The Annals of Applied Probability
2019, Vol. 29, No. 2, 717–738
https://doi.org/10.1214/17-AAP1368

CHANGE-POINT DETECTION FOR LÉVY PROCESSES

BY JOSÉ E. FIGUEROA-LÓPEZ AND SVEINN ÓLAFSSON

Washington University in St. Louis and University of California, Santa Barbara

Since the work of Page in the 1950s, the problem of detecting an abrupt change in the distribution of stochastic processes has received a great deal of attention. In particular, a deep connection has been established between Lorden's minimax approach to change-point detection and the widely used CUSUM procedure, first for discrete-time processes, and subsequently for some of their continuous-time counterparts. However, results for processes with jumps are still scarce, while the practical importance of such processes has escalated since the turn of the century. In this work, we consider the problem of detecting a change in the distribution of continuous-time processes with independent and stationary increments, that is, Lévy processes, and our main result shows that CUSUM is indeed optimal in Lorden's sense. This is the most natural continuous-time analogue of the seminal work of Moustakides [*Ann. Statist.* **14** (1986) 1379–1387] for sequentially observed random variables that are assumed to be i.i.d. before and after the change-point. From a practical perspective, the approach we adopt is appealing as it consists in approximating the continuous-time problem by a suitable sequence of change-point problems with equispaced sampling points, and for which a CUSUM procedure is shown to be optimal.

1. Introduction. Quickest detection is the problem of detecting, with as little delay as possible, a change in the probability distribution of a sequence of random measurements, and it has a wide range of applications in various branches of science and engineering, such as signal processing, supply chain management, cybersecurity and finance (see [20] and references therein). The main result of this paper is an extension of a well-known discrete-time quickest detection result of Moustakides [13], to an important class of continuous-time stochastic processes with jumps: Lévy processes.

In the discrete-time setting, the change-point problem involves a sequence $(X_n)_{n\geq 1}$ of random observations whose statistical properties change at some unknown point in time τ. In the simplest case, the pre-change observations $X_1, X_2, \dots, X_{\tau-1}$ are assumed to be independently drawn from one distribution, while the post-change observations $X_\tau, X_{\tau+1}, \dots$ are independently drawn from a different distribution. The objective is then to detect the change-point τ as soon as

Received November 2016; revised June 2017.
MSC2010 subject classifications. Primary 62L10, 60G51; secondary 60G40, 62C20.
Key words and phrases. Change-point, sequential detection, optimal stopping, CUSUM, Lévy processes.

possible, and the set of feasible detection strategies corresponds to the set of (extended real-valued) stopping times with respect to the observed sequence, with the understanding that a stopping time T decides that the change-point τ has occurred at time k when $T = k$. Naturally, the frequency of false alarms needs to be taken into account, so the design of detection procedures typically involves optimizing a trade-off between two types of performance indices, one quantifying the delay between the time a change occurs and the time it is detected, that is, the random variable $(T - \tau + 1)^+$, and the other being a measure of the frequency of false alarms, that is, events of the type $\{T < \tau\}$.

There are two main formulations of this optimization problem. The first of these is a Bayesian formulation in which the change-point is endowed with a prior distribution, usually a geometric distribution in discrete-time models or an exponential distribution in continuous-time models. This framework was first proposed with a linear delay penalty by Kolmogorov and Shiryayev [1], where the expected delay $\mathbb{E}(T - \tau + 1)^+$ was to be minimized subject to an upper bound on the probability of a false alarm, $\mathbb{P}(T < \tau)$. In applications there is typically limited information about the distribution of the change-point, and the second formulation is a more conservative minimax approach, where the change-point is assumed to be deterministic but unknown. In this setting, first proposed in the linear delay penalty case by Lorden [12], the delay penalty is a worst-worst case measure of expected delay, taken over all possible realizations of the observations leading up to the change-point, and over all possible values of the change-point [see equation (3.1) for details], and false alarms are constrained by a lower bound on the mean time between such events.

In this work, we are concerned with the latter formulation, which whenever it can be optimized, tends to give rise to the CUSUM (cumulative sum) stopping rule, first proposed by Page [15] as a continuous inspection scheme in the 1950s. CUSUM is one of the most widely used detection schemes in practice, and is based on the first time the accumulated likelihood (or log-likelihood) breaches a certain barrier [see equations (3.3)–(3.4)]. For a sequence of independent observations as described above, the asymptotic optimality of CUSUM, as the mean time between false alarms tends to infinity, was shown by Lorden [12] in 1971, and 15 years later, Moustakides [13] proved its optimality for any finite bound on the false alarm rate. Similar procedures were subsequently applied in [19] with Lorden's linear criterion replaced by exponentially penalized detection delays.

For continuous-time processes, the optimality of the CUSUM procedure for detecting a change in the drift of a Brownian motion was shown independently by several authors (see [4, 14] and [22]). More generally, its optimality for detecting a change in the drift of Itô processes was shown in [14], and more recently in [6], it was finally established for arbitrary processes with continuous paths. In both cases, the optimality was established under a convenient modification of Lorden's criterion, based on the Kullback–Leibler divergence that coincides with Lorden's criterion when the quadratic variation of the process is proportional to time.

For continuous-time processes with jumps, the current body of work is much more limited. In fact, to our knowledge the only available optimality result is for a proportional change in the intensity of doubly stochastic Poisson processes [7], with Lorden's expected delay criterion replaced by the expected number of jumps until detection, motivated by applications in actuarial science. This result includes the important case of a change in the jump intensity of a homogeneous Poisson process, for which the delay measure proposed in [7] coincides with Lorden's criterion. We also mention a recent nonparametric result for jump processes [5], based on the empirical tail integral of the jump-measure, and a separate stream of literature concerning change-point detection for Poisson processes in the Bayesian setting described above (see [3, 16, 18] and references therein).

The proofs of the aforementioned results do not appear to extend in an obvious way to more general jump processes. For instance, a fundamental step in the methodology of [14] for continuous processes, as well as in [7] for doubly stochastic Poisson processes, is to use stochastic calculus to characterize the CUSUM performance functions [i.e., the average run length as described in Remark 3.5(i) below] in terms of the solutions of certain differential equations, or delayed differential equations (DDE). In particular, the proof in [7] uses scale functions from the theory of Lévy processes to deal with the aforementioned DDEs, and resolves a long-standing discontinuity problem in the methodology of Moustakides (cf. [20], Section 6.4.4) using the concept of a discontinuous local time, both of which may prove difficult to extend to more general jump processes (see [7] for a further discussion, and [2] for another application of scale function in sequential testing).

In this work, we show that CUSUM is indeed optimal for detecting a change in the statistical properties of processes with independent and stationary increments, that is, Lévy processes. This result is in some sense the most natural continuous-time counterpart of the discrete-time problem considered by Moustakides in [13]. In addition to being of theoretical interest, it also has practical implications, as Lévy processes form a tractable and flexible family of stochastic models with jumps, that is well suited to model random phenomena that exhibit erratic and discontinuous behavior. Indeed, since the turn of the century, Lévy processes have found numerous applications in areas as diverse as finance and insurance, physics and biology.

Our approach to the problem has two main steps. First, we consider a continuous-time problem where the change-point is assumed to take values in a discrete set, and for which the methodology of Moustakides [13] can be adapted. We show that a discretized version of the CUSUM procedure is optimal in this case, which is of practical interest in its own right, for instance in financial markets where the change-point may be assumed to occur at the beginning of a new business day. The second step consists in increasing the sampling frequency, and using a limiting procedure to establish the optimality of CUSUM for the continuous-time detection problem with no restriction on the value of the change-point. This latter part of the proof is novel and relatively general; it relies on little more than standard pathwise properties of Lévy processes, and unlike the approach in [7], does not require separate analysis depending on whether there is a rise or a decline in the jump intensity, in addition to including changes in more general Lévy processes. The trade-off is that one does not obtain as a byproduct semiexplicit expressions for the CUSUM performance functions that are at the center of the methodology developed in [7, 14] and described above. On the other hand, we believe that our approach can be extended in various important ways, such as to incorporate exponential delay penalties (cf. [19]), and to derive optimal stopping times for more general point processes, such as Hawkes' processes. This is left for further research.

The remainder of this paper has two main sections. Section 2 introduces the probabilistic framework and the notation needed to study change-point detection for Lévy processes. Section 3 then reviews Lorden's change-point problem for discrete-time processes, as introduced in [12], before defining the analogous continuous-time problem and presenting our optimal change-detection results for Lévy processes. Proofs of ancillary results are deferred to the Appendix.

2. Probabilistic framework. Let $X^0 := (X_t^0)_{t\geq 0}$ and $X^1 := (X_t^1)_{t\geq 0}$ be Lévy processes on $\mathbb{R}$, defined on the same complete filtered probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$, with generating triplets given by $(\sigma^{(0)}, b^{(0)}, \nu^{(0)})$ and $(\sigma^{(1)}, b^{(1)}, \nu^{(1)})$, relative to the truncation function $\mathbf{1}_{\{|x|\leq 1\}}$ (see [21], Section 8). In other words, X^0 and X^1 have independent and stationary increments, and trajectories that are almost surely càdlàg (right-continuous with left limits). It is assumed that $(\sigma^{(0)}, b^{(0)}, \nu^{(0)}) \neq (\sigma^{(1)}, b^{(1)}, \nu^{(1)})$, and that we continuously observe the stochastic process $X^{(\tau)} := (X_t^{(\tau)})_{t\geq 0}$, defined by

$$X_t^{(\tau)} = \begin{cases} X_t^0, & t < \tau, \\ X_t^1 - X_\tau^1 + X_{\tau-}^0, & t \geq \tau, \end{cases}$$

where $\tau \in \bar{\mathbb{R}}_0^+ := [0, \infty) \cup \{\infty\}$, referred to as the change-point of the process, is assumed to be unknown and deterministic. It follows that $dX_t^{(\tau)} = dX_t^0 \mathbf{1}_{\{t<\tau\}} + dX_t^1 \mathbf{1}_{\{t\geq\tau\}}$, so the pre-change and post-change distributions of the process are determined by X^0 and X^1. We also set $X^{(\infty)} := X^0$ and $X^{(0)} := X^1$, which correspond, respectively, to the cases of a change-point at time zero and no change-point. Finally, observe that for any $\tau \in (0, \infty)$, $X^{(\tau)}$ is almost surely continuous at τ, and $(X_{t\wedge\tau}^{(\tau)})_{t\geq 0}$ and $(X_{t+\tau}^{(\tau)} - X_\tau^{(\tau)})_{t\geq 0}$ are independent (stopped) Lévy processes with the same generating triplets as X^0 and X^1, respectively.

Change-point detection revolves around detecting the change-point τ as quickly and as reliably as possible, using sequential detection schemes, that is to say, a set of admissible stopping times. In order to formalize a framework for this problem, let us introduce the space of càdlàg functions $\omega : [0, \infty) \to \mathbb{R}$, denoted by $\Omega = \mathbb{D}([0, \infty), \mathbb{R})$, along with the canonical process $X := (X_t)_{t\geq 0}$, defined by

$$X_t(\omega) := \omega(t), \qquad (\omega, t) \in \Omega \times [0, \infty), \tag{2.1}$$

and let $\mathcal{F}_t$ (resp., $\mathcal{F}$) be the smallest σ-field that makes $(X_s)_{s\le t}$ [resp., $(X_s)_{s\ge 0}$] measurable. As customary, let $\mathcal{F}_{t-} := \sigma(\bigcup_{s<t}\mathcal{F}_t)$, for $t>0$, and $\mathcal{F}_{0-}\equiv\mathcal{F}_0$, where $\mathcal{F}_0$ is the trivial σ-algebra. Next, for each $\tau\in\bar{\mathbb{R}}_0^+$, define the probability measure $\mathbb{P}_\tau$ on the space $(\Omega,\mathcal{F})$ as

$$\mathbb{P}_\tau(A) := \widetilde{\mathbb{P}}\big(\tilde\omega\in\widetilde\Omega : X_\cdot^{(\tau)}(\tilde\omega)\in A\big), \qquad A\in\mathcal{F}, \tag{2.2}$$

and denote by $\mathbb{E}_\tau$ the expected value w.r.t. to $\mathbb{P}_\tau$. Finally, by including in $\mathcal{F}_0$ the null sets of the measure $\mathbb{P}_\tau$ in $\mathcal{F}$, denoted by $\mathcal{N}_\tau$, we make $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge 0},\mathbb{P}_\tau)$ a complete filtered probability. Under assumptions (i)–(iii) below, $\mathcal{N}_\tau$ is the same set for each $\tau\in\bar{\mathbb{R}}_0^+$.

Note that for the canonical process X, Borel sets $B_1,\dots,B_n$, and time points $t_1,\dots,t_n$, we have

$$\mathbb{P}_\tau(X_{t_1}\in B_1,\dots,X_{t_n}\in B_n) = \widetilde{\mathbb{P}}\big(X_{t_1}^{(\tau)}\in B_1,\dots,X_{t_n}^{(\tau)}\in B_n\big),$$

so the distribution of X under $\mathbb{P}_\tau$ is the same as the distribution of $X^{(\tau)}$ under $\widetilde{\mathbb{P}}$. In particular, under $\mathbb{P}_\tau$ with $\tau\in(0,\infty)$, the processes $(X_{t\wedge\tau})_{t\ge 0}$ and $(X_{t+\tau}-X_\tau)_{t\ge 0}$ are independent (stopped) Lévy processes with generating triplets $(\sigma^{(0)},b^{(0)},\nu^{(0)})$ and $(\sigma^{(1)},b^{(1)},\nu^{(1)})$, respectively. The process X can therefore be referred to as the observed process, with the data-generating probability measure unknown.

It is also assumed that the probability measures $\mathbb{P}_\infty$ and $\mathbb{P}_0$ induced on the path space Ω by the Lévy processes $X^{(\infty)}$ and $X^{(0)}$, sometimes termed the *in-control* and *out-of-control* measures, are mutually absolutely continuous. Equivalently, it is assumed that their generating triplets satisfy the following conditions (see [21], Theorem 33.1):

(i) The Brownian volatilities are equal: $\sigma^{(0)}=\sigma^{(1)}$.
(ii) The Lévy measures $\nu^{(0)}$ and $\nu^{(1)}$ are equivalent and satisfy

$$\int_{\mathbb{R}_0}\big(e^{\varphi(x)/2}-1\big)^2\nu^{(0)}(dx)<\infty, \tag{2.3}$$

where $e^{\varphi(x)}=d\nu^{(1)}/d\nu^{(0)}$ is the Radon–Nikodým derivative of $\nu^{(1)}$ w.r.t. $\nu^{(0)}$.
(iii) The drift parameters $b^{(0)}$ and $b^{(1)}$ are such that

$$b^{(1)}-b^{(0)}-\int_{|x|\le 1}x\big(\nu^{(1)}-\nu^{(0)}\big)(dx)=\alpha\big(\sigma^{(0)}\big)^2, \tag{2.4}$$

for some $\alpha\in\mathbb{R}$, and $\alpha=0$ if $\sigma^{(0)}=0$.

Under these conditions, each member of the family of measures $\{\mathbb{P}_\tau,\tau\in\bar{\mathbb{R}}_0^+\}$ is absolutely continuous with respect to $\mathbb{P}_\infty$. It follows that for each $\tau\ge 0$ the likelihood ratio process

$$L_t^{(\tau)} := \frac{d\mathbb{P}_\tau|_{\mathcal{F}_t}}{d\mathbb{P}_\infty|_{\mathcal{F}_t}}, \qquad t\ge 0, \tag{2.5}$$

is well defined, with $L_t^{(\tau)} = 1$ for $t \le \tau$, while for $t \ge \tau$ it can be written in terms of the likelihood ratios $L_\tau^{(0)}$ and $L_t^{(0)}$ (see the Appendix for a justification):

$$L_t^{(\tau)} = \frac{d\mathbb{P}_0|_{\mathcal{F}_t}}{d\mathbb{P}_\infty|_{\mathcal{F}_t}} \bigg/ \frac{d\mathbb{P}_0|_{\mathcal{F}_\tau}}{d\mathbb{P}_\infty|_{\mathcal{F}_\tau}} = \frac{L_t^{(0)}}{L_\tau^{(0)}}, \qquad t \ge \tau. \tag{2.6}$$

Moreover, the likelihood ratio process

$$L_t^{(0)} = e^{U_t}, \qquad t \ge 0, \tag{2.7}$$

is a $\mathbb{P}_\infty$-martingale, and the log-likelihood ratio U_t takes the following form (see [21], Theorem 33.2):

$$\begin{aligned} U_t &= \alpha X_t^c - \frac{1}{2}\alpha^2 \big(\sigma^{(0)}\big)^2 t - \alpha b^{(0)} t \\ &\quad + \lim_{\varepsilon \downarrow 0} \bigg(\sum_{0 \le s \le t : |\Delta X_s| > \varepsilon} \varphi(\Delta X_s) - t \int_{|x|>\varepsilon} \big(e^{\varphi(x)} - 1\big) \nu^{(0)}(dx) \bigg), \end{aligned} \tag{2.8}$$

where $(X_t^c)_{t \ge 0}$ is the continuous part of X (i.e., a Brownian motion with drift), and φ and α are as in equations (2.3)–(2.4). We remark that $(U_t)_{t \ge 0}$ is a Lévy process under $\mathbb{P}_\infty$ and $\mathbb{P}_0$, with generating triplets given explicitly in terms of those of X under $\mathbb{P}_\infty$ and $\mathbb{P}_0$ (see [21], Section 33). In particular, the Lévy measures are given by $\nu^{(0)} \circ \varphi^{-1}$ and $\nu^{(1)} \circ \varphi^{-1}$, respectively. Furthermore, under the measures $\mathbb{P}_\tau$, with $\tau \in (0, \infty)$, the processes $(U_{t \wedge \tau})_{t \ge 0}$ and $(U_{t+\tau} - U_\tau)_{t \ge 0}$ are independent (stopped) Lévy processes, with the same generating triplets as $(U_t)_{t \ge 0}$ under $\mathbb{P}_\infty$ and $\mathbb{P}_0$, respectively.

As mentioned above, a natural class of detection strategies corresponds to the set of stopping times with respect to the filtration generated by the observed process. Hence, for each $\gamma > 0$ we define

$$\mathcal{T}_\gamma := \{T \in \mathcal{T} : \mathbb{E}_\infty(T) \ge \gamma\}, \tag{2.9}$$

where $\mathcal{T}$ is the set of stopping times on Ω with respect to $(\mathcal{F}_t)_{t \ge 0}$, taking values in $\bar{\mathbb{R}}_0^+$. Also, for $\Delta > 0$, let $\mathcal{T}(\Delta)$ and $\mathcal{T}_\gamma(\Delta)$ denote the corresponding subsets of $\Delta \bar{\mathbb{Z}}_0^+$-valued stopping times.[1] Since $\mathbb{P}_\infty$ is a probability measure under which $\tau = \infty$, that is, under which there is no change-point, the purpose of the constraint $\mathbb{E}_\infty(T) \ge \gamma$ in (2.9) is to serve as a lower bound on the mean time between false alarms (i.e., premature detection). Such a condition is needed since, as explained in the Introduction, change-point detection involves a trade-off between the delay until detection (i.e., the time while a change goes undetected) and the frequency of false alarms. This naturally gives rise to an optimization problem, and since our strategy to solve the continuous-time problem consists in approximating it by a sequence of discrete-time problems, the following section sets out with a discussion on Lorden's change-point problem in discrete time, and then introduces the corresponding problem for continuous-time stochastic processes.

[1] Let $\mathbb{Z}_0^+ := \{0, 1, \dots\}$ and $\bar{\mathbb{Z}}_0^+ := \mathbb{Z}_0^+ \cup \{\infty\}$.

3. Lorden's change-point problem. The minimax approach to change-point detection, wherein the change-point is assumed to be deterministic but unknown, was originally proposed by Lorden [12] in 1971. In this setting, detection delay is penalized *linearly* via its worst-case expected value, and the frequency of false alarms is constrained by a lower bound on the expected time between such events. In what follows we make this precise for discrete-time processes, and recall the seminal result of Moustakides [13], before moving on to the continuous-time case and presenting our optimal change-detection result for Lévy processes.

3.1. *Discrete time.* To define Lorden's change-point problem for discrete-time stochastic processes, we need the following notation:

(i) On the sample space $\hat{\Omega} := \mathbb{R}^{\mathbb{N}}$, consider the canonical process $\hat{X}_k(\hat{\omega}) := \hat{\omega}(k)$, for $\hat{\omega} \in \hat{\Omega}$ and $k \geq 1$, and the natural filtration $(\hat{\mathcal{F}}_k)_{k\geq 0}$ defined by $\hat{\mathcal{F}}_0 := \{\hat{\Omega}, \varnothing\}$, $\hat{\mathcal{F}}_k := \sigma(\hat{X}_1, \dots, \hat{X}_k)$, for $k \geq 1$, and $\hat{\mathcal{F}}_\infty := \sigma(\hat{X}_k : k \geq 1)$.

(ii) For equivalent probability distributions Q_0 and Q_1 on $\mathbb{R}$, let $(\hat{\mathbb{P}}_k)_{k\geq 1}$ be a family of probability measures on $\hat{\Omega}$ such that, under $\hat{\mathbb{P}}_k$, $(\hat{X}_i)_{i\geq 1}$ are independent with $\hat{X}_1, \dots, \hat{X}_{k-1}$ having distribution Q_0 and $\hat{X}_k, \hat{X}_{k+1}, \dots$ having distribution Q_1. Let $\hat{\mathbb{P}}_\infty$ be a probability measure under which $(\hat{X}_i)_{i\geq 1}$ is i.i.d. with distribution Q_0, and denote by $\hat{\mathbb{E}}_k$ (resp., $\hat{\mathbb{E}}_\infty$) the expected value w.r.t. $\hat{\mathbb{P}}_k$ (resp., $\hat{\mathbb{P}}_\infty$).

(iii) Let $\hat{\mathcal{T}}$ be the set of $\bar{\mathbb{Z}}_0^+$-valued stopping times $\hat{T}$ on $\hat{\Omega}$ with respect to the filtration $(\hat{\mathcal{F}}_k)_{k\geq 0}$, and, for $\gamma > 0$, let $\hat{\mathcal{T}}_\gamma := \{\hat{T} \in \hat{\mathcal{T}} : \hat{\mathbb{E}}_\infty(\hat{T}) \geq \gamma\}$ be the subset of stopping times satisfying a lower bound on the mean time between false alarms.

In this setting, $\hat{\mathbb{P}}_k$ is a probability measure under which the change-point $\hat{\tau}$ is equal to k, that is, under which k is the first instant that the sequence is governed by the post-change distribution Q_1. In particular, $\hat{\mathbb{P}}_1$ is a measure under which the sequence is i.i.d. with distribution Q_1 (i.e., $\hat{\tau} = 1$) and $\hat{\mathbb{P}}_\infty$ is a measure under which the sequence is i.i.d. with distribution Q_0 (i.e., $\hat{\tau} = \infty$).

As a set of detection strategies, we consider all stopping times $\hat{T} \in \hat{\mathcal{T}}$, and the performance of a given stopping time is evaluated in the sense of Lorden [12], with a linear penalty on detection delay,[2]

$$\hat{d}(\hat{T}) := \sup_{k\geq 1} \operatorname{ess\,sup} \hat{\mathbb{E}}_k\big((\hat{T} - (k-1))^+ | \hat{\mathcal{F}}_{k-1}\big). \tag{3.1}$$

That is, detection delay is penalized via its worst-case expected value under each of the measures $\hat{\mathbb{P}}_k$, where the worst case is taken over all realizations of the process up to (and including) time $k-1$. The desire to make $\hat{d}(\hat{T})$ small must be balanced

[2] The essential supremum of a random variable X, defined on a generic probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is defined as $\operatorname{ess\,sup} X := \operatorname{ess\,sup}_{\omega\in\Omega} X(\omega) = \inf\{u \in \mathbb{R} : \mathbb{P}(X \geq u) = 0\}$, with the convention that $\inf \varnothing = \infty$.

with a constraint on the rate of false alarms, so Lorden's change-point detection problem is defined as the following optimization problem:

$$\hat{\Pi}^d_\gamma(Q_0, Q_1) := \inf_{\hat{T}\in\hat{\mathcal{T}}_\gamma} \hat{d}(\hat{T}), \tag{3.2}$$

where $\gamma > 0$, and the infimum is taken over all stopping times $\hat{T}$ that satisfy the constraint $\hat{\mathbb{E}}_\infty(\hat{T}) \geq \gamma$ on the mean time between false alarms.

The solution to this optimization problem is the widely used CUSUM procedure, as stated in the following theorem, originally due to Moustakides [13]. His methodology is based on reframing the problem so that it can be solved using the techniques of Markovian optimal stopping theory. The key step is to establish a convenient lower bound on the detection delay of a generic stopping time, and then proving that the lower bound is attained by CUSUM stopping times.

THEOREM 3.1 (Moustakides (1986) [13]). *Let $h \geq 0$ and define the CUSUM stopping time by*

$$\hat{T}^c_h := \inf\{k \geq 0 : \hat{S}_k \geq h\}, \tag{3.3}$$

where $\hat{S}_0 = 0$ and

$$\hat{S}_k := \max_{1\leq j\leq k} \prod_{i=j}^{k} \hat{L}(\hat{X}_i) = \max(\hat{S}_{k-1}, 1)\hat{L}(\hat{X}_k), \qquad k \geq 1, \tag{3.4}$$

where $\hat{L} := dQ_1/dQ_0$ is the Radon–Nikodým derivative of Q_1 with respect to Q_0. Then $\hat{T}^c_h$ solves the optimization problem (3.1)–(3.2), with $\gamma = \hat{\mathbb{E}}_\infty(\hat{T}^c_h)$.

REMARK 3.2. (i) Note that $h > 0$ implies $\gamma = \hat{\mathbb{E}}_\infty(\hat{T}^c_h) \geq 1$, since $\hat{S}_0 = 0$, so at least one sample is needed for the barrier h to be breached. Hence, the theorem can equivalently be formulated for a fixed rate of false alarms $\gamma \geq 1$, assuming the existence of a barrier $h > 0$ such that $\hat{\mathbb{E}}_\infty(\hat{T}^c_h) = \gamma$. For $0 < \gamma < 1$, the optimal rule is to stop at $k = 0$ w.p. $1 - \gamma$, or stop at $k = 1$ w.p. γ. This stopping time outperforms any CUSUM rule, even after randomizing with $k = 0$. That is, if $\hat{T}^{c,p}_h = \hat{T}^c_h$ w.p. p, and $\hat{T}^{c,p}_h = 0$ w.p. $1 - p$, for some $h > 0$ and $0 < p < 1$ such that $\mathbb{E}_\infty(\hat{T}^{c,p}_h) = \gamma$, then $\hat{d}(\hat{T}^{c,p}_h) = \hat{d}(\hat{T}^c_h) \geq 1 > \gamma$. To be precise, these stopping times based on a randomization do not belong to the set of admissible stopping times $\hat{\mathcal{T}}_\gamma$, but that can simply be resolved by extending the probability space (see [9], Chapter 5) to include a random variable $\hat{X}_0 \in \hat{\mathcal{F}}_0$ that is uniformly distributed on [0, 1], and that is independent of $(\hat{X}_k)_{k\geq 1}$ under each of the measures $\hat{\mathbb{P}}_k$.

(ii) The optimality of CUSUM hinges on the linear delay penalty in (3.1). This type of penalty is suitable for many applications, such as the monitoring of manufacturing processes, where the cost of discarded items grows linearly. However, in other applications, it may be of interest to use a nonlinear cost function, such as in finance where the cost of an undetected change may increase exponentially. In this case, the CUSUM test can be arbitrarily unfavorable relative to the optimal test, if the rate at which delay penalty accumulates is too high relative to the rate at which information to discriminate between the pre-change and post-change distributions accumulates. However, in [19] it is shown that a simple and intuitive adaptation of the CUSUM procedure is optimal when (3.1) is replaced by an exponential cost of delay function.

An important implication of Theorem 3.1 is that CUSUM is optimal in Lorden's sense when sequentially observing evenly spaced increments of a continuous-time stochastic process like X, defined in (2.1), which under each of the measures $\mathbb{P}_\tau$, defined in (2.2), has independent and stationary increments before and after the change-point τ. To formalize this idea, we need to add to the notation introduced in Section 2:

(i) For $\Delta > 0$, denote by $Q_0^{(\Delta)}$ and $Q_1^{(\Delta)}$ the distributions of X_Δ under $\mathbb{P}_\infty$ and $\mathbb{P}_0$, respectively.

(ii) Define the filtration $(\breve{\mathcal{F}}_{k\Delta})_{k\geq 0}$ generated by the Δ-increments of the process X: $\breve{\mathcal{F}}_0 := \{\Omega, \varnothing\}$, $\breve{\mathcal{F}}_{k\Delta} := \sigma(\Delta_i X : 1 \leq i \leq k)$ for $k \geq 1$ and $\breve{\mathcal{F}}_\infty := \sigma(\Delta_k X : k \geq 1)$, where $\Delta_i X := X_{i\Delta} - X_{(i-1)\Delta}$, for $1 \leq i \leq k$.

(iii) Let $\breve{\mathcal{T}}(\Delta)$ be the set of $\Delta\bar{\mathbb{Z}}_0^+$-valued stopping times $\breve{T}$ on Ω with respect to $(\breve{\mathcal{F}}_{k\Delta})_{k\geq 0}$, and as before, let $\breve{\mathcal{T}}_\gamma(\Delta)$ be the subset of those stopping times that satisfy the false alarm constraint $\mathbb{E}_\infty(\breve{T}) \geq \gamma$.

Note that under the measure $\mathbb{P}_{k\Delta}$, with $k \geq 0$, the sequence of increments $(\Delta_i X)_{i\geq 1}$ consists of independent random variables whose marginal distribution changes from $Q_0^{(\Delta)}$ to $Q_1^{(\Delta)}$ after the kth increment. That is, under $\mathbb{P}_{k\Delta}$, the random variables $\Delta_1 X, \ldots, \Delta_k X$ have distribution $Q_0^{(\Delta)}$, while the random variables $\Delta_{k+1} X, \Delta_{k+2} X, \ldots$ have distribution $Q_1^{(\Delta)}$. Similarly, under $\mathbb{P}_\infty$ the sequence $(\Delta_i X)_{i\geq 1}$ is i.i.d. with distribution $Q_0^{(\Delta)}$.

It then follows from Theorem 3.1 that the CUSUM stopping time

$$\breve{T}_h^c(Q_0^{(\Delta)}, Q_1^{(\Delta)}) := \inf\{k\Delta \geq 0 : \breve{S}_{k\Delta} \geq h\} = \Delta \inf\{k \geq 0 : \breve{S}_{k\Delta} \geq h\},$$

where $h \geq 0$, $\breve{S}_0 = 0$, and

$$\breve{S}_{k\Delta} := \max_{1\leq j\leq k} \prod_{i=j}^{k} \frac{dQ_1^{(\Delta)}}{dQ_0^{(\Delta)}}(\Delta_i X) = \max(\breve{S}_{(k-1)\Delta}, 1)\frac{dQ_1^{(\Delta)}}{dQ_0^{(\Delta)}}(\Delta_k X), \qquad k \geq 1,$$

solves the Lorden-type optimization problem defined by

$$\breve{\Pi}_\gamma^d(Q_0^{(\Delta)}, Q_1^{(\Delta)}) := \inf_{\breve{T}\in\breve{\mathcal{T}}_\gamma(\Delta)} \breve{d}(\breve{T}, \Delta), \tag{3.5}$$

where

$$\breve{d}(\breve{T}, \Delta) := \sup_{k\geq 0} \operatorname{ess\,sup} \mathbb{E}_{k\Delta}\big((\breve{T} - k\Delta)^+ | \breve{\mathcal{F}}_{k\Delta}\big), \tag{3.6}$$

and $\gamma = \mathbb{E}_\infty(\breve{T}_h^c(Q_0^{(\Delta)}, Q_1^{(\Delta)}))$. In the following section (see Proposition 3.6 therein), we extend this result to a setting where rather than observing the discrete increments $(\Delta X_i)_{i\geq 1}$, one observes the entire trajectory of the process X, but the change-point is still assumed to take values in the discrete set $\Delta\bar{\mathbb{Z}}_0^+$.

3.2. *Continuous time*. Now we return to the continuous-time framework, as introduced in Section 2. Recall that under the probability measure $\mathbb{P}_\tau$, the distribution of the observed process X, defined in (2.1), undergoes an abrupt shift at the change-point τ, and $\tau \in \bar{\mathbb{R}}_0^+$ is assumed to be deterministic but unknown. The continuous-time analogue of Lorden's change-point detection problem (3.2) can then be defined as the optimization problem

$$\Pi_\gamma^c := \inf_{T\in\mathcal{T}_\gamma} d^c(T), \tag{3.7}$$

where the infimum is taken over all stopping times T with respect to the filtration generated by the observed process, that satisfy a lower bound on the mean time between false alarms, given by $\mathbb{E}_\infty(T) \geq \gamma$, and

$$d^c(T) := \sup_{\tau\geq 0} \operatorname{ess\,sup} \mathbb{E}_\tau\big((T - \tau)^+ | \mathcal{F}_\tau\big), \tag{3.8}$$

so detection delay is penalized linearly via its worst-case expected value under each of the measures $\mathbb{P}_\tau$.

The following theorem is our main result and it shows that the continuous-time Lorden problem (3.7)–(3.8) is solved by the continuous-time analogue of the CUSUM stopping time. The remarks that follow then discuss some extensions of the theorem, and provide examples for specific types of Lévy processes.

THEOREM 3.3. *Let $h \geq 1$ and define the CUSUM stopping time by*

$$T_h^c := \inf\{t \geq 0 : S_t \geq h\}, \tag{3.9}$$

where the CUSUM process $(S_t)_{t\geq 0}$ is defined by

$$S_t := \sup_{0\leq\tau\leq t} L_t^{(\tau)}, \qquad t \geq 0, \tag{3.10}$$

where $L_t^{(\tau)}$ is the likelihood ratio defined in (2.5). Then T_h^c solves Lorden's optimization problem (3.7)–(3.8) with $\gamma = \mathbb{E}_\infty(T_h^c)$.

REMARK 3.4. (i) This theorem encompasses previously established results on a change in the drift of a Brownian motion (see, e.g., [14]), and a change in the jump-intensity of a homogeneous Poisson process (cf. [7]). Moreover, in a unified framework it also includes changes in the statistical properties of more general Lévy processes, such as compound Poisson processes, jump-diffusions and Lévy processes with infinite jump activity.

(ii) In Section 2, we assumed the processes X^0 and X^1 to be càdlàg, but the theorem extends to any processes with independent and stationary increments that are continuous in probability, since such processes have unique càdlàg modifications that are identical in distribution to the original processes (cf. [21], Section 11).

(iii) The extension to multidimensional Lévy processes is also straightforward. The proof goes through without any significant changes if X^0 and X^1 are Lévy processes on $\mathbb{R}^d$ for some $d > 1$, with generating triplets $(A^{(0)}, b^{(0)}, \nu^{(0)})$ and $(A^{(1)}, b^{(1)}, \nu^{(1)})$, where the Brownian covariance matrices satisfy $A^{(0)} = A^{(1)}$, the drift parameters are such that $b^{(1)} - b^{(0)} - \int_{|x|\le 1} x(\nu^{(1)} - \nu^{(0)})(dx) = A^{(0)}\alpha$ for some $\alpha \in \mathbb{R}^d$, with $\alpha = 0$ if $A^{(0)} = 0$, and the Lévy measures $\nu^{(0)}$ and $\nu^{(1)}$ are equivalent and satisfy the integrability condition (2.3).

(iv) Our strategy of proof is based on approximating the continuous-time problem by discrete-time problems, and

$$\tilde{d}^c(T) := \sup_{\tau\ge 0} \operatorname{ess\,sup} \mathbb{E}_\tau\big((T-\tau)^+ | \mathcal{F}_{\tau-}\big), \qquad T \in \mathcal{T},$$

can be viewed as a natural continuous-time limit of Lorden's criterion (3.1), where $\hat{\mathcal{F}}_{k-1}$ is the information set *prior* to the change-point. However, it turns out that $\tilde{d}^c(T)$ coincides with Lorden's measure $d^c(T)$, for any $T \in \mathcal{T}$, due to the quasi-left-continuity of the filtration $(\mathcal{F}_t)_{t\ge 0}$.

REMARK 3.5. (i) The proof of Theorem 3.3 [see equation (3.24) below] shows that the CUSUM stopping time is an equalizer rule in the sense that its performance does not depend on the value of the change-point τ:

$$d^c\big(T_h^c\big) := \sup_{\tau\ge 0} \operatorname{ess\,sup} \mathbb{E}_\tau\big((T-\tau)^+ | \mathcal{F}_\tau\big) = \mathbb{E}_0\big(T_h^c\big).$$

The quantities $\mathbb{E}_0(T_h^c)$ and $\mathbb{E}_\infty(T_h^c)$ are generally referred to as the average run lengths (ARL) under the *out-of-control* and *in-control* regimes $\mathbb{P}_0$ and $\mathbb{P}_\infty$, respectively, and are standard measures of the performance of the CUSUM procedure.

(ii) The CUSUM process (3.10) is also known as the maximum likelihood ratio process, and by using (2.6) it is easy to see that $\hat{\tau} := \sup\{t \le T_h^c : S_t = 1\}$ is the maximum likelihood estimate for the change-point τ, based on the observed process up to time T_h^c. The CUSUM procedure thus combines detection and estimation, which is one reason for its sustained popularity in practical applications. It can also be viewed as a sequential procedure for testing the *in-control* null hypothesis H_0 against the *out-of-control* alternative H_1, with a change announced as soon as the maximum likelihood ratio test statistic (3.10) breaches a prescribed barrier. This barrier reflects the trade-off between a large ARL under H_0 and a small ARL under H_1, which are analogous to Type I and Type II error probabilities in conventional hypothesis testing.

(iii) Another useful representation of the CUSUM stopping time is

$$T_h^c = \inf\{t \geq 0 : \log(S_t) \geq \log(h)\} = \inf\{t \geq 0 : Y_t \geq \bar{h}\}, \tag{3.11}$$

where $\bar{h} := \log(h) \geq 0$ for $h \geq 1$, and, from (2.6)–(2.7), it follows that the process $(Y_t)_{t\geq 0}$ has the form

$$Y_t := \sup_{s\leq t}(U_t - U_s) = U_t - \inf_{0\leq s\leq t} U_s, \tag{3.12}$$

where $(U_t)_{t\geq 0}$ is the log-likelihood process defined in (2.8). This shows that the CUSUM stopping time is the first hitting time to $[\bar{h}, \infty)$ of the process $(U_t)_{t\geq 0}$ reflected at its running minimum. This is also referred to as the draw-up process of $(U_t)_{t\geq 0}$, and it has, along with the corresponding draw-down process, received considerable attention in the financial risk management literature (see [11] and references therein).

(iv) The expression (2.8) for U_t can be written more concisely for specific Lévy processes:

(a) Let X be a standard Brownian motion with a change in drift from 0 to a nonzero $\mu \in \mathbb{R}$. Then

$$U_t = \mu X_t - \frac{1}{2}\mu^2 t, \qquad t \geq 0, \tag{3.13}$$

so the process $(U_t)_{t\geq 0}$ is a Brownian motion with drift shifting from $-\mu^2/2 < 0$ to $\mu^2/2 > 0$ at the change-point τ, which in turn drives the process $(Y_t)_{t\geq 0}$ to the barrier $\bar{h}$.

(b) Let X be a compound Poisson process with a linear drift $b \in \mathbb{R}$ and a change in Lévy measures from $\nu^{(0)}$ to $\nu^{(1)}$. Then

$$U_t = \sum_{0\leq s\leq t} \varphi(\Delta X_s) - (\lambda^{(1)} - \lambda^{(0)})t, \qquad t \geq 0, \tag{3.14}$$

where $\lambda^{(i)} = \nu^{(i)}(\mathbb{R})$, $i = 0, 1$, are the pre-change and post-change jump intensities of X, and $\varphi = \log(d\nu^{(1)}/d\nu^{(0)})$. Furthermore, if $d\nu^{(1)}/d\nu^{(0)} \equiv \lambda^{(1)}/\lambda^{(0)}$, that is, only the overall jump intensity changes, then

$$U_t = \log\left(\frac{\lambda^{(1)}}{\lambda^{(0)}}\right) N_t - (\lambda^{(1)} - \lambda^{(0)})t, \qquad t \geq 0,$$

where $(N_t)_{t\geq 0}$ is a counting process with jump-intensity shifting from $\lambda^{(0)}$ to $\lambda^{(1)}$ at the change-point τ. If $\lambda^{(1)} < \lambda^{(0)}$ the process $(Y_t)_{t\geq 0}$ drifts continuously through the barrier $\bar{h}$, but if $\lambda^{(1)} > \lambda^{(0)}$ it crosses the barrier by jumping and may overshoot it.

(c) Let X be a jump-diffusion process, $X = X^c + X^j$ where X^c is a standard Brownian motion with a drift shifting from 0 to $\mu \neq 0$, and X^j a compound Poisson process with a Lévy measure changing from $\nu^{(0)}$ to $\nu^{(1)}$. In that case,

$$U_t = \mu X_t^c - \frac{1}{2}\mu^2 t + \sum_{0\le s\le t} \varphi(\Delta X_s^j) - (\lambda^{(1)} - \lambda^{(0)})t, \qquad t \ge 0,$$

which is simply the sum of the log-likelihood ratios in (3.13) and (3.14). In other words, information to distinguish between the pre-change and post-change distributions accumulates independently from the continuous component and the jump component, which is simply a consequence of their independence. This extends to Lévy processes with infinite jump activity for which the three components of the Lévy–Itô decomposition—the continuous component, the "small-jump" component and the "large-jump" component—are all independent (see, e.g., [21], Chapter 4).

(d) Let X be a pure-jump Lévy process with infinite jump activity. Then

$$U_t = \int_0^t \int_{\mathbb{R}\setminus\{0\}} \varphi(x)\bar{N}(dx, ds) + \beta t, \qquad t \ge 0,$$

where $\bar{N}$ is a compensated Poisson random measure with intensity measure $\nu^{(0)}(dx)\,dt$ under $\mathbb{P}_\infty$ and $\nu^{(1)}(dx)\,dt$ under $\mathbb{P}_0$, and condition (2.3) implies that the stochastic integral $(U_t - \beta t)_{t\ge 0}$ is a square-integrable zero-mean martingale. The drift β under $\mathbb{P}_\infty$ is

$$\beta^{(0)} = -\int_{\mathbb{R}\setminus\{0\}} (e^{\varphi(x)} - 1 - \varphi(x))\nu^{(0)}(dx) < 0,$$

while under $\mathbb{P}_0$ it is

$$\begin{aligned}\beta^{(1)} &= \beta^{(0)} + \int_{\mathbb{R}\setminus\{0\}} \varphi(x)(\nu^{(1)} - \nu^{(0)})(dx) \\ &= \int_{\mathbb{R}\setminus\{0\}} (e^{\varphi(x)}(\varphi(x) - 1) + 1)\nu^{(0)}(dx) > 0,\end{aligned}$$

which in turn pushes $(Y_t)_{t\ge 0}$ toward the barrier $\bar{h}$ when the change-point τ is passed. Note that condition (2.3) ensures that the integrals appearing in the drift coefficients are well defined.

As previously mentioned, the proof of Theorem 3.3 is based on considering a sequence of discrete-time problems. More precisely, the first step is to show that a "discretized" version of the CUSUM stopping time T_h^c solves a change-point problem where the change-point is restricted to take values in the discrete set $\Delta\bar{\mathbb{Z}}_0^+$, for some $\Delta > 0$. This gives rise to an optimization problem similar to the one in (3.5)–(3.6), but rather than conditioning on $\check{\mathcal{F}}_{k\Delta} = \sigma(\Delta_i X : 1 \le i \le k)$, the σ-algebra generated by the Δ-increments of the observed process, we condition on

$\mathcal{F}_{k\Delta} = \sigma(X_t, t \leq k\Delta)$, the σ-algebra generated by the paths of the process itself. The following proposition formalizes this idea, which is a nontrivial and somewhat unexpected extension of the result of Moustakides [13] for sequentially observed random variables.

PROPOSITION 3.6. *Let $h \geq 0$ and*

$$T_h^c(\Delta) := \Delta \inf\{k \geq 0 : S_k(\Delta) \geq h\}, \tag{3.15}$$

where $S_0(\Delta) = 0$, and

$$S_k(\Delta) := \sup_{0 \leq m < k} L_{k\Delta}^{(m\Delta)}, \qquad k \geq 1. \tag{3.16}$$

Then $T_h^c(\Delta)$ solves the optimization problem

$$\Pi_\gamma^c(\Delta) := \inf_{T \in \mathcal{T}_\gamma(\Delta)} d(T, \Delta), \tag{3.17}$$

where

$$d(T, \Delta) := \sup_{k \geq 0} \operatorname{ess\,sup} \mathbb{E}_{k\Delta}\big((T - k\Delta)^+ | \mathcal{F}_{k\Delta}\big), \tag{3.18}$$

and $\gamma = \mathbb{E}_\infty(T_h^c(\Delta))$.

REMARK 3.7. (i) This proposition serves as a stepping stone in the proof of Theorem 3.3, but it is also of importance in its own right. It states that the CUSUM stopping time (3.15) is optimal when continuously monitoring a process whose distribution undergoes a change at an unknown time τ that is assumed to belong to a discrete set of times, and the change is also declared at one of those times. For example, in financial applications the change may reasonably be assumed to take place at the beginning of a new business day, and in quality control a similar thing can be said about the change from the *in-control* state to the *out-of-control* state.

(ii) Remark 3.2(i) following Theorem 3.1 also applies here. That is, $\gamma = \mathbb{E}_\infty(T_h^c(\Delta)) \geq \Delta$ for any $h > 0$, so the theorem can equivalently be stated for a fixed $\gamma \geq \Delta$, assuming the existence of a barrier h such that $\mathbb{E}_\infty(T_h^c(\Delta)) = \gamma$. On the other hand, for $0 < \gamma < \Delta$ the optimal stopping rule is to randomize between $k = 0$ and $k = \Delta$, with probabilities $1 - \gamma$ and γ, respectively.

(iii) As in the discrete-time case [see Eq. (3.4)], it is easy to check that (2.6) implies the following recursive formula for the CUSUM process (3.16):

$$S_k(\Delta) = \max\big(S_{k-1}(\Delta), 1\big)L_k(\Delta), \qquad k \geq 1, \tag{3.19}$$

where for brevity we have defined $L_k(\Delta) := L_{k\Delta}^{((k-1)\Delta)}$.

The proof of the above proposition follows similar steps as the proof of Theorem 3.1, using the methodology developed by Moustakides in [13], and is deferred to the Appendix.

Before proving Theorem 3.3, we introduce two lemmas. The first one says that the CUSUM stopping time T_h^c coincides with the first hitting time of the CUSUM process to the open set (h, ∞), and that T_h^c changes continuously as the barrier h is increased. The second lemma states that the discretized CUSUM stopping time $T_h^c(\Delta)$ converges to T_h^c, as the step size Δ is reduced. Note that since $\{\mathbb{P}_\tau, \tau \in \bar{\mathbb{R}}_0^+\}$ is a family of equivalent probability measures, almost surely in the following lemmas actually holds with respect to any of those measures. Similarly, since the Lévy measures $\nu^{(0)}$ and $\nu^{(1)}$ are assumed to be equivalent, condition (3.20) holds for both $\nu^{(0)}$ and $\nu^{(1)}$ or neither of them.

LEMMA 3.8. *Let $h > 1$ and assume that under the measures $\mathbb{P}_0$ and $\mathbb{P}_\infty$, the Lévy measure of the log-likelihood ratio process $U := (U_t)_{t\geq 0}$, defined in* (2.8), *does not have an atom at $\bar{h} = \log(h)$. That is,*

$$(3.20)\qquad \left(\nu^{(i)} \circ \varphi^{-1}\right)(\bar{h}) = \nu^{(i)}\left(\left\{x \in \mathbb{R} : \varphi(x) = \bar{h}\right\}\right) = 0, \qquad i = 0, 1,$$

where $\varphi = \log(d\nu^{(1)}/d\nu^{(0)})$. Then the following assertions hold true almost surely, under the measures $\mathbb{P}_0$ and $\mathbb{P}_\infty$, for the CUSUM stopping time T_h^c defined in (3.9):

(i) $T_h^c = \tau_h := \inf\{t \geq 0 : S_t > h\}$.
(ii) $T_{h-\varepsilon}^c \leq T_h^c$, *for any $\varepsilon > 0$.*
(iii) $T_{h-\varepsilon}^c \to T_h^c$, *as $\varepsilon \downarrow 0$.*

PROOF. To prove (i), recall the representation (3.11)–(3.12) for T_h^c in terms of the drawup process $Y := (Y_t)_{t\geq 0}$, that is, the process $U := (U_t)_{t\geq 0}$ reflected at its running minimum, and observe that the paths of U can be decomposed into independent excursions from its running minimum, potentially interlaced by time intervals where the process can be described as drifting at its minimum.[3] Then note that $Y_{T_h^c} \geq \bar{h}$, and that if U is not a compound Poisson process, then the process Y can cross the barrier $\bar{h}$ in two different ways, which we now proceed to describe.

First, Y is said to creep through the barrier if $Y_{\tau_h^c} = \bar{h}$. If U is an infinite variation Lévy process, then $T_h^c = \tau_h$ follows from the strong Markov property and the point 0 being regular for $(0, \infty)$, which makes $T_h^c < \tau_h$ impossible. If U is a bounded variation Lévy process, the same argument can be used because the point 0 is regular for $(0, \infty)$ when the drift $d^{(i)} = b^{(i)} - \int_{|x|\leq 1} x\nu^{(i)}(dx)$ of the process is positive, which is also a necessary condition for a bounded variation process to creep through a barrier with a positive probability (see Theorems 6.5 and 7.11 in [10]).

[3]Such intervals, contributing to the Lebesgue measure of the time the process spends at its minimum, are not restricted to processes with a compound Poisson jump component. For instance, any spectrally positive and bounded variation Lévy process X, with generating triplet $(0, b, \nu)$, can drift at its minimum if $d = b - \int_{|x|\leq 1} x\nu(dx) < 0$, because in that case $X_t/t \to d$ a.s. as $t \to 0$ (cf. [21], page 323).

Second, Y can cross the barrier by jumping. However, since $Y = 0$ during the intermediate times when U is at its minimum, condition (3.20) ensures that Y cannot jump straight to the barrier $\bar{h}$, so in that case $T_h^c = \tau_h$. Similarly, during an excursion of U from its running minimum, Y breaches the barrier $\bar{h}$ by overshooting it, so $T_h^c = \tau_h$. This is because for a Lévy process X that is not a compound Poisson process, $\{X_{\hat{\tau}_x} = x, X_{\hat{\tau}_x-} < x\}$ is a null event, where for $x > 0$, $\hat{\tau}_x = \inf\{t \geq 0 : X_t \geq x\}$. In other words, X cannot strike a given barrier from a position strictly below it. This follows from [10], Lemma 5.8 when X is a subordinator, while for a general Lévy process X it holds because the range of the running maximum process, $\bar{X}_t := \sup_{0\leq s\leq t} X_t$, coincides almost surely with the range of the ascending ladder heights process of X, which is a subordinator and cannot jump to the level x from below it (cf. [10], page 219).

Finally, we consider the case when U is a compound Poisson process, which happens when X is a compound Poisson process with the same pre- and post-change jump intensity, but a different jump size distribution (see Eq. 3.14). In this case condition (3.20) ensures that Y cannot hit the barrier $\bar{h}$ starting from zero, but Y can potentially do so in a finite number of jumps. Then $\bar{h}$ is said to be Δ-accessible (cf. [17]), but the number of such points is finite or countable, so we can find a sequence $(\varepsilon_n)_{n\geq 1}$ such that $\varepsilon_n \downarrow 0$ and $\bar{h} - \varepsilon_n$ is not Δ-accessible. For such points it is clear that $T_{h-\varepsilon_n}^c = \tau_{h-\varepsilon_n}$, and by part (iii) of this lemma we have $T_{h-\varepsilon_n}^c \to T_h^c$ as $n \to \infty$, and it follows that $T_h^c = \tau_h$.

To show (ii), note that $[h, \infty) \subset [h - \varepsilon, \infty)$ for any $\varepsilon > 0$, so $T_{h-\varepsilon}^c$ is an increasing sequence of stopping times, and $T_{h-\varepsilon}^c \leq T_h^c$, for all $\varepsilon > 0$. Thus, the limit $T = \lim_{\varepsilon\to 0} T_{h-\varepsilon}^c$ is a stopping time and $T \leq T_h^c$. Due to quasi-left-continuity of Lévy processes, we have $Y_{T_{h-\varepsilon}^c} \to Y_T$ almost surely, as $\varepsilon \to 0$, and since $Y_{T_{h-\varepsilon}^c} \in [h - \varepsilon, \infty)$ it follows that $Y_T \in [h, \infty)$ and, therefore, $T_h^c \leq T$. We conclude that $T = T_h^c$, so $T_{h-\varepsilon}^c \to T_h^c$, as $\varepsilon \to 0$, which proves (iii). Note that condition (3.20) is not needed for (ii) and (iii) to be satisfied. $\square$

LEMMA 3.9. *Let $h > 1$ and $(\Delta_n)_{n\geq 1}$ be such that $\Delta_n \mathbb{Z}_0^+ \subset \Delta_{n+1}\mathbb{Z}_0^+$ for all $n \geq 1$, and assume that condition* (3.20) *is satisfied. Then the following assertions hold true almost surely, under the measures $\mathbb{P}_0$ and $\mathbb{P}_\infty$, for the CUSUM stopping time T_h^c defined in* (3.9), *and the stopping times $(T_h^c(\Delta_n))_{n\geq 1}$ defined in* (3.15):

(i) $T_h^c \leq T_h^c(\Delta_{n+1}) \leq T_h^c(\Delta_n)$, $n \geq 1$.
(ii) $T_h^c(\Delta_n) \to T_h^c$, $n \to \infty$.

PROOF. Assertion (i) is clear from the definitions of T_h^c and $T_h^c(\Delta_n)$. To show (ii), recall as in the proof of the previous lemma, the representation (3.11)–(3.12) for T_h^c in terms of the draw-up process Y, and write $T_h^c(\Delta_n)$ in a similar way as

$$T_h^c(\Delta_n) = \inf\{k\Delta_n \geq 0 : S_k(\Delta_n) \geq h\} = \inf\{k\Delta_n \geq 0 : Y_{k\Delta_n}^{(\Delta_n)} \geq \bar{h}\},$$

where $\bar{h} = \log(h)$, and the discretized draw-up process is defined by

$$Y_{k\Delta_n}^{(\Delta_n)} := U_{k\Delta_n} - \inf_{0\le m<k} U_{m\Delta_n}, \qquad k \ge 0.$$

Since U is a Lévy process, its trajectories are càdlàg, and it follows that the trajectories of Y and $M_t := \inf_{s\le t} U_s$, are càdlàg as well. The process $(Y_{k\Delta_n}^{(\Delta_n)})_{k\ge 0}$ can also be extended to a piecewise constant càdlàg process by defining

$$Y_t^{(\Delta_n)} := Y_{k_t^{(n)}\Delta_n}^{(\Delta_n)}, \qquad t \ge 0,$$

where[4] $k_t^{(n)} := \lfloor t/\Delta_n \rfloor$, and we now show that

$$\forall t \in \bigcup_{n\ge 1} \Delta_n \mathbb{Z}_0^+ : \qquad Y_t^{(\Delta_n)} \xrightarrow{\text{a.s.}} Y_t, \qquad n \to \infty. \tag{3.21}$$

Indeed, for a fixed $t_0 \in \bigcup_{n\ge 1} \Delta_n \mathbb{Z}_0^+$ we have $k_{t_0}^{(n)} \Delta_n = t_0$ for n big enough, so $U_{k_{t_0}^{(n)}\Delta_n} = U_{t_0}$. The definition of $Y_{t_0}^{(\Delta_n)}$ then shows that a sufficient condition for the convergence $Y_{t_0}^{(\Delta_n)} \xrightarrow{\text{a.s.}} Y_{t_0}$ is given by

$$M_{t_0}^{(\Delta_n)} := \inf_{0\le m<k_{t_0}^{(n)}} U_{m\Delta_n} \xrightarrow{\text{a.s.}} \inf_{0\le s\le t_0} U_s = M_{t_0}, \qquad n \to \infty.$$

The definition of M_{t_0} and the right continuity of the process U show that for any $\varepsilon > 0$, there exist $s_\varepsilon \in [0, t_0]$ and $N_\varepsilon \in \mathbb{N}$ such that $s_\varepsilon \in \Delta_n \mathbb{Z}_0^+$ for all $n \ge N_\varepsilon$, and such that $U_{s_\varepsilon} < M_{t_0} + \varepsilon$. It follows that $M_{t_0}^{(\Delta_n)} < M_{t_0} + \varepsilon$, for all $n \ge N_\varepsilon$, which implies that $M_{t_0}^{(\Delta_n)} \xrightarrow{\text{a.s.}} M_{t_0}$, as $n \to \infty$ and, therefore, $Y_{t_0}^{(\Delta_n)} \xrightarrow{\text{a.s.}} Y_{t_0}$. The convergence (3.21) then follows from the fact that a countable union of almost sure events is also almost sure.

Now we show that $T_h^c(\Delta_n) \xrightarrow{\text{a.s.}} T_h^c$, as $n \to \infty$, that is, that the hitting time of $(Y_t^{(\Delta_n)})_{t\ge 0}$ to the set $[\bar{h}, \infty)$ converges to the corresponding hitting time of Y. By Lemma 3.8(i) and the right continuity of Y, for any $\varepsilon > 0$ there exists $t_\varepsilon \in [T_h^c, T_h^c + \varepsilon)$ such that $t_\varepsilon \in \Delta_n \mathbb{Z}_0^+$ for any n greater than some $N_{t_\varepsilon} \in \mathbb{N}$, and such that $Y_{t_\varepsilon} > \bar{h}$. By (3.21), $Y_{t_\varepsilon}^{(\Delta_n)} \to Y_{t_\varepsilon}$, as $n \to \infty$, so there exists $N_\varepsilon \in \mathbb{N}$ such that $Y_{t_\varepsilon}^{(\Delta_n)} > \bar{h}$, for $n \ge N_\varepsilon$. Thus, $T_h^c(\Delta_n) < T_h^c + \varepsilon + \Delta_n$ for any $N \ge N_\varepsilon$, which implies that $T_h^c(\Delta_n) \to T_h^c$, as $n \to \infty$. $\square$

We are now ready to prove Theorem 3.3, and thus show that the CUSUM stopping time T_h^c solves the continuous-time version of Lorden's change-point problem.

[4]For $x \in \mathbb{R}$, $\lfloor x \rfloor := \sup\{z \in \mathbb{Z} : z \le x\}$ and $\lceil x \rceil := \inf\{z \in \mathbb{Z} : z \ge x\}$.

PROOF OF THEOREM 3.3. Let $(\Delta_n)_{n\geq 1}$ and $(T_h^c(\Delta_n))_{n\geq 1}$ be as in Lemma 3.9, and assume that condition (3.20) is satisfied. In the Appendix, we show that

$$\begin{aligned} &d\big(T_h^c(\Delta_n), \Delta_n\big) \\ &\quad = \sup_{k\geq 1} \operatorname{ess\,sup} \mathbb{E}_{(k-1)\Delta_n}\big(\big(T_h^c(\Delta_n) - (k-1)\Delta_n\big)^+ | \mathcal{F}_{(k-1)\Delta_n}\big) \\ &\quad = \mathbb{E}_0\big(T_h^c(\Delta_n)\big), \end{aligned} \tag{3.22}$$

and a similar identity can be established for T_h^c. To see that, first note that from (2.6) it follows that $L_t^{(\tau)} = L_s^{(\tau)} \cdot L_t^{(s)}$, for any $\tau \leq s \leq t$, so

$$S_t = \max\Big(S_\tau L_t^{(\tau)}, \sup_{\tau\leq s\leq t} L_t^{(s)}\Big). \tag{3.23}$$

Since $(U_t)_{t\geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by X, it is clear from (2.7) and (2.8) that $L_t^{(s)} = e^{U_t - U_s}$, for $s \in [\tau, t]$, is measurable with respect to the filtration generated by $(X_s - X_\tau)_{\tau\leq s\leq t}$, and independent of $\mathcal{F}_\tau$. Hence, (3.23) shows that for fixed $(X_s - X_\tau)_{\tau\leq s\leq t}$, S_t is a nondecreasing function of S_τ, which implies that on $\{T_h^c \geq \tau\}$, T_h^c is a nonincreasing function of $S_\tau \in \mathcal{F}_\tau$. Thus, since $S_\tau \geq 1$,

$$\begin{aligned} d^c\big(T_h^c\big) &= \sup_{\tau>0} \operatorname{ess\,sup} \mathbb{E}_\tau\big(\big(T_h^c - \tau\big)^+ | \mathcal{F}_\tau\big) \\ &= \sup_{\tau>0} \operatorname{ess\,sup} \mathbb{E}_\tau\big(\big(T_h^c - \tau\big)^+ | S_\tau = 1\big) \\ &= \mathbb{E}_0\big(T_h^c\big), \end{aligned} \tag{3.24}$$

where the third equality follows from the homogeneous Markov property of $(S_t)_{t\geq 0}$. Using (3.22) and (3.24), as well as Lemma 3.9(i), now yields

$$d^c\big(T_h^c\big) = \mathbb{E}_0\big(T_h^c\big) \leq \liminf_n \mathbb{E}_0\big(T_h^c(\Delta_n)\big) = \liminf_n d\big(T_h^c(\Delta_n), \Delta_n\big), \tag{3.25}$$

and, furthermore, by Lemma 3.9(i) and (ii), and the monotone convergence theorem,

$$\gamma_n := \mathbb{E}_\infty\big(T_h^c(\Delta_n)\big) \searrow \mathbb{E}_\infty\big(T_h^c\big) = \gamma, \qquad n \to \infty. \tag{3.26}$$

Next, for a fixed $T \in \mathcal{T}_\gamma$, define the stopping times

$$T_n := \left\lceil \frac{T}{\Delta_n} \right\rceil \Delta_n + \left\lceil \frac{\gamma_n - \gamma}{\Delta_n} \right\rceil \Delta_n, \qquad n \geq 1,$$

which belong to $\mathcal{T}_{\gamma_n}(\Delta_n)$, so by Proposition 3.6,

$$d\big(T_h^c(\Delta_n), \Delta_n\big) \leq d(T_n, \Delta_n), \qquad n \geq 1. \tag{3.27}$$

Moreover, using $T_n \le T + (1+\eta_n)\Delta_n$, where $\eta_n := \lceil (\gamma_n - \gamma)/\Delta_n \rceil$, we have

$$\begin{aligned} d(T_n, \Delta_n) &= \sup_{m\ge 0} \operatorname{ess\,sup} \mathbb{E}_{m\Delta_n}\big((T_n - m\Delta_n)^+ | \mathcal{F}_{m\Delta_n}\big) \\ &\le \sup_{m\ge 0} \operatorname{ess\,sup} \mathbb{E}_{m\Delta_n}\big((T - m\Delta_n)^+ | \mathcal{F}_{m\Delta_n}\big) + (1+\eta_n)\Delta_n \\ &\le \sup_{\tau\ge 0} \operatorname{ess\,sup} \mathbb{E}_{\tau}\big((T-\tau)^+ | \mathcal{F}_{\tau}\big) + (1+\eta_n)\Delta_n \\ &= d^c(T) + (1+\eta_n)\Delta_n \\ &\to d^c(T), \qquad n \to \infty, \end{aligned}$$

since $\eta_n\Delta_n \le \gamma_n - \gamma + \Delta_n \to 0$, as $n \to \infty$, because of (3.26). This implies that

$$\limsup_n d(T_n, \Delta_n) \le d^c(T), \tag{3.28}$$

which together with (3.25) and (3.27), shows that

$$\begin{aligned} d^c\big(T_h^c\big) &\le \liminf_n d\big(T_h^c(\Delta_n), \Delta_n\big) \\ &\le \limsup_n d\big(T_h^c(\Delta_n), \Delta_n\big) \\ &\le \limsup_n d(T_n, \Delta_n) \\ &\le d^c(T). \end{aligned}$$

In other words, for a given $T \in \mathcal{T}_\gamma$ we have $d^c(T_h^c) \le d^c(T)$, which concludes the proof when condition (3.20) of Lemma 3.9 is satisfied.

If (3.20) is not satisfied, consider a sequence $(\varepsilon_n)_{n\ge 1}$ such that $\varepsilon_n \downarrow 0$ as $n \to \infty$, and such that the Lévy measures $\nu^{(i)} \circ \varphi^{-1}$ do not have an atom at $\bar{h} - \varepsilon_n$, that is, $\nu^{(i)}(\varphi^{-1}(\bar{h} - \varepsilon_n)) = 0$, $i = 0, 1$, for all $n \ge 1$. This is possible because Lévy measures are σ-finite and, therefore, have at most countably many atoms. In that case, we have shown that $d^c(T^c_{h-\varepsilon_n}) \le d^c(T)$, for any $T \in \mathcal{T}_{\gamma_{h-\varepsilon_n}}$, with $\gamma_{h-\varepsilon_n} := \mathbb{E}_\infty(T^c_{h-\varepsilon_n}) \le \gamma$. Since $\mathcal{T}_\gamma \subseteq \mathcal{T}_{\gamma_{h-\varepsilon_n}}$, it follows that $d^c(T^c_{h-\varepsilon_n}) \le d^c(T)$ is in particular true for any $T \in \mathcal{T}_\gamma$. To complete the proof, it is therefore sufficient to show that $d^c(T^c_{h-\varepsilon_n}) \to d^c(T^c_h)$ as $n \to \infty$, which follows from Lemma 3.8(iii) and the dominated convergence theorem, because $d^c(T^c_{h-\varepsilon_n}) = \mathbb{E}_0(T^c_{h-\varepsilon_n})$ and $d^c(T^c_h) = \mathbb{E}_0(T^c_h)$, by (3.24). $\square$

APPENDIX: ADDITIONAL PROOFS

PROOF OF (2.6). The definition of $L_t^{(\tau)}$ entails that

$$\mathbb{P}_\tau(B) = \mathbb{E}_\infty\big(\mathbf{1}_B L_t^{(\tau)}\big), \qquad \forall B \in \mathcal{F}_t,$$

so to prove (2.6) it is sufficient to show that

$$\mathbb{P}_\tau(B) = \mathbb{E}_\infty\left(\mathbf{1}_B \frac{L_t^{(0)}}{L_\tau^{(0)}}\right), \qquad \forall B \in \mathcal{F}_t. \tag{A.1}$$

We first show this for $B \in \mathcal{F}_t$ of the form

$$B = \{X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k, \ldots, X_{t_n} \in A_n\},$$

for some $n \geq 1$, $0 \leq t_1 < \cdots < t_{k-1} \leq \tau < t_k < \cdots < t_n \leq t$, and Borel sets $A_1, \ldots, A_n$. By invoking the monotone class theorem, (A.1) can then be shown to hold for any $B \in \mathcal{F}_t$. The details can be found in [8]. □

PROOF OF PROPOSITION 3.6. We first remark that it is sufficient to consider stopping times $T \in \mathcal{T}_\gamma(\Delta)$ that satisfy the constraint $\mathbb{E}_\infty(T) = \gamma$ with equality. First, if T satisfies $\mathbb{E}_\infty(T) = \infty$, then it can be excluded by choosing a sufficiently large integer n such that $\gamma \leq \mathbb{E}_\infty(T \wedge n\Delta) < \infty$, and $d(T \wedge n\Delta, \Delta) \leq d(T, \Delta)$. Second, if T satisfies $\gamma < \mathbb{E}_\infty(T) < \infty$, then we can consider a stopping time $T^{(p)}$ such that $T^{(p)} = T$ w.p. p, and $T^{(p)} = 0$ w.p. $1 - p$, where $p = \gamma/\mathbb{E}_\infty(T)$. Then $\mathbb{E}_\infty(T^{(p)}) = \gamma$, and $d(T^{(p)}, \Delta) \leq d(T, \Delta)$, so $T^{(p)}$ outperforms T, while satisfying the false alarm constraint.

After this simplifying observation, the proof rests on the following two results. The first one gives a convenient lower bound for the performance of a generic stopping time, which the CUSUM stopping time $T_h^c(\Delta)$ satisfies with equality, while the second one shows that $T_h^c(\Delta)$ is the solution to a key optimization problem. The proofs of these intermediary results follow closely the methodology developed by Moustakides in [13] and are therefore omitted, but can be found in full detail in [8] (see Lemma 3.8 and Proposition 3.9 therein).

LEMMA A.1. *Let $T \in \mathcal{T}(\Delta)$ such that $0 < \mathbb{E}_\infty(T) < \infty$. Then*

$$d(T, \Delta) \geq \bar{d}(T, \Delta) := \Delta \frac{\mathbb{E}_\infty(\sum_{k=0}^{T/\Delta-1} \max(S_k(\Delta), 1))}{\mathbb{E}_\infty(\sum_{k=0}^{T/\Delta-1} (1 - S_k(\Delta))^+)}, \tag{A.2}$$

with equality if $T = T_h^c(\Delta)$ for some $h > 0$.

PROPOSITION A.2. *Let $0 < h < \infty$, $\gamma = \mathbb{E}_\infty(T_h^c(\Delta))$, and $g : [0, \infty) \to \mathbb{R}$ be a nonincreasing and continuous function. Then T_h^c satisfies*

$$\sup_T \mathbb{E}_\infty\left(\sum_{k=0}^{T/\Delta-1} g(S_k(\Delta))\right) = \mathbb{E}_\infty\left(\sum_{k=0}^{T_h^c(\Delta)/\Delta-1} g(S_k(\Delta))\right), \tag{A.3}$$

where the supremum is taken over all stopping times $T \in \mathcal{T}(\Delta)$ that satisfy $\mathbb{E}_\infty(T) = \gamma$.

With the above results at our disposal, we can easily proof Proposition 3.6. The case $h = 0$ is trivial. For $h > 0$, take $g(x) = -\max(x, 1)$ and $g(x) = (1 - x)^+$ in (A.3) to see that $T_h^c(\Delta)$ simultaneously minimizes the numerator and maximizes the denominator of (A.2), over all stopping times $T \in \mathcal{T}(\Delta)$ with $\mathbb{E}_\infty(T) = \gamma$. From this, it follows that, for any such stopping time,

$$d(T, \Delta) \geq \bar{d}(T, \Delta) \geq \bar{d}\big(T_h^c(\Delta), \Delta\big) = d\big(T_h^c(\Delta), \Delta\big),$$

which shows that $T_h^c(\Delta)$ solves the optimization problem (3.17)–(3.18). $\square$

PROOF OF (3.22). From the recursive formula (3.19), it follows that for $n \geq k$, and for fixed $(L_m(\Delta))_{k<m\leq n}$, $S_n(\Delta)$ is an increasing function of $\max(S_k(\Delta), 1)$. Therefore, on the event $\{T_h^c(\Delta) \geq k\Delta\}$, $T_h^c(\Delta)$ is a nonincreasing function of $\max(S_k(\Delta), 1)$. Using that, and the homogeneous Markov property of $(S_k(\Delta))_{k\geq 1}$, we obtain for any $k \geq 0$:

$$\begin{aligned} d_k\big(T_h^c(\Delta), \Delta\big) &= \operatorname{ess\,sup} \mathbb{E}_{k\Delta}\big(\big(T_h^c(\Delta) - k\Delta\big)^+ | \mathcal{F}_{k\Delta}\big) \\ &= \operatorname{ess\,sup} \mathbb{E}_{k\Delta}\big(\big(T_h^c(\Delta) - k\Delta\big)^+ | S_k(\Delta) \leq 1\big) \\ &= \operatorname{ess\,sup} \mathbb{E}_0\big(T_h^c(\Delta)\big) \\ &= \mathbb{E}_0\big(T_h^c(\Delta)\big), \end{aligned}$$

from which it follows that $d(T_h^c(\Delta), \Delta) = \sup_{k\geq 1} d_k(T_h^c(\Delta), \Delta) = \mathbb{E}_0(T_h^c(\Delta_n))$. $\square$

REFERENCES

[1] ŠIRJAEV, A. N. (1963). On optimum methods in quickest detection problems. *Theory Probab. Appl.* **8** 22–46.

[2] ALBRECHER, H., ASADI, P. and IVANOVS, J. (2014). Exact boundaries in sequential testing for phase-type distributions. *J. Appl. Probab.* **51A** 347–358. MR3317368

[3] BAYRAKTAR, E., DAYANIK, S. and KARATZAS, I. (2006). Adaptive Poisson disorder problem. *Ann. Appl. Probab.* **16** 1190–1261. MR2260062

[4] BEIBEL, M. (1996). A note on Ritov's Bayes approach to the minimax property of the cusum procedure. *Ann. Statist.* **24** 1804–1812. MR1416661

[5] BÜCHER, A., HOFFMANN, M., VETTER, M. and DETTE, H. (2017). Nonparametric tests for detecting breaks in the jump behaviour of a time-continuous process. *Bernoulli* **23** 1335–1364. MR3606768

[6] CHRONOPOULOU, A. and FELLOURIS, G. (2013). Optimal sequential change detection for fractional diffusion-type processes. *J. Appl. Probab.* **50** 29–41. MR3076770

[7] EL KAROUI, N., LOISEL, S. and SALHI, Y. (2017). Minimax optimality in robust detection of a disorder time in doubly-stochastic Poisson processes. *Ann. Appl. Probab.* **27** 2515–2538. MR3693532

[8] FIGUEROA-LÓPEZ, J. E. and ÓLAFSSON, S. Change-point detection for Lévy processes. Preprint. Available at arXiv:1611.01561.

[9] KALLENBERG, O. (1997). *Foundations of Modern Probability*. Springer, New York. MR1464694

[10] KYPRIANOU, A. E. (2013). *Introductory Lectures on Fluctuations of Lévy Processes with Applications*, 2nd ed. Springer, Berlin.

[11] LANDRIAULT, D., LI, B. and ZHANG, H. (2017). On magnitude, asymptotics and duration of drawdowns for Lévy models. *Bernoulli* **23** 432–458. MR3556778

[12] LORDEN, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Stat.* **42** 1897–1908. MR0309251

[13] MOUSTAKIDES, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist.* **14** 1379–1387. MR0868306

[14] MOUSTAKIDES, G. V. (2004). Optimality of the CUSUM procedure in continuous time. *Ann. Statist.* **32** 302–315. MR2051009

[15] PAGE, E. S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115. MR0088850

[16] PESKIR, G. and SHIRYAEV, A. N. (2002). Solving the Poisson disorder problem. In *Advances in Finance and Stochastics* 295–312. Springer, Berlin. MR1929384

[17] PICARD, J. (1995/97). Density in small time for Levy processes. *ESAIM Probab. Stat.* **1** 357–389. MR1486930

[18] PLUMMER, P. J. and CHEN, J. (2014). A Bayesian approach for locating change points in a compound Poisson process with application to detecting DNA copy number variations. *J. Appl. Stat.* **41** 423–438. MR3291244

[19] POOR, H. V. (1998). Quickest detection with exponential penalty for delay. *Ann. Statist.* **26** 2179–2205. MR1700227

[20] POOR, H. V. and HADJILIADIS, O. (2009). *Quickest Detection*. Cambridge Univ. Press, Cambridge. MR2482527

[21] SATO, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. Translated from the 1990 Japanese original. Revised by the author. MR1739520

[22] SHIRYAEV, A. N. (1996). Minimax optimality of the method of cumulative sums (CUSUM) in the continuous case. *Russian Math. Surveys* **51** 750–751.

DEPARTMENT OF MATHEMATICS
WASHINGTON UNIVERSITY IN ST. LOUIS
ST. LOUIS, MISSOURI 63130
USA
E-MAIL: figueroa@math.wustl.edu

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY
UNIVERSITY OF CALIFORNIA, SANTA BARBARA
SANTA BARBARA, CALIFORNIA 93106
USA
E-MAIL: olafsson@pstat.ucsb.edu