Change-level detection for Lévy subordinators

Zeina Al Masry[a], Landy Rabehasaina[b], Ghislain Verdier[c]

[a] [1]*FEMTO-ST institute, Université de Bourgogne Franche-Comté, CNRS, ENSMM, Besançon, France*
[b]*Laboratoire de Mathématiques de Besançon, Université de Bourgogne Franche-Comté, CNRS UMR 6623, Besançon, France*
[c]*LMAP, Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, Pau, France*

Abstract

Let $\boldsymbol{X} = (X_t)_{t\geq 0}$ be a process behaving as a general increasing Lévy process (subordinator) prior to hitting a given unknown level m_0, then behaving as another different subordinator once this threshold is crossed. This paper addresses the detection of this unknown threshold $m_0 \in [0, +\infty]$ from an observed trajectory of the process. These kind of model and issue are encountered in many areas such as reliability and quality control in degradation problems. More precisely, we construct, from a sample path and for each $\epsilon > 0$, a so-called detection level L_ϵ by considering a CUSUM inspired procedure. Under mild assumptions, this level is such that, while m_0 is infinite (i.e. when no changes occur), its expectation $\mathbb{E}_\infty(L_\epsilon)$ tends to $+\infty$ as ϵ tends to 0, and the expected overshoot $\mathbb{E}_{m_0}([L_\epsilon - m_0]^+)$, while the threshold m_0 is finite, is negligible compared to $\mathbb{E}_\infty(L_\epsilon)$ as ϵ tends to 0. Numerical illustrations are provided when the Lévy processes are gamma processes with different shape parameters.

Keywords: Change detection, Lévy process, subordinator, CUSUM, sequential testing.

1. Introduction

Let $\boldsymbol{X} = (X_t)_{t\geq 0}$ be a continuous stochastic process. More precisely, we here assume that $\boldsymbol{X}$ is a monotone increasing Lévy process prior to hitting m_0, then it is another increasing Lévy process once this threshold is crossed. The process $\boldsymbol{X} = (X_t)_{t\geq 0}$ is observed on-line and $m_0 \in [0, +\infty]$ is unknown. In this context, it naturally appears the need to detect the change, if it took place. One could think about processes that describe gradual deterioration due to continuous use such as erosion, corrosion, concrete creep, crack propagation (see [1] in the case of a gamma process). A unit system could have an accelerating degradation whenever its degradation level crosses m_0. This latter is usually unknown and an abrupt change in the degradation could be one of the causes that complicate the determination of the failure time of the system. Recall that industries seek to maintain their equipment's available while minimizing their total maintenance cost including the unavailability cost. For that purpose, many maintenance policies have been proposed in the literature. The maintenance decision could be based on many characteristics such as the reliability level, the cost function, the remaining useful life or the change time. We here are interested in the change time characteristic.

Recall that the theory of change detection consists in developing tools to detect the change as soon as possible and by taking into consideration the false alarm constraints. Many works in the literature have studied the online change detection for continuous-time stochastic processes which are Lévy processes. In [2] and [3], the authors studied the problem of Poisson disorder problem which seeks to get a stopping time which is close to the time of disorder (or change-point) when the intensity of an observed Poisson process changes from one value to another one at a certain (unknown) time. A revisited version of the latter work has been proposed by [4] by providing a complete solution of the Poisson disorder problem. Recently, [5] proposed to solve the quickest drift change detection problem for a Lévy process under the Bayesian set-up by assuming an exponential a priori distribution of the change point. Furthermore, the authors in [6, 7, 8] have studied the online change detection for a gamma process in the framework of the condition-based maintenance strategy. They used the classical CUSUM rule to determine the change time. As for [9], they considered the change detection problem for continuous-time Lévy processes by approximating an adapted sequence of change-point problems and where the optimality of a CUSUM rule is shown. To sum up, in the previous works, the proposed techniques were based on an a priori distribution for the change-time or a deterministic unknown change-time.

The aim of this paper is to propose a detection level rule to ensure a quick detection when the degradation of a unit system crosses m_0 while minimizing the false alarm rate. For that purpose, we consider a procedure inspired from a CUSUM detection rule applied to intervals between jumps larger than some given constant, rather than on the increments of the process as it is the case in the usual setup. Moreover, in the classic methodology of on-line change detection, the change is related to the temporal aspect. As a natural consequence, the performance criteria of the change detection rules result from the mean time between false alarms when there is no change and the mean time before the detection of a change. These quantities are named the Average Run Lengths (and denoted respectively ARL_∞ and ARL_0). However, in our case, the change is no more related to a temporal aspect but rather on what we could call a spatial aspect: the change takes place when the system reaches a given level m_0. Consequently, we here consider, instead of the Average Run Length, a kind of Average Run Level ($ARLev$) criterion for the evaluation of the detection rules. Roughly speaking, we are interested by the accumulated level since the change rather than the delay before the detection. To motivate this approach, if we consider again the context of the accumulative deterioration of a system, one can imagine that the level of the accumulated degradation since the change occurred is just as important as the delay for detection (even if, obviously, the two are related).

The remainder of the paper is as follows. Section 2 provides a quick presentation of the CUSUM procedure. The proposed methodology along with the main results are presented in Section 3. In Section 4, the proofs of the main results are exposed. Finally, numerical results are given in Section 5. In particular, the present detection procedure is compared to a naive CUSUM approach and is shown to perform better on some examples.

2. A review of the CUSUM procedure for discrete time observed sequences

In the classic online change-point detection problem, it is assumed that a sequence of i.i.d. random variables $Z_1, Z_2, \dots$, with probability density function (pdf) f_1, is observed sequentially, until a change occurs at an unknown instant denoted $K \in \mathbb{N}$. After the change, the observations $Z_K, Z_{K+1}, \dots$ are again i.i.d. but with a pdf f_2 such that $f_2 \neq f_1$. K is called the change time. We can write:

$$Z_k = Z_k^1 \mathbb{1}_{[k<K]} + Z_k^2 \mathbb{1}_{[k \geq K]}, \quad k \in \mathbb{N}, \tag{1}$$

where $(Z_n^1)_{n\in\mathbb{N}}$ and $(Z_n^2)_{n\in\mathbb{N}}$ are two i.i.d. sequences with respectively common pdf f_1 and f_2.

Whenever a new observation is collected, a decision must be made: either there is no evidence against the hypothesis of no change and the system is declared under control and the monitoring continues, waiting for the next observation, or there is clear evidence that a change occurred in the past and an alarm is issued resulting in the monitoring to stop. This decision is taken from a change-detection rule, whose aim is to detect as soon as possible the change from the two operating modes, guaranteeing a low false alarm rate.

The unknown change time K can be either deterministic or random. In the latter case, K is supposed to be a random variable which can be dependent on the observations or completely independent of the observations. The change-point detection rules then rely on a prior distribution, that means a sequence of probabilities $\pi_n = P[K = n | Z_1, \ldots, Z_n]$ for $n \in \mathbb{N}$. In this bayesian setting, [10] obtains the asymptotic optimality of a rule based on a likelihood ratio in the case where the prior distribution of the change time is supposed to be geometric and independent of the observations. Since then, its works were extended to non-independent observations, to more general prior distributions and considering several optimality criteria (see for example [11], [12] or even [13] for an overview). Up to our knowledge, in all the previous papers, a prior distribution (usually the geometric distribution) on the change time should be assumed. As for our case, the procedure is different since we are here interested in the level detection time. As a consequence, the latter cannot have an imposed prior distribution.

In the non-Bayesian setting, the most popular change-detection rule is probably the CUSUM, initially proposed by [14] in 1954. The CUSUM consists in constructing some likelihood ratio between two hypothesis:

- H_1: $K = +\infty$
- H_2: $K < \infty$

At time n, the CUSUM statistic is defined by

$$g_n = \max_{1 \leq k \leq n} \sum_{i=k}^{n} \log \frac{f_2(Z_i)}{f_1(Z_i)}. \tag{2}$$

The CUSUM stopping rule is then defined by

$$\tau_{CUSUM} = \inf\{n \geq 1 : g_n \geq \gamma\} \tag{3}$$

where $\gamma > 0$ refers to a given threshold. It is standard, as $\gamma > 0$, that the test statistic in the definition (3) of τ_{CUSUM} can be written, equivalently, in the following recursive form:

$$g_{n+1} = \left(g_n + \log \frac{f_2(Z_{n+1})}{f_1(Z_{n+1})} \right)^+ \tag{4}$$

with $g_0 = 0$ and $a^+ = \max(a, 0)$, see [15, (2.2.9) p.38], or in the following form

$$g_n = \left\{ \max_{1 \leq k \leq n} \sum_{i=k}^{n} \log \frac{f_2(Z_i)}{f_1(Z_i)} \right\} \vee 0 \tag{5}$$

which will be used throughout the paper.

Let us define the set $\mathcal{F}$ of all monitoring schemes, i.e. of all stopping times adapted to the filtration induced by $(Z_i)_{i \in \mathbb{N}}$. For all positive real number h, let us define $\mathcal{F}_h$ by

$$\mathcal{F}_h = \left\{ T \in \mathcal{F} : E_T^1 := \mathbb{E}^{(\infty)}(T) \geq h \right\}, \tag{6}$$

i.e. the set of all monitoring schemes such that the mean time before a false alarm is larger than h. Let also the worst mean delay for detection be defined as

$$E_T^2 = \sup_{K \geq 1} \text{ess sup} \mathbb{E}^{(K)}[(T - K + 1)^+ | Z_1, \ldots, Z_{K-1}]. \tag{7}$$

In Equations (6) and (7), $\mathbb{E}^{(\infty)}$ and $\mathbb{E}^{(K)}$ respectively refer to the expectation with respect to the probability distribution $\mathbb{P}^{(\infty)}$ of no change and probability distribution $\mathbb{P}^{(K)}$ of a change at time K.

In [16], the asymptotic optimality of the CUSUM is obtained by showing that

$$E^1_{\tau_{CUSUM}} \geq \exp(\gamma), \tag{8}$$

meaning that τ_{CUSUM} in (3) belongs to $\mathcal{F}_h$ by considering $\gamma = \log h$, and that

$$E^2_{\tau_{CUSUM}} \sim \frac{\log h}{I} \text{ as } h \to \infty \tag{9}$$

$$\sim \inf_{T \in \mathcal{F}_h} E_T^2 \text{ as } h \to \infty, \tag{10}$$

where I refers to the Kullback-Leibler (KL) distance between the two distributions with respective densities f_1 and f_2, which is defined as follows

$$I = KL(f_2 || f_1) = \int \log\left(\frac{f_2(x)}{f_1(x)}\right) f_2(x) dx. \tag{11}$$

Among the most significant extended results of [16] on the CUSUM rule, we may cite [17] who obtains the optimality of the CUSUM for a fixed value for h (i.e. a non asymptotic optimality result) and [18] who considers dependent observations, and also several other optimality criteria.

3. Description of the model and main results

Let us now consider an increasing stochastic process $\boldsymbol{X} = (X_t)_{t \geq 0}$ which behaves as one of two given processes $\boldsymbol{X^1}$ or $\boldsymbol{X^2}$, depending on whether it is below or above a certain (unknown) threshold $m_0 \in [0, +\infty]$. More precisely, both processes $\boldsymbol{X^1} = (X_t^1)_{t \geq 0}$ and $\boldsymbol{X^2} = (X_t^2)_{t \geq 0}$ are Lévy processes with respective characteristic exponents

$$\psi_j(\theta) = i a_j \theta + \int_{\mathbb{R}} (1 - \exp(i\theta x) + i\theta x \mathbb{1}_{[|x|<1]}) Q^j(dx)$$

where $\theta \in \mathbb{R}$ and $Q^j(.), j = 1, 2$ refers to the Lévy measure that verifies

$$\mathbb{E}[\exp(i\theta X_t^j)] = \exp(-t\psi_j(\theta)), \quad \forall t \geq 0,$$

see [19, Section 1.1 p.4]. Let us consider the following assumptions

($\mathbf{A_1}$) $Q^j(-\infty, 0) = 0$, $d_j = -(a_j + \int_{(0,1)} x Q^j(dx)) = 0$ and $\int_{(0,\infty)} (1 \wedge x) Q^j(dx) < \infty$, $j = 1, 2$, i.e. both processes $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ are driftless subordinators (see [19, Lemma 2.14, p.55]),

($\mathbf{A_2}$) $\int_{(0,\infty)} x^2 Q^j(dx) < \infty$ (this implies that $\int_{(0,\infty)} x Q^j(dx) < \infty$ based on ($\mathbf{A_1}$)),

($\mathbf{A_3}$) $Q^j((0,+\infty)) = +\infty$ for $j=1$ or $j=2$, and possibly both at the same time,

($\mathbf{A_4}$) $\liminf_{\epsilon\to 0}\left|\frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} - 1\right| > 0$, where $\overline{Q}^j(\epsilon) = Q^j((\epsilon,+\infty))$ for $j=1,2$.

These assumptions may be interpreted as follows. ($\mathbf{A_3}$) means that at least one of the two processes ($\boldsymbol{X^1}$ or $\boldsymbol{X^2}$) admits infinitely many small jumps in a finite time interval. Assumption ($\mathbf{A_4}$) is used to discriminate between $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$. Indeed, $\overline{Q}^j(\epsilon)$ represents the average number of jumps greater than ϵ by unit of time [20, Remark 1.16 p.318]. Intuitively speaking, this assumption says that the order of magnitude of the number of jumps larger than ϵ is different when ϵ tends to 0. Assumptions ($\mathbf{A_1}$), ($\mathbf{A_2}$) and ($\mathbf{A_4}$) are supposed to hold throughout the paper. In the sequel, the study of the proposed detection level rule will be done when Assumption ($\mathbf{A_3}$) holds but also when it does not hold, i.e. when $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ are both Compound Poisson processes. Also, both processes $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ being subordinators (see ($\mathbf{A_1}$)), we recall that they may be characterized by their Laplace exponent $\phi_j(\theta)$, instead of the characteristic exponent $\psi_j(\theta)$, $j=1,2$, that verify $\mathbb{E}[\exp(-\alpha X_t^j)] = \exp(-t\phi_j(\alpha))$ for all $t\geq 0$ and $\alpha \geq 0$. Since the subordinators are driftless processes, those Laplace exponent have here the simple expression

$$\phi_j(\alpha) = \int_{(0,+\infty)} (1-e^{-\alpha x}) Q^j(dx), \quad \alpha \geq 0, \quad j=1,2,$$

see [21, Section 2.2 page 9].

The latter processes allow for the process $\boldsymbol{X}$ to be now written as

$$X_t = X_t^1 \mathbb{1}_{[t\leq \tau_{m_0}]} + (X^2_{t-\tau_{m_0}} + X^1_{\tau_{m_0}})\mathbb{1}_{[t>\tau_{m_0}]}, \quad \forall t \geq 0, \tag{12}$$

with the crossing time of level m_0 of the process $\boldsymbol{X^1}$ defined as

$$\tau_{m_0} = \inf\{t\geq 0 | X_t^1 \geq m_0\}, \tag{13}$$

where $m_0 \in [0,\infty]$ is unknown.
Moreover, following [19, Section 2.4, p.44], and because $\mathbf{X}$ is driftless, it will be convenient to note that, thanks to ($\mathbf{A_2}$) and more precisely to the fact that $\int_{(0,\infty)} xQ^j(dx)$ is finite, each process $\mathbf{X^j}$, $j=1,2$, may be expressed in function of a Poisson random measure N_j on $([0,+\infty)\times(0,+\infty), \mathcal{B}([0,+\infty))\times\mathcal{B}((0,+\infty)), \eta^j)$ as follows

$$X_t^j = \int_{[0,t]}\int_{(0,+\infty)} x N_j(ds\times dx), \quad t\geq 0, \tag{14}$$

where η^j refers to a measure on $([0,+\infty)\times(0,+\infty), \mathcal{B}([0,+\infty))\times\mathcal{B}((0,+\infty))$ given by $\eta^j(ds\times dx) = ds\, Q^j(dx)$.
Throughout the paper, $\mathbb{E}_{m_0}(.)$ refers to the expectation under the assumption that the process $\boldsymbol{X}$ change its behavior when it exceeds m_0 and $\mathbb{E}_\infty(.)$ to the expectation where $m_0=\infty$, i.e. when there is no change in the behavior of $\boldsymbol{X}$, so that $X_t = X_t^1$ for all $t\geq 0$, see (12).
The aim of the paper is to determine, by sequentially observing a sample path $t\geq 0 \mapsto X_t$, a detection rule which achieves the two following goals:

1. To guarantee a quick detection of the crossing of level m_0 if it occurs, such that the overshoot of the underlying process at the time of detection is not too large with respect to the fixed threshold m_0.

2. To provide a low false alarm rate if there is no change, that means $m_0 = +\infty$.

This will be achieved by proving the forthcoming main Theorems 2 and 3. Theorem 2 holds when $(\mathbf{A_3})$ is satisfied and Theorem 3 holds when $(\mathbf{A_3})$ is not satisfied.

3.1. Construction of the ϵ-detection rule

Let $\epsilon > 0$ and $(T_i^\epsilon)_{i\in\mathbb{N}}$ the sequence constructed from the observed trajectory X_t, $t \geq 0$, in the following way:

$$\begin{aligned} T_0^\epsilon &= 0, \\ T_{i+1}^\epsilon &= \inf\{t > T_i^\epsilon |\ \Delta X_t = X_t - X_{t-} > \epsilon\}, \quad i \geq 0. \end{aligned} \tag{15}$$

Thus, the T_i^ϵ, $i \in \mathbb{N}$, are the successive times when an observed jump of the process $\mathbf{X}$ is larger than a given ϵ. The corresponding inter-arrival times are then defined by:

$$\eta_i^\epsilon = T_i^\epsilon - T_{i-1}^\epsilon, \quad i \geq 1. \tag{16}$$

Intuitively, due to the Lévy nature of the processes $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$, η_i^ϵ seems to be exponentially distributed with parameter $\bar{Q}^1(\epsilon)$ if $\tau_{m_0} > T_i^\epsilon$ or parameter $\bar{Q}^2(\epsilon)$ if $\tau_{m_0} < T_{i-1}^\epsilon$. In other words, η_i^ϵ looks either distributed as $\mathcal{E}(\bar{Q}^1(\epsilon))$ or $\mathcal{E}(\bar{Q}^2(\epsilon))$ whether we are before or after having crossed the threshold. Although this latter statement is not rigorous and is only intuitively correct, this however motivates the use of the following CUSUM type rule. More precisely, we introduce the CUSUM statistic $(G_n^\eta)_{n\in\mathbb{N}}$ associated to the sequence $(\eta_i^\epsilon)_{i\in\mathbb{N}}$ as in (4) by $G_0^\eta = 0$ and

$$G_{n+1}^\eta = \left(G_n^\eta + \phi_\epsilon(\eta_{n+1}^\epsilon)\right)^+, \ n \geq 0 \tag{17}$$

where

$$x \in [0, +\infty) \mapsto \phi_\epsilon(x) := \log\frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} + (-\bar{Q}^2(\epsilon) + \bar{Q}^1(\epsilon))x := a_\epsilon + b_\epsilon x \tag{18}$$

is the logarithm of the likelihood ratio of the exponential distributions with respective parameters $\bar{Q}^2(\epsilon)$ and $\bar{Q}^1(\epsilon)$.

Remark 1. One can verify easily that $\phi_\epsilon(.)$ is linear, and is increasing (resp. decreasing) when $\bar{Q}^2(\epsilon) \leq \bar{Q}^1(\epsilon)$ (resp. $\geq$).

Finally, we set the associated stopping rule related to a threshold $\gamma(\epsilon)$ (which will be made explicit later on), by

$$\tau_{CUSUM}^{\epsilon,\eta} := \inf\{n \geq 0 |\ G_n^\eta \geq \gamma(\epsilon)\}, \tag{19}$$

and the associated "pseudo-level":

$$M_\epsilon = \sum_{i=1}^{\tau_{CUSUM}^{\epsilon,\eta}} \Delta X_{T_i^\epsilon} = \sum_{i=1}^{\tau_{CUSUM}^{\epsilon,\eta}} \left(X_{T_i^\epsilon} - X_{T_i^\epsilon -}\right). \tag{20}$$

The idea behind the above construction is that the smaller ϵ is, the closer M_ϵ is from $\boldsymbol{X}$ at the detection time defined by

$$d^\epsilon = \sum_{i=1}^{\tau_{CUSUM}^{\epsilon,\eta}} \eta_i^\epsilon. \tag{21}$$

In this case, we will denote by L_ϵ the detection level defined as follows:

$$L_\epsilon = X_{d^\epsilon}. \tag{22}$$

Figure 1 displays all the quantities above that we have used for the construction of the detection rule. One should note the importance of comparing the proposed model with the

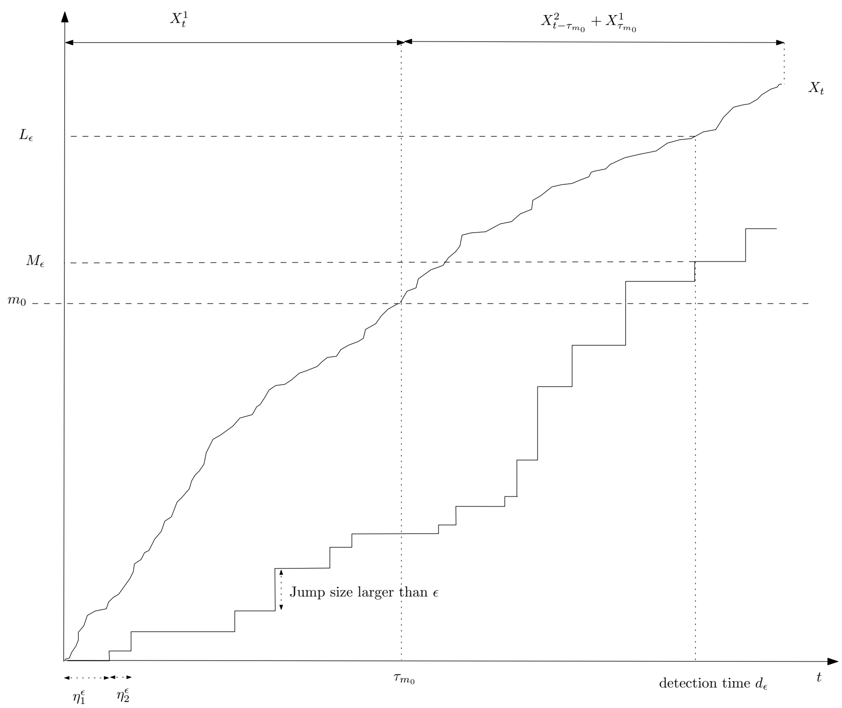

Figure 1: Detection time and level

classical one described in Section 2 where the model features an unknown (but not random) instant change of the process behavior. In our case, the process behavior changes once it exceeds a certain threshold. The main idea of this paper is to exchange the role of time and space, so that the analog of time detection τ_{CUSUM} in (3) is the level detection L_ϵ in (22).

3.2. Main results

The three following theorems are the main contributions of the paper.

Theorem 1. *Assume that* $(\mathbf{A_1}) - (\mathbf{A_4})$ *hold and define the CUSUM statistics* $\tau^{\epsilon,\eta}_{CUSUM}$ *in* (19) *with* $\gamma(\epsilon) := \log h(\epsilon)$ *and*

$$h(\epsilon) := \begin{cases} [\bar{Q}^2(\epsilon) I^\epsilon]^2 & \textit{if} \quad \bar{Q}^2(0) = +\infty, \\ [\bar{Q}^1(\epsilon)]^\beta & \textit{if} \quad \bar{Q}^2(0) < +\infty \ (\textit{and consequently} \ \ \bar{Q}^1(0) = +\infty), \end{cases} \tag{23}$$

where $\beta > 2$ *is arbitrary, and*

$$I^\epsilon = \log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} - 1 + \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}, \forall \epsilon > 0. \tag{24}$$

Then, for all $\epsilon > 0$, there exist two positive quantities c_1^ϵ and c_2^ϵ such that the following properties are satisfied:

(P_1^ϵ) $\mathbb{E}_{m_0}([M_\epsilon - m_0]^+) \leq c_1^\epsilon$, *for all* $\epsilon > 0$,

(P_2^ϵ) $\mathbb{E}_\infty(M_\epsilon) \geq c_2^\epsilon$, *for all* $\epsilon > 0$,

(P_3) $c_1^\epsilon = o(c_2^\epsilon)$ *and* $\lim_{\epsilon \to 0} c_2^\epsilon = +\infty$.

In the case where the following assumption holds

$$\limsup_{\epsilon \to 0} \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} < \infty, \tag{25}$$

then **(P_3)** *can be made more precise as*

(P_3') $\limsup_{\epsilon \to 0} c_1^\epsilon < \infty$ *and* $\lim_{\epsilon \to 0} c_2^\epsilon = +\infty$.

Theorem 2. *Under the assumptions of Theorem 1, it holds that*

$$\lim_{\epsilon \to 0} \mathbb{E}_{m_0}(L_\epsilon - M_\epsilon) = 0, \quad \forall m_0 \geq 0, \tag{26}$$

and the properties **(P_1^ϵ)**, **(P_2^ϵ)**, **(P_3)** *and* **(P_3')** *are still valid by substituting M_ϵ by L_ϵ.*

Remark 2. *Note that $\mathbb{E}_{m_0}([L_\epsilon - m_0]^+)$ corresponds to the mean overshoot of the process* $\mathbf{X}$ *at detection time above the threshold m_0. In a degradation context, this corresponds to the mean degradation above m_0. $\mathbb{E}_\infty(L_\epsilon)$ is the mean detection level when there is no regime change, which will be denoted as the Average Run Level ($ARLev_\infty$). This latter quantity is the analog of the Average Run Length (ARL) in the usual temporal context for detection rules. Properties* **(P_1^ϵ)** *and* **(P_2^ϵ)** *in Theorem 2, applied to L_ϵ, combined with property* **(P_3)** *ensure that, as ϵ tends to 0, the mean overshoot of the process is negligible against the mean level when there is no regime change $\mathbb{E}_\infty(L_\epsilon)$. Even better,* **($P_3'$)** *guarantees that the mean overshoot is bounded, which is useful in practical situations. Moreover, in some specific cases (such as when $\boldsymbol{X}^1$ is a gamma process or an inverse Gaussian process), it can be easily verified from the proofs of Theorems 1 and 2 that the quantities c_ϵ^1 and c_ϵ^2 could be explicitly expressed for a fixed ϵ.*

When the assumption ($\mathbf{A_3}$) is not satisfied, the processes $\mathbf{X}^1$ and $\mathbf{X}^2$ are compound Poisson processes. In that case, the detection procedure is slightly different. The jump and inter-arrival times of the process $\mathbf{X}$ are defined as

$$\begin{aligned} T_0 &:= 0, \\ T_{i+1} &:= \inf\{t > T_i |\ \Delta X_t = X_t - X_{t-} > 0\}, \quad i \geq 0, \\ \eta_i &:= T_i - T_{i-1}, \quad i \geq 1. \end{aligned}$$

The sequence $(T_i)_{i\in\mathbb{N}}$ corresponds to the jump times of $\mathbf{X}^1$ when T_i is less than τ_{m_0} (i.e. before the process crossed the level m_0), and to the jump times of $\mathbf{X}^2$ when T_i is larger than τ_{m_0}. In other word, this corresponds to $(T_i^\epsilon)_{i\in\mathbb{N}}$ in (15) with $\epsilon = 0$. We aim here at devising a "classical" CUSUM rule to the sequence $(\eta_i)_{i\in\mathbb{N}}$ defined above. More precisely we let the associated CUSUM statistic $(G_n)_{n\in\mathbb{N}}$, defined in (17), with $G_0 = 0$ and

$$G_{n+1} = (G_n + \phi_0(\eta_{n+1}))^+, \ n \geq 0,$$

where we recall from (18) that $\phi_0(x) = \log \frac{\bar{Q}^2(0)}{\bar{Q}^1(0)} + (-\bar{Q}^2(0) + \bar{Q}^1(0))x$, and

$$\tau^{\eta}_{CUSUM} := \inf\{n \geq 0 |\ G_n \geq \log h\}, \quad h > 1. \tag{27}$$

Finally, the associated level is defined as

$$L = \sum_{i=1}^{\tau^{\eta}_{CUSUM}} \Delta X_{T_i} = \sum_{i=1}^{\tau^{\eta}_{CUSUM}} \left(X_{T_i} - X_{T_i^-} \right). \tag{28}$$

The equivalent of Theorems 1 and 2 in the present case is as follows.

Theorem 3. *Let us suppose that Assumptions* $(\mathbf{A_1}) - (\mathbf{A_2})$ *hold and that* $\mathbf{X}^1$ *and* $\mathbf{X}^2$ *are compound Poisson processes with finite different intensities* $\bar{Q}^1(0)$ *and* $\bar{Q}^2(0)$*. Then there exists* c_1^h *and* c_2^h *such that the following properties are satisfied*

($\boldsymbol{P_1^0}$*)* $\mathbb{E}_{m_0}([L - m_0]^+) \leq c_1^h$ *for all* $h > 1$*,*

($\boldsymbol{P_2^0}$*)* $\mathbb{E}_\infty(L) \geq c_2^h$ *for all* $h > 1$*,*

($\boldsymbol{P_4}$*)* $c_1^h = o(c_2^h)$ *as* $h \to +\infty$ *and* $\lim_{h \to +\infty} c_2^h = +\infty$*.*

Remark 3. *Similarly to Remark 2, Theorem 3 identifies the detection level by detecting a change in the respective intensities. When the intensities are equal, then the jumps necessarily have distinct distributions whenever the process* X_t *is less or larger than* m_0*. An intuitive solution for detecting the change would be to perform a CUSUM rule on the successive jump sizes of the process. However, theoretical results on the detection level with this latter solution seem to be difficult to prove, and this case remains an open problem.*

3.3. Example: Gamma Processes

Setting $A : \mathbb{R}_+ \to \mathbb{R}_+$ to be a measurable, increasing and right-continuous function with $A(0) = 0$ and $b > 0$, let us recall that a standard (non homogeneous) gamma process $\boldsymbol{Y} = (Y_t)_{t \geq 0}$, with $A(.)$ as shape function and b as scale parameter (denoted by $\boldsymbol{Y} \sim \Gamma(A(.), b)$), is a stochastic process with independent, non-negative and gamma distributed increments such that $Y_0 = 0$ almost surely. The pdf of an increment $Y_t - Y_s$ (with $0 < s < t$) is given by

$$f(x) = \frac{b^{A(t)-A(s)}}{\Gamma(A(t) - A(s))} x^{A(t)-A(s)-1} \exp(-bx), \forall x \geq 0.$$

Gamma processes are largely used in reliability, notably to model the cumulative deterioration of a system (see [1] for an overview).

When the shape function is linear, i.e. $A(t) = \gamma t$, the gamma process is said to be homogeneous. A homogeneous gamma process is a subordinator.

We consider the case in which $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ are two homogeneous gamma processes, so that

$$\boldsymbol{X^1} \sim \Gamma(\gamma_1 ., b_1) \quad \text{and} \quad \boldsymbol{X^2} \sim \Gamma(\gamma_2 ., b_2)$$

where b_1 and b_2 refer to the scale parameters of $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$, respectively.

The Lévy measures of the process $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ are given by:

$$Q^j(dx) = \gamma_i \frac{1}{x} e^{-b_j x} dx, \quad \text{for} \quad j = 1, 2.$$

If $\gamma_1 \neq \gamma_2$, the change consists in a modification in the mean rate of degradation or the variability of the system considered. One can check that the latter condition is equivalent to the assumption $(\mathbf{A_4})$ and that, when $b_1 = b_2 = b$, an explicit form of ϕ_ϵ can be obtained as follows, for $i \geq 1$:

$$\phi_\epsilon(\eta_i^\epsilon) = \log \frac{\gamma_2}{\gamma_1} + (-\gamma_2 + \gamma_1) g(0, b\epsilon) \eta_i^\epsilon,$$

where $g(.,.)$ is the upper gamma incomplete function.

4. Proof of the main results

The idea of the proof of Theorems 1 and 2 is to approximate the original process $\boldsymbol{X}$ by a somewhat simpler jump process $\boldsymbol{X^\epsilon}$ of which behavior changes too when it crosses m_0 and such that X_t^ϵ converges towards X_t point-wise as $\epsilon \to 0$. The construction of this process $\boldsymbol{X^\epsilon}$ is given in Subsection 4.1. The change detection for the approximating process is easier to deal with. The point of Section 4.2 is to study some properties of the stopping procedure related to $\boldsymbol{X^\epsilon}$ and $\tau_{CUSUM}^{\epsilon,\eta}$ defined in (19).

4.1. The approximating process and its associated CUSUM statistics

Construction of $\boldsymbol{X^\epsilon}$

Let $\epsilon > 0$. The idea here is to approximate $\boldsymbol{X^j}, j = 1, 2$, by $\boldsymbol{X^{j,\epsilon}}$ which is given by (14) by removing the jumps less or equal to ϵ. Consequently, similarly to the representation (14), we can express $\boldsymbol{X^{j,\epsilon}}$ as follows

$$X_t^{j,\epsilon} = \int_{[0,t]} \int_{(\epsilon,+\infty)} x N_j(ds \times dx), \quad t \geq 0. \tag{29}$$

Moreover, it is standard that $X_t^{j,\epsilon}$ is a compound Poisson process (see [19, Lemma 2.8, p.44]) that can be written as

$$X_t^{j,\epsilon} = \sum_{i=1}^{N_X^{j,\epsilon}(t)} \Delta_i^{j,\epsilon}, \forall t \geq 0,$$

with underlying Poisson process $\boldsymbol{N}_X^{j,\epsilon} = (N_X^{j,\epsilon}(t))_{t\geq 0}$, intensity $\overline{Q^j}(\epsilon)$ and jumping times denoted by $(T_i^{j,\epsilon})_{i\geq 0}$. The distribution of the associated increments $\Delta_i^{j,\epsilon} = X_{T_i^{j,\epsilon}}^{j,\epsilon} - X_{T_{i-1}^{j,\epsilon}}^{j,\epsilon}$ is given by

$$\frac{1}{\overline{Q^j}(\epsilon)} Q^j(dx) \mathbb{1}_{[x>\epsilon]}. \tag{30}$$

In the following, we will let $\Delta^{j,\epsilon}$, $j = 1, 2$, be generic random variables with same distribution as the $\Delta_i^{j,\epsilon}$, $i \in \mathbb{N}$.

Once the processes $\boldsymbol{X^{j,\epsilon}}$ are constructed, then the process $\boldsymbol{X^\epsilon}$ can be defined in a similar way of Equation (12)

$$X_t^\epsilon = X_t^{1,\epsilon} \mathbb{1}_{[t \leq \tau_{m_0}^\epsilon]} + (X_{t-\tau_{m_0}^\epsilon}^{2,\epsilon} + X_{\tau_{m_0}^\epsilon}^{1,\epsilon}) \mathbb{1}_{[t > \tau_{m_0}^\epsilon]} \tag{31}$$

with the first crossing time of level m_0 of process $\boldsymbol{X}^{1,\epsilon}$ defined as

$$\tau_{m_0}^\epsilon = \inf\{t \geq 0 | X_t^{1,\epsilon} \geq m_0\}. \tag{32}$$

At this point, we may note that the pseudo-level in (20) may be conveniently expressed as

$$M_\epsilon = X_{d^\epsilon}^\epsilon \tag{33}$$

where we recall that d^ϵ is the detection time defined by (21).

CUSUM statistics associated to $\boldsymbol{X^\epsilon}$

We then define the jumping times of the process $\boldsymbol{X^\epsilon}$ by

$$U_i^\epsilon = T_i^{1,\epsilon}\mathbb{1}_{[i\leq \mathcal{N}^{1,\epsilon}]} + (T_{i-\mathcal{N}^{1,\epsilon}}^{2,\epsilon} + T_{\mathcal{N}^{1,\epsilon}}^{1,\epsilon})\mathbb{1}_{[i>\mathcal{N}^{1,\epsilon}]}, i \geq 0, \tag{34}$$

where

$$\mathcal{N}^{1,\epsilon} = \inf\left\{ n \in \mathbb{N} \;\middle|\; \sum_{i=1}^{n} \Delta_i^{1,\epsilon} \geq m_0 \right\}. \tag{35}$$

Note that $\mathcal{N}^{1,\epsilon}$ refers to the index of the jumping time where $\boldsymbol{X^{1,\epsilon}}$ exceeds m_0, i.e. that $\mathcal{N}^{1,\epsilon} = \inf\{i \in \mathbb{N} | T_i^{1,\epsilon} \geq \tau_{m_0}^\epsilon\}$. We also define the index of the first jumping time of the process $\boldsymbol{X}^{1,\epsilon}$ following τ_{m_0} by

$$\mathcal{D}^\epsilon = \inf\{i \in \mathbb{N} | T_i^{1,\epsilon} \geq \tau_{m_0}\}. \tag{36}$$

We define the inter-arrival times of the jump processes $\boldsymbol{X^{j,\epsilon}}$ and $\boldsymbol{X^\epsilon}$ as

$$\begin{aligned} \eta_i^{j,\epsilon} &= T_i^{j,\epsilon} - T_{i-1}^{j,\epsilon}, \quad j = 1, 2, \; i \geq 1, \\ \delta_i^\epsilon &= U_i^\epsilon - U_{i-1}^\epsilon, \quad i \geq 1. \end{aligned}$$

Thus, $(\eta_i^\epsilon)_{i\in\mathbb{N}}$ defined in (16) is associated to the observed process $\boldsymbol{X}$, $(\eta_i^{j,\epsilon})_{i\in\mathbb{N}}$ referred to the inter-arrival times of the process $\boldsymbol{N}_X^{j,\epsilon}$, $j = 1, 2$, and $(\delta_i^\epsilon)_{i\in\mathbb{N}}$ is related to the constructed process $\boldsymbol{X^\epsilon}$. A trajectory of $\boldsymbol{X}$ and $\boldsymbol{X^\epsilon}$ as well as the corresponding stopping times τ_{m_0} and $\tau_{m_0}^\epsilon$ on crossing level m_0 are illustrated in Figure 2. An illustration of the construction of the above sequences is given in Figure 3. As in (17) for the sequence $(\eta_i^\epsilon)_{i\in\mathbb{N}}$, let us now introduce the CUSUM statistic $(G_n^\delta)_{n\in\mathbb{N}}$ associated to the sequence $(\delta_i^\epsilon)_{i\in\mathbb{N}}$ by $G_0^\delta = 0$ and

$$G_{n+1}^\delta = \left(G_n^\delta + \phi_\epsilon(\delta_{n+1}^\epsilon)\right)^+, \; n \geq 0, \tag{37}$$

and the resulting stopping rule related to the threshold $\gamma(\epsilon)$

$$\tau_{CUSUM}^{\epsilon,\delta} := \inf\{n \geq 0 | \; G_n^\delta \geq \gamma(\epsilon)\}. \tag{38}$$

Outline of Theorems 1, 2, 3 proofs

The proofs of Theorems 1, 2 have different methodologies than that of Theorem 3. More precisely, the sketch of the proof of Theorems 1, 2 is the following:

1. We perform a detection of the random index $\mathcal{N}^{1,\epsilon}$ defined in (35) through the CUSUM procedure $\tau_{CUSUM}^{\epsilon,\delta}$.
2. We prove that $\tau_{CUSUM}^{\epsilon,\eta}$ (that depends on the observed process $\boldsymbol{X}$) is stochastically smaller than $\tau_{CUSUM}^{\epsilon,\delta}$ (that depends on the unobserved process $\boldsymbol{X^\epsilon}$). This part is one of the main difficulty of the proof and is the object of the forthcoming Proposition 1 in Section 4.2.
3. Since X_t^ϵ point-wise converges to X_t as $\epsilon \to 0$, we argue that the so-called pseudo-level M_ϵ and detection level L_ϵ defined respectively in (20) and (22), which depend on the observed inter-arrival times $(\eta_i^\epsilon)_{i\in\mathbb{N}}$ and the associated jumps $(\Delta X_{T_i^\epsilon})_{i\in\mathbb{N}}$, satisfy the properties $(\boldsymbol{P_1^\epsilon})$, $(\boldsymbol{P_2^\epsilon})$ and $(\boldsymbol{P_3})$ or $(\boldsymbol{P_3'})$.

The proof of Theorem 3 is on the other hand simpler, as we let $\epsilon = 0$ i.e. we directly deal with the observed process $\boldsymbol{X}$ and construct consequently the CUSUM statistic that detects the index $\mathcal{N}^{1,\epsilon}$.

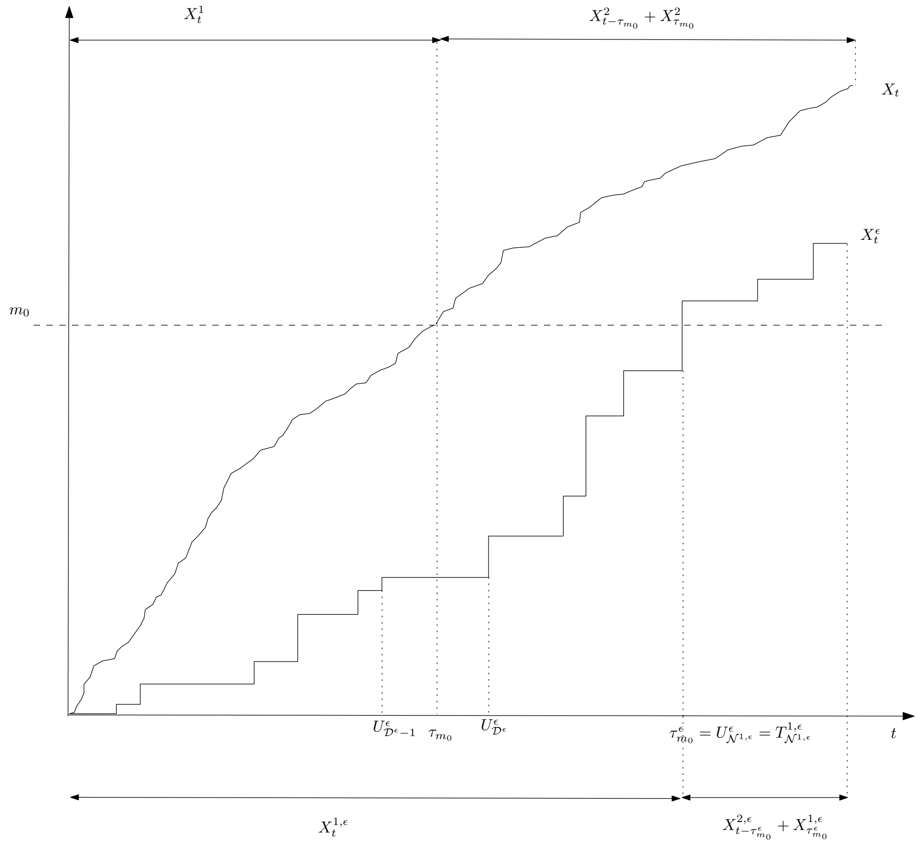

Figure 2: Trajectories for $\boldsymbol{X}$ and $\boldsymbol{X}^\epsilon$ and switching times on reaching level m_0.

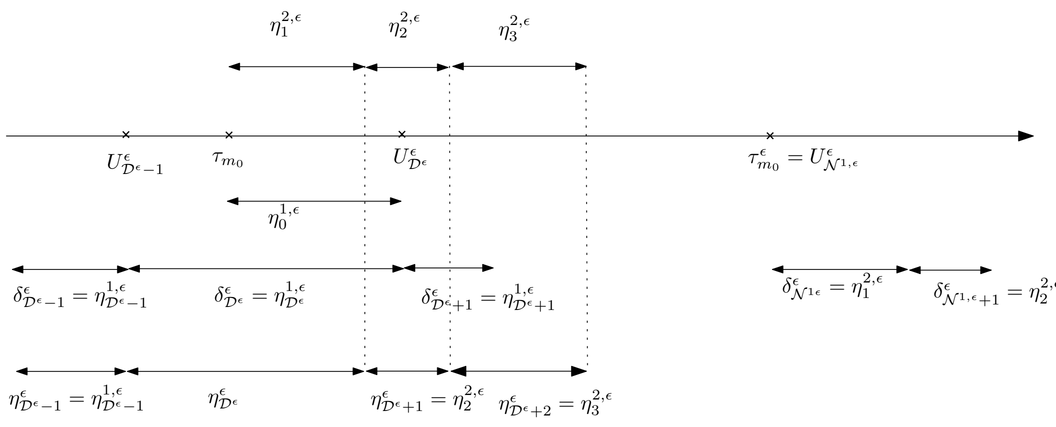

Figure 3: Illustration of definitions for the sequences $(\eta_i^\epsilon)_{i\in\mathbb{N}}$, $(\delta_i^\epsilon)_{i\in\mathbb{N}}$ and $(\eta_i^{j,\epsilon})_{i\in\mathbb{N}}$, $j=1,2$. Recall that $(\eta_i^\epsilon)_{i\in\mathbb{N}}$ is related to the observed process $\boldsymbol{X}$ and $(\delta_i^\epsilon)_{i\in\mathbb{N}}$ is related to the unobserved process $\boldsymbol{X}^\epsilon$.

4.2. Intermediary results

Firstly, one should note that, in view of the explicit expression (5) for the CUSUM statistic, the following equalities hold

$$\{G_n^\delta < \gamma(\epsilon)\} = \left\{ \max_{1\le k\le n} \sum_{i=k}^n \phi_\epsilon(\delta_i^\epsilon) \vee 0 < \gamma(\epsilon) \right\}, \quad n \in \mathbb{N}, \tag{39}$$

$$\{G_n^\eta < \gamma(\epsilon)\} = \left\{ \max_{1\le k\le n} \sum_{i=k}^n \phi_\epsilon(\eta_i^\epsilon) \vee 0 < \gamma(\epsilon) \right\}, \quad n \in \mathbb{N}. \tag{40}$$

One can easily prove by induction on $n \ge m$ the following relation between G_n^δ and G_m^δ

$$G_n^\delta = \left[\max_{m+1\le j\le n} \sum_{r=j}^n \phi_\epsilon(\delta_r^\epsilon) \right] \vee 0 \vee \left[G_m^\delta + \sum_{r=m+1}^n \phi_\epsilon(\delta_r^\epsilon) \right]. \tag{41}$$

One should note that $\tau_{CUSUM}^{\epsilon,\eta}$, which is the CUSUM stopping time based on the observed path $\mathbf{X}$ is based on the sequence $(\eta_i^\epsilon)_{i\in\mathbb{N}}$. Nevertheless, we can see on Figure 3 that the particular time interval $\eta_{\mathcal{D}^\epsilon}^\epsilon$ causes a problem since its distribution is unknown because the regime change does not occur necessarily after a jump of size larger than ϵ (i.e. at time $U_{\mathcal{D}^\epsilon}^\epsilon$), unlike $\tau_{CUSUM}^{\epsilon,\delta}$. The following results, that will be used in the proof of Theorem 1, shows that $\tau_{CUSUM}^{\epsilon,\delta}$ is larger than $\tau_{CUSUM}^{\epsilon,\eta}$ in some sense.
Throughout the following sections and for the sake of readability, we denote $\mathbb{E}_{m_0}$ and $\mathbb{P}_{m_0}$ by $\mathbb{E}$ and $\mathbb{P}$, respectively, when m_0 is finite and when no confusion is possible.
Let us recall that, given X and Y two random variables, X is said to be stochastically smaller than Y, denoted $X \le_{\text{st}} Y$ (see [22, Chapter 1, p.3]) if, for all $x \in \mathbb{R}$:

$$\mathbb{P}(Y \ge x) \ge \mathbb{P}(X \ge x).$$

Proposition 1. *It holds that* $\left[\tau_{CUSUM}^{\epsilon,\eta} |\ \mathcal{N}^{1,\epsilon} = p\right] \le_{st} \left[\tau_{CUSUM}^{\epsilon,\delta} |\ \mathcal{N}^{1,\epsilon} = p\right]$ *for all* $p \in \mathbb{N}^*$.

Proof. Setting $p \in \mathbb{N}^*$, we have to prove the following inequality

$$\mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \ge n,\ \mathcal{N}^{1,\epsilon} = p) \ge \mathbb{P}(\tau_{CUSUM}^{\epsilon,\eta} \ge n,\ \mathcal{N}^{1,\epsilon} = p) \tag{42}$$

for all $n \in \mathbb{N}$. We set throughout the proof

$$\mathcal{G}^\epsilon := G_{\mathcal{D}^\epsilon-1}^\eta + b_\epsilon(\tau_{m_0} - U_{\mathcal{D}^\epsilon-1}^\epsilon) \tag{43}$$

with b_ϵ defined in (18). By construction of the process $\boldsymbol{X}^{\boldsymbol{\epsilon}}$, one has $G_i^\eta = G_i^\delta$ for all $i = 1, \ldots, \mathcal{D}^\epsilon - 1$ (see Figure 3), so that $G_{\mathcal{D}^\epsilon}^\eta$ and $G_{\mathcal{D}^\epsilon}^\delta$ may be expressed as

$$G_{\mathcal{D}^\epsilon}^\eta = \left(\mathcal{G}^\epsilon + \phi_\epsilon(\eta_1^{2,\epsilon})\right)^+, \tag{44}$$

$$G_{\mathcal{D}^\epsilon}^\delta = \left(\mathcal{G}^\epsilon + \phi_\epsilon(\eta_0^{1,\epsilon})\right)^+, \quad \eta_0^{1,\epsilon} := U_{\mathcal{D}^\epsilon}^\epsilon - \tau_{m_0}. \tag{45}$$

A crucial remark is that, since $\boldsymbol{X}^{\boldsymbol{1,\epsilon}}$ has independent increments, $\eta_0^{1,\epsilon}$ defined above is independent from τ_{m_0}, $G_{\mathcal{D}^\epsilon-1}^\delta$ and U_i^ϵ, $i = 1, \ldots, \mathcal{D}^\epsilon - 1$, and hence is in particular independent from $\mathcal{G}^\epsilon$. Also, since $\boldsymbol{X}^{\boldsymbol{1,\epsilon}}$ and $\boldsymbol{X}^{\boldsymbol{2,\epsilon}}$ are independent processes, $\eta_1^{2,\epsilon}$ is independent from

$\mathcal{G}^\epsilon$ and is $\mathcal{E}(\bar{Q}^2(\epsilon))$ distributed. In fact, it will be proved later on that the r.v. $\eta_0^{1,\epsilon}$ is $\mathcal{E}(\bar{Q}^1(\epsilon))$ distributed. Hence, the statement that we want to prove is justified by the following heuristic argument: from (44) and (45), it is clear that both CUSUM statistics G_n^δ and G_n^η coincide up to index $\mathcal{D}^\epsilon - 1$. Starting from index $\mathcal{D}^\epsilon$, these two quantities may then be seen as two CUSUM statistics starting from $\mathcal{G}^\epsilon$, with increments respectively given by $\phi_\epsilon(\eta_0^{1,\epsilon}), \phi_\epsilon(\eta_{\mathcal{D}^\epsilon+1}^{1,\epsilon}), \ldots, \phi_\epsilon(\eta_{\mathcal{N}^\epsilon-1}^{1,\epsilon})$ (cf. Figure 3) which are negative in expectation then switch to $\phi_\epsilon(\eta_1^{2,\epsilon})$, $\phi_\epsilon(\eta_2^{2,\epsilon})$, ..., which are positive in expectation, and $\phi_\epsilon(\eta_1^{2,\epsilon})$, $\phi_\epsilon(\eta_2^{2,\epsilon})$, ..., which are positive in expectation. Consequently, $(G_n^\eta)_{n\in\mathbb{N}}$ will tend to hit the threshold $\gamma(\epsilon)$ before $(G_n^\delta)_{n\in\mathbb{N}}$.
The proof is decomposed into several steps as follows.

Step 1: Decomposition of the event $\{\tau_{CUSUM}^{\epsilon,\delta} \geq n\}$

Let us now decompose the left-hand side of (42) as

$$\begin{aligned}\mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ \mathcal{N}^{1,\epsilon} = p) &= \mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ n \geq \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p) \\ &\quad + \mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ n < \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p) \end{aligned} \tag{46}$$

and consider separately the two terms in the right-hand side of (46). We start by considering $\mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ n < \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p)$. One has

$$\begin{aligned}\left\{\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ n < \mathcal{D}^\epsilon\right\} &= \left\{G_k^\delta < \gamma(\epsilon), k = 1, \ldots, n-1,\ n < \mathcal{D}^\epsilon\right\} \\ &= \left\{G_k^\eta < \gamma(\epsilon), k = 1, \ldots, n-1,\ n < \mathcal{D}^\epsilon\right\} \\ &= \left\{\tau_{CUSUM}^{\epsilon,\eta} \geq n,\ n < \mathcal{D}^\epsilon\right\},\end{aligned}$$

as indeed, for all $i = 1, \ldots, \mathcal{D}^\epsilon - 1$, $\delta_i^\epsilon = \eta_i^\epsilon = \eta_i^{1,\epsilon}$ and using (39) and (40). Hence we have the equality

$$\mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ n < \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p) = \mathbb{P}(\tau_{CUSUM}^{\epsilon,\eta} \geq n,\ n < \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p). \tag{47}$$

The main bulk of the proof concerns the first term in the right-hand side of (46). Conditioning on the crossing time τ_{m_0}, the occurrences of the jump times of $\mathbf{X}^\epsilon$ prior to τ_{m_0} as well as $\mathcal{D}^\epsilon$, and since $\mathcal{N}^{1,\epsilon} \geq \mathcal{D}^\epsilon$, we obtain that

$$\begin{aligned}&\mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n,\ n \geq \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p) \\ &= \int_{t=0}^\infty \sum_{d=1}^{n\wedge p} \int_{t_1 \leq \ldots \leq t_{d-1} \leq t} \mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n, \tau_{m_0} \in dt, \mathcal{D}^\epsilon = d, \mathcal{N}^{1,\epsilon} = p, T_i^{1,\epsilon} \in dt_i, i = 1, \ldots, d-1).\end{aligned} \tag{48}$$

To avoid cumbersome notation, we define

$$\begin{aligned}A &:= \{\tau_{m_0} \in dt, \mathcal{D}^\epsilon = d, \mathcal{N}^{1,\epsilon} = p, T_i^{1,\epsilon} \in dt_i, i = 1, \ldots, d-1\} \\ &= \{\tau_{m_0} \in dt, T_{d-1}^{1,\epsilon} < t, T_d^{1,\epsilon} \geq t, \mathcal{N}^{1,\epsilon} = p, T_i^{1,\epsilon} \in dt_i, i = 1, \ldots, d-1\}\end{aligned} \tag{49}$$

where the last equality comes from the definition (36) of $\mathcal{D}^\epsilon$. We now consider the integrand on the RHS of (48). The rest of the proof is dedicated to prove that $\mathbb{P}(\tau_{CUSUM}^{\epsilon,\delta} \geq n, A) \geq$

$\mathbb{P}(\tau^{\epsilon,\eta}_{CUSUM} \geq n, A)$. To this end, we first observe, according to (38), that

$$\begin{aligned}\{\tau^{\epsilon,\delta}_{CUSUM} \geq n,\ n \geq \mathcal{D}^\epsilon = d\} &= \{G^\delta_k < \gamma(\epsilon),\ k=1,\dots,n-1,\ n \geq \mathcal{D}^\epsilon = d\} \\ &= \{G^\delta_k < \gamma(\epsilon),\ k=1,\dots,d-1\} \bigcap \{G^\delta_d < \gamma(\epsilon)\} \\ &\quad \bigcap \{G^\delta_k < \gamma(\epsilon),\ k=d+1,\dots,n-1\} \bigcap \{\mathcal{D}^\epsilon = d\} \quad (50)\end{aligned}$$

We now consider each event on the RHS of the above equality intersected with A. Thanks to the explicit expression (5) of the CUSUM statistics as well as (39), the first event may be expressed as

$$\{G^\delta_k < \gamma(\epsilon),\ k=1,\dots,d-1\} \bigcap A = \left\{ \max_{1\leq k\leq d-1} \max_{1\leq j\leq k} \sum_{i=j}^{k} \phi_\epsilon(t_i - t_{i-1}) < \gamma(\epsilon) \right\} \bigcap A.$$

Then, thanks to (45), the second event can be written as

$$\{G^\delta_d < \gamma(\epsilon)\} \bigcap A = \left\{ (f_{d-1}(t_1,\dots,t_{d-1},t) + \phi_\epsilon(\eta_0^{1,\epsilon}))^+ < \gamma(\epsilon) \right\} \bigcap A$$

where

$$f_{d-1} : (t_1,\dots,t_{d-1},t) \mapsto f_{d-1}(t_1,\dots,t_{d-1},t) := \max_{1\leq k\leq d-1} \sum_{i=k}^{d-1} \phi_\epsilon(t_i - t_{i-1}) \vee 0 + b_\epsilon(t - t_{d-1}),$$

and we recall that $\eta_0^{1,\epsilon} := U^\epsilon_{\mathcal{D}^\epsilon} - \tau_{m_0}$ refers to the residual time before a jump larger than ϵ after that the process $\mathbf{X}$ crosses the threshold m_0 (see Figure 3). Finally, the third event of (50) intersected with A can be written as follows thanks to (41):

$$\begin{aligned}\{G^\delta_k < \gamma(\epsilon),\ k=d+1,\dots,n\} \bigcap A = \Bigg\{ \max_{d+1\leq k\leq n} \Bigg(\Bigg[\max_{d+1\leq j\leq k} \sum_{r=j}^{k} \phi_\epsilon(\mathbb{1}_{[r\leq p]}\eta_r^{1,\epsilon} + \mathbb{1}_{[r>p]}\eta_{r-p}^{2,\epsilon}) \Bigg] \vee 0 \\ \vee \Bigg[(f_{d-1}(t_1,\dots,t_{d-1},t) + \phi_\epsilon(\eta_0^{1,\epsilon}))^+ + \sum_{r=d+1}^{k} \phi_\epsilon(1_{[r\leq p]}\eta_r^{1,\epsilon} + \mathbb{1}_{[r>p]}\eta_{r-p}^{2,\epsilon}) \Bigg] \Bigg) < \gamma(\epsilon) \Bigg\} \bigcap A.\end{aligned}$$

Hence, we obtain

$$\begin{aligned}&\mathbb{P}(\tau^{\epsilon,\delta}_{CUSUM} \geq n,\ n \geq \mathcal{D}^\epsilon, A) = \mathbb{P}(\tau^{\epsilon,\delta}_{CUSUM} \geq n,\ A) \\ &\quad = \mathbb{P}\left(v(t_1,\dots,t_{d-1}) < \gamma(\epsilon), \Psi(\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}, (\eta_i^{2,\epsilon})_{1\leq i\leq n-p})) < \gamma(\epsilon),\ A \right) \\ &\quad = \mathbb{1}_{\left\{v(t_1,\dots,t_{d-1}) < \gamma(\epsilon)\right\}} \mathbb{P}\left(\Psi(\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}, (\eta_i^{2,\epsilon})_{1\leq i\leq n-p})) < \gamma(\epsilon),\ A \right) \quad (51)\end{aligned}$$

where we define the following two functions

$$v(t_1,\dots,t_{d-1}) = \max_{1\leq k\leq d-1} \left\{ \max_{1\leq j\leq k} \sum_{i=j}^{k} \phi_\epsilon(t_i - t_{i-1}) \right\} \vee 0,$$

$$\Psi : (z_1, \ldots, z_{n-d+1}) \mapsto \max\Bigg\{ (f_{d-1}(t_1, \ldots, t_{d-1}, t) + \phi_\epsilon(z_1))^+, \max_{d+1 \le k \le n} \Bigg(\Bigg[\max_{d+1 \le j \le k} \sum_{r=j}^{k} \phi_\epsilon(z_{r-d+1}) \Bigg] \vee 0$$

$$\vee \Bigg[(f_{d-1}(t_1, \ldots, t_{d-1}, t) + \phi_\epsilon(z_1))^+ + \sum_{r=d+1}^{k} \phi_\epsilon(z_{r-d+1}) \Bigg] \Bigg) \Bigg\}. \quad (52)$$

Note that Ψ depends on (fixed) $t_1, \ldots, t_{d-1}, t$. For the sake of clarity, this dependence is not mentioned, as there is no possible ambiguity in the following.

Step 2: General properties of $\eta_0^{1,\epsilon}$

Let us now argument the fact that $\eta_0^{1,\epsilon}$ is independent from A and is distributed as $\mathcal{E}(\bar{Q}^1(\epsilon))$ i.e.

$$\mathbb{P}(\eta_0^{1,\epsilon} \ge x,\ A) = \exp(-\bar{Q}^1(\epsilon)x)\mathbb{P}(A), \quad \forall x \ge 0. \quad (53)$$

Recalling the definition of $\eta_0^{1,\epsilon}$ in (45), we have $\eta_0^{1,\epsilon} = U^\epsilon_{\mathcal{D}^\epsilon} - \tau_{m_0}$ which can be written as $\eta_0^{1,\epsilon} = T^{1,\epsilon}_{\mathcal{D}^\epsilon} - \tau_{m_0}$, since $\mathcal{N}^{1,\epsilon} \ge \mathcal{D}^\epsilon$ and because of the definition of U_i^ϵ in (34). Additionally, the definition of $\mathcal{D}^\epsilon$ in (36) implies that $\{N_X^{1,\epsilon}(t) = d-1\} = \{\mathcal{D}^\epsilon = d\}$ on $\tau_{m_0} \in dt$. Then, we have for $x \ge 0$,

$$\begin{aligned}\{\eta_0^{1,\epsilon} \ge x, \tau_{m_0} \in dt, \mathcal{D}^\epsilon = d\} &= \{T_d^{1,\epsilon} - t \ge x, \tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1\} \\ &= \{N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0, \tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1\}. \quad (54)\end{aligned}$$

Since $\{\mathcal{N}^{1,\epsilon} = p\} = \left\{\sum_{i=1}^p \Delta_i^{1,\epsilon} \ge m_0 > \sum_{i=1}^{p-1} \Delta_i^{1,\epsilon}\right\}$ and thanks to the definition (49) of the event A, the left-hand side of (53) thus reads

$$\begin{aligned}\mathbb{P}(\eta_0^{1,\epsilon} \ge x,\ A) = \mathbb{P}\Big(& N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0, \tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1, \\ & \sum_{i=1}^{p} \Delta_i^{1,\epsilon} \ge m_0 > \sum_{i=1}^{p-1} \Delta_i^{1,\epsilon}, T_i^{1,\epsilon} \in dt_i, i = 1, \ldots, d-1\Big),\end{aligned}$$

which can be written in integral form, with respect to the first $d-1$ inter-arrival times $\Delta_i^{1,\epsilon}, \ldots, \Delta_{d-1}^{1,\epsilon}$, as follows

$$\int_{z_1,\ldots,z_{d-1} \ge 0} \mathbb{P}\Big(N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0, \tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1,$$

$$\sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p} \Delta_i^{1,\epsilon} \ge m_0 > \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p-1} \Delta_i^{1,\epsilon}, T_i^{1,\epsilon} \in dt_i, \Delta_i^{1,\epsilon} \in dz_i, i = 1, \ldots, d-1\Big). \quad (55)$$

We now observe for fixed $z_1, \ldots, z_{d-1}$, that the event $\{\tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1, T_i^{1,\epsilon} \in dt_i, \Delta_i^{1,\epsilon} \in dz_i, i = 1, \ldots, d-1\}$ depends on X_s^1 for $s \in [0,t]$, whereas $\{N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0, \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p} \Delta_i^{1,\epsilon} \ge m_0 > \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p} \Delta_i^{1,\epsilon}\}$ depends on the increments ΔX_s^1 for $s \ge t$. Since $\boldsymbol{X^1}$ is a Lévy process, its increments after time t are independent from its history up to time t, we hence deduce that (55) can be written as

$$\int_{z_1,\ldots,z_{d-1}\geq 0} \mathbb{P}\Big(N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0, \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p} \Delta_i^{1,\epsilon} \geq m_0 > \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p-1} \Delta_i^{1,\epsilon}\Big)$$
$$\times\, \mathbb{P}\Big(\tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1, T_i^{1,\epsilon} \in dt_i, \Delta_i^{1,\epsilon} \in dz_i, i = 1,\ldots,d-1\Big). \tag{56}$$

Since $\boldsymbol{N_X^{1,\epsilon}}$ is a Poisson process, the random variables $N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t), N_X^{1,\epsilon}(t)$ are independent. Additionally, since $\mathbf{X^{1,\epsilon}}$ is a compound Poisson process, the jump times are independent from the jump size $(\Delta_i^{1,\epsilon})_{i=1,\ldots,p}$ (see Section 4.1). Consequently, (56) can be written as

$$\int_{z_1,\ldots,z_{d-1}\geq 0} \mathbb{P}\Big(N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0\Big) \times \mathbb{P}\Big(\sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p} \Delta_i^{1,\epsilon} \geq m_0 > \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p-1} \Delta_i^{1,\epsilon}\Big)$$
$$\times\, \mathbb{P}\Big(\tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1, T_i^{1,\epsilon} \in dt_i, \Delta_i^{1,\epsilon} \in dz_i, i = 1,\ldots,d-1\Big). \tag{57}$$

Since the intensity of the Poisson process $\boldsymbol{N_X^{1,\epsilon}}$ is $\bar{Q}^1(\epsilon)$, we have $\mathbb{P}\Big(N_X^{1,\epsilon}(t+x) - N_X^{1,\epsilon}(t) = 0\Big) = e^{-\bar{Q}^1(\epsilon)x}$, so that we may write (57) as

$$e^{-\bar{Q}^1(\epsilon)x} \int_{z_1,\ldots,z_{d-1}\geq 0} \mathbb{P}\Big(\sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p} \Delta_i^{1,\epsilon} \geq m_0 > \sum_{i=1}^{d-1} z_i + \sum_{i=N_X^{1,\epsilon}(t)+1}^{p-1} \Delta_i^{1,\epsilon}\Big)$$
$$\times\, \mathbb{P}\Big(\tau_{m_0} \in dt, N_X^{1,\epsilon}(t) = d-1, T_i^{1,\epsilon} \in dt_i, \Delta_i^{1,\epsilon} \in dz_i, i = 1,\ldots,d-1\Big)$$

which is the right-hand side of (53). Similarly, one can easily argument that the vector $(\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}, (\eta_i^{2,\epsilon})_{1\leq i\leq n-p})$ is independent from A and has independent components. Finally, the independence of processes $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ immediately implies that A is independent from $(\eta_1^{2,\epsilon}, \ldots, (\eta_i^{2,\epsilon})_{n-d+1})$.

Step 3: End of proof

Let us assume $\bar{Q}^2(\epsilon) \leq \bar{Q}^1(\epsilon)$. This implies that $(\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}) \leq_{\text{st}} (\eta_1^{2,\epsilon}, \ldots, \eta_{p-d+1}^{2,\epsilon})$. By independence of $\eta_0^{1,\epsilon}$, $(\eta_j^{1,\epsilon})_{j\in\mathbb{N}}$, $(\eta_i^{2,\epsilon})_{i\in\mathbb{N}}$ we thus obtain

$$\begin{aligned}(\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}, (\eta_i^{2,\epsilon})_{1\leq i\leq n-p})) &\overset{\mathcal{L}}{=} (\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}, (\eta_i^{2,\epsilon})_{p-d+2\leq i\leq n-d+1})) \\ &\leq_{\text{st}} (\eta_1^{2,\epsilon}, \ldots, \eta_{n-d+1}^{2,\epsilon}).\end{aligned} \tag{58}$$

The inequality $\bar{Q}^2(\epsilon) \leq \bar{Q}^1(\epsilon)$ entails that ϕ_ϵ defined by (18) is increasing (see Remark 1) which in turn implies that $\Psi : (z_1, \ldots, z_{n-d+1}) \mapsto \Psi(z_1, \ldots, z_{n-d+1})$ (defined in (52)) is an increasing function in each of the variables z_i, $i = 1, \ldots, n-d+1$. Hence, the independence of $(\eta_0^{1,\epsilon}, (\eta_j^{1,\epsilon})_{d+1\leq j\leq p}, (\eta_i^{2,\epsilon})_{1\leq i\leq n-p})$ from A, argued at the end of Step 2, as well as the stochastic order in (58) (see [22, Theorem 1.A.3 (b), p.6]) imply that (51) is upper bounded as follows

$$\begin{aligned}
&\mathbb{P}(\tau^{\epsilon,\delta}_{CUSUM} \ge n,\ A)\\
=\ &\mathbb{1}_{\{v(t_1,\dots,t_{d-1})<\gamma(\epsilon)\}}\mathbb{P}\left(\Psi(\eta^{1,\epsilon}_0, (\eta^{1,\epsilon}_j)_{d+1\le j\le p}, (\eta^{2,\epsilon}_i)_{1\le i\le n-p})) < \gamma(\epsilon),\ A\right)\\
=\ &\mathbb{1}_{\{v(t_1,\dots,t_{d-1})<\gamma(\epsilon)\}}\mathbb{P}\left(\Psi(\eta^{1,\epsilon}_0, (\eta^{1,\epsilon}_j)_{d+1\le j\le p}, (\eta^{2,\epsilon}_i)_{1\le i\le n-p})) < \gamma(\epsilon)\right) \times \mathbb{P}(A)\\
\ge\ &\mathbb{1}_{\{v(t_1,\dots,t_{d-1})<\gamma(\epsilon)\}}\mathbb{P}\left(\Psi(\eta^{2,\epsilon}_1, \dots, \eta^{2,\epsilon}_{n-d+1}) < \gamma(\epsilon)\right) \times \mathbb{P}(A)\\
=\ &\mathbb{P}(\tau^{\epsilon,\eta}_{CUSUM} \ge n,\ A),
\end{aligned} \tag{59}$$

where we recall that $v(t_1,\dots,t_{d-1}) = \max\limits_{1\le k\le d-1}\left\{\max\limits_{1\le j\le k}\sum\limits_{i=j}^{k}\phi_\epsilon(t_i - t_{i-1})\right\} \vee 0$.

Using the inequality (59), we obtain by integrating in (48) that

$$\mathbb{P}(\tau^{\epsilon,\delta}_{CUSUM} \ge n,\ n \ge \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p) \ge \mathbb{P}(\tau^{\epsilon,\eta}_{CUSUM} \ge n,\ n \ge \mathcal{D}^\epsilon,\ \mathcal{N}^{1,\epsilon} = p).$$

Finally, using (46) and (47), we conclude that

$$\mathbb{P}(\tau^{\epsilon,\delta}_{CUSUM} \ge n,\ \mathcal{N}^{1,\epsilon} = p) \ge \mathbb{P}(\tau^{\epsilon,\eta}_{CUSUM} \ge n,\ \mathcal{N}^{1,\epsilon} = p) \tag{60}$$

for all $n \in \mathbb{N}$ and $p \in \mathbb{N}$.
If $\bar{Q}^2(\epsilon) > \bar{Q}^1(\epsilon)$, one verifies this time that $(\eta^{1,\epsilon}_0, (\eta^{1,\epsilon}_j)_{d+1\le j\le p}) \ge_{\mathrm{st}} (\eta^{2,\epsilon}_1, \dots, \eta^{2,\epsilon}_{p-d+1})$, so that (58) is replaced by

$$(\eta^{1,\epsilon}_0, (\eta^{1,\epsilon}_j)_{d+1\le j\le p}, (\eta^{2,\epsilon}_i)_{1\le i\le n-p})) \ge_{\mathrm{st}} (\eta^{2,\epsilon}_1, \dots, \eta^{2,\epsilon}_{n-d+1}).$$

Coupled with the fact that $\Psi : (z_1,\dots,z_{n-d+1}) \mapsto \Psi(z_1,\dots,z_{n-d+1})$ (defined in (52)) is this time a decreasing function in each of the variables z_i, $i = 1,\dots,n-d+1$, one deduces that the (59) as well as the conclusion (60) still hold. This ends the proof. □

As explained in the outline of the proof of theorems at the end of Section 4.1, we now find the adequate candidate for detecting the index $\mathcal{N}^{1,\epsilon}$, defined in (35), of the jump time where $\boldsymbol{X^{1,\epsilon}}$ exceeds m_0.

Since the inter-arrivals $(\eta^{1,\epsilon}_i)_{i\ge 0}$ and $(\eta^{2,\epsilon}_i)_{i\ge 0}$ are independent from $\mathcal{N}^{1,\epsilon}$, we first observe that

$$\mathbb{E}[(\tau^{\epsilon,\delta}_{CUSUM} - \mathcal{N}^{1,\epsilon})^+ | \mathcal{N}^{1,\epsilon} = K] = \mathbb{E}^{(K)}[(\tau^{\epsilon,\delta,\star}_{CUSUM} - K)^+] \tag{61}$$

where $\tau^{\epsilon,\delta,\star}_{CUSUM}$ refers to the stopping time of the CUSUM given by (3) associated to the sequence $(Z_i)_{i\in\mathbb{N}}$ in (1) such that the sequences $(Z^1_i)_{i\in\mathbb{N}}$ and $(Z^2_i)_{i\in\mathbb{N}}$ are respectively $\mathcal{E}(\bar{Q}^1(\epsilon))$ and $\mathcal{E}(\bar{Q}^2(\epsilon))$ distributed and for a threshold $\gamma(\epsilon) = \log h(\epsilon)$. Conditioning on Z_1,...,Z_K results on the following upper bound:

$$\begin{aligned}
\mathbb{E}^{(K)}[(\tau^{\epsilon,\delta,\star}_{CUSUM} - K)^+] &= \mathbb{E}^{(K)}(\mathbb{E}^{(K)}[(\tau^{\epsilon,\delta,\star}_{CUSUM} - K)^+ | Z_1,\dots,Z_K])\\
&\le \mathbb{E}^{(K)}\left(\sup_{p\ge 1}\mathrm{ess\,sup}\,\mathbb{E}^{(p)}[(\tau^{\epsilon,\delta,\star}_{CUSUM} - p)^+ | Z_1,\dots,Z_p]\right)\\
&= E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}},
\end{aligned} \tag{62}$$

where we recall that $E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}}$ is the worst mean delay defined in (7). Equation (9) states that

$$E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}} \sim \frac{\log h(\epsilon)}{I^\epsilon} \text{ as } h(\epsilon) \to \infty \tag{63}$$

where $h(\epsilon) := \exp(\gamma(\epsilon))$ and I^ϵ refers to the KL distance defined in (11) and given by (24). At this point, (63) does not provide much information. Indeed, $E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}}$ behaves like $\frac{\log h(\epsilon)}{I^\epsilon}$ as $h(\epsilon)$ becomes large. However, this fact is not really useful as we wish to rather know how this quantity behaves when $\epsilon \to 0$. An upper bound is hence provided in the following lemma:

Lemma 1. *(a) For all $\epsilon > 0$, the following inequality holds*

$$E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}} \leq \frac{\log h(\epsilon) + \max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} \tag{64}$$

where I^ϵ is given by (24).

(b) Under the assumption $(\mathbf{A_4})$, *it holds that*

$$\liminf_{\epsilon\to 0} I^\epsilon > 0 \quad \textit{and} \quad \limsup_{\epsilon \to 0} \frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} < \infty.$$

Proof. We first prove (a). Let τ be the stopping variable of a one-sided sequential probability ratio tests (SPRT) of $F_1 = \mathcal{E}(\bar{Q}^1(\epsilon))$ (associated to a probability measure $\mathbb{P}_1$ and expectation $\mathbb{E}_1$) *vs* $F_2 = \mathcal{E}(\bar{Q}^2(\epsilon))$ (associated to a probability measure $\mathbb{P}_2$ and expectation $\mathbb{E}_2$) with likelihood boundary ratio $h(\epsilon)$, given by

$$\tau = \inf\{n \geq 1 | S_n \geq \log h(\epsilon)\} \tag{65}$$

where S_n refers to the random walk defined by

$$S_n = \sum_{i=1}^n \xi_i = \sum_{i=1}^n \left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} + (-\bar{Q}^2(\epsilon) + \bar{Q}^1(\epsilon))Y_i\right) = \sum_{i=1}^n \phi_\epsilon(Y_i)$$

and $(Y_i)_{i\in\mathbb{N}}$ is i.i.d and follows an exponential distribution with parameter $\bar{Q}^j(\epsilon)$ under $\mathbb{P}_j$, $j = 1, 2$.

We have $\mathbb{P}_2(\tau < \infty) \leq h(\epsilon)$ (see [23]), so that [16, (11) Theorem 2] reads here

$$E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}} \leq \mathbb{E}_2(\tau). \tag{66}$$

We know from [23] that $\mathbb{E}_2(\tau)$ is equivalent to $\log(h(\epsilon))/I^\epsilon$ when $h(\epsilon)$ is large, however this information is not satisfactory at this point because we want an estimate for $\mathbb{E}_2(\tau)$ when ϵ tends to 0, such that $\lim_{\epsilon\to 0} h(\epsilon) = +\infty$. Note however that this asymptotic implies that $\mathbb{E}_2(\tau)$ is finite (see [24]) so that, since τ is a stopping time adapted to the sequence $(Y_i)_{i\in\mathbb{N}}$, Wald's equation (see [25, Theorem 3.3.2 p. 105]) reads

$$\mathbb{E}_2(S_\tau) = \mathbb{E}_2(\xi_1).\mathbb{E}_2(\tau) = I^\epsilon \mathbb{E}_2(\tau), \tag{67}$$

so that (66) implies the upper bound

$$E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}} \leq \frac{\mathbb{E}_2(S_\tau)}{I^\epsilon}. \tag{68}$$

We now observe that,

- if $\bar{Q}^2(\epsilon) \geq \bar{Q}^1(\epsilon)$, then $\xi_i \leq \log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}$ and hence $\mathbb{E}_1(S_\tau) \leq \log h(\epsilon) + \log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}$,
- if $\bar{Q}^2(\epsilon) \leq \bar{Q}^1(\epsilon)$, then we have, by the memoryless property of the exponential distribution, that $\mathbb{E}_2(S_\tau) = \log h(\epsilon) + \mathbb{E}_2((-\bar{Q}^2(\epsilon) + \bar{Q}^1(\epsilon))Y_i) = \log h(\epsilon) + \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1$.

In both cases, we have

$$\mathbb{E}_2(S_\tau) \leq \log h(\epsilon) + \max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right), \tag{69}$$

so that (68) implies (64).

We now come to prove (b). Let us denote

$$\varphi(x) = -\log x + x - 1, \forall x > 0, \tag{70}$$

so that $I^\epsilon = \varphi(\bar{Q}^1(\epsilon)/\bar{Q}^2(\epsilon))$ by (24). Note that φ is a convex function and it admits a unique minimum for $x = 1$ with $\varphi(1) = 0$. Assumption $(\mathbf{A_4})$ means that $\frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}$ belongs to $H := (0, 1-c)\bigcup(1+c, +\infty)$ for some small $c \in (0,1)$, and for $\epsilon \leq \epsilon_0$ small enough. This implies $I_\epsilon \geq \inf_{x\in H} \varphi(x) := d > 0$ for $\epsilon \leq \epsilon_0$, proving that $\liminf_{\epsilon\to 0} I^\epsilon > 0$.

Then, we show that $\limsup_{\epsilon\to 0} \frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} < \infty$. For that purpose, for $\epsilon \leq \epsilon_0$, we consider the following cases:

- If $\frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} < 1 - c$, then $\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} < 0$ and $\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1 > 0$. Consequently, we obtain that

$$\frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} = \frac{\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1}{\varphi\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right)} \leq \sup_{x\in(0,1-c)} \frac{1/x - 1}{\varphi(1/x)} < +\infty.$$

- If $\frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} > 1 + c$, then $\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} > 0$ and $\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1 < 0$. Hence, this yields to

$$\frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} = \frac{\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}}{\varphi\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right)} \leq \sup_{x>1+c} \frac{\log x}{\varphi(1/x)} < +\infty.$$

This ends the proof of the Lemma. □

4.3. Proof of Theorem 1

Theorem 1 aims to study the properties of M_ϵ defined in (20). The latter can be rewritten as

$$M_\epsilon = \sum_{i=1}^{\tau_{CUSUM}^{\epsilon,\eta}} \Delta_i^{\epsilon,\eta} \text{ where } \Delta_i^{\epsilon,\eta} = \begin{cases} \Delta_i^{1,\epsilon}, i \leq \mathcal{D}^\epsilon - 1, \\ \Delta_{i-\mathcal{D}^\epsilon+1}^{2,\epsilon}, i \geq \mathcal{D}^\epsilon \end{cases} \tag{71}$$

where we recall that $\mathcal{D}^\epsilon$ is defined in (36).

Proof. Using the basic inequality $(a+b)^+ \leq a^+ + b^+$, we obtain the following upper bound:

$$\begin{aligned}
\mathbb{E}([M_\epsilon - m_0]^+) &= \mathbb{E}\left(\left[M_\epsilon - \sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{\epsilon,\eta} + \sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{\epsilon,\eta} - \sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{1,\epsilon} + \sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{1,\epsilon} - m_0\right]^+\right) \\
&\leq \mathbb{E}\left(\left[M_\epsilon - \sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{\epsilon,\eta}\right]^+\right) + \mathbb{E}\left(\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} [\Delta_i^{\epsilon,\eta} - \Delta_i^{1,\epsilon}]^+\right) + \mathbb{E}\left(\left[\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{1,\epsilon} - m_0\right]^+\right) \\
&= A_\epsilon + B_\epsilon + C_\epsilon. \qquad (72)
\end{aligned}$$

In the above decomposition, $\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{1,\epsilon} = X^\epsilon_{\tau^\epsilon_{m_0}}$ is the level of the process $\boldsymbol{X}^{\boldsymbol{\epsilon}}$ at the instant where its behavior changes. Thus C_ϵ refers to the overshoot of the latter process while it crosses m_0. As for the quantities A_ϵ and B_ϵ that involve $\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{\epsilon,\eta}$, it is somehow difficult to provide an intuitive interpretation. The idea is to be able to exploit the fact that $\tau^{\epsilon,\eta}_{CUSUM}$ enables to detect the index $\mathcal{N}^{1,\epsilon}$ where the distribution of the jumps X^ϵ changes (as mentioned in the outline of the theorems proof in Section 4.1). Firstly, we look at A_ϵ:

$$\begin{aligned}
A_\epsilon &= \mathbb{E}\left(\left[\sum_{i=1}^{\tau^{\epsilon,\eta}_{CUSUM}} \Delta_i^{\epsilon,\eta} - \sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{\epsilon,\eta}\right]^+\right) \\
&= \mathbb{E}\left(\left[\sum_{i=\mathcal{N}^{1,\epsilon}+1}^{\tau^{\epsilon,\eta}_{CUSUM}} \Delta^{2,\epsilon}_{i-\mathcal{D}^\epsilon+1}\right] \mathbb{1}_{\{\tau^{\epsilon,\eta}_{CUSUM} \geq \mathcal{N}^{1,\epsilon}+1\}}\right) \\
&= \sum_{d=1}^\infty \sum_{p=d}^\infty \mathbb{E}\left(\left[\sum_{i=p+1}^{\tau^{\epsilon,\eta}_{CUSUM}} \Delta^{2,\epsilon}_{i-d+1}\right] \mathbb{1}_{\{\tau^{\epsilon,\eta}_{CUSUM} \geq p+1\}} \mathbb{1}_{\{\mathcal{D}^\epsilon = d, \mathcal{N}^{1,\epsilon}=p\}}\right). \qquad (73)
\end{aligned}$$

Since $\mathcal{N}^{1,\epsilon}$, $\mathcal{D}^\epsilon$ and τ_{m_0} depend on the process $\boldsymbol{X}^1$ and since $\tau^{\epsilon,\eta}_{CUSUM}$ depends on $(\eta_i^{1,\epsilon})_{i\in\mathbb{N}}$, τ_{m_0} and $(\eta_i^{2,\epsilon})_{i\in\mathbb{N}}$, then $\Delta^{2,\epsilon}_{i-d+1}$ is independent form $\mathcal{N}^{1,\epsilon}$, $\mathcal{D}^\epsilon$ and $\tau^{\epsilon,\eta}_{CUSUM}$. Consequently, we can write (73) as follows

$$\begin{aligned}
A_\epsilon &= \mathbb{E}\left[\Delta^{2,\epsilon}\right] \sum_{d=1}^\infty \sum_{p=d}^\infty \mathbb{E}\left([\tau^{\epsilon,\eta}_{CUSUM} - p]^+ \mathbb{1}_{\{\mathcal{D}^\epsilon = d, \mathcal{N}^{1,\epsilon}=p\}}\right) \\
&= \mathbb{E}\left[\Delta^{2,\epsilon}\right] \mathbb{E}\left(\left[\tau^{\epsilon,\eta}_{CUSUM} - \mathcal{N}^{1,\epsilon}\right]^+\right) \\
&= \mathbb{E}\left[\Delta^{2,\epsilon}\right] \sum_{p=1}^\infty \mathbb{E}\left([\tau^{\epsilon,\eta}_{CUSUM} - p]^+ \mid \mathcal{N}^{1,\epsilon} = p\right) \times \mathbb{P}(\mathcal{N}^{1,\epsilon} = p). \qquad (74)
\end{aligned}$$

By using Proposition 1, A_ϵ in (74) can be upper bounded by

$$A_\epsilon \leq \mathbb{E}\left[\Delta^{2,\epsilon}\right] \sum_{p=1}^\infty \mathbb{E}\left(\left[\tau^{\epsilon,\delta}_{CUSUM} - p\right]^+ \Big| \, \mathcal{N}^{1,\epsilon} = p\right) \times \mathbb{P}(\mathcal{N}^{1,\epsilon} = p). \qquad (75)$$

Additionally, the combination of (61) and (62) yields that

$$\mathbb{E}\left(\left[\tau^{\epsilon,\delta}_{CUSUM} - p\right]^+ \Big| \, \mathcal{N}^{1,\epsilon} = p\right) \leq E^2_{\tau^{\epsilon,\delta,\star}_{CUSUM}}$$

from which, thanks to Lemma 1 (a) and (75), one can easily obtain the following upper bound for A_ϵ:

$$\begin{aligned} A_\epsilon &\leq \mathbb{E}\left[\Delta^{2,\epsilon}\right] \frac{\log h(\epsilon) + \max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} \sum_{p=1}^{\infty} \mathbb{P}(\mathcal{N}^{1,\epsilon} = p) \\ &\leq \mathbb{E}\left[\Delta^{2,\epsilon}\right] \frac{\log h(\epsilon) + \max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon}. \end{aligned} \tag{76}$$

Secondly, we deal with the term B_ϵ. We observe that, from (71), $\Delta_i^{\epsilon,\eta} = \Delta_i^{1,\epsilon}$ for all $i = 1, \ldots, \mathcal{D}^\epsilon - 1$, so that

$$\begin{aligned} B_\epsilon &= \mathbb{E}\left(\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} [\Delta_i^{\epsilon,\eta} - \Delta_i^{1,\epsilon}]^+\right) = \mathbb{E}\left(\sum_{i=\mathcal{D}^\epsilon}^{\mathcal{N}^{1,\epsilon}} [\Delta_{i-\mathcal{D}^\epsilon+1}^{2,\epsilon} - \Delta_i^{1,\epsilon}]^+\right) \\ &\leq \mathbb{E}\left(\sum_{i=\mathcal{D}^\epsilon}^{\mathcal{N}^{1,\epsilon}} \Delta_{i-\mathcal{D}^\epsilon+1}^{2,\epsilon}\right) \leq \mathbb{E}\left(\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{2,\epsilon}\right) \\ &= \mathbb{E}\left(\mathcal{N}^{1,\epsilon}\right) \mathbb{E}\left(\Delta^{2,\epsilon}\right), \end{aligned} \tag{77}$$

the last equality stemming from the independence between $\mathcal{N}^{1,\epsilon}$ and the $\Delta_i^{2,\epsilon}$, $i \in \mathbb{N}$. By Lorden's inequality (see e.g. [26, Proposition 6.2 p.160]), we have

$$\mathbb{E}\left(\mathcal{N}^{1,\epsilon} - 1\right) \leq \frac{m_0}{\mathbb{E}\left(\Delta^{1,\epsilon}\right)} + \frac{\mathbb{E}\left[\left(\Delta^{1,\epsilon}\right)^2\right]}{\left(\mathbb{E}\left[\Delta^{1,\epsilon}\right]\right)^2}, \tag{78}$$

which, plugged into (77), yields

$$B_\epsilon \leq m_0 \frac{\mathbb{E}\left[\Delta^{2,\epsilon}\right]}{\mathbb{E}\left[\Delta^{1,\epsilon}\right]} + \frac{\mathbb{E}\left[\left(\Delta^{1,\epsilon}\right)^2\right]}{\left(\mathbb{E}\left[\Delta^{1,\epsilon}\right]\right)^2} \mathbb{E}\left[\Delta^{2,\epsilon}\right] + \mathbb{E}\left[\Delta^{2,\epsilon}\right]. \tag{79}$$

Finally, we end up by looking to the term C_ϵ. Since $\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{1,\epsilon}$ is larger than m_0, it simplifies as

$$C_\epsilon = \mathbb{E}\left(\sum_{i=1}^{\mathcal{N}^{1,\epsilon}} \Delta_i^{1,\epsilon} - m_0\right).$$

Wald's equality (see [25, Theorem 3.3.2 p. 105]) and (78) provide

$$\begin{aligned} C_\epsilon &= \mathbb{E}\left[\mathcal{N}^{1,\epsilon}\right] \times \mathbb{E}[\Delta_i^{1,\epsilon}] - m_0 \\ &\leq \frac{\mathbb{E}\left[\left(\Delta^{1,\epsilon}\right)^2\right]}{\mathbb{E}\left[\Delta^{1,\epsilon}\right]} + \mathbb{E}(\Delta^{1,\epsilon}). \end{aligned} \tag{80}$$

Gathering (76), (79) and (80), we obtain from (72) that $\mathbb{E}([M_\epsilon - m_0]^+) \leq c_1^\epsilon$ and then ($\boldsymbol{P_1^\epsilon}$)

is satisfied, where

$$c_1^\epsilon := \sum_{i=1}^{6} T_i(\epsilon),$$

$$T_1(\epsilon) := \mathbb{E}\left[\Delta^{2,\epsilon}\right] \frac{\log h(\epsilon) + \max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon}, \quad T_2(\epsilon) := m_0 \frac{\mathbb{E}\left[\Delta^{2,\epsilon}\right]}{\mathbb{E}\left[\Delta^{1,\epsilon}\right]},$$

$$T_3(\epsilon) := \frac{\mathbb{E}\left[\left(\Delta^{1,\epsilon}\right)^2\right]}{\left(\mathbb{E}\left[\Delta^{1,\epsilon}\right]\right)^2}\mathbb{E}\left[\Delta^{2,\epsilon}\right], \quad T_4(\epsilon) := \mathbb{E}\left[\Delta^{2,\epsilon}\right] \quad T_5(\epsilon) := \frac{\mathbb{E}\left[\left(\Delta^{1,\epsilon}\right)^2\right]}{\mathbb{E}\left[\Delta^{1,\epsilon}\right]}, \quad T_6(\epsilon) := \mathbb{E}\left[\Delta^{1,\epsilon}\right].$$

Let us now prove ($\boldsymbol{P_2^\epsilon}$). For that purpose , we set

$$c_2^\epsilon \quad := \quad h(\epsilon)\mathbb{E}\left[\Delta^{1,\epsilon}\right] \tag{81}$$

where we recall that $h(\epsilon)$ is defined by (23). When $m_0 = +\infty$ then the process $\mathbf{X}$ is equal to $\mathbf{X^1}$, $\tau^{\epsilon,\delta}_{CUSUM} = \tau^{\epsilon,\eta}_{CUSUM}$, and the expression of M_ϵ in (71) is simplified to

$$M_\epsilon = \sum_{i=1}^{\tau^{\epsilon,\delta}_{CUSUM}} \Delta_i^{1,\epsilon}.$$

The independence of $\tau^{\epsilon,\delta}_{CUSUM}$ from the sequence $(\Delta_i^{1,\epsilon})_{i\in\mathbb{N}}$ yields

$$\mathbb{E}_\infty(M_\epsilon) = \mathbb{E}_\infty\left(\sum_{i=1}^{\tau^{\epsilon,\delta}_{CUSUM}} \Delta_i^{1,\epsilon}\right) = \mathbb{E}_\infty(\tau^{\epsilon,\delta}_{CUSUM})\mathbb{E}(\Delta^{1,\epsilon}) \geq h(\epsilon)\mathbb{E}(\Delta^{1,\epsilon}) = c_2^\epsilon \tag{82}$$

by (8).
Finally, we prove the last property ($\boldsymbol{P_3}$) or ($\boldsymbol{P_3'}$). First, the expression (30) of the distribution of $\Delta^{j,\epsilon}$, $j = 1, 2$, and in view of assumption ($\mathbf{A_2}$), yields the following estimates for their first and second order moments:

$$\begin{aligned} \mathbb{E}[\Delta^{j,\epsilon}] &= \frac{1}{\bar{Q}^j(\epsilon)}\int_{(\epsilon,\infty)} xQ^j(dx) \sim_{\epsilon\to 0} \frac{1}{\bar{Q}^j(\epsilon)}\int_{(0,\infty)} xQ^j(dx), \\ \mathbb{E}[(\Delta^{j,\epsilon})^2] &= \frac{1}{\bar{Q}^j(\epsilon)}\int_{(\epsilon,\infty)} x^2Q^j(dx) \sim_{\epsilon\to 0} \frac{1}{\bar{Q}^j(\epsilon)}\int_{(0,\infty)} x^2Q^j(dx). \end{aligned} \tag{83}$$

We then consider the different cases given in the statement of the Theorem. In each case, we consider the behaviour as $\epsilon \to 0$ of each term $T_i(\epsilon)$, $i = 1, \dots 6$ in the definition of c_1^ϵ. We note that for $i = 4, 5, 6$, it holds that $\limsup_{\epsilon\to 0} T_i(\epsilon)$ is finite, so that we will mainly focus on the terms $T_1(\epsilon)$, $T_2(\epsilon)$ and $T_3(\epsilon)$ in the following.

Case 1: $\bar{Q}^2(0) = +\infty$ *and* $\limsup_{\epsilon\to 0} \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} < \infty$.

In this case, we have $h(\epsilon) = [\bar{Q}^2(\epsilon)I^\epsilon]^2$ and thus $\mathbb{E}[\Delta^{2,\epsilon}]\frac{\log h(\epsilon)}{I^\epsilon} = O(\frac{\log(\bar{Q}^2(\epsilon)I^\epsilon)}{\bar{Q}^2(\epsilon)I^\epsilon}) \longrightarrow 0$ as $\epsilon \to 0$. Hence, thanks to the bound (b) in Lemma 1, we obtain that

$$\limsup_{\epsilon\to 0} T_1(\epsilon) \leq \limsup_{\epsilon\to 0} \mathbb{E}[\Delta^{2,\epsilon}]\frac{\log h(\epsilon)}{I^\epsilon} + \limsup_{\epsilon\to 0} \mathbb{E}[\Delta^{2,\epsilon}]\frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} = 0,$$

as indeed $\bar{Q}^2(0) = +\infty$ and so $\lim_{\epsilon\to 0} \mathbb{E}[\Delta^{2,\epsilon}] = 0$. We check easily from (83) that $T_2(\epsilon)$ and $T_3(\epsilon)$ are $O\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right)$, hence $\limsup_{\epsilon\to 0} T_2(\epsilon)$ and $\limsup_{\epsilon\to 0} T_3(\epsilon)$ are finite under the present assumption. Thus, we have $\limsup_{\epsilon\to 0} c_1^\epsilon < \infty$. As for c_2^ϵ, we write from (81) that

$$c_2^\epsilon = [\bar{Q}^2(\epsilon)I^\epsilon]^2 \frac{1}{\bar{Q}^1(\epsilon)} \int_{(\epsilon,\infty)} xQ^1(dx) = \bar{Q}^2(\epsilon)[I^\epsilon]^2 \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} \int_{(\epsilon,\infty)} xQ^1(dx),$$

so that

$$\begin{aligned} \liminf_{\epsilon\to 0} c_2^\epsilon &\geq \liminf_{\epsilon\to 0} \bar{Q}^2(\epsilon) \ \liminf_{\epsilon\to 0} [I^\epsilon]^2 \ \liminf_{\epsilon\to 0} \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} \ \liminf_{\epsilon\to 0} \int_{(\epsilon,\infty)} xQ^1(dx) \\ &= \liminf_{\epsilon\to 0} \bar{Q}^2(\epsilon) \ \liminf_{\epsilon\to 0} [I^\epsilon]^2 \ \liminf_{\epsilon\to 0} \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} \ \int_{(0,\infty)} xQ^1(dx) = +\infty \end{aligned}$$

as we used the property that $\limsup_{\epsilon\to 0} \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} < \infty$ implies that $\liminf_{\epsilon\to 0} \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)} > 0$ as well as Lemma 1 (b). Finally, we verified that $(\mathbf{P_3'})$ holds.

Case 2: $\bar{Q}^2(0) = +\infty$ *and* $\limsup_{\epsilon\to 0} \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} = \infty$.

In this case, $h(\epsilon)$ has the same form as in Case 1. We note that this case necessarily implies that $\bar{Q}^1(0) = +\infty$. As in the previous case, the property $\limsup_{\epsilon\to 0} T_1(\epsilon) < \infty$ can be proved similarly. As to $T_2(\epsilon)$ and $T_3(\epsilon)$, both of those terms are $O\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right)$, hence we have

$$c_1^\epsilon = O\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right). \tag{84}$$

Contrary to the Case 1, $\limsup_{\epsilon\to 0} T_2(\epsilon)$ and $\limsup_{\epsilon\to 0} T_3(\epsilon)$ are infinite in the present case. To prove that $\lim_{\epsilon\to 0} c_2^\epsilon = +\infty$, we recall that $I^\epsilon = \varphi(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)})$ where φ is defined in (70). We note that $\lim_{x\to\infty} \varphi(x)/x = 1$, which implies in particular that $\varphi(x) \geq x/2$ for $x \geq K$ large enough. Recalling that $I^\epsilon \geq d > 0$ for $\epsilon \leq \epsilon_0$ small enough (see proof of Lemma 1), we then have, for $\epsilon \leq \epsilon_0$,

$$c_2^\epsilon = \begin{cases} \bar{Q}^2(\epsilon) \frac{\varphi\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right)}{\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}} I^\epsilon \int_{(\epsilon,\infty)} xQ^1(dx) \geq \bar{Q}^2(\epsilon)\frac{1}{2}d \int_{(\epsilon_0,\infty)} xQ^1(dx) & \text{if } \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} \geq K, \\ \bar{Q}^2(\epsilon)\frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}[I^\epsilon]^2 \int_{(\epsilon,\infty)} xQ^1(dx) \geq \bar{Q}^2(\epsilon)\frac{1}{K}d^2 \int_{(\epsilon_0,\infty)} xQ^1(dx) & \text{if } \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} < K, \end{cases}$$

so that, all in all, we have $c_2^\epsilon \geq \bar{Q}^2(\epsilon) \min(\frac{1}{2}d, \frac{1}{K}d^2) \int_{(\epsilon_0,\infty)} xQ^1(dx) \longrightarrow +\infty$ as $\epsilon \to 0$. Furthermore, the definition (81) of c_ϵ^2 combined with (84) implies $\frac{c_1^\epsilon}{c_2^\epsilon} = O\left(\frac{1}{[\bar{Q}^2(\epsilon)I^\epsilon]^2} \frac{[\bar{Q}^1(\epsilon)]^2}{\bar{Q}^2(\epsilon)}\right)$. Now, remembering that $\varphi(x) \geq x/2$ for $x \geq K$, we have, for ϵ small enough

$$\frac{1}{[\bar{Q}^2(\epsilon)I^\epsilon]^2} \frac{[\bar{Q}^1(\epsilon)]^2}{\bar{Q}^2(\epsilon)} \leq \frac{1}{\bar{Q}^2(\epsilon)d^2} K^2 \mathbb{1}_{[\bar{Q}^1(\epsilon)/\bar{Q}^2(\epsilon)\leq K]} + \frac{4}{\bar{Q}^2(\epsilon)} \mathbb{1}_{[\bar{Q}^1(\epsilon)/\bar{Q}^2(\epsilon)> K]},$$

so that $\limsup_{\epsilon\to 0} \frac{1}{[\bar{Q}^2(\epsilon)I^\epsilon]^2} \frac{[\bar{Q}^1(\epsilon)]^2}{\bar{Q}^2(\epsilon)} \leq \limsup_{\epsilon\to 0} \frac{1}{\bar{Q}^2(\epsilon)d^2} K^2 + \limsup_{\epsilon\to 0} \frac{4}{\bar{Q}^2(\epsilon)} = 0$. This entails that $c_1^\epsilon = o(c_2^\epsilon)$. Hence $(\mathbf{P_3})$ holds.

Case 3: $\bar{Q}^1(0) = +\infty$ and $\bar{Q}^2(0) < \infty$.

In this case, $h(\epsilon) = [\bar{Q}^1(\epsilon)]^{\beta}$. Let us first note that, since $\varphi(x) \sim x$ as $x \to \infty$, $I^\epsilon = \varphi\left(\frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)}\right) \sim \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(0)}$ as $\epsilon \to 0$. Hence, $\log h(\epsilon)/I^\epsilon$ is equivalent to $\beta \bar{Q}^2(0) \log \bar{Q}^1(\epsilon)/\bar{Q}^1(\epsilon) \longrightarrow 0$ as $\epsilon \to 0$. Furthermore, we have from Lemma 1 that $\frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon} < \infty$ as $\epsilon \to 0$. Both facts, along with the boundedness of $\mathbb{E}[\Delta^{2,\epsilon}]$, implies that $\limsup_{\epsilon \to 0} T_1(\epsilon) < +\infty$. Again, we check easily from (83) that $T_2(\epsilon)$ and $T_3(\epsilon)$ are $O(\bar{Q}^1(\epsilon))$ as $\epsilon \to 0$. All in all, we have that c_1^ϵ is a $O(\bar{Q}^1(\epsilon))$. Hence

$$\frac{c_1^\epsilon}{c_2^\epsilon} = O\left(\frac{\bar{Q}^1(\epsilon)}{[\bar{Q}^1(\epsilon)]^\beta / \bar{Q}^1(\epsilon)}\right) = O\left([\bar{Q}^1(\epsilon)]^{2-\beta}\right) \longrightarrow 0, \quad \epsilon \to 0.$$

Furthermore, it can be easily verified that $\lim_{\epsilon \to 0} c_2^\epsilon = +\infty$ and consequently $(\mathbf{P_3})$ holds. $\square$

4.4. Proof of Theorem 2

We denote throughout this section the processes $\mathbf{X}^{\mathbf{j},\epsilon-} = (X_t^{j,\epsilon-})_{t\geq 0}$, $j = 1, 2$ as well as $\mathbf{X}^{\epsilon-} = (X_t^{\epsilon-})_{t\geq 0}$ obtained from $\mathbf{X}^{\mathbf{j}}$, $j = 1, 2$ and $\mathbf{X}$ by discarding the jumps of height larger than ϵ. Using the Poisson random measure N_j introduced in Section 3, and with a similar expression as those in (14) and (29) for X_t^j and $X_t^{j,\epsilon}$, this reads

$$X_t^{j,\epsilon-} := \int_{[0,t]} \int_{(0,\epsilon]} x N_j(ds \times dx) = \sum_{s \leq t} \Delta X_s^j \mathbb{1}_{[\Delta X_s^j \leq \epsilon]}, \quad j = 1, 2, \tag{85}$$

$$X_t^{\epsilon-} := \sum_{s \leq t} \Delta X_s \mathbb{1}_{[\Delta X_s \leq \epsilon]} = X_t - X_t^\epsilon, \tag{86}$$

so that one has, similarly to (12):

$$X_t^{\epsilon-} = X_t^{1,\epsilon-} \mathbb{1}_{[t \leq \tau_{m_0}]} + (X_{\tau_{m_0}}^{1,\epsilon-} + X_{t-\tau_{m_0}}^{2,\epsilon-}) \mathbb{1}_{[t > \tau_{m_0}]}. \tag{87}$$

Thus, one can point out that the difference between the detection level L_ϵ and the pseudo-level M_ϵ expressed in (22) and (33) respectively, is given by

$$L_\epsilon - M_\epsilon = X_{d^\epsilon}^{\epsilon-}$$

where we recall that d^ϵ is the detection time defined by (21). Indeed, L_ϵ is the detection level, i.e. the level of the process $\mathbf{X}$ at time d^ϵ, and M_ϵ is the sum of all jumps larger than ϵ of the process X_t between $t = 0$ and $t = d^\epsilon$. So that, by distinguishing the two cases ($d^\epsilon \leq \tau_{m_0}$ and $d^\epsilon > \tau_{m_0}$), this difference may also be written thanks to (87) as

$$0 \leq L_\epsilon - M_\epsilon = X_{d^\epsilon \wedge \tau_{m_0}}^{1,\epsilon-} + X_{d^\epsilon - d^\epsilon \wedge \tau_{m_0}}^{2,\epsilon-}. \tag{88}$$

We next show that the expectation of each term on the right-hand side of (88) tends to 0 when $\epsilon \to 0$. Before showing the latter, we start by mentioning an important result related to martingales associated to the processes $\mathbf{X^1}$ and $\mathbf{X^2}$. Note first that $\int_{(0,\epsilon)} x Q^j(dx)$, $j = 1, 2$ is finite for ϵ small enough (see Assumption $(\mathbf{A_1})$). Hence the following process

$$\left(\int_{[0,t]} \int_{(0,\epsilon]} x N_j(ds \times dx) - t \int_{(0,\epsilon)} x Q^j(dx)\right)_{t\geq 0} = \left(X_t^{j,\epsilon-} - t \int_{(0,\epsilon)} x Q^j(dx)\right)_{t\geq 0} \tag{89}$$

is a martingale adapted to $\mathbf{X}^j$, $j = 1, 2$. This fact may be verified by applying [19, Corollary 4.6 p.97] with function ϕ used in that reference given by $\phi(s, x) := x \mathbb{1}_{(0,\epsilon)}(x)$.

Term $X^{1,\epsilon-}_{d^\epsilon\wedge\tau_{m_0}}$

We first observe that, $(X^{1,\epsilon-}_t)_{t\geq 0}$ in (86) is increasing so that

$$X^{1,\epsilon-}_{d^\epsilon\wedge\tau_{m_0}} \leq X^{1,\epsilon-}_{\tau_{m_0}}.$$

Furthermore, for $m_0 < +\infty$, τ_{m_0} in (13) is a stopping time adapted to $\mathbf{X}^1$ of finite expectation, so that according to Doob's optional stopping time theorem (see [27, p.4]) applied to the martingale (89) we have for all $N \in \mathbb{N}$ that

$$\mathbb{E}\left(X^{1,\epsilon-}_{\tau_{m_0}\wedge N}\right) = \mathbb{E}(\tau_{m_0}\wedge N) \times \int_{(0,\epsilon)} xQ^1(dx),$$

which, by letting $N \to +\infty$ and using the monotone convergence theorem, yields

$$\mathbb{E}\left(X^{1,\epsilon-}_{\tau_{m_0}}\right) = \mathbb{E}(\tau_{m_0}). \int_{(0,\epsilon)} xQ^1(dx) \tag{90}$$

which tends to 0 as $\epsilon \to 0$. Hence, we deduce that

$$X^{1,\epsilon-}_{d^\epsilon\wedge\tau_{m_0}} \stackrel{L^1}{\longrightarrow} 0 \text{ as } \epsilon \to 0. \tag{91}$$

Term $X^{2,\epsilon-}_{d^\epsilon - d^\epsilon\wedge\tau_{m_0}}$

Let us introduce the process $(V_i)_{i\in\mathbb{N}}$ as a time shifted version of the CUSUM statistic $(G^\eta_i)_{i\in\mathbb{N}}$ as $V_0 = \mathcal{G}^\epsilon$ defined in (43) and $V_n = G^\eta_{\mathcal{D}^\epsilon+n-1}$ for $n \geq 1$, so that $(V_i)_{i\in\mathbb{N}}$ satisfies from (44), (17) as well as the relation $\eta^\epsilon_{\mathcal{D}^\epsilon+n} = \eta^{2,\epsilon}_{n+1}$ (see Figure 3 for the illustration of this latter fact) the recursive equation

$$\left\{\begin{array}{rcl} V_{n+1} & = & \left(V_n + \phi_\epsilon(\eta^{2,\epsilon}_{n+1})\right)^+, \ n \geq 0 \\ V_0 & = & \mathcal{G}^\epsilon. \end{array}\right. \tag{92}$$

Let us associate to this CUSUM statistic the corresponding first passage time above the threshold $\gamma(\epsilon) = \log h(\epsilon)$ given by

$$\tau^V_{CUSUM} := \left\{\begin{array}{cc} \inf\{n \geq 0 |\ V_n \geq \log h(\epsilon)\} & \text{if } \tau^{\epsilon,\eta}_{CUSUM} \geq \mathcal{D}^\epsilon, \\ 0 & \text{otherwise.} \end{array}\right.$$

We observe in particular, thanks to the relation between the process $(V_i)_{i\in\mathbb{N}}$ and $(G^\eta_i)_{i\in\mathbb{N}}$, that their corresponding CUSUM statistics verify $\tau^V_{CUSUM} = \tau^{\epsilon,\eta}_{CUSUM} - \mathcal{D}^\epsilon$ when $\tau^{\epsilon,\eta}_{CUSUM} \geq \mathcal{D}^\epsilon$. We also define the random walk $(S^{(2)}_n)_{n\geq 0}$ by $S^{(2)}_n = \sum_{k=1}^n \phi_\epsilon(\eta^{2,\epsilon}_k)$, where ϕ_ϵ is defined in (18), with first passage time above the threshold $\log h(\epsilon)$ given by

$$\tau^{S^{(2)}} := \inf\{n \geq 0 |\ S^{(2)}_n \geq \log h(\epsilon)\}. \tag{93}$$

It can be easily verified by induction that $V_n \geq S^{(2)}_n$ for all n, hence the corresponding first passage time verifies $\tau^V_{CUSUM} \leq \tau^{S^{(2)}}$. Recalling the properties $\eta^\epsilon_{\mathcal{D}^\epsilon+n-1} = \eta^{2,\epsilon}_n$ for $n \geq 1$ and $\tau^V_{CUSUM} = \tau^{\epsilon,\eta}_{CUSUM} - \mathcal{D}^\epsilon$ when $\tau^{\epsilon,\eta}_{CUSUM} \geq \mathcal{D}^\epsilon$, this results, thanks to the definition of d^ϵ in (21), in

$$d^\epsilon - d^\epsilon \wedge \tau_{m_0} = \sum_{k=1}^{\tau^V_{CUSUM}} \eta^{2,\epsilon}_k \leq \sum_{k=1}^{\tau^{S^{(2)}}} \eta^{2,\epsilon}_k := d^{\epsilon,2}. \tag{94}$$

which in turns results in

$$X^{2,\epsilon-}_{d^\epsilon - d^\epsilon \wedge \tau_{m_0}} \leq X^{2,\epsilon-}_{d^{\epsilon,2}}. \tag{95}$$

Since $d^{\epsilon,2}$ is a stopping time adapted to $\mathbf{X}^2$, a similar martingale argument to the one leading to (90) leads to the following

$$\mathbb{E}\left(X^{2,\epsilon-}_{d^{\epsilon,2}}\right) = \mathbb{E}(d^{\epsilon,2}) \times \int_{(0,\epsilon)} xQ^2(dx). \tag{96}$$

We then proceed to study the behaviour of $\mathbb{E}(d^{\epsilon,2})$ as $\epsilon \to 0$. Following (94) and by using Wald's equation (see [25, Theorem 3.3.2 p. 105])), we obtain

$$\mathbb{E}(d^{\epsilon,2}) = \mathbb{E}\left(\tau^{S^{(2)}}\right) \times \mathbb{E}(\eta^{2,\epsilon}). \tag{97}$$

One can easily see the following equality

$$\mathbb{E}\left(\tau^{S^{(2)}}\right) = \mathbb{E}_2(\tau)$$

with the right-sided term which was defined in the proof of Lemma 1, with τ defined by (65). By using (67) and (69), we consequently obtain

$$\mathbb{E}\left(\tau^{S^{(2)}}\right) \leq \frac{\log h(\epsilon) + \max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{I^\epsilon}.$$

Let us suppose that $\bar{Q}^2(0) = \infty$ (see (23)), so that the threshold in (93) is given by $\log h(\epsilon)$ where $h(\epsilon) = [\bar{Q}^2(\epsilon) I^\epsilon]^2$. Multiplying both sides of the above inequality by $\mathbb{E}(\eta^{2,\epsilon}) = O\left(1/\bar{Q}^2(\epsilon)\right)$, we thus get from (97) and (96), that

$$\mathbb{E}\left(X^{2,\epsilon-}_{d^{\epsilon,2}}\right) = O\left(\frac{\log(\bar{Q}^2(\epsilon)I^\epsilon)}{\bar{Q}^2(\epsilon)I^\epsilon}\int_{(0,\epsilon)} xQ^2(dx) + \frac{\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)}{\bar{Q}^2(\epsilon)I^\epsilon}\int_{(0,\epsilon)} xQ^2(dx)\right). \tag{98}$$

Assumption ($\mathbf{A_2}$) ensures that $\int_{(0,\epsilon)} xQ^2(dx)$ tends to 0 as ϵ tends to 0, and Lemma 1 (b) ensures that $\max\left(\log \frac{\bar{Q}^2(\epsilon)}{\bar{Q}^1(\epsilon)}, \frac{\bar{Q}^1(\epsilon)}{\bar{Q}^2(\epsilon)} - 1\right)/I^\epsilon$ is bounded when ϵ tends to 0. Consequently, the term (98) tends to 0 when ϵ tends to 0.

Finally, this implies thanks to (95) that

$$X^{2,\epsilon-}_{d^\epsilon - d^\epsilon \wedge \tau_{m_0}} \xrightarrow{L_1} 0, \quad \epsilon \to 0. \tag{99}$$

One can easily show the same results when $\bar{Q}^2(0) < \infty$ and $h(\epsilon) = [\bar{Q}^1(\epsilon)]^\beta$ (see (23)).

End of proof

Combining (88), (91) and (99), we then deduce (26). By observing that

$$\mathbb{E}([L_\epsilon - m_0]^+) \leq \mathbb{E}(L_\epsilon - M_\epsilon) + \mathbb{E}([M_\epsilon - m_0]^+),$$

and by recalling the form of c_1^ϵ in the proof of Theorem 1, the property ($\boldsymbol{P_1^\epsilon}$) holds by replacing the quantity c_1^ϵ by $\mathbb{E}(L_\epsilon - M_\epsilon) + c_1^\epsilon$. As for the property ($\boldsymbol{P_2^\epsilon}$), it holds because the following holds

$$\mathbb{E}_\infty(L_\epsilon) \geq \mathbb{E}_\infty(M_\epsilon) \geq c_2^\epsilon,$$

where the last inequality stems from Theorem 1. Finally, one can verify that ($\boldsymbol{P_3}$) (or ($\boldsymbol{P_3'}$) when (25) is satisfied) holds because of (26).

4.5. Proof of Theorem 3

The idea of the proof is to take $\epsilon = 0$ in the construction of the approximated process defined in (31), so that both processes $\mathbf{X}$ and $\mathbf{X}^\epsilon$ now coincide. The situation is hence less complicated than in Section 4.1 as here τ_{m_0} is equal to $T_{\mathcal{D}^0}$ for some (random) index $\mathcal{D}^\epsilon = \mathcal{D}^0$ defined in (36) with $\epsilon = 0$. Likewise, when $\epsilon = 0$, the index $\mathcal{N}^{1,\epsilon}$ coincides with $\mathcal{D}^0$, the crossing time $\tau^\epsilon_{m_0}$ is now equal to τ_{m_0} and the quantity $\eta_0^{1,\epsilon}$ defined in (45) with $\epsilon = 0$ is equal to 0. We may then drop the dependence in $\epsilon = 0$ now and denote by $(\Delta^1_i)_{i\in\mathbb{N}}$ and $(\Delta^2_i)_{i\in\mathbb{N}}$ the respective jumps of the compound Poisson processes $\mathbf{X}^1$ and $\mathbf{X}^2$ and M as the pseudo level. We note that, now that $\epsilon = 0$ and $\mathbf{X} = \mathbf{X}^0$, the pseudo-level M is now the same as L defined in (28), which is nicely illustrated in Figure 1. This latter quantity may also be expressed as

$$L = \sum_{i=1}^{\tau^\eta_{CUSUM}} \Delta^\eta_i \text{ where } \Delta^\eta_i = \begin{cases} \Delta^1_i, i \leq \mathcal{D}^0, \\ \Delta^2_{i-\mathcal{D}^0+1}, i > \mathcal{D}^0. \end{cases} \tag{100}$$

As in (72), we upper bound the expected delay as

$$\begin{aligned}\mathbb{E}([L-m_0]^+) = \mathbb{E}([M-m_0]^+) &\leq \mathbb{E}\left(\left[M - \sum_{i=1}^{\mathcal{D}^0}\Delta^\eta_i\right]^+\right) + \mathbb{E}\left(\left[\sum_{i=1}^{\mathcal{D}^0}\Delta^1_i - m_0\right]^+\right) \\ &= A_0(h) + C_0 := c^h_1 \end{aligned}\tag{101}$$

where $h > 1$ is the threshold for the CUSUM rule defined in (27). Similarly to (73) and (74), one obtains

$$A_0(h) = \mathbb{E}(\Delta^2)\,\mathbb{E}\left([\tau^\eta_{CUSUM} - \mathcal{D}^0]^+\right). \tag{102}$$

Moreover, Lemma 1 is still valid when $\epsilon = 0$ by substituting $h(\epsilon)$ by h, so that one can easily derive that, similarly to (75),

$$A_0(h) \leq \mathbb{E}(\Delta^2)\frac{\log h + \max\left(\log\frac{\bar{Q}^2(0)}{\bar{Q}^1(0)}, \frac{\bar{Q}^1(0)}{\bar{Q}^2(0)} - 1\right)}{I^0}. \tag{103}$$

with $I^0 = \log\frac{\bar{Q}^2(0)}{\bar{Q}^1(0)} - 1 + \frac{\bar{Q}^1(0)}{\bar{Q}^2(0)}$. As for C_0, a similar analysis as in (80) yields that

$$\begin{aligned} C_0 &= \mathbb{E}\left[\mathcal{D}^0\right].\mathbb{E}[\Delta^1_i] - m_0 \\ &\leq \frac{\mathbb{E}\left[\left(\Delta^1\right)^2\right]}{\mathbb{E}\left[\Delta^1\right]} + \mathbb{E}(\Delta^1). \end{aligned}\tag{104}$$

Thus, $(\boldsymbol{P^0_1})$ is satisfied. Now, we turn to prove $(\boldsymbol{P^0_2})$. For that purpose, one gets similarly to (82) that

$$\mathbb{E}_\infty(L) = \mathbb{E}_\infty(\tau^\eta_{CUSUM})\,\mathbb{E}(\Delta^1) \geq h\,\mathbb{E}(\Delta^1) := c^h_2$$

with $\mathbb{E}(\Delta^1)$ a finite quantity thanks to Assumption $(\mathbf{A_1})$ as well as $\bar{Q}^1(0) < +\infty$. Finally, since $c^h_1 = O(\log h)$ as $h \to \infty$ because of (103) and (104), one easily verifies that $(\boldsymbol{P_4})$ holds. This ends the proof of Theorem 3.

5. Numerical illustrations on case study: standard gamma processes with level switching

As previously explained, the process $\boldsymbol{X}$ defined as in equation (12) changes from $\boldsymbol{X^1}$ (called in the sequel Regime 1) to $\boldsymbol{X^2}$ (called in the sequel Regime 2) when it crosses the level m_0. In this section, we consider more specifically gamma processes with corresponding shape functions $A_i : t \mapsto \gamma_i t$, $i = 1, 2$ and scale parameter $b = 1$. That means

$$\boldsymbol{X^i} = (X_t^i)_{t\geq 0} \text{ with } X_t^i \sim \Gamma(\gamma_1 t, 1), \quad i = 1, 2.$$

The objective of this experiment is to illustrate the ϵ-detection rule (19) proposed in this paper and to compare its performances with those of the classical CUSUM rule, applied on the increments $(X_{t_{i+1}} - X_{t_i})_{i\in\mathbb{N}}$ of the process $\boldsymbol{X}$, for some temporal discretization $(t_i)_{i\in\mathbb{N}}$ of $[0, +\infty)$ with constant size $t_{i+1} - t_i = s > 0$. Indeed, conditionally to the change time, one may think that the increments of the process are independent and gamma distributed random variables $\Gamma(\gamma_1 s, 1)$ or $\Gamma(\gamma_2 s, 1)$ before and after the regime change which may lead us to believe that a classic CUSUM rule is well adapted. However, not only this latter fact is not true, but in our case, the change time is not deterministic, but random; Worse, it is a stopping time that depends on the trajectory $\boldsymbol{X}$. Moreover the crossing of level m_0 can happen between two "inspections" t_i and the distribution of the resulting increment is therefore unknown (neither $\Gamma(\gamma_1 s, 1)$ or $\Gamma(\gamma_2 s, 1)$). Consequently, the particular situation considered in this paper does not correspond to the conditions for applying the classic CUSUM rule.

Figure 4 illustrates a sample path of the process $\mathbf{X}$ with threshold $m_0 = 10$ as well as the trajectories for the ϵ-detection rule and classical CUSUM statistics.

The process $\mathbf{X}$ (top of Figure 4) is supposed to be observed continuously. To apply the classic CUSUM rule (bottom of Figure 4), and as explained earlier on, a time-discretization step-size s is considered, so that $t_i = is$, $i \in \mathbb{N}$. For each increment $X_{is} - X_{(i-1)s}$, $i = 1, 2, \ldots$, a likelihood ratio is computed between Regime 1 and Regime 2 by noting that, intuitively, $X_{is} - X_{(i-1)s}$ is distributed according to a gamma $\Gamma(\gamma_1 s, 1)$ as long as $is < \tau_{m_0}$, and to a gamma $\Gamma(\gamma_2 s, 1)$ as soon as $(i-1)s > \tau_{m_0}$, with τ_{m_0} the crossing time of level m_0 defined in (13). The likelihood ratios can then be used sequentially to compute the test statistic (4) of the CUSUM rule. The implementation of the ϵ-detection rule (middle of Figure 4) is different. By construction, a new increment of the rule is computed as soon as the process, which is a pure-jump process, observed a jump greater than a fixed value ϵ. The number of increments of the ϵ-detection rule is therefore random and depends also on the functioning mode: in the gamma case considered in this section (see Section 3.3), the expected delay before a new increment is equal to:

$$\mathbb{E}[\eta_i^\epsilon] = \frac{1}{\bar{Q}^i(\epsilon)} = \frac{1}{\gamma_i \int_\epsilon^{+\infty} \frac{1}{y} e^{-y}\, dy}, \tag{105}$$

for $i = 1$ or 2 depending on whether the process is under Regime 1 or Regime 2.

For each of the two change detection rules, it is necessary to determine a threshold $\gamma(\epsilon)$ (for the ϵ-detection rule) or $\gamma(s)$ (for the classic CUSUM rule). In a classic way in statistical process control, the threshold is chosen in practice such that the false alarm "rate" can be controlled. More precisely, the threshold $\gamma(\epsilon)$ of the ϵ-detection rule is chosen such that the Average Run Level under Regime 1 $\mathbb{E}_\infty(L_\epsilon)$ (and denoted $ARLev_\infty$, see Remark 2) is equal to a specific value chosen here equal to 30. It is obtained empirically by a dichotomous

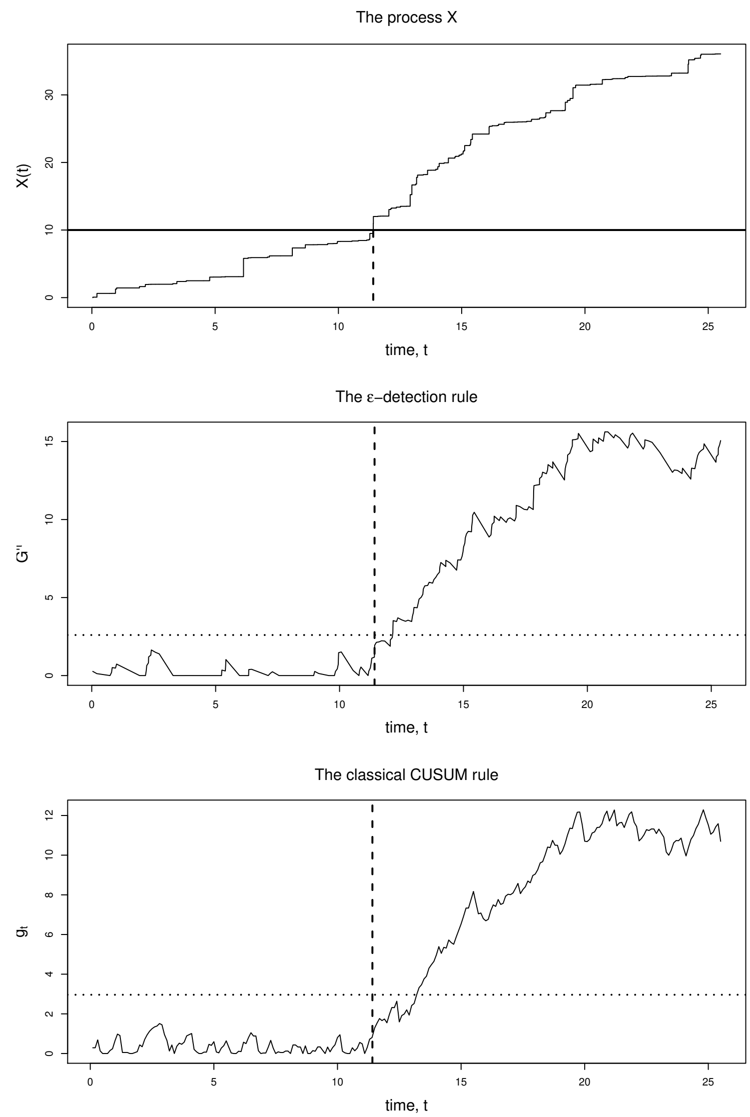

Figure 4: Top: Process X with a change from a gamma process $\Gamma(t, 1)$ to a gamma process $\Gamma(1.5t, 1)$ at level $m_0 = 10$; Middle: The corresponding ϵ-detection rule statistic (plain line) with $\epsilon = 7.2 \times 10^{-4}$ for a theoretical Average Run Level ($\mathbb{E}_\infty(L_\epsilon)$) equal to 30, leading to a threshold equal to $\gamma(\epsilon) = 2.599$ (dotted line) ; Bottom: The classical CUSUM statistic (plain line) with $s = 0.1$ for a theoretical Average Run Level equal to 30, leading to a threshold equal to $\gamma(s) = 2.963$ (dotted line).

approach on $\gamma(\epsilon)$ based on the test statistics of 10000 trajectories of $\mathbf{X}$ simulated under Regime 1 (by adapting Algorithm 1 of [28] to our context). The same approach is followed for $\gamma(s)$.

As expected, the test statistic in Figure 4 is close to zero when the system is under Regime 1 and it increases as soon as the system is under Regime 2. Indeed, the logarithm of the likelihood ratio (18) tends to be negative before the change and positive after the change. The regime change is therefore quickly detected.

To go further, the performances of the ϵ-detection and classic CUSUM rules are compared. For the same $ARLev_\infty$, fixed to the value 30, and leading to the determination of the threshold $\gamma(\epsilon)$ and $\gamma(s)$, we compare the mean overshoot of the process $\mathbf{X}$ at detection time above the threshold m_0, defined by $\mathbb{E}_{m_0}([L_\epsilon - m_0]^+)$ for the ϵ-detection rule (see again Remark 2). In the sequel, we use $m_0 = 0$ for the mean overshoot, that means that the system is under Regime 2 at the beginning of the monitoring. It allows to obtain what is called, in the control chart community, the zero-state ARL (as opposed to the steady-state ARL with a non-null m_0). The corresponding quantity will be denoted $ARLev_0$. Obviously, for the same $ARLev_\infty$, the best rule is the one with the smallest $ARLev_0$. These ARLev will be obtained by Monte Carlo simulations.

For the classic CUSUM, and since we only need the increment values for this approach, the gamma processes are simulated by the "increment" approach (see [1]), which is an exact and efficient simulation technique. For the ϵ-detection rule, and since it is necessary to observe the jumps of the process, we used a series representation of the gamma process (see [29, Section 6], or [30, Proposition 2]) and simulate the processes with the rejection approach (with $B = 30$ for the truncation of the series, see [30, Algorithm 2]).

To make the comparison between the classic CUSUM and the ϵ-detection rule relevant, the number of increments of a run must be approximately the same. Remind that the number of increments of the ϵ-detection rule is random (of which expectation is given by (105)). For a value s, we then determine the value of ϵ which gives, in mean, the same number of increments. For one value for s, we obtain two values for ϵ, ϵ_1 and ϵ_2 depending on whether the process is under Regime 1 or Regime 2. The ϵ-detection rule was then applied with two configurations: either we used the greater value for ϵ (that means we consider the least advantageous case for the ϵ-detection rule), or we used the mean between the two epsilon values.

The Monte Carlo approximations of the $ARLev$ are based on 10000 repetitions. We use for the simulations of the gamma processes, $\gamma_1 = 1$ and several values for γ_2 (1.1, 1.2 and 1.5). The time-discretization step-size s for the classic CUSUM is varying from 2 to 0.05. For each value, ϵ_1 and ϵ_2 are computed, and then ϵ is chosen. For example, if $s = 0.5$, that means that we consider two increments per unit of time for the classic CUSUM and $\gamma_2 = 1.1$, then we have to use $\epsilon_1 = 8.2 \times 10^{-2}$ to obtain, for the ϵ-detection rule, two increments per unit of time in mean when the system is under Regime 1 and $\epsilon_2 = 10^{-1}$ when the system is under Regime 2. We then test two values for ϵ : $\epsilon = 10^{-1}$ which corresponds to the "worst-case scenario" and $\epsilon = 9.1 \times 10^{-2}$, the "medium scenario". The results are presented, for the classic CUSUM, in Table 1 and for the ϵ-detection rule, in Table 2 for the "worst-case scenario" and Table 3 for the "medium scenario". In each table are reported the threshold ($\gamma(s)$ for the classic CUSUM and $\gamma(\epsilon)$ for the ϵ-detection rule) of the rules leading to an empirical $ARlev_\infty$ close to 30 when the system is under Regime 1, the corresponding observed $ARLev_\infty$ and the empirical standard deviation of the Run Level $SDRLev_\infty$. For the comparison of the rules when the system is under Regime 2, we find in Table 1 to 3, the

Table 1: Empirical Average Run Levels for the classic CUSUM for a change from a gamma process $\Gamma(t, 1)$ to a gamma process $\Gamma(\gamma_2 t, 1)$ obtain from 10000 repetitions.

s	γ_2	$\gamma(s)$	$ARLev_\infty$	$SDRLev_\infty$	$ARLev_0$	$SDRLev_0$	$RLevMax_0$
2.00	1.1	0.371	29.99	20.26	23.31	13.25	142.97
2.00	1.2	0.638	29.96	21.13	19.21	10.22	106.56
2.00	1.5	1.110	29.89	23.13	13.54	6.01	70.29
0.50	1.1	0.601	30.05	21.58	21.15	12.76	110.56
0.50	1.2	1.022	29.86	22.28	16.55	9.15	102.78
0.50	1.5	1.773	30.28	24.10	10.81	4.64	45.82
0.10	1.1	1.200	30.07	24.68	14.47	9.56	82.62
0.10	1.2	1.897	29.98	25.59	9.35	5.74	51.35
0.10	1.5	2.963	29.92	27.18	5.03	2.78	26.86
0.05	1.1	1.592	30.00	26.24	11.14	7.57	82.79
0.05	1.2	2.424	29.81	26.74	6.58	4.32	38.87
0.05	1.5	3.592	30.02	28.05	3.19	2.05	18.34

Table 2: Empirical Average Run Levels for the ϵ-detection rule for a change from a gamma process $\Gamma(t, 1)$ to a gamma process $\Gamma(\gamma_2 t, 1)$ obtain from 10000 repetitions. The value for ϵ is computed from the less favourable functioning mode ($\epsilon = \max(\epsilon_1, \epsilon_2)$, "worst-case scenario").

s	ϵ	γ_2	$\gamma(\epsilon)$	$ARLev_\infty$	$SDRLev_\infty$	$ARLev_0$	$SDRLev_0$	$RLevMax_0$
2.00	6.00×10^{-1}	1.1	0.247	34.90	29.00	28.98	23.11	249.41
2.00	6.45×10^{-1}	1.2	0.384	31.39	27.34	22.73	18.30	178.32
2.00	7.60×10^{-1}	1.5	0.642	30.88	28.24	17.52	14.23	168.00
0.50	1.00×10^{-1}	1.1	0.508	29.12	23.24	20.11	15.00	146.02
0.50	1.19×10^{-1}	1.2	0.855	30.05	25.49	16.00	11.40	97.33
0.50	1.75×10^{-1}	1.5	1.389	30.30	26.55	10.40	6.86	55.42
0.10	6.50×10^{-5}	1.1	1.121	29.33	25.44	13.26	9.88	88.42
0.10	1.40×10^{-4}	1.2	1.773	30.04	27.20	8.83	6.43	69.79
0.10	7.20×10^{-4}	1.5	2.599	29.62	27.49	4.46	3.19	35.09
0.05	7.20×10^{-9}	1.1	1.501	29.14	26.28	10.26	7.73	69.82
0.05	3.20×10^{-8}	1.2	2.262	29.81	28.11	6.11	4.61	43.71
0.05	9.10×10^{-7}	1.5	3.208	29.15	28.08	2.82	2.23	26.22

empirical $ARLev_0$, but also the empirical standard deviation $SDRLev_0$ and the maximal level of degradation accumulated before the detection over the 10000 repetitions, denoted $RLevMax_0$.

Table 3: Empirical Average Run Levels for the ϵ-detection rule for a change from a gamma process $\Gamma(t, 1)$ to a gamma process $\Gamma(\gamma_2 t, 1)$ obtain from 10000 repetitions. The value for ϵ is the mean between ϵ_1 and ϵ_2 ($\epsilon = (\epsilon_1 + \epsilon_2)/2$, "medium scenario").

s	ϵ	γ_2	$\gamma(\epsilon)$	$ARLev_0$	$SDRLev_0$	$ARLev_1$	$SDRLev_1$	$RLevMax_1$
2.00	5.775×10^{-1}	1.1	0.233	30.48	25.19	25.69	20.44	193.71
2.00	6.000×10^{-1}	1.2	0.409	31.30	27.13	22.40	17.74	183.73
2.00	6.575×10^{-1}	1.5	0.725	31.18	27.99	16.47	13.06	142.20
0.50	9.100×10^{-2}	1.1	0.510	27.78	22.55	19.22	14.27	133.55
0.50	1.005×10^{-1}	1.2	0.876	29.02	24.61	15.18	10.81	151.06
0.50	1.285×10^{-1}	1.5	1.505	30.18	26.73	9.56	6.29	69.39
0.10	4.550×10^{-5}	1.1	1.137	29.20	25.42	13.06	9.96	112.37
0.10	8.300×10^{-5}	1.2	1.780	29.16	26.52	8.32	6.07	55.89
0.10	3.730×10^{-4}	1.5	2.685	30.48	28.44	4.19	3.03	35.09
0.05	4.200×10^{-9}	1.1	1.566	30.95	27.75	10.64	7.99	69.82
0.05	1.660×10^{-8}	1.2	2.283	29.29	27.40	5.93	4.42	40.78
0.05	4.556×10^{-7}	1.5	3.263	30.30	29.41	2.74	2.16	26.22

Obviously, for each rule, we can see that more the number of increments per unit of time is high (a small s for the classic CUSUM and a small ϵ for the ϵ-detection rule), more efficient is the detection rule. It is an illustration of the results obtained in the previous sections, which show that the ϵ-detection rule is efficient as ϵ tends to 0.

Another interesting point is that, for a comparable number of increments per unit of time, the ϵ-detection rule gives better results (a smaller $ARLev_0$) than the the classic CUSUM, even in the "worst-case scenario", as soon as ϵ is sufficiently small ($\epsilon \leq 0.1$). The difference is even more pronounced when the classic CUSUM is compared to the medium scenario.

References

[1] J. M. van Noortwijk, A survey of the application of gamma processes in maintenance, Reliability Engineering & System Safety 94 (1) (2009) 2–21.

[2] M. H. Davis, A note on the Poisson disorder problem., Banach center publications 1 (1976) 65–72.

[3] G. Peskir, A. N. Shiryaev, Solving the Poisson disorder problem, in: Advances in finance and stochastics, Springer, 2002, pp. 295–312.

[4] E. Bayraktar, H. V. Poor, Quickest detection of a minimum of two Poisson disorder times, SIAM Journal on Control and Optimization 46 (1) (2007) 308–331.

[5] M. Krawiec, Z. Palmowski, Ł. Płociniczak, Quickest drift change detection in Lévy-type force of mortality model, Applied Mathematics and Computation 338 (2018) 432–450.

[6] M. Fouladirad, A. Grall, Condition-based maintenance for a system subject to a non-homogeneous wear process with a wear rate transition, Reliability Engineering & System Safety 96 (6) (2011) 611–618.

[7] M. Fouladirad, A. Grall, On-line change detection and condition-based maintenance for systems with unknown deterioration parameters, IMA Journal of Management Mathematics 25 (2) (2014) 139–158.

[8] M. Fouladirad, A. Grall, L. Dieulle, On the use of on-line detection for maintenance of gradually deteriorating systems, Reliability Engineering & System Safety 93 (12) (2008) 1814–1820.

[9] J. E. Figueroa-López, S. Ólafsson, Change-point detection for Lévy processes, Annals of Applied Probability 29 (2) (2019) 717–738.

[10] A. N. Shiryaev, On optimum methods in quickest detection problems, Theory of Probability & Its Applications 8 (1) (1963) 22–46.

[11] A. G. Tartakovsky, Asymptotic optimality in bayesian changepoint detection problems under global false alarm probability constraint, Theory of Probability & Its Applications 53 (3) (2009) 443–466.

[12] A. G. Tartakovsky, G. V. Moustakides, State-of-the-art in bayesian changepoint detection, Sequential Analysis 29 (2) (2010) 125–145.

[13] A. G. Tartakovsky, V. V. Veeravalli, General asymptotic bayesian theory of quickest change detection, Theory of Probability & Its Applications 49 (3) (2005) 458–497.

[14] E. S. Page, Continuous inspection schemes, Biometrika 41 (1/2) (1954) 100–115.

[15] M. Basseville, I. V. Nikiforov, Detection of abrupt changes: theory and application, Vol. 104, prentice Hall Englewood Cliffs, 1993.

[16] G. Lorden, Procedures for reacting to a change in distribution, The Annals of Mathematical Statistics 42 (6) (1971) 1897–1908.

[17] G. V. Moustakides, Optimal stopping times for detecting changes in distributions, The Annals of Statistics 14 (4) (1986) 1379–1387.

[18] T. L. Lai, Information bounds and quick detection of parameter changes in stochastic systems, IEEE Transactions on Information Theory 44 (7) (1998) 2917–2929.

[19] A. E. Kyprianou, Introductory lectures on fluctuations of Lévy processes with applications, Springer Science & Business Media, 2006.

[20] E. Cinlar, Introduction to stochastic processes, Courier Corporation, 2013.

[21] R. A. Doney, Fluctuation theory for Lévy processes, Springer, Berlin, Heidelberg, 2007.

[22] M. Shaked, J. G. Shanthikumar, Stochastic orders, Springer Science & Business Media, 2007.

[23] A. Wald, Sequential Analysis, Wiley, New York, 1947.

[24] C. Stein, A note on cumulative sums, The Annals of Mathematical Statistics 17 (4) (1946) 498–499.

[25] S. M. Ross, J. J. Kelly, R. J. Sullivan, W. J. Perry, D. Mercer, R. M. Davis, T. D. Washburn, E. V. Sager, J. B. Boyce, V. L. Bristow, Stochastic processes, Vol. 2, Wiley New York, 1996.

[26] S. Asmussen, Applied probability and queues, Vol. 51, Springer Science & Business Media, 2008.

[27] J. Bertoin, Lévy processes, Vol. 121, Cambridge university press Cambridge, 1996.

[28] P. Qiu, Distribution-free multivariate process control based on log-linear modeling, IIE Transactions 40 (7) (2008) 664–677.

[29] J. Rosiński, Series representations of Lévy processes from the perspective of point processes, in: Lévy processes, Springer, 2001, pp. 401–415.

[30] Z. Al Masry, S. Mercier, G. Verdier, Approximate simulation techniques and distribution of an extended gamma process, Methodology and Computing in Applied Probability 19 (1) (2017) 213–235.