This is a peer-reviewed, accepted author manuscript of the following research article: Langston, A, Finkelstein, M & Cha, J 2023, 'On the point process with finite memory and its application to optimal age replacement', *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*.

On the point process with finite memory and its application to optimal age replacement

Amy Langston[a], Maxim Finkelstein[b, c] and Ji Hwan Cha[*, d]

[a]Department of Statistics, Rhodes University, Makhanda, South Africa.
E-mail: a.langston@ru.ac.za
[b]Department of Mathematical Statistics, University of the Free State, Bloemfontein, South Africa
E-mail: FinkelM@ufs.ac.za
[c]Department of Management Science, University of Strathclyde, Glasgow, UK
[d]Department of Statistics, Ewha Womans University, Seoul, 120-750, Rep. of Korea.
E-mail: jhcha@ewha.ac.kr

Abstract

There has been extensive study of various repair models in the literature, mostly under the assumption that these repairs are minimal or imperfect/better than minimal. Although this is often a realistic assumption, it may not be sufficient to model instances where the repair is worse than minimal. The generalized Polya process (GPP) that has been used to describe this type of repair takes into account all previous events/repairs, which is not often the case in practice. Therefore, in this paper, we define a new process with *finite memory* that starts as the GPP but, after a certain number of events or elapsed time, becomes the non-homogeneous Poisson process of repairs (minimal repairs). The corresponding age replacement policy is defined and the optimal solutions that minimize the long-run expected cost rate are analyzed. The detailed numerical examples illustrate our findings.

Keywords: Worse than minimal repair; minimal repair; generalized Polya process, optimal age replacement, long-run expected cost rate.

1 Introduction

Stochastic point processes have been extensively used in reliability literature for modeling the processes of failures and repairs (see, e.g., Cha and Finkelstein (2018)) in repairable items. If each failure of a repairable item is instantaneously repaired, then these processes, obviously, coincide, which will be assumed in our study. For definiteness, we will call the processes of

[*]corresponding author

interest by the term "repair processes". One of the simplest models used to describe a repair process is that of perfect repairs, when an item at each repair is restored to the 'as good as new' state. This can be modeled by the classical renewal process (Barlow and Proschan, 1975). Note that, apart from the renewal points of time, this process is characterized by some memory/history, i.e., the time elapsed since the last renewal. However, it is often unrealistic to assume that repair is perfect as in practice it is imperfect due to various reasons. Indeed, even spare parts in a warehouse cannot be considered as new as they deteriorate with age during storage. Therefore, considerable attention in reliability literature has been given to studying various imperfect repair models and their applications.

The first imperfect repair model, and still one of the most commonly used models in applications, is that for minimal repairs. A minimal repair restores an item to the 'as bad as old' state. This type of repair often occurs in practice (Rausand and Hoyland, 2003) and can be seen in many real-world settings, such as when a minor component is repaired or replaced as part of a larger failed system. Thus, in accordance with the definition, the minimal repair does not change the failure rate that describes a lifetime of an item. It is also well-known that the process of instantaneous minimal repairs can be described by the non-homogeneous Poisson process (NHPP) with the rate equal to the failure rate that describes a lifetime of the 'underlying' item. It also does not have 'memory' as its increments are independent. Using the NHPP as the corresponding point process model usually allows for closed-form analytical results and tractable mathematical properties that can be used to, e.g., solve various optimal maintenance problems.

A large number of imperfect repair models have been reported in the literature in the last several decades. For instance, Kijima (1989), Ansell et al. (2004),Tanwar et al. (2014), Finkelstein (2008), Finkelstein and Cha (2013), Jack (1998), Wang and Pham (2006) study models based on the notion of virtual age. These results were generalized in some recent publications, see e.g., Syamsundar et al. (2021), Liu et al. (2020), and Doyen et al. (2019), to name a few. Some other approaches to imperfect repair modeling can be found, e.g., in Brown and Proschan (1983), de Toledo et al. (2015),Levitin and Lisnianski (2000).

All these papers and numerous others consider the cases when imperfect repair is better than minimal. Although minimal repair or imperfect repair of the described types can be realistic in practice, they still may not be sufficient to fully describe the type of repair being conducted. This happens, e.g., when the repair, in essence, is worse than minimal, which can be realistic in various applications. A worse than minimal repair restores an item to the 'worse than old' state and could be the result of external or internal shocks affecting an item, insufficient quality of the repair action itself, the adverse effects of previous repairs, etc. (see Lee and Cha (2016) for specific real-world examples). It is important to note that after each worse than minimal repair, the corresponding failure rate is larger than it was just before failure (compare with minimal repairs).

As shown in Lee and Cha (2016),Badía et al. (2018),Cha and Finkelstein (2018), the generalized

Polya process (GPP) introduced and extensively studied in Cha (2014), is an effective tool in practice for modeling the worse than minimal repair and considering the corresponding optimal PM strategy. See the examples of the GPP repair in this paper mostly dealing with the case when, due to the failure of a component, others experience a larger stress, and therefore, after repair the stochastic intensity of a system is larger than that before.

It is worth noting that the worse than minimal repair have been considered in the literature in the framework of virtual age models for systems with the given bathtub failure rate (see, e.g., Dijoux (2009)). Indeed, if an age is decreased for the decreasing failure rate, it means that this repair is harmful. The approach in this and subsequent papers in this respect is 'empirical' postulating the shapes of failure rates. On the contrary, our harmful repair is driven by the justified mathematical model with the comprehensive probabilistic analysis.

A recent study by Cha et al. (2023) proposes the extension of the GPP repair model to the case of the GPP process with delay. It models the situation when initially, an item is minimally repaired after failures, however, with the increasing number of minimal repairs, they become worse than minimal after some point (delay). This models deterioration of the quality of the repairs as an item deteriorates beyond a certain 'point of resilience'. The NHPP + GPP process described in this paper is used for obtaining optimal replacement times for items with repair processes that follow the defined pattern.

As it was already discussed, the Poisson process does not have a memory. Although the renewal process has a 'short' memory, this memory does not affect repair actions, as there is no memory at the renewal times. On the other hand, in many practical situations, e.g., in optimal replacement scheduling, not all previous events in the process (history/memory) influence the occurrence of future events. The GPP was introduced to describe this dependence probabilistically (see the next section). However, the GPP takes into account all previously occurred events, whereas in practice, this is often not the case. The new point process proposed in our paper takes into account only a finite number of events or events in a finite interval of time. In the latter case, the corresponding number is not increasing to infinity (almost surely) as time increases to infinity. This is dramatically different from the GGP and the delayed model in Cha et al. (2023). This process will be defined and the properties necessary for application to the optimal age replacement problem will be obtained.

To model the described pattern we introduce the following repair process: it starts as the GPP, but after a certain number of events (or elapsed time) it becomes the NHPP, thus 'fixing' the history after the corresponding change point. Apart from the theoretical value of the proposed model, there exist a number of practical situations for a repair process of this nature. For example, an item undergoes worse than minimal repairs until specific repair equipment is obtained or a trained technician is hired to be able to minimally repair the system. There can be also a shortage of 'proper', needed for repair spares for some time that results in a worse than minimal repair or other environmental or operational restrictions. As the worse than minimal repair is usually cheaper than minimal (for instance it uses cheaper spare parts that have larger failure rates) it can be cost-wise reasonable to perform it for some time (or after several repairs of this kind) and then switch to 'normal' spare parts. The corresponding bivariate optimization problem for obtaining e.g., the optimal time of preventive maintenance and the optimal time of switching to minimal repairs can be considered in principle as the follow-up of the current research.

Another example justifying the setting of the proposed combined model is in the framework of the Maintenance-Induced-Failures (see, e.g., Selvik and Lohne (2022)). These failures can result, e.g., from the new failure modes induced by maintenance that are limited in numbers. Therefore, after starting as GPP, eventually the process 'converges' to minimal repair as described in the paper. Obviously, this should be explored more in the future by practitioners, whereas we provide an innovative model and a possible application.

It should be noted that apart from preventive maintenance, there are other applications where the proposed process can be used. For instance, to model the impact of external shocks on a system that is diminishing after some time due to protection or other measures.

The paper is organized as follows. In Section 2, we provide a preliminary discussion and define the combined process for two relevant settings. The optimal age replacement problem and the derivations for the corresponding long run expected cost rates are presented in Section 3. Section 4 provides numerical illustrations of our findings. Finally, concluding remarks are given in Section 5.

2 Model Descriptions

2.1 Preliminaries

In this paper, we define the following stochastic process with a finite memory. It starts as the GPP, but after a certain number of events (or elapsed time) it becomes the NHPP, thus 'fixing' the history after the change point. We will characterize the point processes of interest through the notion of stochastic intensity (intensity process). Therefore, we will first provide a brief yet sufficient description of this notion to allow for our further presentation. For further mathematical details on stochastic intensity, we refer the reader to Aven and Jensen (1999; 2000).

Let $\{N(t), t \geq 0\}$ be an orderly (i.e., without multiple occurrences) point process whose history (internal filtration) in $[0, t)$ is $H_{t-} \equiv \{N(u), 0 \leq u < t\}$. Here, the history H_{t-} is the set of all point events in $[0, t)$ and can be equivalently defined in terms of $N(t-)$ and the sequential arrival times of the events $0 \leq T_1 \leq T_2 \leq \ldots \leq T_{N(t-)} < t$, where T_i is the time from 0 until the arrival of the ith event in $[0, t)$. The stochastic intensity for $\{N(t), t \geq 0\}$, λ_t, which is a stochastic process itself, can then be defined as the following limit

$$\lambda_t = \lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t) = 1 \mid H_{t-})}{\Delta t} = \lim_{\Delta t \to 0} \frac{E[N(t, t + \Delta t) \mid H_{t-}]}{\Delta t}, \tag{1}$$

where $N(t, t + \Delta t)$ represents the number of events in $[t, t + \Delta t)$ (see e.g., Finkelstein and Cha (2013) and Cha (2014)).
It is well known that for the NHPP with the intensity function (rate) $\lambda(t)$, the stochastic intensity is deterministic and is given by $\lambda_t = \lambda(t)$. The GPP is a relatively new counting process (see Cha (2014) for an extensive discussion of its properties) and can be defined as follows.

Definition 1 ***G eneralized P olya Process***
A counting process $\{N(t), t \geq 0\}$ is said to be a generalized Polya process (GPP) with the set of parameters $(\lambda(t), \alpha, \beta)$, $\alpha \geq 0$, $\beta > 0$, if

i. $N(0) = 0$;

ii. $\lambda_t = (\alpha N(t-) + \beta) \lambda(t)$.

The GPP reduces to the NHPP with intensity function $\lambda(t)$ when considering the parameter set $(\lambda(t), \alpha = 0, \beta = 1)$, and as such, can be understood as a generalization of the NHPP (Cha, 2014). The intepretation of this process is also fairly starightforward from the definition, where $\lambda(t)$ is some baseline function for the process, and α describes how much influence the history (given by the number of prior events) has on the probability of occurrence of an event in an infinitesimal interval of time.

In order to define the new combined process probabilistically, we will require the following supplementary result:

Lemma 1 *For the GPP with set of parameters $(\lambda(t), \alpha, \beta)$, $\alpha > 0, \beta > 0$, the distribution of $N(t)$ is given by*

$$P(N(t) = n) = \frac{\Gamma\left(\frac{\beta}{\alpha} + n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} (1 - \exp\{-\alpha\Lambda(t)\})^n (\exp\{-\alpha\Lambda(t)\})^{\frac{\beta}{\alpha}}, \quad n = 0, 1, 2, \ldots \tag{2}$$

where $\Lambda(t) \equiv \int_0^t \lambda(u)\, du$.

The proof of this is given by the proof of Theorem 1 in Cha (2014).
The corresponding repair/failure process is called the GPP repair process, and is, without loss of generality, a reparameterization of the GPP definition above with $\beta = 1$ (Lee and Cha, 2016). Note that in this case, before the first event, an item's intensity function is equal to its failure rate and with each event, it increases on $\alpha\lambda(t)$. Thus, for our application in Sections 3 and 4, the value $\beta = 1$ will be considered, whereas it will be a positive arbitrary value for the general description of the process to follow. In the next section, two models to describe the finite memory process are defined.

2.2 Model 1: NHPP events start after k GPP events

From the start of operation up to the kth event, the point process behaves as the GPP, and then, just after the kth event, the point process behaves as the NHPP. Thus the filtration (history) is not increasing any more. The resulting combined process can be described as follows.

Definition 2 ***Combined process of Type I***

1. $N(0) = 0$;
2. $\lambda_t = (\alpha N(t-) + \beta)\lambda(t),\ N(t-) = 1, 2, \ldots, k$;
3. $\lambda_t = (\alpha k + \beta)\lambda(t),\ N(t-) = k+1, k+2, \ldots$

An example of the stochastic intensity for the combined process of Type I with $k = 4$ is given in Figure 1 for $\alpha = 2$ and $\beta = 1$. Here, it is assumed that the baseline rate is $\lambda(t) = t^2 + 1$ and that the ith event occurs at time S_i. As shown in the figure, the stochastic intensity jumps at each event occurrence until the kth event occurring at time S_4. Beyond this, the stochastic intensity has the smooth Poisson rate. The baseline rate on the plot is equivalent to the case where $\alpha = 0$ and $\beta = 1$. That is when the process is simply the NHPP with rate $\lambda(t)$.

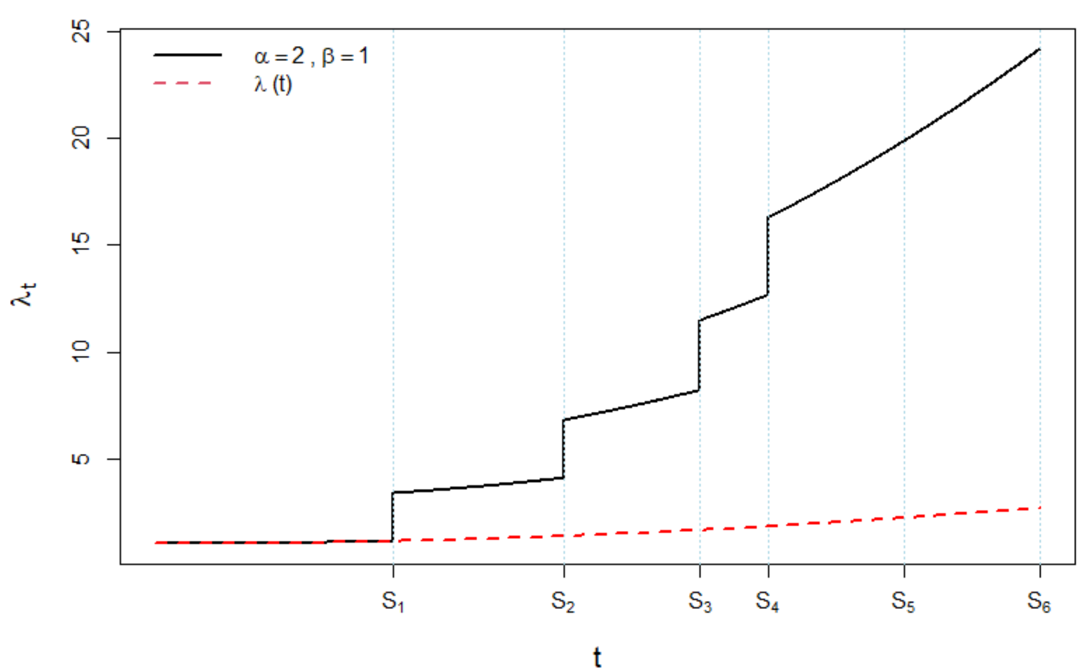

Figure 1: The stochastic intensity for the combined process of Type I with $k = 4$ and baseline failure rate $\lambda(t) = t^2 + 1$ for $\alpha = 2$, $\beta = 1$.

Before proceeding, we need to characterize this process probabilistically, as it is neither pure GPP nor pure NHPP. Therefore, let us derive $P(N(t)=n)$.

1. First, let $n=1,2,\ldots,k-1$.
   Denoting s_n as the nth arrival time, s_k would be considered as the change point for this combined process. Therefore, if we consider the event $\{N(t)=n\}$ when $n=1,2,\ldots,k-1$, we have that $s_n<t\leq s_{n+1}\leq s_k$. Thus, the time instant t is before the change point s_k, and the process is the GPP. Obviously, in this case, $P(N(t)=n)$ is given by (2) for $n=1,2,\ldots,k-1$.

2. Now, consider $n=k$.
   Once again denoting s_n as the nth arrival time, the time instant of interest t for the event $\{N(t)=k\}$ now satisfies $s_k<t\leq s_{k+1}$. Therefore, for $n=k$, the process is in the interval $(s_k,t]$ and is characterized by the NHPP. However, due to property (3) in Definition 2, the stochastic intensity in this interval is still given by $(\alpha k+\beta)\lambda(t)$, which also results in
   $$P(N(t)=k)=\frac{\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)k!}\left(1-\exp\{-\alpha\Lambda(t)\}\right)^k\exp\{-\beta\Lambda(t)\}$$
   from (2).

3. Finally, let $n=k+1,k+2,\ldots$
   Once again, if S_k is the arrival time of the kth event, we observe that
   $$P(N(t)=n)=\int_0^t P(N(t)=n\mid S_k=u)\,f_{S_k}(u)\,\mathrm{d}u,$$
   where $f_{S_k}(u)$ is the pdf of S_k and
   $$\begin{aligned}P(N(t)=n\mid S_k=u)&=P(N(t)=n\mid N(s)<k,0\leq s<u,N(u)=k)\\&=P(N(t)-N(u)=n-k\mid N(s)<k,0\leq s<u,N(u)=k)\\&=P(N(t)-N(u)=n-k).\end{aligned}$$
   Note that the last equality holds due to the independent increment property of the NHPP. That is, in the interval $[u,t)$ the process is the NHPP with rate $(\alpha k+\beta)\lambda(t)$. It then follows that
   $$\begin{aligned}P(N(t)=n\mid S_k=u)&=P(N(t)-N(u)=n-k)\\&=\frac{(W(t)-W(u))^{n-k}}{(n-k)!}\exp\{-(W(t)-W(u))\},\end{aligned}$$
   where $W(t)=(\alpha k+\beta)\Lambda(t)$.
   On the other hand, from Badía et al. (2018), we have that
   $$f_{S_k}(t)=\beta\left(1+\frac{\beta}{\alpha}\right)\left(1+\frac{\beta}{2\alpha}\right)\ldots\left(1+\frac{\beta}{(k-1)\alpha}\right)\lambda(t)\exp\{-\beta\Lambda(t)\}\left(1-\exp\{-\alpha\Lambda(t)\}\right)^{k-1}$$

which can be written as

$$f_{S_k}(t) = \frac{\alpha\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!}\lambda(t)\exp\{-\beta\Lambda(t)\}\left(1-\exp\{-\alpha\Lambda(t)\}\right)^{k-1}.$$

Finally,

$$\begin{aligned} P(N(t)=n) &= \int_0^t P(N(t)=n \mid S_k=u) f_{S_k}(u)\,\mathrm{d}u \\ &= \int_0^t \frac{\alpha\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!}\lambda(u)\exp\{-\beta\Lambda(u)\}\left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1} \\ &\quad \times \frac{(W(t)-W(u))^{n-k}}{(n-k)!}\exp\{-(W(t)-W(u))\}\,\mathrm{d}u, \qquad n=k+1,k+2,\ldots \end{aligned}$$

Given the above, the mean number of events in $[0,t)$ can be obtained as

$$\begin{aligned} E[N(t)] =& \sum_{n=0}^{k-1} n\frac{\Gamma\left(\frac{\beta}{\alpha}+n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)n!}\left(1-\exp\{-\alpha\Lambda(t)\}\right)^n\exp\{-\beta\Lambda(t)\} \\ &+\alpha\int_0^t (W(t)-W(u)+k)\lambda(u)\frac{\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!}\exp\{-\beta\Lambda(u)\}\left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1}\mathrm{d}u. \end{aligned} \tag{3}$$

The proof of this is deferred to the Appendix. Note that $E[N(t)]$ is the main characteristic that is needed for our optimal replacement application.

Remark 1 *The change point k, where the GPP becomes the NHPP, can be random. Let this random variable be denoted by K with the corresponding probability mass function $p(k)$. Then, $E[N(t)]$ is given by*

$$\begin{aligned} E[N(t)] =& \sum_{k=0}^{\infty} p(k)\left[\sum_{n=0}^{k-1} n\frac{\Gamma\left(\frac{\beta}{\alpha}+n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)n!}\left(1-\exp\{-\alpha\Lambda(t)\}\right)^n\exp\{-\beta\Lambda(t)\}\right. \\ &\left.+\alpha\int_0^t (W(t)-W(u)+k)\lambda(u)\frac{\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!}\exp\{-\beta\Lambda(u)\}\left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1}du\right]. \end{aligned}$$

Note that, $K=0$ corresponds to the pure NHPP and $K=\infty$ corresponds to the pure GPP.

2.3 Model 2: NHPP events start after a fixed time c

Another option is when the occurs after some time. For $t \le c$, the point process behaves as the GPP, and for $t > c$, the point process behaves as the NHPP. The resulting combined process can be described as follows.

Definition 3 ***Combined process of Type II***

1. $N(0) = 0$;
2. $\lambda_t = (\alpha N(t-) + \beta)\,\lambda(t)$, $t \le c$;
3. $\lambda_t = (\alpha N(c) + \beta)\,\lambda(t)$, $t > c$.

An example of the stochastic intensity for the combined process of Type II is given in Figure 2 for $\alpha = 2$ and $\beta = 1$. Here, it is assumed that the baseline rate is $\lambda(t) = t^2 + 1$ and that the ith event occurs at time S_i. As shown in the figure, the stochastic intensity jumps at each event occurrence until the event that occurs just prior to time c. In this case, until the event occurring at time S_3. Beyond this, the stochastic intensity has the smooth Poisson rate. The baseline rate on the plot is equivalent to the case where $\alpha = 0$ and $\beta = 1$. That is when the process is simply the NHPP with rate $\lambda(t)$.

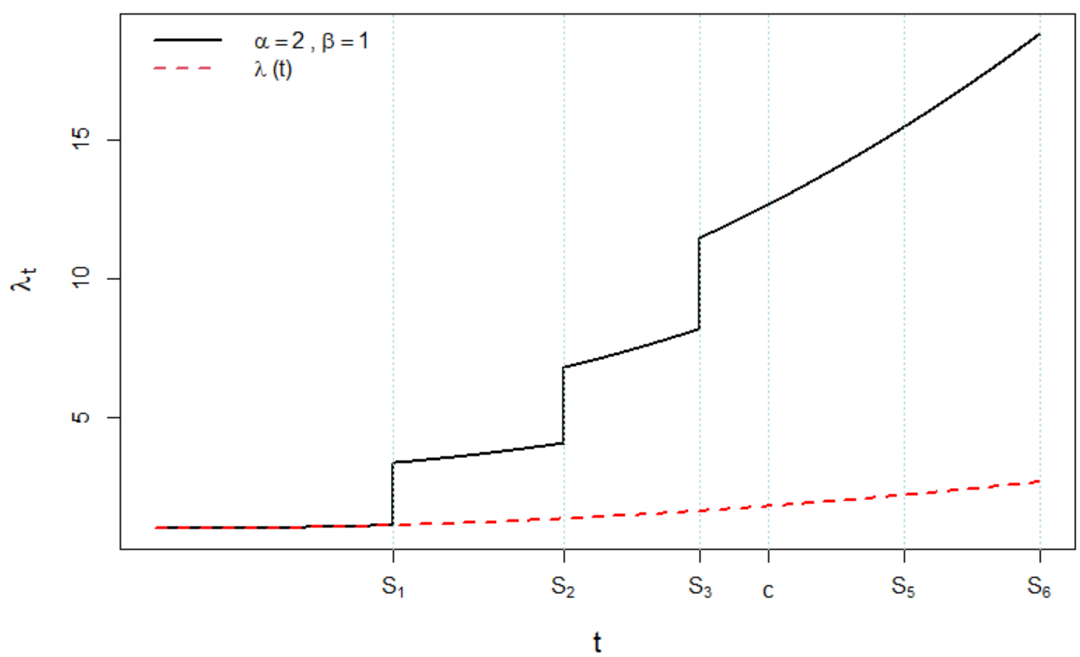

Figure 2: The stochastic intensity for the combined process of Type II with baseline failure rate $\lambda(t) = t^2 + 1$ for $\alpha = 2$, $\beta = 1$.

Once again, we need to characterize this process probabilistically. Therefore, let us derive $P(N(t) = n)$.

1. For $t \le c$, as in Model 1, $P(N(t) = n)$ is given by (2).

2. For $t > c$, observe that

$$P(N(t) = n) = \sum_{j=0}^{n} P(N(t) = n \mid N(c) = j) P(N(c) = j)$$
$$= \sum_{j=0}^{n} P(N(t) - N(c) = n - j \mid N(c) = j) P(N(c) = j).$$

In the interval (c, ∞), the process is the NHPP with intensity function $\lambda_t = (\alpha N(c) + \beta) \lambda(t)$. Therefore,

$$P(N(t) - N(c) = n - j \mid N(c) = j) = \frac{(Z(t) - Z(c))^{n-j}}{(n-j)!} \exp\{-(Z(t) - Z(c))\},$$

where now $Z(t) = (\alpha j + \beta) \Lambda(t)$.
On the other hand, $P(N(c) = j)$ can be obtained from (2). Therefore,

$$P(N(t) = n) = \sum_{j=0}^{n} P(N(t) - N(c) = n - j \mid N(c) = j) P(N(c) = j)$$
$$= \sum_{j=0}^{n} \frac{\Gamma\left(\frac{\beta}{\alpha} + j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} (1 - \exp\{-\alpha\Lambda(c)\})^{j} \exp\{-\beta\Lambda(c)\} \frac{(Z(t) - Z(c))^{n-j}}{(n-j)!} \exp\{-(Z(t) - Z(c))\}$$

Given the above, the mean number of events in $[0, t)$ can be obtained as

$$E[N(t)] = \sum_{n=0}^{\infty} n \frac{\Gamma\left(\frac{\beta}{\alpha} + n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} (1 - \exp\{-\alpha\Lambda(t)\})^{n} \exp\{-\beta\Lambda(t)\}$$
$$= \frac{\beta}{\alpha} (\exp\{\alpha\Lambda(t)\} - 1) \tag{4}$$

for $t \geq c$, and

$$E[N(t)] = \sum_{n=0}^{\infty} n \sum_{j=0}^{n} \frac{\Gamma\left(\frac{\beta}{\alpha} + j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} (1 - \exp\{-\alpha\Lambda(c)\})^{j} \exp\{-\beta\Lambda(c)\} \frac{(Z(t) - Z(c))^{n-j}}{(n-j)!} \exp\{-(Z(t) - Z(c))\}$$
$$= \frac{\beta}{\alpha} (\exp\{\alpha\Lambda(c)\} - 1) + \beta (\Lambda(t) - \Lambda(c)) \exp\{\alpha\Lambda(c)\} \tag{5}$$

for $t < c$. The proof of these is deferred to the Appendix.

3 Optimal age replacement policies

In what follows we will consider the combined process with finite memory to describe a repair process and investigate its application to the optimal age replacement problem. Therefore, let $\{N(t), t \geq 0\}$ be the stochastic repair process of an item with the baseline function (the failure rate of an item) $\lambda(t)$, and $\alpha > 0$, $\beta = 1$.

Assume that an item is replaced whenever it reaches age T. That is, between consequent replacements, the failures/repairs are performed in accordance with one of the described models above. After each replacement, a new cycle begins, and so on.

3.1 Model 1

Recall that for Model 1, the process is the GPP until the kth event, and then it is the NHPP. The corresponding long-run expected cost rate (or, equivalently, the expected cost rate on one cycle) for this periodic setting is:

$$C(T) = \frac{c_g \sum_{j=0}^{k} jP(N(T)=j) + k(c_g - c_m)\sum_{j=k+1}^{\infty} P(N(T)=j) + c_m \sum_{j=k+1}^{\infty} jP(N(T)=j) + c_R}{T}, \tag{6}$$

where c_g is the cost of repair under the GPP (cost of a worse than minimal repair), c_m is the cost of repair under the NHPP (cost of a minimal repair), c_R is the cost of replacement, and $c_g \leq c_m < c_R$.

Remark 2 *If $c_g = c_m$, the long run expected cost rate reduces to*

$$C(T) = \frac{c_g E[N(t)] + c_R}{T},$$

where $E[N(t)]$ is given by (3).

Assume that the failure rate of an item before the first repair (baseline failure rate) $\lambda(t)$ is increasing. For definiteness, assume that $\lim_{t\to\infty} \lambda(t) = \infty$. This is a common assumption made in standard optimal age replacement problems in the literature as it describes an item's natural degradation or aging with time (see Nakagawa (2005)). The corresponding generalization to the case of a finite limit can be also considered in a standard way. Under this assumption, it can be easily shown that

$$\lim_{T\to 0} C(T) = \infty, \quad \lim_{T\to\infty} C(T) = \infty,$$

and $C(T)$ is decreasing in the vicinity of $T = 0$. Therefore, there exists an optimal solution to the problem

$$C(T^*) = \min_{0<T<\infty} C(T).$$

Differentiating (6) and equating to 0 produces an expression that is too cumbersome to allow for the optimal solution to be found analytically. However, the optimal solution to the problem can be found numerically.

3.2 Model 2

Recall that for Model 2, the process is the GPP for $t \leq c$, then it is the NHPP. As the analysis of optimal procedures is similar to that for Model 1, below we only briefly describe the slight modifications for obtaining the long-run expected cost rate $C(T)$.

Therefore, for $T \leq c$, the corresponding long-run expected cost rate is given by

$$C(T) = \frac{c_g E[N(T)] + c_R}{T}, \tag{7}$$

where $E[N(t)]$ is defined by (4), whereas for $T > c$:

$$C(T) = \frac{c_g E[N(c)] + c_m (E[N(T)] - E[N(c)]) + c_R}{T}, \tag{8}$$

where $E[N(t)]$ is now defined by (5) .

Remark 3 *Equation (8) can be explicitly written as*

$$C(T) = \frac{\frac{c_g}{\alpha}(\exp\{\alpha\Lambda(c)\} - 1) + c_m [\Lambda(T) - \Lambda(c)] \exp\{\alpha\Lambda(c)\} + c_R}{T}.$$

4 Numerical illustrations and discussion

4.1 Model 1

Before the first repair, consider the GPP with baseline failure rate $\lambda(t)$ modeled by the power function, which corresponds to the Weibull baseline distribution of an item. More specifically, let $\lambda(t) = \lambda t$, $\lambda = 0.5$, and $c_g = c_m = 1$. The latter is sufficient to illustrate the main properties of the considered optimal model. The corresponding expected cost rate for different values of k is given in Figure 3. Note that since $c_g = c_m$, we can use $C(T)$ in Remark 2 and (3), where $\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u = \frac{\lambda t^2}{2}$.

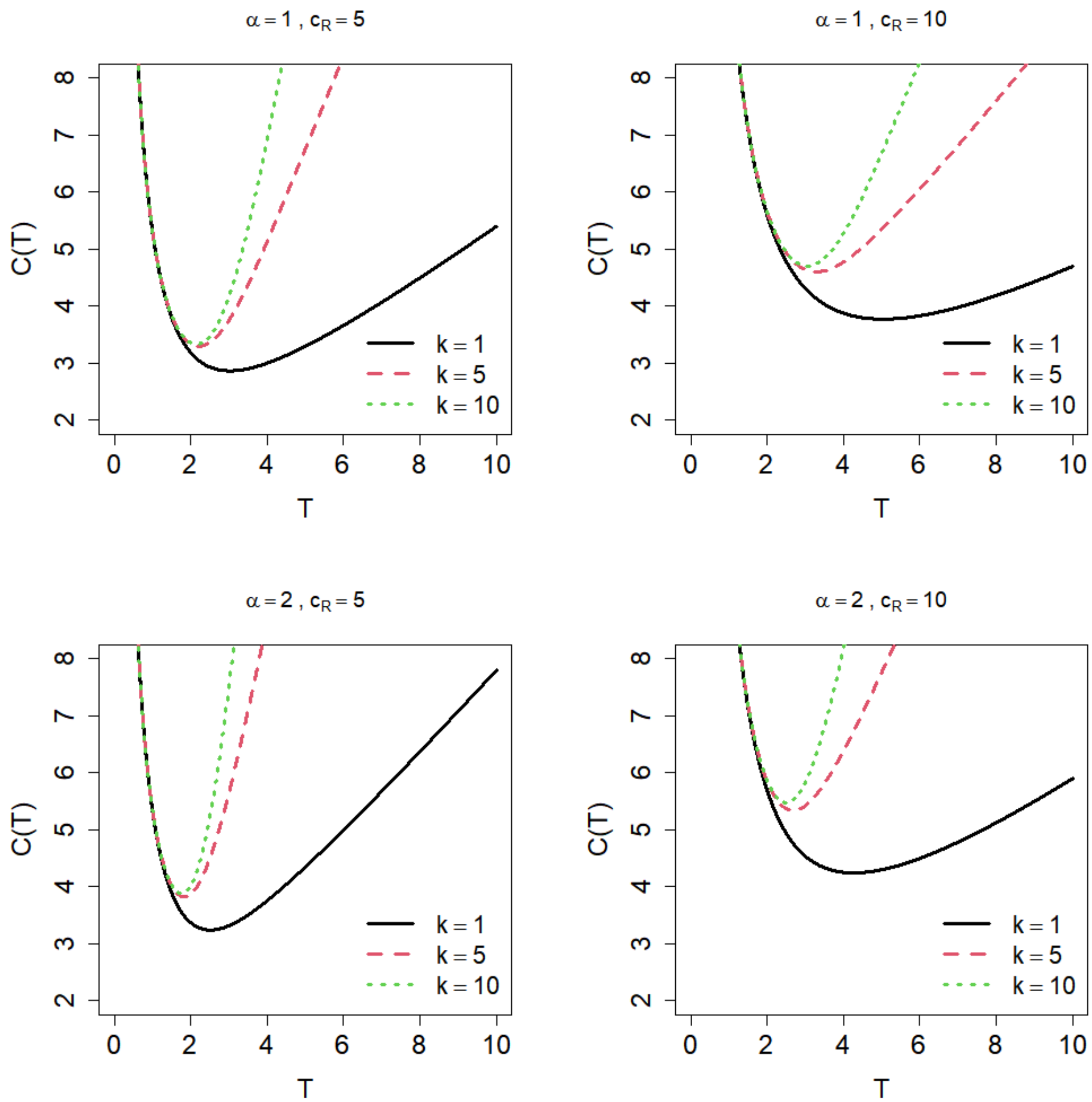

Figure 3: The long-run expected cost rates for $\alpha = 1$, $\alpha = 2$, $\lambda = 0.5$, $c_g = c_m = 1$, $c_R = 5$, $c_R = 10$.

As expected, there is a clear solution to the optimal problem $C(T^*) = \min_{0<T<\infty} C(T)$. We can see that as the number of worse than minimal (GPP) repairs k increases, the optimal replacement time T^* is decreasing. This is intuitively clear as a larger k implies that there are a larger number of GPP repairs, i.e., the failure rate that describes the repairable item (i.e., the corresponding expectation of the intensity process) is larger than that for a smaller k. Furthermore, it can be shown that an item with a larger increasing failure rate should be optimally replaced at a smaller replacement time (this can be loosely described as: the more ageing items need the less optimal PM period). Comparing curves for $\alpha = 1$ and $\alpha = 2$, it can be seen that the optimal replacement time T^* is decreasing with the increase of the GPP parameter α. This goes in line with our foregoing explanation. Similarly, for the increasing shape parameter λ of the Weibull distribution. This will be illustrated and discussed in more detail for Model 2. Further, as the cost of replacement c_R is increasing, the optimal time is increasing. This effect can be seen by comparing curves for $c_R = 5$ and $c_R = 10$. This holds as when the replacement cost is relatively large, replacement should be delayed as far as possible. Similar effects can be seen in Figure 4, where plots for the expected numbers of repairs at the optimal time T^* for varying α and c_R are presented. Increasing α decreases the expected number of repairs before replacement, and increasing c_R increases the expected number of repairs before replacement. The justification for this is as above. An interesting feature of the $E[N(T^*)]$ curves is that they appear to flatten as k gets relatively large. This occurs because beyond some k, the optimal time of replacement T^* will always fall before the switch to the NHPP repairs, and therefore, T^* is fully defined under the GPP repairs. That is, increasing k beyond a certain point will not add anything to the optimal replacement model.

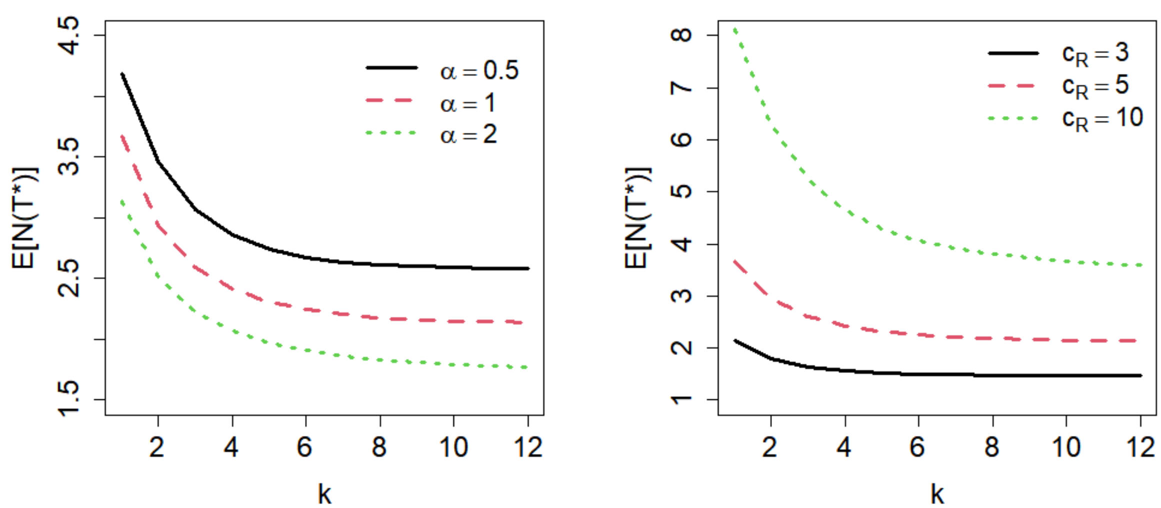

Figure 4: $E[N(T^*)]$ for fixed $\lambda = 0.5$, $c_g = c_m = 1$, $c_R = 5$ for varying α (left) and fixed $\lambda = 0.5$, $c_g = c_m = 1$, $\alpha = 1$ for varying c_R (right).

4.2 Model 2

As in the previous subsection, consider the GPP with the baseline failure rate $\lambda(t)$ modeled by the power function, which corresponds to the Weibull baseline distribution of a system and let $\lambda(t) = \lambda t$, $\lambda = 0.5$, and $c_g = c_m = 1$. The corresponding expected cost rate for the selected values of the change point c is given in Figure 3 using (7) and (8).

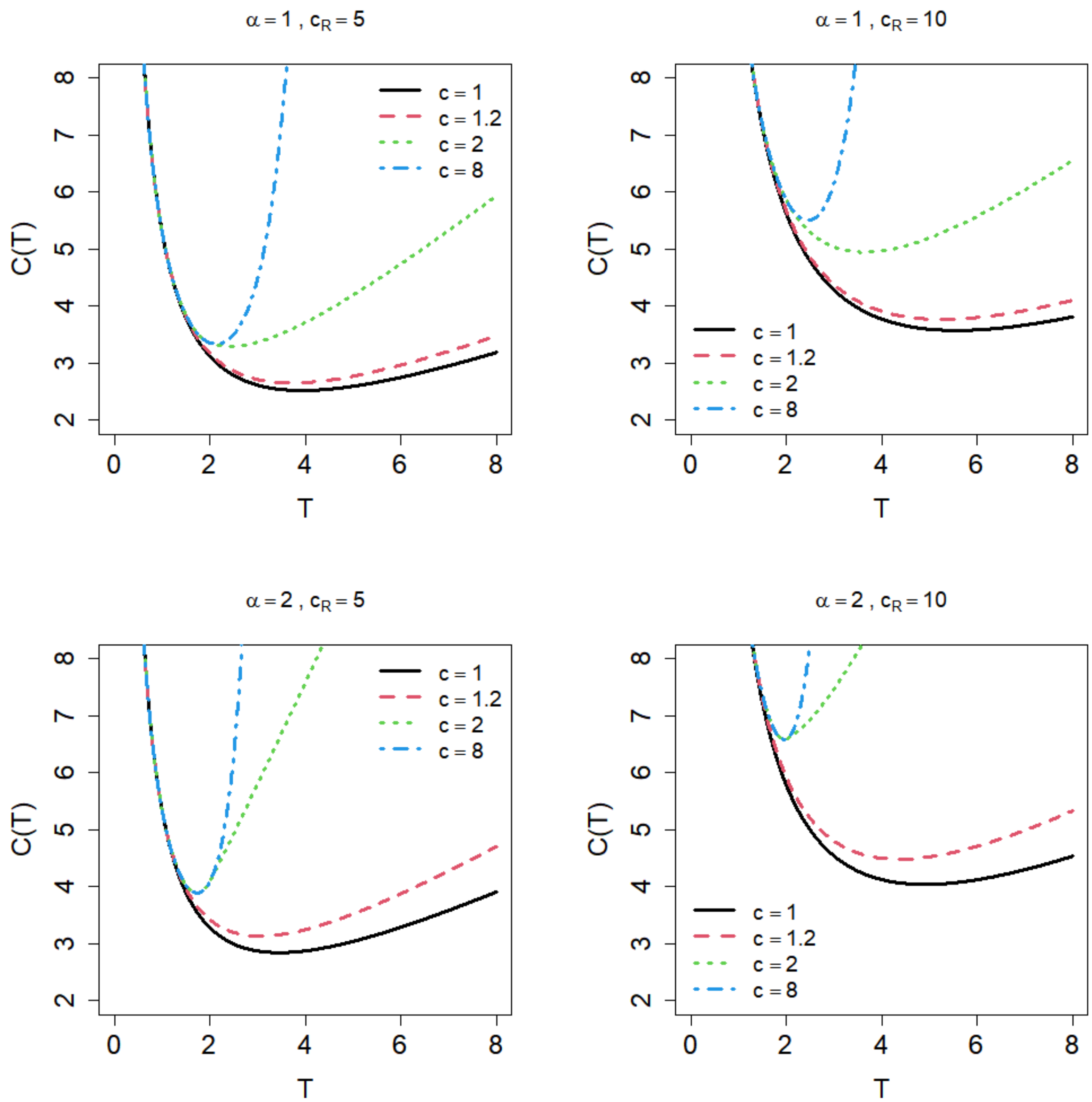

Figure 5: The long-run expected cost rates for $\alpha = 1$, $\alpha = 2$, $\lambda = 0.5$, $c_g = c_m = 1$, $c_R = 5$, $c_R = 10$.

As in Model 1, the optimal replacement time T^* is decreasing for increasing c. A larger c implies that for a longer time, the failure rate is increasing with the number of repairs, and as such, one should replace the system sooner. Further, increasing the GPP parameter α decreases the optimal time T^*. The same effect holds when increasing the shape parameter of the Weibull distribution λ (see Figure 6). This effect follows from general considerations (see e.g., Finkelstein et al. (2016)) in that increasing these parameters increases the aging (or degradation) of the system, which would lead to more frequent preventative maintenance/repair

actions, and as such, to reduce the long-run expected cost rate a replacement would need to be implemented sooner. Further, the optimal time T^* is increasing for increasing c_R.
Similar effects can be seen in Figure 6, where plots for the expected number of repairs at the optimal time T^* for varying α, λ, and c_R are presented. Increasing α and λ decreases the expected number of repairs before replacement, and increasing c_R increases the expected number of repairs before replacement. The justification for this is as above. An interesting feature of the $E\left[N\left(T^*\right)\right]$ curves is that beyond some c, they become constant. Recall that for $t \leq c$, the expectation $E\left[N\left(T^*\right)\right]$ is given by (7), which does not depend on c as follows from (4). Note that, even for the case $c_g = c_m = 1$, the cost rate in (8) depends on c as follows from (5).
Therefore, for all $T^* \leq c$, the expected number of repairs at the optimal replacement time T^* is constant with respect to c.

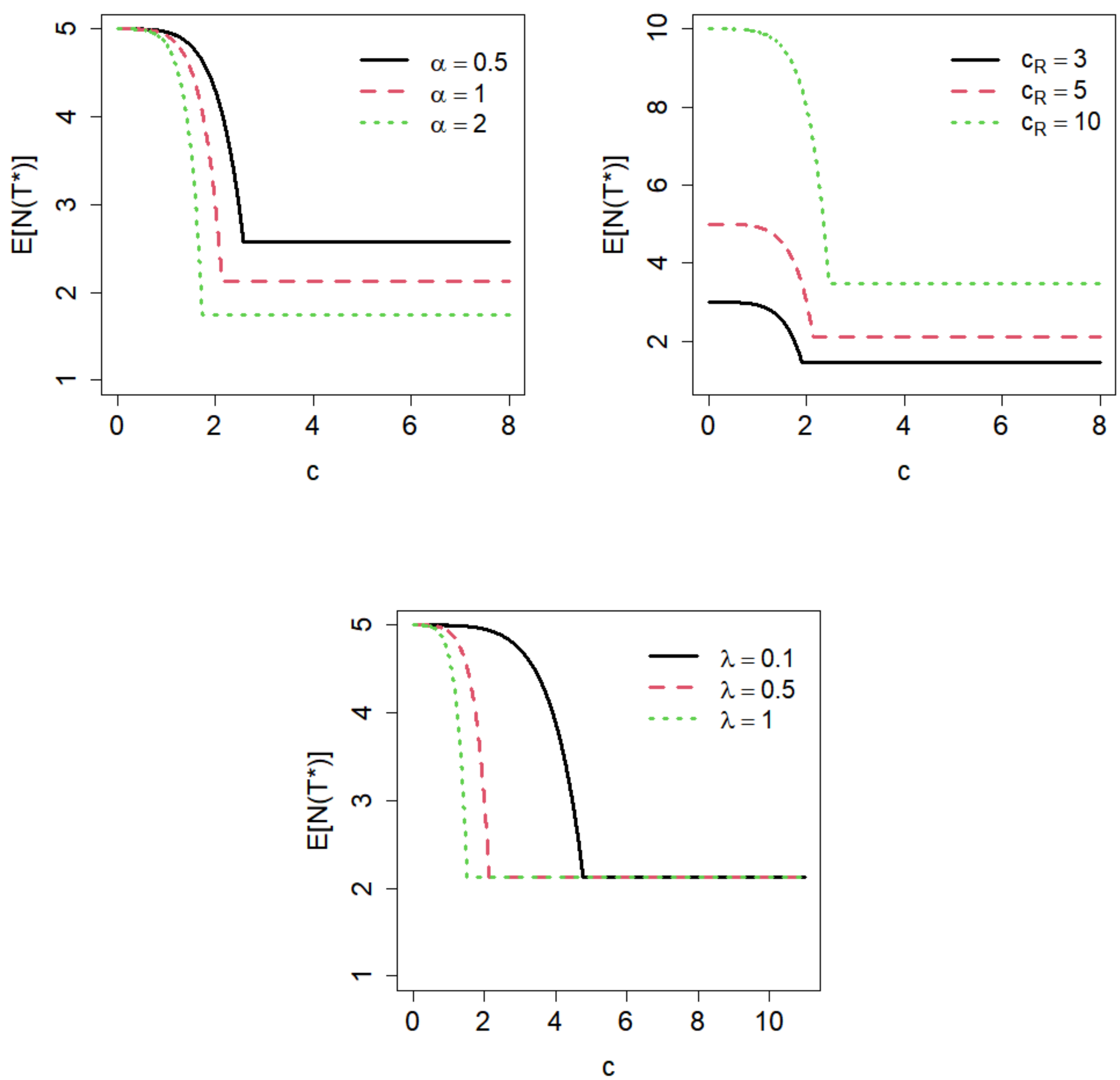

Figure 6: $E[N(T^*)]$ for fixed $\lambda = 0.5$, $c_g = c_m = 1$, $c_R = 5$ for varying α (left) and fixed $\lambda = 0.5$, $c_g = c_m = 1$, $\alpha = 1$ for varying c_R (right), fixed $c_g = c_m = 1$, $\alpha = 1$, $c_R = 5$ for varying λ (below).

5 Conclusions

In the literature, considerable attention has been given to minimal repair and imperfect, better than minimal repair models. In some instances, however, due to the adverse effects of previous repairs, external and internal shocks, insufficient quality of repair, etc., a repair can be worse than minimal. Although this setting can be more realistic in some reliability applications, much less attention has been given to it in the literature.

In this paper,we first propose a new stochastic process with a finite memory, which is more realistic than the 'unlimited' memory of the GGP. Thus, only a finite number of previous events can influence the occurrence of future events. This is done by 'combining' the GGP and the NHPP. The process starts as the GGP that takes into account previous events and then at some point, it becomes the the memoryless NHPP 'fixing' in this way the influence of events that occurred under the GGP.

We discuss the relevant stochastic properties of the new process and derive the necessary expressions for considering an optimal age replacement policy under two settings, namely, where the repair type changes after a certain number of worse than minimal repairs or after a given time. The existence of an optimal replacement time is also discussed, and detailed numerical illustrations explain our findings under various settings. We provide also relevant sensitivity analysis for the main parameters of the considered models.

There are a number of directions that could be considered in future research. In Model 1 and Model 2, the change points k and c, respectively, could be considered as random variables. In these cases, the appropriate distribution should be assumed, and the stochastic properties of these more complex models can be considered. Alternatively, one could consider more complex combined processes to model systems with a 'bounded' memory.

References

Ansell, J., Archibald, T., and Thomas, L. (2004). The elixir of life: using a maintenance, repair and replacement model based on virtual and operating age in the water industry. *IMA Journal of Management Mathematics*, 15(2):151–160.

Aven, T. and Jensen, U. (1999). *Stochastic models in reliability.* Springer.

Aven, T. and Jensen, U. (2000). A general minimal repair model. *Journal of applied probability*, 37:187–197.

Badía, F. G., Berrade, M. D., Cha, J. H., and Lee, H. (2018). Optimal replacement policy under a general failure and repair model: Minimal versus worse than old repair. *Reliability Engineering & System Safety*, 180:362–372.

Barlow, R. E. and Proschan, F. (1975). *Statistical theory of reliability and life testing: probability models.* Holt, Rinehart and Winston.

Brown, M. and Proschan, F. (1983). Imperfect repair. *Journal of Applied Probability*, 20(4):851–859.

Cha, J. H. (2014). Characterization of the generalized Polya process and its applications. *Advances in applied probability*, 46:1148–1171.

Cha, J. H. and Finkelstein, M. (2018). *Point processes for reliability analysis: shocks and repairable systems*. Springer.

Cha, J. H., Finkelstein, M., and Levitin, G. (2023). On the delayed worse-than-minimal repair model and its application to preventive replacement. *IMA Journal of Management Mathematics*, 34(1):101–122.

de Toledo, M. L. G., Freitas, M. A., Colosimo, E. A., and Gilardoni, G. L. (2015). Ara and ari imperfect repair models: Estimation, goodness-of-fit and reliability prediction. *Reliability Engineering & System Safety*, 140:107–115.

Dijoux, Y. (2009). A virtual age model based on a bathtub shaped initial intensity. *Reliability Engineering & System Safety*, 94(5):982–989.

Doyen, L., Drouilhet, R., and Brenière, L. (2019). A generic framework for generalized virtual age models. *IEEE Transactions on Reliability*, 69(2):816–832.

Finkelstein, M. (2008). *Failure rate modelling for reliability and risk*. Springer.

Finkelstein, M. and Cha, J. H. (2013). *Stochastic modeling for reliability: shocks, burn-in and heterogeneous populations*. Springer.

Finkelstein, M., Shafiee, M., and Kotchap, A. N. (2016). Classical optimal replacement strategies revisited. *IEEE Transactions on Reliability*, 65(2):540–546.

Jack, N. (1998). Age-reduction models for imperfect maintenance. *IMA Journal of Management Mathematics*, 9:347–354.

Kijima, M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability*, 26(1):89–102.

Lee, H. and Cha, J. H. (2016). New stochastic models for preventive maintenance and maintenance optimization. *European Journal of Operational Research*, 255(1):80–90.

Levitin, G. and Lisnianski, A. (2000). Optimization of imperfect preventive maintenance for multi-state systems. *Reliability Engineering & System Safety*, 67(2):193–203.

Liu, X., Finkelstein, M., Vatn, J., and Dijoux, Y. (2020). Steady-state imperfect repair models. *European Journal of Operational Research*, 286(2):538–546.

Nakagawa, T. (2005). *Maintenance theory of reliability*. Springer.

Rausand, M. and Hoyland, A. (2003). *System reliability theory: models, statistical methods, and applications*, volume 396. John Wiley & Sons.

Selvik, J. T. and Lohne, H. P. (2022). Modelling the effect of maintenance-induced failures from periodic testing of safety-critical equipment as part of RCM analysis in the oil and gas industry. In Pinto, J. O. P., Kimpara, M. L. M., Reis, R. R., Seecharan, T., Upadhyaya, B. R., and Amadi-Echendu, J. (eds) *15th WCEAM Proceedings. WCEAM 2021. Lecture Notes in Mechanical Engineering*. Springer. Cham.

Syamsundar, A., Naikan, V. N. A., and Wu, S. (2021). Estimating maintenance effectiveness of a repairable system under time-based preventive maintenance. *Computers & Industrial Engineering*, 156:107278.

Tanwar, M., Rai, R. N., and Bolia, N. (2014). Imperfect repair modeling using kijima type generalized renewal process. *Reliability Engineering & System Safety*, 124:24–31.

Wang, H. and Pham, H. (2006). *Reliability and optimal maintenance*. Springer.

Appendix

Derivation of $E[N(t)]$ for Model 1:

$$
\begin{aligned}
E[N(t)] &= \sum_{n=0}^{k} n \frac{\Gamma\left(\frac{\beta}{\alpha}+n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} \left(1-\exp\{-\alpha\Lambda(t)\}\right)^{n} \exp\{-\beta\Lambda(t)\} \\
&\quad + \sum_{n=k+1}^{\infty} n \int_0^t \frac{\alpha\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!} \lambda(u) \exp\{-\beta\Lambda(u)\} \left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1} \frac{(W(t)-W(u))^{n-k}}{(n-k)!} \exp\{-(W(t)-W(u))\} \,\mathrm{d}u \\
&= \sum_{n=0}^{k-1} n \frac{\Gamma\left(\frac{\beta}{\alpha}+n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} \left(1-\exp\{-\alpha\Lambda(t)\}\right)^{n} \exp\{-\beta\Lambda(t)\} \\
&\quad + \sum_{n=k}^{\infty} n \int_0^t \frac{\alpha\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!} \lambda(u) \exp\{-\beta\Lambda(u)\} \left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1} \frac{(W(t)-W(u))^{n-k}}{(n-k)!} \exp\{-(W(t)-W(u))\} \,\mathrm{d}u \\
&= \sum_{n=0}^{k-1} n \frac{\Gamma\left(\frac{\beta}{\alpha}+n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} \left(1-\exp\{-\alpha\Lambda(t)\}\right)^{n} \exp\{-\beta\Lambda(t)\} \\
&\quad + \int_0^t \left(\sum_{n=k}^{\infty} (n+k) \frac{(W(t)-W(u))^{n-k}}{(n-k)!} \exp\{-(W(t)-W(u))\}\right) \frac{\alpha\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!} \lambda(u) \exp\{-\beta\Lambda(u)\} \left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1} \,\mathrm{d}u \\
&= \sum_{n=0}^{k-1} n \frac{\Gamma\left(\frac{\beta}{\alpha}+n\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) n!} \left(1-\exp\{-\alpha\Lambda(t)\}\right)^{n} \exp\{-\beta\Lambda(t)\} \\
&\quad + \alpha \int_0^t (W(t)-W(u)+k) \lambda(u) \frac{\Gamma\left(\frac{\beta}{\alpha}+k\right)}{\Gamma\left(\frac{\beta}{\alpha}\right)(k-1)!} \exp\{-\beta\Lambda(u)\} \left(1-\exp\{-\alpha\Lambda(u)\}\right)^{k-1} \,\mathrm{d}u.
\end{aligned}
$$

Derivation of $E[N(t)]$ for Model 2:

The proof for $t \leq c$ is straightforward and follows from $E[N(t)]$ when $\{N(t), t \geq 0\}$ is the GPP with parameter set $(\lambda(t), \alpha, \beta)$, $\alpha \geq 0$, $\beta > 0$.
For $t > c$,

$$
\begin{aligned}
E[N(t)] &= \sum_{n=0}^{\infty} n \sum_{j=0}^{n} \frac{\Gamma\left(\frac{\beta}{\alpha}+j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} \left(1-\exp\{-\alpha\Lambda(c)\}\right)^{j} \exp\{-\beta\Lambda(c)\} \frac{(Z(t)-Z(c))^{n-j}}{(n-j)!} \exp\{-(Z(t)-Z(c))\} \\
&= \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{\beta}{\alpha}+j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} \left(1-\exp\{-\alpha\Lambda(c)\}\right)^{j} \exp\{-\beta\Lambda(c)\} \sum_{n=j}^{\infty} n \frac{(Z(t)-Z(c))^{n-j}}{(n-j)!} \exp\{-(Z(t)-Z(c))\} \\
&= \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{\beta}{\alpha}+j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} \left(1-\exp\{-\alpha\Lambda(c)\}\right)^{j} \exp\{-\beta\Lambda(c)\} \sum_{n=0}^{\infty} (n+j) \frac{(Z(t)-Z(c))^{n}}{n!} \exp\{-(Z(t)-Z(c))\} \\
&= \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{\beta}{\alpha}+j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} \left(1-\exp\{-\alpha\Lambda(c)\}\right)^{j} \exp\{-\beta\Lambda(c)\} (Z(t)-Z(c)+j) \\
&= \sum_{j=0}^{\infty} \left((\alpha j+\beta)(\Lambda(t)-\Lambda(c))+j\right) \frac{\Gamma\left(\frac{\beta}{\alpha}+j\right)}{\Gamma\left(\frac{\beta}{\alpha}\right) j!} \left(1-\exp\{-\alpha\Lambda(c)\}\right)^{j} \exp\{-\beta\Lambda(c)\} \\
&= \alpha(\Lambda(t)-\Lambda(c)) \frac{\beta}{\alpha} \left(\exp\{\alpha\Lambda(c)\}-1\right) + \beta(\Lambda(t)-\Lambda(c)) + \frac{\beta}{\alpha}\left(\exp\{\alpha\Lambda(c)\}-1\right) \\
&= \frac{\beta}{\alpha}\left(\exp\{\alpha\Lambda(c)\}-1\right) + \beta(\Lambda(t)-\Lambda(c)) \exp\{\alpha\Lambda(c)\}.
\end{aligned}
$$