

Extensions of the generalized Pólya process

Francisco Germán Badía · Sophie Mercier ·
Carmen Sangüesa

Received: date / Accepted: date

Abstract A new self-exciting counting process is here considered, which extends the generalized Pólya process introduced by Cha [Advances in Applied Probability, 46, 1148-1171 (2014)]. Contrary to Cha's original model, where the intensity of the process (linearly) increases at each jump time, the extended version allows for more flexibility in the dependence between the point-wise intensity of the process at some time t and the number of already observed jumps. This allows the "extended Pólya process" to be appropriate, e.g., for describing successive failures of a system subject to imperfect but effective repairs, where the repair can lower the intensity of the underlying counting process. Probabilistic properties of the new process are studied (construction from a homogeneous pure-birth process, conditions of non explosion, computation of distributions, convergence of a sequence of such processes, ...) and its connection with Generalized Order Statistics is highlighted. Positive dependence properties are also provided. The maximum likelihood method is finally considered in a parametric setting and tested on a few simulated data sets, to highlight the practical use of the new process in an application context.

Keywords Counting process · Non-homogeneous pure-birth process · Stochastic order · Reliability theory

Mathematics Subject Classification (2000) 60K10 · 60E15

F. Germán Badía
Department of Statistical Methods, University of Zaragoza, Zaragoza, 50018, Spain
E-mail: gbadia@unizar.es

S. Mercier
CNRS/Univ Pau & Pays Adour/E2S UPPA, Laboratoire de Mathématiques et de leurs Applications - IPRA, UMR5142, 64000, Pau, France
Tel: +33-5-59-40-75-37 E-mail: sophie.mercier@univ-pau.fr ORCID: 0000-0003-3916-024X

C. Sangüesa
Department of Statistical Methods, University of Zaragoza, Zaragoza, 50009, Spain
E-mail: csangues@unizar.es

1 Introduction

The point of the paper is the study of a self-exciting counting process $(N_t)_{t \geq 0}$, with stochastic intensity of the shape

$$\lambda_t = h(N_{t-}) \lambda(t) \quad (1)$$

where $h(\cdot) : \mathbb{N} \rightarrow \mathbb{R}_+^*$ and where $\lambda(\cdot)$ stands for the (deterministic) baseline intensity function of the system, with $\lambda(t) \geq 0$ for all $t > 0$.

The specific case $h(n) = \alpha n + 1$ with $\alpha \geq 0$ has already been considered in several papers from the literature, see, e.g., [Asfaw and Lindqvist(2015), Babykina and Couallier(2010), Cha(2014), Konno(2010), Le Gat(2009), Le Gat(2014)]. This specific case is referred to as Linear Extension of the Yule Process (LEYP) by [Le Gat(2009), Le Gat(2014)]. In [Cha(2014)], the author prefers considering the model $h(n) = \alpha n + \beta$ with $\alpha \geq 0$, $\beta > 0$, which allows him to easily write down the so-called "restarting property" of the process (details later on). This provides an equivalent model as the LEYP which Cha calls Generalized Pólya Process (GPP). We use this terminology in the following. In such a model, $h(\cdot)$ is non decreasing and λ_t makes a positive jump at each arrival time. This can be useful for describing, for instance, shocks to an organism throughout his lifetime as in [Cha and Finkelstein(2016)]. However, it can be restrictive in some other application context. As an example from reliability theory, one can think of a system submitted to instantaneous imperfect repairs at failure times $T_1 < T_2 < \dots < T_n < \dots$, which can be thought as the arrival times of a counting process $(N_t)_{t \geq 0}$. Typically, a repair is expected to improve the state of the system and considering the intensity of the counting process, one can expect that it is lowered by a repair, so that λ_t should make a negative jump at each arrival (repair) time. A GPP model is hence not adapted for modeling the effect of such repairs. Note that other types of counting processes have been used in the reliability literature for modeling imperfect repairs, such as virtual age models [Kijima(1989)], among which first-order Arithmetic Reduction of Age models [Doyen and Gaudoin(2004)] with $\lambda_t = \lambda(t - \rho T_{N_{t-}})$ and $\rho \in [0, 1]$, geometric processes [Lam(2007)] with

$$\lambda_t = \frac{1}{\rho^{N_{t-}}} \lambda\left(\frac{t - T_{N_{t-}}}{\rho^{N_{t-}}}\right)$$

and $\rho > 0$ (see also [Bordes and Mercier(2013)] for an extended version). See [Chauvel et al(2016)] for other models of imperfect repairs. In a similar spirit as GPPs, we here suggest to consider a stochastic intensity which is piecewise proportional to a baseline deterministic failure intensity $\lambda(\cdot)$, with a proportionality coefficient depending of the number of already suffered failures from the system, as defined by (1). When $h(\cdot) < 1$, this allows to model repairs with a certain efficiency. The case $h(\cdot) = 1$ corresponds to As Bad As Old (ABAO) repairs, where the successive failure times are the points of a NHPP with intensity $\lambda(\cdot)$ (written NHPP($\lambda(\cdot)$)). The case $h(\cdot) > 1$ corresponds to worse than ABAO repairs, which deteriorate the state of the system (as a GPP). Note that the model allows to consider a function $h(\cdot)$ such that $h(n) < 1$ for $n \leq n_0$ and $h(n) \geq 1$ for $n > n_0$ which would mean that only the first n_0 repairs are efficient and that they become inefficient later on. The model hence presents more flexibility than GPPs for application purpose.

As we will discuss in Section 2, our model can be seen as a specific non-homogeneous pure birth process. These processes have been extensively used in the literature with applications in shock models (see [Sheu et al(2016)] and the references therein) or insurance models, in which the arrival of claims are time-dependent (see [Landriault et al(2014)] and the references therein). Moreover, considering a slightly generalized version of (1) (see Definition 2) allows us to see Generalized Order Statistics (GOS) of continuous distributions as the arrival points of our model. Generalized order statistics were introduced by [Kamps(1995a)] in order to give a unified setting to model ordered random variables, such as usual order statistics, sequential order statistics or record values. They nowadays constitute an active area of research (see, for instance, [Bedbur and Kamps(2017)] and the references therein).

The article is organized as follows. In Section 2, we provide a first general definition (Definition 2) in which the existence of the stochastic intensity is not required, which is next specialized in Definition 4 to the case where it exists. In this section, we establish the connection between our model and generalized order statistics. We next observe that our model can be constructed from a homogeneous pure-birth process through a change of time scale (Proposition 2). This allows to extend several well-known properties in the context of homogeneous pure birth processes to our setting. In particular, we show a "restarting property" in Proposition 7, where it is observed that a shifted process from an arbitrary time point provides another process of the same family. This extends the same property given by [Cha(2014)] for the generalized Pólya process. We also show a convergence property to a limiting process (Proposition 8), considering a sequence of processes associated with convergent sequences of parameters $(\lambda^{(k)})$ and $(h^{(k)})$. In Section 3, we establish several dependence properties between both increments and arrival times of the process. Finally, the practical effectiveness of the model is demonstrated in Section 4 through the development of a classical maximum likelihood parametric estimation procedure on a few simulated data sets.

2 Definition, construction and first properties

2.1 Definitions

We begin by defining what is called cumulative intensity function in all the following.

Definition 1 A function $A(\cdot)$ is called cumulative intensity function if:

- There exists $M_A \in \mathbb{R}_+ \cup \{\infty\}$ such that $A(\cdot) : [0, M_A) \rightarrow \mathbb{R}_+$,
- $A(\cdot)$ is non-decreasing and continuous on $[0, M_A)$,
- $A(0) = 0$,
- $\lim_{t \rightarrow M_A^-} A(t) = \infty$.

Remark 1 For some specific applications, $A(\cdot)$ will stand for the cumulative hazard rate function of a non-negative continuous random variable in the sequel. When $M_A < \infty$, the restriction of the domain of $A(\cdot)$ to $[0, M_A)$ allows $A(\cdot)$ to remain finite on its whole domain, even in case of a random variable with a bounded support.

We are now ready to define Extended Pólya Processes in their most general version, where the cumulative intensity function is not assumed to be absolutely continuous. They will be referred to as EPP* in the following. The specific case where the cumulative intensity function is absolutely continuous will be seen later on and referred to as EPP.

Definition 2 Let $h(\cdot) : \mathbb{N} \rightarrow \mathbb{R}_+^*$ and let $\Lambda(\cdot)$ be a cumulative intensity function (on $[0, M_\Lambda]$). Let $N = (N_t)_{0 \leq t < M_\Lambda}$ be a counting process with points $(T_n)_{n \in \mathbb{N}}$ where $T_0 = 0$. Then N is called an Extended Pólya Process* with parameter $(\Lambda(\cdot), h(\cdot))$ (written EPP*($\Lambda(\cdot), h(\cdot)$)) if we have:

$$\mathbb{P}(T_1 > t) = e^{-h(0)\Lambda(t)}, \quad t \in [0, M_\Lambda] \quad (2)$$

and

$$\mathbb{P}(T_{n+1} > t | \mathcal{F}_{T_n}) = \begin{cases} e^{-h(n)(\Lambda(t) - \Lambda(T_n))}, & t \in (T_n, M_\Lambda) \\ 1, & t \in [0, T_n] \end{cases} \quad (3)$$

$$= e^{-h(n)(\Lambda(t) - \Lambda(T_n))^+}, \quad t \in [0, M_\Lambda] \quad \text{a.s.} \quad (4)$$

where $\mathcal{F}_{T_n} = \sigma(T_i, i = 1, \dots, n)$ is the σ -field generated by $\{T_i, i = 1, \dots, n\}$ for $n = 1, 2, \dots$, a.s. means "almost surely" and $x^+ = \max(x, 0)$ for all $x \in \mathbb{R}$.

Alternatively to the first line of (3), one may also write:

$$\mathbb{P}(T_{n+1} - T_n > u | \mathcal{F}_{T_n}) = e^{-h(n)(\Lambda(u+T_n) - \Lambda(T_n))} \quad \text{a.s.} \quad (5)$$

for all $u \in (0, M_\Lambda - T_n)$.

This shows that the inter-arrival times of an EPP* are dependent and hence EPP*s enlarge several models from the previous literature such as renewal processes [Asmussen(2003)] or geometric processes and extensions [Lam(2007), Bordes and Mercier(2013)], for which inter-arrival times are independent.

Remark 2 Note that both sets of functional parameters $(C \Lambda(\cdot), h(\cdot)/C)$ and $(\Lambda(\cdot), h(\cdot))$ provide the same Extended Pólya Process*, whatever $C > 0$ is. In a general setting, the model hence requires some additional condition to be possibly identifiable such as $h(0) = 1$ or $\Lambda(1) = 1$, to remove the ambiguity on the constant C . As an example, a parametric framework such as $\Lambda(t) = t^b$ and $h(n) = c n$ would lead to a theoretically identifiable model.

To better connect Extended Pólya Processes with their Generalized version introduced by [Cha(2014)] and with other models from the literature, we now come to the case where Λ is absolutely continuous. We begin with defining so-called intensity functions.

Definition 3 A function $\lambda(\cdot)$ is called an intensity function if:

- There exists $M_\lambda \in \mathbb{R}_+ \cup \{\infty\}$ such that $\lambda(\cdot) : [0, M_\lambda] \rightarrow \mathbb{R}_+$,
- $\lambda(\cdot)$ is a measurable (Borel) function,
- $\int_0^t \lambda(u) du < +\infty$ for all $0 \leq t < M_\lambda$,
- $\int_0^{M_\lambda} \lambda(u) du = \infty$.

Starting from an intensity function $\lambda(\cdot)$, it is easy to check that the function

$$A(t) := \int_0^t \lambda(u) \, du, 0 \leq t < M_\lambda \quad (6)$$

is a cumulative intensity function such that $M_A = M_\lambda$. Conversely, starting from $A(\cdot)$, there exists $\lambda(\cdot)$ such that (6) is true as soon as $A(\cdot)$ is absolutely continuous.

Definition 4 Let $h(\cdot) : \mathbb{N} \rightarrow \mathbb{R}_+^*$ and let $\lambda(\cdot)$ be an intensity function. A counting process $N = (N_t)_{0 \leq t < M_\lambda}$ with stochastic intensity $\lambda_t = h(N_{t-})\lambda(t)$ for all $t \in [0, M_\lambda)$ is called an Extended Pólya Process (on $[0, M_\lambda)$) with parameter $(\lambda(\cdot), h(\cdot))$ (written $\text{EPP}(\lambda(\cdot), h(\cdot))$).

It is easy to check that an $\text{EPP}(\lambda(\cdot), h(\cdot))$ also is an $\text{EPP}^*(A(\cdot), h(\cdot))$ with $A(\cdot)$ provided by (6) and $M_\lambda = M_A$. Starting from an $\text{EPP}(\lambda(\cdot), h(\cdot))$, we shall refer to the corresponding cumulative intensity function A with no further notification. Also, we shall make use of M_λ or M_A indifferently, according to the context.

The case $h(\cdot) = C$ (and $M_\lambda = \infty$) corresponds to a Non Homogeneous Poisson Process (NHPP($C \lambda(\cdot)$)). The case $h(n) = \alpha n + \beta$ with $\alpha \geq 0$ corresponds to a GPP (or LEYP). Two other parametric shapes are also considered in the following, for illustration purpose:

- $h(n) = (1+n)^\alpha$ with $\alpha \in \mathbb{R}$,
- $h(n) = q^n$ with $q > 0$.

According to the parameters α and q , we may have $h(\cdot) > 1$ (case $\alpha > 0$ and $q > 1$, respectively), $h(\cdot) < 1$ (case $\alpha < 0$ and $0 < q < 1$, respectively) or $h(\cdot) = 1$ (case $\alpha = 0$ and $q = 1$, respectively). These different behaviors are illustrated in Figure 1 for $A(t) = t^{1.25}$ and $h(n) = n + 1$ (case (a)), $h(n) = 1.15^n$ (case (b)), $h(n) = 0.9^n$ (case (c)), $h(n) = 1/(1+n)^{0.15}$ (case (d)). Accordingly, the interpretation of an EPP for application purpose may hence be similar as for a GPP (cases (a) and (b)) or provide a model for imperfect but effective repairs (cases (c) and (d)).

Following the vocabulary of [Belzunce et al(2001), p.202], one can note that an EPP on $[0, \infty)$ is a specific non homogeneous pure birth process with intensity functions $r_i(t) = h(i-1)\lambda(t)$ for all $i = 1, 2, \dots$ and $t \geq 0$.

In the specific case where $\lambda(\cdot)$ is a constant function ($\lambda(t) = \lambda$ for all $t \in \mathbb{R}_+$ with $\lambda > 0$), Equation (5) gives:

$$\mathbb{P}(T_{n+1} - T_n > u | \mathcal{F}_{T_n}) = e^{-h(n)\lambda u} \text{ a.s.} \quad (7)$$

for all $u \geq 0$. The inter-arrival times then are independent. Also, this specific EPP appears as a pure birth homogeneous Markov process with generator matrix $A = (a_{i,j})_{i,j \in \mathbb{N}}$ and

$$a_{i,j} = \begin{cases} h(i)\lambda & \text{if } j = i + 1, \\ -h(i)\lambda & \text{if } j = i, \\ 0 & \text{elsewhere.} \end{cases} \quad (8)$$

We now come to the connection between EPPs and Generalized Order Statistics (GOS). We recall that the connection between GOS and non-homogeneous pure-birth processes was previously observed by [Belzunce et al(2005)]. We begin by recalling the definition of GOS, as introduced in [Kamps(1995b), Definition 2.1 page 49].

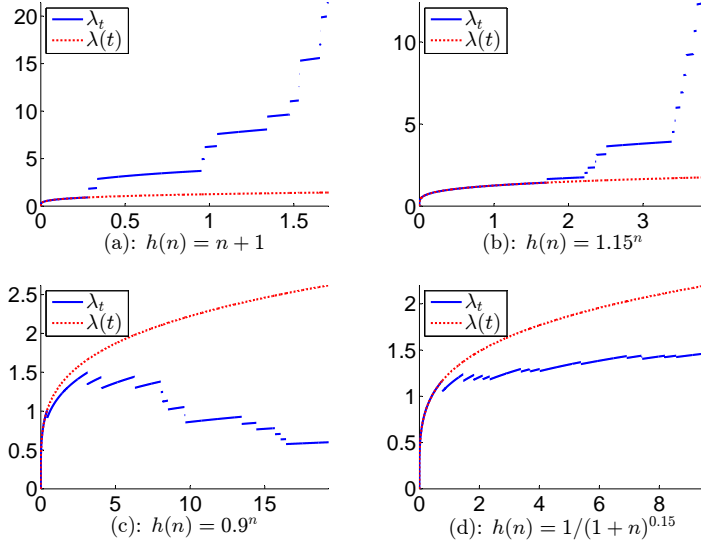


Fig. 1 Plot of $\lambda(t) = t^{1.25}$ and trajectories of λ_t for different choices of $h(n)$.

Definition 5 Let $n \in \mathbb{N}$, $k \geq 1$, $m_1, \dots, m_{n-1} \in \mathbb{R}$, $M_r = \sum_{j=r}^{n-1} m_j$, $1 \leq r \leq n-1$, be parameters such that $\gamma_r = k + n - r + M_r \geq 1$, for all $r \in \{1, \dots, n-1\}$ and let $\tilde{m} = (m_1, \dots, m_{n-1})$ if $n \geq 2$ ($\tilde{m} \in \mathbb{R}$ arbitrary for $n = 1$). We call uniform GOS the random vector $(U_{(1,n,\tilde{m},k)}, \dots, U_{(n,n,\tilde{m},k)})$ with joint probability density function (p.d.f.) given by

$$k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} (1 - u_j)^{m_j} \right) (1 - u_n)^{k-1}, \quad 0 \leq u_1 \leq \dots \leq u_n \leq 1$$

Now, given a cumulative distribution function (c.d.f.) F , we call GOS based on F the random vector

$$(X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)}) = (F^{-1}(U_{(1,n,\tilde{m},k)}), \dots, F^{-1}(U_{(n,n,\tilde{m},k)})),$$

in which F^{-1} denotes the inverse of F (quantile function). Moreover, if F has p.d.f. f , the joint p.d.f. of $(X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)})$ is given by

$$k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} \bar{F}^{m_j}(x_j) f(x_j) \right) \bar{F}(x_n)^{k-1} f(x_n),$$

for all (x_1, \dots, x_n) such that $F^{-1}(0) \leq x_1 \leq \dots \leq x_n \leq F^{-1}(1)$, where $\bar{F} := 1 - F$ denotes the survival function.

Now, let us consider a GOS $(X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)})$ based on a continuous c.d.f. F parameters (n, \tilde{m}, k) as in the previous definition. Assume that $F(0) = 0$ and let

$$M_\Lambda := \sup\{t \geq 0 : F(t) < 1\} \leq \infty$$

and let Λ be the cumulative hazard rate for F , that is

$$\Lambda(t) = -\log(\bar{F}(t)), \quad t \in [0, M_\Lambda]. \quad (9)$$

Then Λ is a cumulative intensity function on $[0, M_\Lambda)$. Now, setting

$$h(i) := \gamma_{i+1}, \quad i = 0, 1, \dots, n-2, \quad \text{and} \quad h(n-1) := k, \quad (10)$$

we can see from (2) and (3) that the arrival times (T_1, \dots, T_n) of an $\text{EPP}^*(\Lambda(\cdot), h(\cdot))$ are identically distributed as those of a Markov chain with initial distribution

$$\mathbb{P}(T_1 > t) = e^{-h(0)\Lambda(t)} = \bar{F}(t)^{h(0)}, \quad t \in [0, M_\Lambda)$$

and transition probabilities

$$\mathbb{P}(T_{i+1} > t | T_i) = \begin{cases} e^{-h(i)(\Lambda(t) - \Lambda(T_i))} = \left(\frac{\bar{F}(t)}{\bar{F}(T_i)}\right)^{h(i)}, & t \in (T_i, M_\Lambda), \\ 1, & t \in [0, T_i], \end{cases}$$

(a.s.) for $i \in \{1, \dots, n-1\}$. Based on the equivalent representation of GOS in terms of a Markov Chain provided in [Kamps(1995b), Remark 2.8 page 56], this means that the GOS of order n for the continuous c.d.f. F with $F(0) = 0$ can be seen as the first n points in an $\text{EPP}^*(\Lambda(\cdot), h(\cdot))$.

Remark 3 Note that considering the reverse construction and starting from the points of an EPP^* , we would recover the points of an extended GOS as proposed in [Kamps(1995b), Remark 2.2 page 50], where $k > 0$ and $\gamma_r > 0$ (whereas $k \geq 1$ and $\gamma_r \geq 1$ is required in [Kamps(1995b), Definition 2.1 page 49] and [Kamps(1995a)]).

2.2 Construction - Conditions of non-explosion

We here provide several probabilistic constructions of an EPP^* together with conditions of non-explosion. As a first step, we check that the assumptions on $(\Lambda(\cdot), h(\cdot))$ given in Definition 2 allow the points $(T_n)_{n \in \mathbb{N}}$ of an EPP^* to be well defined (finite).

Lemma 1 *Let $(\Lambda(\cdot), h(\cdot))$ (and M_Λ) as in Definition 2 and let $(T_n)_{n \in \mathbb{N}}$ be recursively constructed through (2-3) with $T_0 = 0$. Set $N = (N_t)_{0 \leq t < M_\Lambda}$ to be the corresponding counting process. Then:*

1. $T_n < T_{n+1} < M_\Lambda$ a.s. for all $n \in \mathbb{N}$.
2. $\lim_{t \rightarrow M_\Lambda^-} N_t = +\infty$ a.s.

Proof For the first point, let us first show that $T_n < M_\Lambda$ a.s., namely that $\mathbb{P}(T_n \geq M_\Lambda) = 0$ for all $n \in \mathbb{N}^*$. We proceed by induction.

For $n = 1$, we have by monotonous convergence:

$$\mathbb{P}(T_1 \geq M_\Lambda) = \lim_{t \rightarrow M_\Lambda^-} \mathbb{P}(T_1 > t) = \lim_{t \rightarrow M_\Lambda^-} e^{-h(0)\Lambda(t)} = 0$$

because $\lim_{t \rightarrow M_\Lambda^-} \Lambda(t) = \infty$ and $h(0) > 0$.

Now, assume the property to be true for some $n \in \mathbb{N}^*$. Based on (4), we have

$$\mathbb{P}(T_{n+1} > t | \mathcal{F}_{T_n}) = \phi(T_n) \quad (11)$$

with $\phi(T_n) = e^{-h(n)(\Lambda(t) - \Lambda(T_n))^+}$. Taking the expectation in (11), we get:

$$\mathbb{P}(T_{n+1} > t) = \mathbb{E}[\phi(T_n)] = \mathbb{E}\left[e^{-h(n)(\Lambda(t) - \Lambda(T_n))^+}\right].$$

This provides

$$\mathbb{P}(T_{n+1} \geq M_\Lambda) = \lim_{t \rightarrow M_\Lambda^-} \mathbb{P}(T_{n+1} > t) = \lim_{t \rightarrow M_\Lambda^-} \mathbb{E}\left[e^{-h(n)(\Lambda(t) - \Lambda(T_n))^+}\right] = 0$$

by dominated convergence. The property is hence true for all $n \in \mathbb{N}^*$, namely $T_n < M_\Lambda$ a.s. for all $n \in \mathbb{N}$. The fact that $T_n < T_{n+1}$ a.s. now is a direct consequence of (5), taking into account that $T_n < M_\Lambda$ (and thus $\Lambda(T_n) < \infty$).

The second point is a classical consequence of the first point and its proof is omitted.

The points $(T_n)_{n \in \mathbb{N}}$ of an EPP*, as expected, form an almost surely increasing sequence. Our aim now is to provide other different constructions in terms of well-known processes. First of all, considering a cumulative intensity function Λ , we need to introduce its generalized inverse function. The function Λ is non-decreasing with range $\Lambda([0, M_\Lambda]) = \mathbb{R}_+$, whatever M_Λ is. Then, we can introduce its generalized inverse function Λ^{-1} on \mathbb{R}_+ , with:

$$\Lambda^{-1}(s) = \inf\{0 \leq t < M_\Lambda : \Lambda(t) \geq s\} = \sup\{0 \leq t < M_\Lambda : \Lambda(t) < s\} \quad (12)$$

for all $s \geq 0$ with the convention $\sup \emptyset = 0$. The function Λ^{-1} is known to be left-continuous on \mathbb{R}_+ and such that $\Lambda(\Lambda^{-1}(s)) = s$ for all $s \in \mathbb{R}_+$ (based on the right-continuity of Λ and on its range). Also, for all $s \in \mathbb{R}_+$, $t \in [0, M_\Lambda)$, we have $\Lambda(t) < s$ if and only if $t < \Lambda^{-1}(s)$ (based on the right-continuity of Λ). Finally, Λ^{-1} is (strictly) increasing, based on the continuity of Λ . See [Embrechts and Hofert(2013)] or [Boyer and Roux(2016)] for more details on generalized inverse functions.

If Λ is not strictly increasing, we have the problem that the previous inverse function is not a continuous function despite the fact that Λ is continuous. To be more specific, the discontinuity points of the function Λ^{-1} are of the shape $\Lambda(s)$ with $\Lambda^{-1}(\Lambda(s)) < s$. This can happen when Λ is constant on an interval, for instance $\Lambda(s) = k$ for all $s \in [a, b]$. In that case, for all $s \in [a, b]$, we have $\Lambda^{-1}(\Lambda(s)) = \Lambda^{-1}(k) = a$. Now let

$$D := \{s \in [0, M_\Lambda] | \Lambda^{-1}(\Lambda(s)) < s\}. \quad (13)$$

If $[a, b]$ is the only interval on which Λ is constant, then $D = (a, b]$. As Λ^{-1} is increasing, such intervals are countably many so that D can be written as:

$$D = \bigcup_{i \in I} (a_i, b_i] \quad (14)$$

where I is at most countable. Note that if Λ_{D^c} stands for the restriction of Λ to D^c , the function Λ_{D^c} now is a one-to-one function from D^c to \mathbb{R}_+ . As a consequence, for all $s \in \mathbb{R}_+, t \in D^c$, we have $\Lambda(t) > s$ if and only if $t > \Lambda_{D^c}^{-1}(s)$, namely if and only if $t > \Lambda^{-1}(s)$.

Next result shows us that the probability of observing the points of an EPP* process in the set D is 0.

Proposition 1 *Let $(T_n)_{n \in \mathbb{N}}$ be the arrival points of an EPP*($\Lambda(\cdot), h(\cdot)$) and let D be the set defined in (13). We have*

- (a) $\mathbb{P}(T_{n+1} \in D | \mathcal{F}_{T_n}) = 0, \quad n = 0, 1, \dots$
(b) $\mathbb{P}(T_n \in D) = 0, \quad n = 1, 2, \dots$

Proof For part (a), we consider an interval $I_i = (a_i, b_i] \subseteq D, i \in I$. This means, using (3) (or (2) for $n = 0$) that

$$\begin{aligned} \mathbb{P}(T_{n+1} \in I_i | \mathcal{F}_{T_n}) &= \mathbb{P}(T_{n+1} > a_i | \mathcal{F}_{T_n}) - \mathbb{P}(T_{n+1} > b_i | \mathcal{F}_{T_n}) \\ &= e^{-h(n)(\Lambda(a_i) - \Lambda(T_n))} \mathbf{1}_{\{a_i > T_n\}} + \mathbf{1}_{\{a_i \leq T_n\}} \\ &\quad - e^{-h(n)(\Lambda(b_i) - \Lambda(T_n))} \mathbf{1}_{\{b_i > T_n\}} - \mathbf{1}_{\{b_i \leq T_n\}} \\ &= \left(1 - e^{-h(n)(\Lambda(a_i) - \Lambda(T_n))}\right) \mathbf{1}_{\{a_i \leq T_n < b_i\}} \\ &= 0 \end{aligned}$$

for all $n \in \mathbb{N}$, using that $\Lambda(a_i) = \Lambda(b_i)$ for the previous to last line and $\Lambda(T_n) = \Lambda(a_i)$ when $a_i \leq T_n < b_i$ for the last one. Thus, (a) follows as I is a countable set. To show (b), the case $n = 1$ is included in part (a), whereas for $n = 2, 3, \dots$

$$\mathbb{P}(T_n \in D) = \mathbb{E}[\mathbb{P}(T_n \in D | \mathcal{F}_{T_{n-1}})] = 0.$$

Remark 4 By Proposition 1 (b), we obviously have that

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} \{T_n \in D\}\right) = 0.$$

We now provide two different constructions of an EPP*, to be used later on. One construction is very similar to that of [Daley and Vere-Jones(2003), Theorem 7.4.I. page 258] (under different assumptions however). The other one allows to recover the points of a general EPP* from those of an EPP with $\lambda(\cdot) = \mathbf{1}$ (with $\mathbf{1}$ the constant function equal to 1).

Proposition 2 *Let $h(\cdot) : \mathbb{N} \rightarrow \mathbb{R}_+^*$ and let $\Lambda(\cdot)$ be a cumulative intensity function. Let $(T_n)_{n \in \mathbb{N}}$ be the arrival points of a counting process on $[0, M_\Lambda)$, with $T_0 = 0$ and let $(S_n)_{n \in \mathbb{N}}, (V_n)_{n \in \mathbb{N}}$ be the arrival points of two counting processes on \mathbb{R}_+ ,*

with $S_0 = V_0 = 0$. We assume the following link between the points of the three processes:

$$\begin{cases} T_n = \Lambda^{-1}(S_n) \\ T_{n+1} = \Lambda^{-1}\left(\Lambda(T_n) + \frac{V_{n+1} - V_n}{h(n)}\right) \end{cases} \quad (15)$$

for all $n \geq 0$, where we recall that Λ^{-1} stands for the left-continuous generalized inverse function of Λ .

1. We have

$$\begin{cases} S_n = \Lambda(T_n), \\ V_n = \sum_{i=1}^{n-1} h(i) (\Lambda(T_{i+1}) - \Lambda(T_i)) \end{cases} \quad (16)$$

for all $n \in \mathbb{N}$.

2. The three following assertions are equivalent:

- (a) $(T_n)_{n \in \mathbb{N}}$ are the points of an EPP* $(\Lambda(\cdot), h(\cdot))$,
- (b) $(S_n)_{n \in \mathbb{N}}$ are the points of an EPP $(\mathbf{1}, h(\cdot))$,
- (c) $(V_n)_{n \in \mathbb{N}}$ are the points of a homogeneous Poisson process with rate 1 (written HPP (1)).

Proof Assume the points of the three processes to be linked by (15). Then, it is easy to check that the first line of (16) is true, based on the fact that $\Lambda(\Lambda^{-1}(s)) = s$ for all $s \in \mathbb{R}_+$. Next, we know from (15) that

$$\frac{V_{n+1} - V_n}{h(n)} = \Lambda(T_{n+1}) - \Lambda(T_n)$$

for all $n \in \mathbb{N}$, which readily provides the second line of (16). Hence point 1 is proved.

In order to prove point 2, let us now consider the filtrations (\mathcal{F}_{T_n}) , (\mathcal{G}_{S_n}) and (\mathcal{H}_{V_n}) generated by the T_n 's, S_n 's and V_n 's, respectively. Based on (15) and (16), we clearly have $\mathcal{F}_{T_n} = \mathcal{G}_{S_n} = \mathcal{H}_{V_n}$ for all $n \in \mathbb{N}$. As a first step, assume that $(S_n)_{n \in \mathbb{N}}$ are the points of an EPP $(\mathbf{1}, h(\cdot))$. Then, on $\{t > V_n\}$:

$$\begin{aligned} \mathbb{P}(V_{n+1} > t | \mathcal{H}_{V_n}) &= \mathbb{P}(V_n + h(n)(S_{n+1} - S_n) > t | \mathcal{H}_{V_n}, \mathcal{G}_{S_n}) \\ &= \mathbb{P}\left(S_{n+1} - S_n > \frac{t - V_n}{h(n)} \middle| \mathcal{H}_{V_n}, \mathcal{G}_{S_n}\right) \\ &= e^{-h(n)\left(\frac{t - V_n}{h(n)}\right)} \\ &= e^{-(t - V_n)} \end{aligned}$$

This shows that $(V_n)_{n \in \mathbb{N}}$ are the points of an HPP (1) so that assertion 2b implies assertion 2c. The converse implication is similar and its proof is omitted.

Now, assume that $(V_n)_{n \in \mathbb{N}}$ are the points of an HPP (1). Remembering that $\Lambda^{-1}(s) > t$ if and only if $s > \Lambda(t)$, we have on $\{t > T_n\}$:

$$\begin{aligned} \mathbb{P}(T_{n+1} > t | \mathcal{F}_{T_n}) &= \mathbb{P}\left(\Lambda^{-1}\left(\Lambda(T_n) + \frac{V_{n+1} - V_n}{h(n)}\right) > t | \mathcal{F}_{T_n}\right) \\ &= \mathbb{P}\left(\Lambda(T_n) + \frac{V_{n+1} - V_n}{h(n)} > \Lambda(t) | \mathcal{F}_{T_n}, \mathcal{H}_{V_n}\right) \\ &= \mathbb{P}(V_{n+1} - V_n > h(n)(\Lambda(t) - \Lambda(T_n)) | \mathcal{F}_{T_n}, \mathcal{H}_{V_n}) \\ &= e^{-h(n)(\Lambda(t) - \Lambda(T_n))} \end{aligned}$$

because $\Lambda(t) \geq \Lambda(T_n)$. This shows that $(T_n)_{n \in \mathbb{N}}$ are the points of an $\text{EPP}^*(\Lambda(\cdot), h(\cdot))$ and assertion 2c (or 2b) implies assertion 2a. Now, the only thing remaining is to show that assertion 2a implies assertion 2c (or 2b). Then, assume $(T_n)_{n \in \mathbb{N}}$ to be the points of an $\text{EPP}^*(\Lambda(\cdot), h(\cdot))$. Recall that, due to Proposition 1, $\mathbb{P}(T_n \in D) = 0$, where D is defined in (13). But note that on the set D^c , Λ is a one-to-one function and the generalized inverse function coincides with the "true" inverse function. Thus, on $\{t > S_n\}$,

$$\begin{aligned} \mathbb{P}(S_{n+1} > t | \mathcal{G}_{S_n}) &= \mathbb{P}(\Lambda(T_{n+1}) > t, T_{n+1} \in D^c | \mathcal{F}_{T_n}) \\ &= \mathbb{P}(T_{n+1} > \Lambda^{-1}(t), T_{n+1} \in D^c | \mathcal{F}_{T_n}) \\ &= \mathbb{P}(T_{n+1} > \Lambda^{-1}(t) | \mathcal{F}_{T_n}) \\ &= e^{-h(n)(\Lambda(\Lambda^{-1}(t)) - \Lambda(T_n))} \\ &= e^{-h(n)(t - S_n)} \end{aligned}$$

This shows that $(S_n)_{n \in \mathbb{N}}$ are the points of an $\text{EPP}(\mathbf{1}, h(\cdot))$ and concludes this proof.

Remark 5 The previous proposition allows us to construct the points of a general $\text{EPP}^*(\Lambda(\cdot), h(\cdot))$ from those of an $\text{EPP}(\mathbf{1}, h(\cdot))$ or from those of an $\text{HPP}(1)$. Using the fact that the inter-arrival times of an $\text{HPP}(1)$ are standard exponential and can hence be simulated through setting $V_{n+1} - V_n = -\ln(U_n)$ for all $n \in \mathbb{N}$, where $(U_n)_{n \in \mathbb{N}}$ is an i.i.d. sequence of uniform random variables on $[0, 1]$, it is now easy to recursively simulate the points $(T_n)_{n \in \mathbb{N}}$ of a general EPP^* : first set $T_0 = 0$; next construct the T_n 's recursively through

$$T_{n+1} = \Lambda^{-1} \left(\Lambda(T_n) - \frac{\ln(U_n)}{h(n)} \right) \quad (17)$$

for all $n \in \mathbb{N}$.

The following corollary is a direct consequence of Proposition 2.

Corollary 1 *Let $h(\cdot) : \mathbb{N} \rightarrow \mathbb{R}_+^*$ and let $\Lambda_2(\cdot)$ be a cumulative intensity function on $[0, M_{\Lambda_2})$. Let $N^{(1)} = (N_t^{(1)})_{t \geq 0}$ and $N^{(2)} = (N_t^{(2)})_{t \in [0, M_{\Lambda_2})}$ be two counting processes on \mathbb{R}_+ and $[0, M_{\Lambda_2})$, respectively, such that $N_t^{(2)} = N_{\Lambda_2(t)}^{(1)}$ for all $t \in [0, M_{\Lambda_2})$. Then $N^{(1)}$ is an $\text{EPP}(\mathbf{1}, h(\cdot))$ if and only if $N^{(2)}$ is an $\text{EPP}^*(\Lambda_2(\cdot), h(\cdot))$.*

Proof For $i \in \{1, 2\}$, let $(T_n^{(i)})_{n \in \mathbb{N}}$ be the points of $N^{(i)}$. For $n \in \mathbb{N}$ and $t \in [0, M_{\Lambda_2})$, we have $\{N_t^{(2)} \geq n\} = \{N_{\Lambda_2(t)}^{(1)} \geq n\}$ with

$$\{N_t^{(2)} \geq n\} = \{T_n^{(2)} \leq t\}$$

and

$$\{N_{\Lambda_2(t)}^{(1)} \geq n\} = \{T_n^{(1)} \leq \Lambda_2(t)\} = \{\Lambda_2^{-1}(T_n^{(1)}) \leq t\}.$$

This implies

$$\{T_n^{(2)} \leq t\} = \{\Lambda_2^{-1}(T_n^{(1)}) \leq t\}$$

for all $t \in [0, M_{\Lambda_2})$ and $T_n^{(2)} = \Lambda_2^{-1}(T_n^{(1)})$. The results follow from point 2 in Proposition 2.

We now come to conditions of non explosion. Let us first remind that a counting process with points $(T_n)_{n \in \mathbb{N}}$ in $[0, M_\Lambda)$ is said to be explosive as soon as $\mathbb{P}(\sup_n T_n < M_\Lambda) > 0$. Also, based on the fact that an EPP $(\mathbf{1}, h(\cdot))$ is a homogeneous Markov pure birth process, it is known from [Gikhman and Skorokhod(1969), page 323] that the EPP $(\mathbf{1}, h(\cdot))$ with points $(S_n)_{n \in \mathbb{N}}$ is explosive if and only if $\sum_{i=0}^{+\infty} \frac{1}{h(i)} < +\infty$ and that $\mathbb{P}(\sup_n S_n < +\infty)$ can only be either 0 or 1. This allows to get the following result for a general EPP.

Proposition 3 *Let $(T_n)_{n \in \mathbb{N}}$ be the points of an EPP $^*(\Lambda(\cdot), h(\cdot))$. We have the following dichotomy, whatever $\Lambda(\cdot)$ is:*

1. *If $\sum_{i=0}^{+\infty} \frac{1}{h(i)} = +\infty$, then $\sup_n T_n = M_\Lambda$ a.s. (almost sure non explosion). Also $N_t < +\infty$ a.s. for all $0 \leq t < M_\Lambda$.*
2. *If $\sum_{i=0}^{+\infty} \frac{1}{h(i)} < +\infty$, then $\sup_n T_n < M_\Lambda$ a.s. (almost sure explosion). Also, for almost all ω , there exists $T(\omega) < M_\Lambda$ such that $N_t(\omega) = \infty$ for all $t \in (T(\omega), M_\Lambda)$.*

Proof Let us assume the T_n 's to be constructed from the points $S_n, n \in \mathbb{N}$ of an EPP $(\mathbf{1}, h(\cdot))$ as in Proposition 2, with $T_n = \Lambda^{-1}(S_n)$, all $n \in \mathbb{N}$. Then:

$$\mathbb{P}\left(\sup_n T_n = M_\Lambda\right) = \mathbb{P}\left(\sup_n S_n = +\infty\right).$$

The dichotomy between almost sure explosion (point 1) and almost sure non explosion (point 2) hence is a direct consequence of the same result for the EPP $(\mathbf{1}, h(\cdot))$ with points $(S_n)_{n \in \mathbb{N}}$. The induced almost sure infiniteness (finiteness) of N_t under almost sure (non) explosion is classical and hence omitted.

Example 1 If $h(n) = \alpha n + \beta$ with $\alpha \geq 0, \beta > 0$ or if $h(n) = q^n$ with $q \leq 1$, there is almost surely no explosion, whatever $\Lambda(\cdot)$ is. The explosive and non explosive cases are illustrated in Figure 2 for $\Lambda(t) = t^{1.25}$ with $h(n) = 0.75^n$ (left; almost sure no explosion) and $h(n) = 1.5^n$ (right; almost sure explosion).

2.3 First probabilistic properties

We here provide several probabilistic results for an EPP * . We begin with the marginal distribution of an EPP * (probability mass function of N_t) together with the distribution of its arrival times $T_n, n \geq 1$ in the case where all $h(j)$'s are distinct.

Proposition 4 *Let $(N_t)_{t \in [0, M_\Lambda)}$ be an EPP $^*(\Lambda(\cdot), h(\cdot))$ and assume that $h(j), j \geq 0$ are all distinct. Then:*

$$\mathbb{P}(N_t = n) = \alpha_n \times \sum_{j=0}^n \gamma_{j,n} e^{-h(j)\Lambda(t)}$$

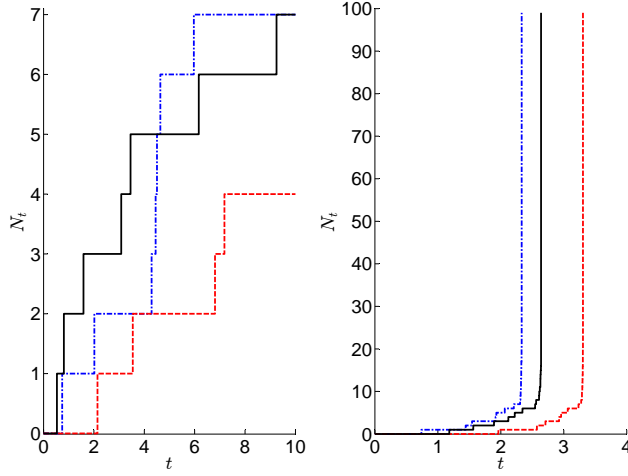


Fig. 2 A few trajectories of $(N_t)_{t \geq 0}$ for $\Lambda(t) = t^{1.25}$ with $h(n) = 0.75^n$ (left) and $h(n) = 1.5^n$ (right)

for all $n \geq 0$ and $0 \leq t < M_\Lambda$ with

$$\alpha_n = \prod_{k=0}^{n-1} h(k), \quad (18)$$

$$\gamma_{j,n} = \frac{1}{\prod_{0 \leq k \leq n, k \neq j} (h(k) - h(j))} \quad (19)$$

for $n \geq 1$ and $\alpha_0 = \gamma_{0,0} = 1$.

Proof In the specific case where $\lambda(\cdot) = \mathbf{1}$, the result can be found in the literature on homogeneous Markov pure birth processes, see, e.g., [Gikhman and Skorokhod(1969), formulas (3-4) page 322]. It is next easy to extend it to the case of a general $\Lambda(\cdot)$ by using the construction of a general EPP* from an EPP $(\mathbf{1}, h(\cdot))$ provided in Corollary 1.

Remark 6 In the explosive case (see Proposition 3), the distribution of N_t is defective in the sense that $\sum_{n=0}^{+\infty} \mathbb{P}(N_t = n) < 1$ and it admits a non zero mass at ∞ .

Example 2 In case of an EPP* with $h(n) = \alpha n + \beta$, we get for $n \geq 1$:

$$\alpha_n = \prod_{k=0}^{n-1} (\alpha k + \beta) = \alpha^n \frac{\Gamma\left(n + \frac{\beta}{\alpha}\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)}$$

and

$$\gamma_{j,n} = \frac{1}{\alpha^n \prod_{0 \leq k \leq n, k \neq j} (k - j)} = \frac{(-1)^j}{\alpha^n n!} \binom{n}{j}.$$

We easily derive that

$$\mathbb{P}(N_t = n) = \frac{\Gamma\left(n + \frac{\beta}{\alpha}\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} \times e^{-\beta\Lambda(t)} \left(1 - e^{-\alpha\Lambda(t)}\right)^n$$

so that N_t has a negative binomial distribution, which can be found in several papers from the literature (in the specific case of an EPP), see e.g. [Asfaw and Lindqvist(2015), Babykina and Couallier(2010), Cha(2014), Konno(2010), Le Gat(2009), Le Gat(2014)].

We now provide the c.d.f. (p.d.f.) of arrival times of an EPP* (EPP), here again in the case where all $h(j)$'s are distinct. When $\lambda(\cdot) = \mathbf{1}$, the result can be found in e.g. [Cox(1970), page 17]. It is next easy to extend it to the case of general EPP*'s by considering that $T_n = \Lambda^{-1}(S_n)$ in the notations of Proposition 2, which entails that $F_{T_n}(t) = F_{S_n}(\Lambda(t))$ and, for an EPP $(\lambda(\cdot), h(\cdot))$, $f_{T_n}(t) = \lambda(t)f_{S_n}(\Lambda(t))$.

Proposition 5 *Let $(T_n)_{n \geq 0}$ be the points of an EPP* $(\Lambda(\cdot), h(\cdot))$ (with $T_0 = 0$) and assume that $h(j)$, $j \geq 0$ are all distinct. For $n \geq 1$, the c.d.f. of T_n is given by*

$$F_{T_n}(t) = 1 - \sum_{i=0}^{n-1} A_{i,n-1} e^{-h(i)\Lambda(t)},$$

for all $t \in [0, M_\Lambda)$, with

$$A_{i,n-1} = \prod_{0 \leq k \leq n-1, k \neq i} \frac{h(k)}{h(k) - h(i)}$$

for $n \geq 1$ and $i \geq 0$, $A_0 = 1$. If, in addition the process is an EPP $(\lambda(\cdot), h(\cdot))$, the p.d.f. is given by

$$f_{T_n}(t) = \alpha_n \lambda(t) \sum_{i=0}^{n-1} \gamma_{i,n-1} e^{-h(i)\Lambda(t)},$$

where α_n and $\gamma_{i,n}$ are given by (18) and (19).

In case where the $h(j)$'s are possibly not all distinct and $\lambda(\cdot) = \mathbf{1}$, there are many different formulations in the literature for the distribution of T_n , e.g. see the section devoted to sums of gamma random variables in [Nadarajah(2008)], with references therein. We here propose two different formulations. The first one from [Efthymoglou and Aalo(1995)] gives the result through a computable integral form, the second one from [Aalo et al(2005)] proposes a full form expression in terms of the confluent Lauricella hypergeometric function of n variables. We recall that this function is defined by

$$\Phi_2^{(n)}(\beta_1, \dots, \beta_n; \gamma; x_1, \dots, x_n) = \sum_{m_1, \dots, m_n=0}^{+\infty} \frac{(\beta_1)_{m_1} \cdots (\beta_n)_{m_n} x_1^{m_1} \cdots x_n^{m_n}}{(\gamma)_{m_1 + \dots + m_n} m_1! \cdots m_n!},$$

where $(\alpha)_k = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$ is the Pochhammer symbol. Here again, the results are derived from those of [Aalo et al(2005), Efthymoglou and Aalo(1995)] by considering $T_n = \Lambda^{-1}(S_n)$ in the notations of Proposition 2.

Proposition 6 Let $(T_n)_{n \geq 0}$ be the points of an EPP* $(\Lambda(\cdot), h(\cdot))$ (with $T_0 = 0$). For $n \geq 1$ and $t \in (0, M_\Lambda)$, the c.d.f. of T_n is given by

$$F_{T_n}(t) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\sin\left(\sum_{i=0}^{n-1} \arctan\left(\frac{s}{h(i)}\right) - s \Lambda(t)\right)}{s \prod_{i=0}^{n-1} \left(1 + \left(\frac{s}{h(i)}\right)^2\right)^{\frac{1}{2}}} ds$$

and, if it is an EPP $(\lambda(\cdot), h(\cdot))$, its p.d.f is given by

$$f_{T_n}(t) = \frac{\lambda(t)}{\pi} \int_0^\infty \frac{\cos\left(\sum_{i=0}^{n-1} \arctan\left(\frac{s}{h(i)}\right) - s \Lambda(t)\right)}{\prod_{i=0}^{n-1} \left(1 + \left(\frac{s}{h(i)}\right)^2\right)^{\frac{1}{2}}} ds,$$

alternatively, we can write

$$f_{T_n}(t) = \alpha_n \frac{\lambda(t) (\Lambda(t))^{n-1}}{(n-1)!} \Phi_2^{(n)}(1, \dots, 1; n; -h(0) \Lambda(t), \dots, -h(n-1) \Lambda(t)),$$

$$F_{T_n}(t) = \alpha_n \frac{(\Lambda(t))^n}{n!} \Phi_2^{(n)}(1, \dots, 1; n+1; -h(0) \Lambda(t), \dots, -h(n-1) \Lambda(t))$$

The previous result allows to compute the probability mass function of N_t as an immediate consequence, using the fact that

$$\mathbb{P}(N_t = n) = F_{T_n}(t) - F_{T_{n+1}}(t).$$

However, this expression does not seem easy to simplify. We now come to other probabilistic properties. As a first, we show that an EPP* has a similar "restarting property" as a GPP [Cha(2014)].

Proposition 7 ("Restarting property") Let $(\bar{N}_t)_{t \in [0, M_\Lambda]}$ be an EPP* with parameter $(\Lambda(\cdot), h(\cdot))$ and $u \in (0, M_\Lambda)$ be fixed. Set $N_t^{(u)} = N_{t+u} - N_u$ for all $t \in [0, M_\Lambda - u)$. Then, given $\mathcal{F}_{u-} = \sigma(N_s, s < u)$, the conditional process $(N_t^{(u)})_{t \in [0, M_\Lambda - u)}$ is an EPP* with parameter $(\Lambda(\cdot|u), h(\cdot|N_{u-}))$ where $\Lambda(t|u) = \Lambda(t+u) - \Lambda(u)$ for all $t \in [0, M_\Lambda - u)$ and $h(k|n) = h(n+k)$ for all $k \in \mathbb{N}$.

Proof Let us first consider the case where $(\bar{N}_t)_{t \in [0, \infty)}$ is an EPP $(\mathbf{1}, h(\cdot))$ with $\Lambda(t) = t$. In this case, it is a homogeneous pure-birth Markov process and the Markov property at time u entails that $(\bar{N}_t^{(u)})_{t \in [0, \infty)}$ is an EPP $(\mathbf{1}, h(\cdot|\bar{N}_{u-}))$.

To show the general case, recall that, due to Corollary 1, $(N_t := \bar{N}_{\Lambda(t)})_{t \in [0, M_\Lambda]}$ is an EPP* $(\Lambda(\cdot), h(\cdot))$. Also,

$$N_t^{(u)} = N_{t+u} - N_u = \bar{N}_{\Lambda(t+u) - \Lambda(u) + \Lambda(u)} - \bar{N}_{\Lambda(u)} = \bar{N}_{\Lambda(t|u)}^{(\Lambda(u))} \quad (20)$$

The restarting property for $(\bar{N}_t)_{t \in [0, \infty)}$ says us that $(\bar{N}_t^{(\Lambda(u))})_{t \in [0, \infty)}$ is an EPP $(\mathbf{1}, h(\cdot|\bar{N}_{\Lambda(u)-}))$. Noting that $\Lambda(\cdot|u)$ is a cumulative intensity function, we conclude by (20) and Corollary 1 again that $(N_t^{(u)})_{t \in [0, M_\Lambda - u)}$ is an EPP* $(\Lambda(\cdot|u), h(\cdot|N_{u-}))$.

The restarting property allows to easily obtain many probabilistic results (see [Cha(2014)] for more details). As an example, considering an EPP($\lambda(\cdot), h(\cdot)$), we can derive the following conditional p.d.f. for T_{n+1} given \mathcal{F}_{T_n} from (3):

$$h(n) \lambda(t_{n+1}) e^{-h(n)(A(t_{n+1})-A(T_n))} \quad (21)$$

for all $t_{n+1} \in (T_n, M_\lambda)$. The restarting property now allows to write down the joint density of the arrival times (T_1, T_2, \dots, T_n) as follows:

$$\begin{aligned} f(t_1, \dots, t_n) &= \prod_{i=1}^n \left(h(i-1) \lambda(t_i) e^{-h(i-1)(A(t_i)-A(t_{i-1}))} \right) \\ &= \prod_{i=1}^{n-1} \left(h(i-1) \lambda(t_i) e^{(h(i)-h(i-1))A(t_i)} \right) h(n-1) \lambda(t_n) e^{-h(n-1)A(t_n)} \quad (22) \end{aligned}$$

for all $0 = t_0 \leq t_1 < t_2 < \dots < t_n < M_\lambda$. Here, we recognize the p.d.f. for the GOS of an absolutely continuous F , when we choose A as given in (9) and $h(\cdot)$ as given in (10).

Remark 7 In [Cha(2014), Thm. 2] it was shown an unconditional restarting property for the generalized Pólya process, which says that if we begin to observe a GPP from an arbitrary point u (without knowledge of the past history), we also observe a GPP. This is not, in general, true for an EPP. We give here the reason, and for the sake of simplicity, we restrict ourselves to the case $A(t) = t$, that is, a pure-birth process. If we know that $N(u) = k$, the conditional restarting property (recall Proposition 7) says that, from this moment we observe a pure birth process with $h_k(\cdot) = h(k+\cdot)$, where h is the original pure-birth function. Now let us consider (T_1^u, \dots, T_n^u) the arrival times of the new process observed from u . Thus, applying (22), and the conditional restarting property, the density of this random vector is given as

$$\sum_{k=0}^{\infty} \prod_{i=1}^{n-1} \left(h_k(i-1) e^{(h_k(i)-h_k(i-1))t_i} \right) h_k(n-1) e^{-h_k(n-1)t_n} P(N(u) = k)$$

observe that, in general, this density has not the "multiplicative" expression given in (22) for an EPP (as a product of univariate t_i functions). This can be only true in the case that $h_k(i) - h_k(i-1)$ does not depend on k (a GPP process). In fact, in this specific case, the unconditional restarting property is true, as mentioned above.

Also, as a consequence of (22), and provided that the intensity λ exists, the conditional p.d.f. of (T_1, \dots, T_n) given $N_t = n$ is given in the following corollary. Note that in [Cha(2014)], the author already observed that in case of GPPs, this conditional p.d.f. was that of an ordinary order statistics (just as for Poisson processes). The following proposition shows that GPPs are the only EPPs for which this property is true.

Corollary 2 Let $(N_t)_{t \in [0, M_\Lambda]}$ be an EPP with parameter $(\lambda(\cdot), h(\cdot))$ and let $(T_n)_{n \in \mathbb{N}}$ be its arrival times. For $n \in \mathbb{N}^*$, the conditional p.d.f. of (T_1, \dots, T_n) given $N_t = n$ is provided by

$$f_{(T_1, \dots, T_n | N_t)}(t_1, \dots, t_n | n) = \frac{e^{-h(n)\Lambda(t)}}{\mathbb{P}(N_t = n)} \prod_{i=1}^n \left(h(i-1)\lambda(t_i) e^{(h(i)-h(i-1))\Lambda(t_i)} \right) \quad (23)$$

for all $0 \leq t_1 < t_2 < \dots < t_n \leq t$ and the conditional distribution of (T_1, \dots, T_n) given $N_t = n$ is **not** an order statistics unless $h(i) - h(i-1)$ is independent of i (constant), namely unless the process is a GPP ($h(i) = \alpha i + \beta$ for all $i \geq 0$).

Proof Expression (23) is a direct consequence of the restarting property (using a similar procedure as in (22)). Then, the conditional distribution of (T_1, \dots, T_n) given $N_t = n$ is an order statistics if and only if $\lambda(u) e^{(h(i)-h(i-1))\Lambda(u)}$ is of the shape $C_i f(u)$, which provides the result.

We now look at the convergence of a sequence $(N^{(k)})_{k \in \mathbb{N}}$ of EPP*'s. We restrict ourselves to the non explosive case, in order to use results concerning convergence of point processes [Jacod and Shiryaev(1987), Thm 3.37 page 354], which require a finite number of points in a bounded interval.

Proposition 8 Let $(h^{(k)}(\cdot))_{k \in \mathbb{N}}$ be a sequence of functions from \mathbb{N} to \mathbb{R}_+^* such that

$$\sum_{i=0}^{+\infty} \frac{1}{h^{(k)}(i)} = +\infty$$

for all $k \in \mathbb{N}$ (non explosion condition), and let $(\Lambda^{(k)}(\cdot))_{k \in \mathbb{N}}$ be a sequence of cumulative intensity functions such that $M_{\Lambda_k} = M_\Lambda$ for all $k \in \mathbb{N}$, with $M_\Lambda > 0$. Assume that:

- (A₁) $\lim_{k \rightarrow +\infty} h^{(k)} = h$, where $h(\cdot) : \mathbb{N} \rightarrow \mathbb{R}_+^*$ is such that $\sum_{i=0}^{+\infty} \frac{1}{h(i)} = +\infty$,
- (A₂) $\lim_{k \rightarrow +\infty} \Lambda^{(k)}(t) = \Lambda(t)$ for all $t \in [0, M_\Lambda]$, with Λ continuous and $\lim_{t \rightarrow M_\Lambda^-} \Lambda(t) = \infty$.

Then, the sequence $(N^{(k)})_{k \in \mathbb{N}}$ of EPP*'s with respective parameters $(\Lambda^{(k)}, h^{(k)})$ weakly converges towards an EPP* $N = (N_t)_{0 \leq t < M_\Lambda}$ with parameters (Λ, h) (namely in the Skorohod topology).

Proof Let us first observe that, under conditions of non explosion, N and $N^{(k)}$'s are finite-valued non decreasing point processes. This allows to use [Jacod and Shiryaev(1987), Thm 3.37 page 354], from which we know that it is enough to show the finite dimensional (fidi) convergence property, namely that $(N_{t_1}^{(k)}, \dots, N_{t_d}^{(k)})$ converges in distribution to $(N_{t_1}, \dots, N_{t_d})$ for all $d \geq 1$ and all $t_1 < \dots < t_d$, where symbol $^{(k)}$ refers to the EPP* $N^{(k)}$, as in all the remaining of this proof. This convergence will follow if we show the convergence of the corresponding probability

mass functions, due to Scheffe's Theorem [Billingsley(1995), Thm. 16.11]. In fact, we have

$$\begin{aligned}
& \mathbb{P}\left(N_{t_1}^{(k)} = n_1, \dots, N_{t_d}^{(k)} = n_d\right) \\
&= \mathbb{P}\left(N_{t_1}^{(k)} = n_1, N_{t_2}^{(k)} - N_{t_1}^{(k)} = n_2 - n_1, \dots, N_{t_d}^{(k)} - N_{t_{d-1}}^{(k)} = n_d - n_{d-1}\right) \\
&= \mathbb{P}\left(N_{t_1}^{(k)} = n_1\right) \prod_{j=2}^d \mathbb{P}\left(N_{t_j}^{(k)} - N_{t_{j-1}}^{(k)} = n_j - n_{j-1} \mid N_{t_1}^{(k)} = n_1, \dots, N_{t_{j-1}}^{(k)} = n_{j-1}\right) \\
&= \mathbb{P}\left(N_{t_1}^{(k)} = n_1\right) \prod_{j=2}^d \mathbb{P}\left(N_{t_j}^{(k)} - N_{t_{j-1}}^{(k)} = n_j - n_{j-1} \mid N_{t_{j-1}}^{(k)} = n_{j-1}\right)
\end{aligned}$$

for all $0 \leq t_1 < \dots < t_d \leq M_\Lambda$ and due to the restarting property, the term j in this product is the probability mass function of an EPP* with $\Lambda(t|t_{j-1}) = \Lambda^{(k)}(t + t_{j-1}) - \Lambda^{(k)}(t_{j-1})$ and $h^{(k)}(n|n_{j-1}) = h^{(k)}(n + n_{j-1})$. Thus, the problem is reduced to prove the unidimensional convergence of the probability mass function of $N_t^{(k)}$ towards that of N_t for $t < M_\Lambda$, when conditions (A1) and (A2) are satisfied. (Obviously, the "restarted" processes also verify these properties if the original does so). Now, assume the sequences $(T_n^{(k)})_{n \in \mathbb{N}}$ with $k \in \mathbb{N}$ to be constructed by (17) from the same uniform sequence $(U_n)_{n \in \mathbb{N}}$, with no loss of generality. Then, under (A1) and (A2), let us prove that

$$T_n^{(k)} \xrightarrow{D} T_n, \quad n = 1, 2, \dots \quad (24)$$

where $T_1, T_2, \dots, T_n, \dots$ are the points of the limiting EPP, and where \xrightarrow{D} denotes convergence in distribution. Let us first look at the case where $\Lambda^{(k)}(t) = t$, $k = 1, 2, \dots$ for all $t \geq 0$. For $k = 1, 2, \dots$, based on (17), the $S_n^{(k)}$'s can be recursively constructed through:

$$S_1^{(k)} = -\frac{\ln(U_0)}{h^{(k)}(n)}, \quad (25)$$

$$S_{n+1}^{(k)} = S_n^{(k)} - \frac{\ln(U_n)}{h^{(k)}(n)} \quad (26)$$

for all $n \in \mathbb{N}$. Thus, it is easy to see that

$$S_n^{(k)} \xrightarrow{\text{a.s.}} S_n, \quad n = 1, 2, \dots, \quad (27)$$

where $\xrightarrow{\text{a.s.}}$ denotes almost sure convergence. In fact, (27) for $n = 1$ is clear from (25) and the general case follows from (26) and induction. Now, take $\Lambda^{(k)}(t)$ arbitrary. Using Corollary 1, we have

$$\lim_{k \rightarrow \infty} \mathbb{P}(T_n^{(k)} \leq t) = \lim_{k \rightarrow \infty} \mathbb{P}\left(S_n^{(k)} \leq \Lambda^{(k)}(t)\right) = \mathbb{P}(S_n \leq \Lambda(t)) = \mathbb{P}(T_n \leq t), \quad (28)$$

in which, in the second equality we have used (27), as this implies that

$$\frac{S_n^{(k)}}{\Lambda^{(k)}(t)} \xrightarrow{\text{a.s.}} \frac{S_n}{\Lambda(t)},$$

and the almost sure convergence implies convergence in distribution. Thus, (24) follows from (28) .

Recalling that

$$\mathbb{P}\left(N_t^{(k)} \leq n\right) = \mathbb{P}\left(T_{n+1}^{(k)} > t\right), n = 0, 1, 2, \dots \quad (29)$$

we deduce from this expression and (27)

$$\lim_{k \rightarrow \infty} \mathbb{P}\left(N_t^{(k)} \leq n\right) = \mathbb{P}\left(N_t \leq n\right)$$

and this shows that $N_t^{(k)}$ converges in distribution towards N_t for all $0 \leq t < M_\Lambda$, thus completing the proof.

Example 3 Let $h^{(k)}(n) = (q_k)^n$ for all $n \in \mathbb{N}$ where $0 \leq q_k \leq 1$ for all $k \in \mathbb{N}$ and $\lim_{k \rightarrow +\infty} q_k = q \in [0, 1]$. We set $h(n) = q^n$ for all $n \in \mathbb{N}$. Then, under assumption (A_2) from the proposition, the sequence $\left(N^{(k)}\right)_{k \in \mathbb{N}}$ weakly converges towards an $\text{EPP}^*(\Lambda, h)$.

Example 4 Let $h^{(k)}(n) = \alpha^{(k)}n + \beta^{(k)}$ where $\alpha^{(k)} \geq 0$, $\beta^{(k)} > 0$, $\lim_{k \rightarrow +\infty} \alpha^{(k)} = \alpha \geq 0$ and $\lim_{k \rightarrow +\infty} \beta^{(k)} = \beta > 0$. Then, under assumption (A_2) , the sequence $\left(N^{(k)}\right)_{k \in \mathbb{N}}$ weakly converges towards an $\text{EPP}^*(\Lambda, h)$, where $h(n) = \alpha n + \beta$.

Remark 8 Note that we can also use the previous proposition to obtain approximations for probabilities in an EPP^* with $h(j)$ not necessarily all distinct by means of EPP^* 's with $h(j)$ all distinct and therefore, we can approximate the expressions given in Proposition 6 by linear combinations of exponentials as in Proposition 5. For instance, let us consider a Pure-Birth process in which $h(n) = \lambda$ for $n = 0, 1, \dots, m$ and $h(n) = \beta$ for $n = m + 1, m + 2, \dots$ ($\lambda \neq \beta$). We can approximate this process by a pure-birth process in which $h^{(\epsilon)}(n) = \lambda + n\epsilon$ for $n = 0, 1, \dots, m$ and $h^{(\epsilon)}(n) = \beta + n\epsilon$ for $n = m + 1, m + 2, \dots$ (ϵ small enough) and use Propositions 4 and 5 to evaluate the corresponding probabilities. In this way, we could give approximations both to the expressions given in Proposition 6 and to the integral expressions given in [Janardan(2005)] for this specific case.

3 Comparison results, aging and positive dependence properties

In this section, we establish stochastic properties of distributions related to an EPP^* . Remembering that EPP 's are specific non-homogeneous pure-birth processes and that the arrival times in an EPP^* contain Generalized Order Statistics (GOS), some of our results extend previous ones from the literature (in our specific setting of EPP^* 's) devoted either to non-homogeneous pure-birth processes [Belzunce et al(2001)] or GOS [Belzunce et al(2005)]. Also, some of our results extend to EPP^* 's some previous ones obtained in [Badía et al(2018)] and [Cha and Finkelstein(2017)] in the context of Generalized Pólya processes.

Lemma 2 *Let us consider two EPP^* 's $\left(N_t^{(i)}\right)_{t \in [0, M_\Lambda]}$, $i = 1, 2$ with parameters $(h^{(i)}, \Lambda)$, $i = 1, 2$ such that $h^{(1)}(\cdot) \leq h^{(2)}(\cdot)$. Let $(T_n^{(i)})_{n=1,2,\dots}$, $i = 1, 2$ be their respective arrival times. Then $T_n^{(1)} \geq_{st} T_n^{(2)}$ for all $n \in \mathbb{N}^*$.*

Proof Let us show the result recursively on n .

Based on (2), we have $\mathbb{P}(T_1^{(1)} > t) \geq \mathbb{P}(T_1^{(2)} > t)$ for all $t > 0$, and the result follows for $n = 1$.

Assume $T_n^{(1)} \geq_{st} T_n^{(2)}$ for some $n \geq 1$ and let $t > 0$. Starting from (3), we have:

$$\begin{aligned} \mathbb{P}\left(T_{n+1}^{(1)} > t\right) &= \mathbb{E}\left[e^{-h_1(n)(\Lambda(t)-\Lambda(T_n^{(1)}))^+}\right] \\ &\geq \mathbb{E}\left[e^{-h_2(n)(\Lambda(t)-\Lambda(T_n^{(1)}))^+}\right] \end{aligned}$$

because $h^{(1)} \leq h^{(2)}$.

Based on the non-decreasingness of $e^{-h_2(n)(\Lambda(t)-\Lambda(u))^+}$ with respect to u and on $T_n^{(1)} \geq_{st} T_n^{(2)}$, we derive that

$$\mathbb{P}\left(T_{n+1}^{(1)} > t\right) \geq \mathbb{E}\left[e^{-h_2(n)(\Lambda(t)-\Lambda(T_n^{(2)}))^+}\right]$$

and the result.

Proposition 9 *If h is increasing, then*

$$(\bar{M}_1 := N_{t_1}, \bar{M}_2 := N_{t_2} - N_{t_1}, \dots, \bar{M}_n := N_{t_n} - N_{t_{n-1}}), \quad t_1 < t_2 < \dots < t_n < M_\Lambda$$

is conditionally increasing in sequence. If h is decreasing, the previous vector is conditionally decreasing in sequence.

Proof Let us denote by T_i^m the i -th arrival time in the process $(N_t^{(u)})_{t \in [0, M_\Lambda - u]}$, given that $N_u = m$, which is an EPP* on $[0, M_\Lambda - u]$ with parameters $(\Lambda(\cdot|u), h(\cdot|m))$. If h is increasing, we have for $m \leq m'$ that

$$h(\cdot|m) = h(\cdot + m) \leq h(\cdot + m') = h(\cdot|m').$$

Based on Lemma 2, we derive that $T_i^{m'} \leq_{st} T_i^m$ (the $\Lambda(\cdot|u)$ is the same). Therefore, consider $k = 1, \dots, n-1$ and $(i_1, \dots, i_k) \leq (j_1, \dots, j_k)$. If we denote by $I_k = i_1 + \dots + i_k$ and $J_k = j_1 + \dots + j_k$, we have with $u = t_k$:

$$\begin{aligned} \mathbb{P}(\bar{M}_{k+1} > i | \bar{M}_1 = i_1, \bar{M}_2 = i_2, \dots, \bar{M}_k = i_k) &= \mathbb{P}(T_{i+1}^{I_k} \leq t_{k+1} - t_k) \\ &\leq \mathbb{P}(T_{i+1}^{J_k} \leq t_{k+1} - t_k) = \mathbb{P}(\bar{M}_{k+1} > i | \bar{M}_1 = j_1, \bar{M}_2 = j_2, \dots, \bar{M}_k = j_k) \end{aligned}$$

Thus $(\bar{M}_1, \dots, \bar{M}_n)$ is conditionally increasing in sequence. If h is decreasing, the previous inequalities are reversed, and therefore $(\bar{M}_1, \dots, \bar{M}_n)$ are conditionally decreasing in sequence.

Remark 9 Notice that $(\bar{M}_1, \dots, \bar{M}_n)$ conditionally increasing in sequence implies that this vector is associated. In particular the increments in such EPP are positive upper orthant dependent and positive lower orthant dependent, that is:

$$\begin{aligned} \mathbb{P}(\bar{M}_1 > x_1, \dots, \bar{M}_n > x_n) &\geq \prod_{i=1}^n \mathbb{P}(\bar{M}_i > x_i) \text{ for all } (x_1, \dots, x_n) \in \mathbb{R}^n \\ \mathbb{P}(\bar{M}_1 \leq x_1, \dots, \bar{M}_n \leq x_n) &\geq \prod_{i=1}^n \mathbb{P}(\bar{M}_i \leq x_i) \text{ for all } (x_1, \dots, x_n) \in \mathbb{R}^n \end{aligned}$$

Note that the positive lower orthant dependency was already established in [Cha and Finkelstein(2017)] in the specific case of a GPP process. Moreover, (M_1, \dots, M_n) conditionally decreasing in sequence implies that this vector is negative upper orthant dependent and negative lower orthant dependent (that is, the previous inequalities are reversed).

An analogous property to Proposition 9 can be shown for the interarrival times in an EPP, but in this case, using shape properties of Λ .

Proposition 10 *Let $X_i := T_i - T_{i-1}$ be the interarrival times in an EPP*. If Λ is concave, then (X_1, \dots, X_n) is conditionally increasing in sequence. If Λ is convex, the previous vector is conditionally decreasing in sequence.*

Proof Using the restarting property, it is clear that

$$\mathbb{P}(X_{n+1} > x | (X_1, \dots, X_n)) = e^{-h(n)(\Lambda(T_n+x) - \Lambda(T_n))} \quad (30)$$

where $T_n = X_1 + \dots + X_n$. Thus, if $\Lambda(x)$ is concave (convex) the previous expression shows that X_{n+1} is stochastically increasing (decreasing) in (X_1, \dots, X_n) , thus providing the result.

As we can see, the increments (inter-arrival times) in an EPP* have positive or negative dependence properties, according to the monotonicity (shape) of $h(\Lambda)$. However, in [Belzunce et al(2003)], the authors show the so-called Multivariate Increasing Failure Rate property for the jump times (T_1, \dots, T_n) of a general non-homogeneous pure-birth process, which includes EPP*. Remembering that Multivariate Increasing Failure Rate property is a positive dependence property implying CIS, this shows that the successive jump (arrival) times of an EPP* are CIS, and hence positively dependent. We next show that they exhibit the stronger Multivariate Totally Positive property of order 2 (MTP2), which is known to imply CIS. For sake of completeness, we recall that a function $f : \mathbb{R}^n \mapsto \mathbb{R}_+$ is said to be MTP2 as soon as

$$f(\mathbf{x})f(\mathbf{y}) \leq f(\mathbf{x} \vee \mathbf{y})f(\mathbf{x} \wedge \mathbf{y}), \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

where \vee and \wedge are the max and min component-wise operations, respectively.

Proposition 11 *The vector (T_1, \dots, T_n) of arrival times in an EPP is MTP2.*

Proof Let

$$A := \{(t_1, \dots, t_n) \in \mathbb{R}_+^n | 0 \leq t_1 < t_2 < \dots < t_n\}. \quad (31)$$

Recalling (22), the joint density function for this vector is given by

$$\mathbf{1}_A(t_1, \dots, t_n) \prod_{i=1}^{n-1} \left(h(i-1) \lambda(t_i) e^{(h(i)-h(i-1))\Lambda(t_i)} \right) h(n-1) \lambda(t_n) e^{-h(n-1)\Lambda(t_n)}$$

The functions inside the product are all MTP2, because they are unidimensional. On the other hand, as A is a lattice (stable by minimum and maximum), then $\mathbf{1}_A$ is clearly MTP2. The conclusion follows as the product of MTP2 functions is also MTP2.

Remark 10 In a general EPP*, the vector (T_1, \dots, T_n) is known to be Conditionally Increasing (CI, see [Müller and Stoyan(2002), p. 125]), which is a stronger property than CIS. This can be easily derived from the previous result as the MTP2 property for (S_1, \dots, S_n) implies their CI property. The CI property for (T_1, \dots, T_n) now follows from $T_i = A^{-1}(S_i)$ and the preservation of the CI property under increasing transforms.

The multivariate likelihood ratio order for the arrival times in a general pure-birth process was studied in Belzunce et al. [Belzunce et al(2001)]. In the next result, our aim is to provide a specific study in the case of EPPs. Let us first recall the expression for the density of (T_1, \dots, T_n) given in (22):

$$f(t_1, \dots, t_n) = \prod_{i=1}^n f_i(t_i), \quad 0 < t_1 < t_2 < \dots < t_n < M_\lambda, \quad (32)$$

(0 otherwise), where

$$f_i(x) := h(i-1)\lambda(x)e^{(h(i)-h(i-1))\Lambda(x)}, \quad i = 1, \dots, n-1 \quad (33)$$

$$f_n(x) := h(n-1)\lambda(x)e^{-h(n-1)\Lambda(x)} \quad (34)$$

for $x \in (0, M_\lambda)$ (0 otherwise). Note that for each $i = 1, 2, \dots, n$, the support of f_i is the same as for λ . Also, remembering the definition of the set D given in (14) and considering the first interval $(a_1, b_1]$ in D (if any), let us introduce

$$m_\lambda = \begin{cases} b_1 & \text{if } a_1 = 0, \\ 0 & \text{otherwise.} \end{cases} \quad (35)$$

Then the function λ is almost surely zero on $[0, m_\lambda]$ and its common support with all $f_i, i = 1, \dots, n$ can be considered to be included in $[m_\lambda, M_\lambda]$.

4 Parametric estimation procedure

The classical maximum likelihood method is here considered in a parametric setting, which is tested on a few simulated data sets. We assume that m independent trajectories of an EPP $(N_t)_{t \geq 0}$ with parameters $(\lambda(\cdot), h(\cdot))$ are observed. For each $1 \leq j \leq m$, the j -th path is observed on $[0, t_f^{(j)}]$, which provides an observation $(n^{(j)}, t_1^{(j)}, \dots, t_{n^{(j)}}^{(j)})$ of $(N_{t_f^{(j)}}, T_1, \dots, T_{N_{t_f^{(j)}}})$. The data are denoted by (\mathbf{n}, \mathbf{t}) with $\mathbf{n} = (n^{(j)})_{1 \leq j \leq m}$ and $\mathbf{t} = (t_i^{(j)})_{1 \leq j \leq m}$ with $t_i^{(j)} = (t_i^{(j)})_{1 \leq i \leq n^{(j)}+1}$, where we set $t_{n^{(j)}+1}^{(j)} = t_f^{(j)}$ for ease of notation.

Setting Θ to be the parameters to estimate and

$$\bar{\lambda}_s^{(j)} = h(n_s^{(j)})\lambda(s),$$

$$\bar{\Lambda}_s^{(j)} = \int_0^s \bar{\lambda}_u^{(j)} du = \sum_{i=0}^{n_s^{(j)}-1} h(i) \left(\Lambda(t_{i+1}^{(j)}) - \Lambda(t_i^{(j)}) \right) + h(n_s^{(j)}) \left(\Lambda(s) - \Lambda(t_{n_s^{(j)}}^{(j)}) \right),$$

for all $s > 0$ and $j = 1, \dots, m$ (where $n_s^{(j)}$ refers to the j -th observation of N_s), the likelihood function is

$$\mathcal{L}(\Theta|\mathbf{n}, \mathbf{t}) = \prod_{j=1}^m \left(\prod_{i=1}^{n^{(j)}} \bar{\lambda}_{t_i^{(j)}}^{(j)} \right) e^{-\bar{\Lambda}_{t_f^{(j)}}^{(j)}},$$

see, e.g., [Daley and Vere-Jones(2003), Prop. 7.2.III.]. This provides the following log-likelihood function:

$$\begin{aligned} l(\Theta|\mathbf{n}, \mathbf{t}) &= \log(\mathcal{L}(\Theta|\mathbf{n}, \mathbf{t})) \\ &= \sum_{j=1}^m \sum_{i=1}^{n^{(j)}} \log\left(\bar{\lambda}_{t_i^{(j)}}^{(j)}\right) - \sum_{j=1}^m \bar{\Lambda}_{t_f^{(j)}}^{(j)} \\ &= \sum_{j=1}^m \sum_{i=1}^{n^{(j)}} \log\left(h(i-1)\lambda\left(t_i^{(j)}\right)\right) - \sum_{j=1}^m \left(\sum_{i=0}^{n^{(j)}} h(i)\left(\Lambda\left(t_{i+1}^{(j)}\right) - \Lambda\left(t_i^{(j)}\right)\right)\right) \end{aligned}$$

remembering that $t_{n^{(j)}+1}^{(j)} = t_f^{(j)}$.

Considering $\Lambda(t) = at^b$ for $t \in \mathbb{R}^+$, we get:

$$\begin{aligned} l(a, b, \theta|\mathbf{n}, \mathbf{t}) &= \sum_{j=1}^m \sum_{i=1}^{n^{(j)}} \log(h(i-1|\theta)) + (\log(a) + \log(b)) \sum_{j=1}^m n^{(j)} \\ &\quad + (b-1) \sum_{j=1}^m \sum_{i=1}^{n^{(j)}} \log\left(t_i^{(j)}\right) - a \sum_{j=1}^m \left(\sum_{i=0}^{n^{(j)}} h(i|\theta) \left(\left(t_{i+1}^{(j)}\right)^b - \left(t_i^{(j)}\right)^b\right)\right). \end{aligned}$$

where θ stands for the parameter(s) of h .

Solving $\frac{\partial l(a, b, \theta|\mathbf{n}, \mathbf{t})}{\partial a} = 0$ provides $a = \hat{a}(b, \theta)$ with

$$\hat{a}(b, \theta) = \frac{\sum_{j=1}^m n^{(j)}}{\sum_{j=1}^m \left(\sum_{i=0}^{n^{(j)}} h(i|\theta) \left(\left(t_{i+1}^{(j)}\right)^b - \left(t_i^{(j)}\right)^b\right)\right)}.$$

Maximizing

$$\begin{aligned} &l(\hat{a}(b, \theta), b, \theta|\mathbf{n}, \mathbf{t}) \\ &\propto g(\theta|\mathbf{n}, \mathbf{t}) - \left(\sum_{j=1}^m n^{(j)}\right) \log\left(\sum_{j=1}^m \left(\sum_{i=0}^{n^{(j)}} h(i|\theta) \left(\left(t_{i+1}^{(j)}\right)^b - \left(t_i^{(j)}\right)^b\right)\right)\right) \\ &\quad + \log(b) \sum_{j=1}^m n^{(j)} + (b-1) \sum_{j=1}^m \sum_{i=1}^{n^{(j)}} \log\left(t_i^{(j)}\right) \end{aligned}$$

(where \propto means equal up to a constant) with

$$g(\theta|\mathbf{n}, \mathbf{t}) = \sum_{j=1}^m \sum_{i=1}^{n^{(j)}} \log(h(i-1|\theta))$$

Table 1 Estimation results, case $\Lambda(t) = t^{1.5}$, $h(n) = q^n$ with $q \in \{0.5, 0.75, 0.9\}$ and $[t_0, t_1] = [5, 10]$.

m	\hat{n}_f	$a = 1$		$b = 1.5$		$q \in \{0.5, 0.75, 0.9\}$		
		Mean(std)	95% IC	Mean(std)	95% IC	Mean(std)	95% IC	
10	4.11	1.051(0.343)	[0.379, 1.723]	1.639(0.307)	[1.037, 2.241]	0.5	0.461(0.095)	[0.275, 0.647]
40	4.11	1.002(0.140)	[0.728, 1.276]	1.533(0.123)	[1.292, 1.774]		0.490(0.043)	[0.406, 0.574]
160	4.12	1.000(0.068)	[0.867, 1.133]	1.506(0.065)	[1.379, 1.633]		0.498(0.022)	[0.455, 0.541]
10	6.84	0.991(0.276)	[0.450, 1.532]	1.615(0.265)	[1.096, 2.134]	0.75	0.722(0.070)	[0.585, 0.859]
40	6.84	1.005(0.135)	[0.740, 1.270]	1.532(0.125)	[1.287, 1.777]		0.741(0.035)	[0.672, 0.810]
160	6.84	0.998(0.067)	[0.867, 1.129]	1.508(0.060)	[1.390, 1.626]		0.748(0.017)	[0.715, 0.781]
10	10.93	0.989(0.259)	[0.481, 1.497]	1.597(0.222)	[1.162, 2.032]	0.9	0.880(0.042)	[0.798, 0.962]
40	10.98	0.991(0.127)	[0.742, 1.240]	1.522(0.109)	[1.308, 1.736]		0.897(0.020)	[0.858, 0.936]
160	10.96	1.000(0.065)	[0.873, 1.127]	1.508(0.055)	[1.400, 1.616]		0.898(0.010)	[0.878, 0.918]

next provides $(\hat{b}, \hat{\theta})$. We finally set $\hat{a} = \hat{a}(\hat{b}, \hat{\theta})$.

We take $h(n) = q^n$ with $q \in [0, 1]$ as a first case (which provides an identifiable model, see Remark 2), so that $\theta = q$ and

$$g(q|\mathbf{n}, \mathbf{t}) = \frac{\log(q)}{2} \sum_{j=1}^m n^{(j)} (n^{(j)} - 1).$$

Numerical experiments are performed, where $(t_f^{(j)})_{1 \leq j \leq m}$ are chosen as i.i.d. observations of a uniform r.v. on $[t_0, t_1]$. Also, $R = 500$ i.i.d. sets of m i.i.d. paths are generated, which provides R estimates. The corresponding empirical mean and square root of the empirical variance (std) are reported in Tables 1 and 2, together with an approximate 95% (normal-based) confidence interval (IC). Finally, the mean number of jumps observed per trajectory (denoted by \hat{n}_f) is also reported, to have an idea of the size of a data set (which corresponds to m trajectories). The only difference between the two tables is that $b = 1.5$ ($\Lambda(\cdot)$ is convex) in Table 1 whereas $b = 0.9$ ($\Lambda(\cdot)$ is concave) in Table 2. As expected, the concave case requires a longer observation period than the convex case to have enough data to get reliable results. Also, we can see that multiplying the number m of observed trajectories by 4 ($40 = 10 \times 4$; $160 = 40 \times 4$) divides the standard deviation by 2, which shows a root-of- m convergence in coherence with a likely asymptotic normality (which is beyond the scope of the present paper). Finally, the bias decreases with the size of the data set.

We now take $h(n) = (1+n)^\alpha$ with $\alpha > 0$ (identifiable model, see Remark 2), so that $\theta = \alpha$ and

$$g(\alpha|\mathbf{n}, \mathbf{t}) = \alpha \sum_{j=1}^m \log(n^{(j)}!).$$

Estimation results are given in Tables 3 and 4, with similar observations as for Tables 1 and 2. The estimation procedure hence seems to behave well, which allows EEPs to be used in an application context.

Table 2 Estimation results, case $\Lambda(t) = t^{0.9}$, $h(n) = q^n$ with $q \in \{0.5, 0.75, 0.9\}$ and $[t_0, t_1] = [15, 25]$.

m	\hat{n}_f	$a = 1$		$b = 0.9$		$q \in \{0.5, 0.75, 0.9\}$		
		Mean(std)	95% IC	Mean(std)	95% IC	Mean(std)	95% IC	
10	3.68	1.027(0.314)	[0.412, 1.642]	0.990(0.196)	[0.606, 1.374]	0.5	0.457(0.102)	[0.257, 0.657]
40	3.69	1.008(0.147)	[0.720, 1.296]	0.918(0.082)	[0.757, 1.079]		0.491(0.051)	[0.391, 0.591]
160	3.70	1.005(0.072)	[0.864, 1.146]	0.906(0.039)	[0.830, 0.982]		0.496(0.026)	[0.445, 0.547]
10	5.89	1.011(0.287)	[0.448, 1.574]	0.979(0.163)	[0.660, 1.298]	0.75	0.710(0.079)	[0.555, 0.865]
40	5.90	1.003(0.132)	[0.744, 1.262]	0.916(0.074)	[0.771, 1.061]		0.741(0.039)	[0.665, 0.817]
160	5.91	1.005(0.067)	[0.874, 1.136]	0.901(0.036)	[0.830, 0.972]		0.748(0.018)	[0.713, 0.783]
10	8.99	0.981(0.258)	[0.475, 1.487]	0.970(0.157)	[0.662, 1.278]	0.9	0.876(0.058)	[0.762, 0.990]
40	9.01	0.999(0.131)	[0.742, 1.256]	0.913(0.064)	[0.788, 1.038]		0.896(0.024)	[0.849, 0.943]
160	8.97	0.999(0.063)	[0.876, 1.122]	0.903(0.033)	[0.838, 0.968]		0.899(0.012)	[0.875, 0.923]

Table 3 Estimation results, case $\Lambda(t) = t^{1.5}$, $h(n) = (1+n)^\alpha$ with $\alpha \in \{0.5, 0.75, 1\}$ and $[t_0, t_1] = [1.5, 2]$.

m	\hat{n}_f	$a = 1$		$b = 1.5$		$\alpha \in \{0.5, 0.75, 1\}$		
		Mean(std)	95% IC	Mean(std)	95% IC	Mean(std)	95% IC	
10	3.47	1.032(0.334)	[0.377, 1.687]	1.642(0.380)	[0.897, 2.387]	0.5	0.387(0.281)	[0.000, 0.938]
40	3.49	1.016(0.160)	[0.702, 1.330]	1.530(0.176)	[1.185, 1.875]		0.461(0.156)	[0.155, 0.767]
160	3.51	0.998(0.072)	[0.857, 1.139]	1.508(0.086)	[1.339, 1.677]		0.496(0.072)	[0.355, 0.637]
10	5.08	1.056(0.322)	[0.425, 1.687]	1.663(0.381)	[0.916, 2.410]	0.75	0.630(0.262)	[0.116, 1.144]
40	5.04	1.012(0.146)	[0.726, 1.298]	1.539(0.164)	[1.218, 1.860]		0.725(0.114)	[0.502, 0.948]
160	5.01	1.002(0.069)	[0.867, 1.137]	1.513(0.082)	[1.352, 1.674]		0.741(0.056)	[0.631, 0.851]
10	9.60	1.017(0.250)	[0.527, 1.507]	1.631(0.336)	[0.972, 2.290]	1	0.935(0.157)	[0.627, 1.243]
40	9.63	1.005(0.128)	[0.754, 1.256]	1.531(0.166)	[1.206, 1.856]		0.983(0.072)	[0.842, 1.124]
160	9.68	1.008(0.061)	[0.888, 1.128]	1.505(0.074)	[1.360, 1.650]		0.995(0.031)	[0.934, 1.056]

Table 4 Estimation results, case $\Lambda(t) = t^{0.9}$, $h(n) = (1+n)^\alpha$ with $\alpha \in \{0.5, 0.75, 1\}$ and $[t_0, t_1] = [2.5, 3]$.

m	\hat{n}_f	$a = 1$		$b = 0.9$		$\alpha \in \{0.5, 0.75, 1\}$		
		Mean(std)	95% IC	Mean(std)	95% IC	Mean(std)	95% IC	
10	3.92	1.067(0.317)	[0.446, 1.688]	0.981(0.217)	[0.556, 1.406]	0.5	0.373(0.262)	[0.000, 0.887]
40	3.83	1.011(0.152)	[0.713, 1.309]	0.924(0.105)	[0.718, 1.130]		0.464(0.141)	[0.188, 0.740]
160	3.84	1.000(0.077)	[0.849, 1.151]	0.905(0.050)	[0.807, 1.003]		0.496(0.070)	[0.359, 0.633]
10	5.56	1.040(0.319)	[0.415, 1.665]	0.983(0.241)	[0.511, 1.455]	0.75	0.652(0.243)	[0.176, 1.128]
40	5.56	1.005(0.143)	[0.725, 1.285]	0.919(0.094)	[0.735, 1.103]		0.729(0.104)	[0.525, 0.933]
160	5.58	1.002(0.064)	[0.877, 1.127]	0.904(0.044)	[0.818, 0.990]		0.745(0.052)	[0.643, 0.847]
10	11.05	0.993(0.266)	[0.472, 1.514]	0.982(0.222)	[0.547, 1.417]	1	0.950(0.151)	[0.654, 1.246]
40	11.13	1.007(0.123)	[0.766, 1.248]	0.917(0.086)	[0.748, 1.086]		0.985(0.063)	[0.862, 1.108]
160	11.05	1.001(0.060)	[0.883, 1.119]	0.905(0.043)	[0.821, 0.989]		0.995(0.030)	[0.936, 1.054]

Acknowledgements This work has been supported by Spanish government research project MTM2015-63978(MINECO/FEDER). The first and third authors acknowledge the support of DGA S11 and E64, respectively.

References

[Aalo et al(2005)] Aalo V, Piboongunon T, Efthymoglou G (2005) Another look at the performance of MRC schemes in Nakagami-m fading channels with arbitrary parameters.

- IEEE Transactions on Communications 53(12):2002–2005
- [Arias-Nicolás et al(2005)] Arias-Nicolás JP, Fernández-Ponce JM, Luque-Calvo P, Suárez-Llorens A (2005) Multivariate dispersion order and the notion of copula applied to the multivariate t-distribution. *Probability in the Engineering and Informational Sciences* 19(3):363–375
- [Asfaw and Lindqvist(2015)] Asfaw ZG, Lindqvist BH (2015) Extending minimal repair models for repairable systems: A comparison of dynamic and heterogeneous extensions of a nonhomogeneous Poisson process. *Reliability Engineering & System Safety* 140:53 – 58
- [Asmussen(2003)] Asmussen S (2003) *Applied probability and queues*, Applications of Mathematics, vol 51, 2nd edn. Springer-Verlag, New York
- [Babykina and Couallier(2010)] Babykina G, Couallier V (2010) *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance*, Birkhäuser Boston, Boston, MA, chap Modelling Recurrent Events for Repairable Systems Under Worse Than Old Assumption, pp 339–354
- [Badía et al(2018)] Badía F, Sangüesa C, Cha J (2018) Stochastic comparisons and ageing properties of generalized Pólya processes. *Journal of Applied Probability* 55(1):233–253
- [Bedbur and Kamps(2017)] Bedbur S, Kamps U (2017) Inference in a two-parameter generalized order statistics model. *Statistics* 51(5):1132–1142
- [Belzunce et al(2001)] Belzunce F, Lillo RE, Ruiz JM, Shaked M (2001) Stochastic comparisons of nonhomogeneous processes. *Probability in the Engineering and Informational Sciences* 15:199–224
- [Belzunce et al(2003)] Belzunce F, Mercader JA, Ruiz JM (2003) Multivariate aging properties of epoch times of nonhomogeneous processes. *Journal of Multivariate Analysis* 84:335–350
- [Belzunce et al(2005)] Belzunce F, Mercader J, Ruiz J (2005) Stochastic comparisons of generalized order statistics. *Probability in the Engineering and Informational Sciences* 19:99–120
- [Belzunce et al(2012)] Belzunce F, Suárez-Llorens A, Sordo MA (2012) Comparison of increasing directionally convex transformations of random vectors with a common copula. *Insurance: Mathematics and Economics* 50:385–390
- [Belzunce et al(2016)] Belzunce F, Martínez-Riquelme C, Mulero J (2016) *An Introduction to Stochastic Orders*, 1st edn. Academic Press
- [Billingsley(1995)] Billingsley P (1995) *Probability and Measure*, 3rd edn. Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York
- [Bordes and Mercier(2013)] Bordes L, Mercier S (2013) Extended geometric processes: semi-parametric estimation and application to reliability. *Journal of the Iranian Statistical Society* 12(1):1–34
- [Boyer and Roux(2016)] Boyer M, Roux P (2016) A common framework embedding network calculus and event stream theory, URL <https://hal.archives-ouvertes.fr/hal-01311502>, working paper or preprint
- [Cha and Finkelstein(2016)] Cha J, Finkelstein M (2016) Justifying the Gompertz curve of mortality via the generalized Pólya process of shocks. *Theoretical Population Biology* 109:54–62
- [Cha(2014)] Cha JH (2014) Characterization of the generalized Pólya process and its applications. *Advances in Applied Probability* 46:1148–1171
- [Cha and Finkelstein(2017)] Cha JH, Finkelstein M (2017) New shock models based on the generalized Pólya process. *European Journal of Operations Research* 251:1148–1171
- [Chauvel et al(2016)] Chauvel C, Dauxois JY, Doyen L, Gaudoin O (2016) Parametric bootstrap goodness-of-fit tests for imperfect maintenance models. *IEEE Transactions on Reliability* 65(3):1343–1359
- [Cox(1970)] Cox D (1970) *Renewal Theory*, 1st edn. Methuen & Co.
- [Daley and Vere-Jones(2003)] Daley DJ, Vere-Jones D (2003) *An introduction to the theory of point processes - Volume I: General theory and structure*, second edition edn. Springer Science & Business Media
- [Doyen and Gaudoin(2004)] Doyen L, Gaudoin O (2004) Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering & System Safety* 84(1):45–56
- [Efthymoglou and Aalo(1995)] Efthymoglou G, Aalo V (1995) Performance of RAKE receivers in Nakagami fading channel with arbitrary fading parameters. *Electronics Letters* 31(18):1610–1612
- [Embrechts and Hofert(2013)] Embrechts P, Hofert M (2013) A note on generalized inverses. *Mathematical Methods of Operations Research* 77(3):423–432

- [Gikhman and Skorokhod(1969)] Gikhman II, Skorokhod AV (1969) Introduction to the theory of random processes. Translated from the Russian by Scripta Technica, Inc, W. B. Saunders Co., Philadelphia, Pa.-London-Toronto, Ont.
- [Jacod and Shiryaev(1987)] Jacod J, Shiryaev AN (1987) Limit theorems for stochastic processes, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol 288. Springer-Verlag, Berlin
- [Janardan(2005)] Janardan K (2005) Integral representation of a distribution associated with a pure birth process. *Communications in Statistics: Theory & Methods* 34(11):2097-2103
- [Kamps(1995a)] Kamps U (1995a) A concept of generalized order statistics. *Journal of Statistical Planning and Inference* 48(11):1-23
- [Kamps(1995b)] Kamps U (1995b) A concept of generalized order statistics. Teubner Skripten zur Mathematischen Stochastik. [Teubner Texts on Mathematical Stochastics], B. G. Teubner, Stuttgart
- [Kijima(1989)] Kijima M (1989) Some results for repairable systems with general repair. *Journal of Applied Probability* 26(1):89-102
- [Konno(2010)] Konno H (2010) On the exact solution of a generalized Pólya process. *Advances in mathematical physics* 2010, article ID 504267
- [Lam(2007)] Lam Y (2007) The geometric process and its applications. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ
- [Landriault et al(2014)] Landriault D, Willmot GE, Xu D (2014) On the analysis of time dependent claims in a class of birth process claim count models. *Insurance: Mathematics and Economics* 58:168-173
- [Le Gat(2009)] Le Gat Y (2009) Une extension du processus de Yule pour la modélisation stochastique des événements récurrents. Application aux défaillances de canalisations d'eau sous pression. PhD thesis, Ecole Nationale du Génie Rural des Eaux et des Forêts (EN-GREF), Paris
- [Le Gat(2014)] Le Gat Y (2014) Extending the Yule process to model recurrent pipe failures in water supply networks. *Urban Water Journal* 11(8):617-630
- [Müller and Stoyan(2002)] Müller A, Stoyan D (2002) Comparison methods for stochastic models and risks. *Wiley Series in Probability and Statistics*, Wiley, Chichester
- [Nadarajah(2008)] Nadarajah S (2008) A review of results on sums of random variables. *Acta Applicandae Mathematicae* 103(2):131-140
- [Nelsen(2006)] Nelsen RB (2006) An introduction to copulas, 2nd edn. *Springer Series in Statistics*, Springer, New York
- [Shaked and Shanthikumar(2007)] Shaked M, Shanthikumar JG (2007) Stochastic orders. *Springer Series in Statistics*, Springer, New York
- [Sheu et al(2016)] Sheu SH, Chen YL, Chang CC, Zhang ZG (2016) A note on a two variable block replacement policy for a system subject to non-homogeneous pure birth shocks. *Applied Mathematical Modelling* 40(5-6):3703-3712