OPTIMAL DETECTION OF A CHANGE-SET IN A SPATIAL POISSON PROCESS

BY B. GAIL IVANOFF[1] AND ELY MERZBACH

University of Ottawa and Bar-Ilan University

We generalize the classic change-point problem to a "change-set" framework: a spatial Poisson process changes its intensity on an unobservable random set. Optimal detection of the set is defined by maximizing the expected value of a gain function. In the case that the unknown change-set is defined by a locally finite set of incomparable points, we present a sufficient condition for optimal detection of the set using multiparameter martingale techniques. Two examples are discussed.

1. Introduction. In this paper, we consider the multiparameter version of the classic optimal detection problem; the goal is to detect the occurrence of a random set on which an observable Poisson process changes its intensity. To be precise, we let $N = \{N_t, t \in \mathbf{R}_+^2\}$ be a nonexplosive point process defined on the positive quadrant of the plane and let $\{\tau_n\}$ be its jump points, numbered in some arbitrary way. Then $N_t = \sum_{n=1}^{\infty} I_{\{\tau_n \le t\}}$ (cf. [6]). Here, "$\le$" denotes the usual partial order on $\mathbf{R}_+^2 : s = (s_1, s_2) \le t = (t_1, t_2) \Leftrightarrow s_1 \le t_1, s_2 \le t_2$. On some random set ξ, the intensity of N changes from μ_0 to μ_1, where $0 < \mu_0 < \mu_1$: specifically, given ξ, N is a Poisson process with intensity

$$\mu_0 I_{\{t \notin \xi\}} + \mu_1 I_{\{t \in \xi\}} = \mu_0 + (\mu_1 - \mu_0) I_{\{t \in \xi\}}.$$

The problem is that the "change-set" ξ is unobservable and we must detect ξ as well as possible, given our observation of the point process N. In particular, our goal is to find a random set $\hat{\xi}$ that maximizes the expected value of a specified valuation or gain function. The random set $\hat{\xi}$ must be adapted

to the underlying information structure: if the information available to us at $t \in \mathbf{R}_+^2$ is represented by the σ-field $\mathcal{F}_t$, then we must have $\{t \in \hat{\xi}\} \in \mathcal{F}_t$.

There are many potential areas of application. For example:

- Environment: The increased occurrence of polluted wells in a rural area could indicate a geographic region that has been subjected to industrial waste.
- Population health: Unusually frequent outbreaks of a disease such as leukemia near a nuclear power plant could signal a region of possible air or ground contamination.
- Astronomy: A cluster of black holes could be the result of an unobservable phenomenon affecting a region in space.
- Quality control: An increased rate of breakdowns in a certain type of equipment might follow the failure of one or more components.
- Archaeology: An increased number of archaeological items such as ancient coins found in a particular region could indicate the location of an event of historical interest.
- Forestry: The spread of an airborne disease through a forest would occur at a higher rate on ξ, the set of points to the northeast of the (unobserved) point (σ) of initial infection if the prevailing winds are from the southwest.

It is this final type of example, illustrated in Figure 1, that motivates the model to be studied in this paper.
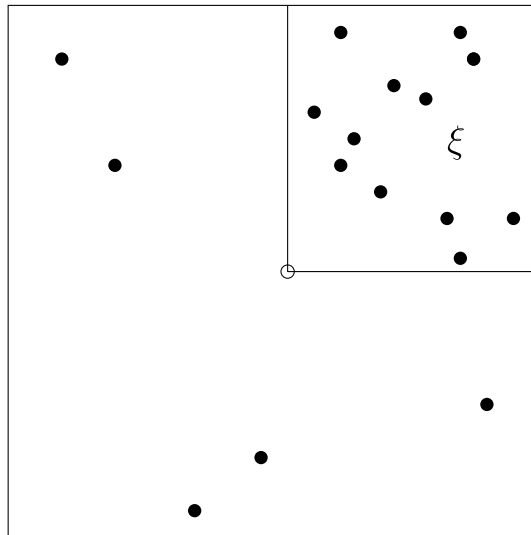
$\bullet$ = observed points of N, $\circ$ = σ, unobserved jump point of L

FIG. 1.   *A change-set ξ generated by a single point σ.*

As will be discussed in the conclusion, this paper represents only a first step in the solution of what we call the "optimal set-detection problem." Here, we consider the case in which the change-set ξ is a random *upper layer* (cf. Section 2) generated by a locally finite set of incomparable points. In general, the optimal solution $\hat{\xi}$ will be a random upper layer which is adapted to the available information structure. This means that the solution is exact in the sense that it is explicitly defined by the observed data points. This problem cannot be solved by one-parameter methods. Indeed, even if the random set is characterized by a single change-point, it will be seen that the optimal solution does not necessarily correspond to a point.

In the one-parameter case, the optimal detection of an exponential change time in a Poisson process was thoroughly studied in [5] using martingale techniques combined with Bayesian arguments (see also [13] for a different approach to the same problem). In the general set-indexed framework, we found only a very few papers addressing the problem of a change-point or a change-set (cf. [4], [10] and [11]). However, none of these papers deal with the question of the existence of an optimal solution to the detection problem. Our approach, inspired by that of [5], makes use of the general theory of set-indexed martingales as developed in [6]. We are then able to solve the problem with a Bayes-type formula.

The paper is structured as follows. In the next section, the model is presented and the optimal detection problem is formally defined. In Section 3, we give the necessary background for the multiparameter martingale approach that is the key for proving the existence of an optimal solution, and develop a semimartingale representation of the gain function. In Section 4, sufficient conditions for the existence of an optimal solution are developed, and then applied to two examples in Section 5. Finally, in Section 6, we discuss possible extensions and directions for further research.

2. The model. In order to better understand the two-dimensional model, we review the change-point problem on $\mathbf{R}_+$ considered in [5]. We have a nonexplosive point process $N = \{N_t, t \in \mathbf{R}_+\}$ on $\mathbf{R}_+$, and a random time $\sigma \geq 0$. Given σ, N is a Poisson process with intensity μ_0 on $[0, \sigma)$ and intensity μ_1 on $\xi = [\sigma, \infty)$ ($\mu_1 > \mu_0 > 0$). Modifying the notation of [5] slightly, the gain function at t is defined by

$$(1) \qquad Z_t = c_0(t \wedge \sigma) - c_1(t - \sigma)^+ + k_0 + k_1 I_{\{t \geq \sigma\}},$$

where $c_0 \geq 0$, $c_1 > 0$ and $k_1 \geq 0$. The parameters can be interpreted as follows: the gain function is piecewise linear, increasing at rate c_0 before the jump point and decreasing at rate c_1 after. When $k_1 > 0$, a penalty equivalent to $-k_1$ is incurred for stopping the process before the change has occurred. The gain is maximized when $t = \sigma$.

Let $\mathcal{F} = (\mathcal{F}_t, t \in \mathbf{R}_+)$ denote the filtration which characterizes the underlying information available (in [5], the process N is always $\mathcal{F}$-adapted). For various filtrations, it is shown in [5] that Z_t has a *smooth semimartingale* (SSM) representation with respect to $\mathcal{F}$:

$$(2) \qquad Z_t = Z_0 + \int_0^t U_s \, ds + M_t,$$

where M is an $\mathcal{F}$-martingale and U is $\mathcal{F}$-progressive (i.e., observable). If U is *monotone* in the sense that $U_t \leq 0 \Rightarrow U_{t+h} \leq 0 \; \forall h > 0$, then it is straightforward to see that (cf. [5], Theorem 1) $\hat{\sigma} := \inf\{t : U_t \leq 0\}$ is an optimal $\mathcal{F}$-stopping rule for Z in terms of expected values: we have

$$(3) \qquad E[Z_{\hat{\sigma}}] = \sup\{E[Z_\tau] : \tau \text{ an } \mathcal{F}\text{-stopping time}\}.$$

To motivate the model on $\mathbf{R}_+^2$, we will rewrite (1) in terms of the single jump point process $L_t = I_{\{\sigma \leq t\}}$ and the random set $\xi = [\sigma, \infty) = \{t : L_t > 0\}$:

$$
\begin{aligned}
Z_t &= c_0 |A_t \cap \xi^c| - c_1 |A_t \cap \xi| + k_0 + k_1 L_t \\
&= k_0 + \int_{A_t} (-c_1 + (c_0 + c_1) X_u) \, du + k_1 L_t,
\end{aligned}
$$
(4)

where $A_t = [0, t]$, $|\cdot|$ denotes Lebesgue measure and $X_t = 1 - I_{\{t \in \xi\}} = I_{\{L_t = 0\}}$.

We are now ready to describe the two-dimensional model. We are given a random Borel set $\xi \subset (0, \infty)^2$. N is a nonexplosive point process on $\mathbf{R}^2$ such that given ξ, N is Poisson with intensity μ_0 on ξ^c and μ_1 on ξ. It is always assumed that $\mu_1 > \mu_0 > 0$. (The case $\mu_0 = 0$ will be briefly discussed at the end of Section 4.) We will assume that the set ξ is generated by a *single line point process* L: that is, L is a nonexplosive point process whose jump points are all incomparable ($s, t \in \mathbf{R}_+^2$ are incomparable if both $s \not\leq t$ and $t \not\leq s$). It is noted in [7] that in two or more dimensions, the single line process is the natural generalization of the single jump process, and in analogy with the change-point model on $\mathbf{R}_+$, we define $\xi := \{t : L_t > 0\}$. We observe that ξ is an *upper layer* (ξ is an upper layer if $t \in \xi \Rightarrow s \in \xi \; \forall s \geq t$). When L has only one jump point σ, we observe that ξ consists of the points to the northeast of σ. This is illustrated in Figure 1. The more general situation in which L is a single line process is illustrated in Figure 2. In this case, ξ consists of all the points to the northeast of one or more jump points of L.

Using notation similar to that used for the one-dimensional problem, for $t \in \mathbf{R}_+^2$ let $A_t = \{s \in \mathbf{R}_+^2 : s \leq t\}$ and $X_t = 1 - I_{\{t \in \xi\}} = I_{\{L_t = 0\}}$. The definition of the gain function at $t \in \mathbf{R}_+^2$ is exactly the same is in (4):

$$
\begin{aligned}
Z_t &= c_0 |A_t \cap \xi^c| - c_1 |A_t \cap \xi| + k_0 + k_1 L_t \\
&= k_0 + \int_{A_t} (-c_1 + (c_0 + c_1) X_u) \, du + k_1 L_t.
\end{aligned}
$$
(5)

Once again, we assume that $c_0 \geq 0$, $c_1 > 0$ and $k_1 \geq 0$, and that $|\cdot|$ denotes Lebesgue measure on $\mathbf{R}_+^2$.
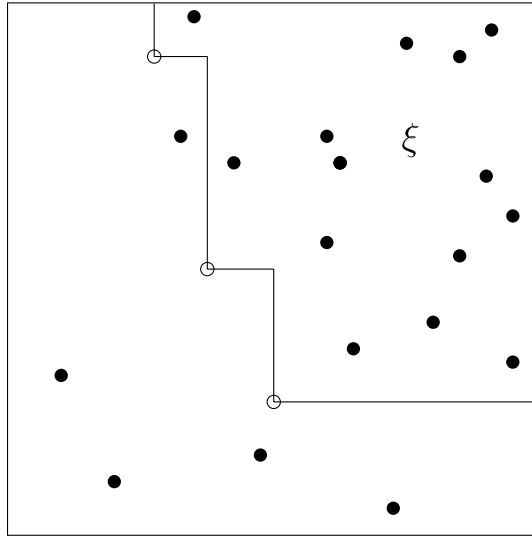
$\bullet$ = observed points of N, $\circ$ = unobserved points of L

FIG. 2. *A change-set ξ generated by a single line process L.*
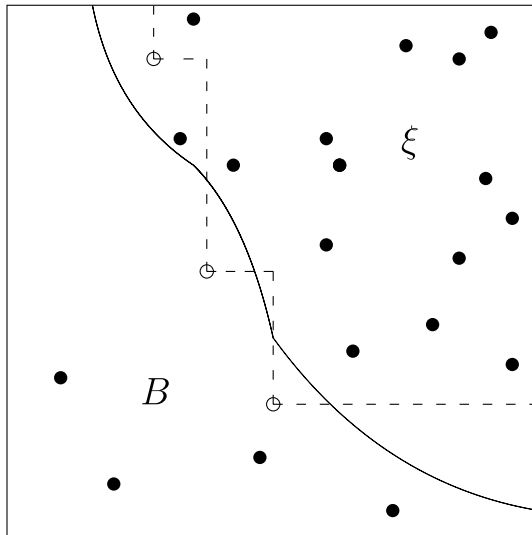
$\bullet$ = observed points of N, $\circ$ = unobserved points of L

FIG. 3. *A lower layer B and the change-set ξ.*

Any point process N can be indexed by the Borel sets in $\mathbf{R}_+^2$. As in the Introduction, if $\{\tau_n\}$ denotes the jump points of N numbered in some arbitrary way, then for any Borel set B, $N(B) := \sum_{n=1}^{\infty} I(\tau_n \in B)$. [Therefore, we have $N_t = N(A_t)$.] Consequently, we can define the gain function more generally over the class of *lower layers* $\mathcal{L}$: a set $B \subseteq \mathbf{R}_+^2$ is a lower layer if $t \in B \Rightarrow A_t \subseteq B \, \forall t \in \mathbf{R}_+^2$. The gain function at $B \in \mathcal{L}$ is defined as

$$Z(B) = c_0 |B \cap \xi^c| - c_1 |B \cap \xi| + k_0 + k_1 L(B)$$

(6)

$$= k_0 + \int_B (-c_1 + (c_0 + c_1) X_u) \, du + k_1 L(B).$$

A lower layer B and the change-set ξ are illustrated in Figure 3; we observe that $L(B) = 2$ in this case.

We see that the gain function defined in (6) is a natural generalization of the one-dimensional gain function (1). The gain evaluated at B increases in proportion to the area of B outside of the change-set ξ, and decreases in proportion to the area inside of ξ. When $k_1 > 0$, there is a penalty incurred that is equivalent to $-k_1$ times the number of points in L that lie outside of (or "after") B. The gain is maximized when $B = \overline{\xi^c}$.

We would like to find a random lower layer that maximizes the expected value of the gain function. The lower layer will depend on the available information, or more precisely, the underlying *filtration.*

A class of σ-fields $\mathcal{F} = \{\mathcal{F}_t, t \in \mathbf{R}_+^2\}$ is a filtration if:

- $\mathcal{F}$ is increasing: $s \leq t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t$, and
- $\mathcal{F}$ is outer-continuous: $\mathcal{F}_t = \bigcap_n \mathcal{F}_{t_n}$ for every decreasing sequence $(t_n) \subset \mathbf{R}_+^2$ with $t_n \downarrow t$.

DEFINITION 2.1 (Cf. [6]). A closed random lower layer ρ is an $\mathcal{F}$-stopping set if

$$\{t \in \rho\} \in \mathcal{F}_t \qquad \forall t \in \mathbf{R}_+^2.$$

The general optimal set-detection problem in two dimensions can now be stated as follows: for a given filtration $\mathcal{F}$, our goal is to maximize $E[Z_\rho]$, where ρ is an $\mathcal{F}$-stopping set. If it can be shown that a stopping set $\hat{\rho}$ exists that satisfies the condition

(7) $E[Z(\hat{\rho})] = \sup\{E[Z(\rho)] : \rho \text{ an } \mathcal{F}\text{-stopping set}\},$

then our optimal estimate of ξ is $\hat{\xi} = \overline{\hat{\rho}^c}$ [$\overline{(\cdot)}$ denotes set closure]. It is trivial that $\hat{\xi}$ is an upper layer, and by outer continuity of $\mathcal{F}$, it is easily seen that $\hat{\xi}$ is also an adapted random set (i.e., $\{t \in \hat{\xi}\} \in \mathcal{F}_t \, \forall t \in \mathbf{R}_+^2$).

In this paper, we will be focussing on the *sequential* estimation problem: that is, we will be assuming that $\mathcal{F}_t = \mathcal{F}_t^N = \sigma\{N_s : s \leq t\}$. If ρ is an $\mathcal{F}^N$-stopping set, then $I(t \in \rho)$ is a function of the number and locations of jump

points of N in the set A_t. For technical reasons, we shall see that in general it is necessary to restrict the detection problem to a bounded rectangle $R = [0, r]^2$. The goal is to find a stopping set $\hat{\rho} \subseteq R$ that is optimal in the following sense:

DEFINITION 2.2. An $\mathcal{F}^N$-stopping set $\hat{\rho}$ is called an optimal solution to the sequential detection problem on R provided that $\hat{\rho}$ satisfies the following equation:

(8) $$E[Z(\hat{\rho})] = \sup\{E[Z(\rho)] : \rho \subseteq R \text{ an } \mathcal{F}^N\text{-stopping set}\}.$$

Restricting our attention to R ensures that $\hat{\rho}$ is bounded and so $E[Z(\hat{\rho})]$ is always well defined. In this case, we have an optimal estimate $\hat{\xi}_R$ of $\xi \cap R$, defined by $\hat{\xi}_R = \overline{R \setminus \hat{\rho}}$.

3. Mathematical preliminaries. In this section we present the mathematical tools needed in the sequel. In [5], Herberts and Jensen make use of martingale techniques to provide a simple and elegant method of finding sufficient conditions for the existence of an optimal solution to the detection problem on $\mathbf{R}_+$. Martingale methods have been extended to more general spaces in [6], and we are able to exploit this theory in a similar way. To motivate the necessary technical details that follow, we first describe our overall plan of attack. Recall that $\mathcal{F}^N$ denotes the filtration representing the data that can be observed, and below $\mathcal{G}$ will denote a larger filtration containing additional information, some of which cannot be observed.

Plan of attack:

- The gain function Z can be rewritten as a (two-parameter) semimartingale (Definition 3.7):

$$Z_B = k_0 + \int_B U_t \, dt + k_1 M_B,$$

where M is a weak martingale (Definition 3.1) with respect to a filtration $\mathcal{G}$ and U is $\mathcal{G}$-adapted but not necessarily observable (cf. Lemma 3.14).
- For the observable filtration $\mathcal{F}^N$ and ρ an $\mathcal{F}^N$-stopping set, we have $E[M_\rho] = 0$ (Lemma 3.6) and if $V_t = E[U_t | \mathcal{F}_t^N]$ (observable), then (Lemma 3.10)

$$E[Z_\rho] = k_0 + E\left[\int_\rho U_t \, dt\right] = k_0 + E\left[\int_\rho V_t \, dt\right].$$

- If V satisfies a monotonicity property on R (cf. Definition 3.8 and Lemma 3.10), then there exists an optimal solution $\hat{\rho}$ to the sequential detection problem on R, and the optimal estimate of $\xi \cap R$ is

$$\hat{\xi}_R = \{t \in R : V_t \le 0\}.$$

Keeping this outline of our approach in mind, we continue with the necessary mathematical details.

3.1. *Martingale preliminaries.* Martingales on $\mathbf{R}_+^2$ can be defined in various ways (cf. [6]), but here we need only the weakest definition. In what follows, T denotes either $\mathbf{R}_+^2$ or a bounded region $R = [0,r]^2$, and $(\Omega, \mathcal{F}, P)$ is a complete probability space equipped with a T-indexed filtration $\mathcal{F} = \{\mathcal{F}_t : t \in T\}$ (without loss of generality, assume that $\mathcal{F}_t$ contains all the P-null sets $\forall t \in T$). A T-indexed process $X = \{X_t : t \in T\}$ is adapted to $\mathcal{F}$ if X_t is $\mathcal{F}_t$-measurable, for all $t \in T$. For any T-indexed process $X = \{X_t : t \in T\}$, for $s = (s_1, s_2) \le (t_1, t_2) = t \in T$, define the increment of X over the rectangle $(s, t] = (s_1, t_1] \times (s_2, t_2]$ in the usual way:

$$X(s, t] = X_{(t_1, t_2)} - X_{(s_1, t_2)} - X_{(t_1, s_2)} + X_{(s_1, s_2)}.$$

DEFINITION 3.1.   Let $M = \{M_t : t \in T\}$ be an integrable process on T, adapted to a filtration $\mathcal{F} = \{\mathcal{F}_t : t \in T\}$. M is a weak $\mathcal{F}$-(sub)martingale if M is equal to 0 on the axes, and for every $s \le t \in T$,

$$E[M(s, t]|\mathcal{F}_s] = (\ge)0.$$

(A process X is integrable if $E[|X_t|] < \infty \, \forall t \in T$.)

DEFINITION 3.2.   Let $v = \{v_t : t \in T\}$ be a function on T. We say that v is increasing (decreasing) if:

- v is 0 on the axes,
- v is outer continuous with inner limits: that is, v is continuous from above and with limits from the other three quadrants at each $t \in T$, and
- for every $s \le t \in T$, $v(s, t] \ge (\le)0$.

A process $V = \{V_t : t \in T\}$ is increasing (decreasing) if for each $\omega \in \Omega$, the function $V.(\omega)$ is increasing (decreasing).

COMMENT 3.3.   An increasing function v can be regarded as the distribution of a measure on $\mathbf{R}_+^2$. Therefore, $v(B)$ is well defined for any Borel set B, where we use $v.$ and $v(\cdot)$ to denote, respectively, the function and the generated measure. Likewise, a decreasing function generates a negative measure, and we will use similar notation.

DEFINITION 3.4.   Let L be a weak $\mathcal{F}$-submartingale. An increasing process Λ is a compensator for L if Λ is $\mathcal{F}$-adapted and $M = L - \Lambda$ is a weak martingale.

COMMENT 3.5.   As defined above, the compensator of a submartingale need not be unique (any increasing process is trivially a compensator for itself). A type of predictability is required for uniqueness (cf. [6]), but this point is not of importance here.

In light of Comment 3.3, the following lemma is a special case of of Lemma 3.3.5 of [6].

LEMMA 3.6.   *If M is a weak martingale which can be expressed as the difference of two increasing integrable processes, and ρ is a stopping set such that $\rho \subseteq R = [0, r]^2$, then $M(\rho)$ is well defined and $E[M(\rho)] = 0$.*

DEFINITION 3.7.   Let $Z = \{Z_t : t \in T\}$ be a process on T, adapted to a filtration $\mathcal{F} = \{\mathcal{F}_t : t \in T\}$. Z is a smooth semimartingale with respect to $\mathcal{F}$ ($\mathcal{F}$-SSM) if it satisfies a decomposition of the form

$$(9) \qquad Z_t = Z_{(0,0)} + \int_0^{t_1} \int_0^{t_2} U_{(s_1, s_2)} \, ds_2 \, ds_1 + M_t$$

for each $t = (t_1, t_2) \in T$, where U is an outer continuous process with inner limits adapted to $\mathcal{F}$ and M is a weak $\mathcal{F}$-martingale. We denote the $\mathcal{F}$-SSM as $Z = (U, M)$.

In order to show that an optimal solution exists to the sequential detection problem, we will require a monotonicity property.

DEFINITION 3.8.   A function $v = \{v_t : t \in T\}$ is monotone on T if $v_s \leq 0 \Rightarrow v_t \leq 0 \, \forall t \geq s \in T$. A process V is monotone if $V.(\omega)$ is monotone for each $\omega \in \Omega$.

COMMENT 3.9.

1. Note that any decreasing function is monotone, but the converse is not true.
2. If a process V is decreasing in each parameter separately on T, then V is monotone on T but not necessarily decreasing in the sense of Definition 3.2.
3. Note that if V is monotone, then $V_t > 0 \Rightarrow V_s > 0 \, \forall s \leq t$.
4. If V is monotone and adapted to a filtration $\mathcal{F}$, the set

$$(10) \qquad \hat{\rho} = \{t \in T : V_s > 0 \, \forall s \ll t\}$$

is an $\mathcal{F}$-stopping set (cf. Definition 2.1). [$s = (s_1, s_2) \ll (t_1, t_2) = t \Leftrightarrow s_i < t_i$ if $t_i > 0$, and $s_i = 0$ if $t_i = 0$, $i = 1, 2$.] Clearly, $\hat{\rho}$ is a random closed

lower layer, and the fact that V is adapted ensures that $\{t \in \hat{\rho}\} \in \mathcal{F}_t$: taking any sequence $(t_n) \uparrow t$ with $t_n \ll t$, by monotonicity it follows that

$$\{t \in \hat{\rho}\} = \bigcap_n \{V_{t_n} > 0\} \in \bigcup_n \mathcal{F}_{t_n} \subseteq \mathcal{F}_t.$$

In [5], the solution to the optimal stopping problem is based on a SSM representation of the form (2), which in turn is based on a projection theorem. The question of the existence of optional and predictable projections in higher dimensions is a delicate one, usually requiring a strong assumption of conditional independence on the underlying filtration [denoted (F4) in the two-dimensional literature]. For details, see [12], for example. In practice, one can generally show directly that a suitable projection exists without relying on a general existence theorem, and for our purposes the following lemma will be adequate.

LEMMA 3.10. *Let U be a bounded T-indexed process adapted to a filtration $\mathcal{G}$ such that U is outer-continuous with inner limits. If $\mathcal{F}$ is a subfiltration of $\mathcal{G}$ (i.e., $\mathcal{F}_t \subseteq \mathcal{G}_t \forall t$), and if a version of $V_t = E[U_t|\mathcal{F}_t]$ exists that is outer-continuous with inner limits, then for any $\mathcal{F}$-stopping set $\rho \subseteq R = [0,r]^2$,*

$$(11) \qquad E\left[\int_\rho U_t \, dt\right] = E\left[\int_\rho V_t \, dt\right].$$

In addition, if V is monotone on R, then the $\mathcal{F}$-stopping set $\hat{\rho} \subseteq R$ defined by

$$(12) \qquad \hat{\rho} = \{t \in R : V_s > 0 \, \forall s \ll t\}$$

is optimal in the sense that

$$E\left[\int_{\hat{\rho}} U_t \, dt\right] = \sup\left\{E\left[\int_\rho U_t \, dt\right] : \rho \subseteq R, \rho \text{ an } \mathcal{F}\text{-stopping set}\right\}.$$

PROOF. First, the assumption that U and V have sample paths that are regular (outer-continuous with inner limits) and that U (and hence V) is bounded ensures that the integrals and expectations in (11) are well defined.

Next, let $T_n := \{(\frac{i}{2^n}r, \frac{j}{2^n}r) : 0 \leq i, j \leq 2^n\}$ denote the "dyadics" of order n in R. The class of rectangles $\mathcal{C}_n$ partitions R, where $C \in \mathcal{C}_n$ if C is of the form $C = A_t \setminus (\bigcup_{s \in T_n, s \not\geq t} A_s)$ for some $t \in T_n$. Let $t_{C-} = \inf\{t \in C\}$ denote the lower left corner of C. We now define the "discrete" approximation ρ_n of ρ by

$$\rho_n = \bigcup_{C \in \mathcal{C}_n : t_{C-} \in \rho} C.$$

It is straightforward that $\rho_n \subseteq R$ is an $\mathcal{F}$-stopping set, that (ρ_n) is decreasing in n and $\rho = \bigcap_n \rho_n$. Boundedness and uniform integrability ensure that $E[\int_\rho U_t \, dt] = \lim_n E[\int_{\rho_n} U_t \, dt]$ and $E[\int_\rho V_t \, dt] = \lim_n E[\int_{\rho_n} V_t \, dt]$. To complete the proof of the first statement in the theorem, observe that by boundedness of U,

$$E\left[\int_{\rho_n} U_t \, dt\right] = E\left[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} \int_C U_t \, dt\right]$$

$$= E\left[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} E\left[\int_C U_t \, dt \Big| \mathcal{F}_{t_{C-}}\right]\right]$$

$$= E\left[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} E\left[\int_C E[U_t | \mathcal{F}_t] \, dt \Big| \mathcal{F}_{t_{C-}}\right]\right]$$

$$= E\left[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} E\left[\int_C V_t \, dt \Big| \mathcal{F}_{t_{C-}}\right]\right]$$

$$= E\left[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} \int_C V_t \, dt\right] = E\left[\int_{\rho_n} V_t \, dt\right].$$

The third equality above follows by Fubini and the assumption that V has regular sample paths, and since $t \in C \Rightarrow t \geq t_{C-}$. [The assumption that V has a version with regular sample paths ensures that V is jointly $\mathcal{F} \times \mathcal{B}(\mathbf{R}_+^2)$-measurable, where $\mathcal{B}(\mathbf{R}_+^2)$ denotes the Borel sets in R.]

Next, assume that V is monotone. To prove optimality of $\hat{\rho}$, let $\rho \subseteq R$ be any other stopping set in R. We have

$$E\left[\int_{\hat{\rho}} U_t \, dt - \int_\rho U_t \, dt\right] = E\left[\int_{\hat{\rho}\backslash\rho} V_t \, dt - \int_{\rho\backslash\hat{\rho}} V_t \, dt\right] \geq 0,$$

since $V > 0$ on $\hat{\rho}^o$ (the interior of $\hat{\rho}$) and $V \leq 0$ on $\hat{\rho}^c$. $\square$

3.2. *Smooth semimartingale representation of the gain function.* We begin this section with an analysis of the single line process L_t: L is a non-explosive point process whose jump points are all incomparable. Single line processes and their compensators were discussed in [7], to which the reader may refer for more detail. Heuristically, if $\mathcal{F}_s^L = \sigma(L_u : u \leq s)$, then a process Λ will be an $\mathcal{F}^L$-compensator of L if

$$\Lambda((s_1, s_2), (s_1 + ds_1, s_2 + ds_2)])$$

$$\approx I_{\{L_s = 0\}} E[L((s_1, s_2), (s_1 + ds_1, s_2 + ds_2)]) | L_s = 0],$$

since L cannot have any jump points in $((s_1, s_2), (s_1 + ds_1, s_2 + ds_2)])$ if $L_s > 0$ and $\{L_s = 0\}$ is an atom of $\mathcal{F}_s^L$. Define the (deterministic) increasing

function $\Lambda_t^{(s)} := E[L(s,t)|L_s = 0]$, for $t \geq s$, and when it exists, let

$$\lambda_s = \lim_{t_1 \downarrow s_1, t_2 \downarrow s_2} \frac{\Lambda_t^{(s)}}{(t_1 - s_1)(t_2 - s_2)}.$$

In particular, if λ_s exists for every $s \in T$ and is Lebesgue measurable, then

$$(13) \qquad \Lambda_t = \int_{A_t} \lambda_u I_{\{L_u = 0\}} \, du.$$

In what follows (and as will be seen to be the case in our examples), we will assume that a representation of the form (13) exists for the compensator Λ of L, and we will refer to the deterministic function λ as the *weak hazard function* of L. It will always be assumed that λ is continuous.

To better understand the weak hazard, we observe that if $E[L]$ of L is absolutely continuous with respect to Lebesgue measure with Radon–Nikodym derivative $\tilde{\lambda}$, then for every $u \in T$ with $P(L_u = 0) > 0$, $\lambda_u = \tilde{\lambda}_u/P(L_u = 0)$. To see this, simply observe that for each $t \in T$,

$$(14) \qquad \int_{A_t} \tilde{\lambda}_u \, du = E[L_t] = E[\Lambda_t] = \int_{A_t} \lambda_u P(L_u = 0) \, du.$$

Returning to the gain function (6), let M denote the weak martingale $L - \Lambda$ and recall that $X_u = I_{\{L_u = 0\}}$. For any lower layer $B \subseteq T$,

$$
\begin{aligned}
Z(B) &= k_0 + \int_B (-c_1 + (c_0 + c_1)X_u) \, du + k_1 L(B) \\
(15) & \\
&= k_0 + \int_B (-c_1 + (c_0 + c_1 + k_1 \lambda_u)X_u) \, du + k_1 M(B).
\end{aligned}
$$

We note that X is outer-continuous with inner limits by definition and that λ is assumed to be continuous, and so we now have an $\mathcal{F}^L$-SSM representation of the gain function: $Z = (U, M)$, where $U_t := -c_1 + (c_0 + c_1 + k_1 \lambda_t)X_t$.

COMMENT 3.11. As a simple illustration, if the point process L and the set $\xi = \{t : L_t > 0\}$ are unobservable and no other information is available (i.e., N is not observed and $\mathcal{F}_t = \{\varnothing, \Omega\} \, \forall t \in T$), then for $R = [0, r]^2$, we are looking for a deterministic set $\hat{B} \subseteq R$ that maximizes

$$
\begin{aligned}
E[Z(B)] &= E\left[k_0 + \int_B (-c_1 + (c_0 + c_1 + k_1 \lambda_u)X_u) \, du + k_1 M(B) \right] \\
(16) & \\
&= k_0 + \int_B (-c_1 + (c_0 + c_1 + k_1 \lambda_u)P(L_u = 0)) \, du.
\end{aligned}
$$

Letting $V_u = [-c_1 + (c_0 + c_1 + k_1 \lambda_u)P(L_u = 0)]$, it is easily seen that V is deterministic and an optimal solution for the detection problem exists if V

is monotone, in which case

(17)
$$\hat{B} = \{t \in R : V_u > 0 \, \forall u \ll t\}$$
$$= \left\{ t \in R : P(L_u = 0) > \frac{c_1}{(c_0 + c_1 + k_1 \lambda_u)} \forall u \ll t \right\}.$$

The optimal estimate of $\xi \cap R$ is

$$\hat{\xi}_R = \left\{ t \in R : P(L_t = 0) \le \frac{c_1}{(c_0 + c_1 + k_1 \lambda_t)} \right\}.$$

EXAMPLE 3.12 (The single jump process). Suppose $L_t = I_{\{Y \in A_t\}}$, where Y is a T-valued random variable with distribution F and continuous density f. Then we have $\lambda_u = \frac{f_u}{1 - F_u}$. To verify that the representation (13) is satisfied with this definition, observe first that $E[L(s,t]|\mathcal{F}_s^L] = \frac{F(s,t]}{1 - F_s} I_{\{L_s = 0\}}$. Next,

$$E\left[\int_{(s,t]} \frac{f_u}{1 - F_u} I_{\{L_u = 0\}} \, du \Big| \mathcal{F}_s\right] = \int_{(s,t]} \frac{f_u}{1 - F_u} P(L_u = 0 | \mathcal{F}_s) \, du$$
$$= \int_{(s,t]} \frac{f_u}{1 - F_u} \cdot \frac{1 - F_u}{1 - F_s} I_{\{L_s = 0\}} \, du$$
$$= \frac{F(s,t]}{1 - F_s} I_{\{L_s = 0\}}.$$

Thus, the increasing process $\Lambda_t = \int_{A_t} \lambda_s I(L_s = 0) \, ds$ is a $\mathcal{F}^L$-compensator for L, verifying (13).

It should be noted that in the literature on bivariate survival analysis, the definition of the hazard function is $\frac{f_u}{S_u}$ where $S_u = P(Y \ge u)$. For this reason, we refer to our hazard $\lambda = \frac{f}{1 - F}$ as the "weak" hazard.

Returning to Comment 3.11, when no information is available, V is decreasing and (17) defines an optimal deterministic solution if f is decreasing in each parameter.

EXAMPLE 3.13 (First line of a Poisson process). Consider a homogeneous Poisson process J on T with rate γ. If Δ_J denotes the set of jump points of J, then the *first line* of J is the single line point process L with (incomparable) jump points

$$\Delta_L = \min(\Delta_J) = \{\tau \in \Delta_J : \tau' \not\le \tau \, \forall \tau' \in \Delta_J \text{ such that } \tau' \ne \tau\}.$$

In this case, $\xi = \{t : L_t > 0\} = \{t : J_t > 0\}$. As is shown in [7], the weak hazard of L is γ.

Considering the situation in Comment 3.11 when no information is available, we have $V_u = -c_1 + (c_0 + c_1 + k_1 \gamma) e^{-\gamma u_1 u_2}$, which is clearly monotone.

In this case, the optimal solution given in (17) becomes

$$\hat{B} = \left\{ t \in R : e^{-\gamma t_1 t_2} \geq \frac{c_1}{(c_0 + c_1 + k_1 \gamma)} \right\}$$

$$= \left\{ t \in R : t_1 t_2 \leq \frac{\ln(c_0 + c_1 + k_1 \gamma) - \ln(c_1)}{\gamma} \right\}$$

and

$$\hat{\xi}_R = \left\{ t \in R : t_1 t_2 \geq \frac{\ln(c_0 + c_1 + k_1 \gamma) - \ln(c_1)}{\gamma} \right\}.$$

We are now ready to return to the sequential detection problem, and consider the case in which the process N is observed (recall that N is a Poisson process with rate μ_0 on ξ^c and μ_1 on ξ). We denote the full filtration $\mathcal{F}^{L,N} = \{\mathcal{F}_t^{L,N} : t \in T\}$, where $\mathcal{F}_t^{L,N} = \sigma\{L_s, N_s, s \leq t\}$, and (as before) the subfiltrations $\mathcal{F}^L = \{\mathcal{F}_t^L : t \in T\}$ and $\mathcal{F}^N = \{\mathcal{F}_t^N : t \in T\}$ where $\mathcal{F}_t^L = \sigma\{L_s : s \leq t\}$ and $\mathcal{F}_t^N = \sigma\{N_s : s \leq t\}$. Although we defined the weak hazard of L with respect to $\mathcal{F}^L$, it is easy to see that given the full filtration $\mathcal{F}^{L,N}$, $L - \Lambda$ is still a weak $\mathcal{F}^{L,N}$-martingale. This follows because on $\{L_s = 0\} = \{s \in \xi^c\}$, N is a Poisson process with rate μ_0 on A_s and so $N|_{A_s}$ (N restricted to A_s) adds no additional information about the behavior of L_t for $t > s$. Formally, we have

$$E[L(s,t)|\mathcal{F}_s^{L,N}] = I_{\{L_s=0\}} \Lambda_t^{(s)} = E[L(s,t)|\mathcal{F}_s^L].$$

Therefore, from this discussion we have the following lemma and we are ready to proceed with finding an optimal solution to the sequential detection problem.

LEMMA 3.14.   *Equation (15) defines an $\mathcal{F}^{L,N}$-SSM representation of the gain function Z: $Z = (U, M)$ where $U_t := -c_1 + (c_0 + c_1 + k_1 \lambda_t) X_t$.*

4. Optimal solution to the sequential detection problem.   We consider the $\mathcal{F}^{L,N}$-SSM representation of the gain function (15):

$$Z(B) = k_0 + \int_B (-c_1 + (c_0 + c_1 + k_1 \lambda_u) X_u) \, du + k_1 M(B).$$

In order to find sufficient conditions for the existence of an optimal solution in the sequential case, we will be appealing to Lemma 3.10, with $\mathcal{G} = \mathcal{F}^{L,N}$, $\mathcal{F} = \mathcal{F}^N$ and $U_t = -c_1 + (c_0 + c_1 + k_1 \lambda_t) X_t$. In order to find $V_t = E[U_t|\mathcal{F}_t^N]$, it is enough to determine

$$E[X_t|\mathcal{F}_t^N] = P(L_t = 0|\mathcal{F}_t^N).$$

As in [5], we use a Bayesian argument. The first step is to determine the conditional likelihood $\ell_{N|L}(t)$ of $N|_{A_t}$ given L and use this to find the likelihood $\ell_N(t)$ of $N|_{A_t}$. Next we find the conditional likelihood $\ell_{N|L_t=0}(t)$ of $N|_{A_t}$ on the set $\{L_t = 0\}$. Finally, we have

(18)
$$E[X_t|\mathcal{F}_t^N] = P(L_t = 0|\mathcal{F}_t^N)$$
$$= \frac{\ell_{N|L_t=0}(t) \times P(L_t = 0)}{\ell_N(t)}.$$

When computing the likelihood $\ell_{N|L}$, in fact it is equivalent to condition on the random upper layer $\xi = \{u : L_u > 0\}$. To see this, let $(\mathcal{U}, d_H)$ denote the collection of closed upper layers in T endowed with the Hausdorff metric. It is shown in [8] that $(\mathcal{U}, d_H)$ is a complete separable metric space and that ξ can be regarded as the unique jump point in a single jump process $\tilde{L}$ on $\mathcal{U}$; in addition, L determines and is determined by $\tilde{L}$. In particular, $L_t > 0 \Leftrightarrow t \in \xi \Leftrightarrow E_t \subseteq \xi$, where $E_t = \{s \in T : s \geq t\}$. Let μ_ξ denote the measure induced by ξ on $\mathcal{U}$.

Given L, or equivalently ξ, N is a Poisson process with rate μ_0 on ξ^c and μ_1 on ξ. Using the well-known likelihood for the Poisson process (cf. [3], page 22), we have

(19)
$$\ell_{N|L}(t) = \ell_{N|\xi}(t)$$
$$= e^{-\mu_0|A_t\setminus\xi|}\mu_0^{N(A_t\setminus\xi)}e^{-\mu_1|A_t\cap\xi|}\mu_1^{N(A_t\cap\xi)}$$
$$= e^{-\mu_0|A_t|}\mu_0^{N_t}e^{-(\mu_1-\mu_0)|A_t\cap\xi|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap\xi)}.$$

By considering separately the events $\{L_t = 0\} = \{t \notin \xi\} = \{E_t \not\subseteq \xi\} = \{A_t \cap \xi = \varnothing\}$ and $\{L_t > 0\} = \{t \in \xi\} = \{E_t \subseteq \xi\}$, we use (19) obtain

(20)
$$\ell_N(t) = P(L_t = 0)e^{-\mu_0|A_t|}\mu_0^{N_t}$$
$$+ e^{-\mu_0|A_t|}\mu_0^{N_t}\int_{\{D\in\mathcal{U}:E_t\subseteq D\}}e^{-(\mu_1-\mu_0)|A_t\cap D|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap D)}d\mu_\xi(D)$$
$$= e^{-\mu_0|A_t|}\mu_0^{N_t}[P(L_t = 0) + e^{-(\mu_1-\mu_0)|A_t|}Q_t],$$

where

(21)
$$Q_t = \int_{\{D\in\mathcal{U}:E_t\subseteq D\}}e^{(\mu_1-\mu_0)|A_t\setminus D|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap D)}d\mu_\xi(D).$$

Before continuing, we observe that since $\mu_1 > \mu_0$, Q is increasing in each parameter separately because each term in the integrand is increasing in each component for D fixed, and the range of integration is increasing since the set E_t decreases with each component.

Next, if $L_t = 0$, then $N|_{A_t}$ is Poisson with rate μ_0, and so

$$\ell_{N|L_t=0} = e^{-\mu_0|A_t|} \mu_0^{N_t}. \tag{22}$$

Substituting (20) and (22) in (18), we obtain

$$
\begin{aligned}
E[X_t | \mathcal{F}_t^N] &= \frac{e^{-\mu_0|A_t|} \mu_0^{N_t} P(L_t = 0)}{e^{-\mu_0|A_t|} \mu_0^{N_t} [P(L_t = 0) + e^{-(\mu_1 - \mu_0)|A_t|} Q_t]} \\
&= \frac{1}{1 + q_t Q_t},
\end{aligned}
\tag{23}
$$

where $q_t = \frac{e^{-(\mu_1 - \mu_0)|A_t|}}{P(L_t = 0)}$. [If $P(L_t = 0) = 0$, (23) remains formally valid since $E[X_t | \mathcal{F}_t^N] = 0$ and $q_t = \infty$.] We are now ready to state our main result:

THEOREM 4.1.   *Let L be a single line process with continuous weak hazard λ, and define the function q by*

$$q_t = \frac{e^{-(\mu_1 - \mu_0) t_1 t_2}}{P(L_t = 0)} \qquad for \ t = (t_1, t_2) \in \mathbf{R}_+^2.$$

An optimal solution to the sequential detection problem on $R = [0, r]^2$ exists if λ and q are decreasing and increasing, respectively, in each component on R. In this case V is monotone on R, and the optimal solution is given by (12):

$$\hat{\rho} = \{t \in R : V_s > 0 \, \forall s \ll t\},$$

where

$$V_t = -c_1 + (c_0 + c_1 + k_1 \lambda_t) \frac{1}{1 + q_t Q_t}.$$

PROOF.   We review our results so far. We have the $\mathcal{F}^{L,N}$-SSM representation of the gain function $Z(B) = k_0 + \int_B U_t \, dt + k_1 M(B)$, where $U_t = -c_1 + (c_0 + c_1 + k_1 \lambda_t) X_t$. U is bounded since λ is decreasing in each component and X is an indicator function. By the argument immediately preceding the theorem, we have that

$$V_t = E[U_t | \mathcal{F}_t^N] = -c_1 + (c_0 + c_1 + k_1 \lambda_t) \frac{1}{1 + q_t Q_t}. \tag{24}$$

To see that V has a version which is outer-continuous with inner limits (o.c.i.l.), recall that λ is assumed to be continuous and observe that q is o.c.i.l. by definition. Turning next to Q, we see that the integrand in (21) is o.c.i.l. and increasing in each component in t, as is

$$\mu_\xi(\{D \in \mathcal{U} : E_t \subseteq D\}) = P(L_t > 0).$$

Therefore, it follows that Q, and hence V are o.c.i.l. Therefore, Lemmas 3.6 and 3.10 imply that for any $\mathcal{F}^N$-stopping set $\rho \subseteq R$,

$$(25) \qquad E[Z(\rho)] = k_0 + E\left[\int_\rho U_t \, dt\right] = k_0 + E\left[\int_\rho V_t \, dt\right].$$

To show that an optimal solution $\hat{\rho}$ exists [as in (12)], it is sufficient to show that V is monotone (again, by Lemma 3.10). Since we have already seen that Q is increasing in each component on R, the assumption that λ and q are decreasing and increasing, respectively, in each component imply that V is monotone on R. This completes the proof. $\square$

COMMENT 4.2. It has been pointed out by an anonymous referee that the case $\mu_0 = 0$ relates to a so-called support estimation problem. In this case, the random set ξ denotes the support of a Poisson process with rate μ_1. The gain function can be defined exactly as before, and the analysis proceeds in very much the same way. Now we know that $N_t > 0 \Rightarrow t \in \xi \Rightarrow L_t > 0$, and equation (18) becomes

$$
\begin{aligned}
(26) \qquad E[X_t | \mathcal{F}_t^N] &= P(L_t = 0 | \mathcal{F}_t^N) \\
&= P(L_t = 0 | N_t = 0) I(N_t = 0) \\
&= \frac{P(L_t = 0)}{P(N_t = 0)} I(N_t = 0).
\end{aligned}
$$

Continuing with the same sort of arguments used previously, if $\mu_0 = 0$, equation (23) becomes

$$(27) \qquad E[X_t | \mathcal{F}_t^N] = \frac{1}{1 + q_t \dot{Q}_t} I(N_t = 0),$$

where q_t is defined as before with $\mu_0 = 0$, and

$$(28) \qquad \dot{Q}_t = \int_{\{D \in \mathcal{U} : E_t \subseteq D\}} e^{\mu_1 |A_t \setminus D|} \, d\mu_\xi(D).$$

It is easy now to see that the statement of Theorem 4.1 is still valid in this case, with V replaced by $\dot{V}$, where

$$\dot{V}_t = -c_1 + (c_0 + c_1 + k_1 \lambda_t) \frac{1}{1 + q_t \dot{Q}_t} I(N_t = 0).$$

5. Examples. In this section, we apply Theorem 4.1 to our two examples. We will see that in some sense they are are both analogous to the univariate model of [5], in which the change-point is exponentially distributed. There are two natural generalizations in $\mathbf{R}_+^2$: first, L is the single jump process in which the components of the jump are independent univariate exponential

random variables, and second, L is the first line of a Poisson process, noting that an exponential random variable can be regarded as the "first line" of a Poisson process on $\mathbf{R}_+$. Although at first glance the single jump process looks more straightforward, we shall see that in fact the analysis is far more complex than in the case of the first line of a Poisson process.

EXAMPLE 5.1 (The single jump process). Referring to Example 3.12, we have $\lambda_t = \frac{f_t}{1-F_t}$ and $q_t = \frac{e^{-(\mu_1-\mu_0)t_1t_2}}{1-F_t}$. Here we will consider the case in which the components (Y_1, Y_2) of the jump Y are independent identically distributed exponential random variables with parameter γ. In this case,

$$\lambda_t = \frac{f_t}{1-F_t} = \frac{\gamma e^{-\gamma t_1}\gamma e^{-\gamma t_2}}{1-(1-e^{-\gamma t_1})(1-e^{-\gamma t_2})} = \frac{\gamma^2}{e^{\gamma t_1} + e^{\gamma t_2} - 1},$$

and is decreasing in each component. Next, we consider q_t:

$$q_t = \frac{e^{-(\mu_1-\mu_0)t_1t_2}}{1-(1-e^{-\gamma t_1})(1-e^{-\gamma t_2})} = \frac{e^{-(\mu_1-\mu_0)t_1t_2}}{e^{-\gamma t_1} + e^{-\gamma t_2} - e^{-\gamma(t_1+t_2)}}.$$

To find sufficient conditions to ensure that q is increasing in t_1 and t_2 on some set $R = [0, r]^2$, we will assume that $\gamma > \mu_1 - \mu_0$ and to simplify the discussion (without loss of generality, by suitably rescaling the time parameters if necessary) that $\mu_1 - \mu_0 = 1$. Now rewrite $q_t = 1/g_t$ where

$$g_t = g_{(t_1,t_2)} = e^{-t_1(\gamma-t_2)}(1 - e^{-\gamma t_2}) + e^{t_1 t_2}e^{-\gamma t_2}.$$

We will show that if $r \leq \frac{\ln\gamma}{\gamma}$, then $\frac{d}{dt_1}g_{(t_1,t_2)} \leq 0$ for $(t_1, t_2) \in R = [0, r]^2$. By symmetry, the same is true for $\frac{d}{dt_2}g_{(t_1,t_2)}$ for $t \in R$. Therefore, g is decreasing and $q = 1/g$ is increasing in each component on R, and an optimal solution exists for the sequential detection model.

To complete the example, we observe that

$$\frac{d}{dt_1}g_{(t_1,t_2)} = e^{-t_1(\gamma-t_2)}(-(\gamma - t_2)(1 - e^{-\gamma t_2}) + t_2 e^{\gamma(t_1-t_2)}) \leq 0$$

if and only if

$$(\gamma - t_2)(1 - e^{-\gamma t_2}) \geq t_2 e^{\gamma(t_1-t_2)}$$

or equivalently,

(29) $$e^{\gamma t_1} \leq \frac{\gamma - t_2}{t_2}(e^{\gamma t_2} - 1).$$

The left-hand side of (29) is bounded above by γ since $t_1 \leq r \leq \frac{\ln\gamma}{\gamma}$. The right-hand side of (29) is bounded below by γ since $t_2 \leq r \leq \frac{\ln\gamma}{\gamma} \leq \gamma - 1$ when $\gamma \geq 1$, and so $\frac{\gamma-t_2}{t_2}(e^{\gamma t_2} - 1) = \frac{e^{\gamma t_2}-1}{t_2}(\gamma - t_2) \geq \gamma(\gamma - t_2) \geq \gamma$. Therefore, it is sufficient that $t_1, t_2 \leq r \leq \frac{\ln\gamma}{\gamma}$.

EXAMPLE 5.2 (First line of a Poisson process). From the discussion in Example 3.13, if L is the first line of a Poisson process with rate γ, then $\lambda \equiv \gamma$, and so trivially is decreasing in each component. We have $q_t = \frac{e^{-(\mu_1-\mu_0)t_1 t_2}}{e^{-\gamma t_1 t_2}} = e^{(\gamma-(\mu_1-\mu_0))t_1 t_2}$, which is increasing in each component if $\gamma \geq \mu_1 - \mu_0$. Therefore, an optimal solution to the sequential detection problem exists on any bounded set $R = [0,r]^2$ if $\gamma \geq \mu_1 - \mu_0$, and is defined by (12). In fact, this is exactly the same as the sufficient condition for the univariate detection problem proven in [5] and [9].

6. Conclusion. As indicated in the Introduction, the sequential detection model considered here is only one of many scenarios that should be analyzed in the general context of the "optimal set-detection problem." Indeed, the model can be extended in many possible ways.

- The information structure: In addition to the sequential information model, Herberts and Jensen [5] consider what they call the "ex-post" analysis. This would correspond to observing N on all of R, and then trying to optimize the expectation of the valuation function. (Formally, this corresponds to $\mathcal{F}_t = \mathcal{F}_{(r,r)} \forall t \in R$.) Several variants or combinations of the ex-post and sequential schemes can be studied.
- The underlying space: We worked here on a bounded subset of R_+^2. It would be of interest to consider change-point problems on higher-dimensional Euclidean spaces or more general partially ordered sets as in [8].
- The change mechanism: Here the change occurs at either a single random point or at the first line of a more general point process. The example involving the first line of a Poisson process turned out to be (perhaps surprisingly) the more natural analog of the one-dimensional exponential change-point problem. Consideration should be given to more general single jump and first line processes, as well as to more general random sets (not necessarily upper layers). For example, the case in which L is the first line of an inhomogeneous Poisson process with intensity $\gamma(\cdot)$ is considered in [2] where it is proven that an optimal solution exists if $\inf_{u \in R} \gamma(u) \geq \mu_1 - \mu_0$.
- The observed process: On $\mathbf{R}_+$, the process subject to the change can be a more general process, such as the Brownian motion process (cf. [1]). Here too, we can consider more general processes such as the set-indexed Brownian motion (cf. [6]).
- The parameters: In our analysis, it is implicitly assumed that the parameters of the various processes are all known. How does one approach the problem when one or more parameters must be estimated?
- The gain function: Different valuation functions can be chosen, thereby changing the notion of optimality. For example, with a change generated

by a single jump at Y, instead of two cost parameters c_0 and c_1 associated respectively with E_Y^c and E_Y, we could have different costs in each of the four quadrants defined by Y. Another variation considered in [2] is to replace L_t in (5) with $I(L_t > 0)$. Although this does not change the valuation when the change is generated by a single jump, the analysis becomes more complex when L is the first line of a Poisson process.

- Number of changes: Here we deal with only one change-set. However, we can imagine that several changes occur on a decreasing sequence of random upper layers, for example. This would correspond to multiple change points on $\mathbf{R}_+$.

REFERENCES

[1] CARLSTEIN, E., MÜLLER, H.-G.and SIEGMUND, D. (1994). *Change-point Problems.* IMS, Hayward, CA. MR1477909

[2] COLLINGWOOD, J. (2009). Application of martingale methods to a change set problem. M. Sc. thesis, Univ. Ottawa.

[3] DALEY, D. J. and VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*, 2nd ed. Springer, New York. MR1950431

[4] HAHUBIA, T. and MNATSAKANOV, R. (1996). On the mode-change problem for random measures. *Georgian Math. J.* **3** 343–362. MR1397816

[5] HERBERTS, T. and JENSEN, U. (2004). Optimal detection of a change point in a Poisson process for different observation schemes. *Scand. J. Statist.* **31** 347–366. MR2087830

[6] IVANOFF, B. G. and MERZBACH, E. (2000). *Set-indexed Martingales. Monographs on Statistics and Applied Probability* **85**. Chapman & Hall/CRC, Boca Raton, FL. MR1733295

[7] IVANOFF, B. G. and MERZBACH, E. (2006). What is a multi-parameter renewal process? *Stochastics* **78** 411–441. MR2281679

[8] IVANOFF, B. G., MERZBACH, E. and PLANTE, M. (2007). A compensator characterization of point processes on topological lattices. *Electron. J. Probab.* **12** 47–74 (electronic). MR2280258

[9] JENSEN, U. and HSU, G. (2002). Optimal stopping by means of point process observations with applications in reliability. *Calcutta Statist. Assoc. Bull.* **52** 117–141.

[10] KHMALADZE, E., MNATSAKANOV, R. and TORONJADZE, N. (2006). The change-set problem for Vapnik–Červonenkis classes. *Math. Methods Statist.* **15** 224–231. MR2256476

[11] KHMALADZE, E., MNATSAKANOV, R. and TORONJADZE, N. (2006). The change set problem and local covering numbers. *Math. Methods Statist.* **15** 289–308. MR2278291

[12] MERZBACH, E. and ZAKAI, M. (1980). Predictable and dual predictable projections of two-parameter stochastic processes. *Z. Wahrsch. Verw. Gebiete* **53** 263–269. MR586018

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

DEPARTMENT OF MATHEMATICS & STATISTICS   DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OTTAWA   BAR-ILAN UNIVERSITY
585 KING EDWARD   52900 RAMAT-GAN
OTTAWA, ONTARIO K1N 6N5   ISRAEL
CANADA   E-MAIL: merzbach@macs.biu.ac.il
E-MAIL: givanoff@uottawa.ca