

Approximate counting with m counters: A probabilistic analysis

Research Article

Guy Louchard^{1*}, Helmut Prodinger^{2**}

1. Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium

2. University of Stellenbosch, Mathematics Department, 7602 Stellenbosch, South Africa

Abstract: Motivated by a recent paper by Cichoń and Macyna [1], who introduced m counters (instead of just one) in the approximate counting scheme first analysed by Flajolet [2], we analyse the moments of the sum of the m counters, using techniques that proved to be successful already in several other contexts [11].

2010 MSC: 60C05, 05A16, 68Q24

Keywords: Approximate counting, Moments, Constant and fluctuating components, Complex analysis, Product of Fourier series

1. Introduction

Approximate counting is a technique that was first analysed by Flajolet [2]; some subsequent papers [6, 12–14] added to the analysis.

A counter C is kept, and each time an item arrives and needs to be counted, a random experiment is performed; if the current value of the counter is i , then with probability 2^{-i} the counter is increased by 1, otherwise it keeps its value; at the beginning, the counter value is $C = 1$. After n random increments, the value of the counter is typically close to $\log_2 n$, and the cited papers contain exact and asymptotic values for average and variance. For instance, Flajolet [2] gives the dominant constant part of mean and variance and the periodic part of the mean.

Recently, Cichoń and Macyna [1] used this idea as follows: Instead of one counter, they keep m counters, where $m \geq 1$ is an integer. For our subsequent analysis we will assume that m is fixed. When a new element arrives (and needs to be counted), it is randomly (with probability $\frac{1}{m}$) assigned to one

* E-mail: louchard@ulb.ac.be

** E-mail: hprodinger@sun.ac.za (Corresponding Author)

of the m counters, and then the random experiment is performed as usual. The parameter that Cichoń and Macyna are interested in is the total number of changes of any counter. They also consider another strategy which is not discussed in the present paper. In other words, if we (for convenience) assume that the initial setting of a counter to the value 1 counts as a change, Cichoń and Macyna are interested in the sum of the values of the m counters.

The paper [16] provided the first analysis of Cichoń and Macyna's scheme: Based on *exact expressions*, asymptotics for expectation and variance are derived with Rice's method. There is a price to be paid for dealing with these exact expressions, as there are computational hardships to be dealt with. Let us also mention Fuchs, Lee and Prodinger [4] who analyze this algorithm via the Poisson-Laplace-Mellin method. In the present paper, the approach is different: Going to approximations immediately, one loses exact expressions, but on the other hand the computations become much more manageable, so that one can go to higher moments, which we do here, mostly, to show the power of the method.

The new interest in approximate counting that Cichoń and Macyna's scheme initiated, motivated us to provide this paper: We had a long report [10] on asymptotics of the moments of extreme-value related distribution functions, but only a shortened version of it was published [11]; especially the analysis of classical approximate counting had to be left out. We present it here, together with additional material that deals with the m counters (instead of just one, as in the classical case). We also present new simplifications of products of some Fourier series.

Let $J(m, n)$ be the random variable (RV): "total value of the counter after n items have arrived"; we can write $J(m, n) = \sum_1^m J_i(n)$ where $J_i(n)$ is related to the i th counter. When we have only one counter, we can just write $J(n)$.

The most important motivation of the paper is to compute the asymptotic distribution and the moments of $J(m, n)$. The asymptotic distribution is related to the extreme-value Gumbel distribution function (DF): $\exp(-\exp(-x))$. The moments are usually given by a constant part and a small fluctuating part. There we use Laplace and Mellin transforms and singularity analysis.

Our aim is to derive an (almost) purely mechanical computation of constant and fluctuating components, with the help of computer algebra systems (we use Maple here). As an example, we provide the first four moments, (even the third moment is very rarely computed in the literature) but the treatment is completely automatic (with some human guidance of course). The fourth moment is particularly interesting: it presents a wide variety of combinatorial and mathematical constants as well as several types of Fourier series (including products of them).

A last but not least motivation is to simplify the analytic treatment by using only easy complex analysis: only simple poles are needed, and 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.

We have uniform integrability for the moments of our RV's. To show that the limiting moments are equivalent to the moments of the limiting distributions, we need a suitable rate of convergence. This is related to a uniform integrability condition (see Loève [8, Section 11.4]).

The total error term related to our asymptotics of moments is detailed in [11]; it is given by $\mathcal{O}(n^{-C})$, where C is some constant.

Another technical point of interest are the periodic oscillations that always occur in approximate counting and related questions: When one goes for higher moments, there are many extra terms coming in, making the Fourier coefficients very complicated. In particular, there are high powers of some (simple) periodic functions. We present a technique to bring the Fourier coefficients of these into some standard form, using residue calculus.

To summarize, we had several motivations to write this paper:

- show the power of the method we introduced in [11]
- present, with great details, the analysis of classical Approximate counting
- give new simplifications of product of some Fourier series

- analyze, with precision, Approximate counting with m counters

The paper is organized as follows: Definitions, notations and known properties are recalled in Section 2. Classical approximate counter (1 counter) is analyzed in Section 3. The case of m counters is considered in Section 4. The asymptotic moments of $J(n)$ are given in Section 5. Section 6 concludes the paper.

2. Definitions, notations and known properties

Let us first give the notations we will need throughout the paper. They will be used in Section 3, but we prefer to summarize them here instead of introducing them one by one.

$$\begin{aligned}
 L &:= \ln 2, & \log &:= \log_2, & \varepsilon &:= \text{small real } > 0, & \tilde{\alpha} &:= \alpha/L, \\
 Q(j) &:= \prod_{k=1}^j (1 - 2^{-k}), & Q(0) &= 1, & Q &:= Q(\infty) = 0.288788095087\dots, \\
 R(j) &:= \frac{(-1)^{j+1}}{2^{j(j+1)/2} Q(j)}, & R(0) &= -1, & \chi_l &:= \frac{2\pi i l}{L}, \\
 \rho_k &:= \sum_{l \neq 0} \Gamma(\chi_l) \psi(\chi_l)^k e^{-2l\pi i \log n}, & k &= 1, 2, 3, & \rho_4 &:= \sum_{l \neq 0} \Gamma(\chi_l) \psi(1, \chi_l) e^{-2l\pi i \log n}.
 \end{aligned}$$

These functions appear in many analyses of algorithms (see, for instance, Flajolet [2], Flajolet and Sedgewick [3], Hwang et al. [5], Louchard [9], Louchard and Prodinger [11]).

The following facts will be frequently used:

$$\begin{aligned}
 Q(i) &\geq Q, \\
 (1 - u)^n &= e^{-nu} [1 - nu^2/2 + \mathcal{O}(nu^3)], \quad u \in]0, 1[.
 \end{aligned}$$

For the integer-valued RV J (from now on, we drop i and n from $J_i(n)$ in order to ease the notation; J is now related to one counter, with n items), we set

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

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

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

and, of course,

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

Asymptotically, the distribution will be a periodic function of $\log n$ in the following sense: fix n ;

$$\mathbb{P}(J(n) \leq j) \sim F(\eta), \quad \eta = j - \log n = j - \lfloor \log n \rfloor - \{\log n\}.$$

Set $\log n' = \log n + 1$ (hence $n' = 2n$, the base could of course be changed). Then

$$\mathbb{P}(J(n') \leq j + 1) \sim F(j + 1 - \log n') = F(\eta).$$

The asymptotic distribution of $J(n')$ is the same as the one of $J(n)$, only shifted by 1. All n with the same $\{\log n\}$ lead to the same points of $F(\eta)$. 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. This type of convergence is not uncommon in the Analysis of Algorithms. Many examples are given in [11].

Next, we must check that

$$\mathbb{E}(J^k) = \sum_j j^k p(j) \sim \sum_j (\eta + \log n)^k f(\eta), \tag{1}$$

by computing a suitable rate of convergence. This is related to a uniform integrability condition (see Loève [8, Section 11.4].)

3. Classical approximate counting (one counter). Analysis of $J(n)$

In this section, we provide the asymptotic distribution and the first four asymptotic moments of the RV $J(n)$ based on n items. From Flajolet [2, Proposition 1], we have

$$p(j) = \sum_{k=0}^{j-1} 2^k \frac{-R(k)}{Q(j-1-k)} (1 - 1/2^{j-k})^n. \tag{2}$$

Letting $\eta = j - \log n$, it is proved in [2] that

$$f(\eta) = \sum_{k=0}^{\infty} 2^k \frac{-R(k)}{Q} \exp(-2^{-\eta+k}).$$

(Compare also [6, 13, 14].) It has been pointed out in [9] that it has some similarities with the Digital Search Tree distribution. Actually, as noticed by S. Janson (private communication), the distribution for approximate counting is the same as for unsuccessful search in Digital Search Trees (not only asymptotically).

The rate of the convergence problem is completely solved in Flajolet [2]. Also, we obtain by summing

$$\begin{aligned} P(j) &= - \sum_{k=0}^{j-1} \sum_{u=k+1}^j 2^k \frac{R(k)}{Q(u-1-k)} (1 - 1/2^{u-k})^n, \\ F(\eta) &= - \sum_{k=0}^{\infty} 2^k \frac{R(k)}{Q} \sum_{i=k}^{\infty} \exp(-2^{-\eta+i}). \end{aligned} \tag{3}$$

We note that the algorithm can be generalized by changing the base. The analysis is quite similar, and we won't provide details here.

The first three asymptotic moments are given in our unpublished report [10]. We take the opportunity to present them here, with some complements. In particular we analyze the product of Fourier series, which leads to convolutions of the coefficients. In order to show the power of the methods we use, we also give the fourth moment. Using the techniques we described in our published paper [11], we proceed as follows.

3.1. Some preliminary identities

Some preliminary identities are necessary. They appear in the asymptotic moments, but we prefer to avoid overloading Sections 3.3 and 3.4.

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

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

(for a simple proof, see Knuth [7, Ex 5.1.1–16]). It holds for general q , once the definition of $Q(i)$ is adapted by replacing $\frac{1}{2}$ by q . Set

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

It is not hard to see that, with $z = -1$, $q = 1/2$, we get the following expansions

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

with the abbreviations

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

Now we want to compute sums of type

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

We obtain

$$\Pi^{*'} = \Sigma_1 \Pi^*, \quad \Pi' = \Pi^* + (1 + z)\Sigma_1 \Pi^*,$$

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

$$\begin{aligned}U_2 &= -Q(-1 + 2C_1), \\ U_3 &= Q(1 - 6C_1 - 3C_2 + 3C_1^2), \\ U_4 &= -Q(-1 + 14C_1 + 18C_2 - 18C_1^2 + 8C_3 - 12C_1C_2 + 4C_1^3), \\ U_5 &= Q(1 - 30C_1 - 75C_2 + 75C_1^2 - 80C_3 + 120C_1C_2 - 40C_1^3 - 30C_4 + 40C_1C_3 \\ &\quad + 15C_2^2 - 30C_1^2C_2 + 5C_1^4).\end{aligned}\tag{4}$$

Also

$$U_0 = 0;\tag{5}$$

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

$$\sum_{i=0}^{\infty} 2^{(k+1)i} R(i) = 0.\tag{6}$$

3.2. Slow increase property

It is necessary that the functions we need decrease exponentially in the direction $i\infty$. Indeed, (see ([11] for details), some integration in the complex plane (along a rectangle) must be convergent.

It will appear that all functions we use here are analytic (in some domain), depending on classical functions such as Γ , ζ , $\psi(k, s)$ (the $(k + 1)$ -gamma function).

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 $s = \sigma + it$,

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

We will also use the function $H_2(-Ls)$, with

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

To analyze this function, we use the “sum splitting technique” as described in Knuth [7, p. 131], and used in Flajolet [2]: let $\sigma > -1$ and $\rho(x)$ be an increasing function.

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

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

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

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

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

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

$$H_2(\alpha) = -\theta(2e^\alpha) = (e^\alpha - 1)\theta(e^\alpha).$$

This easily leads to

$$\begin{aligned} H_2(\alpha) &\sim -Q(\alpha - L) + \mathcal{O}((\alpha - L)^2), \\ H_2(\alpha) &\sim 3Q(\alpha - 2L) + \mathcal{O}((\alpha - 2L)^2), \\ H_2(\alpha) &\sim -21Q(\alpha - 3L) + \mathcal{O}((\alpha - 3L)^2), \end{aligned} \tag{7}$$

and similar asymptotics for $\alpha = kL, k \geq 4$.

3.3. The asymptotic moments

The detailed proofs of relations (8) to (17) are given in [11]. Let an (integer-valued) RV K be such that $\Pr(K - \log n \leq \eta) \sim F(\eta)$, 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

$$\varphi(\alpha) = \mathbb{E}(e^{\alpha Z}) = 1 + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} m_k = e^{\alpha m_1} \lambda(\alpha), \tag{8}$$

say, with

$$\lambda(\alpha) = 1 + \frac{\alpha^2}{2} \sigma^2 + \sum_{k=3}^{\infty} \frac{\alpha^k}{k!} \mu_k. \tag{9}$$

Also

$$\varphi(\alpha) = \int_{-\infty}^{+\infty} e^{\alpha\eta} F'(\eta) d\eta. \tag{10}$$

We have here, for J ,

$$\phi(\alpha) := \int_{-\infty}^{\infty} e^{\alpha\eta} f(\eta) d\eta = -\frac{1}{QL} H_2(\alpha) \Gamma(-\tilde{\alpha}), \quad \Re(\alpha) < 0.$$

But we need a larger range for α . But, by (7), we can continue $\phi(\alpha)$ analytically for all α : all singularities of $\Gamma(-\tilde{\alpha})$ are cancelled by the successive roots of $H_2(\alpha)$. Moreover, this implies

$$e^{\alpha\eta} f(\eta) \rightarrow 0, \quad \eta \rightarrow \infty, \quad \eta \rightarrow -\infty. \tag{11}$$

Now, we have

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{\alpha\eta} [F'(\eta) - F'(\eta - 1)] d\eta &= (1 - e^\alpha) \varphi(\alpha) \\ &= [e^{\alpha\eta} [F(\eta) - F(\eta - 1)]]_{-\infty}^{\infty} - \alpha \int_{-\infty}^{+\infty} e^{\alpha\eta} [F(\eta) - F(\eta - 1)] d\eta \\ &= [e^{\alpha\eta} f(\eta)]_{-\infty}^{\infty} - \alpha \int_{-\infty}^{+\infty} e^{\alpha\eta} f(\eta) d\eta \\ &= -\alpha \phi(\alpha), \quad \text{by (11)}. \end{aligned}$$

This gives

$$\varphi(\alpha) = \frac{\alpha}{e^\alpha - 1} \phi(\alpha) = -\frac{H_2(\alpha) \Gamma(1 - \tilde{\alpha})}{Q(1 - e^\alpha)}.$$

This leads to (we give only the first two expressions for m_i)

$$\begin{aligned} m_1 &= \frac{\gamma}{L} - C_1, \\ m_2 &= \frac{\pi^2 + 6\gamma^2}{6L^2} - \frac{2\gamma C_1}{L} - C_2 + C_1^2 - C_1, \\ \mu_2 &= \sigma^2 = \frac{\pi^2}{6L^2} - C_1 - C_2, \\ \mu_3 &= \frac{2\zeta(3)}{L^3} - 2C_3 - 3C_2 - C_1, \\ \mu_4 &= 3C_1^2 - C_1 - 7C_2 - 12C_3 + 3C_2^2 - \frac{\pi^2 C_1}{L^2} + 6C_1 C_2 - \frac{\pi^2 C_2}{L^2} + \frac{3\pi^4}{20L^4} - 6C_4. \end{aligned}$$

Let w, κ 's (with or without subscripts) denote periodic functions of $\log n$, with period 1 and 0 mean (for w) or non-zero small mean (for κ) and small (of order 10^{-4}) amplitude. Actually, these functions depend on the fractional part of $\log n$: $\{\log n\}$.

The moments of $J(n) - \log n$ are asymptotically given by $\tilde{m}_i + w_i$. The constant part \tilde{m}_i comes from the residue of some function at 0 (see [11]) and w_i comes from the residues of the same function at some complex poles. The generating function of \tilde{m}_i is given by

$$\phi(\alpha) := \int_{-\infty}^{\infty} e^{\alpha\eta} f(\eta) d\eta = 1 + \sum_{i=1}^{\infty} \frac{\alpha^i}{i!} \tilde{m}_i = \varphi(\alpha) \frac{e^\alpha - 1}{\alpha}. \tag{12}$$

This leads to

$$\begin{aligned}\tilde{m}_1 &= \frac{1}{2} - C_1 + \frac{\gamma}{L}, \\ \tilde{m}_2 &= \frac{\pi^2 + 6\gamma^2}{6L^2} + \frac{\gamma - 2\gamma C_1}{L} + \frac{1}{3} - 2C_1 - C_2 + C_1^2.\end{aligned}$$

More generally, the centered moments of $J(n)$ are asymptotically given by $\boldsymbol{\mu}_i = \tilde{\mu}_i + \kappa_i$, with the asymptotic *dominant* constant centered moment generating function given by

$$\Theta(\alpha) := 1 + \sum_{k=2}^{\infty} \frac{\alpha^k}{k!} \tilde{\mu}_k = \frac{2}{\alpha} \sinh(\alpha/2) \lambda(\alpha). \quad (13)$$

The neglected part is of order $1/n^\beta$ with $0 < \beta < 1$. We derive, with (4), the following result about these centered moments.

Theorem 3.1. *The asymptotic dominant constant parts of the centered moments of $J(n)$ are given by*

$$\begin{aligned}\tilde{\mu}_2 &= \frac{\pi^2}{6L^2} + \frac{1}{12} - C_1 - C_2, \\ \tilde{\mu}_3 &= \frac{2\zeta(3)}{L^3} - 2C_3 - 3C_2 - C_1, \\ \tilde{\mu}_4 &= \frac{1}{80} + 3C_1^2 - \frac{3C_1}{2} - \frac{15C_2}{2} - 12C_3 - 6C_4 + \frac{3\pi^4}{20L^4} \\ &\quad + \frac{\pi^2}{12L^2} + 6C_1C_2 - \frac{\pi^2C_1}{L^2} - \frac{\pi^2C_2}{L^2} + 3C_2^2.\end{aligned}$$

Note that

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

Now

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

with

$$w_k = \frac{1}{L} \sum_{l \neq 0} \Upsilon_k^*(\chi_l) e^{-2l\pi i \log n}, \quad (14)$$

and

$$\Upsilon_k^*(s) = L \phi^{(k)}(\alpha) \Big|_{\alpha = -Ls}. \quad (15)$$

With (5), we check that we have no singularity of $\phi^{(k)}$, $k > 0$, at $\alpha = 0$. The fundamental strip for (15) is $\Re(s) \in (-1, 0)$. We first obtain

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

\tilde{m}_1 , $\tilde{\mu}_2$ and w_1 are identical to the expressions given in Flajolet [2].

To compute the periodic components κ_i (with non-zero small mean) to be added to the centered moments $\tilde{\mu}_i$, we first set

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

Now, 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

$$\begin{aligned} \phi_p(\alpha) &= \phi(\alpha) + \sum_{k=0}^{\infty} \sum_{l \neq 0} \phi^{(k)}(\alpha) \Big|_{\alpha=-L\chi_l} e^{-2l\pi i \log n} \frac{\alpha^k}{k!} \\ &= \phi(\alpha) + \sum_{l \neq 0} \phi(\alpha - L\chi_l) e^{-2l\pi i \log n}. \end{aligned} \tag{16}$$

Finally, we compute

$$\Theta_p(\alpha) = \phi_p(\alpha) e^{-\alpha m_1} = 1 + \sum_{k=2}^{\infty} \frac{\alpha^k}{k!} (\tilde{\mu}_k + \kappa_k) = \Theta(\alpha) + \sum_{k=2}^{\infty} \frac{\alpha^k}{k!} \kappa_k, \tag{17}$$

leading to the (exponential) generating function of κ_k . By expansion and taking differences, we have a result about the oscillating parts.

Theorem 3.2. *The asymptotic oscillating parts of the centered moments of $J(n)$ are given by*

$$\begin{aligned} \kappa_2 &= -w_1^2 - \frac{2\gamma w_1}{L} + \frac{2}{L^2} \rho_1, \\ \kappa_2^2 &= w_1^4 + \frac{4\gamma^2 w_1^2}{L^2} + \frac{4}{L^4} \rho_1^2 + \frac{4\gamma w_1^3}{L} - \frac{4w_1^2}{L^2} \rho_1 - \frac{8w_1}{L^3} \rho_1, \\ \kappa_3 &= \frac{4L^2 w_1^2 + 12w_1 L \gamma + 6\gamma^2 - \pi^2}{2L^2} w_1 - \frac{6(\gamma + w_1 L)}{L^3} \rho_1 - \frac{3}{L^3} \rho_2 - \frac{3}{L^3} \rho_4, \\ \kappa_4 &= w_1 \left[-w_1/2 + \frac{12\gamma C_2}{L} + \frac{12\gamma C_1}{L} - 3w_1^3 + \frac{\pi^2 w_1}{L^2} - \frac{8\zeta(3)}{L^3} \right. \\ &\quad \left. + 6C_1 w_1 + 6w_1 C_2 - \frac{12\gamma^2 w_1}{L^2} - \frac{4\gamma^3}{L^3} - \frac{12w_1^2 \gamma}{L} - \frac{\gamma}{L} \right] \\ &\quad + \frac{L^2 - 12C_2 L^2 - 12C_1 L^2 + 24\gamma w_1 L + 12w_1^2 L^2 + 12\gamma^2}{L^4} \rho_1 \\ &\quad + \frac{12}{L^4} \sum_{l \neq 0} e^{-2l\pi i \log n} \psi(\chi_l) \psi(1, \chi_l) \Gamma(\chi_l) + \frac{12(w_1 L + \gamma)}{L^4} \rho_4 \\ &\quad + \frac{4}{L^4} \sum_{l \neq 0} e^{-2l\pi i \log n} \psi(2, \chi_l) \Gamma(\chi_l) + \frac{4}{L^4} \rho_3 + \frac{12(w_1 L + \gamma)}{L^4} \rho_2. \end{aligned}$$

All algebraic manipulations of this paper are mechanically performed by Maple.¹

¹ $\Gamma'(x) = \Gamma(x)\psi(x)$, $\Gamma''(x) = \Gamma(x)\psi(1, x) + \Gamma(x)\psi^2(x)$, $\Gamma'''(x) = \Gamma(x)\psi(2, x) + 3\Gamma(x)\psi(1, x)\psi^2(x) + \Gamma(x)\psi^3(x)$ etc.

Note that the non-zero κ_i mean must be added to the *dominant* constant parts of the centered moments. This will be considered in the next section.

3.4. The corrections

Products of Fourier series *may* have a constant term, even if the factors do not. This term must be included in the *dominant* constant part of our moments. This is the object of the present subsection.

We denote by $[f]_k$ the coefficient of $e^{-2k\pi i \log n}$ in the Fourier expansion of f .

In [11], we have proved the following relations

$$c_1[0] := [w_1^2]_0 = \frac{1}{L^2} \sum_{k \neq 0} \Gamma(-\chi_k) \Gamma(\chi_k) = -\frac{2}{L} D - \frac{11}{12} + \frac{\pi^2}{6L^2}$$

with

$$D := \sum_{l \geq 1} \frac{(-1)^l}{l(2^l - 1)}.$$

The coefficient $c_1[k]$ of $e^{-2k\pi i \log n}$ in the Fourier expansion of w_1^2 is given by

$$\begin{aligned} c_1[k] &= \frac{1}{L^2} \sum_{j \neq 0, k} \Gamma(-\chi_j) \Gamma(\chi_k + \chi_j) \\ &= \frac{2}{L} \sum_{l \geq 1} \frac{(-1)^l \Gamma(\chi_k + l)}{l!(2^l - 1)} + \frac{2}{L^2} \Gamma(\chi_k) (\psi(\chi_k) + \gamma), \end{aligned}$$

$$c_2[0] := [w_1^3]_0 = 1 + \frac{2\zeta(3)}{L^3} + \frac{1}{L} D - \frac{6}{L^2} D_1 - \frac{2 \log 3}{L} - \frac{2}{L} D_2,$$

with

$$\begin{aligned} D_1 &= \sum_{l \geq 1} \frac{(-1)^l H_{l-1}}{l(2^l - 1)}, \\ D_2 &= \sum_{l, j \geq 1} \frac{(-1)^{l+j}}{(l+j)(2^l - 1)} \left[\frac{1}{2^j - 1} + \frac{1}{2^{j+l} - 1} \right] \binom{l+j}{j}. \end{aligned}$$

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

The coefficient $c_2[k]$ of $e^{-2k\pi i \log n}$ in the Fourier expansion of w_1^3 is given by

$$\begin{aligned} c_2[k] &= - \left\{ \frac{2}{L^3} \left[2L \sum_{l \geq 1} \frac{(-1)^l}{l!(2^l - 1)} \Gamma(l + \chi_k) (\psi(l + \chi_k) + \gamma) - L^2 \sum_{l \geq 1} \frac{(-1)^l \Gamma(l + \chi_k)}{l!} \frac{2^l}{(2^l - 1)^2} \right. \right. \\ &\quad \left. \left. + \frac{1}{12} (18\Psi(1, \chi_k) + 18(\psi(\chi_k) + \gamma)^2 - 11L^2 - \pi^2) \Gamma(\chi_k) \right] \right. \\ &\quad \left. + \frac{2}{L^2} \sum_{l \geq 1} \frac{(-1)^l}{l!(2^l - 1)} \left[L \sum_{h \geq 0} \frac{(-1)^h}{h!(2^{h+l} - 1)} \Gamma(h + l + \chi_k) + L \sum_{h \geq 1} \Gamma(l + h + \chi_k) \frac{(-1)^h}{h!(2^h - 1)} \right. \right. \\ &\quad \left. \left. + L\Gamma(l + \chi_k) 2^{-l} - L\Gamma(l + \chi_k) - (l-1)! \Gamma(\chi_k) + \Gamma(\chi_k) \left(\psi(\chi_k) + \gamma + \frac{L}{2} \right) \right] \right. \\ &\quad \left. + \left[\frac{\pi^2}{6L^3} + \frac{1}{12L} - \frac{1}{L} - \frac{2}{L^2} \sum_{l \geq 1} \frac{(-1)^{l-1}}{l(2^l - 1)} \right] \Gamma(\chi_k) \right\}. \end{aligned}$$

For a complete description of the Fourier coefficients of the oscillations occurring in the third and fourth moment, we need the following expressions (note that Maple splits the higher derivatives of the Gamma function; if one could rework that, one could reduce the number of necessary expressions):

$$\begin{aligned} c_3[0] &:= [w_1\rho_1]_0, & c_4[0] &:= [w_1^4]_0, & c_5[0] &:= [(\rho_1)^2]_0, & c_6[0] &:= [w_1^2\rho_1]_0, \\ c_7[0] &:= [w_1\rho_4]_0, & c_8[0] &:= [w_1\rho_2]_0, & c_3[k] &:= [w_1\rho_1]_k, & c_4[k] &:= [w_1^4]_k, \\ c_5[k] &:= [(\rho_1)^2]_k, & c_6[k] &:= [w_1^2\rho_1]_k, & c_7[k] &:= [w_1\rho_4]_k, & c_8[k] &:= [w_1\rho_2]_k. \end{aligned}$$

We will show now how to “compute” some Fourier coefficients we need. “Computing” is perhaps a very ambitious word, it might be better replaced by “rewriting”. We have

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

and higher powers have convolutions as coefficients:

$$w_1^4 = \frac{1}{L^4} \sum_{k \neq 0} \sum_{k_1+k_2+k_3+k_4=k} \Gamma(\chi_{k_1})\Gamma(\chi_{k_2})\Gamma(\chi_{k_3})\Gamma(\chi_{k_4}) e^{-2\pi i k x},$$

and none of the k_1, \dots, k_4 is allowed to be zero. The only thing that we are able to achieve is to have only one Gamma-term in the (multiple) sum, where a typical term might look like

$$\sum_{j_1+\dots+j_t=j} C_{j_1}^{(1)} \dots C_{j_t}^{(t)} \Gamma^{(s)}(\chi_k + j) e^{-2\pi i k x}.$$

Such a representation is not a priori *better* than the straight-forward convolution, but we will sketch now how to achieve them. One (small) advantage is that the zeroth term can be explicitly determined, and extracted, and what is left is then oscillating around zero.

As an example, we consider

$$W := w_1^2 - [w_1^2]_0 = \sum_{k \neq 0} c_1[k] e^{-2\pi i k x},$$

with

$$c_1[k] = \frac{2}{L} \sum_{l \geq 1} \frac{(-1)^l \Gamma(\chi_k + l)}{l!(2^l - 1)} + \frac{2}{L^2} \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right).$$

Let us discuss W^2 . It is clear that the convolution of W with itself contains already several terms. For instance,

$$\begin{aligned} [W^2]_0 &= \frac{4}{L^2} \sum_{j,l \geq 1} \frac{(-1)^{j+l}}{j!(2^j - 1)l!(2^l - 1)} \sum_{k \neq 0} \Gamma(-\chi_k + j) \Gamma(\chi_k + l) \\ &+ \frac{8}{L^3} \sum_{l \geq 1} \frac{(-1)^l}{l!(2^l - 1)} \sum_{k \neq 0} \Gamma(-\chi_k + l) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right) \\ &+ \frac{4}{L^4} \sum_{k \neq 0} \left(\Gamma'(-\chi_k) + \gamma \Gamma(-\chi_k) \right) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right). \end{aligned}$$

We would like to demonstrate how to rewrite the k -sums in these expressions. The survey paper [15] has many similar examples. The approach we found most versatile is via residue calculus. One writes a suitable function, computes all the residues in the complex plane, and the sum of them is zero. There are of course some technical subtleties, like showing that integral tends to zero for larger and larger radii,

and also there are usually some series that do not converge absolutely. The suitable limit of them is the *Abel* limit, i.e., consider a power series in x , and let x tend to a point at the boundary of convergence. Here, we want to concentrate on the computational part only. The function that is suitable for the first part is

$$\frac{L}{2^z - 1} \Gamma(j - z) \Gamma(l + z).$$

A first contribution comes from the poles at $z = \chi_k$:

$$S_1 = \sum_{k \in \mathbb{Z}} \Gamma(j - \chi_k) \Gamma(l + \chi_k).$$

The next contribution stems from the poles at $z = j, j + 1, \dots$:

$$S_2 = (l + j - 1)! L \sum_{h \geq 0} \frac{(-1)^{h-1}}{2^{j+h} - 1} \binom{l + j + h - 1}{h}.$$

Now we look at the poles at $z = -l, -l - 1, \dots$:

$$\sum_{h \geq 0} \frac{L}{2^{-l-h} - 1} \Gamma(j + l + h) \frac{(-1)^h}{h!}.$$

This series does not converge absolutely. It is best to pull out the “bad” part, which leads to two contributions:

$$S_3 = (l + j - 1)! L \sum_{h \geq 0} \frac{(-1)^{h-1}}{2^{l+h} - 1} \binom{l + j + h - 1}{h},$$

$$S_4 = (l + j - 1)! L \sum_{h \geq 0} (-1)^{h-1} \binom{l + j + h - 1}{h}.$$

As announced, we must interpret S_4 as a limit:

$$S_4 = -(l + j - 1)! L \lim_{x \rightarrow 1} \sum_{h \geq 0} (-x)^h \binom{l + j + h - 1}{h} = -(l + j - 1)! L 2^{-l-j}.$$

Altogether we found

$$\begin{aligned} \sum_{k \neq 0} \Gamma(j - \chi_k) \Gamma(l + \chi_k) &= -(j - 1)! (l - 1)! - (l + j - 1)! L \sum_{h \geq 0} \frac{(-1)^{h-1}}{2^{j+h} - 1} \binom{l + j + h - 1}{h} \\ &\quad - (l + j - 1)! L \sum_{h \geq 0} \frac{(-1)^{h-1}}{2^{l+h} - 1} \binom{l + j + h - 1}{h} + (l + j - 1)! L 2^{-l-j}. \end{aligned}$$

Now, let us look at

$$\sum_{k \neq 0} \Gamma(-\chi_k + l) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right).$$

The proper function is

$$\frac{L}{2^z - 1} \Gamma(-z + l) \left(\Gamma'(z) + \gamma \Gamma(z) \right).$$

The poles at $z = \chi_k$, $k \neq 0$, lead to

$$S_1 = \sum_{k \neq 0} \Gamma(-\chi_k + l) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right).$$

The pole at $z = 0$ leads to

$$S_2 = \Gamma(l) \left(\frac{\pi^2 - L^2}{12} - \frac{\gamma^2}{2} - \frac{L\gamma}{2} \right) - \Gamma'(l) \left(\gamma + \frac{L}{2} \right) - \frac{1}{2} \Gamma''(l).$$

The poles at $z = l, l+1, \dots$ lead to

$$\begin{aligned} S_3 &= L \sum_{h \geq 0} \frac{1}{2^{l+h} - 1} \frac{(-1)^h}{h!} \left(\Gamma'(l+h) + \gamma \Gamma(l+h) \right) \\ &= L \sum_{h \geq 0} \frac{1}{2^{l+h} - 1} \frac{(-1)^h (l+h-1)!}{h!} H_{l+h}. \end{aligned}$$

The poles at $z = -1, -2, \dots$ lead to

$$\begin{aligned} & - \sum_{h \geq 1} \frac{L}{1 - 2^{-h}} \Gamma(h+l) \frac{(-1)^h \gamma}{h!} \\ &= -L\gamma \sum_{h \geq 1} \frac{1}{2^h - 1} \frac{(-1)^h (l+h-1)!}{h!} + L\gamma(1 - 2^{-l}). \end{aligned}$$

Therefore

$$\begin{aligned} \sum_{k \neq 0} \Gamma(-\chi_k + l) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right) &= -\Gamma(l) \left(\frac{\pi^2 - L^2}{12} - \frac{\gamma^2}{2} - \frac{L\gamma}{2} \right) + \Gamma'(l) \left(\gamma + \frac{L}{2} \right) + \frac{1}{2} \Gamma''(l) \\ & - L \sum_{h \geq 0} \frac{1}{2^{l+h} - 1} \frac{(-1)^h (l+h-1)!}{h!} H_{l+h} + L\gamma \sum_{h \geq 1} \frac{1}{2^h - 1} \frac{(-1)^h (l+h-1)!}{h!} - L\gamma(1 - 2^{-l}). \end{aligned}$$

Finally, let us consider

$$\sum_{k \neq 0} \left(\Gamma'(-\chi_k) + \gamma \Gamma(-\chi_k) \right) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right).$$

The function of interest is

$$\frac{L}{2^z - 1} \left(\Gamma'(-z) + \gamma \Gamma(-z) \right) \left(\Gamma'(z) + \gamma \Gamma(z) \right).$$

The poles at $z = \chi_k$, $k \neq 0$, lead to

$$S_1 = \sum_{n \neq 0} \left(\Gamma'(-\chi_k) + \gamma \Gamma(-\chi_k) \right) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right).$$

The pole at $z = 0$ leads to

$$S_2 = -\frac{11\pi^4}{360} - \frac{L^2\pi^2}{72} - \frac{L^4}{720}.$$

The poles at $z = 1, 2, \dots$ lead to

$$S_3 = L\gamma \sum_{h \geq 1} \frac{(-1)^{h-1} H_h}{h(2^h - 1)}.$$

The poles at $z = -1, -2, \dots$ lead to

$$\begin{aligned} S_3 &= -L\gamma \sum_{h \geq 1} \frac{(-1)^h H_h}{h(1-2^{-h})} \\ &= L\gamma \sum_{h \geq 1} \frac{(-1)^{h-1} H_h}{h(2^h-1)} - L\gamma \sum_{h \geq 1} \frac{(-1)^{h-1} H_h}{h}. \end{aligned}$$

Altogether

$$\begin{aligned} \sum_{k \neq 0} \left(\Gamma'(-\chi_k) + \gamma \Gamma(-\chi_k) \right) \left(\Gamma'(\chi_k) + \gamma \Gamma(\chi_k) \right) &= \frac{11\pi^4}{360} + \frac{L^2\pi^2}{72} + \frac{L^4}{720} \\ &\quad - 2L\gamma \sum_{h \geq 1} \frac{(-1)^{h-1} H_h}{h(2^h-1)} + L\gamma \sum_{h \geq 1} \frac{(-1)^{h-1} H_h}{h}. \end{aligned}$$

The last alternating sum has a closed form evaluation:

$$\sum_{h \geq 1} \frac{(-1)^h H_h}{h} = \frac{L^2}{2} - \frac{\pi^2}{12}.$$

This gives us the constant term in w_1^4 :

$$\begin{aligned} c_4[0] &= [w_1^4]_0 = \left(\frac{2}{L} \sum_{h \geq 1} \frac{(-1)^{h-1}}{h(2^h-1)} - \frac{11}{12} + \frac{\pi^2}{6L^2} \right)^2 \\ &\quad + \frac{4}{L^2} \sum_{j, l \geq 1} \frac{(-1)^{j+l}}{j!(2^j-1)l!(2^l-1)} \\ &\quad \times \left[-(j-1)!(l-1)! - (l+j-1)!L \sum_{h \geq 0} \frac{(-1)^{h-1}}{2^{j+h}-1} \binom{l+j+h-1}{h} \right. \\ &\quad \left. - (l+j-1)!L \sum_{h \geq 0} \frac{(-1)^{h-1}}{2^{l+h}-1} \binom{l+j+h-1}{h} + (l+j-1)!L2^{-l-j} \right] \\ &\quad + \frac{8}{L^3} \sum_{l \geq 1} \frac{(-1)^l}{l!(2^l-1)} \\ &\quad \times \left[-\Gamma(l) \left(\frac{\pi^2 - L^2}{12} - \frac{\gamma^2}{2} - \frac{L\gamma}{2} \right) + \Gamma'(l) \left(\gamma + \frac{L}{2} \right) + \frac{1}{2} \Gamma''(l) - L\gamma(1-2^{-l}) \right. \\ &\quad \left. - L \sum_{h \geq 0} \frac{1}{2^{l+h}-1} \frac{(-1)^h (l+h-1)!}{h!} H_{l+h} + L\gamma \sum_{h \geq 1} \frac{1}{2^h-1} \frac{(-1)^h (l+h-1)!}{h!} \right] \\ &\quad + \frac{4}{L^4} \left[\frac{11\pi^4}{360} + \frac{L^2\pi^2}{72} + \frac{L^4}{720} - 2L\gamma \sum_{h \geq 1} \frac{(-1)^{h-1} H_h}{h(2^h-1)} - L\gamma \left(\frac{L^2}{2} - \frac{\pi^2}{12} \right) \right]. \end{aligned}$$

Although a few simplifications in this expression are still possible, it is clear that the complexity of the expressions does not make it attractive enough to write more similar evaluations.

We can now compute the corrected values.

Theorem 3.3. *Taking the contribution of products of Fourier series into account, the corrected asymp-*

otic constant and oscillating parts of the centered moments of $J(n)$ are given by

$$\begin{aligned} \tilde{\mu}_{2,c} &= \tilde{\mu}_2 - c_1[0] = 1 - C_1 - C_2 + 2\frac{D}{L}, \\ \tilde{\mu}_{3,c} &= \tilde{\mu}_3 + 2c_2[0] + \frac{6}{L}\gamma c_1[0] - \frac{6}{L^2}c_3[0] \\ &= -C_1 + \frac{6\zeta(3)}{L^3} - 2C_3 + 2 + 2\frac{D}{L} - \frac{12}{L^2}D_1 - \frac{12}{L^2}\gamma D - \frac{11}{2L}\gamma + \frac{\gamma\pi^2}{L^3} - 3C_2 - \frac{4}{L}D_2 - \frac{4}{L}\log(3) \\ &\quad - \frac{6}{L^2}c_3[0], \\ \tilde{\mu}_{4,c} &= \tilde{\mu}_4 - \frac{1}{2}c_1[0] - 3c_4[0] + \frac{\pi^2}{L^2}c_1[0] + 6C_1c_1[0] + 6C_2c_1[0] \\ &\quad - \frac{12}{L^2}\gamma^2c_1[0] - \frac{12}{L}\gamma c_2[0] + \frac{24}{L^3}\gamma c_3[0] + \frac{12}{L^2}c_6[0] + \frac{12}{L^3}c_7[0] + \frac{12}{L^3}c_8[0] \\ \kappa_{2,c} &= -\sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} - \frac{2\gamma w_1}{L} + \frac{2}{L^2}\rho_1, \\ \kappa_{3,c} &= 2\sum_{l \neq 0} c_2[l]e^{-2l\pi i \log n} + \frac{6}{L}\gamma \sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} + \frac{(6\gamma^2 - \pi^2)w_1}{2L^2} \\ &\quad - \frac{6\gamma}{L^3}\rho_1 - \frac{6}{L^2}\sum_{l \neq 0} c_3[l]e^{-2l\pi i \log n} - \frac{3}{L^3}\rho_2 - \frac{3}{L^3}\rho_4, \\ \kappa_{4,c} &= \left[\frac{12\gamma C_2}{L} + \frac{12\gamma C_1}{L} - \frac{8\zeta(3)}{L^3} - \frac{4\gamma^3}{L^3} - \frac{\gamma}{L} \right] w_1 + \frac{L^2 - 12C_2L^2 - 12C_1L^2 + 12\gamma^2}{L^4}\rho_1 \\ &\quad + \frac{12}{L^4}\sum_{l \neq 0} e^{-2l\pi i \log n} \psi(\chi_l)\psi(1, \chi_l)\Gamma(\chi_l) + \frac{12\gamma}{L^4}\rho_4 \\ &\quad + \frac{4}{L^4}\sum_{l \neq 0} e^{-2l\pi i \log n} \psi(2, \chi_l)\Gamma(\chi_l) + \frac{4}{L^4}\sum_{l \neq 0} \rho_3 + \frac{12\gamma}{L^4}\rho_2 \\ &\quad - \frac{1}{2}\sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} - 3\sum_{l \neq 0} c_4[l]e^{-2l\pi i \log n} + \frac{\pi^2}{L^2}\sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} + 6C_1\sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} \\ &\quad + 6C_2\sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} - \frac{12}{L^2}\gamma^2\sum_{l \neq 0} c_1[l]e^{-2l\pi i \log n} - \frac{12}{L}\gamma\sum_{l \neq 0} c_2[l]e^{-2l\pi i \log n} \\ &\quad + \frac{24}{L^3}\gamma\sum_{l \neq 0} c_3[l]e^{-2l\pi i \log n} + \frac{12}{L^2}\sum_{l \neq 0} c_6[l]e^{-2l\pi i \log n} + \frac{12}{L^3}\sum_{l \neq 0} c_7[l]e^{-2l\pi i \log n} + \frac{12}{L^3}\sum_{l \neq 0} c_8[l]e^{-2l\pi i \log n}. \end{aligned}$$

Note that $\tilde{\mu}_{2,c}$ fits with the result given in [16].

4. m counters: Asymptotic independence of the m counters

In this section, we analyze the asymptotic properties of the RV $J(m, n)$.

We will prove that, asymptotically, the counters are independent with n/m items each. We must analyze the random variable $J(m, n) = \sum_1^m J_i(n)$ where $J_i(n)$ has the distribution $p_\eta(j)$ with η is now given by ν_i : the number of items arriving in counter i . The quantity ν_i is $\text{bin}[\tilde{n}, \tilde{n}(1 - 1/m)]$ with $\tilde{n} := n/m$. Actually, $\{\nu_1, \dots, \nu_m\}$ is given by a multinomial distribution. We know that we can construct a “box”

$$[\tilde{n} - \tilde{n}^\theta, \tilde{n} + \tilde{n}^\theta]^m, \quad \frac{1}{2} < \theta < 1, \tag{18}$$

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 will analyze

$$J(m, n) - m \log(\tilde{n}) = \sum_{i=1}^m X_i, \quad \text{with } X_i := J_i - \log(\tilde{n}).$$

The rate of convergence is analyzed as follows.

4.1.

Let us first assume that ν_i is exactly given by its mean \tilde{n} . As mentioned in the previous section, the rate of convergence problem is solved in Flajolet [2].

4.2.

Now we assume that we are inside the box (18). We drop the $\tilde{}$ sign from \tilde{n} for convenience. Let $0 < \varepsilon < 1$. We must bound

$$S_2 := \sum_j j^k |p_n(j) - p_{n+\xi n^\theta}(j)|,$$

with $|\xi| < 1$. Note that $n + \xi n^\theta = n[1 + \xi n^{\theta-1}]$. Set $1 > \beta > \theta$.

- For $j < \beta \log n$, we have

$$\begin{aligned} 2^{-\eta} &> n^{1-\beta}, \\ |p_n(j)| &\leq \frac{1}{Q^2} \sum_{k=0}^{\infty} \frac{1}{2^{k(k-1)/2}} \exp(-n^{1-\beta} 2^k) = \mathcal{O}(\exp(-n^{1-\beta})), \\ |p_{n+\xi n^\theta}(j)| &\leq \frac{1}{Q^2} \sum_{k=0}^{\infty} \frac{1}{2^{k(k-1)/2}} \exp(-n^{1-\beta}(1 + \xi n^{\theta-1}) 2^k) = \mathcal{O}(\exp(-n^{(1-\beta)(1-\varepsilon)})), \\ |p_n(j) - p_{n+\xi n^\theta}(j)| &= \mathcal{O}(\exp(-n^{(1-\beta)(1-\varepsilon)})). \end{aligned}$$

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

$$\begin{aligned} 2^{-\eta} &> \frac{1}{n}, \\ (1 - 1/2^j)^n - (1 - 1/2^j)^{n+\xi n^\theta} &= (1 - 1/2^j)^n (1 - (1 - 1/2^j)^{\xi n^\theta}) = \mathcal{O}(n^\theta/2^j). \end{aligned}$$

We use again the “sum splitting technique.” Set $r = \sqrt{2 \log n}$.

1. Truncating the sum in (2) to $k \geq r$ leads to an error E_1 :

$$E_1 \leq \frac{1}{Q^2} \sum_{k=r}^{\infty} \frac{1}{2^{k(k-1)/2}} \left[\exp(-\frac{2^k}{n}) + \exp(-\frac{1 + \xi n^{\theta-1}}{n} 2^k) \right] = \mathcal{O}\left(\frac{1}{n}\right)$$

2. The remaining sum $k \leq r$ leads to

$$\begin{aligned} E_2 &\leq \frac{1}{Q^2} \sum_{k=0}^r \frac{1}{2^{k(k-1)/2}} \left[(1 - 1/2^{j-k})^n - (1 - 1/2^{j-k})^{n+\xi n^\theta} \right] \\ &= \sum_{k=0}^r \mathcal{O}\left(\frac{n^\theta}{2^{j-k}}\right) = \mathcal{O}\left(\frac{1}{n^{(\beta-\theta)(1-\varepsilon)}}\right). \end{aligned}$$

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

1.

$$E_1 = \mathcal{O}\left(\frac{2^{-x}}{n}\right).$$

2.

$$E_2 \leq \frac{1}{Q^2} \sum_{k=0}^r \frac{1}{2^{k(k-1)/2}} \frac{n^\theta}{2^{j-k}} = \mathcal{O}\left(\frac{n^\theta}{n^{2(1-\varepsilon)} 2^{x(1-\varepsilon)}}\right).$$

Now we come to S_2 . We get

$$\begin{aligned} S_2 &= \mathcal{O}\left((\beta \log n)^{k+1} \exp(-n^{(1-\beta)(1-\varepsilon)})\right) + (2 \log n)^{k+1} \mathcal{O}\left(\frac{1}{n^{(\beta-\theta)(1-\varepsilon)}}\right) \\ &+ \mathcal{O}\left(\sum_{x \geq 0} (2 \log n + x)^k \mathcal{O}\left(\frac{2^{-x}}{n}\right)\right) = \mathcal{O}\left(\frac{1}{n^{(\beta-\theta)(1-\varepsilon)}}\right); \end{aligned}$$

(not with the same ε , of course).

4.3.

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

$$S_3 := \sum_j p_\nu(j) j^k \exp(-Cn^{2\theta-1})$$

is small. We notice that $1 < \nu/n < m$. But, by the rate of convergence proved in [2], $\sum_j p_\nu(j) j^k$ is asymptotically bounded by $\mathcal{O}((\log \nu)^k) = \mathcal{O}((\log(nm))^k)$ and S_3 is asymptotically small.

4.4.

The last case to consider is $0 < \nu < n - n^\theta$. We have here $0 < \nu/n < 1$. The analysis proceeds like in the previous subsection. We therefore omit the details.

In conclusion, as S_2 and S_3 are asymptotically small, we can assume that ν_i can be deterministically chosen as \tilde{n} for all i , and that the counters are asymptotically independent.

5. m counters: Asymptotic moments

If J_i are iid RV, with asymptotic centered moments $\mu_k = \tilde{\mu}_k + \kappa_k$, then $J = \sum_{i=1}^m J_i$ has asymptotic distribution given by the convolution $f(\eta)^{(m)}$, mean

$$m \log(\tilde{n}) + m \mathbf{m}_1$$

and asymptotic centered moments $\mu_k(m)$ given by

$$\begin{aligned} \mu_2(m) &= m \mu_2, \\ \mu_3(m) &= m \mu_3, \\ \mu_4(m) &= m[\mu_4 + 3(m-1)\mu_2^2]. \end{aligned}$$

For $\mu_2(m)$ and $\mu_3(m)$, we immediately use μ_2 and μ_3 . For $\mu_4(m)$, we have $\mu_2 = \tilde{\mu}_2 + \kappa_2$, so $\mu_2^2 = \tilde{\mu}_2^2 + \kappa_2^2 + 2\tilde{\mu}_2\kappa_2$. Hence

$$\begin{aligned}\tilde{\mu}_4(m) &= m[\tilde{\mu}_4 + 3(m-1)\tilde{\mu}_2^2], \\ \kappa_4(m) &= m[\kappa_4 + 3(m-1)(\kappa_2^2 + 2\tilde{\mu}_2\kappa_2)].\end{aligned}$$

Also, we have

$$\begin{aligned}[\kappa_2^2]_0 &= c_4[0] + \frac{4\gamma^2 c_1[0]}{L^2} + \frac{4c_5[0]}{L^4} + \frac{4\gamma c_2[0]}{L} - \frac{4c_6[0]}{L^2} - \frac{8c_3[0]}{L^3}, \\ [\kappa_2^2]_k &= c_4[k] + \frac{4\gamma^2 c_1[k]}{L^2} + \frac{4c_5[k]}{L^4} + \frac{4\gamma c_2[k]}{L} - \frac{4c_6[k]}{L^2} - \frac{8c_3[k]}{L^3}.\end{aligned}$$

The corrected moments must now be computed. The interesting case is the fourth moment, since here the dependency on m is more involved: we obtain our last theorem.

Theorem 5.1. *Taking the contribution of products of Fourier series into account, the asymptotic constant and oscillating parts of the corrected fourth centered moment of J are given by*

$$\begin{aligned}\tilde{\mu}_{4,c}(m) &= m[\tilde{\mu}_{4,c} + 3(m-1)\tilde{\mu}_2^2] + 3m(m-1)[[\kappa_2^2]_0 - 2\tilde{\mu}_2 c_1[0]], \\ \kappa_{4,c}(m) &= m\left[\kappa_{4,c} + 3(m-1)\sum_{l \neq 0} [\kappa_2^2]_l e^{-2l\pi i \log n} + 3(m-1)2\tilde{\mu}_2 \kappa_{2,c}\right].\end{aligned}$$

6. Conclusion

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 asymptotic distributions and higher moments.

Acknowledgment: We would like to thank two referees for many useful suggestions that improved the paper.

References

- [1] J. Cichoń and W. Macyna, *Approximate counters for flash memory*, 17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), Toyama, Japan, 2011.
- [2] P. Flajolet, *Approximate counting: A detailed analysis*, BIT, 25, 113–134, 1985.
- [3] P. Flajolet and R. Sedgewick, *Digital search trees revisited*, SIAM J. Comput., 3, 748–767, 1986.
- [4] M. Fuchs, C. K. Lee, and H. Prodinger, *Approximate counting via the Poisson-Laplace-Mellin method*, DMTCS, proc AQ, 13–28, 2012.
- [5] H.-K. Hwang, M. Fuchs and V. Zacharovas, *Asymptotic variance of random symmetric digital search trees*, Discrete Math. Theor. Comput. Sci., 12, 103–166, 2010.
- [6] P. Kirschenhofer and H. Prodinger, *Approximate counting: An alternative approach*, RAIRO Theor. Inform. Appl., 25, 43–48, 1991.
- [7] D. E. Knuth, *The art of computer programming, volume 3: Sorting and searching*, Addison-Wesley, 1973, (Second edition, 1998).

- [8] M. Loève, *Probability theory, 3rd ed*, D. Van Nostrand, 1963.
- [9] G. Louchard, *Exact and asymptotic distributions in digital and binary search trees*, RAIRO Theor. Inform. Appl., 21(4), 479–496, 1987.
- [10] G. Louchard and H. Prodinger, *Asymptotics of the moments of extreme-value related distribution functions*, Technical report, 2005. Long version: <http://www.ulb.ac.be/di/mcs/louchard/moml.ps>.
- [11] G. Louchard and H. Prodinger, *Asymptotics of the moments of extreme-value related distribution functions*, Algorithmica, 46, 431–467, 2006.
- [12] G. Louchard and H. Prodinger, *Generalized approximate counting revisited*, Theoret. Comput. Sci., 391, 109–125, 2008.
- [13] H. Prodinger, *Hypothetic analyses: Approximate counting in the style of Knuth, path length in the style of Flajolet*, Theoret. Comput. Sci., 100, 243–251, 1992.
- [14] H. Prodinger, *Approximate counting via Euler transform*, Mathematica Slovaca, 44, 569–574, 1994.
- [15] H. Prodinger, *Periodic oscillations in the analysis of algorithms and their cancellations*, J. Iran. Stat. Soc., 3, 251–270, 2004.
- [16] H. Prodinger, *Approximate counting with m counters: a detailed analysis*, Theoret. Comput. Sci., 439, 58–68, 2012.