Occupancy Schemes Associated to Yule Processes

Philippe Robert, Florian Simatos

OCCUPANCY SCHEMES ASSOCIATED TO YULE PROCESSES

PHILIPPE ROBERT AND FLORIAN SIMATOS

Abstract. An occupancy problem with an infinite number of bins and a random probability vector for the locations of the balls is considered. The respective sizes of bins are related to the split times of a Yule process. The asymptotic behavior of the landscape of first empty bins, i.e., the set of corresponding indices represented by point processes, is analyzed and convergences in distribution to mixed Poisson processes are established. Additionally, the influence of the random environment, the random probability vector, is analyzed. It is represented by two main components: an i.i.d. sequence and a fixed random variable. Each of these components has a specific impact on the qualitative behavior of the stochastic model. It is shown in particular that for some values of the parameters, some rare events, which are identified, play an important role on average values of the number of empty bins in some regions.

Contents

1. Introduction 1
2. A Bins and Balls Problem in Random Environment 3
3. Convergence of Point Processes 4
4. Asymptotic Behavior of the Indices of the First Empty Bins 11
5. Rare Events 13
6. Generalizations 18
References 19

1. Introduction

Occupancy schemes in terms of bins and balls offer a very flexible and elegant way to formulate various problems in computer science, biology and applied mathematics for example. One of the earliest models investigated in the literature consists in throwing m balls at random into n identical bins. Asymptotic behavior of occupancy variables have been analyzed when n grows to infinity, with different scalings in n for the variable m. The books by Johnson and Kotz [11] and Kolchin *et al.* [14] are classical references on this topic. See also Chapter 6 of Barbour *et al.* [3] for a recent presentation of these problems.

An extension of these models is when there is an infinite number of bins and a probability vector (p_n) on $\mathbb{N}$ describing the way balls are sent: for $n \geq 0$, p_n is the probability that a ball is sent into the nth bin. In one of the first studies in this setting, Karlin [12] analyzed the asymptotic behavior of the number of occupied bins. More recently Hwang and Janson [10] proves in a quite general framework central limit results for these quantities. In this setting, some additional variables are also of interest like the sets of indices of occupied or empty bins, adding a

Date: October 17, 2008.
Key words and phrases. Occupancy Schemes; Poisson Processes; Convergence of Point Processes; Random Environment; Rare Events.

geometric component to these problems. For specific probability vectors (p_n) Csáki and Földes [4] and Flajolet and Martin [6] investigated the index of the first empty bin. See the recent survey Gnedin *et al.* [8] for more references on the occupancy problem with infinitely many bins.

A further extension of these stochastic models consists in considering *random* probability vectors. Gnedin [7] (and subsequent papers) analyzed the case where (p_n) decays geometrically fast according to some random variables, i.e., for $n \geq 1$, $p_n = \prod_{i=1}^{n-1} Y_i(1 - Y_n)$ where (Y_i) are i.i.d. random variables on $(0, 1)$. Various asymptotic results on the number of occupied bins in this case have been obtained. The random vector can be seen as a "random environment" for the bins and balls problem, it complicates significantly the asymptotic results in some cases. In particular, the indices of the urns in which the balls fall are no longer independent random variables as in the deterministic case.

The general goal of this paper is to investigate in detail the impact of this randomness for a bins and balls problem associated to a Yule process, see Athreya and Ney [2] for the definition of a Yule process. This (quite natural) stochastic model has its origin in network modeling, see Simatos *et al.* [23] for a detailed presentation. It can be described as follows: the non-decreasing sequence (t_n) of split times of the Yule process defines the bins, the nth bin, $n \geq 1$, being the interval $(t_{n-1}, t_n]$. The locations of balls are represented by independent exponential random variables with parameter ρ. The main problem investigated here concerns the asymptotic description of the set of indices of first empty bins when the number of balls goes to infinity. Mathematically, it is formulated as a convergence in distribution of rescaled point processes having Dirac masses at the indices of empty bins.

For $n \geq 1$, if P_n is the probability that a ball falls into the nth bin, it is easily seen that, for a large n, P_n has a power law decay, it can be expressed as $V E_n / n^{\rho+1}$ where (E_n) are i.i.d. exponential random variables with parameter 1 and V some independent random variable related to the limit of a martingale. The randomness of the probability vector (P_n) has two components: one which is a part of an i.i.d. sequence, changing from one bin to another, and the other being "fixed once for all" inducing a dependency structure. As it will be seen, the two components have separately a significant impact on the qualitative behavior of this model.

Convergence in Distribution and Rare Events. Because the variables (E_n) can be arbitrarily small with positive probability, empty bins are likely to be created earlier (i.e., with smaller indices) than for a deterministic probability vector with the same power law decay. It is shown in fact that, for the convergence in distribution, the first empty bins occur around indices of the order of $n^{1/(\rho+2)}$ instead of $(n/\log n)^{1/(1+\rho)}$ in the deterministic case.

The variable V has a more subtle impact, when $\rho > 1$ it is shown that, due to some heavy tail property of V^{-1}, rare events affect the asymptotic behavior of *averages* of some of the characteristics. For $\alpha \in [1/(2\rho+1), 1/(\rho+2))$, despite that the number of empty bins with indices of order n^α converges *in distribution* to 0, the corresponding average converges to $+\infty$. When $\rho < 1$, the average is converging to 0. A phase transition phenomenon at $\rho = 1$ has been identified through simulations in a related context, communication networks, in Saddi and Guillemin [22]. It is not apparent as long as convergence in distribution is concerned but it shows up when average quantities are considered. This phenomenon is due to rare events related to the total size of the $\lfloor \rho \rfloor$ first bins: On these events, the indices of the first empty bins are of the order $n^{1/(2\rho+1)} \ll n^{1/(\rho+2)}$ and a lot them are created at this occasion. See Proposition 6 and Corollary 2 for a precise statement of this result. Concerning the generality of the results obtained, it is believed that some of them hold in a more general setting, for the underlying branching process for example, see Section 6.

Point Processes. Technically, one mainly uses point processes on $\mathbb{R}_+$ to describe the asymptotic behavior of the indices of the first empty bins and not only the index of the first one (or the subsequent ones) as it is usually the case in the literature. It turns out that it is quite appropriate in our setting to get a full description of the set of the first empty bins and, moreover, it reduces the technicalities of some of the proofs. One of the arguments for the proofs of the convergence results is a simple convergence result of two-dimensional point processes to Poisson point process with some intensity measure. A one-dimensional equivalent of this point of view is implicit in most of the papers of the literature, in Hwang and Janson [10] in particular. See Robert and Simatos [20] for a presentation of an extension of this approach in a more general framework.

The paper is organized as follows. Section 2 introduces the stochastic model investigated. The main results concerning convergence of related point processes in $\mathbb{R}_+^2$ are presented in Section 3. Convergence results for the indices of empty bins are proved in Section 4. Section 5 investigates in detail the case $\rho \geq 1$. Section 6 presents some possible extensions.

2. A Bins and Balls Problem in Random Environment

The stochastic model is described in detail and some notations are introduced.

The Bins. Let (E_i) be a sequence of i.i.d. exponential random variables with parameter 1. Define the non-decreasing sequence (t_n) by, for $n \geq 1$,

$$t_n = \sum_{i=1}^{n} \frac{1}{i} E_i.$$

It is easy to check that for $x \geq 0$,

$$\mathbb{P}(t_n \leq x) = \mathbb{P}(\max(E_1, E_2, \dots, E_n) \leq x) = (1 - e^{-x})^n. \tag{1}$$

The nth bin will be identified by the interval $(t_{n-1}, t_n]$.

If $H_n = 1 + 1/2 + \cdots + 1/n$ is the nth harmonic number, since $(t_n - H_n)$ is a square integrable martingale whose increasing process is given by

$$\mathbb{E}\left((t_n - H_n)^2\right) = \sum_{i=1}^{n} \frac{1}{i^2},$$

then $(M_n) \stackrel{\text{def.}}{=} (t_n - \log n)$ is almost surely converging to some finite random variable M_∞. See Neveu [17] or Williams [24]. By using Equation (1), it is not difficult to get that the distribution of M_∞ is given by

$$\mathbb{P}(M_\infty \leq x) = \exp\left(-e^{-x}\right), \quad x \in \mathbb{R}. \tag{2}$$

An alternative description of the sequence (t_n) is provided by the split times of a Yule (branching) process starting with one individual. See Athreya and Ney [2].

The Balls. The locations of the balls are given by an independent sequence (B_j) of i.i.d. exponential random variables with parameter ρ for some $\rho > 0$.

Conditionally on the point process (t_n) associated with the location of bins, the probability that a given ball falls into the nth bin $(t_{n-1}, t_n]$ is given by

$$P_n = \mathbb{P}\left[B_1 \in (t_{n-1}, t_n] \,\middle|\, (t_n)\right] = e^{-\rho t_{n-1}} - e^{-\rho t_n} = e^{-\rho t_{n-1}}\left(1 - e^{-\rho E_n/n}\right).$$

This quantity can be rewritten as

$$P_n = \frac{1}{n^{\rho+1}} W_n^\rho Z_n, \text{ with } Z_n = n\left(1 - e^{-\rho E_n/n}\right) \text{ and } W_n = e^{-M_{n-1}}. \tag{3}$$

The variables W_n and Z_n are independent random variables with different behavior.

(1) The variables (Z_n) are independent and converge in distribution to an exponentially distributed random variable with parameter $1/\rho$.
(2) The random variables (W_n) converge almost surely to the finite random variable $W_\infty = \exp(-M_\infty)$ which is exponentially distributed with parameter 1.

This suggests an asymptotic representation of the sequence (P_n) as

$$P_n \sim \frac{1}{n^{\rho+1}} W_\infty^\rho \widetilde{E}_n, \tag{4}$$

where $(\widetilde{E}_n)$ is an i.i.d. sequence of exponential random variables with mean ρ independent of W_∞. The sequence (P_n) has a power law decay with a random coefficient consisting of the product of two terms: a fixed random variable W_∞^ρ and the other being an element of an i.i.d. sequence. As it will be seen, these two terms have a significant impact on the bins and balls problem studied in this paper.

3. Convergence of Point Processes

One of the main result, Theorem 2 in the next section, which establishes convergence results for the indices of the first empty bins is closely related to the asymptotic behavior of the point process $\{(i/n^{1/(2+\rho)}, nP_i), i \geq 1\}$ on $\mathbb{R}_+^2$. For this reason, some results on convergence of point processes in $\mathbb{R}_+^2$ are first proved. The point process associated to the (nP_i) appears quite naturally, especially in view of the Poisson transform used in the proof of Theorem 2. This is also a central variable in Hwang and Janson [10] in some cases.

An important tool to study point processes in $\mathbb{R}_+^d$ for some $d \geq 1$ is the Laplace transform: If $\mathcal{N}$=$\{t_n, n \geq 1)$ is a point process and f a function in $C_c^+(\mathbb{R}_+^d)$, the set of non-negative continuous functions with a compact support, it is defined as $\mathbb{E}(\exp(-\mathcal{N}(f)))$, where

$$\mathcal{N}(f) \stackrel{\text{def.}}{=} -\sum_{n\geq 1} f(t_n).$$

This functional uniquely determines the distribution of $\mathcal{N}$ and it is a key tool to establish convergence results. See Neveu [18] and Dawson [5] for a comprehensive presentation of these questions. In the following, the quantity $\mathcal{N}(A)$ denotes the number of t_n's in the subset A of $\mathbb{R}_+^d$.

The main results of this section establish convergence in distribution to mixed Poisson point processes, i.e., distributed as a Poisson point process with a parameter which is a random variable. A natural tool in this domain is the Chen-Stein approach which gives the convergence in distribution and, generally, quite good bounds on the convergence rate. See Chapter 10 of Barbour *et al.* [3] for example. This has been used in Simatos *et al.* [23], when the probability vector is deterministic. For some of the results of this section, this approach can probably also be used. Unfortunately, due to the almost surely converging sequence (W_n) creating a dependency structure, it does not seem that the main convergence result, Theorem 1, can be proved in a simple way by using Chen-Stein's method. The main problem being of conditioning on the variable W_∞ and keeping at the same time upper bounds on the total variation distance converging to 0.

Condition C. A sequence of independent random variables (X_i) satisfies Condition C if there exist some $\alpha>0$ and $\eta>0$ such that, for all $i\geq 1$,

$$|\mathbb{P}(X_i\leq x)-\alpha x|\leq Cx^2, \text{ when } 0\leq x\leq \eta. \tag{5}$$

The following proposition is a preliminary result that will be used to prove the main convergence results for the indices of the first empty bins.

Proposition 1 (Convergence to a Poisson process)**.** *For $\delta>0$ and $n\geq 1$, let $\mathcal{P}_n$ be the point process on $\mathbb{R}_+^2$ defined by*

$$\mathcal{P}_n \stackrel{def.}{=} \left\{\left(\frac{i}{n^{1/(\delta+1)}},\frac{n}{i^\delta}X_i\right), i\geq 1\right\},$$

where (X_i) a sequence of non-negative independent random variables satisfying Condition C. Then the sequence of point processes $(\mathcal{P}_n)$ converges in distribution to a Poisson point process $\mathcal{P}$ in $\mathbb{R}_+^2$ with intensity measure $x^\delta\,dx\,dy$ on $\mathbb{R}_+^2$. In particular, its Laplace transform is given by

$$\mathbb{E}(\exp[-\mathcal{P}(f)])=\exp\left(-\alpha\int_{\mathbb{R}_+^2}\left(1-e^{-f(x,y)}\right)x^\delta\,dx\,dy\right),\quad f\in C_c^+(\mathbb{R}_+^2). \tag{6}$$

See Robert [19] for the definition and the main properties of Poisson processes in general state spaces.

Proof. There exists some $\eta_0>0$ such that $\mathbb{P}(X_i\leq x)\leq 2\alpha x$ for $0\leq x\leq\eta_0$ and all $i\geq 1$. Let $f\in C_c^+(\mathbb{R}_+^2)$ be such that f is differentiable with respect to the second variable. There is some $K>0$ so that the support of f is included in $[0,K]\times[0,K]$, define $g(x,y)=1-\exp(-f(x,y))$, then by independence of the variables X_i, $i\geq 1$,

$$\log\mathbb{E}\left(e^{-\mathcal{P}_n(f)}\right)=\sum_{i=1}^{+\infty}\log\left(1-\mathbb{E}\left[g\left(\frac{i}{n^{1/(\delta+1)}},\frac{n}{i^\delta}X_i\right)\right]\right).$$

Since

$$\mathbb{E}\left[g\left(\frac{i}{n^{1/(\delta+1)}},\frac{n}{i^\delta}X_i\right)\right]\leq\mathbb{P}\left(X_i\leq K\frac{i^\delta}{n}\right)\mathbb{1}_{\{i\leq Kn^{1/(\delta+1)}\}},$$

the elementary inequality $|\log(1-y)+y|\leq 3y^2/2$ valid for $0\leq y\leq 1/2$ shows that there exists some $n_0\geq 1$ such that

$$\left|\log\mathbb{E}\left(e^{-\mathcal{P}_n(f)}\right)+\sum_{i=1}^{+\infty}\mathbb{E}\left[g\left(\frac{i}{n^{1/(\delta+1)}},\frac{n}{i^\delta}X_i\right)\right]\right| \leq\frac{6(\alpha K)^2}{n^2}\sum_{i=1}^{\lfloor Kn^{1/(\delta+1)}\rfloor}i^{2\delta}\leq 6\alpha^2K^{2\delta+3}\frac{1}{n^{1/(\delta+1)}}$$

holds for $n \geq n_0$. It is therefore enough to study the asymptotics of the series of the left hand side of the above inequality. For $x \geq 0$, by using Fubini's Theorem, one gets

$$\mathbb{E}\left(g\left(x, \frac{n}{i^\delta} X_i\right)\right) = -\int_0^{+\infty} \frac{\partial g}{\partial y}(x,y)\mathbb{P}\left(X_i \leq y i^\delta / n\right) \, dy.$$

By using again Condition C, one obtains that the log of the Laplace transform of $\mathcal{P}_n$ has the same asymptotic behavior as

$$-\alpha \frac{1}{n^{1/(\delta+1)}} \sum_{i=1}^{+\infty} \int_0^{+\infty} \frac{\partial g}{\partial y}\left(\frac{i}{n^{1/(\delta+1)}}, y\right) y \left(\frac{i}{n^{1/(\delta+1)}}\right)^\delta \, dy$$

which is a Riemann sum converging to

$$-\alpha \int_{\mathbb{R}_+^2} \frac{\partial g}{\partial y}(x,y) y x^\delta \, dxdy = \alpha \int_{\mathbb{R}_+^2} \left(1 - e^{-f(x,y)}\right) x^\delta \, dxdy.$$

This shows in particular that for any compact set H of $\mathbb{R}_+^2$, then

$$\sup_{n \geq 1} \mathbb{E}(\mathcal{P}_n(H)) < +\infty,$$

the sequence $(\mathcal{P}_n)$ is therefore tight for the weak topology, see Dawson [5].

By the convergence result, if $\mathcal{P}$ is any limiting point of the sequence $(\mathcal{P}_n)$, for any function $f \in C_c^+(\mathbb{R}_+^2)$ such that $y \to f(x,y)$ is differentiable, then the Laplace transform of $\mathcal{P}$ at f is given by the right hand side of Equation (6). By density of these functions f in $C_c^+(\mathbb{R}_+^2)$ for the uniform topology, this implies that $\mathcal{P}$ is indeed a Poisson point process with intensity measure $x^\delta \, dxdy$ on $\mathbb{R}_+^2$. The proposition is proved. □

The above result can be (roughly) restated as follows: for the indices of the order of $n^{1/(\delta+1)}$, the points nX_i/i^δ, lying in some finite fixed interval converge to an homogeneous Poisson point process. The following proposition gives an asymptotic description of the indices of the points nX_i/i^δ but for indices of the order of n^κ with $\kappa > 1/(\delta+1)$. It shows that, on finite intervals, these points accumulate at rate $n^{(1+\delta)\kappa-1}$ according to the Lebesgue measure with some density.

Proposition 2 (Law of Large Numbers)**.** *If, for $\kappa > 1/(1+\delta)$ and for $n \in \mathbb{N}$, $\mathcal{P}_n^\kappa$ is the point process on $\mathbb{R}_+$ defined by*

$$\mathcal{P}_n^\kappa(f) = \frac{1}{n^{(1+\delta)\kappa-1}} \sum_{i=1}^{+\infty} f\left(\frac{i}{n^\kappa}, \frac{n}{i^\delta} X_i\right), \quad f \in C_c^+(\mathbb{R}_+^2),$$

where (X_i) is a sequence of non-negative independent random variables satisfying Condition C, then the sequence $(\mathcal{P}_n^\kappa)$ converges in distribution to the deterministic measure $\mathcal{P}_\infty^\kappa$ defined by

$$\mathcal{P}_\infty^\kappa(f) = \alpha \int_{\mathbb{R}_+^2} f(x,y) x^\delta \, dxdy, \quad f \in C_c^+(\mathbb{R}_+^2).$$

Proof. Let $f \in C_c^+(\mathbb{R}_+^2)$ be such that f is differentiable with respect to the second variable. As before, the convergence result is proved for such a function f, the generalization to an arbitrary function $f \in C_c^+(\mathbb{R}_+^2)$ follows the same lines as the previous proof (relative compactness argument and identification of the limit). Let $K>0$ such that the support of f is included in $[0,K]\times[0,K]$. One has

$$\mathbb{E}\left(\mathcal{P}_n^\kappa(f)\right) = -\frac{1}{n^{(1+\delta)\kappa-1}}\sum_{i=1}^{+\infty}\int_0^{+\infty}\frac{\partial f}{\partial y}\left(\frac{i}{n^\kappa},y\right)\mathbb{P}(X_i\le yi^\delta/n)\,dy,$$

as in the previous proof, by using Condition (5), one gets that

$$\mathbb{E}\left(\mathcal{P}_n^\kappa(f)\right) \sim -\alpha\frac{1}{n^\kappa}\sum_{i=1}^{+\infty}\int_0^{+\infty}\frac{\partial f}{\partial y}\left(\frac{i}{n^\kappa},y\right)y\left(\frac{i}{n^\kappa}\right)^\delta\,dy,$$

therefore,

$$\lim_{n\to+\infty}\mathbb{E}\left(\mathcal{P}_n^\kappa(f)\right) = -\alpha\int_0^{+\infty}\int_0^{+\infty}\frac{\partial f}{\partial y}(x,y)\,yx^\delta\,dxdy = \alpha\int_{\mathbb{R}_+^2}f(x,y)x^\delta\,dxdy.$$

By independence of the X_i's the second moment of the difference

$$\begin{aligned}&\mathcal{P}_n^\kappa(f)-\mathbb{E}\left(\mathcal{P}_n^\kappa(f)\right)\\ &\quad= -\frac{1}{n^{(1+\delta)\kappa-1}}\sum_{i=1}^{+\infty}\int_0^{+\infty}\frac{\partial f}{\partial y}\left(\frac{i}{n^\kappa},y\right)\left[\mathbb{1}_{\{X_i\le yi^\delta/n\}}-\mathbb{P}(X_i\le yi^\delta/n)\right]dy,\end{aligned}$$

can be expressed as

$$\begin{aligned}&n^{2((1+\delta)\kappa-1)}\times\mathbb{E}\left(\left[\mathcal{P}_n^\kappa(f)-\mathbb{E}\left(\mathcal{P}_n^\kappa(f)\right)\right]^2\right)\\ &=\sum_{i=1}^{+\infty}\mathbb{E}\left(\left[\int_0^{+\infty}\frac{\partial f}{\partial y}\left(\frac{i}{n^\kappa},y\right)\left[\mathbb{1}_{\{X_i\le yi^\delta/n\}}-\mathbb{P}(X_i\le yi^\delta/n)\right]dy\right]^2\right)\\ &\le K\sum_{i=1}^{+\infty}\int_0^{+\infty}\left[\frac{\partial f}{\partial y}\left(\frac{i}{n^\kappa},y\right)\right]^2\mathbb{E}\left(\left[\mathbb{1}_{\{X_i\le yi^\delta/n\}}-\mathbb{P}(X_i\le yi^\delta/n)\right]^2\right)dy\\ &\le K\sum_{i=1}^{+\infty}\int_0^{+\infty}\left[\frac{\partial f}{\partial y}\left(\frac{i}{n^\kappa},y\right)\right]^2\mathbb{P}(X_i\le yi^\delta/n)\,dy,\end{aligned}$$

by Cauchy-Shwartz's Inequality. The last term is, with the same arguments as for the asymptotics of $\mathbb{E}\left(\mathcal{P}_n^\kappa(f)\right)$, equivalent to

$$Kn^{(1+\delta)\kappa-1}\times\int_{\mathbb{R}_+^2}\left[\frac{\partial f}{\partial y}(x,y)\right]^2yx^\delta\,dxdy.$$

In particular, the sequence $(\mathcal{P}_n^\kappa(f))$ converges in L_2 (and therefore in distribution) to $\mathcal{P}_\infty^\kappa(f)$. The proposition is proved.

The main convergence result can now be established. □

Theorem 1. *If, for $n\ge1$, $\mathcal{P}_n$ is the point process on $\mathbb{R}_+^2$ defined by*

$$\mathcal{P}_n=\left\{\left(\frac{i}{n^{1/(\rho+2)}},nP_i\right),i\ge1\right\},$$

then the sequence $(\mathcal{P}_n)$ converges in distribution and the relation

$$\lim_{n\to+\infty}\mathbb{E}\left(e^{-\mathcal{P}_n(f)}\right)=\mathbb{E}\left[\exp\left(-\frac{W_\infty^{-\rho}}{\rho}\int_{\mathbb{R}_+^2}\left(1-e^{-f(x,y)}\right)x^{\rho+1}\,dx\,dy\right)\right]\tag{7}$$

holds for any $f\in C_c^+(\mathbb{R}_+^2)$.

In other words the point process $\mathcal{P}_n$ converges in distribution to a mixed Poisson point process: conditionally on W_∞, it is a Poisson process with intensity measure $W_\infty^{-\rho} x^{\rho+1}\, dx\, dy/\rho$.

Proof. The proof proceeds in several steps. The main objective of these steps is to decouple the sequences (W_i) and (Z_i) defining the (P_i) and then to apply Proposition 1.

Step 1. One defines the sequences

$$P_i^1 = \frac{1}{i^{\rho+1}} \widetilde{W}_\infty^\rho \widetilde{Z}_i, \quad i \geq 1, \quad P_i^2 = \frac{1}{i^{\rho+1}} \widetilde{W}_{\beta_n}^\rho \widetilde{Z}_i, \quad i \geq 1,$$

where (β_n) is some sequence of integers converging to $+\infty$. The sequences of random variables $(\widetilde{W}_i, 1 \leq i \leq +\infty)$ and $(\widetilde{Z}_i)$ are assumed to be independent and to have, respectively, the same distribution as $(W_i, 1 \leq i \leq +\infty)$ and (Z_i) defined by Equation (3). Recall that the sequence $(\widetilde{W}_i)$ converges almost surely to $\widetilde{W}_\infty$. These sequences define point processes in the following way, for $j = 1$ and 2,

$$\mathcal{P}_n^j = \left\{ \left(\frac{i}{n^{1/(\rho+2)}}, nP_i^j \right), i \geq 1 \right\}.$$

If f is a non-negative continuous function with compact support on $\mathbb{R}_+^2$, because, conditionally on $\widetilde{W}_\infty$, the variables $(\widetilde{W}_\infty Z_i)$ satisfy Condition C with $\alpha = \widetilde{W}_\infty^{-\rho}/\rho$, Proposition 1, with $\delta = \rho + 1$, shows that

$$\lim_{n\to+\infty} \mathbb{E}\left(e^{-\mathcal{P}_n^1(f)} \,\middle|\, \widetilde{W}_\infty \right) = \exp\left(-\frac{\widetilde{W}_\infty^{-\rho}}{\rho} \int_{\mathbb{R}_+^2} \left(1 - e^{-f(x,y)}\right) x^{\rho+1}\, dx\, dy \right).$$

Because of the boundedness of these quantities, by Lebesgue's Theorem, the same result holds for the expected values. Therefore, the sequence $(\mathcal{P}_n^1)$ converges in distribution to the point process $\mathcal{P}$ on $\mathbb{R}_+^2$ whose Laplace transform is given by Equation (7).

Let $K \geq 2$ be such that the support of f is a subset of $[0, K]^2$ and $\varepsilon > 0$. Since the limiting point process $\mathcal{P}$ is almost surely a Radon measure, there exists some $m \in \mathbb{N}$ such that $\mathbb{P}(\mathcal{P}_n^1([0, 2K]^2) \geq m) \leq \varepsilon$ for all $n \geq 1$. By uniform continuity, there exists $0 < \eta < 1/2$ such that $|f(u) - f(v)| \leq \varepsilon/m$ for u, $v \in \mathbb{R}_+^2$ such that $\|u - v\| \leq \eta$. For $n \geq 1$, if

$$\mathcal{A} \stackrel{\text{def}}{=} \{|\widetilde{W}_{\beta_n}^\rho/\widetilde{W}_\infty^\rho - 1| \geq \eta/2K\} \cup \{\mathcal{P}_n^1([0, 2K]^2) \geq m\}$$

then

$$\begin{aligned} &\left|\mathbb{E}\left(\exp\left[-\mathcal{P}_n^2(f)\right]\right) - \mathbb{E}\left(\exp\left[-\mathcal{P}_n^1(f)\right]\right)\right| \leq \mathbb{P}(\mathcal{A}) \\ &+ \mathbb{E}\left(\left(\exp\left[\sum_{i\geq 1}\left| f\left(\frac{i}{n^{1/(\rho+2)}}, \frac{\widetilde{W}_{\beta_n}^\rho}{\widetilde{W}_\infty^\rho} nP_i^1\right) - f\left(\frac{i}{n^{1/(\rho+2)}}, nP_i^1\right)\right|\right] - 1\right) \mathbb{1}_{\mathcal{A}^c}\right) \\ &\qquad\qquad \leq \mathbb{P}(|W_{\beta_n}^\rho/W_\infty^\rho - 1| \geq \eta/2K) + 2\varepsilon, \end{aligned}$$

hence, by the almost sure convergence of (W_n) to W_∞, the right hand side of the last relation can be arbitrarily small as n goes to infinity. One concludes that the sequence $(\mathcal{P}_n^2)$ also converges in distribution to the point process $\mathcal{P}$.

Step 2. For $n\geq 1$, define

$$\beta_n=\left\lfloor n^{1/(\rho+2)}/\log n\right\rfloor,$$

then it will be shown that the point processes

$$\mathcal{Q}_n=\left\{\left(\frac{i}{n^{1/(\rho+2)}},nP_i\right),1\leq i\leq \beta_n\right\}$$

converge to the measure identically null. It is sufficient to prove that for any $f\in C_c^+(\mathbb{R}_+)$, the sequence $(\mathcal{Q}_n(f))$ converges in distribution to 0. For a fixed i, the sequence (nP_i) converges in distribution to infinity, since f is continuous with compact support and therefore bounded, one obtains that, in the definition of $\mathcal{Q}_n$, it can be assumed that the indices i are restricted to the set $\{\lceil\rho\rceil,\ldots,\beta_n\}$.

Let K be such that the support of f is included in $[0,K]^2$, if $u_n=\log\log n$, for $i\geq\lceil\rho\rceil$,

$$\mathbb{E}\left(f(i/n^{1/(\rho+2)},nP_i)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq u_n\}}\right)\leq\|f\|_\infty\mathbb{P}\left(t_{\lfloor\rho\rfloor}\leq u_n,nP_i\leq K\right),$$

since $P_i=e^{-\rho t_{\lfloor\rho\rfloor}}e^{-\rho(t_{i-1}-t_{\lfloor\rho\rfloor})}\left(1-e^{-\rho E_i/i}\right)$,

$$\mathbb{E}\left(f(i/n^{1/(\rho+2)},nP_i)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq u_n\}}\right)\\ \leq\|f\|_\infty\mathbb{P}\left[\left(\frac{1-e^{-\rho E_i/i}}{\rho/i}\right)\leq\frac{i}{\rho}Ke^{\rho u_n}e^{\rho(t_{i-1}-t_{\lfloor\rho\rfloor})}/n\right].$$

By using the elementary inequality, if E_1 is exponentially distributed with mean 1,

$$\mathbb{P}\left(\frac{1}{y}\left(1-e^{-yE_1}\right)\leq x\right)\leq e\left(1-e^{-x}\right),\quad y\leq 1,x\geq 0,\tag{8}$$

one gets that, for $i>\rho$,

$$\begin{aligned}\mathbb{E}&\left(f(i/n^{1/(\rho+2)},nP_i)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq u_n\}}\right)\\ &\leq e\|f\|_\infty\,\mathbb{E}\left(1-\exp\left[-\frac{i}{n\rho}Ke^{\rho u_n}e^{\rho(t_{i-1}-t_{\lfloor\rho\rfloor})}\right]\right)\\ &\leq eK\|f\|_\infty\frac{ie^{\rho u_n}}{n\rho}\mathbb{E}\left(e^{\rho(t_{i-1}-t_{\lfloor\rho\rfloor})}\right)\\ &=eK\|f\|_\infty\frac{ie^{\rho u_n}}{n\rho}e^{\rho\sum_{k=\lceil\rho\rceil}^{i-1}1/k}e^{\sum_{k=\lceil\rho\rceil}^{i-1}-\log(1-\rho/k)-\rho/k}.\end{aligned}$$

Thus, there exists some finite constant C such that, for $i>\rho$,

$$\mathbb{E}\left(f(i/n^{1/(\rho+2)},nP_i)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq u_n\}}\right)\leq C\frac{i^{\rho+1}e^{\rho u_n}}{n}=C\frac{i^{\rho+1}(\log n)^\rho}{n},$$

consequently,

$$\mathbb{E}\left(\mathcal{Q}_n(f)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq u_n\}}\right)\leq C\frac{\beta_n^{\rho+2}(\log n)^\rho}{n}\leq C\frac{1}{(\log n)^2}.$$

This relation and the inequality

$$\mathbb{E}\left(1-e^{-\mathcal{Q}_n(f)}\right)\leq\mathbb{P}(t_{\lfloor\rho\rfloor}>u_n)+\mathbb{E}\left(\mathcal{Q}_n(f)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq u_n\}}\right)$$

give the desired result.

Step 3. The proof of the theorem can be now completed. By Equation (3), for $i \geq 1$, $P_i = W_i^\rho Z_i / i^{\rho+1}$, by using Step 2 and the same techniques as in Step 1 together with the fact that, for $\eta > 0$, the probability of the event

$$\left\{\sup\left(\left|W_i^\rho / W_{\beta_n}^\rho - 1\right| : i \geq \beta_n\right) \geq \eta\right\}$$

converges to 0 as n gets large, it is not difficult to show that the sequences of point processes

$$\left\{\left(\frac{i}{n^{1/(\rho+2)}}, \frac{n}{i^{\rho+1}} W_i^\rho Z_i\right), i \geq 1\right\} \text{ and } \left\{\left(\frac{i}{n^{1/(\rho+2)}}, \frac{n}{i^{\rho+1}} W_{\beta_n}^\rho Z_i\right), i \geq \beta_n\right\}$$

have the same limit in distribution. Because W_{β_n} is independent of $(Z_i, i \geq \beta_n)$, the last point process has the same distribution as $\mathcal{P}_n^2$ (up to the first β_n terms which are negligible similarly as in Step 2). By Step 1, the convergence in distribution is therefore proved. □

The following proposition strengthens the statement of Proposition 1, it will be used to prove the main asymptotic result on the indices of empty bins.

Proposition 3. *If $f : \mathbb{R}_+^2 \to \mathbb{R}_+$ is a continuous function such that*

(1) *there exists K such that $f(x, y) = 0$ for any $x \leq K$ and $y \in \mathbb{R}_+$,*
(2) *for all $x \in \mathbb{R}_+$, the function $y \to f(x, y)$ is differentiable and*

$$y \to y \left\|\frac{\partial f}{\partial y}\right\|_y \stackrel{def.}{=} y \sup_{x \in \mathbb{R}_+} \left|\frac{\partial f}{\partial y}\right|(x, y)$$

is integrable on $\mathbb{R}_+$,

then Convergence (7) also holds for f.

Proof. For M, $L \geq 0$ and i, $n \in \mathbb{N}$, one has

$$\mathbb{E}\left(f\left(\frac{i}{n^{\rho+2}}, nP_i\right) \mathbb{1}_{\{nP_i \geq M, t_{\lfloor\rho\rfloor} \leq L\}}\right)$$
$$= -\int_0^{+\infty} \frac{\partial f}{\partial y}\left(\frac{i}{n^{\rho+2}}, y\right) \mathbb{P}(M \leq nP_i \leq y, t_{\lfloor\rho\rfloor} \leq L)\, dy.$$

By using similar arguments as in the end of the proof of the above theorem, one gets

$$\begin{aligned}
&\mathbb{E}\left(f\left(\frac{i}{n^{\rho+2}},\ nP_i\right) \mathbb{1}_{\{nP_i \geq M, t_{\lfloor\rho\rfloor} \leq L\}}\right) \\
&\qquad \leq e \int_M^{+\infty} \left\|\frac{\partial f}{\partial y}\right\|_y \mathbb{E}\left(1 - \exp\left[-\frac{i}{n\rho} y e^{\rho L} e^{\rho(t_{i-1} - t_{\lfloor\rho\rfloor})}\right]\right) dy \\
&\qquad \leq \frac{i e^{\rho L}}{n\rho} e \mathbb{E}\left(e^{\rho(t_{i-1} - t_{\lfloor\rho\rfloor})}\right) \int_M^{+\infty} y \left\|\frac{\partial f}{\partial y}\right\|_y dy \\
&\qquad \leq C \frac{i^{\rho+1} e^{\rho L}}{n} \int_M^{+\infty} y \left\|\frac{\partial f}{\partial y}\right\|_y dy,
\end{aligned}$$

for some fixed constant C. Define $k_n = \lfloor Kn^{1/(\rho+2)} \rfloor$, by summing up these terms, this gives the relation

$$\mathbb{E}\left(\sum_{i\geq 1} f\left(\frac{i}{n^{\rho+2}}, nP_i\right) \mathbb{1}_{\{M\leq nP_i, t_{\lfloor\rho\rfloor}\leq L\}}\right) \leq C\frac{k_n^{\rho+2}e^{\rho L}}{n}\int_M^{+\infty} y \left\|\frac{\partial f}{\partial y}\right\|_y dy \leq CK^{\rho+2}e^{\rho L}\int_M^{+\infty} y \left\|\frac{\partial f}{\partial y}\right\|_y dy. \tag{9}$$

Define $f_0(x,y) = f(x,y)\mathbb{1}_{\{y\leq M\}}$, by using a convolution kernel on the variable y, there exist sequences (g_p^+) and (g_p^-) in $C_c^+(\mathbb{R}_+)$ converging pointwisely to f_0 for all $y \neq M$ such that $g_p^- \leq f_0 \leq g_p^+$. See Rudin [21] for example. Proposition 1 gives that

$$\mathbb{E}(\exp(-\mathcal{P}(g_p^+))) \leq \liminf_{n\to+\infty} \mathbb{E}(\exp(-\mathcal{P}_n(f_0))) \leq \limsup_{n\to+\infty} \mathbb{E}(\exp(-\mathcal{P}_n(f_0))) \leq \mathbb{E}(\exp(-\mathcal{P}(g_p^-))),$$

and Expression (6) shows that, as p goes to infinity, the left and right hand side terms of this relation converge to the Laplace transform of $\mathcal{P}$ at f_0. Therefore, Convergence (7) holds at f_0. Since

$$\begin{aligned}0 &\leq \mathbb{E}\left(e^{-\mathcal{P}_n(f)}\right) - \mathbb{E}\left(e^{-\mathcal{P}_n(f_0)}\right) \\ &\leq P(t_{\lfloor\rho\rfloor} \geq L) + \mathbb{E}\left[\left(1 - e^{-(\mathcal{P}_n(f)-\mathcal{P}_n(f_0))}\right)\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq L\}}\right] \\ &\leq P(t_{\lfloor\rho\rfloor} \geq L) + \mathbb{E}\left[(\mathcal{P}_n(f) - \mathcal{P}_n(f_0))\,\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq L\}}\right],\end{aligned}$$

and the last term being the left hand side of Relation (9), one can choose L and M sufficiently large so that this difference is arbitrarily small. The proposition is proved. □

4. Asymptotic Behavior of the Indices of the First Empty Bins

It is assumed that a large number n of balls are thrown in the bins according to the probability distribution (P_i) defined by Equation (3). The purpose of this section is to establish limit theorems to describe the limiting distribution of the set of indices of bins having a fixed number of balls.

Theorem 2. *The point process of rescaled indices of empty bins associated to the probability vector (P_i) when n balls have been used*

$$\mathcal{N}_n = \left\{\frac{i}{n^{1/(\rho+2)}} : i \geq 1,\ \text{the } i\text{th bin is empty}\right\}$$

converges in distribution as n goes to infinity to a point process $\mathcal{N}_\infty$ whose distribution is given by

$$\mathbb{E}\left(e^{-\mathcal{N}_\infty(g)}\right) = \mathbb{E}\left[\exp\left(-\frac{W_\infty^{-\rho}}{\rho}\int_{\mathbb{R}_+}\left(1-e^{-g(x)}\right)x^{\rho+1}\,dx\right)\right], \tag{10}$$

for $g \in C_c^+(\mathbb{R}_+)$. Equivalently $(\mathcal{N}_n)$ converges in distribution to the point process

$$\left(W_\infty^{\rho/(\rho+2)}\, t_i^{1/(\rho+2)}\right),$$

where (t_i) *is a standard Poisson process with parameter* $[\rho(\rho+2)]^{-1/(\rho+2)}$.

It can also be shown that the same result holds when the indices of bins containing k balls are considered. If $\mathcal{N}_{k,n}$ is the corresponding point process, the limiting point process does not in fact depend on k and, moreover, the sequence $(\mathcal{N}_{k,n}, k \geq 0)$ converges in distribution to $(\mathcal{N}_{k,\infty}, k \geq 0)$ and, conditionally on W_∞, the random variables $\mathcal{N}_{k,\infty}$, $k \geq 0$ are independent with the same distribution.

Proof. Poissonization. A closely related model is first analyzed when U_n balls are used, U_n being an independent Poisson random variable with mean n, $\mathcal{N}_n^0$ denotes the corresponding point process. For this model, conditionally on the sequence (P_i), the number of balls in the bins are independent Poisson random variables with respective parameters (nP_i). In a first step, the convergence in distribution of the sequence $(\mathcal{N}_n^0)$ of point processes is investigated. Let $g \in C_c^+(\mathbb{R}_+)$,

$$\mathbb{E}\left(e^{-\mathcal{N}_n^0(g)}\right) = \mathbb{E}\left(\exp\left[\sum_{i=1}^{+\infty}\log\left[1-e^{-nP_i}\left(1-e^{-g(i/n^{1/(\rho+2)})}\right)\right]\right]\right),$$

if one defines $f(x,y) = -\log\left[1-e^{-y}\left(1-e^{-g(x)}\right)\right]$, then

$$\mathbb{E}\left(\exp\left[-\mathcal{N}_n^0(g)\right]\right) = \mathbb{E}\left(\exp\left[-\mathcal{P}_n(f)\right]\right),$$

where $\mathcal{P}_n$ is the point process defined in Theorem 1. By using Proposition 3, one gets the relation

$$\begin{aligned}\lim_{n\to+\infty}\mathbb{E}\left(e^{-\mathcal{N}_n^0(g)}\right) &= \mathbb{E}\left[\exp\left(-\frac{W_\infty^{-\rho}}{\rho}\int_{\mathbb{R}_+^2}\left(1-e^{-f(x,y)}\right)x^{\rho+1}\,dx\,dy\right)\right] \\ &= \mathbb{E}\left[\exp\left(-\frac{W_\infty^{-\rho}}{\rho}\int_{\mathbb{R}_+}\left(1-e^{-g(x)}\right)x^{\rho+1}\,dx\right)\right].\end{aligned}$$

For $0<\alpha<1$, it is not difficult to check that the same result holds for the case when U_{n+n^α} balls are used, $\mathcal{N}_n^1$ denotes the associated point process. For $x>0$, the monotonicity property $\mathcal{N}_a([0,x]) \leq \mathcal{N}_b([0,x])$ for $b \leq a$ gives the relation

$$\mathbb{P}(\mathcal{N}_n([0,x]) \leq k) \leq \mathbb{P}\left(\mathcal{N}_n^1([0,x]) \leq k\right) + \mathbb{P}(U_{n+n^\alpha} \leq n).$$

The central limit theorem for Poisson processes shows that for $\alpha \in (1/2,1)$, the quantity $\mathbb{P}(U_{n+n^\alpha} \leq n)$ converges to 0 as n gets large, therefore if $k \geq 0$,

$$\limsup_{n\to+\infty}\mathbb{P}(\mathcal{N}_n([0,x]) \leq k) \leq \lim_{n\to+\infty}\mathbb{P}\left(\mathcal{N}_n^1([0,x]) \leq k\right).$$

By using a similar argument with the lim inf, one gets that the sequences $(\mathcal{N}_n)$ and $(\mathcal{N}_n^0)$ converge in distribution and have the same limit. The proposition is proved. □

Corollary 1. *If* ν_n *is the index of the first empty bin when* n *balls are thrown, then*

$$\lim_{n\to+\infty}\mathbb{P}\left(\frac{\nu_n}{n^{1/(\rho+2)}} \geq x\right) = \mathbb{E}\left(\exp\left(-\frac{x^{\rho+2}W_\infty^{-\rho}}{\rho(\rho+2)}\right)\right), \quad x \geq 0.$$

Comparison with Deterministic Power Law Decay. For $\delta>1$, one considers the bins and balls problem with the probability vector $Q=(Q_i, i\geq 1)=\left(\alpha/i^\delta\right)$. Note that for the problems analyzed in this paper, only the asymptotic behavior of the sequence (Q_i) matters. The equivalent of Theorem 2 can be obtained directly from Theorem 1 of Simatos *et al.* [23].

Proposition 4. *As n goes to infinity, the point process*

$$\left\{\frac{i(\log n)^{1/\delta-1}}{(\alpha\delta n)^{1/\delta}}-\frac{1+\delta}{\delta}\log\log n: \textit{ the } i\textit{th bin is empty}\right\}$$

converges in distribution to a Poisson point process with the intensity measure $(\alpha\delta)^{1/\delta}e^x\,dx$ on $\mathbb{R}$.

The probability vector considered in the above theorem has an asymptotic expression of the form $(P_i)=(W_\infty^\rho E_i/i^{\rho+1})$. In this case, empty bins show up for indices of the order of $n^{1/(\rho+2)}$, i.e., much earlier than for the deterministic case where the exponent of n is $1/\delta=1/(\rho+1)$ (if one ignores the log). This can be explained simply by the fact that some of the i.i.d. exponential random variables (E_i) can be very small thereby creating an additional possibility of having empty bins.

In this picture, the variable W_∞ does not seem to have an influence on the qualitative behavior of these occupancy schemes other than creating some dependency structure for the vector (P_i). The next section shows that this variable has nevertheless an important role if one looks at the averages of the number of empty bins.

5. Rare Events

By Equation (7) of Theorem 1, for $x>0$, the limiting number (in distribution) of empty bins whose index is less than $xn^{1/(\rho+2)}$ has an average value given by

$$\frac{x^{\rho+2}}{\rho(\rho+2)}\mathbb{E}\left(W_\infty^{-\rho}\right)=\frac{x^{\rho+2}}{\rho(\rho+2)}\int_0^{+\infty}\frac{1}{u^\rho}e^{-u}\,du$$

by Equation (2) and since $W_\infty=\exp(-M_\infty)$.This quantity is infinite when $\rho\geq 1$. The purpose of this section is to investigate this phenomenon which has a significant impact on the system at the origin of this model. It is assumed throughout this section that $\rho\geq 1$.

Definition 1. If $\phi:\mathbb{N}\to\mathbb{R}_+$ is a non-decreasing function, for $n\geq 1$, $\mathcal{N}_n^\phi$ denotes the point process defined by

$$\mathcal{N}_n^\phi=\left\{\frac{i}{\phi(n)}: i\geq 1, \text{ the } i\text{th bin is empty}\right\}.$$

For $i>\lfloor\rho\rfloor$, the quantity P_i can be written as $P_i=\exp(-\rho t_{\lfloor\rho\rfloor})D_iZ_i/i^{\rho+1}$ with

$$D_i\stackrel{\text{def.}}{=}\exp\left(-\rho\left[M_i-M_{\lfloor\rho\rfloor}-\log\lfloor\rho\rfloor\right]\right).$$

The sequence (D_i) converges almost surely to a finite limit D_ρ given by

$$D_\rho\stackrel{\text{def.}}{=}\exp\left(-\rho\left[M_\infty-M_{\lfloor\rho\rfloor}-\log\lfloor\rho\rfloor\right]\right), \tag{11}$$

and, since $\exp(\rho E_i/i)$ is integrable for $i > \rho$, a similar result holds for the expected values

$$\lim_{i\to+\infty} \mathbb{E}(1/D_i) = \mathbb{E}(1/D_\rho) < +\infty.$$

With this definition, the asymptotic representation of (P_i) can be given as $P_i = \exp(-\rho t_{\lfloor\rho\rfloor})D_\rho E/i^{\rho+1}$ where E is an independent exponential random variable with parameter 1. In a similar way as before, this representation can be shown to be valid for the results obtained in this section.

For $0 \le p \le 1$ and $n \ge 1$, the elementary inequality

$$\left|e^{-np} - (1-p)^n\right| \le \frac{p^2}{2} n e^{-np} \le \frac{2e^2}{n}$$

gives directly the following lemma which will be used repeatedly in this section.

Lemma. *For a non-decreasing function ϕ, $x \ge 0$, and $n \ge 1$, then*

$$\left|\mathbb{E}\left(\mathcal{N}_n^\phi([0,x])\right) - \sum_{i=1}^{\lfloor x\phi(n)\rfloor} \mathbb{E}\left(e^{-nP_i}\right)\right| \le 2e^2\frac{\lfloor x\phi(n)\rfloor}{n}.$$

When $\phi(n) \ll n$, this lemma implies that to study the asymptotic behavior of $(\mathbb{E}\left(\mathcal{N}_n^\phi([0,x])\right))$, it is enough to analyze the convergence of the corresponding sum of the $\mathbb{E}\left(e^{-nP_i}\right)$. For the moment, $k \in \mathbb{N}$ is fixed, if $n \ge 1$, $i > \rho$, then

$$\begin{aligned}\mathbb{E}\left(e^{-nP_i}\right) &= \mathbb{E}\left[\exp\left(-nD_\rho e^{-\rho t_{\lfloor\rho\rfloor}} E/i^{\rho+1}\right)\right] \\ &= \mathbb{E}\left(\frac{i^{\rho+1}/n}{i^{\rho+1}/n + e^{-t_{\lfloor\rho\rfloor}}D_\rho}\right),\end{aligned}$$

by summing up these terms, if $\varepsilon_{k,n} \stackrel{\text{def.}}{=} k/n^{1/(\rho+1)}$, one gets that

$$\sum_{i=\lfloor\rho\rfloor+1}^{k} \mathbb{E}\left(e^{-nP_i}\right) = n^{1/(\rho+1)} \int_0^{\varepsilon_{k,n}} \mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1} + e^{-t_{\lfloor\rho\rfloor}}D_\rho}\right) dv + O\left(\varepsilon_{k,n}\right),$$

which gives the relation

$$\sum_{i=1}^{k} \mathbb{E}\left(e^{-nP_i}\right) = n^{1/(\rho+1)}\varepsilon_{k,n}^{\rho+2} \int_0^1 \mathbb{E}\left(\frac{v^{\rho+1}}{\varepsilon_{k,n}^{\rho+1}v^{\rho+1} + e^{-t_{\lfloor\rho\rfloor}}D_\rho}\right) dv + O\left(\varepsilon_{k,n}\right),$$

with a change of variable. By using Equation (1) and again a change of variable, one obtains the relation

$$\begin{aligned}(12)\quad \sum_{i=1}^{k} \mathbb{E}\left(e^{-nP_i}\right) &= \frac{\lfloor\rho\rfloor}{\rho} n^{1/(\rho+1)}\varepsilon_{k,n}^{(2\rho+1)/\rho} \\ &\times \int_0^{1/\varepsilon^{\rho+1}} u^{1/\rho-1}(1-\varepsilon_{k,n}^{\frac{\rho+1}{\rho}}u^{1/\rho})^{\lfloor\rho\rfloor-1}du \int_0^1 \mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1}+uD_\rho}\right) dv + O\left(\varepsilon_{k,n}\right).\end{aligned}$$

This quantity is now analyzed according to the values of ρ.

Case $\rho > 1$.
If $k_n = \lfloor xn^\alpha\rfloor$ with $1/(2\rho+1) \le \alpha < 1/(\rho+1)$, then $\varepsilon_{k_n,n} \sim xn^{(\alpha(\rho+1)-1)/(\rho+1)}$ and,

by Relation (12),

$$\lim_{n\to+\infty} \frac{1}{n^{((2\rho+1)\alpha-1)/\rho}} \sum_{i=1}^{k_n} \mathbb{E}\left(e^{-nP_i}\right) = x^{(2\rho+1)/\rho}\,\frac{\lfloor\rho\rfloor}{\rho}\int_0^{+\infty} u^{1/\rho-1}du \int_0^1 \mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1}+uD_\rho}\right)dv.$$

Case $\rho = 1$.
Equation (12) is for this case

$$\sum_{i=1}^{k} \mathbb{E}\left(e^{-nP_i}\right) = \sqrt{n}\varepsilon_{k,n}^3 \int_0^{1/\varepsilon_{k,n}^2} du \int_0^1 \mathbb{E}\left(\frac{v^2}{v^2+uD_1}\right)dv + O\left(\varepsilon_{k,n}\right).$$

If k_n=$\lfloor xn^{1/3}/\log^\beta n\rfloor$ with $\beta\in\mathbb{R}$, then $\varepsilon_{k_n,n}\sim x/(n^{1/6}(\log n)^\beta)$ and for $\beta\leq 1/3$,

$$\lim_{n\to+\infty} \frac{1}{(\log n)^{(1-3\beta)}} \sum_{i=1}^{k_n} \mathbb{E}\left(e^{-nP_i}\right) = \frac{1}{9}x^3\mathbb{E}\left(\frac{1}{D_1}\right).$$

The following proposition has therefore been proved.

Proposition 5 (Average of the Number of Empty Bins)**.** *For α, $\beta>0$, for $n\in\mathbb{N}$, denote by $p_{\alpha,\beta}(n)=n^\alpha(\log n)^{-\beta}$, and by convention $p_\alpha=p_{\alpha,0}$.*

(1) *If $\rho>1$ and $1/(2\rho+1)\leq\alpha<1/(\rho+1)$,*

$$\lim_{n\to+\infty} \frac{1}{n^{(\alpha(2\rho+1)-1)/\rho}} \mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x])\right) = x^{(2\rho+1)/\rho}\,\frac{\lfloor\rho\rfloor}{\rho}\int_0^{+\infty} u^{1/\rho-1}du \int_0^1 \mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1}+uD_\rho}\right)dv.$$

(2) *If $\rho=1$ and $\beta\leq 1/3$,*

$$\lim_{n\to+\infty} \frac{1}{(\log n)^{(1-3\beta)}} \mathbb{E}\left(\mathcal{N}^{p_{1/3,\beta}}([0,x])\right) = \frac{1}{9}x^3\mathbb{E}\left(\frac{1}{D_1}\right)$$

A Double Threshold. For the convergence in distribution of the sequence of point processes $(\mathcal{N}_n^\phi)$, Theorem 2 has shown that the correct scaling ϕ for the order of magnitude of the indices of the first empty bins is given by $\phi(n)=n^{1/(\rho+2)}$, $n\geq 1$. For the average number of points in a finite interval, the above proposition states that, for $\rho>1$, the correct scaling is in fact $\phi(n)=n^{1/(2\rho+1)}\ll n^{1/(\rho+2)}$.

For $\alpha>0$, with the notations of the above proposition, one concludes that under the condition $\rho>1$ and for $1/(2\rho+1)<\alpha<1/(\rho+2)$, the following limit results hold

$$\mathcal{N}_n^{p_\alpha} \overset{\text{dist.}}{\to} 0 \text{ and } \lim_{n\to+\infty}\mathbb{E}\left(\mathcal{N}_n^{p_\alpha}[0,x]\right)=+\infty, \quad \forall x>0.$$

This suggests that, in this case, with a high probability, all the bins with index less than $n^{1/(\rho+2)}$ have a large number of balls. But also that there exists some rare event for which a very large number of empty bins with indices of an order slightly greater than $n^{1/(2\rho+1)}$ are created. The following proposition shows that the total size of the first $\lfloor\rho\rfloor$ bins is the key variable to explain this phenomenon. It should be of the order of $\log n$ in order to have sufficiently many empty bins in the appropriate region.

Proposition 6. *For $\rho > 1$ and if $p_\alpha(n) = n^\alpha$, for $\alpha \in [1/(2\rho+1), 1/(\rho+2))$ and*

$$\delta_0(\alpha) \overset{def.}{=} \frac{1-\alpha(\rho+2)}{\rho-1} \quad \text{and} \quad \delta_1(\alpha) \overset{def.}{=} \frac{1-\alpha(\rho+1)}{\rho},$$

then, for $a \in \mathbb{R}$ and $x > 0$,

(1) *If $\delta < \delta_0(\alpha)$, then*

$$\lim_{n\to+\infty} \mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x])\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq \delta\log n\}}\right) = 0.$$

(2) *If $\delta \in [\delta_0(\alpha), \delta_1(\alpha)[$, then*

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x])\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq \delta\log n+a\}}\right)}{n^{(\rho+2)\alpha+\delta(\rho-1)-1}} = \frac{x^{\rho+2}}{(\rho+2)}\frac{\lfloor\rho\rfloor}{(\rho-1)}\mathbb{E}\left(\frac{1}{D_\rho}\right)e^{(\rho-1)a}.$$

(3) *If $\delta \geq \delta_1(\alpha)$,*

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x])\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq \delta\log n+a\}}\right)}{n^{((2\rho+1)\alpha-1)/\rho}}$$
$$= x^{(2\rho+1)/\rho}\frac{\lfloor\rho\rfloor}{\rho}\int_{e^{-\rho a}\mathbb{1}_{\{\delta=\delta_1(\alpha)\}}}^{+\infty} u^{1/\rho-1}du\int_0^1 \mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1}+uD_\rho}\right)dv,$$

where D_ρ is the random variable defined by Equation (11).

Proof. To begin with, it is assumed that $\delta \in [\delta_0(\alpha), \delta_1(\alpha))$. If $k \geq 1$, $b > 0$, $\varepsilon_{k,n} = k/n^{1/(\rho+1)}$, $k = \lfloor xn^\alpha \rfloor$ and $b = \delta\log n + a$, in the same way as for Equation (12), one gets

$$\sum_{i=1}^k \mathbb{E}\left(e^{-nP_i}\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq b\}}\right) = \frac{\lfloor\rho\rfloor}{\rho}n^{1/(\rho+1)}\varepsilon_{k,n}^{(2\rho+1)/\rho}$$
$$\times\int_{e^{-\rho b}/\varepsilon_{k,n}^{\rho+1}}^{1/\varepsilon_{k,n}^{\rho+1}} u^{1/\rho-1}(1-\varepsilon_{k,n}^{\frac{\rho+1}{\rho}}u^{1/\rho})^{\lfloor\rho\rfloor-1}du\int_0^1\mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1}+uD_\rho}\right)dv + O(\varepsilon_{k,n})$$

(13)

$$= \frac{\lfloor\rho\rfloor}{\rho}n^{\frac{1}{\rho+1}}\varepsilon_{k,n}^{\frac{2\rho+1}{\rho}}\int_{e^{-\rho b}/\varepsilon_{k,n}^{\rho+1}}^{1/\varepsilon_{k,n}^{\rho+1}} u^{1/\rho-2}du\int_0^1\mathbb{E}\left(\frac{uv^{\rho+1}}{v^{\rho+1}+uD_\rho}\right)dv + O(\varepsilon_{k,n}).$$

Note that

$$e^{-\rho b}/\varepsilon_{k,n}^{\rho+1} \sim n^{1-\rho\delta-\alpha(\rho+1)}e^{-\rho a} \nearrow +\infty,$$

hence the range of the first integral goes to infinity as n gets large. Since

$$\int_0^1\mathbb{E}\left(\frac{uv^{\rho+1}}{v^{\rho+1}+uD_\rho} - \frac{v^{\rho+1}}{D_\rho}\right)dv = \int_0^1\mathbb{E}\left(\frac{v^{2(\rho+1)}}{(v^{\rho+1}+uD_\rho)D_\rho}\right)dv,$$

by Lebesgue's Theorem, this integral is arbitrarily small as u gets large, this implies the equivalence

$$\sum_{i=1}^k \mathbb{E}\left(e^{-nP_i}\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq b\}}\right) \sim \frac{\lfloor\rho\rfloor}{\rho(\rho+2)}\mathbb{E}\left(\frac{1}{D_\rho}\right)n^{1/(\rho+1)}\varepsilon_{k,n}^{(2\rho+1)/\rho}\int_{e^{-\rho b}/\varepsilon_{k,n}^{\rho+1}}^{1/\varepsilon_{k,n}^{\rho+1}} u^{1/\rho-2}du.$$

If C is the multiplicative constant of the right hand side of the above relation, then

$$\sum_{i=1}^{k} \mathbb{E}\left(e^{-nP_i} \mathbb{1}_{\{t_{\lfloor\rho\rfloor} \le b\}}\right) \sim \frac{C\rho}{\rho-1} \frac{k^{\rho+2}}{n} \left(e^{b(\rho-1)}-1\right),$$

this gives the equivalence

$$\sum_{i=1}^{k} \mathbb{E}\left(e^{-nP_i} \mathbb{1}_{\{t_{\lfloor\rho\rfloor} \le b\}}\right) \sim x^{\rho+2} \frac{C\rho}{\rho-1} e^{a(\rho-1)} n^{(\rho+2)\alpha+\delta(\rho-1)-1}.$$

The proof of this case is completed.

The case $\delta \ge \delta_1(\alpha)$ uses Equation (13). The term $e^{-\rho b}/\varepsilon_{k,n}^{\rho+1}$ converges to $e^{-\rho a}$ if $\delta = \delta_1(\alpha)$ and 0 otherwise. This gives directly the desired convergence.

Finally, if $\delta < \delta_0(\alpha)$, for any $a \in \mathbb{R}$, there exists n_0 so that if $n \ge n_0$, then $\delta \log n \le \delta_0(\alpha) \log n + a$, in particular

$$\mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x]) \mathbb{1}_{\{t_{\lfloor\rho\rfloor} \le \delta \log n\}}\right) \le \mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x]) \mathbb{1}_{\{t_{\lfloor\rho\rfloor} \le \delta_0(\alpha) \log n + a\}}\right)$$

hence

$$\limsup_{n\to+\infty} \mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x]) \mathbb{1}_{\{t_{\lfloor\rho\rfloor} \le \delta \log n\}}\right) \le \frac{x^{\rho+2}}{(\rho+2)} \frac{\lfloor\rho\rfloor}{(\rho-1)} \mathbb{E}\left(\frac{1}{D_\rho}\right) e^{(\rho-1)a}.$$

One concludes by letting a go to $-\infty$. □

As a consequence of the above proposition, for $\alpha \in [1/(2\rho+1), 1/(\rho+2))$, the average of the variable $\mathcal{N}_n^{p_\alpha}([0,x])$ converges to infinity only when the total size $t_{\lfloor\rho\rfloor}$ of the first $\lfloor\rho\rfloor$ bins is of the order $\delta \log n$ for a sufficiently large δ. The following corollary gives a more precise formulation.

Corollary 2. *For $\rho > 1$ and if $p_\alpha(n) = n^\alpha$, for $\alpha \in [1/(2\rho+1), 1/(\rho+2))$*

$$\delta_1(\alpha) = (1-\alpha(\rho+1))/\rho,$$

then, for a, $b > 0$,

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x]) \mathbb{1}_{\{\delta_1(\alpha)\log n - a \le t_{\lfloor\rho\rfloor} \le \delta_1(\alpha) \log n + b\}}\right)}{\mathbb{E}\left(\mathcal{N}_n^{p_\alpha}([0,x])\right)} = \psi(-a,b)$$

where, for y, $z \in \mathbb{R}$, $\psi(y,z) = \phi(y,z)/\phi(-\infty,+\infty)$ and

$$\phi(y,z) = x^{(2\rho+1)/\rho} \frac{\lfloor\rho\rfloor}{\rho} \int_{[e^{-\rho z}, e^{-\rho y}]} u^{1/\rho-1} du \int_0^1 \mathbb{E}\left(\frac{v^{\rho+1}}{v^{\rho+1}+uD_\rho}\right) dv.$$

A rough (non-rigorous) interpretation of this result could be as follows: on the event where "most" (i.e., for the averages) of empty bins are created in the interval $[0, xn^\alpha]$, the random variable $t_{\lfloor\rho\rfloor} - \delta_1(\alpha)\log n$ converges in distribution to some random variable X on $\mathbb{R}$, such that $\mathbb{P}(X \le a) = \psi(-\infty, a)$.

The following analogous result is proved in a similar way for the critical case $\rho = 1$.

Proposition 7. *For $\rho = 1$ and with the notations of the above proposition then, for $0 < \beta < 1/3$, $x > 0$, and for $0 \le a \le 1/3$,*

$$\lim_{n\to+\infty} \frac{1}{(\log n)^{1-3\beta}} \mathbb{E}\left(\mathcal{N}_n^{p_{1/3,\beta}}([0,x]) \mathbb{1}_{\{t_{\lfloor\rho\rfloor} \le a \log n\}}\right) = \frac{a}{3} x^3 \mathbb{E}\left(\frac{1}{D_1}\right),$$

and for $a > 1/3$,

$$\lim_{n\to+\infty} \frac{1}{(\log n)^{1-3\beta}} \mathbb{E}\left(\mathcal{N}_n^{p_{1/3,\beta}}([0,x])\mathbb{1}_{\{t_{\lfloor\rho\rfloor}\leq a\log n\}}\right) = \frac{1}{9}x^3\mathbb{E}\left(\frac{1}{D_1}\right),$$

where D_1 is the random variable defined by Equation (11)*.*

6. Generalizations

The problem analyzed in the present paper can be generalized towards two directions. On one hand, the sequence (t_n) can stem from a general branching process instead of the particular Yule one; on the other hand, the locations of balls can have a general distribution. This section discusses these possible extensions.

Exponential Balls and General Branching Process. Let (t_n) be the birth instants of a general supercritical branching process $(Z(t))$. See Kingman [13] and Nerman [16] for example. Let α be the Malthusian parameter, and W the almost sure limit of $(e^{-\alpha t}Z(t))$. Under reasonable technical assumptions, Härnqvist [9] has shown the following result:

Theorem 3. *Define the point process Ψ_t^* by*

$$\Psi_t = \sum_{k\geq 1} \mathbb{1}_{\{t\leq t_k\}}\delta_{t_k e^{\alpha t}},$$

as t gets large, Ψ_t converges in distribution to a mixed Poisson process whose parameter is distributed as γW for some constant $\gamma > 0$.

From this result, it is possible to prove that the process $(n(t_{n+k}-t_n), k\geq 1)$ converges in distribution, as n goes to infinity, to a Poisson process: clearly

$$\Psi_{t_n} = \sum_{k\geq 1} \delta_{(t_{n+k}-t_n)e^{\alpha t_n}},$$

and provided that, up to a multiplicative constant, $e^{\alpha t_k}/k$ converges to W, the point process $\sum_{k\geq 1}\delta_{n(t_{n+k}-t_n)}$ should converge to a Poisson random variable with a deterministic parameter. In this case the probability that a ball falls into the nth bin which is given by

$$P_n = e^{-\rho t_{n-1}}(1-e^{-\rho(t_n-t_{n-1})}),$$

has therefore the following asymptotic behavior

$$P_n \sim n^{-\rho/\alpha}W^{\rho/\alpha}E_i,$$

where (E_i) are i.i.d. exponential random variables. In the Bellman-Harris case, following Athreya and Kaplan [1], it is possible to show that W and (E_i) are independent, so that in this case, the asymptotic behavior of (P_n) is exactly the same as in the case of a Yule process. One can conjecture that this independence property still holds in the general case.

The main obstacle to generalize the results of this paper, even in the Bellman-Harris case, is that although W and (E_i) are independent, t_{n-1} and $t_n - t_{n-1}$ are not independent. In the proof of Proposition 1, this independence plays a crucial role, it has therefore to be generalized to variables which are only asymptotically independent. Additionally, since the heavy tail property of the limiting variable $W_\infty^{-\rho}$ is also true in the general case, see e.g., Liu [15], a similar rare events phenomenon to the one described in Section 5 is plausible in this case.

General Balls and Yule Process. When the underlying branching process is changed, the above discussion suggests that the asymptotic behavior of the sequence (P_n) remains essentially the same as for a Yule process. The situation changes significantly when the law of the location X of a ball is changed, in this case with the same notations as before for the Yule process,

$$P_n = \mathbb{P}(t_{n-1} < X \leq t_{n-1} + E_n/n).$$

The tail distribution of X then plays a key role. Consider for instance a power law, i.e., $\mathbb{P}(X \geq x)$ behaves as $\delta x^{-\beta}$ for some β and $\delta > 0$: then

$$P_{n+1} \sim t_n^{-\beta} - (t_n + E_{n+1}/(n+1))^{-\beta} \sim \frac{\beta\delta E_{n+1}}{n t_n^{\beta+1}} \sim \frac{\beta\delta E_n}{n(\log n)^{\beta+1}},$$

and it can be seen that the random variable W_∞ may not play a role anymore in the asymptotic behavior of the system.

References

[1] K. B. Athreya and N. Kaplan. Convergence of the Age Distribution in the One-Dimensional Supercritical Age-Dependent Branching Process. *Annals of Probability*, 4(1):38–50, February 1976.

[2] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[3] A. D. Barbour, Lars Holst, and Svante Janson. *Poisson approximation*. The Clarendon Press Oxford University Press, New York, 1992. Oxford Science Publications.

[4] Endre Csáki and Antónia Földes. On the first empty cell. *Studia Scientarum Mathematicarum Hungarica*, 11:373–382, 1976.

[5] Donald A. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of *Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.

[6] Philippe Flajolet and G. Nigel Martin. Probabilistic counting algorithms for data base applications. *Journal of Computer and System Sciences*, 21:182–209, 1985.

[7] Alexander V. Gnedin. The Bernoulli Sieve. *Bernoulli*, 10(1):79–96, 2004.

[8] Alexander V. Gnedin, Ben Hansen, and Jim Pitman. Notes on the occupancy problem with infinitely many boxes: General asymptotics and power laws. *Probability Surveys*, 4:146–171, 2007.

[9] Martin Härnqvist. Limit theorems for point processes generated in a general branching process. *Advances in Applied Probability*, 13:650–668, 1981.

[10] Hsien-Kuei Hwang and Svante Janson. Local limit theorems for finite and infinite urn models. *The Annals of Probability*, 36(3):992–1022, 2008.

[11] Norman L. Johnson and Samuel Kotz. *Urn models and their application : an approach to modern discrete probability theory*. New York: Wiley, 1977.

[12] Samuel Karlin. Central limit theorems for certain infinite urn schemes. *Journal of Mathematics and Mechanics*, 17:373–401, 1967.

[13] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Annals of Probability*, 3(5):790–801, 1975.

[14] Valentin F. Kolchin, Boris A. Sevast′yanov, and Vladimir P. Chistyakov. *Random allocations*. V. H. Winston & Sons, Washington, D.C., 1978.

[15] Quansheng Liu. Asymptotic properties and absolute continuity of laws stable by a random weighted mean. *Stochastic Processes and their Applications*, 95:83–107, 2001.

[16] Olle Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57(3):365–395, 1981.

[17] Jacques Neveu. *Martingales à temps discret*. Masson et Cie, éditeurs, Paris, 1972.

[18] Jacques Neveu. Processus ponctuels. In P.-L. Hennequin, editor, *École d'Été de Probabilités de Saint-Flour*, volume 598 of *Lecture Notes in Math.*, pages 249–445. Springer-Verlag, Berlin, 1977.

[19] Philippe Robert. *Stochastic Networks and Queues*, volume 52 of *Stochastic Modelling and Applied Probability Series*. Springer, New-York, June 2003.

[20] Philippe Robert and Florian Simatos. A point process approach to occupancy problems. In Preparation, 2008.

[21] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[22] Walid Saddi and Fabrice Guillemin. Measurement based modeling of edonkey peer-to-peer file sharing system. In *Proc. International Teletraffic Congress 20*, Ottawa, June 2007.

[23] Florian Simatos, Philippe Robert, and Fabrice Guillemin. A queueing system for modeling a file sharing principle. In *SIGMETRICS'08*, 2008.

[24] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

(Ph. Robert, F. Simatos) INRIA Paris — Rocquencourt, Domaine de Voluceau, BP 105, 78153 Le Chesnay, France.

E-mail address: Philippe.Robert@inria.fr

E-mail address: Florian.Simatos@inria.fr

URL: http://www-rocq.inria.fr/~robert