

## A BINOMIAL SPLITTING PROCESS IN CONNECTION WITH CORNER PARKING PROBLEMS

MICHAEL FUCHS,\* *National Chiao Tung University*

HSIEN-KUEI HWANG,\*\* *Academia Sinica*

YOSHIAKI ITOH,\*\*\* *Institute of Statistical Mathematics*

HOSAM H. MAHMOUD,\*\*\*\* *The George Washington University*

### Abstract

This paper studies a special type of binomial splitting process. Such a process can be used to model a high dimensional corner parking problem as well as determining the depth of random PATRICIA (practical algorithm to retrieve information coded in alphanumeric) tries, which are a special class of digital tree data structures. The latter also has natural interpretations in terms of distinct values in independent and identically distributed geometric random variables and the occupancy problem in urn models. The corresponding distribution is marked by a logarithmic mean and a bounded variance, which is oscillating, if the binomial parameter  $p$  is not equal to  $\frac{1}{2}$ , and asymptotic to one in the unbiased case. Also, the limiting distribution does not exist as a result of the periodic fluctuations.

*Keywords:* Binomial distribution; parking problem; periodic fluctuation; asymptotic approximation; digital tree; de-Poissonization

2010 Mathematics Subject Classification: Primary 60C05; 68W40  
Secondary 60F05

### 1. Introduction

In this paper we study the random variable  $X_n$ , defined recursively by

$$X_n \stackrel{D}{=} X_{I_n} + 1 \quad \text{for } n \geq 1, \quad (1)$$

with  $X_0 = 0$ , where  $(X_n)$  and  $(I_n)$  are independent and

$$\mathbb{P}(I_n = k) = \binom{n}{k} \frac{p^k q^{n-k} - p^k (q-p)^{n-k}}{1 - q^n} \quad \text{for } k = 0, \dots, n-1,$$

Received 26 July 2013; revision received 24 November 2013.

\* Postal address: Department of Applied Mathematics, National Chiao Tung University, Hsinchu, 300, Taiwan.  
Email address: mfuchs@math.nctu.edu.tw

\*\* Postal address: Institute of Statistical Science and Institute of Information Science, Academia Sinica, Taipei, 115, Taiwan.

\*\*\* Postal address: Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo, 190-8562, Japan.

\*\*\*\* Postal address: Department of Statistics, The George Washington University, Washington, DC 20052, USA.

where, throughout this paper,  $0 < p \leq q := 1 - p$ . In particular,

$$p = \frac{1}{2} \implies \mathbb{P}(I_n = k) = \binom{n}{k} \frac{1}{2^n - 1},$$

$$p = \frac{1}{3} \implies \mathbb{P}(I_n = k) = \binom{n}{k} \frac{2^{n-k} - 1}{3^n - 2^n} \quad \text{for } k = 0, \dots, n - 1.$$

Note that, for convenience, we retain the case  $k = 0$ , but drop  $k = n$ .

The random variable  $X_n$  originally arose out of the analysis of a special type of parking problem with ‘corner preference’ (described in Section 2). It also arose in a leader election algorithm, in which a truncated binomial number of contestants were advanced at each stage during a contest. Namely,  $X_n$  is the number of rounds taken until the election comes to an end; see [13] and [14] for a broad framework for these types of problems.

Moreover, it turns out that when  $p = \frac{1}{2}$  the distribution of  $X_n$  is identical to that of two of the parameters in a random symmetric PATRICIA (practical algorithm to retrieve information coded in alphanumeric) trie: the depth and the length of the left arm. The depth is the distance between a uniformly chosen leaf node and the root and the left arm is the path that starts at the root and keeps going left until no more nodes can be found on the left; see, for example, [15], [21]. Also, the depth, or the left arm, of random PATRICIA tries is identically distributed as the number of distinct values in some random sequences, see [1], and the number of occupied urns in some urn models, see [7].

In this paper we prove that the random variable  $X_n$  has both a logarithmic mean and a bounded variance for large  $n$ . Also, the distributions do not approach a fixed limit law due to the fluctuations present in the distributions. For a similar context, see [12], [13], [14] and the references therein.

### 2. A corner preference parking problem in discrete space

The parking problem has a long history in the discrete probability literature and it is closely connected to many applications and models in chemistry, physics, biology and computer algorithms; see, for example, [2] and [22]. Most analytic results for the numerous variants in the literature are for the one dimensional setting with very few authors dealing with higher dimensions due to the intrinsic complexity of the corresponding problems.

First, we will give an explanation of a simple discrete parking problem. Integral translates of the cube  $[0, \ell]^n$  are ‘parked’ into the  $n$ -dimensional hypercube  $[0, L]^n$ , where  $L > \ell \geq 1$ . A precise mathematical formulation of this is as follows. Represent cubes by their corner which has the shortest distance to the origin. Moreover, set

$$Z_{L-\ell}^n := \{\mathbf{a} = (a_1, \dots, a_n) : a_j = 0, 1, \dots, L - \ell\} \quad \text{for } 1 \leq j \leq n,$$

and define a distance  $\rho(\mathbf{x}, \mathbf{y}) := \max_{1 \leq j \leq n} |x_j - y_j|$  between two points  $\mathbf{x}, \mathbf{y} \in Z_{L-\ell}^n$ . Initially, choose one point uniformly at random from  $Z_{L-\ell}^n$  and record it as  $\mathbf{a}(1)$ . Then choose another point uniformly at random and record it as  $\mathbf{a}(2)$  if  $\ell \leq \rho(\mathbf{a}(1), \mathbf{a}(2))$ ; otherwise, reject it and repeat the same procedure. If  $\mathbf{a}(1), \mathbf{a}(2), \dots, \mathbf{a}(k)$  are already recorded then choose the next point uniformly at random and record it as  $\mathbf{a}(k+1)$  if  $\ell \leq \rho(\mathbf{a}(k+1), \mathbf{a}(j))$  for  $j = 1, 2, \dots, k$ ; otherwise, reject it and repeat the same procedure. We continue this procedure until it is impossible to add more points among the  $(L - \ell + 1)^n$  points. Since analytic development of this model remains challenging, simulations have been carried out in order to determine the jamming density of this model; see, for example, [9] and [10].

We further restrict the parking so that it operates towards one corner, which we call ‘corner preference parking’. More precisely, let

$$S(\mathbf{a}) := \{\mathbf{x} : \mathbf{x} \in \mathbb{Z}_{L-\ell}^n, 0 \leq x_j \leq a_j, j = 1, \dots, n\},$$

$$U(\mathbf{a}, \ell) := \{\mathbf{x} : \mathbf{x} \in \mathbb{Z}_{L-\ell}^n, \rho(\mathbf{x}, \mathbf{a}) < \ell\} \text{ and } S_U(\mathbf{a}, \ell) := S(\mathbf{a}) \setminus U(\mathbf{a}, \ell).$$

The corner preference parking problem then starts from  $S_U(\mathbf{a}(\mathbf{1}), \ell)$  with

$$\mathbf{a}(\mathbf{1}) = (L - \ell, L - \ell, \dots, L - \ell) \in \mathbb{Z}_{L-\ell}^n.$$

Subsequently, we place, sequentially at random, the integral translates of cube  $[0, \ell]^n$  into the cube  $[0, L]^n$ , so that any car placed will be closer to the ‘corner’ (origin) than the previously placed cubes. This is repeated until there is no possible space to park. By ‘closer to the corner’ we mean that the coordinates of the point representing the car are all, at most, as large as the car parked immediately before it. The process continues until saturation is achieved.

We now take  $L = 2m$ , and  $\ell = m$ , where  $m \geq 1$ . Assume that all possible ‘parking positions’ are equally likely at each stage. Let the random variable  $Y_n$  be the number of cars parked (after the first car) at the time of saturation in our  $n$ -dimensional corner parking problem. The distribution of  $Y_n$  can be explicitly characterized.

**Lemma 1.** *The random variable  $Y_n$  can be recursively enumerated, with  $Y_0 = 0$ , by*

$$\mathbb{E}(u^{Y_n}) = u \sum_{1 \leq k \leq n} \binom{n}{k} \frac{m^k - (m - 1)^k}{(m + 1)^n - m^n} \mathbb{E}(u^{Y_{n-k}}) \text{ for } n \geq 1. \tag{2}$$

This corresponds to (1) with  $p = 1/(m + 1)$ .

*Proof.* There are  $(m + 1)^n$  possible positions of integral translates of hypercubes (cars)  $[0, m]^n$  to park in the hypercube  $[0, 2m]^n$ . After parking the first car at the top right corner the number of possible positions for the second car  $\mathbf{a}(2)$  is equal to  $(m + 1)^n - m^n$ , which is the denominator in (2).

Suppose that  $n - k$  coordinates of  $\mathbf{a}(2)$  assume the value  $m$ , and each of the remaining  $k$  coordinates may assume any of the values  $\{0, 1, \dots, m - 1\}$ . Observe that at least one of the  $k$  coordinates should be 0 for the second car to park without overlapping with the first car. Since the number of cases with each of the  $k$  coordinates taking a value in  $\{1, 2, \dots, m - 1\}$  is  $(m - 1)^k$ , the number of all possible positions for the second car (under the mentioned restriction) is  $m^k - (m - 1)^k$ . We have  $\binom{n}{k}$  choices of the  $k$  coordinates. Thus, under this restriction, the second car has a total of  $\binom{n}{k}(m^k - (m - 1)^k)$  possible positions to park. After this placement the problem is reduced to that of an  $(n - k)$ -dimensional one; see Figure 1 for an illustration with  $L = 4$  and  $\ell = 2$ .

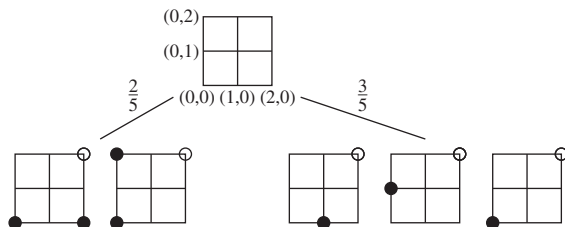


FIGURE 1: The five  $(3^2 - 2^2)$  different two dimensional configurations of corner preference parking, when  $L = 4$  and  $\ell = 2$ . In this case  $\mathbb{E}(u^{X_2}) = \frac{3}{5}u + \frac{2}{5}u^2$ .

### 2.1. Depths of PATRICIA tries

Tries (a mixture of tree and *retrieval*) are one of the most useful tree structures in storing alphabetical or digital data in computer algorithms. The underlying construction principle is simply ‘0-bit directing to the left’ and ‘1-bit directing to the right’. PATRICIA tries are a variant of tries where all nodes with only a single child are compressed; see Figure 2 for a plot of tries and PATRICIA tries and the book [17] for more information. Note that, unlike tries, the number of internal nodes of which is not necessarily a constant, a PATRICIA trie of  $n$  keys always has  $n - 1$  internal nodes for branching purposes. This is a standard property of a binary tree.

To study the shapes of random PATRICIA tries we assume that the input is a sequence of  $n$  independent and identically distributed (i.i.d.) random variables, each composed of an infinite sequence of Bernoulli random variables with mean  $p$ ,  $0 < p < 1$ . Under such a Bernoulli model, we construct random PATRICIA tries and the shape parameters become random variables.

Consider the depth  $Z_n$  of a random PATRICIA trie of  $n$  keys under the Bernoulli model, where the depth denotes the distance between the root and a randomly chosen key (from the leaves where keys are stored), where the  $n$  keys are equally likely to be selected. Then we have the recurrence relation for the probability generating function of  $Z_n$ :

$$\mathbb{E}(u^{Z_n}) = u \sum_{1 \leq k < n} \frac{\binom{n}{k} p^k q^{n-k}}{1 - p^n - q^n} \left( \frac{k}{n} \mathbb{E}(u^{Z_k}) + \frac{n-k}{n} \mathbb{E}(u^{Z_{n-k}}) \right) \quad \text{for } n \geq 2,$$

with  $Z_0 = Z_1 = 0$ . In the unbiased case  $p = \frac{1}{2}$ , this reduces to

$$\mathbb{E}(u^{Z_n}) = u \sum_{0 \leq k \leq n-2} \frac{\binom{n-1}{k}}{2^{n-1} - 1} \mathbb{E}(u^{Z_{k+1}}),$$

which implies that

$$Z_{n+1} \stackrel{D}{=} X_n \quad \text{for } n \geq 1, \quad p = \frac{1}{2}.$$

Another identically distributed random variable is the length  $W_n$  of the ‘left arm’, which is the path starting from the root and going always to the left until reaching a key node. Then, under the Bernoulli model,

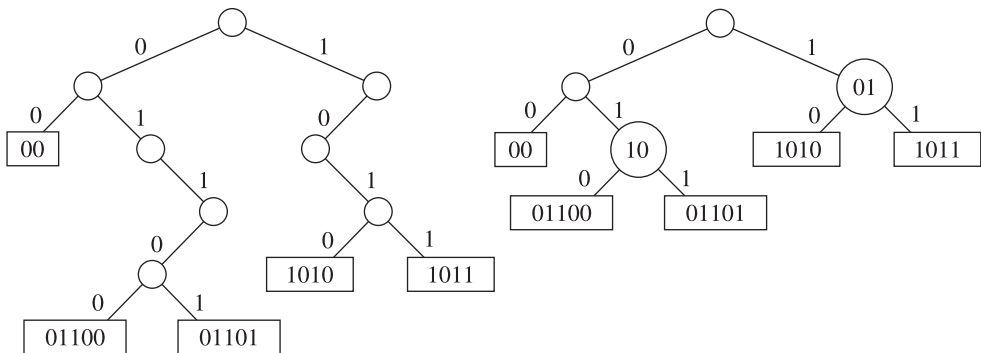


FIGURE 2: A trie (left) of  $n = 5$  records and the corresponding PATRICIA tries (right) where the circles represent internal nodes and the rectangles holding the records are external nodes. The compressed bits are also indicated on the nodes.

$$\mathbb{E}(u^{W_n}) = u \sum_{1 \leq k \leq n} \frac{\binom{n}{k} p^k q^{n-k}}{1 - q^n} \mathbb{E}(u^{W_{n-k}}) \quad \text{for } n \geq 2, \tag{3}$$

with  $W_0 = 0$  and  $W_1 = 1$ . It is obvious that  $W_n \stackrel{D}{=} X_n$ , when  $p = \frac{1}{2}$ .

**2.2. Distinct values and urn models**

The left arm  $W_n$  has two other different interpretations. One is in terms of the number of distinct letters in a sequence of i.i.d. geometric random variables with success probability  $p$  for which we have exactly the recurrence (3); see [1]. Alternatively, if we consider the urn model where the  $j$ th urn has probability of  $pq^j$  of receiving a ball, then the number of occupied urns also follows the same distribution; see [7].

**3. Mean of  $X_n$**

**3.1. Generating functions**

The exponential generating function for the moment generating function of  $X_n$ ,

$$P(z, y) := \sum_{n \geq 0} \frac{\mathbb{E}(e^{X_n y})}{n!} z^n,$$

satisfies, by (1), the functional equation

$$P(z, y) = e^y (e^{qz} - e^{(q-p)z}) P(pz, y) + P(qz, y).$$

It follows that the exponential generating function of the mean  $f_1(z) := \sum_{n \geq 0} \mathbb{E}(X_n) z^n / n!$  satisfies

$$f_1(z) = (e^{qz} - e^{(q-p)z}) f_1(pz) + f_1(qz) + e^z - e^{qz},$$

with  $f_1(0) = 0$ . By iteration, we obtain

$$\begin{aligned} f_1(z) = & \sum_{\substack{k \geq 0 \\ 0 \leq j \leq k}} (\exp(p^{k-j} q^j z) - \exp(p^{k-j} q^{j+1} z)) \\ & \times \sum_{0 \leq i_1 \leq \dots \leq i_{k-j} \leq j} \prod_{0 \leq \ell < k-j} (\exp(p^\ell q^{i_\ell+1} z) - \exp((q-p)p^\ell q^{i_\ell+1} z)), \end{aligned}$$

which does not seem useful for further manipulation.

For our asymptotic purposes it is technically more convenient to consider the Poisson generating function,

$$\tilde{P}(z, y) := e^{-z} P(z, y),$$

which then satisfies the equation

$$\tilde{P}(z, y) = e^y (1 - e^{-pz}) \tilde{P}(pz, y) + e^{-pz} \tilde{P}(qz, y). \tag{4}$$

It follows that the Poisson generating function for the  $m$ th moment

$$\tilde{f}_m(z) := e^{-z} \sum_{n \geq 0} \frac{\mathbb{E}(X_n^m)}{n!} z^n,$$

satisfies the equation

$$\tilde{f}_m(z) = (1 - e^{-pz}) \sum_{0 \leq \ell \leq m} \binom{m}{\ell} \tilde{f}_\ell(pz) + e^{-pz} \tilde{f}_m(qz) \quad \text{for } m \geq 0,$$

where  $\tilde{f}_0(z) = 1$ .

In particular, we have

$$\tilde{f}_1(z) = (1 - e^{-pz})\tilde{f}_1(pz) + e^{-pz}\tilde{f}_1(qz) + 1 - e^{-pz}, \tag{5}$$

$$\tilde{f}_2(z) = (1 - e^{-pz})\tilde{f}_2(pz) + e^{-pz}\tilde{f}_2(qz) + 2(1 - e^{-pz})\tilde{f}_1(pz) + 1 - e^{-pz}. \tag{6}$$

### 3.2. Asymptotics of the mean

Let

$$\phi(z) = e^{-z}(\tilde{f}_1(p^{-1}qz) - \tilde{f}_1(z)), \tag{7}$$

and let  $\phi^*(s)$  denote its Mellin transform (see [4])

$$\phi^*(s) := \int_0^\infty e^{-t} t^{s-1} (\tilde{f}_1(p^{-1}qt) - \tilde{f}_1(t)) dt, \tag{8}$$

which is well defined in the half-plane  $\text{Re}(s) > -1$  (see Appendix A for growth properties of  $\tilde{f}_1$ ).

**Theorem 1.** *The expected value of  $X_n$  satisfies*

$$\mathbb{E}(X_n) = \log_{1/p} n + \frac{\gamma + \phi^*(0)}{\log(1/p)} - \frac{1}{2} + Q(\log_{1/p} n) + O(n^{-1}).$$

Here  $\gamma$  denotes Euler’s constant and

$$Q(u) := \sum_{k \in \mathbb{Z} \setminus \{0\}} Q_k e^{-2k\pi i u}, \quad Q_k := -\frac{\Gamma(\chi_k) - \phi^*(\chi_k)}{\log(1/p)}, \tag{9}$$

where  $\chi_k := 2k\pi i / \log(1/p)$ , and  $\Gamma$  denotes the Gamma function.

The asymptotic expansion simplifies when  $p = \frac{1}{2}$ ; indeed, in this case, we have the closed-form expression

$$\tilde{f}_1(z) = \sum_{k \geq 1} (1 - e^{-z/2^k}) \quad \text{for } \text{Re}(z) > 0.$$

**Corollary 1.** *In the symmetric case when  $p = \frac{1}{2}$ , the expected value of  $X_n$  satisfies asymptotically*

$$\mathbb{E}(X_n) = \log_2 n + \frac{\gamma}{\log 2} - \frac{1}{2} - \frac{1}{\log 2} \sum_{k \neq 0} \Gamma(\chi_k) n^{-\chi_k} + O(n^{-1}),$$

where  $\chi_k = 2k\pi i / \log 2$ .

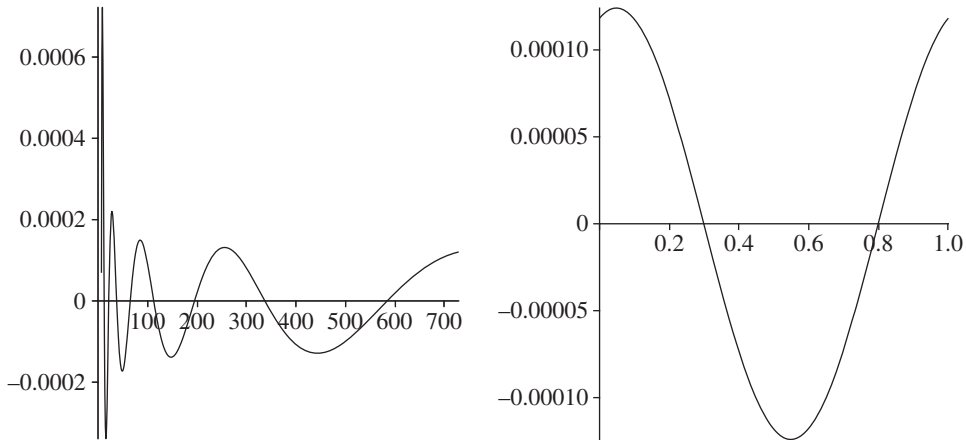


FIGURE 3: Fluctuation of the periodic function  $Q(\log_3 n)$  for  $p = \frac{1}{3}$  as approximated by  $\mu_n - H_n / \log 3 + \frac{1}{2} - \phi^*(0) / \log 3$  (left) and the first five terms of the Fourier series (9) (right).

For numerical purposes, the value of  $\phi^*(\chi_k)$  can be computed by the series expression

$$\phi^*(\chi_k) = \sum_{j \geq 1} \frac{\mu_j}{j!} \Gamma(\chi_k + j) (q^j - 2^{-j-\chi_k}) \quad \text{for } k = 0, 1, \dots$$

Approximate plots of the periodic function  $Q(u)$  for  $p = \frac{1}{3}$  based on exact values of  $\mu_n := \mathbb{E}(X_n)$  and on its Fourier series are shown in Figure 3.

*Outline of the proof of Theorem 1.* Theorem 1 is proved by a two stage purely analytic approach based on a Mellin transform and analytic de-Poissonization (see, for example, [3] and [11]). We outline the major steps and arguments used here, leaving the major technical justification in Appendix A.

Our starting point is the functional equation (5), which is rewritten as

$$\tilde{f}_1(z) = \tilde{f}_1(pz) + \phi(pz) + 1 - e^{-pz},$$

where  $\phi(z)$  is defined in (7). While  $\phi$  also involves  $\tilde{f}$ , we show that it is exponentially small for large complex parameter, and thus the asymptotics of  $\tilde{f}_1(z)$  can be readily derived by standard inverse Mellin transform arguments (growth order of the integrand at infinity and calculus of residues).

Once the asymptotics of  $\tilde{f}_1(z)$  for large  $|z|$  are known, we can apply the Cauchy integral formula

$$\mathbb{E}(X_n) = \frac{n!}{2\pi i} \oint_{|z|=n} z^{-n-1} e^z \tilde{f}_1(z) dz,$$

and the saddle-point method to derive the asymptotics of the mean. Roughly, the growth order of  $\tilde{f}_1$  is small which means that the saddle-point (where the derivative of the integrand becomes zero) lies near  $n$ . The specialization of the saddle-point method here (with integration contour  $|z| = n$ ) has many interesting properties and is often referred to as the analytic de-Poissonization (see the survey paper by Jacquet and Szpankowski [11]).

It turns out that such a Mellin and de-Poissonization process can be manipulated in a rather systematic and operational manner by introducing the notion of JS-admissible functions in

which we combine ideas from [6] and [11] (see also [5] and [8]). We can easily apply the same approach to characterize the asymptotics of the variance and the limiting distribution.

**3.3. Proof of Theorem 1**

Let

$$\mathcal{S}_\varepsilon := \left\{ z : |\arg(z)| \leq \frac{\pi}{2} - \varepsilon \right\} \quad \text{for } \varepsilon > 0. \tag{10}$$

By Proposition 3 (in Appendix A),  $\tilde{f}_1(z)$  is polynomially bounded for large  $|z|$  in the sector  $\mathcal{S}_\varepsilon$ . This means that  $\phi(z) + 1 - e^{-pz} = O(1)$  for  $|z| \geq 1$  in  $\mathcal{S}_\varepsilon$ . Consequently,  $\tilde{f}_1(z) = O(|\log z|)$  in the same range of  $z$ . Moreover, since  $\tilde{f}_1(z) \sim z$ , as  $z \rightarrow 0$ , we see that the Mellin transform

$$\tilde{f}_1^*(s) := \int_0^\infty \tilde{f}_1(z)z^{s-1} dz,$$

exists in the strip  $-1 < \text{Re}(s) < 0$ , and defines an analytic function there.

It follows from (5) that

$$\tilde{f}_1^*(s) = \frac{\Gamma(s) - \phi^*(s)}{1 - p^s} \quad \text{for } -1 < \text{Re}(s) < 0,$$

and  $\phi^*$  is defined in (8).

By the Mellin inversion formula,

$$\tilde{f}_1(z) = \frac{1}{2\pi i} \int_{-1/2-i\infty}^{-1/2+i\infty} \frac{\Gamma(s) - \phi^*(s)}{1 - p^s} z^{-s} ds.$$

We need the growth property of  $\phi^*(\sigma \pm it)$  for large  $|t|$ .

**Lemma 2.** For  $\sigma > -1$ ,

$$|\phi^*(\sigma \pm it)| = O(e^{-(\pi/2-\varepsilon)|t|}) \quad \text{as } |t| \rightarrow \infty.$$

*Proof.* This follows from the fact that  $\tilde{f}_1(z)$  is an entire function, the estimate  $\tilde{f}_1(z) = O(|\log z|)$  for  $z \in \mathcal{S}_\varepsilon$  and the exponential smallness lemma; see [4, Proposition 5].

Moreover, since

$$|\Gamma(\sigma \pm it)| = O(|t|^{\sigma-1/2} e^{-\pi|t|/2}),$$

for finite  $\sigma$  and  $|t| \rightarrow \infty$ , we can move the line of integration to the right, summing the residues of all poles encountered. The result is

$$\tilde{f}_1(z) = \log_{1/p} z + C + Q(\log_{1/p} z) + \frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} \frac{\Gamma(s) - \phi^*(s)}{1 - p^s} z^{-s} ds, \tag{11}$$

where (defining  $\chi_k := 2k\pi i / \log(1/p)$ )

$$C := -\frac{1}{2} + \frac{\gamma + \phi^*(0)}{\log(1/p)}. \tag{12}$$



Note that, by definition, we have

$$\phi^*(\chi_k) = \sum_{j \geq 1} \frac{\mu_j}{j!} \Gamma(j + \chi_k) (q^j - 2^{-j-\chi_k}) \quad \text{for } k \in \mathbb{Z},$$

the series being absolutely convergent by the growth order of  $\mu_j$ . In particular, when  $p = \frac{1}{3}$

$$\phi^*(0) = \sum_{j \geq 1} \frac{\mu_j}{j} \left( \frac{2^j}{3^j} - \frac{1}{2^j} \right) \approx 0.581\,309\,808\,352\,813\,440\,19\dots,$$

so that  $C \approx 0.554\,535\,330\,802\,526\,966\,05\dots$

To evaluate the remainder integral in (11), we expand the factor  $1/(1 - p^s)$  (since  $\text{Re}(s) > 0$ ) into a geometric series, and integrate term by term, giving

$$\frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} \frac{\Gamma(s) - \phi^*(s)}{1 - p^s} z^{-s} ds = \sum_{k \geq 0} e^{-p^{-k}z} (1 - \tilde{f}_1(qp^{-k-1}z) + \tilde{f}_1(p^{-k}z)).$$

Thus the remainder is indeed exponentially small.

We summarize these derivations as follows.

**Proposition 1.** *For  $z$  lying in the sector  $\mathcal{S}_\varepsilon$ ,  $\tilde{f}_1(z)$  satisfies the asymptotic and exact equation*

$$\tilde{f}_1(z) = \log_{1/p} z + C + Q(\log_{1/p} z) + \sum_{k \geq 0} e^{-p^{-k}z} (1 - \tilde{f}_1(qp^{-k-1}z) + \tilde{f}_1(p^{-k}z)), \quad (13)$$

where  $C$  and  $Q$  are given in (12) and (9), respectively.

Theorem 1 follows from a standard de-Poissonization argument (see Appendix A) and (13)

$$\mathbb{E}(X_n) = \tilde{f}_1(n) - \frac{n}{2} \tilde{f}_1''(n) + O(n^{-2}).$$

#### 4. The variance

For the asymptotics of the variance, it proves advantageous to consider suitable Poissonized variance at the generating function level, which, in the case of  $X_n$ , can be handled by the following form

$$\tilde{V}(z) := \tilde{f}_2(z) - \tilde{f}_1(z)^2,$$

where the Poisson generating function of the second moment  $\tilde{f}_2(z)$  satisfies (6). Then  $\tilde{V}(z)$  satisfies the functional equation, with  $\tilde{V}(0) = 0$

$$\tilde{V}(z) = (1 - e^{-pz})\tilde{V}(pz) + e^{-pz}\tilde{V}(qz) + g_V(z), \quad (14)$$

where

$$g_V(z) := e^{-pz}(1 - e^{-pz})(1 + \tilde{f}_1(pz) - \tilde{f}_1(qz))^2.$$

Unlike  $g$  and  $g_2$ , which is  $O(1)$  for large  $z$ ,  $g_V$  is exponentially small for large  $z$ .

When  $p = \frac{1}{2}$ , we see that (14) has the closed form solution

$$\tilde{V}(z) = 1 - e^{-z}.$$

When  $p \neq q$ , define

$$\phi_V^*(s) := \int_0^\infty e^{-z} z^{s-1} (\tilde{V}(p^{-1}qz) - \tilde{V}(z) + (1 - e^{-z})(1 + \tilde{f}_1(z) - \tilde{f}_1(p^{-1}qz))^2) dz.$$

By following exactly the same analysis as that for  $\tilde{f}_1$ , we obtain

$$\begin{aligned} \tilde{V}(z) = Q_V(\log_{1/p} z) + \sum_{k \geq 0} e^{-p^{-k}z} \{ \tilde{V}(p^{-k-1}qz) - \tilde{V}(p^{-k}z) \\ + (1 - e^{-p^{-k}z})(1 + \tilde{f}_1(p^{-k}z) - \tilde{f}_1(p^{-k-1}qz))^2 \}, \end{aligned} \tag{15}$$

for  $\text{Re}(z) > 0$ , where

$$Q_V(u) = \frac{1}{\log(1/p)} \sum_{k \in \mathbb{Z}} \phi_V^*(\chi_k) e^{-2k\pi i u}.$$

Note that  $Q_V(u) = 1$  when  $p = \frac{1}{2}$ .

**Theorem 2.** *If  $p = \frac{1}{2}$ , then the variance of  $X_n$  satisfies*

$$\mathbb{V}(X_n) = 1 + O(n^{-1}),$$

*if  $p \neq q$ , then the variance of  $X_n$  is bounded and asymptotically periodic in nature*

$$\mathbb{V}(X_n) = Q_V(\log_{1/p} n) + O(n^{-1}).$$

*Proof.* By the definition of  $\tilde{V}$

$$\begin{aligned} \mathbb{V}(X_n) &= n! [z^n] e^z \tilde{f}_2(z) - (\mathbb{E}(X_n))^2 \\ &= n! [z^n] e^z \tilde{V}(z) - n! [z^n] e^z \tilde{f}_1(z)^2 - (n! [z^n] e^z \tilde{f}_1(z))^2, \end{aligned}$$

which, by the asymptotic nature of the Poisson–Charlier expansions (see Appendix A), is asymptotic to

$$\mathbb{V}(X_n) = \tilde{V}(n) + O(n\tilde{V}''(n) + n\tilde{f}'_1(n)^2) = \tilde{V}(n) + O(n^{-1}),$$

and the theorem follows from (15).

Figure 4 illustrates the periodic fluctuations of the variance when  $p = \frac{1}{3}$ . See also [20] for a similar situation where the variance is not oscillating when  $p = \frac{1}{2}$ .

For computational purposes we use the series expression

$$\begin{aligned} \phi_V^*(\chi_k) &= \sum_{j \geq 1} \frac{\mathbb{E}(X_j^2)}{j!} \Gamma(j + \chi_k) (q^j - 2^{-j-\chi_k}) + \Gamma(\chi_k) (1 - 2^{-\chi_k}) \\ &\quad + \sum_{j \geq 1} \frac{\mu_j^{[2]}}{j!} \Gamma(j + \chi_k) (2 \cdot 3^{-j-\chi_k} - 4^{-j-\chi_k} - q^j 2^{-j-\chi_k}) \\ &\quad + 2 \sum_{j \geq 1} \frac{\mu_j}{j!} \Gamma(j + \chi_k) (2^{-j-\chi_k} - 3^{-j-\chi_k} - q^j + q^j (1+p)^{-j-\chi_k}) \\ &\quad - 2 \sum_{j \geq 1} \frac{\mu_j^{[11]}}{j!} \Gamma(j + \chi_k) ((1+p)^{-j-\chi_k} - (1+2p)^{-j-\chi_k}), \end{aligned}$$

where  $\mu_n^{[2]} := n! [z^n] f_1(z)^2$  and  $\mu_n^{[11]} := n! [z^n] f_1(pz) f_1(qz)$ .

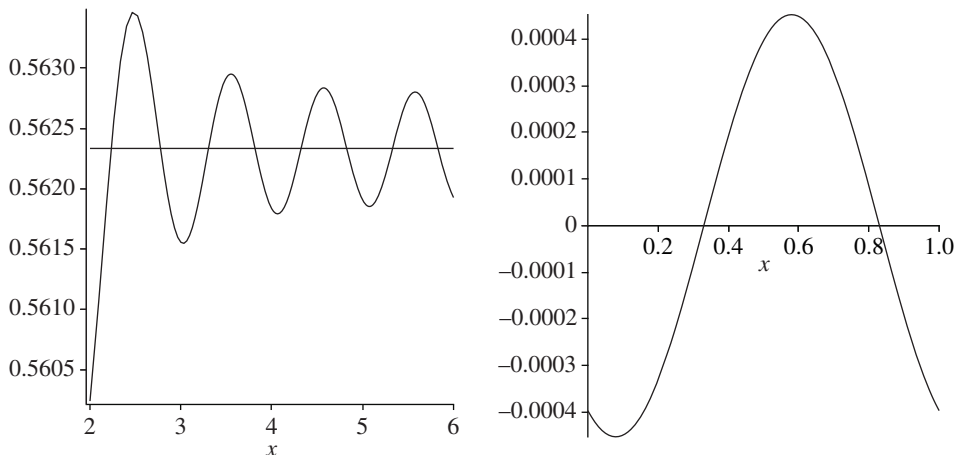


FIGURE 4: Fluctuation of the periodic function  $Q_V(\log_3 n)$ ,  $p = \frac{1}{3}$ , as approximated by  $\mathbb{V}(X_n) + c_0/n - c_0/(2n^2)$  in logarithmic scale (left) and the first four oscillating terms of its Fourier series (right). Here the number  $c_0 = 1/\log(1/p)^2$  and the two additional terms  $c_0/n - c_0/(2n^2)$  are chosen for a better numerical correction and graphical display.

### 5. Asymptotic distribution

We now show that the distribution of  $X_n$  is asymptotically fluctuating and no convergence to a fixed limit law is possible. We focus on deriving an asymptotic approximation to the probability  $\mathbb{P}(X_n = k)$ , which then leads to an effective estimate for the corresponding distribution functions. The method of proof we use here relies on the same analytic de-Poissonization procedure we used for the first two moments, and requires a uniform estimate with respect to  $k$ ; see [12], [16] for a similar analysis.

We begin by considering

$$\tilde{A}_k(z) = e^{-z} \sum_{n \geq 0} \mathbb{P}(X_n = k) \frac{z^n}{n!},$$

which satisfies the obvious bound  $\tilde{A}_k(x) \leq 1$  for real  $x \geq 0$ . Moreover, from (4), it follows that

$$\begin{aligned} \tilde{A}_0(z) &= e^{-z} \\ \tilde{A}_{k+1}(z) &= (1 - e^{-pz})\tilde{A}_k(pz) + e^{-pz}\tilde{A}_{k+1}(qz), \quad k \geq 0. \end{aligned} \tag{16}$$

Iterating (16) gives

$$\tilde{A}_{k+1}(z) = \sum_{j \geq 0} e^{-(1-q^j)z} (1 - e^{-pq^jz}) \tilde{A}_k(pq^jz) \quad \text{for } k \geq 0.$$

We then deduce the explicit expression, for  $k \geq 1$

$$\tilde{A}_k(z) = \sum_{j_1, \dots, j_k \geq 0} \exp \left\{ - \left( 1 - q \sum_{1 \leq r \leq k} p^{r-1} q^{j_1 + \dots + j_r} \right) z \right\} \prod_{1 \leq r \leq k} (1 - \exp\{-p^r q^{j_1 + \dots + j_r} z\}). \tag{17}$$

Now define the normalizing function

$$\Omega(z) := \prod_{j \geq 0} (1 - e^{-p^{-j}z}).$$

For convenience, let

$$\eta(n) := \{\log_{1/p} n\}$$

denote the fractional part of  $\log_{1/p} n$ .

**Theorem 3.** *The distribution of  $X_n$  satisfies*

$$\mathbb{P}(X_n = \lfloor \log_{1/p} n \rfloor + k) = \sum_{j \geq 0} \hat{R}_j(p^{-\eta(n)+k-j}) + O\left(\frac{1}{n}\right), \tag{18}$$

uniformly in  $k \in \mathbb{Z}$ , where  $\hat{R}_k(z) := \Omega(z)e^{-pz}\tilde{A}_k(qz)$  and  $\tilde{A}_k(z)$  is given in (17).

*Proof.* Write  $\hat{A}_k(z) := \Omega(z)\tilde{A}_k(z)$ . Then by (16), we have

$$\hat{A}_{k+1}(z) = \hat{A}_k(pz) + \hat{R}_{k+1}(z) \quad \text{for } k \geq 0,$$

which, after iteration, leads to

$$\hat{A}_k(z) = \sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}z) \quad \text{for } k \geq 0,$$

or, equivalently,

$$\tilde{A}_k(z) = \frac{1}{\Omega(z)} \sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}z) \quad \text{for } k \geq 0.$$

Since, by definition

$$\mathbb{P}(X_n = k) = n! [z^n] e^z \tilde{A}_k(z),$$

we need the following uniform estimates (which are needed to justify the de-Poissonization; see Appendix A).

**Lemma 3.** *The functions  $\tilde{A}_k(z)$  are uniformly JS-admissible, namely, for  $|\arg(z)| \leq \varepsilon$ ,  $0 < \varepsilon < \pi/2$ ,*

$$\tilde{A}_k(z) = O(|z|^{\varepsilon'}), \tag{19}$$

uniformly in  $z$  and  $k \geq 0$ , and, for  $\varepsilon \leq |\arg(z)| \leq \pi$ ,

$$e^z \tilde{A}_k(z) = O(e^{(1-\varepsilon')|z|}), \tag{20}$$

uniformly in  $z$  and  $k \geq 0$ . Here  $0 < \varepsilon' < 1$  and the involved constants in both cases are absolute.

*Proof.* Consider first  $|\arg(z)| \leq \varepsilon$ . Choose  $K > 0$  large enough such that  $1 + 2e^{-p\operatorname{Re}(z)} \leq 1 + \varepsilon'$  for all  $z$  with  $|z| > K$ . Moreover, choose  $C > 0$  such that for  $1 \leq |z| \leq K$

$$|\tilde{A}_k(z)| \leq e^{|z| - \operatorname{Re}(z)} \leq C \quad \text{for } k \geq 0.$$

We use a simple induction to show that

$$|\tilde{A}_k(z)| \leq C|z|^{\log_{1/q}(1+\varepsilon')} \quad \text{for } k \geq 0. \tag{21}$$

A similar inductive proof is used in [11] where it is referred to as induction over increasing domains. The claim (21) holds for  $k = 0$ . Next we assume (21) has been proved for  $k$  and we prove it for  $k + 1$ . The case  $1 \leq |z| \leq K$  follows from the definition of  $C$ . If  $K < |z| \leq K/q$ , we can use (16) and the induction hypothesis, and obtain

$$\begin{aligned} |\tilde{A}_{k+1}(z)| &\leq (1 + e^{-p \operatorname{Re}(z)})|\tilde{A}_k(pz)| + e^{-p \operatorname{Re}(z)}|\tilde{A}_{k+1}(qz)| \\ &\leq C(1 + \varepsilon')|qz|^{\log_{1/q}(1+\varepsilon')} = C|z|^{\log_{1/q}(1+\varepsilon')}. \end{aligned}$$

Continuing successively the same argument with  $K/q^j < |z| \leq K/q^{j+1}$  for  $j \geq 1$ , the upper bound (21) follows for all  $z$ . This concludes the proof of (19).

To prove (20), let  $A_k(z) := e^z \tilde{A}_k(z)$ . Then (16) becomes

$$A_{k+1}(z) = (e^{qz} - e^{(q-p)z})A_k(pz) + A_{k+1}(qz) \quad (k \geq 0).$$

Note that we have the (trivial) bound  $|A_k(z)| \leq e^{|z|}$ . Substituting this into the functional equation above yields

$$|A_{k+1}(z)| \leq (e^{q \cos \varepsilon |z|} + e^{(q-p) \cos(\varepsilon)|z|})e^{p|z|} + e^{q|z|},$$

from which (20) follows.

Now, by a standard de-Poissonization argument (see Appendix A for details and references), we obtain that

$$\mathbb{P}(X_n = k) = \frac{1}{\Omega(n)} \sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}n) + O\left(\frac{1}{n^{1-\varepsilon}}\right), \tag{22}$$

uniformly in  $k$ , where  $\varepsilon > 0$  is an arbitrarily small constant.

Note that we have the identity

$$\frac{1}{\Omega(n)} \sum_{k \geq 0} \sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}n) = 1. \tag{23}$$

This is seen as follows.

$$\begin{aligned} \sum_{k \geq 0} \sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}n) &= \sum_{j \geq 0} \sum_{k \geq 0} \hat{R}_j(p^k n) \\ &= \sum_{k \geq 0} \Omega(p^k n) e^{-p^{k+1}n} \sum_{j \geq 0} \tilde{A}_j(q p^k n) \\ &= \sum_{k \geq 0} \Omega(p^k n) (1 - (1 - e^{-p^{k+1}n})) \\ &= \sum_{k \geq 0} (\Omega(p^k n) - \Omega(p^{k+1}n)) \\ &= \Omega(n), \end{aligned}$$

which completes the proof of (23).

Now

$$\Omega(n) = \prod_{j \geq 0} (1 - e^{-p^{-j}n}) = 1 + O(e^{-n}).$$

This and (22) implies that

$$\sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}n) = O(1),$$

uniformly in  $k$ . Thus

$$\mathbb{P}(X_n = k) = \sum_{0 \leq j \leq k} \hat{R}_j(p^{k-j}n) + O\left(\frac{1}{n^{1-\varepsilon}}\right),$$

uniformly in  $k$ .

Finally, observe that

$$\sum_{j \geq k+1} \hat{R}_j(p^{k-j}n) \leq \sum_{j \geq k+1} e^{-p^{k+1-j}n} = \sum_{j \geq 0} e^{-p^{-j}n} = O(e^{-n}).$$

Thus

$$\mathbb{P}(X_n = k) = \sum_{j \geq 0} \hat{R}_j(p^{k-j}n) + O\left(\frac{1}{n^{1-\varepsilon}}\right),$$

uniformly in  $k$ .

Since the mean is asymptotic to  $\log_{1/p} n$ , we replace  $k$  by  $\lfloor \log_{1/p} n \rfloor + k$ . Then

$$\mathbb{P}(X_n = \lfloor \log_{1/p} n \rfloor + k) = \sum_{j \geq 0} \hat{R}_j(p^{-\eta(n)+k-j}) + O\left(\frac{1}{n^{1-\varepsilon}}\right),$$

uniformly in  $k$ , where  $\eta(n) = \{\log_{1/p} n\}$ . Because of the periodicity, the limiting distribution of  $X_n - \lfloor \log_{1/p} n \rfloor$ , in general, does not exist. However, if we instead consider a subsequence  $n_j$  of positive integers such that  $\eta(n_j) \rightarrow \theta \in (0, 1)$ , as  $j \rightarrow \infty$ , then the limit law does exist. The series on the right hand side sums (over all  $k$ ) asymptotically to 1 by (23).

Finally, the finer error term  $O(n^{-1})$  in (18) is obtained by refining the same procedure by including, say, one more term in the asymptotic expansion. This completes the proof.

A similar analysis can be carried out for the distribution function of  $X_n$  (we only have to divide (4) by  $1 - e^y$ ). This then yields the following estimate for the distribution.

**Corollary 2.** *The distribution function of  $X_n$  satisfies*

$$\mathbb{P}(X_n - \lfloor \log_{1/p} n \rfloor \leq k) = \sum_{j \geq 0} \hat{S}_j(p^{-\eta(n)+k-j}) + O\left(\frac{1}{n}\right),$$

uniformly in  $k \in \mathbb{Z}$ , where  $\hat{S}_k(z) = \sum_{j \leq k} \hat{R}_j(z)$ .

When  $p = \frac{1}{2}$ , we have the representation

$$\sum_{k \geq 0} \tilde{A}_k(z)u^k = \prod_{j \geq 1} (1 + (u - 1)(1 - e^{-z/2^j})),$$

which gives

$$\tilde{A}_k(z) = e^{-z} \sum_{1 \leq j_1 < \dots < j_k} \prod_{1 \leq r \leq k} (e^{z/2^{j_r}} - 1);$$

compare with (16). This expression was already derived in [21] where different expressions of the asymptotic distributions were given.

### Appendix A. Analytic de-Poissonization and JS-admissible functions

We develop the required tools for justifying the growth order of the functions involved in this paper, as well as systematic means of justifying the de-Poissonization procedure, based on the notion of JS-admissible functions (combining ideas from Jacquet and Szpankowski [11] and the classical paper by Hayman [6]). The following material, which is different from that given in [11], has been modified from [5], which provides more details.

**Definition 1.** An entire function  $\tilde{f}$  is said to be JS-admissible, denoted by  $\tilde{f} \in \mathcal{J}\mathcal{S}$ , if the following two conditions hold for  $|z| \geq 1$ .

(I) There exist  $\alpha, \beta \in \mathbb{R}$  such that uniformly for  $|\arg(z)| \leq \varepsilon$ ,

$$\tilde{f}(z) = O(|z|^\alpha (\log_+ |z|)^\beta), \quad \text{where } \log_+ x := \log(1 + x).$$

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

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

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

For convenience, we also write  $\tilde{f} \in \mathcal{J}\mathcal{S}_{\alpha,\beta}$  to indicate the growth order of  $\tilde{f}$  inside the sector  $|\arg(z)| \leq \varepsilon$ .

Note that if  $\tilde{f}$  satisfies condition (I), then, by Cauchy’s integral representation for derivatives (or by Ritt’s theorem; see [18, Ch. 1, Section 4.3]), we have,

$$\tilde{f}^{(k)}(z) = O(|z|^{\alpha-k} (\log_+ |z|)^\beta).$$

Moreover, by Cauchy’s integral representation, we also have

$$a_n = \frac{n!}{2\pi i} \oint_{|z|=n} z^{-n-1} e^z \tilde{f}(z) dz \approx \tilde{f}(n) \frac{n!}{2\pi i} \oint_{|z|=n} z^{-n-1} e^z dz = \tilde{f}(n),$$

since the saddlepoint  $z = n$  of the factor  $z^{-n} e^z$  is unaltered by the comparatively smoother function  $\tilde{f}(z)$ .

The latter analytic viewpoint provides the additional advantage of allowing an expansion to be obtained by using the Taylor expansion of  $\tilde{f}$  at  $z = n$ , yielding

$$a_n = \sum_{j \geq 0} \frac{\tilde{f}^{(j)}(n)}{j!} \tau_j(n), \tag{24}$$

where

$$\tau_j(n) := n! [z^n](z - n)^j e^z = \sum_{0 \leq \ell \leq j} \binom{j}{\ell} (-1)^{j-\ell} \frac{n! n^{j-\ell}}{(n - \ell)!}, \quad j = 0, 1, \dots,$$

and  $[z^n]\phi(z)$  denotes the coefficient of  $z^n$  in the Taylor expansion of  $\phi(z)$ . We call such an expansion *the Poisson–Charlier expansion* since the  $\tau_j$ 's are essentially the Charlier polynomials  $C_j(\lambda, n)$  defined by

$$C_j(\lambda, n) := \lambda^{-n} n! [z^n](z - 1)^j e^{\lambda z},$$

so that  $\tau_j(n) = n^j C_j(n, n)$ . For the other terms used in the literature and for more properties, see [8] and the references therein. In particular, the expansion (24) is absolutely convergent when  $\tilde{f}$  is entire.

**Proposition 2.** *Assume  $\tilde{f} \in \mathcal{G}_{\alpha, \beta}$ . Let  $f(z) := e^z \tilde{f}(z)$ . Then the Poisson–Charlier expansion (24) of  $f^{(n)}(0)$  is also an asymptotic expansion in the sense that*

$$\begin{aligned} a_n &:= f^{(n)}(0) \\ &= n! [z^n]f(z) \\ &= n! [z^n]e^z \tilde{f}(z) \\ &= \sum_{0 \leq j < 2k} \frac{\tilde{f}^{(j)}(n)}{j!} \tau_j(n) + O(n^{\alpha-k} (\log n)^\beta) \quad \text{for } k = 1, 2, \dots \end{aligned}$$

The polynomial growth of condition (I) is sufficient for all our uses; see [11] for more general versions.

The real advantage of introducing admissibility is that it opens the possibility of developing closure properties as we now briefly discuss.

**Lemma 4.** *Let  $m$  be a nonnegative integer and  $\alpha \in (0, 1)$ .*

- (i)  $z^m, e^{-\alpha z} \in \mathcal{G}$ .
- (ii) If  $\tilde{f} \in \mathcal{G}$ , then  $\tilde{f}(\alpha z), z^m \tilde{f} \in \mathcal{G}$ .
- (iii) If  $\tilde{f}, \tilde{g} \in \mathcal{G}$ , then  $\tilde{f} + \tilde{g} \in \mathcal{G}$ .
- (iv) If  $\tilde{f} \in \mathcal{G}$ , then the product  $\tilde{P} \tilde{f} \in \mathcal{G}$ , where  $\tilde{P}$  is a polynomial of  $z$ .
- (v) If  $\tilde{f}, \tilde{g} \in \mathcal{G}$ , then  $\tilde{h} \in \mathcal{G}$ , where  $\tilde{h}(z) := \tilde{f}(\alpha z) \tilde{g}((1 - \alpha)z)$ .
- (vi) If  $\tilde{f} \in \mathcal{G}$ , then  $\tilde{f}' \in \mathcal{G}$ , and thus  $\tilde{f}^{(m)} \in \mathcal{G}$ .

*Proof.* This is straightforward and is omitted.

Specific to our needs are the following transfer principles, first the real version and then the complex one.

**Lemma 5.** *Let  $\tilde{f}(z)$  and  $\tilde{g}(z)$  be entire functions satisfying*

$$\tilde{f}(z) = (1 - e^{-pz})\tilde{f}(pz) + e^{-pz}\tilde{f}(qz) + \tilde{g}(z), \tag{25}$$

with  $\tilde{f}(0) = \tilde{g}(0) = 0$ . If  $\tilde{g}(x) = O(x^\alpha (\log_+ x)^\beta)$  for real large  $x$ , where  $\alpha, \beta \in \mathbb{R}$ , then

$$\tilde{f}(x) = \begin{cases} O(x^\alpha (\log_+ x)^\beta) & \text{if } \alpha > 0, \\ \left\{ \begin{array}{ll} O((\log_+ x)^{\beta+1}) & \text{if } \beta > -1 \\ O(\log_+ \log_+ x) & \text{if } \beta = -1 \\ O(1) & \text{if } \beta < -1 \end{array} \right\} & \text{if } \alpha = 0, \\ O(1) & \text{if } \alpha < 0. \end{cases} \tag{26}$$



*Proof.* The idea behind this proof is that  $\tilde{f}(x)$  behaves asymptotically like the following recurrence

$$\phi(x) = \phi(px) + \tilde{g}(x),$$

with  $\phi(0) = 0$ . To that end, we need only to show that  $\tilde{f}(x)$  grows, at most, polynomially for large  $x$ . This is easily achieved by noting that  $f$  is bounded above by the function defined by the *trie-recurrence*

$$\lambda(x) = \lambda(px) + \lambda(qx) + \nu(x),$$

with  $\lambda(0) = \nu(0) = 0$ , where

$$\nu(x) := \begin{cases} Kx^{\bar{\alpha}} & \text{if } x > 1, \\ Kx & \text{if } 0 \leq x \leq 1, \end{cases}$$

$K > 0$  being a large constant and  $\bar{\alpha} := \max\{\lfloor \alpha \rfloor, 0\} + 1$ . Note that the exact solution of  $\lambda$  is given by

$$\lambda(x) = \sum_{j, \ell \geq 0} \binom{j + \ell}{j} \nu(p^j q^\ell x).$$

From this we then deduce that  $\lambda(x) = O(x^{\bar{\alpha}})$  for large  $x$ . Accordingly,  $\tilde{f}$  is polynomially bounded. The more precise estimates (26) then follow from standard Mellin arguments (by subtracting the first few  $\bar{\alpha} + 1$  terms of the Taylor expansion of  $\lambda(x)$  and then considering the Mellin transform of the now truncated  $\lambda$  term, which exists in the strip  $-\bar{\alpha} - 1 < \text{Re}(s) < -\bar{\alpha}$ ).

**Proposition 3.** *Let  $\tilde{f}(z)$  and  $\tilde{g}(z)$  be entire functions satisfying (25). Then*

$$\tilde{f} \in \mathcal{F}\mathcal{S} \quad \text{if and only if} \quad \tilde{g} \in \mathcal{F}\mathcal{S}.$$

*Proof.* The necessity part follows from Lemma 4. We prove the sufficiency, namely if  $\tilde{g} \in \mathcal{F}\mathcal{S}$ , then  $\tilde{f} \in \mathcal{F}\mathcal{S}$ .

Throughout the proof we write  $z = re^{i\theta}$ ,  $r \geq 0$  and  $-\pi \leq \theta \leq \pi$ . Consider first the region where  $\varepsilon \leq |\theta| \leq \pi$ . By assumption,  $|e^z \tilde{g}(z)| \leq Ke^{(1-\varepsilon_1)r}$ . Define

$$M(r) := \max_{\varepsilon \leq |\theta| \leq \pi} |f(z)| \quad (r \geq 0).$$

Then by the functional equation

$$f(z) = (e^{qz} - e^{(q-p)z})f(pz) + f(qz) + e^z \tilde{g}(z),$$

we have

$$M(r) \leq |e^{qz} - e^{(q-p)z}|M(pr) + M(qr) + Ke^{(1-\varepsilon_1)r}.$$

By using Pittel’s inequality (see [19, Appendix])

$$|e^z - 1| \leq (e^r - 1)e^{-r(1-\cos\theta)/2}, \quad r \geq 0, \quad |\theta| \leq \pi,$$

we have

$$\begin{aligned} |e^{qz} - e^{(q-p)z}| &= |e^{(q-p)z}| |e^{pz} - 1| \\ &\leq e^{(q-p)r \cos\theta} (e^{pr} - 1) e^{-pr(1-\cos\theta)/2} \\ &= (e^{qr} - e^{(q-p)r}) e^{-(q-p/2)r(1-\cos\theta)} \\ &\leq e^{-\varepsilon_2 r} (e^{qr} - e^{(q-p)r}), \end{aligned}$$

for  $\varepsilon \leq |\theta| \leq \pi$ . Let  $\varepsilon' := \min\{\varepsilon_1, \varepsilon_2\}$ . It follows that

$$M(r) \leq e^{-\varepsilon' r} (e^{qr} - e^{(q-p)r})M(pr) + M(qr) + Ke^{(1-\varepsilon')r}.$$

Let  $\tilde{M}(r) := M(r)e^{-(1-\varepsilon')r}$ . Then

$$\tilde{M}(r) \leq e^{-\varepsilon' pr} (1 - e^{-pr})\tilde{M}(pr) + e^{-(1-\varepsilon')pr}\tilde{M}(qr) + K.$$

By the same bounding argument that was used in Lemma 5, we see that  $\tilde{M}(r) = O(1)$ , and thus  $M(r) = O(e^{(1-\varepsilon')r})$ . [Technically, we define a function, say  $\phi(r)$ , satisfying the functional equation

$$\phi(r) = e^{-\varepsilon' pr} (1 - e^{-pr})\phi(pr) + e^{-(1-\varepsilon')r}\phi(qr) + K,$$

prove  $\phi(r) = O(1)$  and then  $M(r) \leq \phi(r)$ .]

We now consider the sector  $|\theta| \leq \varepsilon$ . Since  $\tilde{g}(z) = O(|z|^\alpha (\log_+ |z|)^\beta)$  in this sector, we can then show that  $\tilde{f}$  grows at most polynomially and is thus JS-admissible, the details being omitted here. This completes the proof.

### Acknowledgements

We thank the referee for helpful comments. M. Fuchs was partially supported by the NSC under grant number NSC-102-2115-M-009-002. A significant part of the research of H-K. Hwang was done while visiting the ISM (Institute of Statistical Mathematics), Tokyo; he thanks the ISM for its hospitality and support. Y. Itoh thanks the Institute of Statistical Science, Academia Sinica, Taipei, for its hospitality and support.

### References

- [1] ARCHIBALD, M., KNOPFMACHER, A. AND PRODINGER, H. (2006). The number of distinct values in a geometrically distributed sample. *Europ. J. Combinatorics* **27**, 1059–1081.
- [2] EVANS, J. W. (1993). Random and cooperative sequential adsorption. *Rev. Modern Phys.* **65**, 1281–1330.
- [3] FLAJOLET, P. AND SEDGEWICK, R. (2009). *Analytic Combinatorics*. Cambridge University Press.
- [4] FLAJOLET, P., GOURDON, X. AND DUMAS, P. (1995). Mellin transforms and asymptotics: harmonic sums. *Theoret. Comput. Sci.* **144**, 3–58.
- [5] FUCHS, M., HWANG, H.-K. AND ZACHAROVAS, V. (2014). An analytic approach to the asymptotic variance of trie statistics and related structures. *Theoret. Comput. Sci.* **527**, 1–36.
- [6] HAYMAN, W. K. (1956). A generalisation of Stirling's formula. *J. Reine Angew. Math.* **196**, 67–95.
- [7] HWANG, H.-K. AND JANSON, S. (2008). Local limit theorems for finite and infinite urn models. *Ann. Prob.* **36**, 992–1022.
- [8] HWANG, H.-K., FUCHS, M. AND ZACHAROVAS, V. (2010). Asymptotic variance of random symmetric digital search trees. *Discrete Math. Theoret. Comput. Sci.* **12**, 103–165.
- [9] ITOH, Y. AND SOLOMON, H. (1986). Random sequential coding by Hamming distance. *J. Appl. Prob.* **23**, 688–695.
- [10] ITOH, Y. AND UEDA, S. (1983). On packing density by a discrete random sequential packing of cubes in a space of  $m$  dimension. *Proc. Inst. Statist. Math.* **31**, 65–69 (in Japanese).
- [11] JACQUET, P. AND SZPANKOWSKI, W. (1998). Analytical de-Poissonization and its applications. *Theoret. Comput. Sci.* **201**, 1–62.
- [12] JANSON, S. AND SZPANKOWSKI, W. (1997). Analysis of an asymmetric leader election algorithm. *Electron. J. Combinatorics* **4**, Research Paper 17.
- [13] JANSON, S., LAVAUULT, C. AND LOUCHARD, G. (2008). Convergence of some leader election algorithms. *Discrete Math. Theoret. Comput. Sci.* **10**, 171–196.
- [14] KALPATHY, R. AND MAHMOUD, H. (2014). Perpetuities in fair leader election algorithms. *Adv. Appl. Prob.* **46**, 203–216.
- [15] KIRSCHENHOFER, P. AND PRODINGER, H. (1987). Asymptotische Untersuchungen über charakteristische Parameter von Suchbäumen. *Zahlentheoretische Analysis II, Lecture Notes in Math.*, **1262**, 93–107, Springer, Berlin.

- [16] KIRSCHENHOFER, P., PRODINGER, H. AND SZPANKOWSKI, W. (1996). Analysis of a splitting process arising in probabilistic counting and other related algorithms. *Random Structures Algorithms* **9**, 379–401.
- [17] MAHMOUD, H. M. (1992). *Evolution of Random Search Trees*. John Wiley, New York.
- [18] OLVER, F. W. J. (1974). *Asymptotics and Special Functions*. Academic Press, New York.
- [19] PITTEL, B. (1986). Paths in a random digital tree: limiting distributions. *Adv. Appl. Prob.* **18**, 139–155. (Correction: **24** (1992), 759.)
- [20] PRODINGER, H. (2004). Periodic oscillations in the analysis of algorithms and their cancellations. *J. Iranian Statist. Soc.* **3**, 251–270.
- [21] RAIS, B., JACQUET, P. AND SZPANKOWSKI, W. (1993). Limiting distribution for the depth in PATRICIA tries. *SIAM J. Discrete Math.* **6**, 197–213.
- [22] DUTOUR SIKIRIĆ, M. AND ITOH, Y. (2011). *Random Sequential Packing of Cubes*. World Scientific, Hackensack, NJ.