $\tau < 1$. Let $\beta < 1$.

- for $j < \beta \log n$, $p(j)$ is small. Indeed,
$$p(j) \le \exp(-n^{1-\beta}).$$
- for $\beta \log n \le j < 3/2 \log n$, we have
$$S_{1,2} < \frac{2}{n}\frac{3}{2}\log n \quad n^{\tau/2}.$$
- for $j = 3/2\log n + x$, $x \ge 0$, we have
$$S_{1,2} < \sum_{x\ge 0} 2^{\tau x}\frac{1}{n^{3/2}}e^{-Lx}n^{\tau/2}.$$

4.8.5

Now we assume that we are inside the box (28) and we must bound

$$S_{1,1} := \sum_j (2^j/n)^\tau |p_1(j) - p_2(j)|,$$
$$S_{1,2} := \sum_j (2^j/n)^\tau |P_1(j) - P_2(j)|,$$

where $p_1(j), P_1(j)$ are based on n and $p_2(j), P_2(j)$ are based on $n + \xi n^\theta, |\xi| < 1$. We have

$$\begin{aligned} \eta_1 &= j - \log n, \\ \eta_2 &= j - \log n - \xi n^{\theta-1} + \mathcal{O}(\xi^2 n^{2(\theta-1)}). \end{aligned}$$

We will choose $\theta < \beta, 1/2 < \theta < 1 - \tau/2, 1/2 < \beta < 1$.

- for $j < \beta \log n$, $p_2(j)$ is small. Indeed,
$$p_2(j) \le \exp(-(n+\xi n^\theta)/n^\beta).$$
- for $\beta \log n \le j < 3/2 \log n$, we have
$$|P_1(j) - P_2(j)| \le 2/n + |A_3(\eta_1) - A_3(\eta_2)|.$$
But
$$|A_3(\eta_1) - A_3(\eta_2)| = \mathcal{O}(F'(\eta_1)\xi n^{\theta-1}) = \mathcal{O}(n^{\theta-1}).$$
So
$$S_{1,2} < 3/2 \log n \, n^{\tau/2} n^{\theta-1}$$
- for $j = 3/2\log n + x$, $x \ge 0$, we have
$$\begin{aligned} |P_1(j) - P_2(j)| &= \mathcal{O}(e^{-Lx}/n^{3/2}) + \mathcal{O}(e^{-Lx}e^{L\xi n^{\theta-1}}/[(n+\xi n^\theta)\sqrt{n}]) \\ &+ \mathcal{O}(e^{-Lx}/\sqrt{n}) + \mathcal{O}(e^{-Lx}e^{L\xi n^{\theta-1}}/\sqrt{n}). \end{aligned}$$
So
$$S_{1,2} < \sum_{x\ge 0} n^{\tau/2}2^{\tau x}e^{-Lx}e^{L\xi n^{\theta-1}}/\sqrt{n}.$$

4.8.6

Now we consider the case $\nu > n + n^\theta$. We will show that

$$S_3 := \sum_j p(j)(2^j/n)^\tau \exp(-Cn^{2\theta-1}), \text{ with } p(j) \text{ based on } \nu,$$

is small.

- for $j < \beta \log n$, we have
$$S_3 < \beta \log n (n^\beta/n)^\tau \exp(-Cn^{2\theta-1}).$$
- for $\beta \log n \le j < \log n$, we have
$$S_3 < \sum_{j=\beta \log n}^{\log n} (2^j/n)^\tau \exp(-Cn^{2\theta-1}) = \mathcal{O}(\log n \exp(-Cn^{2\theta-1})).$$
- for $j = \log n + \eta$, $\eta \ge 0$, we obtain, with $\rho = \log m$,
$$\begin{aligned} \eta &= \log n + \eta - \log \nu, e^{-L\eta} \le me^{-L\eta}, \\ S_3 &\le m^\tau \exp(-Cn^{2\theta-1}) + \sum_{\eta \ge \rho} m 2^{\tau\eta} e^{-L\eta} \exp(-Cn^{2\theta-1}). \end{aligned}$$

5 Direct limits. The moments.

We compute here the asymptotic form of the moments, in a mechanized way. Let us first recall the notations we use:

$$\begin{aligned} \varphi(\alpha) &= \int_{-\infty}^{+\infty} e^{\alpha\eta} F'(\eta) d\eta = e^{\alpha m_1} \lambda(\alpha), \\ m_1 &= \int_{-\infty}^{+\infty} \eta F'(\eta) d\eta, \\ \lambda(\alpha) &= 1 + \frac{\alpha^2}{2}\sigma^2 + \sum_{k=3}^{\infty} \frac{\alpha^k}{k!}\mu_k. \end{aligned}$$

In particular, the Gumbel DF: $F(\eta) = \exp(-e^{-\eta})$ leads to

$$\begin{aligned} \varphi(\alpha) &= \Gamma(1-\alpha), \Re(\alpha) < 1, \\ m_1 &= \gamma, \\ m_2 &= \pi^2/6 + \gamma^2, \\ m_3 &= 2\zeta(3) + \pi^2\gamma/2 + \gamma^3, \\ \phi(\alpha) &= \varphi(\alpha)(e^\alpha - 1)/\alpha \text{ , there is no singularity at } \alpha = 0, \\ \tilde{m}_1 &= \gamma + 1/2, \\ \tilde{m}_2 &= \gamma + 1/3 + \pi^2/6 + \gamma^2, \\ \tilde{m}_3 &= 1/4 + 2\zeta(3) + \pi^2\gamma/2 + \gamma^3 + \gamma + \pi^2/4 + 3\gamma^2/2, \\ \lambda(\alpha) &= e^{-\alpha m_1}\varphi(\alpha), \\ \sigma^2 &= \pi^2/6, \\ \mu_3 &= 2\zeta(3), \\ \mu_4 &= 3\pi^4/20, \\ \Theta(\alpha) &= \frac{2}{\alpha}\sinh(\alpha/2)\lambda(\alpha), \\ \tilde{\mu}_2 &= \pi^2/6 + 1/12, \\ \tilde{\mu}_3 &= 2\zeta(3). \end{aligned}$$

The fundamental strip for

$$\frac{2^s}{1-2^s}\Upsilon^*_k(s)$$

is $\Re(s) \in \langle -1/L, 0\rangle$.

For further use, we also need

$$\int_{-\infty}^{+\infty} e^{\alpha\eta} e^{\eta}[\exp(-e^{-\eta}) - 1]d\eta = \frac{\Gamma(1-\alpha)}{\alpha(1+\alpha)}.$$

5.1 The depth of the Trie.

We have, from (20),

$$\begin{aligned} F(\eta) &= \exp(-e^{-L\eta}), \\ \varphi(\alpha) &= \Gamma(1-\tilde{\alpha}), \text{ with } \tilde{\alpha} := \alpha/L, \ \Re(\alpha) < L. \end{aligned}$$

This leads to

$$\begin{aligned} m_1 &= \gamma/L, \\ m_2 &= (\pi^2/6 + \gamma^2)/L^2, \\ m_3 &= (2\zeta(3) + \pi^2\gamma/2 + \gamma^3)/L^3, \\ \tilde{m}_1 &= \gamma/L + 1/2, \\ \tilde{m}_2 &= \gamma/L + 1/3 + (\pi^2/6 + \gamma^2)/L^2, \\ \tilde{m}_3 &= 1/4 + (2\zeta(3) + \pi^2\gamma/2 + \gamma^3)/L^3 + \gamma/L + (\pi^2/2 + 3\gamma^2)/(2L^2), \\ \sigma^2 &= \pi^2/(6L^2), \\ \mu_3 &= 2\zeta(3)/L^3, \\ \mu_4 &= 3\pi^4/(20L^4), \\ \tilde{\mu}_2 &= \pi^2/(6L^2) + 1/12, \\ \tilde{\mu}_3 &= 2\zeta(3)/L^3. \end{aligned}$$

Let us now turn to the fluctuating components. The fundamental strip for (6) is $\Re(s) \in \langle -1, 0\rangle$. First of all, (7) and (9) lead to

$$w_1 = -\sum_{l\neq 0} \Gamma(\chi_l) e^{-2l\pi i \log n}/L. \tag{33}$$

Equations (11) and (12) lead, after the usual simplifications necessary to help Maple, to

$$\kappa_2 = -2\gamma w_1/L - w_1^2 + 2\sum_{l\neq 0}\Gamma(\chi_l)\psi(\chi_l)e^{-2l\pi i\log n}/L^2,$$

where ψ is the digamma function. $\tilde{m}_1$, $\tilde{\mu}_2$, w_1 and κ_2 are identical to the expressions given in Knuth [29], Kirschenhofer and Prodinger [25], (again after some simplifications). Next we obtain

$$\begin{aligned} \kappa_3 = &-6(w_1 + \gamma/L)\sum_{l\neq 0}\Gamma(\chi_l)\psi(\chi_l)e^{-2l\pi i\log n}/L^2 - 3\sum_{l\neq 0}\Gamma(\chi_l)\psi^2(\chi_l)e^{-2l\pi i\log n}/L^3 \\ &-3\sum_{l\neq 0}\Gamma(\chi_l)\psi(1,\chi_l)e^{-2l\pi i\log n}/L^3 + 2w_1^3 + (3\gamma^2 - \pi^2/2)w_1/L^2 + 6\gamma w_1^2/L, \end{aligned}$$

where $\psi(1,x)$ is the trigamma function.

Fluctuations in the higher moments.

The ubiquitous fluctuating functions are usually written as a Fourier series with mean zero (coefficient with index 0 missing). However, looking at w_1^2, w_1^3, ..., these functions do not have mean zero, even though w_1 has. In the following, we demonstrate how the mean can be extracted and the Fourier coefficients can be explicitly expressed. We show this for the second power, where the computations had appeared already elsewhere, but then deal with the third power; these computations are new (and quite involved).

The following approach ("residue calculus") to evaluate $[\delta^2]_0$ (which is the constant term[1] in the Fourier expansion of the square of the periodic function $\delta(x)$) seems to be the most appropriate one. See the survey [50]. We start with the following example:

$$\delta(x) = \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2\pi ikx} = -w_1(x).$$

Find a function $F(z)$ so that $[\delta^2]_0$ is (apart from a few extra terms) the sum of the residues along the imaginary axis. Here, take

$$F(z) = \frac{L}{2^z-1}\Gamma(-z)\Gamma(z).$$

For convenience, we define

$$\int_{(x)} := \int_{x-i\infty}^{x+i\infty}.$$

If we set

$$I_1 = \frac{1}{2\pi i}\int_{(1/2)} F(z)dz,$$

then by shifting and collecting residues,

$$I_1 = \frac{1}{2\pi i}\int_{(-1/2)} F(z)dz + \sum_{k\neq 0}\Gamma(-\chi_k)\Gamma(\chi_k) - \frac{\pi^2}{6} - \frac{L^2}{12}.$$

What happens here is often called *closing the box*, see e. g. [53, 44]. One integrates along a rectangle with corners $\pm\frac{1}{2}\pm iM$. One can evaluate it by collecting the residues inside the rectangle. And one can let the parameter M go to infinity. In this type of problems, the integrals along the horizontal lines disappear, and we can express one integral along a vertical line by an integral along another vertical line, plus a few residues.

Now one writes[2]

$$\frac{1}{e^z-1} = -1 - \frac{1}{e^{-z}-1}$$

and gets, by a simple change of variable $z := -z$,

$$I_1 = -\frac{L}{2\pi i}\int_{(-1/2)}\Gamma(-z)\Gamma(z)dz - I_1 + \sum_{k\neq 0}\Gamma(-\chi_k)\Gamma(\chi_k) - \frac{\pi^2}{6} - \frac{L^2}{12}.$$

The integral

$$I_2 = -\frac{L}{2\pi i}\int_{(-1/2)}\Gamma(-z)\Gamma(z)dz$$

[1] We will use notations like $[\delta^2]_k$ to indicate the kth Fourier coefficient.

[2] As S. Janson has pointed out, it is not (always) necessary to take the term '-1' out, as the resulting series (over residues on the left) would still (though not absolutely) converge. However, over the years, we encountered similar problems, where one would get series like $1-2+3-4+\cdots$, so it is just "safer" to proceed as we decided to do here; compare the survey [50].

can be computed by collecting the residues left to the line $\Re z = -\frac{1}{2}$, viz.

$$I_2 = -\frac{L}{2\pi i}\int_{(-1/2)} \Gamma(-z)\Gamma(z)dz = -L\sum_{l\ge 1}\frac{(-1)^l}{l!}(l-1)! = L^2.$$

Altogether we have

$$2I_1 = L^2 + \sum_{k\neq 0}\Gamma(-\chi_k)\Gamma(\chi_k) - \frac{\pi^2}{6} - \frac{L^2}{12}.$$

On the other hand, integral I_1 is also the sum of the negative residues right of the line $\Re z = \frac{1}{2}$, i. e.,

$$I_1 = -L\sum_{l\ge 1}\frac{(-1)^l}{l!(2^l-1)}(l-1)! = -L\sum_{l\ge 1}\frac{(-1)^l}{l(2^l-1)}.$$

Combining these results, we get

$$-2L\sum_{l\ge 1}\frac{(-1)^l}{l(2^l-1)} = L^2 + \sum_{k\neq 0}\Gamma(-\chi_k)\Gamma(\chi_k) - \frac{\pi^2}{6} - \frac{L^2}{12}.$$

This is the identity we wanted.

With not much more effort one can also compute the coefficients $[\delta^2]_k$, for $k \neq 0$. For this, one works with the function

$$F(z) = \frac{L}{2^z-1}\Gamma(-z-\chi_k)\Gamma(z).$$

One obtains

$$\begin{aligned} a_k := [\delta^2]_k &= \frac{1}{L^2}\sum_{j\neq 0,\,\neq k}\Gamma(-\chi_j)\Gamma(-\chi_k+\chi_j) \\ &= \frac{2}{L}\sum_{l\ge 1}\frac{(-1)^l\Gamma(-\chi_k+l)}{l!(2^l-1)} + \frac{2}{L^2}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big). \end{aligned}$$

Now we want the (Fourier-)coefficients of the third power of $\delta(x)$. We start with

$$\begin{aligned} [\delta^3(x)]_0 &= \frac{1}{L}\sum_{k\neq 0} a_k\Gamma(\chi_k) \\ &= \frac{2}{L^3}\sum_{k\neq 0}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big)\Gamma(\chi_k) + \frac{2}{L^2}\sum_{l\ge 1}\frac{(-1)^l}{l!(2^l-1)}\sum_{k\neq 0}\Gamma(-\chi_k+l)\Gamma(\chi_k). \end{aligned} \tag{34}$$

Let us first consider

$$\Xi := \sum_{k\neq 0}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big)\Gamma(\chi_k).$$

Define, as usually, the function

$$F(z) = \frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)$$

and set

$$I := \frac{1}{2\pi i}\int_{(1/2)}\frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)dz.$$

Then

$$
\begin{aligned}
I &= \frac{1}{2\pi i}\int_{(-1/2)} \frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)dz \\
&\quad + \sum_{k\neq 0}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big)\Gamma(\chi_k)+\zeta(3) \\
&= \Xi+\zeta(3)+\frac{L}{2\pi i}\int_{(-1/2)} \Big[-1-\frac{1}{2^{-z}-1}\Big]\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)dz \\
&= \Xi+\zeta(3)-\frac{L}{2\pi i}\int_{(-1/2)} \Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)dz \\
&\quad -\frac{L}{2\pi i}\int_{(-1/2)} \frac{1}{2^{-z}-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)dz \\
&= \Xi+\zeta(3)-L\sum_{l\ge 1}\frac{(-1)^l H_{l-1}}{l}-L\sum_{l\ge 1}\frac{(-1)^l H_{l-1}}{(2^l-1)l} \\
&= \frac{1}{2\pi i}\int_{(1/2)} \frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z)dz \\
&= -L\sum_{l\ge 1}\frac{1}{2^l-1}\frac{(-1)^{l-1}H_{l-1}}{l}+L^2\sum_{l\ge 1}\frac{2^l}{(2^l-1)^2}\frac{(-1)^{l-1}}{l}.
\end{aligned}
$$

From this we find

$$
\Xi = -\zeta(3)+\frac{L^3}{2}-2L\sum_{l\ge 1}\frac{(-1)^{l-1}H_{l-1}}{l(2^l-1)}+L^2\sum_{l\ge 1}\frac{(-1)^{l-1}2^l}{l(2^l-1)^2}.
$$

(Note that, since this is a tiny quantity, we find an approximation of $\zeta(3)$ in this way.)

For our application in (34), it must come with a factor $2/L^3$:

$$
1-\frac{2\zeta(3)}{L^3}-\frac{4}{L^2}\sum_{l\ge 1}\frac{(-1)^{l-1}H_{l-1}}{l(2^l-1)}+\frac{2}{L}\sum_{l\ge 1}\frac{(-1)^{l-1}2^l}{l(2^l-1)^2}.
$$

Now let us concentrate on the other term in (34). For $l\ge 1$, look at

$$
\Xi = \sum_{k\neq 0}\Gamma(-\chi_k+l)\Gamma(\chi_k).
$$

Take

$$
F = \frac{L}{2^z-1}\Gamma(l-z)\Gamma(z)
$$

and

$$
I = \frac{1}{2\pi i}\int_{(1/2)} \frac{L}{2^z-1}\Gamma(l-z)\Gamma(z)dz.
$$

Then the usual computation gives

$$
\begin{aligned}
I &= \Xi - (l-1)!H_{l-1} - (l-1)!\frac{L}{2} \\
&\quad - \frac{L}{2\pi i}\int_{(-1/2)} \Gamma(l-z)\Gamma(z)dz - \frac{L}{2\pi i}\int_{(-1/2)} \frac{1}{2^{-z}-1}\Gamma(l-z)\Gamma(z)dz \\
&= \Xi - (l-1)!H_{l-1} - (l-1)!\frac{L}{2} \\
&\quad - L\sum_{j\geq l}(-1)^{j-l}\frac{(j-1)!}{(j-l)!} - L\sum_{j\geq 1}\frac{1}{2^j-1}\frac{(l+j-1)!(-1)^j}{j!} \\
&= \Xi - (l-1)!H_{l-1} - (l-1)!\frac{L}{2} \\
&\quad - L(l-1)!\sum_{j\geq 0}(-1)^j\binom{j+l-1}{j} - L(l-1)!\sum_{j\geq 1}\frac{(-1)^j}{2^j-1}\binom{l+j-1}{j} \\
&= \Xi - (l-1)!H_{l-1} - (l-1)!\frac{L}{2} \\
&\quad - L(l-1)!2^{-l} - L(l-1)!\sum_{j\geq 1}\frac{(-1)^j}{2^j-1}\binom{l+j-1}{j}.
\end{aligned}
$$

But also

$$
I = L\sum_{j\geq l}\frac{1}{2^j-1}(-1)^{j-l}\frac{(j-1)!}{(j-l)!} = L(l-1)!\sum_{j\geq l}\frac{1}{2^j-1}(-1)^{j-l}\binom{j-1}{l-1},
$$

and so

$$
\begin{aligned}
\Xi &= (l-1)!H_{l-1} + (l-1)!\frac{L}{2} + L(l-1)!2^{-l} \\
&\quad + L(l-1)!\sum_{j\geq 1}\frac{(-1)^j}{2^j-1}\binom{l+j-1}{j} + L(l-1)!\sum_{j\geq l}\frac{1}{2^j-1}(-1)^{j-l}\binom{j-1}{l-1}.
\end{aligned}
$$

For our application, it comes with a factor $\frac{2}{L^2}\sum_{l\geq 1}\frac{(-1)^l}{l!(2^l-1)}$:

$$
\begin{aligned}
&\frac{2}{L^2}\sum_{l\geq 1}\frac{(-1)^l H_{l-1}}{l(2^l-1)} + \frac{1}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)} + \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)}2^{-l} \\
&+ \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)}\sum_{j\geq 1}\frac{(-1)^j}{2^j-1}\binom{l+j-1}{j} \\
&+ \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)}\sum_{j\geq l}\frac{1}{2^j-1}(-1)^{j-l}\binom{j-1}{l-1}.
\end{aligned}
$$

The grand summing up:

$$
\begin{aligned}
[\delta^3(x)]_0 &= 1 - \frac{2\zeta(3)}{L^3} - \frac{4}{L^2}\sum_{l\geq 1}\frac{(-1)^{l-1}H_{l-1}}{l(2^l-1)} + \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^{l-1}2^l}{l(2^l-1)^2} \\
&+ \frac{2}{L^2}\sum_{l\geq 1}\frac{(-1)^l H_{l-1}}{l(2^l-1)} + \frac{1}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)} + \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)}2^{-l} \\
&+ \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)}\sum_{j\geq 1}\frac{(-1)^j}{2^j-1}\binom{l+j-1}{j} \\
&+ \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l}{l(2^l-1)}\sum_{j\geq l}\frac{1}{2^j-1}(-1)^{j-l}\binom{j-1}{l-1}.
\end{aligned}
$$

After several simplifications this becomes

$$[\delta^3(x)]_0 = -1 - \frac{2\zeta(3)}{L^3} - \frac{1}{L}\sum_{l\ge1}\frac{(-1)^l}{l(2^l-1)} + \frac{6}{L^2}\sum_{l\ge1}\frac{(-1)^l H_{l-1}}{l(2^l-1)} + \frac{2\log 3}{L}$$
$$+ \frac{2}{L}\sum_{l,j\ge1}\frac{(-1)^{l+j}}{(l+j)(2^l-1)}\left[\frac{1}{2^j-1}+\frac{1}{2^{j+l}-1}\right]\binom{l+j}{j}.$$

This has been checked numerically and gives the tiny value 9.428177×10^{-25}.

Now attack $[\delta^3(x)]_j$, for $j\neq0$.

$$\begin{aligned}
[\delta^3(x)]_j &= \frac{1}{L}\sum_{k\neq j} a_k\Gamma(\chi_{k-j})\\
&= \frac{1}{L}\sum_{k\neq j,\neq0} a_k\Gamma(\chi_{k-j}) + \frac{1}{L}a_0\Gamma(-\chi_j)\\
&= \frac{1}{L}\sum_{k\neq j,\neq0}\left[\frac{2}{L^2}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big) + \frac{2}{L}\sum_{l\ge1}\frac{(-1)^l\Gamma(-\chi_k+l)}{l!(2^l-1)}\right]\Gamma(\chi_{k-j})\\
&\quad + \frac{1}{L}\left[\frac{\pi^2}{6L^2}+\frac{1}{12}-1-\frac{2}{L}\sum_{l\ge1}\frac{(-1)^{l-1}}{l(2^l-1)}\right]\Gamma(-\chi_j)\\
&= \frac{2}{L^3}\sum_{k\neq j,\neq0}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big)\Gamma(\chi_{k-j})\\
&\quad + \frac{2}{L^2}\sum_{k\neq j,\neq0}\sum_{l\ge1}\frac{(-1)^l\Gamma(-\chi_k+l)}{l!(2^l-1)}\Gamma(\chi_{k-j})\\
&\quad + \left[\frac{\pi^2}{6L^3}+\frac{1}{12L}-\frac{1}{L}-\frac{2}{L^2}\sum_{l\ge1}\frac{(-1)^{l-1}}{l(2^l-1)}\right]\Gamma(-\chi_j).
\end{aligned}$$

The third sum stands as it is, so let us start with the first one,

$$\Xi := \sum_{k\neq j,\neq0}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big)\Gamma(\chi_{k-j})$$

Consider

$$F(z) = \frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)$$

and set

$$I := \frac{1}{2\pi i}\int_{(1/2)}\frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)dz.$$

Then

$$
\begin{aligned}
I &= \frac{1}{2\pi i}\int_{(-1/2)} \frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)dz + \Xi \\
&\quad - \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 + L^2 - \pi^2\Big)\Gamma(-\chi_j) \\
&= -\frac{1}{2\pi i}\int_{(-1/2)} \frac{L}{2^{-z}-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)dz \\
&\quad - \frac{L}{2\pi i}\int_{(-1/2)} \Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)dz \\
&\quad + \Xi - \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 + L^2 - \pi^2\Big)\Gamma(-\chi_j) \\
&= -\frac{1}{2\pi i}\int_{(-1/2)} \frac{L}{2^{-z}-1}\Gamma(-z-\chi_j)\big(\psi(-z-\chi_j)+\gamma\big)\Gamma(z)dz \\
&\quad - \frac{L}{2\pi i}\int_{(-1/2)} \Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)dz \\
&\quad + \Xi - \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 + L^2 - \pi^2\Big)\Gamma(-\chi_j) \\
&= -L\sum_{l\geq 1} \frac{(-1)^l}{l!(2^l-1)}\Gamma(l-\chi_j)\big(\psi(l-\chi_j)+\gamma\big) \\
&\quad - L\sum_{l\geq 0} \frac{(-1)^l}{l!}\Gamma(l-\chi_j)(\psi(l-\chi_j)+\gamma) + L\Gamma(-\chi_j)(\psi(-\chi_j)+\gamma) \\
&\quad + \Xi - \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 + L^2 - \pi^2\Big)\Gamma(-\chi_j) \\
&= -L\sum_{l\geq 1} \frac{(-1)^l}{l!(2^l-1)}\Gamma(l-\chi_j)\big(\psi(l-\chi_j)+\gamma\big) \\
&\quad - L\sum_{l\geq 0} \frac{(-1)^l}{l!}\Gamma(l-\chi_j)\psi(l-\chi_j) - L\gamma\Gamma(-\chi_j) + L\Gamma(-\chi_j)(\psi(-\chi_j)+\gamma) \\
&\quad + \Xi - \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 + L^2 - \pi^2\Big)\Gamma(-\chi_j).
\end{aligned}
$$

One sum must be interpreted as follows:

$$
\begin{aligned}
\sum_{l\geq 0}\frac{(-1)^l}{l!}\Gamma(l-\chi_j)\psi(l-\chi_j) &= \sum_{l\geq 0}\frac{(-1)^l}{l!}\frac{d}{dx}\Gamma(x+l-\chi_j)\Big|_{x=0} \\
&= \frac{d}{dx}\sum_{l\geq 0}\frac{(-1)^l}{l!}\Gamma(x+l-\chi_j)\Big|_{x=0} \\
&= \frac{d}{dx}\Gamma(x-\chi_j)\sum_{l\geq 0}\binom{-x+\chi_j}{l}\Big|_{x=0} \\
&= \frac{d}{dx}\Gamma(x-\chi_j)2^{-x}\Big|_{x=0} \\
&= \Gamma(-\chi_j)\big(\psi(-\chi_j)-L\big).
\end{aligned}
$$

So

$$
\begin{aligned}
I &= -L\sum_{l\geq 1} \frac{(-1)^l}{l!(2^l-1)}\Gamma(l-\chi_j)\big(\psi(l-\chi_j)+\gamma\big) \\
&\quad + \Xi - \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 - 11L^2 - \pi^2\Big)\Gamma(-\chi_j).
\end{aligned}
$$

But also

$$
\begin{aligned}
I &= \frac{1}{2\pi i}\int_{(1/2)} \frac{L}{2^z-1}\Gamma(-z)\big(\psi(-z)+\gamma\big)\Gamma(z-\chi_j)dz \\
&= L\sum_{l\geq 1}\frac{(-1)^l\Gamma(l-\chi_j)}{l!}\left[\frac{\psi(l-\chi_j)+\gamma}{2^l-1}-\frac{2^lL}{(2^l-1)^2}\right].
\end{aligned}
$$

Combining these two expressions for I we find

$$
\begin{aligned}
\Xi &= 2L\sum_{l\geq 1}\frac{(-1)^l}{l!(2^l-1)}\Gamma(l-\chi_j)\big(\psi(l-\chi_j)+\gamma\big)-L^2\sum_{l\geq 1}\frac{(-1)^l\Gamma(l-\chi_j)}{l!}\frac{2^l}{(2^l-1)^2} \\
&+\frac{1}{12}\Big(18\Psi(1,-\chi_j)+18\big(\psi(-\chi_j)+\gamma\big)^2-11L^2-\pi^2\Big)\Gamma(-\chi_j).
\end{aligned}
$$

For our application, it gets a factor $2/L^3$, but now let us engage into the remaining term:

$$
\spadesuit=\sum_{k\neq 0,\neq j}\Gamma(l-\chi_k)\Gamma(\chi_k-\chi_j).
$$

Set

$$
F=\frac{L}{2^z-1}\Gamma(l-z)\Gamma(z-\chi_j),
$$

then

$$
\begin{aligned}
I &= \frac{1}{2\pi i}\int_{(1/2)} \frac{L}{2^z-1}\Gamma(l-z)\Gamma(z-\chi_j)dz \\
&= \frac{1}{2\pi i}\int_{(-1/2)} \frac{L}{2^z-1}\Gamma(l-z)\Gamma(z-\chi_j)dz \\
&\quad+\spadesuit+(l-1)!\Gamma(-\chi_j)-\Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big) \\
&= -\frac{L}{2\pi i}\int_{(-1/2)} \frac{1}{2^{-z}-1}\Gamma(l-z)\Gamma(z-\chi_j)dz \\
&\quad-\frac{L}{2\pi i}\int_{(-1/2)} \Gamma(l-z)\Gamma(z-\chi_j)dz \\
&\quad+\spadesuit+(l-1)!\Gamma(-\chi_j)-\Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big) \\
&= -\frac{L}{2\pi i}\int_{(-1/2)} \frac{1}{2^{-z}-1}\Gamma(l-\chi_j-z)\Gamma(z)dz \\
&\quad-\frac{L}{2\pi i}\int_{(-1/2)} \Gamma(l-z)\Gamma(z-\chi_j)dz \\
&\quad+\spadesuit+(l-1)!\Gamma(-\chi_j)-\Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big) \\
&= -L\sum_{h\geq 1}\Gamma(l+h-\chi_j)\frac{(-1)^h}{h!(2^h-1)} \\
&\quad-L\sum_{h\geq l}\frac{(-1)^{h-l}}{(h-l)!}\Gamma(h-\chi_j)+L\Gamma(L-\chi_j) \\
&\quad+\spadesuit+(l-1)!\Gamma(-\chi_j)-\Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big) \\
&= -L\sum_{h\geq 1}\Gamma(l+h-\chi_j)\frac{(-1)^h}{h!(2^h-1)} \\
&\quad-L\Gamma(l-\chi_j)2^{-l}+L\Gamma(l-\chi_j) \\
&\quad+\spadesuit+(l-1)!\Gamma(-\chi_j)-\Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big).
\end{aligned}
$$

But also

$$\begin{aligned}
I &= \frac{1}{2\pi i}\int_{(1/2)} \frac{L}{2^z-1}\Gamma(l-z)\Gamma(z-\chi_j)dz \\
&= L\sum_{h\ge l}\frac{(-1)^{h-l}}{(h-l)!(2^h-1)}\Gamma(h-\chi_j) \\
&= L\sum_{h\ge 0}\frac{(-1)^{h}}{h!(2^{h+l}-1)}\Gamma(h+l-\chi_j).
\end{aligned}$$

So

$$\begin{aligned}
&-L\sum_{h\ge 1}\Gamma(l+h-\chi_j)\frac{(-1)^h}{h!(2^h-1)} - L\Gamma(l-\chi_j)2^{-l} + L\Gamma(l-\chi_j) \\
&\quad + \spadesuit + (l-1)!\Gamma(-\chi_j) - \Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big) \\
&= L\sum_{h\ge 0}\frac{(-1)^{h}}{h!(2^{h+l}-1)}\Gamma(h+l-\chi_j),
\end{aligned}$$

or

$$\begin{aligned}
\spadesuit &= L\sum_{h\ge 0}\frac{(-1)^{h}}{h!(2^{h+l}-1)}\Gamma(h+l-\chi_j) + L\sum_{h\ge 1}\Gamma(l+h-\chi_j)\frac{(-1)^h}{h!(2^h-1)} \\
&\quad + L\Gamma(l-\chi_j)2^{-l} - L\Gamma(l-\chi_j) - (l-1)!\Gamma(-\chi_j) + \Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big).
\end{aligned}$$

For our application it gets a factor

$$\frac{2}{L^2}\sum_{l\ge 1}\frac{(-1)^l}{l!(2^l-1)}$$

The grand summing up:

$$\begin{aligned}
[\delta^3(x)]_j &= \frac{2}{L^3}\Big[2L\sum_{l\ge 1}\frac{(-1)^l}{l!(2^l-1)}\Gamma(l-\chi_j)\big(\psi(l-\chi_j)+\gamma\big) - L^2\sum_{l\ge 1}\frac{(-1)^l\Gamma(l-\chi_j)}{l!}\frac{2^l}{(2^l-1)^2} \\
&\quad + \frac{1}{12}\Big(18\Psi(1,-\chi_j) + 18\big(\psi(-\chi_j)+\gamma\big)^2 - 11L^2 - \pi^2\Big)\Gamma(-\chi_j)\Big] \\
&\quad + \frac{2}{L^2}\sum_{l\ge 1}\frac{(-1)^l}{l!(2^l-1)}\Big[L\sum_{h\ge 0}\frac{(-1)^{h}}{h!(2^{h+l}-1)}\Gamma(h+l-\chi_j) + L\sum_{h\ge 1}\Gamma(l+h-\chi_j)\frac{(-1)^h}{h!(2^h-1)} \\
&\quad + L\Gamma(l-\chi_j)2^{-l} - L\Gamma(l-\chi_j) - (l-1)!\Gamma(-\chi_j) + \Gamma(-\chi_j)\Big(\psi(-\chi_j)+\gamma+\frac{L}{2}\Big)\Big] \\
&\quad + \Big[\frac{\pi^2}{6L^3} + \frac{1}{12L} - \frac{1}{L} - \frac{2}{L^2}\sum_{l\ge 1}\frac{(-1)^{l-1}}{l(2^l-1)}\Big]\Gamma(-\chi_j).
\end{aligned}$$

We did not make an effort to simplify this ungainly expression any further.

5.2 The search cost in Digital Search Tree.

We have, from (23),

$$\begin{aligned}
F(\eta) &= 2^{\eta-1}\sum_{i=0}^{\infty} R(i)/2^i[\exp(-2^{-(\eta-i-1)})-1]/Q, \\
\varphi(\alpha) &= -1/QH_1(\alpha)\Gamma(1-\tilde{\alpha})/(\tilde{\alpha}+1),\ -L<\Re(\alpha)<L,
\end{aligned}$$

with

$$H_1(\alpha) := \sum_{i=0}^{\infty} e^{\alpha(i+1)} R(i).$$

But it is easy to check that the residue of $\phi^{(k)}(\alpha)$, $k > 0$ at $\alpha = -L$ is always 0.

This leads, with (A.1), to (the C_i are defined in Appendix A)

$$\begin{aligned}
m_1 &= (\gamma - 1)/L + 1 - C_1, \\
m_2 &= (\pi^2/6 + \gamma^2 - 2\gamma + 2)/L^2 + 2(\gamma - \gamma C_1 - 1 + C_1)/L + 1 - 3C_1 - C_2 + C_1^2, \\
\tilde{m}_1 &= 3/2 - C_1 + (\gamma - 1)/L, \\
\tilde{m}_2 &= (-2\gamma + \pi^2/6 + \gamma^2 + 2)/L^2 + (3\gamma - 3 - 2\gamma C_1 + 2C_1)/L + 7/3 - 4C_1 - C_2 + C_1^2, \\
\sigma^2 &= (1 + \pi^2/6)/L^2 - C_1 - C_2, \\
\mu_3 &= 2(\zeta(3) - 1)/L^3 - 2C_3 - 3C_2 - C_1, \\
\tilde{\mu}_2 &= (\pi^2/6 + 1)/L^2 + 1/12 - C_1 - C_2, \\
\tilde{\mu}_3 &= 2(\zeta(3) - 1)/L^3 - 2C_3 - 3C_2 - C_1.
\end{aligned}$$

Let us now turn to the fluctuating components. The fundamental strip for (6) is $\Re(s) \in \langle -1, 0 \rangle$. First of all, (7) and (9) lead to

$$w_1 = \sum_{l \neq 0} \Gamma(\chi_l - 1) e^{-2l\pi i \log n}/L.$$

Equations (11) and (12) lead, using for instance $\psi(z+1) = \psi(z) + 1/z$, to

$$\kappa_2 = -w_1^2 - 2(\gamma - 1)w_1/L - 2\sum_{l \neq 0} \Gamma(\chi_l - 1)\psi(\chi_l - 1) e^{-2l\pi i \log n}/L^2.$$

The quantities $\tilde{m}_1$, $\tilde{\mu}_2$, w_1 and κ_2 are identical to the expressions given in Kirschenhofer and Prodinger [25]. Next, we obtain, using $\psi(1, z+1) = \psi(1, z) - 1/z^2$,

$$\begin{aligned}
\kappa_3 = {} & w_1(4L^2 w_1^2 + 12 w_1 L\gamma - 12 w_1 L - 12\gamma + 6\gamma^2 - \pi^2)/(2L^2) \\
& + 6(\gamma - 1 + w_1 L) \sum_{l \neq 0} \Gamma(\chi_l - 1)\psi(\chi_l - 1) e^{-2l\pi i \log n}/L^3 \\
& + 3\sum_{l \neq 0} \Gamma(\chi_l - 1)\psi^2(\chi_l - 1) e^{-2l\pi i \log n}/L^3 + 3\sum_{l \neq 0} \Gamma(\chi_l - 1)\psi(1, \chi_l - 1) e^{-2l\pi i \log n}/L^3.
\end{aligned}$$

5.3 The duration of the Leader Election Algorithm.

We have, from (25),

$$\begin{aligned}
F(\eta) &= 2^{-\eta} \sum_{i=1}^{\infty} \exp(-i 2^{-\eta}), \\
\varphi(\alpha) &= -\alpha \zeta(1 - \tilde{\alpha})\Gamma(1 - \tilde{\alpha})/L, \ \Re(\alpha) < L.
\end{aligned}$$

This leads, with

$$\gamma_k := \lim_{n \to \infty} \left[\sum_{i=1}^{n} \frac{(\ln i)^k}{i} - \frac{(\ln n)^{k+1}}{k+1} \right],$$

also called the Stieltjes constants, to

$$\begin{aligned}
m_1 &= 0,\\
m_2 &= (\pi^2/6-\gamma^2-2\gamma_1)/L^2,\\
m_3 &= (-2\gamma^3-6\gamma\gamma_1+2\zeta(3)-3\gamma_2)/L^3,\\
\tilde{m}_1 &= 1/2,\\
\tilde{m}_2 &= 1/3+(-\gamma^2+\pi^2/6-2\gamma_1)/L^2,\\
\tilde{m}_3 &= 1/4+(-3/2\gamma^2+\pi^2/4-3\gamma_1)/L^2+(-2\gamma^3-6\gamma\gamma_1+2\zeta(3)-3\gamma_2)/L^3,\\
\sigma^2 &= -(\gamma^2-\pi^2/6+2\gamma_1)/L^2,\\
\mu_3 &= -(2\gamma^3+6\gamma\gamma_1-2\zeta(3)+3\gamma_2)/L^3,\\
\tilde{\mu}_2 &= -(\gamma^2-\pi^2/6+2\gamma_1)/L^2+1/12,\\
\tilde{\mu}_3 &= -(2\gamma^3+6\gamma\gamma_1-2\zeta(3)+3\gamma_2)/L^3.
\end{aligned}$$

Let us now turn to the fluctuating components. The fundamental strip for (6) is $\Re(s)\in\langle -1,0\rangle$. First of all, (7) and (9) lead to

$$w_1=-\sum_{l\neq 0}\Gamma(\chi_l)\chi_l\zeta(1+\chi_l)e^{-2l\pi i\log n}/L.$$

Equations (11) and (12) lead to

$$\begin{aligned}
\kappa_2 &= -w_1^2+2\sum_{l\neq 0}\Gamma(\chi_l)\chi_l\zeta(1+\chi_l)\psi(\chi_l)e^{-2l\pi i\log n}/L^2\\
&+ 2\sum_{l\neq 0}\Gamma(\chi_l)\zeta(1+\chi_l)e^{-2l\pi i\log n}/L^2+2\sum_{l\neq 0}\Gamma(\chi_l)\chi_l\zeta'(1+\chi_l)e^{-2l\pi i\log n}/L^2.
\end{aligned}$$

The terms $\tilde{m}_1$, $\tilde{\mu}_2$, w_1 and κ_2 are identical to the expressions given in Prodinger [48], Fill, Mahmoud and Szpankowski [5]. Next, we obtain,

$$\begin{aligned}
\kappa_3 &= 2w_1^3+(3\gamma^2w_1-\pi^2w_1/2+6\gamma_1w_1)/L^2+\sum_{l\neq 0}\Big[\Big(-6\Gamma(\chi_l)\chi_l\zeta'(1+\chi_l)w_1-6\Gamma(\chi_l)\zeta(1+\chi_l)w_1\\
&- 6\Gamma(\chi_l)\chi_l\zeta(1+\chi_l)\psi(\chi_l)w_1\Big)/L^2+\Big(-6\Gamma(\chi_l)\zeta'(1+\chi_l)-3\Gamma(\chi_l)\chi_l\zeta''(1+\chi_l)\\
&- 6\Gamma(\chi_l)\chi_l\zeta'(1+\chi_l)\psi(\chi_l)-6\Gamma(\chi_l)\zeta(1+\chi_l)\psi(\chi_l)\\
&- 3\Gamma(\chi_l)\chi_l\zeta(1+\chi_l)\psi^2(\chi_l)-3\Gamma(\chi_l)\chi_l\zeta(1+\chi_l)\psi(1,\chi_l)\Big)/L^3\Big]e^{-2l\pi i\log n}.
\end{aligned}$$

5.4 The Adaptive sampling algorithm.

We have, from (26),

$$F(r,\eta)=\sum_{l=0}^{\infty}\exp(-2^{-(\eta-l)})\frac{2^{-r(\eta-l)}}{r!}\left[1-\exp(-2^{-(\eta-l)})\sum_{k=0}^{b-r}\frac{2^{-k(\eta-l)}}{k!}\right].$$

We are interested in the moments of $Z:=r2^{X_n}/n$ and $W:=X_n-\log n$. So we first compute

$$\varphi(r,\alpha)=-\alpha\int_{-\infty}^{+\infty}e^{\alpha\eta}F(r,\eta)d\eta=-\alpha/(1-e^{\alpha})\left[\frac{\Gamma(r-\tilde{\alpha})}{Lr!}-\sum_{k=0}^{b-r}\frac{\Gamma(r+k-\tilde{\alpha})}{Lr!k!2^{r+k-\tilde{\alpha}}}\right],\ \Re(\alpha)<rL.$$

This leads to

$$\phi(r,\alpha)=\frac{\Gamma(r-\tilde{\alpha})}{Lr!}-\sum_{k=0}^{b-r}\frac{\Gamma(r+k-\tilde{\alpha})}{Lr!k!2^{r+k-\tilde{\alpha}}}.$$

The fundamental strip for (6) is $\Re(s)\in\langle -r,0\rangle$.

5.4.1 Atom at $Z = 0$.

We have (the RV R is defined in Section 4.4)

$$\Pi_0 := \Pr(Z = 0) = \Pr(R = 0).$$

This leads to

$$\Pi_0 = \phi(0,0) + w_0,$$

with

$$\phi(0,0) = \lim_{\alpha=0} \phi(0,\alpha) = 1 - \sum_{k=1}^{b} \frac{1}{Lk2^k},$$

which is already computed in Louchard [34]. Furthermore,

$$w_0 = \sum_{l\neq 0} \lim_{\alpha\to -L\chi_l} \phi(0,\alpha) e^{-2l\pi i \log n} = -\sum_{l\neq 0}\sum_{k=1}^{b} \frac{\Gamma(k+\chi_l)}{Lk!2^{k+\chi_l}} e^{-2l\pi i \log n}.$$

5.4.2

For Z, we must have $d \leq b$. Using limits (indeed, we must take care of singularities when $r \leq d$), (15) and (16) lead to

$$\begin{aligned}
\tilde{m}_{1,d} &= \sum_{r=1}^{b} r^d \lim_{\alpha\to Ld} \phi(r,\alpha), \\
w_{1,d} &= \sum_{r=1}^{b}\sum_{l\neq 0} r^d \lim_{\alpha\to L(d-\chi_l)} \phi(r,\alpha) e^{-2l\pi i \log n}.
\end{aligned}$$

Note that $w_{1,1} = 0$; $\tilde{m}_{1,d}, w_{1,d}$ are given, for $d = 1, 2$ in Flajolet [8]. The dominant part $\tilde{m}_{1,d}$ was already computed in Louchard [34]:

$$\tilde{m}_{1,d} = 1 + \frac{1}{L}\sum_{i=1}^{d} \frac{i^d(-1)^{d-i}B(d-i)}{i!(d-i)!} + \frac{1}{L}\sum_{k=1}^{d-1} \frac{\left\{ {d \atop k} \right\}}{d-k}\left[\frac{1}{(d-k)!} - \frac{(b-d)!}{(b-k)!}\right](1-2^{d-k}),$$

with

$$B(k) := \sum_{j=1}^{k} \frac{(-1)^j}{j}.$$

Now we can use (4) and (13).

To simplify $\tilde{m}_{1,d}$, we rewrite one of the sums a bit:

$$\begin{aligned}
\mathsf{SUM} &= \sum_{i=1}^{d} \frac{i^d(-1)^{d-i}B(d-i)}{i!(d-i)!} = \sum_{i=0}^{d-1} \frac{(d-i)^d(-1)^i}{i!(d-i)!} \sum_{1\leq j\leq i} \frac{(-1)^j}{j} \\
&= \sum_{\lambda=0}^{d} \left\{ {d \atop \lambda} \right\} \sum_{i=0}^{d-\lambda} \frac{(d-i)!(-1)^i}{(d-i-\lambda)!i!(d-i)!} \sum_{1\leq j\leq i} \frac{(-1)^j}{j} = \sum_{\lambda=0}^{d} \left\{ {d \atop \lambda} \right\} \sum_{i=0}^{d-\lambda} \frac{(-1)^i}{(d-i-\lambda)!i!} \sum_{1\leq j\leq i} \frac{(-1)^j}{j} \\
&= \sum_{\lambda=0}^{d} \left\{ {d \atop \lambda} \right\} \sum_{1\leq j\leq d-\lambda} \frac{(-1)^j}{j} \frac{1}{(d-\lambda)!} \sum_{i=j}^{d-\lambda} (-1)^i \binom{d-\lambda}{i} \\
&= \sum_{\lambda=0}^{d-1} \left\{ {d \atop \lambda} \right\} \sum_{1\leq j\leq d-\lambda} \frac{(-1)^j}{j} \frac{1}{(d-\lambda)!} (-1)^j \binom{d-\lambda-1}{j-1} \\
&= \sum_{\lambda=0}^{d-1} \left\{ {d \atop \lambda} \right\} \frac{1}{(d-\lambda)(d-\lambda)!} \sum_{1\leq j<d} \binom{d-\lambda}{j} = \sum_{\lambda=0}^{d-1} \left\{ {d \atop \lambda} \right\} \frac{2^{d-\lambda}-1}{(d-\lambda)(d-\lambda)!}.
\end{aligned}$$

Then

$$\tilde{m}_{1,d} = 1 + \frac{1}{L}\sum_{k=1}^{d-1}\left\{{d \atop k}\right\}\frac{2^{d-k}-1}{(d-k)(d-k)!} + \frac{1}{L}\sum_{k=1}^{d-1}\frac{\left\{{d \atop k}\right\}}{d-k}\left[\frac{1}{(d-k)!} - \frac{(b-d)!}{(b-k)!}\right](1-2^{d-k})$$
$$= 1 + \frac{(b-d)!}{L}\sum_{k=1}^{d-1}\left\{{d \atop k}\right\}\frac{2^{d-k}-1}{(d-k)(b-k)!}.$$

(The extra term 1 corresponds to the index $k=d$, if we interpret it as a limit.)

Finally, we can also work out an explicit form for $w_{1,d}$:

$$w_{1,d} = \sum_{l\neq 0}\sum_{r=1}^{b} r^d \lim_{\alpha\to d-\chi_l}\left[\frac{\Gamma(r-\alpha)}{Lr!} - \sum_{k=0}^{b-r}\frac{\Gamma(r+k-\alpha)}{Lr!k!2^{r+k-\alpha}}\right]e^{-2l\pi i\log n}.$$

It is nicer to work with falling factorials instead of powers, which will bring in Stirling numbers at the very end. So we look at (for $j\le d$):

$$\begin{aligned}
S &= \sum_{r=j}^{b} r^{\underline{j}}\left[\frac{\Gamma(r-d+\chi_l)}{Lr!} - \sum_{k=0}^{b-r}\frac{\Gamma(r+k-d+\chi_l)}{Lr!k!2^{r+k-d}}\right]\\
&= \frac{1}{L}\sum_{r=j}^{b}\left[\frac{\Gamma(r-d+\chi_l)}{(r-j)!} - \sum_{k=0}^{b-r}\frac{\Gamma(r+k-d+\chi_l)}{(r-j)!k!2^{r+k-d}}\right]\\
&= \frac{1}{L}\sum_{r=0}^{b-j}\frac{\Gamma(r+j-d+\chi_l)}{r!} - \frac{1}{L}\sum_{r=0}^{b-j}\sum_{k=0}^{b-j-r}\frac{\Gamma(r+k+j-d+\chi_l)}{r!k!2^{r+k+j-d}}\\
&= \frac{1}{L}\sum_{r=0}^{b-j}\frac{\Gamma(r+j-d+\chi_l)}{r!} - \frac{1}{L2^{j-d}}\sum_{h=0}^{b-j}\sum_{k=0}^{h}\frac{\Gamma(h+j-d+\chi_l)}{(h-k)!k!2^h}\\
&= \frac{1}{L}\sum_{r=0}^{b-j}\frac{\Gamma(r+j-d+\chi_l)}{r!} - \frac{1}{L2^{j-d}}\sum_{h=0}^{b-j}\frac{\Gamma(h+j-d+\chi_l)}{h!}\\
&= \frac{1}{L}\Big(1-\frac{1}{2^{j-d}}\Big)\Gamma(j-d+\chi_l)\sum_{r=0}^{b-j}\binom{r+j-1-d+\chi_l}{r}\\
&= \frac{1}{L}\Big(1-\frac{1}{2^{j-d}}\Big)\Gamma(j-d+\chi_l)\binom{b-d+\chi_l}{b-j}
\end{aligned}$$

and find eventually

$$w_{1,d} = \sum_{l\neq 0}\frac{1}{L}\sum_{j=1}^{d-1}\left\{{d \atop j}\right\}\Big(1-\frac{1}{2^{j-d}}\Big)\Gamma(j-d+\chi_l)\binom{b-d+\chi_l}{b-j}e^{-2l\pi i\log n}.$$

5.4.3

For W, (3) and (10) lead to

$$\begin{aligned}
\tilde{m}_j &= \sum_{r=0}^{b}\phi^{(j)}(r,0),\\
w_j &= \sum_{r=0}^{b}\sum_{l\neq 0}\phi^{(j)}(r,\alpha)\Big|_{\alpha=-L\chi_l}e^{-2l\pi i\log n}.
\end{aligned}$$

This gives

$$
\begin{aligned}
\tilde{m}_1 &= 3/2+(\gamma-L)/L+\sum_{k=1}^{b}\frac{\psi(k)-L}{L^2k2^k}+\sum_{r=1}^{b}\left[-\frac{1}{Lr2^r}+\frac{\Psi(r)(2^{-r}-1)}{L^2r}\right] \\
&+ \sum_{r=1}^{b}\sum_{k=1}^{b-r}\frac{-[L-\Psi(r+k)]2^{-(r+k)}\Gamma(r+k)}{L^2\Gamma(k+1)\Gamma(r+1)}, \\
\tilde{m}_2 &= 1/3+\gamma/L+\gamma^2/L^2+\pi^2/(6L^2) \\
&+ \sum_{r=1}^{b}\left[-\frac{1}{Lr2^r}+\frac{-\Psi(1,r)(2^{-r}-1)}{L^3r}+\frac{\Psi(r)[\Psi(r)(1-2^{-r})+2^{1-r}L]}{L^3r}\right] \\
&+ \sum_{k=1}^{b}\frac{-\Psi(1,k)-\Psi(k)^2+2L\Psi(k)}{L^3k2^k} \\
&- \sum_{r=1}^{b}\sum_{k=1}^{b-r}\frac{[-2\Psi(r+k)L+L^2+\Psi(1,r+k)+\Psi(r+k)^2]2^{-(r+k)}\Gamma(r+k)}{L^3\Gamma(k+1)\Gamma(r+1)}, \\
w_1 &= \sum_{r=0}^{b}\sum_{l\neq 0}\frac{[-L2^{-r}+\Psi(r+\chi_l)(2^{-r}-1)]\Gamma(r+\chi_l)}{L^2\Gamma(r+1)}e^{-2l\pi i\log n} \\
&+ \sum_{r=0}^{b}\sum_{k=1}^{b-r}\sum_{l\neq 0}\frac{2^{-(r+k)}[-L+\Psi(r+k+\chi_l)]\Gamma(r+k+\chi_l)}{L^2\Gamma(r+1)\Gamma(k+1)}e^{-2l\pi i\log n} \\
w_2 &= \sum_{l\neq 0}\left[\sum_{r=0}^{b}\frac{[-L^22^{-r}-\Psi(1,r+\chi_l)(2^{-r}-1)-\Psi(r+\chi_l)^2(2^{-r}-1)+\Psi(r+\chi_l)L2^{1-r}]\Gamma(r+\chi_l)}{L^3\Gamma(r+1)}\right. \\
&\left.-\sum_{r=0}^{b}\sum_{k=1}^{b-r}\frac{[L^2-2\Psi(r+k+\chi_l)+\Psi(1,r+k+\chi_l)+\Psi(r+k+\chi_l)^2]2^{-(r+k)}\Gamma(r+k+\chi_l)}{L^2\Gamma(r+1)\Gamma(k+1)}\right]e^{-2l\pi i\log n}.
\end{aligned}
$$

Moments of order 3 are rather long, so we don't give them here (but they could be computed with Maple).

Now we can use (4) and (13).

5.5 Approximate Counting.

We have, from (27),

$$
\begin{aligned}
F(\eta) &= -\sum_{k=0}^{\infty}2^k\frac{R(k)}{Q}\sum_{i=k}^{\infty}\exp(-2^{-\eta+i}), \\
\varphi(\alpha) &= -1/QH_2(\alpha)\Gamma(1-\tilde{\alpha})/(1-e^{\alpha}),\ \Re(\alpha)<L,
\end{aligned}
$$

and

$$
H_2(\alpha) := \sum_{k=0}^{\infty}e^{\alpha k}2^kR(k).
$$

With (B.3), we check that we have no singularity of $\phi^{(k)}$, $k>0$ at $\alpha=0$.

5.5.1

Let us first analyze the moments of $X_n - \log n$. This leads, with (B.2), to

$$\begin{aligned}
m_1 &= \gamma/L - C_1, \\
m_2 &= (\pi^2/6+\gamma^2)/L^2 - 2\gamma C_1/L - C_2 + C_1^2 - C_1, \\
\tilde{m}_1 &= 1/2 - C_1 + \gamma/L, \\
\tilde{m}_2 &= (\pi^2/6+\gamma^2)/L^2 + (\gamma - 2\gamma C_1)/L + 1/3 - 2C_1 - C_2 + C_1^2, \\
\sigma^2 &= \pi^2/(6L^2) - C_1 - C_2, \\
\mu_3 &= 2\zeta(3)/L^3 - 2C_3 - 3C_2 - C_1, \\
\tilde{\mu}_2 &= \pi^2/(6L^2) + 1/12 - C_1 - C_2, \\
\tilde{\mu}_3 &= 2\zeta(3)/L^3 - 2C_3 - 3C_2 - C_1.
\end{aligned}$$

Let us now turn to the fluctuating components. The fundamental strip for (6) is $\Re(s) \in \langle -1, 0\rangle$. First of all, (7) and (9) lead to

$$w_1 = -\sum_{l\neq 0} \Gamma(\chi_l) e^{-2l\pi i \log n}/L.$$

Equations (11) and (12) lead to

$$\kappa_2 = -w_1^2 - 2\gamma w_1/L + 2\sum_{l\neq 0}\Gamma(\chi_l)\psi(\chi_l)e^{-2l\pi i\log n}/L^2.$$

$\tilde{m}_1$, $\tilde{\mu}_2$ and w_1 are identical to the expression given in Flajolet [7]. Next, we obtain,

$$\begin{aligned}
\kappa_3 &= w_1(4L^2w_1^2 + 12w_1L\gamma + 6\gamma^2 - \pi^2)/(2L^2) \\
&- 6(\gamma + w_1 L)\sum_{l\neq 0}\Gamma(\chi_l)\psi(\chi_l)e^{-2l\pi i \log n}/L^3 \\
&- 3\sum_{l\neq 0}\Gamma(\chi_l)\psi^2(\chi_l)e^{-2l\pi i\log n}/L^3 - 3\sum_{l\neq 0}\Gamma(\chi_l)\psi(1,\chi_l)e^{-2l\pi i \log n}/L^3.
\end{aligned}$$

5.5.2

Now we turn to the moments of $2^{X_n}/n$. With (16) and (B.5) we obtain

$$\tilde{m}_{1,1} = 1,\ \tilde{m}_{1,2} = 3/2,\ \tilde{m}_{1,3} = 7/2,\ \tilde{\mu}_{1,2} = 1/2,\ \tilde{\mu}_{1,3} = 1.$$

$\tilde{m}_{1,1}$, $\tilde{\mu}_{1,2}$ correspond to the dominant terms of the expressions given in Flajolet [7].

With (B.4), it is immediate that the periodic components are identically 0.

5.6 LogLog counting.

We obtain, with (3),

$$\phi(\alpha) = \left[\Gamma(1-\tilde{\alpha}/m)\frac{e^{\alpha/m}-1}{\alpha/m}\right]^m,\ \Re(\alpha) < mL.$$

With (15), we must have $d < m$. The fundamental strip for s is $\Re(s) \in \langle -m, 0\rangle$. Equation (15) gives

$$\tilde{m}_{1,k} = H_3(k),$$

with

$$H_3(k) := \left[\Gamma(-k/m)(1-e^{kL/m})/L\right]^m.$$

Equation (17) gives

$$\begin{aligned}\tilde{\mu}_{1,2} &= -H_3(1)^2 + H_3(2),\\ \tilde{\mu}_{1,3} &= 2H_3(1)^3 + H_3(3) - 3H_3(2)H_3(1).\end{aligned}$$

Equation (16) gives

$$w_{1,1} = \sum_{l\neq 0} H_4(1,\chi_l)e^{-2l\pi i\log(n/m)},$$

with

$$H_4(k,\chi_l) := \Big[\Gamma((\chi_l - k)/m)(1-e^{L(k-\chi_l)/m})/L\Big]^m.$$

$\tilde{m}_{1,1}$, $\tilde{\mu}_{1,2}$ are given in Durand, Flajolet [4].

Equation (18) gives

$$\begin{aligned}\kappa_{1,2} &= -2H_3(1)w_{1,1} - w_{1,1}^2 + \sum_{l\neq 0} H_4(2,\chi_l)e^{-2l\pi i\log(n/m)},\\ \kappa_{1,3} &= 6H_3(1)^2 w_{1,1} + 6H_3(1)w_{1,1}^2 - 3H_2(2)w_{1,1} + 2w_{1,1}^3\\ &+ \sum_{l\neq 0}\{-3H_4(2,\chi_l)H_3(1) - 3H_4(2,\chi_l)w_{1,1} + H_4(3,\chi_l)\}e^{-2l\pi i\log(n/m)}.\end{aligned}$$

5.7 The number of Registers in arithmetic expressions evaluation.

We have, from (30),

$$f(\eta) = 2\sum_{k\text{ odd }>0}^{\infty}(2k^2 4^\eta - 1)e^{-k^2 4^\eta}.$$

It is indeed easier in this case to start from $f(\eta)$ and to compute

$$\phi(\alpha) = \int_{-\infty}^{+\infty} e^{\alpha\eta} f(\eta)d\eta.$$

Setting $y = 2^\eta$ and using Mellin, this gives

$$\phi(\alpha) = \Gamma(\xi/2)(\xi-1)\zeta(\xi)(1-2^{-\xi})/L\Big|_{\xi=\tilde{\alpha}}, \quad -\infty < \Re(\alpha) < \infty.$$

This Mellin transform is given in Louchard [31]. See also Biane and Yor [2], Biane, Pitman and Yor [1]. It is easy to check that there is no singularity at $\alpha = 0$, $\alpha = L$, $\alpha = -2kL$, $k > 0$.

The fundamental strip for (6) is $\Re(s) \in \langle -\infty, 0\rangle$.

5.7.1

Denote the number of registers by $r(n)$. Let us first analyze $r^*(n) := r(n) - \log_4 n$. After all computations (we must be careful about the presence of $\log_4$), we derive

$$\begin{aligned}\tilde{m}_1 &= (\ln\pi - 1 - \gamma/2)L + 1/2,\\ \tilde{m}_2 &= (-2\ln\pi - \gamma\ln\pi + \gamma + \pi^2/8 - 3/4\gamma^2 + \ln^2\pi - 2\gamma_1)/L^2 + (\ln\pi - 1 - \gamma/2)/L + 1/3,\\ \tilde{m}_3 &= (-3/8\pi^2 + 9/4\gamma^2 - 3/16\pi^2\gamma + 7/4\zeta(3) - 5/8\gamma^3 - 3\gamma_1\gamma + 6\gamma_1 - 3\gamma_2 - 3\ln^2\pi - 3/2\gamma\ln^2\pi\\ &+ 3\gamma\ln\pi - 9/4\gamma^2\ln\pi - 6\gamma_1\ln\pi + 3/8\pi^2\ln\pi + \ln^3\pi)/L^3\\ &+ (-9/8\gamma^2 + 3/2\gamma - 3\ln\pi + 3/2\ln^2\pi - 3\gamma_1 + 3/16\pi^2 - 3/2\gamma\ln\pi)/L^2\\ &+ (\ln\pi - 1 - \gamma/2)/L + 1/4,\\ \tilde{\mu}_2 &= (-1 + \pi^2/8 - \gamma^2 - 2\gamma_1)/L^2 + 1/12,\\ \tilde{\mu}_3 &= (-2 - 3\gamma_2 - 6\gamma_1\gamma - 2\gamma^3 + 7/4\zeta(3))/L^3.\end{aligned}$$

$\tilde{m}_1$ was computed in Flajolet and Prodinger [11], $\tilde{m}_2$ was computed in Louchard [31].

Let us now turn to the fluctuating components. First of all, (7) and (9) lead to

$$w_1 = -\sum_{l\neq 0} \Gamma(-\chi_l/2)\zeta(-\chi_l)(\chi_l+1)e^{-l\pi i \log n}/L.$$

Note that, in the exponent of the Fourier component, we have now $-l\pi i \log n$. Equations (11) and (12) lead to

$$\begin{aligned}
\kappa_2 &= -2w_1 - w_1^2 - \sum_{l\neq 0}\Gamma(-\chi_l/2)[2\zeta'(-\chi_l)(\chi_l+1) + \zeta(-\chi_l)\psi(-\chi_l/2)(\chi_l+1)\\
&+ 2\zeta(-\chi_l)\chi_l]e^{-l\pi i\log n}/L^2 + w_1(-2\ln\pi+\gamma)/L,\\
\kappa_3 &= 3/8w_1(16+10\gamma^2 - 8\gamma\ln\pi + 8\ln\pi^2 - 8\gamma L - 16L\\
&+ 16L\ln\pi + 8\gamma - 16\ln\pi + 8L^2 - \pi^2 + 16\gamma_1)/L^2\\
&+ 3w_1^2(2\ln\pi - \gamma - 2 + 2L)/L + 2w_1^3\\
&+ \sum_{l\neq 0}\Gamma(-\chi_l/2)\left[-3/4\zeta(-\chi_l)\psi(-\chi_l/2)^2(\chi_l+1) - 3\zeta'(-\chi_l)\psi(-\chi_l/2)(\chi_l+1)\right.\\
&+ 3\zeta(-\chi_l)\psi(-\chi_l/2) + 3/2\zeta(-\chi_l)\psi(-\chi_l/2)(\chi_l+1)(2Lw_1 - \gamma - 2 + 2L + 2\ln\pi)\\
&- 3/4\zeta(-\chi_l)\psi(1,-\chi_l/2)(\chi_l+1) - 3\zeta(-\chi_l)(2\ln\pi - \gamma - 2 + 2L + 2Lw_1)\\
&+ \left.\zeta'(-\chi_l)[6 + 3(2\ln\pi + 2Lw_1 - \gamma + 2L - 2)(\chi_l+1)] - 3\zeta''(-\chi_l)(\chi_l+1)\right] e^{-l\pi i\log n}/L^3.
\end{aligned}$$

w_1 is identical to the expressions given in Flajolet and Prodinger [11].

5.7.2

Now we turn to the size of the greatest perfect tree embedded in a binary tree corresponding to $r(n)$ registers. This is given by

$$2^{r(n)} = \sqrt{n}2^{r^*(n)},$$

and

$$Z := 2^{r^*(n)} = (4^{r(n)}/n)^{1/2}.$$

So we proceed as in Section 3.2, again taking into account the presence of $\log_4$. This gives, for the moments of Z

$$\begin{aligned}
\tilde{m}_{1,1} &= \sqrt{\pi}/(2L),\\
\tilde{m}_{1,2} &= \pi^2/(8L),\\
\tilde{m}_{1,3} &= 7/8\sqrt{\pi}\zeta(3)/L,\\
\tilde{\mu}_{1,2} &= -\pi/(4L^2) + \pi^2/(8L),\\
\tilde{\mu}_{1,3} &= \pi^{3/2}/(4L^3) + 7/8\sqrt{\pi}\zeta(3)/L - 3/16\pi^{5/2}/L^2,\\
w_{1,1} &= -1/2\sum_{l\neq 0}\Gamma(1/2-\chi_l/2)\chi_l\zeta(1-\chi_l)e^{-l\pi i\log n}/L,\\
\kappa_{1,2} &= -\sqrt{\pi}w_1/L - w_1^2 + 3/8\sum_{l\neq 0}\Gamma(-\chi_l/2)\chi_l\zeta(2-\chi_l)(\chi_l-1)e^{-l\pi i\log n}/L,\\
\kappa_{1,3} &= 3\sqrt{\pi}w_1^2/L + 3/2\pi w_1(1/L - \pi^2/4)/L + 2w_1^3\\
&+ \sum_{l\neq 0}\Gamma(-\chi_l/2)\zeta(2-\chi_l)\chi_l e^{-l\pi i\log n}9/16(2w_1L+\sqrt{\pi})(1-\chi_l)/L^2\\
&+ \sum_{l\neq 0}\Gamma(1/2-\chi_l/2)\zeta(3-\chi_l)\chi_l e^{-l\pi i\log n}(7/16\chi_l^2 + 7/8 - 21/16\chi_l).
\end{aligned}$$

$\tilde{m}_{1,1}$ was computed in Flajolet [6], $\tilde{m}_{1,2}$ was computed in Louchard [31], $w_{1,1}$ was given in Flajolet [6].

5.8 Distinctness of compositions of an integer.

The approach is different here. From Hwang and Yeh [17], Hitczenko and Louchard [14], we know that

$$\mathbb{E}(X_n) = \log n - 1/2 + \gamma/L + w_{1,T} + \mathcal{O}(1/n),$$

where $w_{1,T}$ is the Trie's w_1, Equation (33). Moreover, setting, as in the Trie case,

$$P(j) := (1 - 1/2^j)^n,$$

we also derive from [14], with asymptotic independence of the urns representing each integer,

$$\mathbb{E}(e^{\alpha X_n}) = \exp\left[\sum_{j=1}^{\infty} \ln\Big(1 + (e^\alpha - 1)(1 - P(j))\Big)\right] = \exp\left[\sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l}(e^\alpha - 1)^l V_l\right],$$

with

$$V_l := \sum_{j=1}^{\infty}(1 - P(j))^l.$$

We obtain

$$\begin{aligned} V_l &= \sum_{j=1}^{\infty}\left\{\sum_{k=0}^{l}(-1)^k \binom{l}{k} P(j)^k\right\} \\ &= \sum_{j=1}^{\infty}\left\{\sum_{k=0}^{l}(-1)^k \binom{l}{k} P(j)^k - \sum_{k=0}^{l}\binom{l}{k}(-1)^k\right\} \\ &= \sum_{k=1}^{l}\binom{l}{k}(-1)^{k+1} S_k, \text{ with} \\ S_k &:= \sum_{j=1}^{\infty}(1 - P(j)^k). \end{aligned}$$

Setting $\eta = j - \log(nk)$, this leads to

$$P(j)^k \sim F(\eta) = \exp(-2^{-\eta}),$$

and S_k is the mean of the RV with distribution function $P(j)^k$, minus 1 (as the sum starts here at $j = 1$). From Section 5.1, we immediately obtain ($\tilde{m}_{1,T}$ denotes here the Tries' $\tilde{m}_1$ and similarly for $w_{1,T}$)

$$S_k = \tilde{m}_{1,T} - 1 + \log(nk) + \beta_{1,k} + \mathcal{O}(1/n),$$

with

$$\begin{aligned} \tilde{m}_{1,T} - 1 &= -1/2 + \gamma/L, \\ \beta_{1,k} &= -\sum_{l \neq 0} \Gamma(\chi_l) e^{-2l\pi i \log(nk)}/L. \end{aligned}$$

Note that $\beta_{1,4} = \beta_{1,2} = \beta_{1,1} = w_{1,T}$. This gives

$$\begin{aligned} V_l &= \log n - 1/2 + \gamma/L + B_l + \beta_l, \text{ with} \\ B_l &:= \sum_{k=2}^{l}\binom{l}{k}(-1)^{k+1}\log k, \\ \beta_l &= \sum_{k=1}^{l}\binom{l}{k}(-1)^{k+1}\beta_{1,k}, \\ \beta_1 &= \beta_{1,1} = w_{1,T}, \end{aligned}$$

and, finally

$$\begin{aligned}
&\phi_p(\alpha)e^{\alpha\log n} = \mathbb{E}(e^{\alpha X_n})\\
&= \exp\left[\alpha(\log n - 1/2 + \gamma/L) + \sum_{l=2}^{\infty}\frac{(-1)^{l+1}}{l}(e^\alpha - 1)^l B_l + \sum_{l=1}^{\infty}\frac{(-1)^{l+1}}{l}(e^\alpha - 1)^l \beta_l + \mathcal{O}(1/n)\right].
\end{aligned}$$

From this, we derive

$$\begin{aligned}
\tilde{m}_1 &= \gamma/L - 1/2,\\
\tilde{m}_2 &= 5/4 - \gamma/L + \gamma^2/L^2,\\
\tilde{\mu}_2 &= 1,\\
\tilde{\mu}_3 &= -3 + 2\log 3,\\
\kappa_2 &= 0,\\
\kappa_3 &= -2w_1 + 2\beta_{1,3}.
\end{aligned}$$

Note that the variance has no periodic component. $\tilde{m}_l$, $\tilde{\mu}_l$ were given in [14].

For interest, we report about a direct approach, that requires to prove directly that the periodic components vanish. This is taken from [51].

The variance of the insertion cost of a random Patricia trie constructed from n random data was computed in [25] as

$$\frac{\pi^2}{6L^2} + \frac{1}{12} - \frac{2}{L}\sum_{l\ge1}\frac{(-1)^{l-1}}{l(2^l-1)} + \sigma_V(\log_2 n) + o(1),$$

with[3]

$$\begin{aligned}
\sigma_V(x) = &-\frac{2}{L^2}\sum_{k\neq0}(1+\chi_k)\Gamma(-1-\chi_k)\big(\psi(-1-\chi_k)+\gamma\big)e^{2\pi ikx}\\
&-\frac{2}{L}\sum_{k\neq0}(1+\chi_k)\Gamma(-1-\chi_k)g(\chi_k)e^{2\pi ikx} - \big(\delta_E(x)\big)^2,
\end{aligned}$$

and

$$\begin{aligned}
\delta_E(x) &= \frac{1}{L}\sum_{k\neq0}(1+\chi_k)\Gamma(-1-\chi_k)e^{2\pi ikx},\\
g(x) &:= \sum_{l\ge1}\binom{x}{l}\frac{1}{2^l-1}.
\end{aligned}$$

For interest, the variance is computed from the recursion

$$K_n(u) = 2^{1-n}u\sum_{k=0}^{n}\binom{n}{k}K_k(u) - 2^{1-n}(u-1)K_n(u),\ n\ge2,\quad K_0(u)=K_1(u)=1$$

via

$$\frac{K_n''(1)}{n} + \frac{K_n'(1)}{n} - \left(\frac{K_n'(1)}{n}\right)^2.$$

[3] There is a typo; the paper [24] contains the correct version. This paper discusses M-ary Patricia tries, and only for the instance $M=2$ (binary Patricia tries) do the cancellations occur.

These results can be rewritten, using $\Gamma(x+1) = x\Gamma(x)$ and $\psi(x+1) = \psi(x) + \frac{1}{x}$:

$$\sigma_V(x) = \frac{2}{L^2}\sum_{k\neq 0}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big)e^{2\pi ikx} + \frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)g(\chi_k)e^{2\pi ikx} - \big(\delta_E(x)\big)^2,$$

and

$$\delta_E(x) = -\frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2\pi ikx}.$$

Now if we look at the coefficient of $e^{2\pi ikx}$, for $k \neq 0$, we find

$$\frac{2}{L^2}\Gamma(-\chi_k)\big(\psi(-\chi_k)+\gamma\big) - \frac{2}{L}\Gamma(-\chi_k)g(\chi_k) - \frac{1}{L^2}\sum_{j\neq 0,\neq k}\Gamma(-\chi_j)\Gamma(-\chi_{k-j}).$$

According to the analysis in the previous sections, these coefficients are all equal to zero.

Patricia tries offered a surprise in 1986 when it turned out that the constant

$$\frac{\pi^2}{6L^2} + \frac{1}{12} - \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^{l-1}}{l(2^l-1)}$$

is just 1.00000000000*1237*... . That was nicely explained by Johannes Schoißengeier [27].

A similar phenomenon [51] occurs for words $x_1 \ldots x_n$, where the letters follow (independent) geometric random variables X with $\Pr\{X = i\} = 2^{-i}$ for $i = 1, 2, 3, \ldots$. Then the variance of the parameter "number of different letters in a random word of length n" is $1 + o(1)$. The cancellation is equivalent to the cancellation in the Patricia instance.

5.9 The Probabilistic Counting and first empty part in Compositions of integers.

We have, from (32),

$$F(\eta) = 1 - \prod_{k=0}^{\infty}[1 - \exp(-e^{-L(\eta-k)})].$$

But as

$$\sum_{j=0}^{\infty}(-1)^{\nu(j)}q^j = \prod_{k=0}^{\infty}(1 - q^{2^k}),$$

we obtain

$$F(\eta) = 1 - \sum_{l=0}^{\infty}(-1)^{\nu(l)}\exp(-le^{-L\eta}),$$

which is the expression Flajolet and Martin used in [10]. This leads to

$$\varphi(\alpha) = -N(-\tilde{\alpha})\Gamma(1-\tilde{\alpha}),\ \Re(\alpha) < L.$$

where $N(s)$ is the analytic continuation of

$$\sum_{j\geq 1}(-1)^{\nu(j)}/j^s.$$

We will need the values of the derivatives of $N(s)$. As shown in [10], we can group the terms of $N(s)$ together. For instance, we can use the fact that

$$\begin{aligned}
\nu(8j) &= \nu(j),\\
\nu(8j+1) &= \nu(8j+2) = \nu(8j+4) = \nu(j)+1,\\
\nu(8j+3) &= \nu(8j+5) = \nu(8j+6) = \nu(j)+2,\\
\nu(8j+7) &= \nu(j)+3.
\end{aligned}$$

We obtain

$$
\begin{aligned}
N(s) &= -1^{-s} - 2^{-s} + 3^{-s} - 4^{-s} + 5^{-s} + 6^{-s} - 7^{-s} \\
&+ \sum_{j=1}^{\infty} \frac{(-1)^{\nu(j)}}{(8j)^s} \left[1 - \frac{1}{(1+1/(8j))^s} - \frac{1}{(1+2/(8j))^s} + \frac{1}{(1+3/(8j))^s}\right. \\
&- \left.\frac{1}{(1+4/(8j))^s} + \frac{1}{(1+5/(8j))^s} + \frac{1}{(1+6/(8j))^s} - \frac{1}{(1+7/(8j))^s}\right].
\end{aligned}
$$

Differentiating the sum in $N(s)$ and setting $s = 0$ leads for instance to

$$
\sum_{j=1}^{\infty} (-1)^{\nu(j)} \ln\left(\frac{(8j+1)(8j+2)(8j+4)(8j+7)}{8j(8j+3)(8j+5)(8j+6)}\right).
$$

This leads to

$$
\begin{aligned}
N(0) &= -1, \\
N'(0) &= -.4874506\ldots, \\
N''(0) &= .8433214\ldots, \\
N'''(0) &= -.8683385\ldots
\end{aligned}
$$

5.9.1

Let us first analyze $X_n - \log n$. We derive

$$
\begin{aligned}
m_1 &= (\gamma + N'(0))/L, \\
m_2 &= (\pi^2/6 + \gamma^2 + 2N'(0)\gamma - N''(0))/L^2, \\
m_3 &= (2\zeta(3) + \pi^2\gamma/2 + \gamma^3 + N'(0)\pi^2/2 + 3N'(0)\gamma^2 - 3N''(0)\gamma + N'''(0))/L^3, \\
\tilde{m}_1 &= (\gamma + N'(0))/L + 1/2, \\
\tilde{m}_2 &= (\gamma + N'(0))/L + 1/3 + (\pi^2/6 + \gamma^2 + 2N'(0)\gamma - N''(0))/L^2, \\
\sigma^2 &= (-N'(0)^2 + \pi^2/6 - N''(0))/L^2, \\
\mu_3 &= (2N'(0)^3 + 3N''(0)N'(0) + N'''(0) + 2\zeta(3)/L^3, \\
\tilde{\mu}_2 &= (-6N'(0)^2 + \pi^2 - 6N''(0))/(6L^2) + 1/12, \\
\tilde{\mu}_3 &= (2N'(0)^3 + 3N''(0)N'(0) + N'''(0) + 2\zeta(3))/L^3.
\end{aligned}
$$

Let us now turn to the fluctuating components. The fundamental strip for (6) is $\Re(s) \in \langle -1, 0 \rangle$. First of all, (7) and (9) lead to

$$
w_1 = \sum_{l \neq 0} N(\chi_l)\Gamma(\chi_l)e^{-2l\pi i \log n}/L.
$$

Equations (11) and (12) lead to

$$
\kappa_2 = -w_1^2 - 2\sum_{l \neq 0} [N(\chi_l)(\gamma + N'(0)) + N'(\chi_l) + N(\chi_l)\psi(\chi_l)]\Gamma(\chi_l)e^{-2l\pi i \log n}/L^2.
$$

Next we obtain

$$
\begin{aligned}
\kappa_3 &= \sum_{l \neq 0} \left\{3\psi(1,\chi_l)N(\chi_l)/L^3 + \psi(\chi_l)[6N(\chi_l)w_1/L^2 + 6N'(\chi_l)/L^3 + 6N(\chi_l)(\gamma + N'(0))/L^3]\right. \\
&+ \left. 3N(\chi_l)\psi^2(\chi_l)/L^3 + 3(2(N'(0)+\gamma)N'(\chi_l) + N''(\chi_l))/L^3\right\} \Gamma(\chi_l)e^{-2l\pi i \log n} \\
&+ 6w_1 \sum_{l \neq 0} N'(\chi_l)\Gamma(\chi_l)e^{-2l\pi i \log n}/L^2 \\
&- 1/2w_1(6\gamma^2 + 12\gamma N'(0) - 6N''(0) + L^2 + \pi^2)/L^2 - 6w_1^2(\gamma + N'(0))/L - 4w_1^3.
\end{aligned}
$$

$\tilde{m}_1, \tilde{\mu}_2, w_1$ are given in [10] (with some discrepancies).

5.9.2

Now we analyze the moments $\mathbb{E}(\Delta^k)$ of $\Delta := m/n2^{\sum J_i/m}$ (m is again a design parameter). Equation (15) gives

$$\tilde{m}_{1,k} = H_5(k),$$

with

$$H_5(k) := \Big[N(-k/m)\Gamma(-k/m)(e^{kL/m} - 1)/L\Big]^m .$$

Equation (17) gives

$$\begin{aligned} \tilde{\mu}_{1,2} &= -H_5(1)^2 + H_5(2), \\ \tilde{\mu}_{1,3} &= 2H_5(1)^3 + H_5(3) - 3H_5(2)H_3(1). \end{aligned}$$

Equation (16) gives

$$w_{1,1} = \sum_{l\neq 0} H_6(1, \chi_l) e^{-2l\pi i \log(n/m)},$$

with

$$H_6(k, \chi_l) := \Big[N((\chi_l - k)/m)\Gamma((\chi_l - k)/m)(e^{L(k-\chi_l)/m} - 1)/L\Big]^m .$$

$\tilde{m}_{1,1}$, $\tilde{m}_{1,2}$ are given in Flajolet and Martin [10].

Equation (18) gives

$$\begin{aligned} \kappa_{1,2} &= -2H_5(1)w_{1,1} - w_{1,1}^2 + \sum_{l\neq 0} H_6(2, \chi_l) e^{-2l\pi i \log(n/m)}, \\ \kappa_{1,3} &= 6H_5(1)^2 w_{1,1} + 6H_5(1)w_{1,1}^2 - 3H_5(2)w_{1,1} + 2w_{1,1}^3 \\ &+ \sum_{l\neq 0} \{-3H_6(2, \chi_l)H_5(1) - 3H_6(2, \chi_l)w_{1,1} + H_6(3, \chi_l)\} e^{-2l\pi i \log(n/m)}. \end{aligned}$$

We see the complete analogy with LogLog.

6 Markov chains. The rate of convergence and moments.

6.1 Limiting Markov chains.

6.1.1 Column convex polyominoes.

From Louchard [35] and [36], we know that the transition matrix is asymptotically given by

$$\Pi(i, j) = \rho^j (i + j - 1)(aj + b)/(ai + b),$$

where ρ is the dominant root (with smallest module) of

$$4z^3 - 7z^2 + 5z - 1,$$

i.e.

$$\begin{aligned} \rho &:= -C_0^{1/3} + \frac{11}{144C_0^{1/3}} + 7/2, \text{ with } C_0 = \frac{71}{1728} + \frac{\sqrt{177}}{288}, \\ \rho &= .3119570552789534\ldots \end{aligned}$$

and

$$\begin{aligned} a &= (7\rho^2 - 13\rho + 5)/(41\rho^2 - 35\rho + 11) = .3993062361\ldots \\ b &= (17\rho^2 - 11\rho + 3)/(41\rho^2 - 35\rho + 11) = .3003468819\ldots \end{aligned}$$

The stationary measure $\boldsymbol{\pi}$ of $\boldsymbol{\Pi}$ is given by

$$\pi(i) = (ai + b)^2 \rho^i / a.$$

The chain is irreducible, ergodic and reversible.

Let $\mathcal{M}(n)$ denote the maximum thickness. From [36], we know that, setting

$$\begin{aligned}
\eta &= j - \Lambda_1(n), \text{ with} \\
\Lambda_1(n) &= \log[\Lambda_2(n)] / \log(\rho^{-1}), \\
\Lambda_2(n) &= \rho a n \mu_1 [\log n / \log(\rho^{-1})]^2 / (1 - \rho), \\
\mu_1 &= \frac{11\rho^2 - 9\rho + 5}{4(12\rho^2 - 4\rho + 5)}.
\end{aligned}$$

Note that $\Lambda_2(n)$ is non-linear in n. We have

$$\begin{aligned}
P(j) &:= \Pr[\mathcal{M}(n) \le j] \sim F(\eta), \eta = \mathcal{O}(1), \text{ with} \\
F(\eta) &= \exp[-\exp[-\eta \log(\rho^{-1})]].
\end{aligned}$$

$P(j)$ is asymptotically a periodic function of $\{\Lambda_1(n)\}$. Let us define $T_k :=$ the time necessary to obtain a thickness of at least length k. We set $h_i := \mathbb{E}_i(T_k)$, $h_i^{(2)} := \mathbb{E}_i(T_k^2)$. In our case, the Markov chain is reversible, so we obtain, from Keilson [20], that

$$\left| \Pr_i[T_k \ge x] \sim e^{-x/h_i} \right| \le C \left[\frac{h_i^{(2)}}{2h_i^2} - 1 \right]^{1/4},$$

for some constant C, uniformly on x. Now, by (C.9),

$$\frac{h_i^{(2)}}{2h_i^2} - 1 = \mathcal{O}(\rho^j).$$

We must bound $|P(j) - F(\eta)|$. Let $0 < \beta < 1$. We will consider three ranges

- for $j < \beta\Lambda_1(n)$, $P(j)$ is small. Indeed

$$P(j) \le \exp[-\Lambda_2(n)^{1-\beta}].$$

- for $\beta\Lambda_1(n) \le j < 2\Lambda_1(n)$, we have

$$|P(j) - F(\eta)| = \mathcal{O}[\exp[-\beta \log \Lambda_2(n)/4]] = \mathcal{O}[1/\Lambda_2(n)^{\beta/4}].$$

- for $j = 2\Lambda_1(n) + x/\log(\rho^{-1})$, we have

$$|P(j) - F(\eta)| = \mathcal{O}[\exp[-[2\log \Lambda_2(n) + x]/4]] = \mathcal{O}[e^{-x/4}/\Lambda_2(n)^{1/2}].$$

We can now proceed as in the Trie case. We omit the details.

6.1.2 Carlitz Compositions.

From Louchard and Prodinger [39], we know that the transition matrix is asymptotically given by

$$\Pi(i,j) = \rho^j (1 + \rho^i)/(1 + \rho^j), \quad j \ne i,$$

where ρ is the dominant root (with smallest modulus) of

$$1 - \sum_{j=1}^{\infty} \frac{z^j}{1 + z^j},$$

i.e.,

$$\rho = .571349793158088\ldots.$$

The stationary measure $\boldsymbol{\pi}$ of $\boldsymbol{\Pi}$ is given by

$$\pi(i) = -\frac{\rho^i}{(1+\rho^i)^2 h_w},$$

with

$$h(w,z) := 1 - \sum_{j=1}^{\infty} \frac{wz^j}{1+wz^j},\ h_w := \partial h/\partial w|_{w=1,z=\rho}.$$

The chain is irreducible, ergodic and reversible. Let $\mathcal{M}(n)$ denote the maximum part size. From [39], we know that, setting

$$\begin{aligned}
\eta &= j - \Lambda_1(n), \text{ with} \\
\Lambda_1(n) &= \log[\Lambda_2(n)]/\log(\rho^{-1}), \\
\Lambda_2(n) &= -n/[(1-\rho)h_z].
\end{aligned}$$

Note that $\Lambda_2(n)$ is now linear in n. We have

$$\begin{aligned}
P(j) &:= \Pr[\mathcal{M}(n) \le j) \sim F(\eta), \eta = \mathcal{O}(1), \text{ with} \\
F(\eta) &= \exp[-\exp[-\eta \log(\rho^{-1})]].
\end{aligned}$$

$P(j)$ is asymptotically a periodic function of $\{\log n\}$. We can now proceed exactly as in Section 6.1.1 (using now (D.15)). We omit the details.

6.2 Exact Markov chains.

6.2.1 Runs of geometrically distributed random variables.

From Louchard [37], we know that the transition matrix is exactly given by

$$\Pi(i,j) = p(j)/(1-p(i)), j \neq i, \text{ with } p(i) = pq^{i-1}.$$

The stationary measure $\boldsymbol{\pi}$ of $\boldsymbol{\Pi}$ is given by

$$\pi(i) = \alpha p(i)(1-p(i)), \text{ with } \alpha = (1+q)/(2q).$$

The chain is irreducible, ergodic and reversible. Let $\mathcal{M}(n)$ denote the maximum run length. From [37], we know that, setting

$$\begin{aligned}
\eta &= j - \Lambda_1(n), \text{ with} \\
\Lambda_1(n) &= \log[\Lambda_2(n)]/\log(p^{-1}), \\
\Lambda_2(n) &= npq.
\end{aligned}$$

Note that $\Lambda_2(n)$ is linear in n. We have

$$\begin{aligned}
P(j) &:= \Pr[\mathcal{M}(n) \le j) \sim F(\eta), \quad \eta = \mathcal{O}(1), \quad \text{with} \\
F(\eta) &= \exp[-\exp[-\eta \log(p^{-1})]].
\end{aligned}$$

$P(j)$ is asymptotically a periodic function of $\{\log n\}$. We can now proceed exactly as in Section 6.1.1. We use the results proved in Louchard [37]:

$$\begin{aligned}
&\left| h_i^{(2)}/(2h_i^2) - 1 \right| = \mathcal{O}(p^k), \\
&M(i,j) = [\![i=j]\!] - p(i)[\![i=j]\!] \\
&\qquad + \frac{1}{p(i)}[-\alpha[p(i)p(j) - p(i)^2 p(j) - p(i)p(j)^2] - p(i)p(j)h_{ww}\alpha^2/2](1-p(j)), \\
&h_{ww} = \frac{-2q(2q^2-q-1)}{(1+q)(q^2+q+1)}.
\end{aligned}$$

We omit the details.

6.2.2 Monotone runs. Runs of uniformly distributed integer random variables.

From Louchard [38], we know that the transition matrix is given by

$$\Pi(i,j) = \frac{\alpha^{i-1}}{s}[\alpha^{-s} - \alpha^{-j}], \quad \text{with} \quad \alpha := 1 - 1/s.$$

The stationary measure $\boldsymbol{\pi}$ of $\boldsymbol{\Pi}$ is given by

$$\pi(i) = 2\frac{s-i}{s(s-1)}.$$

The chain is not reversible, but it is proved in [38] that

$$|P(j) - F(\eta)| = \mathcal{O}(j^{s-2}/s^j). \tag{35}$$

Let $\mathcal{M}(n)$ denote the maximum run length. From [38], we know that, setting

$$\begin{aligned}
\eta &= j - \Lambda_1(n), \text{ with} \\
\Lambda_1(n) &= \log[\Lambda_2(n)]/\log s, \\
\Lambda_2(n) &= n[\log n/\log s]^{s-1}/(s^2(s-2)!).
\end{aligned}$$

Note that $\Lambda_2(n)$ is non-linear in n. We have

$$\begin{aligned}
P(j) &:= \Pr[\mathcal{M}(n) \leq j) \sim F(\eta), \eta = \mathcal{O}(1), \text{ with} \\
F(\eta) &= \exp[-\exp[-\eta \log s]].
\end{aligned}$$

$P(j)$ is asymptotically a periodic function of $\{\Lambda_1(n)\}$. Now we use (35). We know that j^{s-2}/s^j is decreasing for sufficiently large j.

We must bound $|P(j) - F(\eta)|$. Let $0 < \beta < 1$. We will consider three ranges

- for $j < \beta\Lambda_1(n) + (s-2)\log\log n/\log s$, $P(j)$ is small. Indeed

$$P(j) \leq \exp[-\Lambda_2(n)^{1-\beta} - (s-2)\log\log n] = \exp[-\Lambda_2(n)^{1-\beta}]/\log n^{s-2}.$$

- for $\beta\Lambda_1(n) + (s-2)\log\log n/\log s \leq j < 2\Lambda_1(n) + (s-2)\log\log n/\log s$, we have

$$|P(j) - F(\eta)| = \mathcal{O}[1/\Lambda_2(n)^\beta].$$

- for $j = 2\Lambda_1(n) + (s-2)\log\log n/\log s + x/\log s$, we have

$$|P(j) - F(\eta)| = \mathcal{O}[e^{-x}x^{s-2}/\Lambda_2(n)^2].$$

We can now proceed as in the Trie case. We omit the details.

7 Preliminary singularity analysis. The $M/G/1$ queue.

We are interested in the maximum size $\mathcal{M}(t)$ of a queue of length t. Let

$$\begin{aligned}
u_a &:= \text{ interarrival times, assumed to be i.i.d. exponential RV,}\\
1/\lambda &:= \text{ mean of } u_a,\\
u_b &:= \text{ service times, assumed to be i.i.d. RV,}\\
B(t) &:= \text{ DF of } u_b,\\
1/\mu &:= \text{ mean of } u_b,\\
\rho &= \lambda/\mu < 1,\\
\beta(s) &:= \int_0^\infty e^{-st} dB(t), s \geq 0,\\
\xi &:= \text{ the root greater than 1 and nearest to 1 of } \beta[(1-\xi)\lambda] = \xi,\\
\eta &= j - \log t/\log \xi - \log C_1/\log \xi, \text{ linear in } \log t,\\
C_2 &:= 1 - \int_0^\infty e^{(1-\xi)\lambda t} \lambda t dB(t),\\
C_1 &= \lambda(1-\rho)^2(\xi-1)/C_2,\\
F(\eta) &= \exp[-\exp[-\eta \log \xi]],\\
T_j &:= \text{ hitting time into state } j,\\
\varphi_j(s) &:= \mathbb{E}(e^{-sT_j}).
\end{aligned}$$

From Louchard, Kenyon and Schott [23], we know that, with $C_.$ some constants and $P_.(l)$ polynomials in l,

$$\varphi_l(s) \sim \frac{1 + \xi^{-l}[C_3 l + \mathcal{O}(1)] + \mathcal{O}(\xi^{-2l} P_1(l))}{1 - s/s^*}, \quad l \to \infty,$$

where s^* is the dominant root of the denominator of $\varphi_l(s)$, which is given by

$$s^* = \lambda(1-\rho)^2(\xi-1)/(\xi^{l-1}C_2)[1 + \xi^{-l}[C_4 l + \mathcal{O}(1)] + \mathcal{O}(\xi^{-2l} P_2(l))].$$

This leads to

$$P(j) := \Pr[\mathcal{M}(t) \leq j] \sim F(\eta), \quad \eta = \mathcal{O}(1),$$

and

$$|P(j) - F(\eta)| = \mathcal{O}(F(\eta) j/\xi^j).$$

$P(j)$ is asymptotically a periodic function of $\{\log t\}$. We can now proceed as in Section 6.2.2. The $G/M/1$ queue is also considered in [23] and can be analyzed in the same manner.

8 Conclusion.

Identifying the Gumbel distribution (or variants) in many problems from algorithms and discrete structures, an approximation at an early stage leads to simplified computations of moments. In this way, all moments can be (in principle) computed, in a fairly automatic fashion. This is in sharp contrast to many earlier papers, where even the expectation and variance are a challenge, as long as exact (and ungainly) expressions make the computations extremely heavy. Usually, there are periodic oscillations involved, and we demonstrate how to deal with them as well.

9 Acknowledgements.

We are grateful to S. Janson for helpful remarks on a preliminary version of this paper. The first author wishes to thank the The John Knopfmacher Centre for Applicable Analysis and Number Theory for its warm hospitality. In particular, Arnold's choice of restaurants was quite appreciated.

Appendices.

A Identities related to the Digital search tree.

Using a classical Euler identity, we will derive several summation formulae. This identity is

$$\prod_{k=0}^{\infty}(1+q^k z) = \sum_{i=0}^{\infty} z^i q^{i(i-1)/2} / \prod_{k=1}^{i}(1-q^k).$$

(for a simple proof, see Knuth [29, Ex 5.1.1-16]). Set

$$\begin{aligned} \Pi &:= \prod_{k=0}^{\infty}(1+q^k z), \\ \Sigma_k &:= (k-1)! \sum_{i=0}^{\infty} q^{ki}/(1+zq^i)^k. \end{aligned}$$

Setting also

$$C_k := \sum_{j=1}^{\infty} 1/(2^j-1)^k,$$

we easily check that, with $z=-1/2$, $q=1/2$ we derive

$$\Pi \to Q,\ \Sigma_1 \to 2C_1,\ \Sigma_2 \to 4C_2,\ \Sigma_3 \to 16C_3.$$

Now we want to compute sums of type

$$S_k := -\sum_{i=0}^{\infty}(i+1)^k R(i).$$

For instance, set $T_1 := z\Pi$. Differentiating (wrt z) leads to

$$T_2 := T_1' = \Pi(1+z\Sigma_1),$$

and setting $z=-1/2$, $q=1/2$, we obtain

$$S_1 = Q(1-C_1).$$

Similarly, set $T_3 := zT_2$. With the same procedure, Maple gives us all relations we need:

$$\begin{aligned} S_2 &= Q(1-3C_1-C_2+C_1^2), \\ S_3 &= -Q(-1+7C_1+6C_2-6C_1^2+2C_3-3C_1C_2+C_1^3). \end{aligned} \tag{A.1}$$

Also $S_0 = Q$; this is relation (A.3) in Louchard [32].

B Identities related to the Approximate counting.

Set

$$\begin{aligned} \Pi^* &:= \prod_{k=1}^{\infty}(1+q^k z), \\ \Pi &= (1+z)\Pi^*, \\ \Sigma_k &:= (k-1)! \sum_{i=1}^{\infty} q^{ki}/(1+zq^i)^k. \end{aligned}$$

We easily check that, with $z=-1$, $q=1/2$ we derive

$$\Pi^* \to Q,\ \Sigma_1 \to C_1,\ \Sigma_2 \to C_2,\ \Sigma_3 \to 2!C_3,\ \Sigma_4 \to 3!C_4.$$

with

$$C_k := \sum_{j=1}^{\infty} 1/(2^j-1)^k.$$

Now we want to compute sums of type

$$U_k := \sum_{i=0}^{\infty} i^k 2^i R(i).$$

We obtain

$$\begin{aligned}
\Pi^{*'} &= \Sigma_1 \Pi^*,\\
\Pi' &= \Pi^* + (1+z)\Sigma_1\Pi^*,
\end{aligned}$$

and setting $z=-1$, $q=1/2$, we derive $U_1 = Q$. Similarly, set $T_1 := z\Pi'$, etc. With the same procedure, we obtain:

$$\begin{aligned}
U_2 &= -Q(-1+2C_1),\\
U_3 &= Q(1-6C_1-3C_2+3C_1^2),\\
U_4 &= -Q(-1+14C_1+18C_2-18C_1^2+8C_3-12C_1C_2+4C_1^3).
\end{aligned} \tag{B.2}$$

Also

$$U_0 = 0, \tag{B.3}$$

this is Equation (21) in Flajolet [7]. More generally, setting $z=-2^k$, $q=1/2$ in Π, we derive

$$\sum_{i=0}^{\infty} 2^{(k+1)i} R(i) = 0. \tag{B.4}$$

Finally, we note that

$$H_2(\alpha) = -\left.\Pi\right|_{z=-\exp(\alpha),\ q=1/2}.$$

This leads to

$$\begin{aligned}
H_2(\alpha) &\sim -Q(\alpha-L)+\mathcal{O}((\alpha-L)^2),\\
H_2(\alpha) &\sim 3Q(\alpha-2L)+\mathcal{O}((\alpha-2L)^2),\\
H_2(\alpha) &\sim -21Q(\alpha-3L)+\mathcal{O}((\alpha-3L)^2).
\end{aligned} \tag{B.5}$$

C Column convex polyominoes.

C.1 Hitting times.

Set

$$\varphi_1(i,k) := \sum_{j=k}^{\infty} \Pi(i,j) = g_1(i,k)\varepsilon,$$

with

$$\varepsilon := \rho^k,\ g_1(i,k) = \frac{ak^2}{(1-\rho)(ai+b)}[1+\mathcal{O}(1/k)],\ k\to\infty.$$

Let us also define

$$H(k) := \sum \pi(i)\varphi_1(i,k) = G(k)\varepsilon,$$

with

$$G(k) := \sum_i \pi(i) g_1(i,k) = \frac{ak^2}{1-\rho}[1+\mathcal{O}(1/k)].$$

We compute also, for $j \geq k$,

$$\varphi_2(i,j,k) := \frac{\Pi(i,j)}{\varphi_1(i,k)} = \rho^{j-k}(1-\rho)j^2/k^2[1+\mathcal{O}(1/k)],$$

and

$$\sum_{j\geq k} \varphi_2(i,j,k) = 1.$$

Note that the dominant term is independent of i.

We obtain

$$\begin{aligned} h_i &= 1+\sum_{u<k} \Pi(i,u)h_u \\ &= 1+\sum_u \Pi(i,u)h_u - \sum_{u\geq k} \Pi(i,u)h_u \\ &= 1+\sum_u \Pi(i,u)h_u - \varphi_1(i,k)\sum_{u\geq k}\frac{\Pi(i,u)}{\varphi_1(i,k)}h_u. \end{aligned}$$

This leads to

$$\begin{aligned} D_i+\varepsilon\psi(i)+\varepsilon^2\chi(i) &= \varepsilon+\sum_u \Pi(i,u)[D(u)+\varepsilon\psi(u)+\varepsilon^2\chi(i)+\mathcal{O}(\varepsilon^3)]] \\ &- \varphi_1(i,k)\sum_{u\geq k}\frac{\Pi(i,u)}{\varphi_1(i,k)}[D(u)+\varepsilon\psi(u)+\mathcal{O}(\varepsilon^2)] \\ &= \varepsilon+\sum_u \Pi(i,u)[D(u)+\varepsilon\psi(u)+\varepsilon^2\chi(i)] \\ &- g_1(i,k)\varepsilon\sum_{u\geq k}\varphi_2(i,u,k)[D(u)+\varepsilon\psi(u)]+\mathcal{O}(\varepsilon^3). \end{aligned} \tag{C.6}$$

Identification of constant terms in (C.6) leads to $[\mathbf{I}-\mathbf{\Pi}]\mathbf{D}=0$, which confirms that $\mathbf{D}$ is independent of i. Now we premultiply (C.6) by $\boldsymbol{\pi}$. The ε term leads to

$$0 = 1 - DG(k),$$

or

$$D = 1/G(k) = \frac{1-\rho}{ak^2}[1+\mathcal{O}(1/k)].$$

Now the ε term of (C.6) leads to

$$\psi(i) = 1+\mathbf{\Pi}\boldsymbol{\psi}(i) - Dg_1(i,k),$$

or

$$[\mathbf{I}-\mathbf{\Pi}]\boldsymbol{\psi} = \boldsymbol{\delta},$$

with

$$\delta(i) = 1 - Dg_1(i,k).$$

This gives

$$\boldsymbol{\psi} = \mathbf{M}\boldsymbol{\delta}+\gamma_1, \quad \text{where} \quad \gamma_1 := \boldsymbol{\pi}\boldsymbol{\psi}$$

and $\mathbf{M} := \sum_{n\geq 0}(\mathbf{\Pi}^n - \mathbf{1}\times\boldsymbol{\pi})$. $\mathbf{M}$ is the Drazin inverse of $\mathbf{I}-\mathbf{\Pi}$. We refer to Campbell and Meyer [3] for a detailed definition and analysis of the Drazin inverse. We have $\mathbf{M} = \mathbf{Z} - \mathbf{1}\times\boldsymbol{\pi}$ where $\mathbf{Z} := [I - \mathbf{\Pi} + \mathbf{1}\times\boldsymbol{\pi}]^{-1} = \sum_{n\geq 0}[\mathbf{\Pi} - \mathbf{1}\times\boldsymbol{\pi}]^n$ is the potential used in Kemeny, Snell and Knapp [21]. First, by (C.11) (see this Appendix next subsection), we know that $\mathbf{M}\boldsymbol{\delta}$ is finite .

Next, to derive γ_1, we premultiply again (C.6) by $\boldsymbol{\pi}$. When $k\to\infty$, (apart from D, we neglect all $\mathcal{O}(1/k)$ error terms), we obtain

$$0 = \varepsilon - \varepsilon G(k)[D + \varepsilon\sum_{u\geq k}\varphi_2(i,u,k)\psi(u)] + \mathcal{O}(\varepsilon^3).$$

The ε term leads of course again to D. The ε^2 term leads to

$$0 = -G(k)\sum_{u\geq k}\varphi_2(i,u,k)\psi(u),$$

which allows the determination of γ_1.

We must now compute the first terms in the expansion of $h_i^{(2)} := \mathbb{E}_i(T_k^2)$. We set

$$h_i^{(2)} = \frac{D^{(2)}(i)}{\varepsilon^2} + \frac{\psi^{(2)}(i)}{\varepsilon} + \chi^{(2)}(i) + \mathcal{O}(\varepsilon).$$

We derive

$$\begin{aligned}
h_i^{(2)} &= \mathbb{E}[(1+T_i)^2] = 1 + 2\sum_{u<k}\Pi(i,u)h_u + \sum_{u<k}\Pi(i,u)h_u^{(2)} \\
&= 1 + 2\sum_u \Pi(i,u)h_u + \sum_u \Pi(i,u)h_u^{(2)} \\
&- 2\sum_{u\geq k}\Pi(i,u)h_u - \sum_{u\geq k}\Pi(i,u)h_u^{(2)} \\
&= 1 + 2\sum_u \Pi(i,u)[D/\varepsilon + \psi(u) + \varepsilon\chi(u)] \\
&+ \sum_u \Pi(i,u)[D^{(2)}(u)/\varepsilon^2 + \psi^{(2)}(u)/\varepsilon + \chi^{(2)}(u)] \\
&- 2g_1(i,k)\varepsilon[D/\varepsilon + \sum_{u\geq k}\varphi_2(i,u,k)\psi(u)] \\
&- g_1(i,k)\varepsilon\sum_{u\geq k}\varphi_2(i,u,k)[D^{(2)}(u)/\varepsilon^2 + \psi^{(2)}(u)/\varepsilon] + \mathcal{O}(\varepsilon). \qquad \text{(C.7)}
\end{aligned}$$

Multiply (C.7) by ε^2. The constant term leads to $[\mathbf{I}-\mathbf{\Pi}]\mathbf{D^{(2)}} = 0$, which shows that $\mathbf{D^{(2)}}$ is independent of i. The ε term leads to

$$\psi^{(2)}(i) = 2D + \mathbf{\Pi}\boldsymbol{\psi}^{\mathbf{(2)}}(i) - D^{(2)}g_1(i,k),$$

or

$$[\mathbf{I}-\mathbf{\Pi}]\boldsymbol{\psi}^{\mathbf{(2)}} = \boldsymbol{\delta}^{\mathbf{(2)}}, \qquad \text{(C.8)}$$

where

$$\delta^{(2)}(i) = 2D - D^{(2)}g_1(i,k).$$

Now we premultiply (C.7) by $\varepsilon^2\boldsymbol{\pi}$. This leads to

$$0 = 2D - D^{(2)}G(k).$$

Hence

$$D^{(2)} \equiv 2D^2, \qquad \text{(C.9)}$$

as it should. Next (C.8) leads to

$$\boldsymbol{\psi}^{(\mathbf{2})} = \mathbf{M}\boldsymbol{\delta}^{(\mathbf{2})} + \gamma_2, \quad \text{where} \quad \gamma_2 := \boldsymbol{\pi}\boldsymbol{\psi}^{(\mathbf{2})}.$$

First, by (C.11) , we know that $\mathbf{M}\boldsymbol{\delta}^{(\mathbf{2})}$ is finite .

Next, to derive γ_2, we premultiply again (C.7) by $\varepsilon^2\boldsymbol{\pi}$. This leads to

$$\begin{aligned} 0 &= \varepsilon^2 + 2\varepsilon D + 2\gamma_1\varepsilon^2 + 2\varepsilon^3\boldsymbol{\pi}\boldsymbol{\chi} \\ &- 2\varepsilon^2 DG(k) - 2\varepsilon^3 G(k)\sum_{u\geq k}\varphi_2(i,u,k)\psi(u) - \varepsilon^3 G(k)[D^{(2)}/\varepsilon^2 + \sum_{u\geq k}\varphi_2(i,u,k)\psi^{(2)}(u)/\varepsilon] + \mathcal{O}(\varepsilon^3). \end{aligned}$$

The ε term leads of course again to $D^{(2)} = 2D^2$. The ε^2 term leads to

$$0 = 1 + 2\gamma_1 - 2DG(k) - G(k)\sum_{u\geq k}\varphi_2(i,u,k)\psi^{(2)}(u),$$

which allows the determination of γ_2.

C.2 Analysis of M.

Let us first recall a few results from Louchard [35] and [36]: let us denote by $\phi(w,\theta,z|i)$ the GF of the total number of CCA, where we mark the width by w, the area by z, the size of the last column by θ, and we start with a first column of size i. Then

$$\phi = \theta z w[f_0 + f_1 D_1 + f_2 D_2],$$

where $D_1 = N_1/h, D_2 = N_2/h, N_1, N_2, f_0, f_1, f_2$ are rational functions. The PGF of the last column size is asymptotically given,for large area, by

$$G(\theta) = \frac{\theta\rho[f_1(\theta,1)N_1(1,\rho) + f_2(\theta,1)N_2(1,\rho)]}{N_1(1,\rho)}.$$

This corresponds to a distribution

$$\rho^j(aj+b)/a.$$

Now we must compute $\mathbf{M} := \sum_{n\geq 0}(\boldsymbol{\Pi}^n - \mathbf{1}\times\boldsymbol{\pi})$. We see that $\mathbf{S} := \sum_{n\geq 1}(\boldsymbol{\Pi}^n - \mathbf{1}\times\boldsymbol{\pi})$ can be written as

$$S(i,j) = \lim_{w\to 1}\left[\sum_{n\geq 1} w^{n+1}\Pi^n(i,j) - \frac{w^2}{1-w}\frac{\rho^j(aj+b)^2}{a}\right].$$

But

$$\Pi(i,j) = \frac{\rho^i(j+i-1)\rho^j(aj+b)}{\rho^i(ai+b)},$$

similarly,

$$\Pi^2(i,j) = \frac{\rho^i}{\rho^i(ai+b)}\sum_l (i+l-1)\rho^l(l+j-1)\rho^j[aj+b].$$

We obtain

$$S(i,j) = \lim_{w\to 1}\left[\frac{1}{\rho^i(ai+b)}[\theta^j][\phi(w,\theta,\rho|i) - w\rho^i\theta^i][aj+b] - \frac{w^2}{1-w}\frac{\rho^j(aj+b)^2}{a}\right].$$

This can be simplified as

$$S(i,j) = \lim_{w\to 1}\left[\frac{1}{\rho^i(ai+b)}[\theta^j][\tilde{\phi}(w,\theta,\rho|i)][aj+b] - \frac{w}{1-w}\frac{\rho^j(aj+b)^2}{a}\right].$$

where

$$\tilde{\phi}(w,\theta,\rho|i) = [\theta\rho[f_1 D_1 + f_2 D_2]|i] = \varphi_1/h, \text{ say}$$

with

$$\varphi_1 = [\theta\rho[f_1 N_1 + f_2 N_2]|i].$$

Set $w = 1-\varepsilon$. This leads to

$$\begin{aligned} S(i,j) &= \lim_{\epsilon\to 0}\left[-\frac{\rho^j(aj+b)^2}{a\varepsilon} + \frac{\rho^j(aj+b)^2}{a}\right. \\ &+ \left.\frac{1}{\rho^i(ai+b)}[\theta^j]\left[-\frac{\varphi_1(1,\theta,\rho)}{h_w\varepsilon} + \frac{\varphi_{1,w}(1,\theta,\rho)}{h_w} - \frac{\varphi_1(1,\theta,\rho)h_{ww}}{2h_w^2}\right]\right][aj+b], \end{aligned} \tag{C.10}$$

with

$$\begin{aligned} h_w &= (-35\rho + 41\rho^2 + 11)/16, \\ h_{ww} &= 2\rho^3. \end{aligned}$$

But we know that

$$\varphi(1,\theta,\rho)/h_w = -(ai+b)\rho^i G(\theta).$$

So the singularity in (C.10) is removed. Finally, Maple gives us

$$\begin{aligned} M(i,j) &= S(i,j) + [\![j=k]\!] - \pi(j) \\ &= [\![j=k]\!] + \frac{\rho^j(aj+b)^2 h_{ww}}{2ah_w} + \frac{\rho^j(aj+b)}{(ai+b)h_w}\varphi_3, \end{aligned} \tag{C.11}$$

with

$$\varphi_3 = (-7\rho^2 + 13\rho - 5)ij/8 + (-21\rho^2 + 7\rho + 1)(i+j)/16 + (69\rho^2 - 87\rho + 15)/16.$$

D Carlitz Compositions.

D.1 Hitting times.

Set

$$\varphi_1(i,k) := \sum_{j=k}^{\infty}\Pi(i,j) = g_1(i,k)\varepsilon + g_2(i,k)\varepsilon^2 + \mathcal{O}(\varepsilon^3),$$

with

$$\varepsilon := \rho^k, g_1(i,k) = \frac{1+\rho^i}{1-\rho}, g_2(i,k) = -\frac{1+\rho^i}{1-\rho^2}.$$

Let us also define

$$H(k) := \sum \pi(i)\varphi_1(i,k) = G_1(k)\varepsilon + G_2(k)\varepsilon^2 + \mathcal{O}(\varepsilon^3),$$

with

$$G_.(k) := \sum_i \pi(i) g_.(i,k).$$

This gives

$$G_1(k) = -\frac{1}{(1-\rho)h_w}, G_2(k) = \frac{2}{(1-\rho^2)h_w}.$$

We compute also, for $j \geq k$,

$$\varphi_2(i,j,k) := \frac{\Pi(i,j)}{\varphi_1(i,k)} = \rho^{j-k}(1-\rho) + \mathcal{O}(\varepsilon),$$

and

$$\sum_{j\geq k}\varphi_2(i,j,k)=1.$$

Note that the dominant term is independent of i.

We obtain

$$\begin{aligned}
h_i &= 1+\sum_{u<k}\Pi(i,u)h_u \\
&= 1+\sum_{u}\Pi(i,u)h_u-\sum_{u\geq k}\Pi(i,u)h_u \\
&= 1+\sum_{u}\Pi(i,u)h_u-\varphi_1(i,k)\sum_{u\geq k}\frac{\Pi(i,u)}{\varphi_1(i,k)}h_u.
\end{aligned}$$

This leads to

$$\begin{aligned}
D_i+\varepsilon\psi(i)+\varepsilon^2\chi(i) &= \varepsilon+\sum_{u}\Pi(i,u)[D(u)+\varepsilon\psi(u)+\varepsilon^2\chi(i)+\mathcal{O}(\varepsilon^3)] \\
&- \varphi_1(i,k)\sum_{u\geq k}\frac{\Pi(i,u)}{\varphi_1(i,k)}[D(u)+\varepsilon\psi(u)+\mathcal{O}(\varepsilon^2)] \\
&= \varepsilon+\sum_{u}\Pi(i,u)[D(u)+\varepsilon\psi(u)+\varepsilon^2\chi(i)] \\
&- [\varepsilon g_1(i,k)+\varepsilon^2 g_2(i,k)]\sum_{u\geq k}\varphi_2(i,u,k)[D(u)+\varepsilon\psi(u)]+\mathcal{O}(\varepsilon^3). \quad \text{(D.12)}
\end{aligned}$$

Identification of constant terms in (D.12) leads to $[\mathbf{I}-\mathbf{\Pi}]\mathbf{D}=0$, which confirms that $\mathbf{D}$ is independent of i. Now we premultiply (D.12) by $\boldsymbol{\pi}$. The ε term leads to

$$0=1-DG_1(k),$$

or

$$D=-h_w(1-\rho).$$

Now the ε term of (D.12) leads to

$$\psi(i)=1+\mathbf{\Pi}\boldsymbol{\psi}(i)-Dg_1(i,k),$$

or

$$[\mathbf{I}-\mathbf{\Pi}]\boldsymbol{\psi}=\boldsymbol{\delta},$$

with

$$\delta(i)=1-Dg_1(i,k).$$

This gives

$$\boldsymbol{\psi}=\mathbf{M}\boldsymbol{\delta}+\gamma_1,\qquad \text{where}\qquad \gamma_1:=\boldsymbol{\pi}\boldsymbol{\psi},$$

and $\mathbf{M}:=\sum_{n\geq 0}(\mathbf{\Pi}^n-\mathbf{1}\times\boldsymbol{\pi})$. First, by (D.16)(see the next Appendix Section), we know that $\mathbf{M}\boldsymbol{\delta}$ is finite .

Next, to derive γ_1, we premultiply again (D.12) by $\boldsymbol{\pi}$. We obtain

$$0=\varepsilon-[\varepsilon G_1(k)+\varepsilon^2 G_2(k)][D+\varepsilon\sum_{u\geq k}\varphi_2(i,u,k)\psi(u)]+\mathcal{O}(\varepsilon^3).$$

The ε term leads of course again to D. The ε^2 term leads to

$$0=-G_1(k)\sum_{u\geq k}\varphi_2(i,u,k)\psi(u)-DG_2(k),$$

which allows the determination of γ_1.

We must now compute the first terms in the expansion of $h_i^{(2)} := \mathbb{E}_i(T_k^2)$. We set

$$h_i^{(2)} = \frac{D^{(2)}(i)}{\varepsilon^2} + \frac{\psi^{(2)}(i)}{\varepsilon} + \chi^{(2)}(i) + \mathcal{O}(\varepsilon).$$

We derive

$$\begin{aligned}
h_i^{(2)} &= \mathbb{E}[(1+T_i)^2] = 1 + 2\sum_{u<k} \Pi(i,u)h_u + \sum_{u<k} \Pi(i,u)h_u^{(2)} \\
&= 1 + 2\sum_{u} \Pi(i,u)h_u + \sum_{u} \Pi(i,u)h_u^{(2)} \\
&- 2\sum_{u\geq k} \Pi(i,u)h_u - \sum_{u\geq k} \Pi(i,u)h_u^{(2)} \\
&= 1 + 2\sum_{u} \Pi(i,u)[D/\varepsilon + \psi(u) + \varepsilon\chi(u)] \\
&+ \sum_{u} \Pi(i,u)[D^{(2)}(u)/\varepsilon^2 + \psi^{(2)}(u)/\varepsilon + \chi^{(2)}(u)] \\
&- 2[\varepsilon g_1(i,k) + \varepsilon^2 g_2(i,k)][D/\varepsilon + \sum_{u\geq k} \varphi_2(i,u,k)\psi(u)] \\
&- [\varepsilon g_1(i,k) + \varepsilon^2 g_2(i,k)][D^{(2)}(u)/\varepsilon^2 + \sum_{u\geq k} \varphi_2(i,u,k)\psi^{(2)}(u)/\varepsilon] + \mathcal{O}(\varepsilon). \qquad \text{(D.13)}
\end{aligned}$$

Multiply (D.13) by ε^2. The constant term leads to $[\mathbf{I} - \mathbf{\Pi}]\mathbf{D^{(2)}} = 0$, which shows that $\mathbf{D^{(2)}}$ is independent of i. The ε term leads to

$$\psi^{(2)}(i) = 2D + \mathbf{\Pi}\boldsymbol{\psi}^{\mathbf{(2)}}(i) - D^{(2)}g_1(i,k),$$

or

$$[\mathbf{I} - \mathbf{\Pi}]\boldsymbol{\psi}^{\mathbf{(2)}} = \boldsymbol{\delta}^{\mathbf{(2)}}, \qquad \text{(D.14)}$$

where

$$\delta^{(2)}(i) = 2D - D^{(2)}g_1(i,k).$$

Now we premultiply (D.13) by $\varepsilon^2\boldsymbol{\pi}$. This leads to

$$0 = 2D - D^{(2)}G_1(k).$$

Hence

$$D^{(2)} \equiv 2D^2, \qquad \text{(D.15)}$$

as it should. Next (D.14) leads to

$$\boldsymbol{\psi}^{\mathbf{(2)}} = \mathbf{M}\boldsymbol{\delta}^{\mathbf{(2)}} + \gamma_2, \quad \text{where} \quad \gamma_2 := \boldsymbol{\pi}\boldsymbol{\psi}^{\mathbf{(2)}}.$$

First, by (D.16), we know that $\mathbf{M}\boldsymbol{\delta}^{\mathbf{(2)}}$ is finite.

Next, to derive γ_2, we premultiply again (D.13) by $\varepsilon^2\boldsymbol{\pi}$. This leads to

$$\begin{aligned}
0 &= \varepsilon^2 + 2\varepsilon D + 2\gamma_1\varepsilon^2 + 2\varepsilon^3\boldsymbol{\pi}\boldsymbol{\chi} \\
&- 2[\varepsilon G_1(k) + \varepsilon^2 G_2(k)]\varepsilon^2[D/\varepsilon + \sum_{u\geq k} \varphi_2(i,u,k)\psi(u)] \\
&- [\varepsilon G_1(k) + \varepsilon^2 G_2(k)]\varepsilon^2[D^{(2)}/\varepsilon^2 + \sum_{u\geq k} \varphi_2(i,u,k)\psi(u)/\varepsilon] + \mathcal{O}(\varepsilon^3).
\end{aligned}$$

The ε term leads of course again to $D^{(2)} = 2D^2$. The ε^2 term leads to

$$0 = 1 + 2\gamma_1 - 2DG_1(k) - G_2(k)D^{(2)} - G_1(k)\sum_{u\geq k} \varphi_2(i,u,k)\psi^{(2)}(u),$$

which allows the determination of γ_2.

D.2 Analysis of M.

This has been done in Hitczenko, Louchard [14]. This leads to

$$\begin{aligned} M(i,j) &= -\frac{\rho^i}{(1+\rho^i)}[\![i=j]\!] - \frac{\rho^i\rho^j}{(1+\rho^i)(1+\rho^j)^2 h_w} + \frac{\rho^j}{(1+\rho^j)^3 h_w} \\ &- \frac{h_{ww}}{2h_w^2}\frac{\rho^j}{(1+\rho^j)^2} + [\![i=j]\!], \end{aligned} \tag{D.16}$$

with

$$\begin{aligned} h_w &= -\sum_{i=1}^{\infty}\frac{\rho^i}{(1+\rho^i)^2}, \\ h_{ww} &= 2\sum_{i=1}^{\infty}\frac{\rho^{2i}}{(1+\rho^i)^3}. \end{aligned}$$

We will also need

$$h_z = -\sum_{i=1}^{\infty}\frac{i\rho^i}{\rho(1+\rho^i)^2}.$$

References

[1] P. Biane, J. Pitman, and M. Yor. Probability laws related to the Jacobi theta and Riemann zeta functions, and the Brownian excursions. *Bulletin (New series) of the American Mathematical Society*, 38(4):435–465, 2001.

[2] P. Biane and M. Yor. Valeurs principales associées aux temps locaux browniens. *Bulletin de la Société Mathématique de France*, 2e série,111:23–101, 1987.

[3] S. L. Campbell and C. D. Meyer. *Generalized Inverse of Linear Transformations*. Pitman, 1979.

[4] M. Durand and P. Flajolet. Loglog counting of large cardinalities. Technical report, INRIA, 2003.

[5] J.A. Fill, H. Mahmoud, and W. Szpankowski. On the distribution for the duration of a randomized leader election algorithm. *Annals of Applied Probability*, 1:1260–1283, 1996.

[6] P. Flajolet. *Analyse d'algorithmes de manipulation d'arbres et de fichiers*. Université Pierre et Marie Curie, Paris, 1981.

[7] P. Flajolet. Approximate counting: a detailed analysis. *BIT*, 25:113–134, 1985.

[8] P. Flajolet. On adaptive sampling. *Computing*, 34:391–400, 1990.

[9] P. Flajolet, X. Gourdon, and P. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144:3–58, 1995.

[10] P. Flajolet and G.N. Martin. Probabilistic counting algorithms for data base applications. *Journal of Computer and System Sciences*, 31:182–209, 1985.

[11] P. Flajolet and H. Prodinger. Register allocation for unary-binary tress. *SIAM Journal on Computing*, 15(3):629–640, 1986.

[12] P. Flajolet, J.-C. Raoult, and J. Vuillemin. The number of registers required for evaluating arithmetic expressions. *Theoret. Comput. Sci.*, 9(1):99–125, 1979.

[13] P. Flajolet and R. Sedgewick. Digital search trees revisited. *SIAM Journal on Computing*, 3:748–767, 1986.

[14] P. Hitczenko and G. Louchard. Distinctness of compositions of an integer: a probabilistic analysis. *Random Structures and Algorithms*, 19(3,4):407–437, 2001.

[15] F. Hubalek. On the variance of the internal path length of generalized digital trees- the Mellin convolution approach. *Theoretical Computer Science*, 242:143–168, 2000.

[16] F. Hubalek, H.-K. Hwang, H. Mahmoud, W. Lew, and H. Prodinger. A multivariate view of random bucket digital search trees. *Journal of Algorithms*, 44:121–158, 2002.

[17] H.K. Hwang and Y.N. Yeh. Measures of distinctness for random partitions and compositions of an integer. *Advances in Applied Mathematics*, 19:378–414, 1999.

[18] P. Jacquet and M. Régnier. Trie partitioning process: limiting distributions. volume 214 of *Lecture Notes in Computer Science*, pages 196–210, 1986.

[19] S. Janson. Rounding of continuous random variables and oscillatory asymptotics. Technical Report 2004:47, Uppsala University, 2004.

[20] J. Keilson. *Markov Chain Models-Rarity and Exponentiality.* Springer-Verlag, 1979.

[21] J. G. Kemeny, J. L. Snell, and A. W. Knapp. *Denumerable Markov Chains.* Van Nostrand, 1966.

[22] R. Kemp. The average number of registers needed to evaluate a binary tree optimally. *Acta Inform.*, 11(4):363–372, 1978/79.

[23] C. Kenyon, G. Louchard, and R. Schott. Data structures maxima. *SIAM Journal on Computing*, 26(4):1006–1042, 1997.

[24] P. Kirschenhofer and H. Prodinger. Asymptotische Untersuchungen über charakteristische Parameter von Suchbäumen. In E. Hlawka, editor, *Zahlentheoretische Analysis II*, volume 1262 of *Lecture Notes in Mathematics*, pages 93–107, 1987.

[25] P. Kirschenhofer and H. Prodinger. Further results on digital search trees. *Theoretical Computer Science*, 58:143–154, 1988.

[26] P. Kirschenhofer and H. Prodinger. Approximate counting: An alternative approach. *RAIRO Theoretical Informatics and Applications*, 25:43–48, 1991.

[27] P. Kirschenhofer, H. Prodinger, and J. Schoißengeier. Zur Auswertung gewisser numerischer Reihen mit Hilfe modularer Funktionen. In E. Hlawka, editor, *Zahlentheoretische Analysis II*, volume 1262 of *Lecture Notes in Mathematics*, pages 108–110, 1987.

[28] P. Kirschenhofer, H. Prodinger, and W. Szpankowski. Analysis of a splitting process arising in probabilistic counting and other related algorithms. *Random Structures and Algorithms*, 9:379–401, 1996.

[29] D. E. Knuth. *The Art of Computer Programming*, volume 3: Sorting and Searching. Addison-Wesley, 1973. Second edition, 1998.

[30] M. Loève. *Probability Theory.* D. Van Nostrand, 1963.

[31] G. Louchard. Brownian motion and algorithm complexity. *BIT*, 26:17–34, 1986.

[32] G. Louchard. Exact and asymptotic distributions in digital and binary search trees. *RAIRO Theoretical Informatics and Applications*, 21(4):479–496, 1987.

[33] G. Louchard. Digital search trees revisited. *Cahiers du CERO*, 36:259–278, 1995.

[34] G. Louchard. Probabilistic analysis of adaptative sampling. *Random Structures and Algorithms*, 10:157–168, 1997.

[35] G. Louchard. Probabilistic analysis of column-convex and directed diagonally-convex animals. *Random Structures and Algorithms*, 11:151–178, 1997.

[36] G. Louchard. Probabilistic analysis of column-convex and directed diagonally-convex animals. II: Trajectories and shapes. *Random Structures and Algorithms*, 15:1–23, 1999.

[37] G. Louchard. Runs of geometrically distributed random variables: a probabilistic analysis. *Journal of Computational and Applied Mathematics*, 142(1):137–153, 2002.

[38] G. Louchard. Monotone runs of uniformly distributed integer random variables: a probabilistic analysis. 2003. submitted,http://www.ulb.ac.be/di/mcs/louchard/lungo.ps.

[39] G. Louchard and H. Prodinger. Probabilistic analysis of Carlitz compositions. *Discrete Mathematics and Theoretical Computer Science*, 5(1):71–96, 2002.

[40] G. Louchard and H. Prodinger. The moments problem of extreme-value related distribution functions. 2004. http://www.ulb.ac.be/di/mcs/louchard/moml.ps.

[41] G. Louchard, H. Prodinger, and M.D. Ward. The number of distinct values of some multiplicity in sequences of geometrically distributed random variables. *Discrete Mathematics and Theoretical Computer Science*, AD:231–256, 2005. 2005 International Conference on Analysis of Algorithms.

[42] G. Louchard and W. Szpankowski. Average profile and limiting distribution for a phrase size in the Lempel-Ziv parsing algorithm. *IEEE Transactions on Information Theory*, 41:478–488, 1995.

[43] G. Louchard, W. Szpankowski, and J. Tang. Average profile of generalized digital search trees and the generalized Lempel-Zif code. *SIAM Journal on Computing*, pages 935–954, 1999.

[44] H. M. Mahmoud. *Evolution of Random Search Trees*. John Wiley, New York, 1992.

[45] H. Mohamed and P. Robert. A probabilistic analysis of some tree algorithms. Technical report, INRIA, 2005. To appear in the Annals of Applied Probability.

[46] B. Pittel. Paths in a random digital tree: limiting distributions. *Advances in Applied Probability*, 18:139–155, 1986.

[47] H. Prodinger. Hypothetic analyses: Approximate counting in the style of Knuth, path length in the style of Flajolet. *Theoretical Computer Science*, 100:243–251, 1992.

[48] H. Prodinger. How to select a loser. *Discrete Mathematics*, 120:149–159, 1993.

[49] H. Prodinger. Approximate counting via Euler transform. *Mathematica Slovaka*, 44:569–574, 1994.

[50] H. Prodinger. Periodic oscillations in the analysis of algorithms—and their cancellations. *Journal of the Iranian Statistical Society*, 2004.

[51] H. Prodinger. Compositions and Patricia tries: no fluctuations in the variance! *SODA*, 2004 (5 pages).

[52] P. Robert. On the asymptotic behavior of some algorithms. Technical report, INRIA, 2004. To appear in Random Structures and Algorithms.

[53] W. Szpankowski. *Average case Analysis of Algorithms on Sequences*. Wiley, 2001.