Indian J. Pure Appl. Math., **51**(1): 321-344, March 2020
 DOI: 10.1007/s13226-020-0403-y

ON NEW MECHANISMS LEADING TO HEAVY-TAILED DISTRIBUTIONS RELATED TO THE ONES OF YULE-SIMON

Thierry E. Huillet

Laboratoire de Physique Théorique et Modélisation
CNRS-UMR 8089 *et Université de Cergy-Pontoise,*
2 *Avenue Adolphe Chauvin*, 95302, *Cergy-Pontoise, France*
e-mail: Thierry.Huillet@u-cergy.fr

(*Received* 6 *February* 2018; *after final revision* 22 *January* 2019;
accepted 6 *February* 2019)

Scientists reinvent stochastic mechanisms leading to the emergence of a distribution discovered by H.A. Simon, in the context of the study of word frequencies occurring in a textbook. Simon distributions are heavy-tailed as a result of a reinforcement mechanism that produced them, related to the modern notion of preferential attachment. The Simon distribution is a particular case of a distribution recently introduced, itself extending the Sibuya distribution. We exhibit some of the remarkable statistical properties of such a family of distributions, in particular the one of being discrete self-decomposable. Using this and after placing this problem in context, additional stochastic processes where such distributions naturally arise are investigated, in particular a Markov chain model with catastrophes.

"Felix qui potuit rerum cognoscere causas." (Virgil, Georgics, II).

Key words : Yule-Simon; Sibuya; self-decomposability; Gauss hypergeometric function; urn model; death process with immigration; Markov chain with catastrophes.

2010 Mathematics Subject Classification : 60E05 (60E07, 60J10)

1. INTRODUCTION

Part of the remarkable work [24] was to reinvent mechanisms (stochastic processes) leading to the emergence of the Simon distribution discovered in [25], in the context of the study of word frequencies occurring in a textbook. Simon distributions are heavy-tailed as a result of a reinforcement mechanism

that produced them, related to the modern notion of preferential attachment, [8]. They participate to 'the richer get richer' myth and reality. The Simon distribution is a particular case of a generalized distribution recently introduced in [15], itself extending the Sibuya distribution, [23]. We exhibit some of the remarkable statistical properties of such a generalized family, in particular the one of being discrete self-decomposable. Using this and after placing this problem in context, additional stochastic processes where such distributions naturally arise are investigated. This includes: (i) A first hitting time generation in an inhomogeneous success run problem; (ii) A Pólya-Eggenberger urn model generation where the replacement of balls depends on the initial composition of the urn; (iii) A pure death process balanced by incoming immigrants. Being self-decomposable, the distributions under study occur as the limiting equilibrium size of some population whose transient growth results from a trade-off between a pure-death mechanism which tends to shrink the population size, against an input immigration mechanism which tends to have it increased. This both in discrete or continuous time; (iv) An inflating Markov chain model with catastrophe whereby the collapse probabilities constitute a decreasing sequence of the current population size (more specifically inversely proportional to it), a mechanism favoring large equilibrium population size. There is an intimate relationship of this Markov chain with the Sibuya distributions.

Such processes complete the picture of classical processes where the Simon distribution is involved, namely the original Simon model and the Yule-Simon species mutation model.

2. Simon Discrete Model With Tail Index $\alpha > 0$, [25]

We shall first recall the Simon construction of the Simon distribution with parameter $\alpha > 1$ before switching to the Yule-Simon construction involving Simon distributions with any parameter $\alpha > 0$. The critical value $\alpha = 1$ plays a key-role in the classical theory, [26].

2.1 *The Simon construction, [25].*

(We adopt here a different but equivalent image than the original one concerning word frequencies): Every day a naturalist is dispatched from his lab to sample species in Nature. Once he/she meets a species, either new or already sampled, he/she returns back to his lab before proceeding to a new sampling campaign the next day. He/she records the sampled species together with their occurrences.

After n campaigns, let $N_n(k)$ be the number of species sampled k times, with $n = \sum_{k=1}^{n} kN_n(k)$ and $P_n = \sum_{k=1}^{n} N_n(k)$, the number of distinct species discovered in the process. Let $x_n(k) := \mathbf{E}N_n(k)$ and $p_n := \sum_k x_n(k)$, the expected number of distinct species discovered by time n. How can $N_n(k)$ be built up? Fix the $N_n(k)$s and consider the $(n+1)^{\text{th}}$ campaign. Suppose the following

process is at stake when moving from step n to $n+1$:

- there is a probability π to sample a new species so in this case, $N_n(1) \to N_{n+1}(1) = N_n(1)+1$. And $P_n \overset{d}{\sim}$bin(n,π) with $\mathbf{E}P_n = p_n = n\pi$.

- With probability $1-\pi$, the outcome of the $(n+1)^{\text{th}}$ campaign is a species already visited, and it will be species k with probability $kN_n(k)/n$ (a reinforcement property enhancing species visited often). If $k \neq 1$ therefore, $N_n(k)$ grows by one unit with probability $(1-\pi)(k-1)N_n(k-1)/n$, decreases by one unit $kN_n(k)/n$ or stays alike if a new species is sampled or if it is not new but one different from species $\{k-1,k\}$. Taking the average,

$$\begin{cases} x_{n+1}(k) - x_n(k) = (1-\pi)(k-1)x_n(k-1)/n - (1-\pi)kx_n(k)/n \text{ if } k \neq 1 \\ x_{n+1}(1) - x_n(1) = \pi - (1-\pi)x_n(1)/n. \end{cases}$$

Putting $\alpha = 1/(1-\pi) > 1$, the solutions are of the form $x_n(k) = nx(k)$ with

$$x(k) = \frac{k-1}{k+\alpha}x(k-1), k \geq 2,$$

entailing $x(1) = \frac{\pi\alpha}{1+\alpha}$ and

$$x(k) = \pi\alpha B(k, \alpha+1),$$

where $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is the beta function. The sequence

$$\pi_k^+ := \frac{x_n(k)}{\sum_k x_n(k)} = \frac{nx(k)}{p_n} = \frac{x(k)}{\pi} = \alpha B(k, \alpha+1), k \geq 1 \tag{1}$$

is a probability distribution on the positive integers $\mathbb{N} = \{1,2,...\}$ (of some random variable say K^+) known as the Simon distribution. Due to the reinforcement property involved in its generation, π_k^+ is heavy-tailed $\pi_k^+ \underset{k\to\infty}{\sim} \alpha\Gamma(\alpha+1)k^{-(\alpha+1)}$, so with tail index α. It obeys the recursion $\pi_{k+1}^+/\pi_k^+ = k/(k+\alpha+1) < 1$, $\pi_1^+ = \alpha/(\alpha+1) < 1$. Its probability generating function (pgf), denoted by $\varphi^+(z)$ in the sequel, is seen to be

$$\varphi^+(z) := \sum_{k\geq 1}\pi_k^+ z^k =: \mathbf{E}\left(z^{K^+}\right) = \frac{\alpha z}{\alpha+1}F(1,1;2+\alpha;z), \tag{2}$$

involving a special Gauss hypergeometric function[1]. Coming back to the random variables, it can be shown that

Proposition 1 — [19].

$$\frac{N_n(k)}{\sum_{k\geq 1}N_n(k)} = \frac{N_n(k)}{P_n} \underset{n\to\infty}{\to} \pi_k^+ \tag{3}$$

[1]By $F(a,b;c;z)$, we mean ${}_2F_1(a,b;c;z)$, the Gauss hypergeometric function.

for each k, in probability (and thus in distribution).

The latter Simon construction does not exhibit a distribution $\pi_k^+ = \mathbf{P}\left(K^+ = k\right)$ for which $\alpha \in (0,1]$ which clearly exists (and has infinite mean). Indeed, if π is small approaching 0, α can only approach 1 from above. We now come to an alternative model yielding a Simon distribution with parameter $\alpha > 0$

2.2 *Yule mutations model with $\alpha > 0$, [32].*

A genus starts with a single species. As a result of specific mutations, new species arrive according to a linear pure birth Yule process with birth rate λ and they all belong to the same genus. Concomitantly, inside a genus, a species of a novel genus can be created at rate μ and the new genus, once it has appeared, behaves like all the previous ones. The latter event thus results from generic mutations. Consequently, the number of species inside a genus at time t after its appearance contains $K_\lambda^+(t)$ species with

$$\mathbf{P}\left(K_\lambda^+(t) = k\right) = e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{k-1}, k \in \mathbb{N}$$

Taking into account the generation of new genera, putting $\alpha = \mu/\lambda > 0$, the number of species K^+ in a randomly chosen genus is thus

$$\mathbf{P}\left(K^+ = k\right) = \mu \int_0^\infty dt e^{-\mu t} e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{k-1} = \alpha B\left(k, \alpha + 1\right) = \pi_k^+,$$

so with Simon distribution with parameter $\alpha > 0$. It is indeed well-known that the Yule-Simon pgf $\varphi^+(z)$ may be written as an exponential(μ) mixture of a geometric distribution with success parameter $e^{-\lambda t}$, viz:

$$\begin{aligned} \varphi^+(z) &= \mu \int_0^\infty dt \cdot e^{-\mu\tau} \frac{e^{-\lambda t} z}{1 - (1 - e^{-\lambda t}) z} \\ &\text{or} \\ \pi_k^+ &= \mu \int_0^\infty dt e^{-\mu t} e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{k-1}. \end{aligned} \tag{4}$$

This model was designed by Yule to interpret the taxonomic data of Willis, [31]. The Yule occurrence of a Simon distribution permits $\alpha > 0$. In particular, $\alpha = 1$ corresponds to $\lambda = \mu$ and in this case, $\varphi^+(z) = \frac{z}{2} F(1,1;3;z) = 1 + \frac{1-z}{z}\log(1-z)$ corresponding to $\pi_k^+ = 1/[k(k+1)]$, $k \geq 1$, [18].

2.3 *First hitting time generation, $\alpha > 0$*

With $[a]_k := a\,(a+1)\,...\,(a+k-1)$ the rising factorials of a, we have

$$\pi_k^+ = \alpha B\,(k, \alpha+1) = \frac{\alpha\,(k-1)!}{[\alpha+1]_k} = \frac{\alpha}{\alpha+k}\prod_{l=1}^{k-1}\left(1-\frac{\alpha}{\alpha+l}\right).$$

This shows that the Yule-Simon random variable (rv) K^+ is an integral-valued random variable with support $\mathbb{N}$ which can also be generated as:

$$K^+ = \inf\left(l \geq 1 : \mathcal{B}_\alpha\,(l) = 1\right), \tag{5}$$

where $(\mathcal{B}_\alpha\,(l)\,;\ l \geq 1)$ is a sequence of independent Bernoulli rvs obeying $\mathbf{P}\,(\mathcal{B}_\alpha\,(l) = 1) = \alpha/\,(\alpha+l)$, with $\alpha > 0$. It is thus the first epoch of a success in a Bernoulli trials sequence when the probability of success is $\alpha/\,(\alpha+l)$, in particular inversely proportional to the number of the trial. From this success run representation, it is clear that K^+ either takes on either small values close to 1 (the mode of K^+ is at $k=1$) or very large values (responsible of its heavy-tailedness).

The Yule-Simon rv K^+ cannot be infinitely divisible (meaning compound Poisson) because $\pi_0^+ = 0$**.** But considering $\pi_k = \pi_{k+1}^+$, $k \geq 0$, from the Yule construction,

$$\pi_k = \alpha\int_0^1 u^k\,(1-u)^\alpha\,du$$

admits a Hausdorff representation (insuring the complete monotonicity of π_k), showing that upon shifting the Simon variable by -1, the new variable is infinitely divisible or compound Poisson, (see Theorem 10.4 of [27]).

2.4 *Pólya-Eggenberger urn model interpretation, α rational, [13].*

In [24], it is loosely stated that the Simon distribution has to do with an urn problem. This was also observed in [20]. Here we give an alternative description in terms of a Pólya-Eggenberger urn model for which the replacement rule of balls depends on the initial composition of the urn.

With $\mathbb{N}_0 := \{0, 1, 2, ...\}$, the $\mathbb{N}_0-$valued rv $K = K^+ - 1$, with pgf ($\alpha > 0$),

$$\varphi\,(z) := \mathbf{E}\left(z^K\right) = \frac{\alpha}{\alpha+1}\cdot F\,(1, 1; \alpha+2; z) = \frac{\varphi^+\,(z)}{z},$$

is in the class of three-parameters hypergeometric family of pgfs studied for instance in [4]. When $\alpha > 0$ and α is a rational number, K has the following Pólya-Eggenberger urn model interpretation: Take an urn with initially b_0 black balls and w_0 white balls. Balls are drawn at random one at a time from the urn and each selected ball is returned to the urn along with $r-1$ additional balls of the same color, $r \geq 2$. Repeat the sampling procedure. Suppose the number of balls which are returned

is $r = w_0$ and put $\alpha := b_0/r = b_0/w_0 > 0$. It then follows from the inverse sampling procedure described in ([4] page 290) and adapted to the parameter set of the shifted Yule-Simon(α) distribution that:

Proposition 2 — K represents the number of white balls that are drawn till the first black ball is selected in the inversed sampling process with $r = w_0$ returned balls.

PROOF : We have

$$\pi_k = \mathbf{P}(K = k) = \pi^+_{k+1} = \alpha B(k+1, \alpha+1), k \geq 0.$$

It can be checked that π_k obeys the recursion: $\pi_0 = \alpha/(\alpha+1)$ and

$$\pi_{k+1}/\pi_k = (k+1)/(k+\alpha+2), \ k \geq 0.$$

If, in the urn model, a black ball is first drawn, an event with probability $\pi_0 = b_0/(b_0+w_0) = \alpha/(\alpha+1)$, then $K = 0$; if a white ball is first drawn, followed by a black ball, then, as required from the recursion on π_k, $K = 1$ with probability

$$\begin{aligned} \pi_1 &= \left(1 - \frac{\alpha}{\alpha+1}\right) \frac{b_0}{b_0+w_0+r} = \frac{1}{1+\alpha}\frac{b_0}{b_0+2r} \\ &= \frac{1}{\alpha+1}\frac{\alpha}{\alpha+2} = \frac{1}{\alpha+2}\pi_0. \end{aligned}$$

If two white balls are first drawn, followed by a black ball, then $K = 2$ with the required probability

$$\begin{aligned} \pi_2 &= \frac{w_0}{b_0+w_0}\frac{w_0+r}{b_0+w_0+r}\frac{b_0}{b_0+w_0+2r} \\ &= \frac{r}{b_0+r}\frac{2r}{b_0+2r}\frac{b_0}{b_0+3r} \\ &= \frac{1}{\alpha+1}\frac{2}{\alpha+2}\frac{\alpha}{\alpha+3} = \frac{2}{\alpha+3}\pi_1. \end{aligned}$$

This reasoning can be extended by recurrence to yield the correct $\pi_k = \alpha B(k+1, \alpha+1)$. □

The sequence $\pi^+_k = \mathbf{P}(K^+ = k) = \alpha B(k, \alpha+1)$, $k \geq 1$, is the Yule-Simon probability mass function (pmf) on $\mathbb{N}$, as from (1). The distribution of $K = K^+ - 1$ is thus the distribution of a shifted Yule-Simon distribution with $\pi_k = \pi^+_{k+1}$. It is heavy-tailed with tail index α and $\alpha > 0$ is any rational number.

The heavy-tailed character of the distribution of K results from the fact that if a black ball is not drawn in the very first steps, there are many white balls in the urn due to previous white ball returns,

lowering the chance to subsequently draw a black ball in the future steps. This reinforcement property is responsible for K to take on very large values (or very small ones).

Remarks : In the latter Pólya-Eggenberger urn model, a continuum for the set of values of α can of course be achieved if the initial proportions of black and white balls both tend to ∞ with proper prescribed ratio: $b_0,\, w_0 \to \infty$ while $b_0 = [w_0\alpha]$. Note also that the latter Pólya-Eggenberger urn model is somehow special in the sense that the number of balls returned to the urn after sampling is related (here $= w_0$) to the initial composition of the urn. $\diamondsuit$

2.5 *Shifted Simon rv*

The pgf of the shifted Yule-Simon(α) rv K is $\varphi^{+}(z)/z$ and thus it is in the class of hypergeometric pgfs studied in [4]. As a result, with $G(\alpha) \sim$gamma$(\alpha, 1)$, and $G(1)$, $G'(1)$, $G(\alpha)$ mutually independent Gamma rvs (two of which are exponentially distributed), we have the Poisson mixture representation [6, 7, 23] of the Simon rv, say K^{+}, with $\mathbf{E}\left(z^{K^{+}}\right) = \varphi^{+}(z)$, as

$$K^{+} \overset{d}{=} 1 + \text{Poi}\left(\frac{G(1)\, G'(1)}{G(\alpha)}\right). \tag{6}$$

Therefore (see Eqs (16), (17) below for the definition of discrete-self-decomposability and Proposition 11 of [15]):

Proposition 3 — The shifted Simon(α) rv $K := K^{+} - 1$ is discrete-self-decomposable with mode at the origin.

As a SD rv, K is the limit law of a continuous-time branching process with immigration (BPI) where all the immigrants are subject, upon appearance, to a pure death Greenwood process [10].

3. Related Models: Sibuya and Extended Sibuya Distributions

3.1 *Extended Sibuya rv*

Recently, in [15], an extended Sibuya random variable with two parameters was introduced. With $\nu > -1$ and $0 < \alpha < \nu + 1$, consider the random variable with support $\mathbb{N}$ defined as follows:

$$K^{+} = \inf\left(l \geq 1 : \mathcal{B}_{\alpha,\nu}(l) = 1\right). \tag{7}$$

Here $(\mathcal{B}_{\alpha,\nu}(l);\ l \geq 1)$ is a sequence of independent Bernoulli rvs now obeying $\mathbf{P}\left(\mathcal{B}_{\alpha,\nu}(l) = 1\right) = \alpha/(\nu + l)$. And ν is a 'base' parameter. The Yule-Simon model of Eq. (5) is recovered when $\nu = \alpha$. With $G(1)$, $G(\nu + 1 - \alpha)$, $G(\alpha)$ mutually independent and Gamma distributed with corresponding

parameters, K^+ admits the Poisson mixture representation

$$K^+ \stackrel{d}{=} 1 + \text{Poi}\left(\frac{G(1)\,G(\nu+1-\alpha)}{G(\alpha)}\right). \tag{8}$$

Its pgf now is

$$\varphi^+(z) := \mathbf{E}\left(z^{K^+}\right) = \frac{\alpha}{\nu+1} z \cdot F(1, \nu+1-\alpha; \nu+2; z). \tag{9}$$

We note that, consistently with Eq. (7),

$$\mathbf{P}\left(K^+ = k\right) =: \pi_k^+ = \frac{\alpha}{\nu+k} \prod_{l=1}^{k-1} \left(1 - \frac{\alpha}{\nu+l}\right), k \in \mathbb{N}. \tag{10}$$

Some properties of the extended Sibuya rv are:

$(\mathbf{P}1)$ It has heavy tails with tail index $0 < \alpha < \nu+1$.

$(\mathbf{P}2)$ Using Stirling formula: $\pi_k^+ \underset{k\to\infty}{\sim} \alpha\Gamma(\nu+1)\, k^{-(\alpha+1)}/\Gamma(\nu+1-\alpha)$ and for all $k \geq 1$

$$\pi_{k+1}^+/\pi_k^+ = (\nu-\alpha+k)/(\nu+k+1) < 1, \pi_1^+ = \alpha/(\nu+1) \tag{11}$$

(π_k^+ is monotone decreasing). As a result, π_1^+ is the maximal value of the π_k^+: the law of K^+ has its mode at $k=1$.

$(\mathbf{P}3)$ When shifting K^+, the new random variable $K = K^+ - 1$, we get

$$\begin{array}{l} \mathbf{P}(K=k) =: \pi_k = \pi_{k+1}^+ = \frac{\alpha}{\nu+k+1} \prod_{l=1}^{k} \left(1 - \frac{\alpha}{\nu+l}\right), k \in \mathbb{N}_0, \text{ obeying} \\ \pi_{k+1}/\pi_k = (\nu-\alpha+k+1)/(\nu+k+2),\ k \geq 0 \text{ and } \pi_0 = \alpha/(\nu+1). \end{array} \tag{12}$$

$(\mathbf{P}4)$ The rv K is SD as a result of its Poisson-mixture representation (see also Proposition 11 of [15], for a different proof). From Eq. (9), the pgf of K is

$$\varphi(z) := \mathbf{E}\left(z^K\right) = \frac{\alpha}{\nu+1} \cdot F(1, \nu+1-\alpha; \nu+2; z), \tag{13}$$

with finite mean $\mathbf{E}(K) = \varphi'(1) = \frac{\nu+1-\alpha}{\alpha-1}$ only if $\nu+1 > \alpha > 1$ (note that $\mathbf{E}(K) < 1$ if and only if $\alpha > 1 + \nu/2$).

The latter pgf (13) is thus also in the class of the three-parameters Gauss hypergeometric family of pgfs studied in [4]. When $\nu+1$ is the reciprocal of an integer or when $\nu+1$ is an integer dividing w_0+b_0, K also has an alternative Pólya-Eggenberger urn model interpretation, with initially b_0 black

balls and w_0 white balls. Balls are iteratively drawn at random, one at a time, from the urn and each selected ball is returned to the urn along with $r-1$ additional balls of the same color, $r \geq 2$. The number of balls which are returned now is $r = (w_0+b_0)/(\nu+1)$ and we put $\alpha := b_0/r$. Note $0 < \alpha = (\nu+1)(1 - w_0/(w_0+b_0)) < \nu+1$. We get:

Proposition 4 — K with distribution (12), represents the number of white balls that are drawn till the first black ball is selected in the above inverse sampling process with $r = (w_0+b_0)/(\nu+1)$ returned balls.

PROOF : The proof proceeds by recurrence (now using Eq. (12)), following the steps of Proposition 2. Proposition 2 is obtained when $\nu = \alpha$. □

$(\mathbf{P}5)$ The Yule-Simon distribution is obtained when $\nu = \alpha$ with $0 < \alpha < \nu+1 = \alpha+1$, so involving no restriction on the tail index $\alpha > 0$.

$(\mathbf{P}6)$ When $\alpha = 1$, K is defined for all $\nu > 0$. Imposing $\alpha = 1$ in the urn model, $r = b_0$ and the condition on ν is $\nu = w_0/b_0$. In this case, $\mathbf{E}\left(z^K\right) = \frac{1}{\nu+1} \cdot F(1,\nu;\nu+2;z)$ with

$$\mathbf{P}(K=k) =: \pi_k = \nu\left[\frac{1}{\nu+k} - \frac{1}{\nu+k+1}\right], k \in \mathbb{N}_0$$

with $\pi_k/\pi_{k-1} = (\nu+k-1)/(\nu+k+1) < 1$, $\pi_0 = 1/(\nu+1)$. Such a rv K is heavy-tailed with tail index 1 and with mode at the origin.

$(\mathbf{P}7)$ Conditioning the general Sibuya rv K, with parameters α, ν, on the event $K \geq 1$ gives rise to:

$$\begin{aligned}\mathbf{P}(K=k \mid K \geq 1) &= \frac{1}{1-\frac{\alpha}{\nu+1}} \frac{\alpha}{\nu+k+1} \prod_{l=1}^{k}\left(1-\frac{\alpha}{\nu+l}\right) \\ &= \frac{\alpha}{\nu+1+k} \prod_{l=1}^{k-1}\left(1-\frac{\alpha}{\nu+1+l}\right), k \geq 1.\end{aligned}$$

From Eq. (10), it is the probability mass function of a rv K^+ with new tail and base parameters $\alpha, \nu+1$.

3.2 *Sibuya distribution, [23].*

The case $\nu = 0$ deserves some special interest so with the restriction $\alpha < 1$. In this case,

$$K^+ = \inf\left(l \geq 1 : \mathcal{B}_\alpha(l) = 1\right), \tag{14}$$

where $(\mathcal{B}_\alpha(l);\ l\geq 1)$ is a sequence of independent Bernoulli rvs now obeying $\mathbf{P}(\mathcal{B}_\alpha(l)=1)=\alpha/l$. We have

$$K^+\overset{d}{=}1+\text{Poi}\left(\frac{G(1)\,G(1-\alpha)}{G(\alpha)}\right)$$

and

$$\varphi^+(z):=\mathbf{E}\left(z^{K^+}\right)=\alpha z\cdot F(1,1-\alpha;2;z)=1-(1-z)^\alpha$$

defining a Sibuya rv with $\mathbf{P}(K^+=k)=:\pi_k^+=\frac{\alpha}{k}\prod_{l=1}^{k-1}\left(1-\frac{\alpha}{l}\right)=\alpha\frac{[1-\alpha]_{k-1}}{k!}$, $k\in\mathbb{N}$.

Defining $K=K^+-1$, K also has an alternative Pólya-Eggenberger urn model interpretation where the number of balls which are returned in the urn is $r=w_0+b_0$ and $\alpha:=b_0/r=b_0/(w_0+b_0)\in(0,1)$. Using the recursion property of the pmf of K given in $(\mathbf{P}2)$ below, it can easily be checked, following the path of Proposition 2, that:

Proposition 5 — K represents the number of white balls that are drawn till the first black ball is selected in the sampling process with $r=w_0+b_0$ returned balls.

Note that

$$\varphi(z):=\mathbf{E}\left(z^K\right)=\alpha\cdot F(1,1-\alpha;2;z)=\frac{1}{z}\left(1-(1-z)^\alpha\right).$$

Some properties of the Sibuya rv are:

$(\mathbf{P}1)$ It has heavy tails with index $\alpha<1$. Using Stirling formula: $\pi_k^+\underset{k\to\infty}{\sim}\alpha k^{-(\alpha+1)}/\Gamma(1-\alpha)$ and $\pi_{k+1}^+/\pi_k^+=(k-\alpha)/(k+1)<1$, $\pi_1^+=\alpha$ (π_k^+ is monotone decreasing).

$(\mathbf{P}2)$ As a Poisson mixture, the rv K is SD. Its pmf $\mathbf{P}(K=k):=\pi_k=\pi_{k+1}^+$ obeys the recursion: $\pi_{k+1}/\pi_k=(k+1-\alpha)/(k+2)$.

$(\mathbf{P}3)$ Scale-free property: With $(B_l(u))_{l\geq 1}$ an independent and identically distributed (iid) sequence of Bernoulli rvs with success probability $u\in(0,1)$, let $u\circ K^+=\sum_{l=1}^{K^+}B_l(u)$ denote the Bernoulli(u)-thinning of K^+. One can easily check (while computing the pgfs of the rvs of both sides) that:

Proposition 6 — K^+ obeys the scale-free property:

$$\forall u\in(0;1):u\circ K^+\mid\left(u\circ K^+\geq 1\right)\overset{d}{=}K^+,\tag{15}$$

as a fixed point of a transformation involving thinning and conditioning, [1].

$(\mathbf{P}4)$ The Bernoulli(u) thinned version of K^+ is $u\circ K^+$. Its pgf is $\varphi_u(z):=\mathbf{E}\left(z^{u\circ K^+}\right)=1-u^\alpha(1-z)^\alpha$ and it is stable under composition [22], viz.,

$$\varphi_u(\varphi_u(z))=1-u^{\alpha+\alpha^2}(1-z)^{\alpha^2}.$$

See [12] for the use of this remarkable property in the context of branching processes [11].

4. On the Consequences of Being SD

We here briefly address the notions of discrete self-decomposability which have been met previously.

4.1 *Discrete additive self-decomposability*

Let now $K \geq 0$ be an integer-valued random variable. There exists a discrete version of the notion of self-decomposability [28].

Definition 1 — The pgf $\varphi(z) := \mathbf{E}z^K$ is the one of a discrete self-decomposable variable $K \geq 0$ if for any $u \in (0,1)$, there is a pgf $\varphi_u(z)$ (depending on u) such that

$$\varphi(z) = \varphi(1 - u(1-z)) \cdot \varphi_u(z) \tag{16}$$

This is the standard (discrete) version of self-decomposability of probability distributions on the integers, through a functional equation. It follows from the definition of self-decomposable distributions that if $\varphi(z)$ is the pgf of the random variable K, then K can be additively decomposed as

$$K \stackrel{d}{=} u \circ K' + K_u, \tag{17}$$

where the u-thinned random variable $u \circ K$, for $u \in (0,1]$, is defined above. K and K' have the same distribution and $u \circ K'$ is independent of the remaining random variable K_u whose pgf is $\varphi_u(z)$. Consequently, we clearly have

$$\sum_{m=0}^{n} K_m \underset{n \to +\infty}{\stackrel{d}{\to}} K,$$

where $K_m \stackrel{d}{=} u^m \circ K_u$ are independent random variables with pgfs: $\varphi_{K_m}(z) = \varphi_{K_u}(1 - u^m(1-z))$. Discrete self-decomposable random variables are thus obtained as limits in law for sums of independent scaled discrete random variables. A slightly different way to see this is as follows. Consider the discrete-time integral-valued Ornstein-Uhlenbeck process

$$K(n+1) = u \circ K(n) + K_u(n+1), K(0) \text{ random},$$

where $(K_u(n); n \geq 1)$ is an iid driving sequence, each distributed like K_u. Then $K(n)$ is a discrete perpetuity [30], clearly with

$$K(n) = u^n \circ K(0) + \sum_{m=1}^{n} u^{n-m} \circ K_u(m) \stackrel{d}{\to} K \stackrel{d}{=} \sum_{m=0}^{\infty} K_m \text{ as } n \to \infty.$$

In such models typically, a population whose fate is to die out predictably at shrinking rate u is regenerated by the incoming of immigrants in random number K_u. Note that u appears in the law of the incoming immigrants, translating a subtle balance between shrinkage and regeneration of the population.

The following representation result is also known to hold true (see [22], Lemma 2.13): The random variable K with pgf $\varphi(z)$ is discrete self-decomposable if and only if, with $R(z)$ the canonical function, defined through

$$\varphi(z) = e^{-\int_z^1 R(z')dz'},$$

the function $h(z) := 1-(1-z)R(z)$ is absolutely monotone and $h(0) = 0$. So, K is discrete self-decomposable if and only, for such a pgf h, its pgf is of the form

$$\varphi(z) = e^{-\int_z^1 \frac{1-h(z')}{1-z'}dz'}. \tag{18}$$

This also means that the coefficients $r_k := \left[z^k\right] R(z)$, $k \geq 1$, constitute a non-increasing sequence of k, ([27], Theorem 4.13). As a result, the associated probability system $\mathbf{P}(K=k) := p_k$, $k \geq 0$ is unimodal with mode at the origin if and only if $r_0 = \left[z^0\right] R(z) = \frac{p_1}{p_0} \leq 1$, (see [27], Theorem 4.20). The self-decomposable (SD) subclass of infinitely divisible distributions therefore focuses on unimodal distributions, with mode at the origin if $p_1 < p_0$, or with two modes at $\{0,1\}$ if $r_0 = 1$.

4.2 *Self-decomposable rvs and pure-death branching processes with immigration, [29], [21].*

We now emphasize that the self-decomposable models just discussed (when they have mode(s) by the origin) have an equivalent interpretation in terms of a continuous-time branching process with immigration producing special self-decomposable rvs as time goes to ∞.

Consider a continuous-time homogeneous compound Poisson process $P_\lambda(t)$, $t \geq 0$, $P_\lambda(0) = 0$, so with pgf

$$\mathbf{E}_0\left(z^{P_\lambda(t)}\right) = \exp -\lambda t\left(1-h(z)\right),$$

where $h(z)$ (with $h(0) = 0$) is the pgf of the sizes of the batches arriving at the jump times of $P_\lambda(t)$ having rate $\lambda > 0$. Let now

$$\phi_t(z) = 1 - e^{-\mu t}(1-z),$$

be the pgf of a pure-death branching process at rate $\mu > 0$ started with one particle at $t = 0$. The lifetime distribution of the initial particle is thus $1-e^{-\mu t}$. Let K_t with $K_0 = 0$ be a random process counting the current size of some population for which a random number of individuals (the

law of which is determined by $h(z)$) immigrate at the jump times of $P_\lambda(t)$, each of which being independently and immediately subject to the latter pure death process. We have

$$\varphi_t(z) := \mathbf{E}\left(z^{K_t}\right) = \exp -\lambda \int_0^t (1 - h(\phi_s(z)))\, ds,\ \varphi_0(z) = 1,$$

with $\varphi_t(0) = \mathbf{P}(K_t = 0) = \exp -\lambda \int_0^t (1 - h(1 - e^{-s}))\, ds$, the probability that the whole population is extinct at t. As $t \to \infty$,

$$\varphi_t(z) \to \varphi(z) = e^{-\lambda \int_0^\infty \left(1 - h\left(1 - e^{-\mu s}(1-z)\right)\right) ds} = e^{-\frac{\lambda}{\mu} \int_z^1 \frac{1-h(x)}{1-x} dx}. \tag{19}$$

So, $K := K_\infty$, as the limiting population size of this pure-death process with immigration, is a self-decomposable rv because $R(z) = h(z)$ is absolutely monotone, [29]. With $K \overset{d}{=} u \circ K' + K_u$ defining the rv K_u, we have that, for all u,

$$\varphi_{K_u}(z) = \frac{\varphi(z)}{\varphi(1 - u(1-z))} = e^{-\frac{\lambda}{\mu} \int_z^{1-u(1-z)} \frac{1-h(v)}{1-v} dv} \text{ is a pgf.}$$

In such models typically, a subcritical population whose fate is to die out (along the stochastic pure-death process striking each individual alive) is permanently regenerated by the incoming of immigrants at random times and in random number. The immigrants come from an external reservoir that feeds the regularly declining population. Both in discrete or continuous time therefore, the occurrence of a self-decomposable limit law is related to a trade-off between a pure-death mechanism which tends to shrink the population size at a constant rate μ, against an input immigration mechanism at rate λ which tends to have it increased.

Examples pertaining to Sibuya rvs:

Let us now identify the pgf $h(z)$ of the incoming immigrants leading to $\varphi(z) = \frac{\alpha}{\nu+1} \cdot F(1, \nu + 1 - \alpha; \nu + 2; z)$, known to be the pgf of the SD extended-Sibuya random variable K. It is

$$h(z) = 1 - \frac{\mu}{\lambda}(1 - z)\left(\log F(1, \nu + 1 - \alpha; \nu + 2; z)\right)', \tag{20}$$

where $'$ denotes the derivative with respect to z. Furthermore, $h(0) = 0$ entails

$$\left(\log F(1, \nu + 1 - \alpha; \nu + 2; 0)\right)' = \frac{\pi_2^+}{\pi_1^+} = \frac{\nu + 1}{\nu + 2}\left(1 - \frac{\alpha}{\nu + 1}\right) = \lambda/\mu < 1,$$

showing how the birth rate λ should be related to the death rate μ (note $\lambda < \mu$) to produce the desired asymptotic distribution of K. This way to generate the extended Sibuya distribution is of a completely different nature than the Yule approach to Simon model. By construction, $h(z)$ defined in Eq. (20) is a pgf.

If $\nu = 0$ (Sibuya), $\varphi(z) = \alpha \cdot F(1, 1-\alpha; 2; z) = z^{-1}\left(1-(1-z)^{\alpha}\right)$ .The pgf of the number of incoming immigrants reduces to

$$h(z) = 1 - \frac{2}{1-\alpha}(1-z)\left(-\frac{1}{z} + \frac{\alpha(1-z)^{\alpha-1}}{1-(1-z)^{\alpha}}\right),$$

with $h(z) \underset{z\to 1}{\sim} 1 - \frac{2\alpha}{1-\alpha}(1-z)^{\alpha}$. By Tauberian theorem, the distribution of the number of incoming immigrants is itself heavy-tailed of index α with mass in the tails given by: $q_k := \left[z^k\right] h(z) \underset{k\to\infty}{\sim} \frac{2\alpha}{1-\alpha} k^{-(\alpha+1)}/\Gamma(1-\alpha)$. ◇

5. A Special Markov Chain with Catastrophe on $\mathbb{N}_0$

We finally discuss a fundamental discrete-time Markov chain intimately related to the extended Sibuya rv.

5.1 *The catastrophe model*

With $\nu > -1$, let $\nu + 1 > \alpha > 0$. Consider the following discrete-time homogeneous Markov chain $(X_n; n \geq 0)$ with state-space $\mathbb{N}_0 = \{0, 1, ...\}$ and transition probabilities given by:

- given $X_n = k \in \{1, 2, ...\}$, the increment of X_n is

$$\begin{array}{l} +1 \text{ with probability : } p_k = 1 - \alpha/(\nu + k) \\ -k \text{ with probability : } q_k = \alpha/(\nu + k). \end{array} \tag{21}$$

- given $X_n = 0$, the increment of X_n is $+1$ with probability p_0 or 0 with probability $q_0 = 1 - p_0$.

In other words, with $(U_n, n \geq 1)$ a sequence of independent identically distributed (i.i.d.) uniform random variables, the dynamics of X_n reads

$$X_{n+1} = (X_n + 1)\,\mathbf{1}(U_{n+1} > q_{X_n}).$$

This irreducible Markov chain is in the class of general Markov chains with catastrophes (say MCC), state-dependent transition probabilities and no holding probability (the probability to remain in state $k \neq 0$ is 0). If $q_0 > 0$ however, there is a holding probability q_0 at the origin whereas if $q_0 = 0$ ($p_0 = 1$), the chain is instantaneously reflected at the origin. Its detailed study was motivated by Eq. (5). In this model, the walker X_n is occasionally bounced back to the origin. When k is large, the drift of this MCC is of order $1 - \alpha - \alpha(1-\nu)/k$. So when $\alpha > 1$, the walker is attracted to the origin: the strength of the attraction goes like $\alpha - 1$ for large k, to leading order. For $\alpha < 1$, the walker is asymptotically repelled from the origin correspondingly. When $\alpha = 1$, its drift is of order

$-(1-\nu)/k$, and the weak drift that the walker feels vanishes when k approaches ∞. This vanishing drift is attractive when $0 < \nu < 1$, repulsive when $\nu > 1$ and null if $\nu = 1$. We will see that the chain is always recurrent whereas, while crossing the critical value $\alpha = 1$ from above, the process X_n switches from positive recurrent to null-recurrent. There is no transience situation here whereby X_n could drift to infinity with positive probability.

In the context of machine replacement in reliability theory, one may classically interpret this MCC as follows: at time $n = 0$, a machine is put into service. This machine has a (discrete) random lifetime, say $\tau_{0,0} > 0$. After $\tau_{0,0}$, a new machine (with lifetime a copy of $\tau_{0,0}$) is installed to replace the old defective one and so on. The failure epochs of the successive machines constitute a renewal process on $\mathbb{N}_0$ generated by $\tau_{0,0}$. In this context, when $p_0 = 1$, X_n clearly represents the age of the machine currently in action (the time till the last machine failed before it was instantaneously replaced by the new one in action at time n). Given the age of the current machine is k, there is an age-dependent probability p_k that the machine will survive one more time unit and a probability q_k that it will not. Assuming $q_k = \alpha/(\nu + k)$ means that a machine that has already survived for long is robust and so very unlikely to fail in a near future (a reinforcement property). Clearly, the law of $\tau_{0,0}$ is $\mathbf{P}(\tau_{0,0} = k+1) = q_k \prod_{l=0}^{k-1} p_l$, $k \geq 1$, in terms of the ps and the $\tau_{0,0}$s are seen to be the times between consecutive visits to 0 for X_n. While assuming $p_0 = 1$, we suppose that the new machine is put into service immediately as soon as the defective one fails (no latency waiting time).

Suppose $X_n \neq 0$. The age τ_n^B of a machine at n is the backward recurrence time of the underlying renewal process (or the time separating n from the previous catastrophic here failure event); of interest also is the forward recurrence time τ_n^F at n (or the time separating n from the forthcoming catastrophic event, equivalently the remaining lifetime of the machine). It will be seen to be related to the time-reversed version $\overleftarrow{X}_n$ of X_n.

One may also view this MCC as the following growth-collapse model: X_n represents the size of some growing population at time n. Given $X_n = k$, at the next time unit, a new individual presents itself for possible integration to the herd. The population can then grow by one unit with size-dependent probability p_k, integrating normally the new element, so: $X_n = k \to X_{n+1} = k+1$ with probability p_k. But there is a 'chance' q_k that the new element is a black sheep at the contact of which the whole population get decimated, so: $X_n = k \to X_{n+1} = 0$ with probability q_k; the black sheep is then the only survivor in the next generation and serves as a founder of a new growing population itself facing subsequent catastrophic events. Clearly the occurrence of a black sheep is a catastrophic renewal event. The contamination (or collapse) probabilities q_k may be either (i) a decreasing or (ii) an increasing sequence of the current population size k. In the former case (i),

large populations are getting more and more immune to black sheep and one expects large population sizes to stabilize. This is the case under study here with $q_k = \alpha/(\nu + k)$. We will see that such a MCC X_n is always recurrent, with heavy-tailed limit law when $\alpha > 1$ and no transience possible where X_n could even drift to infinity. In case (not studied here) where $q_k \underset{k\to\infty}{\sim} \alpha k^{-\beta}$ ($\alpha, \beta > 0$), we would check that X_n is transient if $\beta > 1$, positive recurrent if $\beta < 1$, whereas if $\beta = 1$ as in Eq. (21), X_n is positive recurrent if $\alpha > 1$, null recurrent if $\alpha \leq 1$. In the latter case (ii), not studied here either, large populations are more susceptible and vulnerable to black sheep and so quite unlikely to grow too large and develop much. This would be the case for a 'dual' model with homographic collapse probabilities $q_k = 1 - \alpha/(\nu + k)$, $0 < \alpha < \nu + 1$ obtained while switching the roles of (p_k, q_k) in Eq. (21). Similar 'switching probabilities' ideas arise in the context of birth and death chains under the name of Wall duality, [5]. Such a MCC is always positive recurrent and it now has a light-tailed limiting population size. Think of a forest fire occasionally devastating completely an otherwise regularly growing forest: the largest the size of this forest, the more likely it is that it will face a forest fire, a feedback mechanism limiting the growth of its final size.

Consider then a general catastrophe process X_n possibly reflected at the origin ($p_0 = 1$ and $q_0 = 0$) for which both p_k and $q_k > 0$, for all $k \geq 1$, with $p_k + q_k = 1$ and so with associated irreducible and aperiodic stochastic transition matrix: $P = [P(k, k')]$, $(k, k') \in \mathbb{N}^2$ with $P(k, 0) = q_k$ and $P(k, k+1) = p_k$, $k \geq 1$. Indeed $\gcd(n \geq 1 : P^n(k, k) > 0) = 1$ for each state $k \in \mathbb{N}_0$ and each state is accessible from any other state.

5.2 *Existence and shape of the invariant measure*

We first proceed with a general study of a catastrophic Markov chain before particularizing it to the present MCC case under study.

Let $\mu' := (\mu_0, \mu_1, ..)$ be the row-vector of the invariant measure, whenever it exists. Then μ should solve $\mu' = \mu'\mathbf{P}$, whose formal solution is:

$$\mu_0 = \sum_{k\geq 0} \mu_k q_k \text{ and } \mu_k = \mu_0 \prod_{l=0}^{k-1} p_l,\ k \geq 1. \tag{22}$$

Let $u_k = \prod_{l=0}^{k-1} p_l$. Using the second equation, the first equation is satisfied whenever $\sum_{k\geq 1} q_k \prod_{l=0}^{k-1} p_l = \sum_{k\geq 1}(u_k - u_{k+1}) = p_0$, so also when $u_\infty = \prod_{l=0}^{\infty} p_l = 0$ which is fulfilled if and only if $S_1 := \sum_{l=0}^{\infty} q_l = \infty$.

We first conclude that there exists an invariant measure if and only if $S_1 = \infty$.

- When $S_1 < \infty$ (no non trivial invariant measure), all states $k \geq 0$ are transient. In this case,

the number of returns to state k when started at k is $< \infty$, $\mathbf{P}_k-$almost surely and geometrically distributed with success parameter $1-\alpha_k$ where $\alpha_k = \mathbf{P}(\tau_{k,k} < \infty) < 1$ is the probability that the first return time to state k, say $\tau_{k,k} = \inf(n \geq 1 : X_n = k \mid X_0 = k)$, is finite.

- When $S_1 = \infty$, the chain is (Harris) recurrent: when started at state k, it first hits 0 with probability 1 and returns infinitely often to 0. With $N_{k,l} := \sum_{n\geq 0} \mathbf{1}(X_n = l \mid X_0 = k)$ the number of visits to state l when started at state k, then $N_{k,l} = \infty$, $\mathbf{P}_k-$almost surely. And $\alpha_k := \mathbf{P}(\tau_{k,k} < \infty) = 1$.

Furthermore, with $\mathcal{N}_{k,l} := \sum_{n=0}^{\tau_{k,k}-1} \mathbf{1}(X_n = l)$ the number of visits to state l before the first return time to state k (between consecutive visits to state k), then: $\mathbf{E}(\mathcal{N}_{k,l}) = \frac{\mu_l}{\mu_k}$ and, by the Chacon-Ornstein limit ratio ergodic theorem [17]:

$$\frac{\frac{1}{N}\sum_{n=0}^{N} \mathbf{1}(X_n = l \mid X_0 = i)}{\frac{1}{N}\sum_{n=0}^{N} \mathbf{1}(X_n = k \mid X_0 = i)} \underset{N\to\infty}{\to} \frac{\mu_l}{\mu_k}, \ \mathbf{P}_i - \text{almost surely.} \tag{23}$$

Starting in particular from $k = 0$, a recurrent chain is made of infinitely many iid excursions which are the sample paths of $(X_n; n \geq 0)$ between consecutive visits to state 0. We have: $\mathbf{E}(\mathcal{N}_{0,k}) = \frac{\mu_k}{\mu_0} = \prod_{l=0}^{k-1} p_l$.

If in addition, $S_2 := \sum_{k\geq 1} \prod_{l=1}^{k} p_l < \infty$, then $\mu_0 = \frac{1}{1+p_0 S_2} \in (0,1)$ and the invariant measure is unique and it is a proper invariant probability measure. In this case, $X_n \overset{d}{\to} X_\infty$ (as $n \to \infty$) and, with the empty product being 1, we have

$$\mu_k := \mathbf{P}(X_\infty = k) = \frac{\prod_{l=0}^{k-1} p_l}{1 + p_0 S_2}, \ k \geq 0. \tag{24}$$

By the ergodic theorem $\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}(X_n = k) \underset{N\to\infty}{\to} \frac{1}{\mathbf{E}\tau_{\mathbf{k},\mathbf{k}}} = \mu_k$, so with $\mathbf{E}\tau_{k,k} < \infty$.

When the chain is positive recurrent ($S_2 < \infty$) the expected time elapsed between consecutive visits to 0 is finite and equal (by Kac theorem, [14]) to $\mathbf{E}(\tau_{0,0}) := \lambda = 1/\mu_0 = 1 + p_0 S_2$, whereas this expected time is infinite when the chain is null recurrent.

When $S_2 = \infty$, the measure solution to (22) exists up to an arbitrary multiplicative constant μ_0 but it is not a probability measure as its total mass $\mu_0(1 + p_0 S_2)$ sums to infinity. And only the ratios $\frac{\mu_l}{\mu_k}$ are well-defined.

In this case, although $\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}(X_n = k \mid X_0 = i) \underset{N\to\infty}{\to} 0$, $\mathbf{P}_i-$almost surely, (23) holds. The expected time elapsed between consecutive visits to 0 is infinite, favoring excursions with very large heights.

Let us particularize these facts to our special MCC.

Proposition 7 — For the special MCC under study (21): $S_1 := \sum_{l=0}^{\infty} q_l = q_0 + \sum_{l=1}^{\infty} \frac{\alpha}{\nu+l} = \infty$. This MCC is always recurrent with invariant measure

$$\begin{array}{l} \mu_k = \mu_0 p_0 \prod_{l=1}^{k-1} \left(1 - \frac{\alpha}{\nu+l}\right) \underset{k\to\infty}{\sim} \mu_0 p_0 k^{-\alpha} \text{ if } \alpha \neq 1 \\ \mu_k = \mu_0 p_0 \prod_{l=1}^{k-1} \left(\frac{\nu+l-1}{\nu+l}\right) = \mu_0 p_0 \frac{\nu}{\nu+k-1} \text{ if } \alpha = 1. \end{array} \tag{25}$$

The special MCC is recurrent positive if and only if $\alpha > 1$ ($S_2 < \infty$), null-recurrent if and only if $\alpha \leq 1$ ($S_2 = \infty$) for all $\nu > -1$. When $\alpha > 1$, the MCC is pinned at the origin in that there exists a limiting positive contact fraction μ_0 at 0. Furthermore, the limiting probability mass function of X_n, which is $\mu = (\mu_k, k \geq 0)$, exists and is heavy-tailed with index $\alpha - 1$ (translating a large asymptotic family size). The critical value $\alpha = 1$ separating the null recurrent phase from the positive recurrent one is a depinning/pinning transition point: when $\alpha < 1$, X_n drifts linearly to infinity on average, as $n \to \infty$, whereas if $\alpha > 1$, X_n has a weak limit (with probability mass μ_k).

PROOF : Only the last statement concerning the situation $\alpha < 1$ deserves a special comment. The result follows by Dynkin-Lamperti theorem; see Theorem 8.7.3 of [2]. The depinning/pinning image is borrowed to the theory of interfaces, viewing the sample paths of X_n as an interface above its base line, [3]. □

5.3 *Return time to the origin*

Clearly of interest also are the times $\tau_{0,0}$ between consecutive visits to 0 (the first return times to 0) because if $X_0 = 0$, X_n represents the backward recurrence time (the time separating time n to the previous visit to 0) of a discrete renewal process generated by $\tau_{0,0}$. We have

$$\mathbf{P}(\tau_{0,0} = k+1) = q_k \prod_{l=0}^{k-1} p_l = u_k - u_{k+1}, k \geq 1, \tag{26}$$

which is also $\mathbf{P}(\tau_{0,0} > k) = \prod_{l=0}^{k-1} p_l = u_k$.

For the MCC under study, if $p_0 = 1$ (the MCC is instantaneously reflected at the origin), then $\mathbf{P}(\tau_{0,0} = k+1) = \frac{\alpha}{\nu+k} \prod_{l=1}^{k-1} \left(1 - \frac{\alpha}{\nu+l}\right)$, $k \geq 1$. We recognize q_k^+ on the right hand-side, as from Eq. (10). So $\tau_{0,0} \overset{d}{=} K^+ + 1$ is seen to be the interarrival time of a renewal sequence on the integers $\mathbb{N}$ corresponding to the successive returns to 0 of X_n given $X_0 = 0$. Moreover for this particular MCC with $p_0 = 1$, the distributions of the length and height of each excursion are respectively the ones of $(K^+ + 1, K^+)$. Because $S_1 = \infty$, X_n is recurrent.

• If $S_1 = \infty$ and $S_2 := \sum_{k\geq 1} \prod_{l=1}^{k} p_l < \infty$, X_n is positive recurrent with $\lambda := \mathbf{E}(\tau_{0,0}) = 1/\mu_0 = 1 + S_2 < \infty$. The mean λ of $\tau_{0,0} \overset{d}{=} K^+ + 1 \overset{d}{=} K + 2$ exists only when $1 < \alpha < \nu + 1$ and

it is

$$\lambda = 2 + \mathbf{E}(K) = 2 + \frac{\nu + 1 - \alpha}{\alpha - 1} = \frac{\nu + \alpha - 1}{\alpha - 1} > 2. \tag{27}$$

Proposition 8 — In the positive recurrent case with $\alpha > 1$ (note $\mu_0 = \mu_1$), the equilibrium probability mass function expresses as

$$\mu_k = \frac{\prod_{l=0}^{k-1} p_l}{1 + S_2} = \frac{\mathbf{P}(\tau_{0,0} > k)}{\lambda} = \frac{\alpha - 1}{\nu + \alpha - 1} \prod_{l=1}^{k-1} \left(1 - \frac{\alpha}{\nu + l}\right), k \geq 0. \tag{28}$$

It is the law μ of the limiting ($n \to \infty$) backward recurrence time τ_∞^B of the discrete renewal process generated by $\tau_{0,0} \geq 2$.

PROOF : Let τ_n be the length of the excursion to which instant n belongs whenever $X_n \neq 0$. Then $\tau_n \xrightarrow{d} \tau$ (convergence in distribution) as $n \to \infty$ with

$$\mathbf{P}(\tau = k) = \frac{k\mathbf{P}(\tau_{0,0} = k)}{\lambda},\ k \geq 2,$$

which is the size-biased version of the law of $\tau_{0,0}$. Note that τ is stochastically larger than $\tau_{0,0}$ (the waiting time paradox). If τ_∞^B (τ_∞^F) denotes the limiting ($n \to \infty$), backward (forward) recurrence time τ_n^B (τ_n^F) at time n of the discrete renewal process generated by $\tau_{0,0} \geq 2$, with $\tau_\infty^B + \tau_\infty^F = \tau$, we have $\tau_\infty^B \overset{d}{=} \tau_\infty^F$ (distributional equality) and (see [27], Lemma 5.9 for similar issues)

$$\left(\tau_\infty^B, \tau_\infty^F\right) \overset{d}{=} \left(U \circ (\tau - 1), (1 - U) \circ (\tau - 1)\right). \tag{29}$$

Note $\tau_\infty^B, \tau_\infty^F \geq 1$. Here, with U a uniform random variable independent of τ, $U \circ \tau$ is the $U-$thinning of τ : $U \circ \tau = \sum_{l=1}^{\tau} B_l(U)$ where, given U, $B_l(U); l \geq 1$ are mutually independent and independent of τ rvs with law Bernoulli(U) . Indeed,

$$\mathbf{P}(U \circ (\tau - 1) = k) = \int_0^1 \mathbf{P}(u \circ (\tau - 1) = k)\, du$$

where

$$\mathbf{P}(u \circ (\tau - 1) = k) = \sum_{l \geq k} \binom{l}{k} u^k (1 - u)^{l-k} \mathbf{P}(\tau = l + 1).$$

Thus, with $k \geq 1$

$$\begin{aligned} \mathbf{P}(U \circ (\tau - 1) = k) &= \sum_{l \geq k} \mathbf{P}(\tau = l + 1) \binom{l}{k} \int_0^1 u^k (1 - u)^{l-k}\, du \\ &= \sum_{l \geq k} \frac{1}{l + 1} \mathbf{P}(\tau = l + 1) = \frac{1}{\lambda} \mathbf{P}(\tau_{0,0} > k) = \mu_k. \ \square \end{aligned}$$

Remark : When $\alpha > 1$, $X_n \stackrel{d}{\to} X_\infty \stackrel{d}{\sim} \mu$ as $n \to \infty$. From (28), $\mu_{k+1}/\mu_k = (\nu - \alpha + k)/(\nu + k)$, $k \geq 1$ with $\mu_0 = \mu_1 = (\alpha - 1)/(\nu + \alpha - 1)$. Comparing with (11), we conclude that, the conditional law of X_∞ given $X_\infty \geq 1$ coincides with the one of an extended Sibuya rv of type K^+ but with parameters $(\alpha - 1, \nu - 1)$. Indeed,

$$\begin{array}{l} \mathbf{P}(X_\infty = k \mid X_\infty \geq 1) = \frac{\mu_k}{1-\mu_0} = \frac{\nu+\alpha-1}{\nu}\mu_k = \frac{\alpha-1}{\nu}\prod_{l=1}^{k-1}\left(1 - \frac{\alpha}{\nu+l}\right) \\ = \frac{\alpha-1}{\nu-1+k}\prod_{l=1}^{k-1}\left(1 - \frac{\alpha-1}{\nu-1+l}\right), k \in \mathbf{N}. \diamondsuit \end{array} \tag{30}$$

• If both $S_1 = S_2 = \infty$ ($\alpha \leq 1$), X_n is null recurrent with $\tau_{0,0} < \infty$ almost surely and $\mathbf{E}(\tau_{0,0}) = \infty$. Here indeed,

$$\begin{array}{l} \mathbf{P}(\tau_{0,0} > k) = \prod_{l=1}^{k-1}\left(1 - \frac{\alpha}{\nu+l}\right) \underset{k\to\infty}{\sim} \frac{1}{k^\alpha}, \text{ if } \alpha < 1, \\ \mathbf{P}(\tau_{0,0} > k) = \prod_{l=1}^{k-1}\left(1 - \frac{1}{\nu+l}\right) = \frac{\nu}{\nu+k-1} \underset{k\to\infty}{\sim} \frac{\nu}{k}, \text{ if } \alpha = 1. \end{array} \tag{31}$$

When $\alpha < 1$, as in this null recurrent case, we are left with a renewal process on the set $\mathbb{N}$ whose inter-arrival times are $\tau_{0,0}$, so with infinite mean. By Dynkin-Lamperti theorem (see [9] and [16]), τ_n has no weak limit as $n \to \infty$, but $n^{-1}\tau_n$ has. Similarly, with $\left(\tau_n^B, \tau_n^F\right)$ the backward and forward recurrence times, $n^{-1}\left(\tau_n^B, \tau_n^F\right)$ have a weak limit as $n \to \infty$, jointly and marginally. Recalling $\tau_n^B \stackrel{d}{=} X_n$, we conclude that $n^{-1}\mathbf{E}X_n$ converges to $1 - \alpha$ (see [2]). When $\alpha = 1$, $b_n^{-1}\mathbf{E}X_n$ converges for some increasing sequence b_n, that we conjecture to be $\log n$ without element of proof.

Let us summarize our results emphasizing that the MCC model X_n with transition probabilities (21) is intimately related to the family of Sibuya distributions:

Proposition 9 — In the positive recurrent case with $\alpha > 1$, X_n converges in distribution to a rv X_∞ whose invariant probability mass function is μ, given by Eq. (28). The limiting rv X_∞ of the MCC (21), reflected at the origin, can be interpreted as the backward recurrence time of a renewal process generated by the excursion length $\tau_{0,0}$, with $\tau_{0,0} - 1$ Sibuya distributed like in (10). The rv X_∞ is heavy-tailed with tail index $\alpha - 1$, so with a finite mean value only if $\alpha > 2$. Conditionally given $X_\infty \geq 1$, the law of X_∞ also is, from Eq. (30), the one of a Sibuya rv K^+, now with parameters $(\alpha - 1, \nu - 1)$.

- When $\alpha < 1$ as in the null recurrent case for X_n, $n^{-1}\mathbf{E}X_n$ has a limit and X_n drifts linearly to infinity with n going to ∞, on average. The critical value $\alpha = 1$ is a pinning transition point.

5.4 *First passage times and Green kernel*

Let $\tau_{k,k'} = \inf(n \geq 1 : X_n = k' \mid X_0 = k)$ be the first passage time at state $k' \neq k$ when the process is started at k. We wish here to briefly revisit an exact formal formula for the law of $\tau_{k,k'}$, making use

of the Green kernel of the MCC. Let

$$\phi_{k,k'}(z) := \sum_{l=1}^{\infty} z^l \mathbf{P}\left(\tau_{k,k'} = l\right)$$

be the generating function of the law of $\tau_{k,k'}$. Then, with

$$g_z\left(k,k'\right) := \sum_{n=0}^{\infty} z^n \mathbf{P}_k\left(X_n = k'\right) = \sum_{n=0}^{\infty} z^n P^n\left(k,k'\right) = \left(I - zP\right)^{-1}\left(k,k'\right),$$

the generating function of $P^n(k,k')$ (the Green potential function of the chain), using $P^n(k,k') = \sum_{m=0}^{n} \mathbf{P}\left(\tau_{k,k'} = m\right) P^{n-m}(k',k')$, we get the expression:

$$\phi_{k,k'}(z) = \frac{g_z(k,k')}{g_z(k',k')}. \tag{32}$$

The pgf $\phi_{k,k}(z) = \mathbf{E}\left(z^{\tau_{k,k}}\right)$ of the first-return time $\tau_{k,k}$ to state k satisfies

$$\phi_{k,k}(z) = \frac{g_z(k,k) - 1}{g_z(k,k)} = 1 - \frac{1}{g_z(k,k)} \tag{33}$$

where $g_z(k,k) = 1 + \sum_{n=1}^{\infty} z^n \mathbf{P}_k(X_n = k) = \sum_{n=0}^{\infty} z^n P^n(k,k)$ is the Green kernel at (k,k). In particular, because in our MCC example, $\tau_{0,0} \overset{d}{=} K + 2$, we obtain

Proposition 10 —

$$g_z(0,0) = \frac{1}{1 - \phi_{0,0}(z)} = \frac{1}{1 - \frac{\alpha z^2}{\nu+1} \cdot F(1, \nu+1-\alpha; \nu+2; z)}. \tag{34}$$

5.5 *Reversing time*

We emphasized above that the MCC X_n (with $\alpha > 1$ and $p_0 = 1$) could be interpreted as the backward recurrence time (or age) τ_n^B at n of a renewal process generated by $\tau_{0,0} \overset{d}{=} K^+ + 1$, so with $\tau_n^B \overset{d}{=} X_n$. The question we will finally address is: what is the process whose law at time n is the one of τ_n^F (the forward recurrence time or remaining lifetime)? Suppose again $\alpha > 1$ and $p_0 = 1$. The ergodic MCC X_n converging in distribution to $\tau_\infty^B \overset{d}{=} \tau_\infty^F \overset{d}{\sim} \mu$ is not time-reversible as detailed balance does not hold. Let $\overleftarrow{P} \neq P$ be the (stochastic) transition matrix of the process $\overleftarrow{X}_n$ which is X_n runned backward in time. With $'$ denoting matrix transposition and D_μ =diag$(\mu_0, \mu_1, ...)$, we have

$$\overleftarrow{P} = D_\mu^{-1} P' D_\mu.$$

It can be checked from Eqs. (25) and (26), that the only non-null entries of $\overleftarrow{P}$ are its first row with ($\overleftarrow{P}_{0,0} = 0$ and $\overleftarrow{P}_{0,k} = \mathbb{P}(\tau_{0,0} - 1 = k)$ if $k \geq 1$) and the lower diagonal whose entries are all

ones. Starting from $\overleftarrow{X}_0$, the process $\overleftarrow{X}_n$ decays linearly till it hits 0 and once at state 0, $\overleftarrow{X}_n$ jumps abruptly upward, undergoing a jump of random amplitude $K^+ \stackrel{d}{=} \tau_{0,0} - 1 \geq 1$ before diminishing again step by step to 0, where it starts afresh (Such a process is in the spirit of the processes defined in Subsection 4.1 with systematic death balanced by immigration events). In other words, the dynamics of $\overleftarrow{X}_n$ is

$$\overleftarrow{X}_{n+1} = \left(\overleftarrow{X}_n - 1\right) \mathbf{1}\left(\overleftarrow{X}_n > 0\right) + K^+_{n+1}.\mathbf{1}\left(\overleftarrow{X}_n = 0\right),$$

where $\{K^+_n\}$ is an iid sequence of rvs with common law the one of $K^+ \geq 1$. Clearly now, $\overleftarrow{X}_n$ can be interpreted as the forward recurrence time at n, τ^F_n, of the original process X_n (else $\tau^F_n \stackrel{d}{=} \overleftarrow{X}_n$), with $\overleftarrow{X}_n \stackrel{d}{\to} \overleftarrow{X}_\infty$ as $n \to \infty$. Note that $\mu' = \mu' \overleftarrow{P}$ (μ is also the invariant measure of $\overleftarrow{P}$) and so: $X_\infty \stackrel{d}{=} \overleftarrow{X}_\infty \stackrel{d}{=} \tau^F_\infty \stackrel{d}{\sim} \mu$, as required. To summarize:

Proposition 11 — The limiting rv $\overleftarrow{X}_\infty$ of the time-reversed MCC $\overleftarrow{X}_n$ of (21), once reflected at the origin, can be interpreted as the forward recurrence time of a renewal process generated by the excursion length $\tau_{0,0}$, with $\tau_{0,0} - 1$ Sibuya distributed like in (10).

ACKNOWLEDGEMENT

The author acknowledges partial support from the labex MME-DII (Modèles Mathématiques et Économiques de la Dynamique, de l' Incertitude et des Interactions), ANR11-LBX-0023-01. This work also benefited from the support of the Chair "Modélisation Mathématique et Biodiversité" of Veolia-Ecole Polytechnique-MNHN-Fondation X.

REFERENCES

1. R. Arratia, T. M. Liggett, and M. J. Williamson, Scale-free and power law distribution via fixed points and convergence of (thinning and conditioning) transformations, *Electronic Communications in Probability*, **19**(39) (2014), 10 pp.

2. N. H. Bingham, C. M. Goldie, and I. L. Teugels, *Regular variation,* (Encyclopedia of Mathematics and its Applications, 27), Cambridge University Press, (1987), 491 pp.

3. P. Collet, F. Dunlop, and T. Huillet, Wetting Transitions for a random line in long-range potential, *J. Stat. Phys.*, **160**(6) (2015), 1545-1622.

4. M. F. Dacey, A hypergeometric family of discrete probability distributions: Properties and applications to location models, *Geographical Analysis*, **1**(3) (1969), 219-320.

5. H. Dette, J. A. Fill, J. Pitman, and W. J. Studden, Wall and Siegmund duality relations for birth and death chains with reflecting barrier, Dedicated to Murray Rosenblatt, *J. Theoret. Probab.*, **10**(2) (1997), 349-374.

6. L. Devroye, A note on Linnik distribution, *Statistics and Probability Letters*, **9** (1990), 305-306.

7. L. Devroye, A triptych of discrete distributions related to stable law, *Statistics & Probability Letters*, **18** (1993), 349-351.

8. S. N. Dorogovtsev and J. F. F. Mendes, Evolution of networks, *Advances in Physics*, **51** (2002), 1079.

9. E. B. Dynkin, Limit theorems for sums of independent random quantities, *Izves. Akad. Nauk U.S.S.R.*, **19** (1955), 247-266.

10. M. Greenwood, On the statistical measure of infectiousness, *Journal of Hygiene*, **31**, Cambridge (1931), 336-351.

11. T. E. Harris, *The theory of branching processes*, Die Grundlehren der Mathematischen Wissenschaften, Bd. 119 Springer-Verlag, Berlin, Prentice-Hall, Inc., Englewood Cliffs, N.J. (1963).

12. T. Huillet, On Mittag-Leffler distributions and related stochastic processes, *Journal of Computational and Applied Mathematics*, **296** (2016), 181-211.

13. N. L. Johnson and S. Kotz, *Urn models and their application*, John Wiley, New York, (1977).

14. M. Kac, On the notion of recurrence in discrete stochastic processes, *Bulletin of the American Mathematical Society*, **53** (1947), 1002-1010 (Reprinted in Kac's Probability, Number Theory, and Statistical Physics: Selected Papers, 231-239).

15. T. J. Kozubowski and K. Podgórski, A generalized Sibuya distribution, *Annals of the Institute of Statistical Mathematics*, **70**(4) (2018), 855-887.

16. J. Lamperti, An invariance principle in renewal theory, *Annals of Mathematical Statistics*, **33** (1962), 685-696.

17. P. Lévy, Systèmes markoviens et stationnaires, cas dénombrables, *Annales Scientifiques de l' ENS*, 3ème Série, tome **68** (1951), 327-381.

18. B. Mandelbrot, A note on a class of skew distribution functions, analysis and critique of a paper by H. Simon. *Information and Control*, **2** (1959), 90.

19. F. Polito and L. Sacerdote, Random graphs associated to some discrete and continuous time preferential attachment models, *J. Stat. Phys.*, **162**(6) (2016), 1608-1638.

20. D. de S. Price, A general theory of bibliometric and other cumulative advantage, *Journal of American Society for Information Science*, **27** (1976), 292.

21. M. P. Quine and E. Seneta, A limit theorem for the Galton-Watson process with immigration, *Australian & New-Zealand Journal of Statistics*, **11**(3) (1969), 166-173.

22. K. Schreiber, Discrete self-decomposable distributions, *Dr. rer. nat. Thesis dissertation*, Otto-von-Guericke-Universität Magdeburg, (1999).

23. M. Sibuya, Generalized hypergeometric, digamma and trigamma distributions, *Annals of the Institute of Statistical Mathematics*, **31** (1979), 373-390.

24. M. V. Simkin and V. P. Roychowdhury, Re-inventing Willis, *Physics Reports,* **502**(1) (2011), 1-35.

25. H. A. Simon, On a class of skew distribution functions, *Biometrika,* **42**(3/4) (1955), 425-440.

26. H. A. Simon, Some further notes on a class of skew distribution functions, *Information and Control,* **3**(1) (1960), 80-88.

27. F. W. Steutel and K. van Harn, *Infinite divisibility of probability distributions on the real line*, Chapman and Hall/CRC Pure and Applied Mathematics, (2003).

28. F. W. Steutel and K. van Harn, Discrete analogues of self-decomposability and stability, *Ann. Prob.,* **7** (1979), 893-899.

29. K. van Harn, F. W. Steutel, and W. Vervaat, Self-decomposable discrete distributions and branching processes, *Z. Wahrsch. Verw. Gebiete*, **61** (1982), 97-118.

30. W. Vervaat, On a stochastic difference equation and a representation of non-negative infinitely divisible random variables, *Adv. Appl. Probab.,* **11** (1979), 750-783.

31. J. C. Willis and G. U. Yule, Some statistics of evolution and geographical distribution in plants and animals, and their significance, *Nature,* **109** (1922), 177.

32. G. U. Yule, A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S., *Philosophical Transactions of the Royal Society of London,* **213**(B) (1925), 21-87.