**MATHEMATICS OF OPERATIONS RESEARCH**
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0364-765X | EISSN 1526-5471 | 00 | 0000 | 0001

**INFORMS**
DOI 10.1287/xxxx.0000.0000


# Sequential sensor installation for Wiener disorder detection

Savas Dayanik
Departments of Industrial Engineering and Mathematics, 06800 Bilkent, Ankara, Turkey, sdayanik@bilkent.edu.tr, http://sdayanik.bilkent.edu.tr

Semih O. Sezer
Sabancı University, Faculty of Engineering and Natural Sciences, 34956 Tuzla, Istanbul, Turkey, sezer@sabanciuniv.edu, http://people.sabanciuniv.edu/sezer

We consider a centralized multi-sensor online quickest disorder detection problem where the observation from each sensor is a Wiener process gaining a constant drift at a common unobservable disorder time. The objective is to detect the disorder time as quickly as possible with small probability of false alarms. Unlike the earlier work on multi-sensor change detection problems, we assume that the observer can apply a sequential sensor installation policy. At any time before a disorder alarm is raised, the observer can install new sensors to collect additional signals. The sensors are statistically identical, and there is a fixed installation cost per sensor. We propose a Bayesian formulation of the problem. We identify an optimal policy consisting of a sequential sensor installation strategy and an alarm time, which minimize a linear Bayes risk of detection delay, false alarm, and new sensor installations. We also provide a numerical algorithm and illustrate it on examples. Our numerical examples show that significant reduction in the Bayes risk can be attained compared to the case where we apply a static sensor policy only. In some examples, the optimal sequential sensor installation policy starts with 30% less number of sensors than the optimal static sensor installation policy and the total percentage savings reach to 12%.

*Key words*: Optimal multiple stopping, multi-sensor sequential change detection, Wiener disorder problem
*MSC2000 subject classification*: Primary 62L10; Secondary 62L15, 62C10, 60G40
*OR/MS subject classification*: Primary: reliability: quality control; secondary: decision analysis: sequential; dynamic programming/optimal control: Markov

---

**1. Introduction.** Suppose that we simultaneously collect observations from $\ell$-many identical sensors, each providing continuous signals of the form

$$X_t^{(i)} := W_t^{(i)} + \mu(t-\Theta)^+, \qquad t \ge 0\,,\ i = 1, \ldots, \ell, \tag{1.1}$$

where the $W^{(i)}$'s are independent Wiener processes, $\mu \neq 0$ is a given constant, and $\Theta$ is an *unobservable* random variable with the prior distribution

$$\mathbb{P}\{\Theta = 0\} = \pi \quad \text{and} \quad \mathbb{P}\{\Theta > t\} = (1-\pi)e^{-\lambda t}, \qquad t \ge 0 \tag{1.2}$$

for known $\pi \in [0,1)$ and $\lambda > 0$. The random variable $\Theta$ is the time of the onset of a new regime, and the objective is to detect this disorder as quickly as possible after it happens, based solely on our observations. We assume that at any time before we declare a change, we can install new sensor(s) in order to enrich our observations at a fixed cost of $b$ per sensor. Each additional sensor is identical to those already in place and contributes another process of the form in (1.1).

Let $\delta = (\tau_1, \tau_2, \ldots)$ denote a sensor installation policy where $\tau_i$ is the installation time of the $i$'th additional sensor, and let $\delta(t) := \sum_{i \geq 1} 1_{\{\tau_i \leq t\}}$ be the number of additional sensors deployed by time $t$ for $t \geq 0$. If we raise the disorder alarm at some time $\tau$, then the cost of additional sensors is $b\delta(\tau)$, and together with the costs of detection delay and false alarms we incur total expected cost of

$$R^{\delta,\tau}(\ell,\pi) := \mathbb{E}\left[1_{\{\tau<\Theta\}} + c\,(\tau-\Theta)^+ + b\,\delta(\tau)\right], \tag{1.3}$$

where $c > 0$ is the unit cost of delay. Here, without loss of generality we assume that the cost of false alarm event is one.

In this setup, our objective is to find a sequential installation policy $\delta$ and an alarm time $\tau$ that will minimize the Bayes risk in (1.3). Clearly, the installation time $\tau_i$ of the $i$th sensor, $i \geq 1$ must be a stopping time of the observations generated by the already installed $\ell + i - 1$ sensors, and the alarm time $\tau$ must be a stopping time of the observation filtration associated with the given policy $\delta$.

As the Bayes risk in (1.3) is minimized, our formulation always allows the option to instantaneously install all of the sensors at any time, especially, at the very beginning. It is, however, natural to ask if spreading the sensor installations over time may further reduce the Bayes risk. We later report in Section 6 on several numerical examples that optimal sequential sensor installation policy starts with up to 30% less number of sensors than the optimal static installation policy and total percentage savings can reach up to 12%. Those concrete savings in numerical examples show the importance of the joint Bayesian sequential change detection and sequential sensor installation problem.

The minimization of the Bayes risk formulation in (1.3) tries to find the best strike between the total expected monetary investment into sensors and expected losses due to untimely disorder alarms. The minimization of the same Bayes risk is also the most natural intermediate step to take in order to minimize the expected total detection delay time

$$\mathbb{E}(\tau - \Theta)^+$$

over alarm time $\tau$ and sensor installation policy $\delta$, subject to strict false alarm and budget constraints

$$\mathbb{P}\{\tau < \Theta\} \leq \alpha \qquad \text{and} \qquad \mathbb{E}[b\delta(\tau)] \leq B,$$

respectively, for some fixed acceptable level of false alarm rate $0 < \alpha < 1$ and for some maximum total monetary budget $B$ available for purchasing and installing new sensors. This constrained optimization problem is naturally attacked by solving its Lagrange relaxation, which boils down to the minimization of the Bayes risk in (1.3) for appropriate choices of $c$ and $b$ values.

Various other forms of change detection problems have been studied extensively in the literature due to their important applications including, for example, the intrusion detection in computer networks and security systems, threat detection in national defense, fault detection in industrial processes, detecting a change in risk characteristics of financial instruments, detecting a change in the reliability of mechanical systems, detecting the onset of an epidemic in biomedical signal processing, and others. We refer the reader to the monographs of Basseville and Nikiforov [3], Peskir and Shiryaev [25], Poor and Hadjiliadis [26], and the references therein for those applications and also for an extensive review of the earlier work on sequential change detection.

Considering the applications in environment monitoring and surveillance there has been a growing interest in the *multi-sensor* change detection problems; see, for example, Crow and Schwartz

[11], Blum et al. [7, Section V-D], Veeravalli [37], Tartakovsky and Veeravalli [34, 35, 36], Chamberland and Veeravalli [10], Mei [23], Moustakides [24], Tartakovsky and Polunchenko [33], Raghavan and Veeravalli [27], who consider both *centralized* and *decentralized* versions. In the centralized settings, signals from the sensors are perfectly transmitted to a *fusion center*, where all the information is processed, and a detection decision is made accordingly. In decentralized settings, on the other hand, sensors send quantized versions of their observations to the fusion center, and the detection decision is based on that partial information. Such a formulation is more suitable in applications where sensors are geographically dispersed and there are constraints on the communication (like bandwidth restrictions). In those problems, quantization schemes are also part of the decisions to be made by the observers.

In the above-mentioned works on multi-sensor problems, it is commonly assumed that the total number of sensors are fixed in advance. The sensors are already in place at $t = 0$, and the detection decision is based on the signals received from them only. However, compared to such a static strategy, applying a sequential sensor installation policy starting with $\ell = 0$ sensor may significantly improve the effectiveness of the disorder detection decisions because the observer can install additional sensors if and when additional information proves to be useful. Table 1 in Section 6 displays several numerical examples for which the percentage savings of optimal sequential sensor installation policies over optimal static sensor installation policies are higher than 11%.

To the best of our knowledge, the combined problem of sequential sensor installation and detection has not been addressed in the literature beforehand. In the current paper, we formulate this problem in a centralized Bayesian setting under the assumption that the observations consist of Brownian motions. The formulation with Brownian motions can be useful when the observations collected from different sensors are in the form of continuously vibrating/oscillating signals whose quadratic variations depend linearly on time. In the decentralized version of the problem, which we do not address here, the method developed here clearly does not apply. For instance, when a finite-alphabet quantization scheme for the likelihood ratio process is used at each sensor as in, for example, Veeravalli [37], the observations at the fusion center appear as piecewise-constant processes with jumps at random times. This requires different techniques than those we developed here for the continuous Brownian observations.

The classical disorder detection problem for the Brownian motion in a Bayesian setting is originally introduced and solved by Shiryaev [30, 31]. Its multi-sensor extension is studied by Dayanik et al. [14]. In the meantime, Shiryaev's formulation is revisited by Beibel [6] for an exponential Bayes risk and by Gapeev and Peskir [16] under a finite-horizon constraint. Recently, Sezer [28] and Dayanik [12] considered the extensions of the infinite horizon problem with different observation structures; Sezer [28] assumes that the change time coincides with one of the arrival times of an observable Poisson process whereas Dayanik [12] assumes that the observations are taken at discrete points in time. In the non-Bayesian formulation, on the other hand, the optimality of CUSUM rule is established by Shiryaev [32] and Beibel [5], and the extension of the problem with multiple alternatives on the value of the drift after the change is studied by Hadjiliadis [18] and Hadjiliadis and Moustakides [19]. The latter also establishes the asymptotic optimality of a 2-CUSUM rule when the drift after the change can be one of two given values with opposite signs. The reader may refer to Hadjiliadis et al. [20], Zhang et al. [39], Zhang et al. [40] for recent models with multiple (static) sensors having non-identical change times, and the references therein for earlier related work in the non-Bayesian framework. For the asymptotic optimality of the Shiryaev's procedure in general continuous time models (not necessarily with Brownian observations) Baron and Tartakovsky [2] can be consulted.

The Bayesian formulation and the solution of the combined problem of sequential sensor installation and the Wiener disorder detection are the contributions of the current paper. Here, we solve the problem by transforming it into an optimal multiple stopping problem for the *conditional* *probability process* $\Pi$, which gives the posterior probability distribution of the disorder event. By means of a dynamic programming operator, the optimal multiple stopping problem is turned into a sequence of classical optimal stopping problems. Carmona and Dayanik [8] and Dayanik and Ludkovski [13] used similar approaches to solve multiple-exercise American-type financial options. However, unlike the other optimal multiple stopping problems, the problem we encountered in this study hosts a controlled stochastic process, $\Pi$, whose dynamics are not fixed, but change every time the control is applied; namely, every time a new sensor is installed. The structure of candidate solution was not obvious and guessed after studying the special structure of the family of optimal quickest detection and static installation problems indexed by the number of sensors in use. For each value of the number of sensors currently in use, the candidate solution satisfies a special differential equation in the no-action space and is constructed by continuously pasting the derivative of the candidate value function at the critical boundaries of candidate sequential installation policy. The construction was complicated by the fact that, as the number of sensors in use changes, the corresponding action spaces can be one- or two-sided and the corresponding no-action spaces can be nested or not nested. We have overcome all of those difficulties and given a precise algorithm to construct the candidates for the minimal Bayes risk and optimal sequential change detection and sensor installation policy, and then use their properties to run a verification lemma by applying Itô rule directly. With this outlined solution approach, the paper is significantly different from other papers in the literature on optimal multiple stopping.

For a given initial number of sensors $\ell$, we show that the optimal policy depends on the fixed installation cost $b$ of a new sensor. If $b$ is high, then we never add any new sensor and we simply apply the classical one-sided disorder detection policy of Shiryaev [30, 31]. If the cost $b$ is low, then there exist two threshold points $0 < B^*_\ell < A^*_\ell < 1$ such that it is optimal to continue and collect observations as long as the conditional probability process stays in the interval $(B^*_\ell, A^*_\ell)$. If $\Pi$ reaches $[A^*_\ell, 1]$ first, the problem terminates with disorder detection alarm. Otherwise, at the entrance of $\Pi$ into the interval $[0, B^*_\ell]$, it is optimal to install a new sensor and proceed optimally with $\ell+1$ sensors. Our numerical results indicate that the intervals $(B^*_\ell, A^*_\ell)$'s are not necessarily nested; see for example Figure 5 in Section 6. Therefore, it may sometimes be optimal to add more than one sensor at once.

In Section 2, a formal description of the problem is given. The conditional probability process and its dynamics are also provided. The lengthy derivation of those dynamics is deferred to the Appendix. In Section 3, we revisit the static multi-sensor disorder problem (with a fixed number of sensors already installed at the beginning) and review the structure of its solution. In Section 4, we introduce a dynamic programming operator, which itself turns out to be the value function of a special one-dimensional optimal stopping problem. By means of the dynamic-programming operator we construct the smallest Bayes risk, and describe an optimal policy in Section 5. In Section 6 we conclude with some numerical examples. Finally, in Section 7, we discuss some extensions.

**2. Problem statement.** Let $(\Omega, \mathcal{H}, \mathbb{P})$ be a probability space hosting independent Wiener processes $W^{(1)}, W^{(2)}, \ldots$, and an independent random variable $\Theta$ with zero-modified exponential distribution

$$\mathbb{P}\{\Theta = 0\} = \pi \quad \text{and} \quad \mathbb{P}\{\Theta > t\} = (1-\pi)e^{-\lambda t}, \qquad t \geq 0$$

for some $\pi \in [0,1)$ and $\lambda > 0$. In terms of those stochastic elements let us introduce $\Theta$-modulated processes

$$X^{(i)}_t := W^{(i)}_t + \mu(t-\Theta)^+ \qquad \text{for every } t \geq 0 \text{ and } i \geq 1,$$

and for some known constant $\mu \neq 0$.

Suppose that the time $\Theta$ is unobservable and we would like to detect it using the continuously observed signals $X_i$'s received only from the deployed sensors. Suppose that we start at time $t=0$ with $\ell$-many sensors already in place yielding observations from $X_t^{(1)},\dots,X_t^{(\ell)}$, $t\geq 0$. At times $\tau_1\leq\tau_2\leq\dots$ of our choice, we can deploy new sensors at a fixed cost of $b>0$ per sensor. The deployed sensors provide observations $(X_t^{(\ell+1)}-X_{t\wedge\tau_1}^{(\ell+1)})_{t\geq 0},(X_t^{(\ell+2)}-X_{t\wedge\tau_2}^{(\ell+2)})_{t\geq 0},\dots$, respectively. An admissible sensor installation policy consists of

$$\delta=(\tau_1,\tau_2,\dots) \qquad \text{with} \qquad \delta(t)=\sum_{i=1}^{\infty}1_{[\tau_i,\infty)}(t),\quad t\geq 0,$$

where the sensor installation times $\tau_i$'s are the stopping times of the appropriate filtrations constructed as follows. Let $\delta_0$ be the no-new-sensor policy with $\delta_0(t)\equiv 0$ for every $t\geq 0$. For every $n\geq 1$, let

$$\delta_n=(\tau_1,\tau_2,\dots,\tau_n), \qquad \text{with} \qquad \delta_n(t)=\sum_{i=1}^{n}1_{[\tau_i,\infty)}(t),\quad t\geq 0$$

be an installation policy of at most $n$-sensors. Let $\mathbb{F}^{\ell,\delta_0}\equiv(\mathcal{F}_t^{\ell,\delta_0})_{t\geq 0}$ be defined as

$$\mathcal{F}_t^{\ell,\delta_0}=\left\{\begin{matrix}\{\varnothing,\Omega\}, & \text{if } \ell=0\\ \sigma(X_s^{(1)},\dots,X_s^{(\ell)};0\leq s\leq t), & \text{if } \ell\geq 1\end{matrix}\right\}, \qquad t\geq 0;$$

namely, as the information generated by the sensors already in place. Then, inductively for every $n\geq 1$, $\tau_n$ is a stopping time of filtration $\mathbb{F}^{\ell,\delta_{n-1}}$ and we define the filtration $\mathbb{F}^{\ell,\delta_n}\equiv(\mathcal{F}_t^{\ell,\delta_n})_{t\geq 0}$ by

$$\mathcal{F}_t^{\ell,\delta_n}:=\mathcal{F}_t^{\ell,\delta_{n-1}}\bigvee\sigma\big(X_s^{(\ell+n)}-X_{\tau_n\wedge s}^{(\ell+n)},\,0\leq s\leq t\big) \quad \text{for every } t\geq 0.$$

Therefore, $\tau_n$ is also a stopping time of $\mathbb{F}^{\ell,\delta_n},\mathbb{F}^{\ell,\delta_{n+1}},\dots$ and is independent of $W^{(\ell+n)},W^{(\ell+n+1)},\dots$.

In plain words, $\mathbb{F}^{\ell,\delta_n}$ represents the information stream generated by installing at most $n$ new sensors at times $\tau_1,\dots,\tau_n$ according to the installation strategy $\delta_n$. We define the observation filtration, $\mathbb{F}^{\ell,\delta}\equiv(\mathcal{F}_t^{\ell,\delta})_{t\geq 0}$ associated with an admissible strategy $\delta$ as

$$\mathcal{F}_t^{\ell,\delta}:=\sigma\Big(\bigcup_{n\geq 0}^{\infty}\mathcal{F}_t^{\ell,\delta_n}\Big) \quad \text{for every } t\geq 0.$$

Observe that

$$\{\tau_n\leq t<\tau_{n+1}\}\cap\mathcal{F}_t^{\ell,\delta}=\{\tau_n\leq t<\tau_{n+1}\}\cap\mathcal{F}_t^{\ell,\delta_m} \qquad \text{for every } t\geq 0,\ m\geq n\geq 0, \text{ and } \ell\geq 0, \tag{2.1}$$

and $\tau_n\in\delta$ is a stopping time of $\mathbb{F}^{\ell,\delta}$ for every $n\geq 0$.

Let $\Delta$ denote the collection of admissible sensor installation policies. For every policy $\delta\in\Delta$ and stopping time $\tau$ of the filtration $\mathbb{F}^{\ell,\delta}$ generated under $\delta$, let us define the Bayes risk

$$R^{\delta,\tau}(\ell,\pi):=\mathbb{E}\left[1_{\{\tau<\Theta\}}+c\,(\tau-\Theta)^{+}+b\,\delta(\tau)\right]$$

as the expected total cost of false alarm frequency, detection delay, and sensor installations for some known constant $c>0$. Our objective is to compute the smallest Bayes risk

$$V(\ell,\pi):=\inf_{\delta\in\Delta,\,\tau\in\mathbb{F}^{\ell,\delta}}R^{\delta,\tau}(\ell,\pi), \quad \ell\geq 0,\ \pi\in[0,1), \tag{2.2}$$

and find an optimal sensor installation strategy $\delta^*\in\Delta$ and an optimal alarm time $\tau^*\in\mathbb{F}^{\ell,\delta^*}$ that attain the infimum in (2.2), if such pairs exist.

For every $\ell \geq 0$ and $\delta \in \Delta$, let us define the posterior probability processes

$$\Pi_t^{\ell,\delta} := \mathbb{P}\{\Theta \leq t \,|\, \mathcal{F}_t^{\ell,\delta}\}, \qquad t \geq 0, \tag{2.3}$$

which is a bounded $(\mathbb{P}, \mathbb{F}^{\ell,\delta})$-submartingale with the last element $\Pi_\infty^{\ell,\delta} = 1$ a.s. thanks to (1.2). Using standard arguments of Shiryaev [31, Chapter 4], one can show that Bayes risk above can be written in terms of the process $\Pi^{\ell,\delta}$ as

$$R^{\delta,\tau}(\ell,\pi) = \mathbb{E}^\pi\left[\int_0^\tau c\,\Pi_t^{\ell,\delta}\mathrm{d}t + 1 - \Pi_\tau^{\ell,\delta} + b\,\delta(\tau)\right], \tag{2.4}$$

where $\mathbb{E}^\pi$ denotes the expectation with respect to the measure $\mathbb{P}$ under which the random variable $\Theta$ has the distribution in (1.2); that is, $\Pi_0^{\ell,\delta} = \pi$ for any $\delta \in \Delta$. Therefore, the problem of finding an optimal detection policy is equivalent to finding a pair $(\delta, \tau)$ minimizing the expectation in (2.4).

Following the usual change-of-measure technique as in, for example Dayanik et al. [14, Section 2], it can be shown that the process $\Pi^{\ell,\delta}$ satisfies

$$\mathrm{d}\Pi_t^{\ell,\delta} = \lambda(1-\Pi_t^{\ell,\delta})\mathrm{d}t + \mu\Pi_t^{\ell,\delta}(1-\Pi_t^{\ell,\delta})\sum_{i=1}^\infty 1_{\{i \leq \ell+\delta(t)\}}\left(\mathrm{d}X_t^{(i)} - \mu\Pi_t^{\ell,\delta}\mathrm{d}t\right), \quad t \geq 0. \tag{2.5}$$

The processes $X_{\tau_n+t}^{(i)} - X_{\tau_n}^{(i)} - \mu\int_{\tau_n}^{\tau_n+t}\Pi_s^{\ell,\delta}ds,\ t \geq 0,\ i \leq \ell+n$ are independent $\left(\mathbb{P}, \left\{\mathcal{F}_{\tau_n+t}^{\ell,\delta}\right\}_{t\geq 0}\right)$-Brownian motions. Then the dynamics in (2.5) imply that over the time interval $[\tau_n, \tau_{n+1})$ the process $\Pi^{\ell,\delta}$ behaves as a one-dimensional *diffusion* process with drift and volatility

$$a(\pi) := \lambda(1-\pi) \qquad \text{and} \qquad \sigma^2(\ell+n,\pi) := (\ell+n)\mu^2\pi^2(1-\pi)^2, \quad \text{respectively.} \tag{2.6}$$

**3. Wiener disorder problem with static monitoring.** Suppose that $\ell$-many sensors are already in place and there is no option to install new sensors. This problem has been solved by Shiryaev [30, 31] for $\ell = 1$ (see also Peskir and Shiryaev [25, Section 22]) and the extension to $\ell \geq 2$ is provided in Dayanik et al. [14]. Let $\Pi^\ell$ denote the conditional probability process of the static problem; namely,

$$\Pi_t^\ell \equiv \Pi_t^{\ell,\delta_0} \quad \text{for every } t \geq 0 \text{ and } \ell \geq 0, \qquad \text{and} \qquad \mathbb{F}^\ell = \mathbb{F}^{\ell,\delta_0}.$$

The $\Pi^\ell$ process evolves according to

$$\Pi_0^\ell = \pi, \qquad \mathrm{d}\Pi_t^\ell = \lambda(1-\Pi_t^\ell)\mathrm{d}t + \mu\Pi_t^\ell(1-\Pi_t^\ell)\sum_{i=1}^\ell\left(\mathrm{d}X_t^{(i)} - \mu\Pi_t^\ell\mathrm{d}t\right), \qquad t \geq 0. \tag{3.1}$$

The minimal Bayes risk becomes

$$\begin{aligned} U(\ell,\pi) := \inf_\tau R^{\delta_0,\tau}(\ell,\pi) &= \begin{Bmatrix} 1 - A_\ell - \displaystyle\int_\pi^{A_\ell}\kappa(\ell,z)\mathrm{d}z, & \text{for } \pi < A_\ell \\ 1-\pi, & \text{for } \pi \geq A_\ell \end{Bmatrix} \\ &= \int_\pi^1 \min\{-\kappa(\ell,z), 1\}\mathrm{d}z, \end{aligned} \tag{3.2}$$

where the infimum is taken over the stopping times of $\mathbb{F}^\ell$; $A_\ell$ is the unique root of $\kappa(\ell,\pi) = -1$, and

$$\begin{aligned} \kappa(\ell,\pi) &:= \frac{c}{\lambda}\left[-\frac{\pi}{1-\pi} + \int_0^\pi \frac{1}{(1-z)^2}e^{\frac{2\lambda}{\ell\mu^2}[\alpha(z)-\alpha(\pi)]}\mathrm{d}z\right], \\ \alpha(\pi) &:= \ln\left(\frac{\pi}{1-\pi}\right) - \frac{1}{\pi} \end{aligned} \tag{3.3}$$

with $\alpha(0+) = -\infty$ and $\alpha(1-) = +\infty$. For $\ell = 0$ the solution in (3.2) still holds provided that we set the integral in (3.3) to zero. The mapping $\pi \mapsto \kappa(\ell,\pi)$ is strictly decreasing with boundary conditions $\kappa(\ell, 0+) = 0$ and $\kappa(\ell, 1-) = -\infty$. Therefore, the unique root $A_\ell$ of the equation $\kappa(\ell,\pi) = -1$ always exists and lies in the interval $A_\ell \in [\lambda/(\lambda+c), 1)$.

For every $\ell \geq 0$, the mapping $\pi \mapsto U(\ell,\pi)$ is concave and strictly decreasing on $[0,1]$. It is strictly concave on $[0, A_\ell)$. It satisfies the variational inequalities

$$\begin{aligned} &\left\{ \begin{aligned} U(\ell,\pi) &< 1-\pi \\ \mathcal{L}[U](\ell,\pi) + c\pi &= 0 \end{aligned} \right\}, \qquad && \pi \in (0, A_\ell), \\ &\left\{ \begin{aligned} U(\ell,\pi) &= 1-\pi \\ \mathcal{L}[U](\ell,\pi) + c\pi &> 0 \end{aligned} \right\}, \qquad && \pi \in (A_\ell, 1), \end{aligned} \tag{3.4}$$

in terms of the operator

$$\mathcal{L}[f](\ell,\pi) := a(\pi) f_\pi(\pi) + \frac{1}{2}\sigma^2(\ell,\pi) f_{\pi\pi}(\pi) \qquad \text{for } \pi \in (0,1) \text{ and } \ell \geq 1, \tag{3.5}$$

defined for a smooth function $f(\cdot)$ on $[0,1]$. Using standard verification arguments with the variational inequalities in (3.4), one can show that the first exit time of $\Pi^\ell$ from the interval $[0, A_\ell)$ is an optimal disorder detection time.

It is easy to see from (3.3) that $\ell \mapsto \kappa(\ell,\pi)$ is strictly increasing for $\pi \in (0,1)$. Hence, $\ell \mapsto U(\ell,\pi)$ is non-increasing as expected. That is, when there are more sensors, the optimal Bayes risk is smaller and the observer can better distinguish pre- and post-disorder regimes. In the limiting case as $\ell \to \infty$, we have $\kappa(\ell,\pi) \nearrow 0$ for $\pi < 1$ due to dominated convergence theorem, and $U(\ell,\pi) \searrow 0$ due to bounded convergence theorem. That is, as the number $\ell$ of initially deployed sensors increases, the Bayes risk vanishes.

Lemma 3.1. For $\ell_1 < \ell_2$, we have $\kappa(\ell_2,\cdot) > \kappa(\ell_1,\cdot)$ on $(0,1)$, and therefore $A_{\ell_1} < A_{\ell_2}$. On the interval $\pi \in [A_{\ell_2}, 1]$, it is obvious that $U(\ell_1,\pi) = U(\ell_2,\pi) = 1-\pi$. For $\pi \in [A_{\ell_1}, A_{\ell_2})$, we have $U_\pi(\ell_1,\pi) = -1 < \kappa(\ell_2,\pi) = U_\pi(\ell_2,\pi)$. Hence, $\pi \mapsto U(\ell_1,\pi) - U(\ell_2,\pi)$ is strictly decreasing on this interval. On $(0, A_{\ell_1})$, we have

$$U_\pi(\ell_1,\pi) - U_\pi(\ell_2,\pi) = \kappa(\ell_1,\pi) - \kappa(\ell_2,\pi) < 0,$$

and $\pi \mapsto U(\ell_1,\pi) - U(\ell_2,\pi)$ is strictly decreasing on this interval as well.

Remark 3.1. As $\pi$ decreases, a decision maker will observe two separate regimes with higher probability, in which case we intuitively expect additional sensors to be more useful. Lemma 3.1 above indeed confirms this intuition, and this suggests that it should be optimal to add a new sensor only when the conditional probability process $\Pi^{\ell,\cdot}$ is low enough provided that the cost of a new sensor is not high. In Section 5, we indeed verify that the solution has this structure.

**4. Dynamic programming operator.** At the first decision time an observer will either raise the detection alarm or install an additional sensor. If an alarm is raised first then the detection problem terminates, otherwise it regenerates itself with a new posterior probability and one extra already deployed sensor. Therefore, we expect the value function $V(\cdot,\cdot)$ to satisfy

$$\begin{aligned} V(\ell,\pi) &= \inf_{\tau,\tau_1 \in \mathbb{F}^\ell} \mathbb{E}^\pi \left[ \int_0^{\tau\wedge\tau_1} c\,\Pi_t^\ell \mathrm{d}t + 1_{\{\tau<\tau_1\}}(1-\Pi_\tau^\ell) + 1_{\{\tau_1\leq\tau\}} \left(b + V(\ell+1, \Pi_{\tau_1}^\ell)\right) \right] \\ &= \inf_{\tau \in \mathbb{F}^\ell} \mathbb{E}^\pi \left[ \int_0^\tau c\,\Pi_t^\ell \mathrm{d}t + \min\left\{1-\Pi_\tau^\ell\,,\, b + V(\ell+1,\Pi_\tau^\ell)\right\} \right] \equiv \mathcal{D}[V(\ell+1,\cdot)](\ell,\pi) \end{aligned} \tag{4.1}$$

in terms of the operator

$$\mathcal{D}[f](\ell,\pi) := \inf_{\tau\in\mathbb{F}^\ell} \mathbb{E}^\pi \left[ \int_0^\tau c\,\Pi_t^\ell dt + \min\left\{1-\Pi_\tau^\ell\,,\, b + f(\Pi_\tau^\ell)\right\} \right], \tag{4.2}$$

defined for a bounded Borel function $f(\cdot)$ on $[0,1]$. The problem in (4.2) is a one-dimensional optimal stopping problem for the process $\Pi^\ell$ in (3.1) with running cost $c\pi$ and terminal cost $\min\{1-\pi,\, b+f(\pi)\}$.

LEMMA 4.1. Because $\min\{1-\pi\,,\, b+f(\pi)\} \le 1-\pi$, it is obvious that $\mathcal{D}[f](\ell,\pi) \le U(\ell,\pi)$ for every bounded $f(\cdot)$.

REMARK 4.1. For every $\ell \ge 0$ and $\Pi_0^\ell = \pi \in [0,1)$, the process $\Pi^\ell$ is a submartingale with the last element $\Pi_\infty^\ell = 1$ almost surely. For $\ell = 0$, it drifts deterministically towards the point 1. For $\ell \ge 1$, it is a diffusion process and it can be proven as in Dayanik et al. [14, Appendix A1] that the end-points 0 and 1 of the state space are, respectively, *entry-but-not-exit* and *natural* boundaries. Moreover, using the dynamics in (3.1) we obtain

$$\mathbb{E}^\pi \Pi_{t\wedge\tau_r}^\ell = \pi + \mathbb{E}^\pi \int_0^{t\wedge\tau_r} \Big(\lambda(1-\Pi_s^\ell) + \frac{1}{2}\ell\mu^2(\Pi_s^\ell)^2(1-\Pi_s^\ell)^2\Big)\mathrm{d}s \ge \mathbb{E}^\pi \int_0^{t\wedge\tau_r} \lambda(1-\Pi_s^\ell)\mathrm{d}s,$$

where $\tau_r$ denotes the entrance time of $\Pi^\ell$ into the set $[r,1]$ for $r<1$. For $s \le \tau_r$, $1-\Pi_s \ge 1-r$ and

$$1 \ge \mathbb{E}^\pi \Pi_{t\wedge\tau_r}^\ell \ge \lambda(1-r)\mathbb{E}^\pi[t\wedge\tau_r],$$

which further implies that $\tau_r$ is uniformly integrable in $\pi$ thanks to monotone convergence theorem.

In the remainder of this section, we fix $\ell$ and construct $\mathcal{D}[f](\ell,\cdot)$ for a typical function $f(\cdot)$ satisfying Assumption 4.1 below, which lists the properties that $V(\ell+1,\cdot)$ is expected to possess.

ASSUMPTION 4.1. *Let $f:[0,1]\mapsto\mathbb{R}$ be any fixed function with the following properties:*

*(i) It is a strictly decreasing and concave function bounded as $0 \le f(\cdot) \le U(\ell+1,\cdot)$.*

*(ii) There exists a point $\bar{A} \in [A_\ell,1]$ such that $f(\pi) = 1-\pi$ for every $\pi \in [\bar{A},1]$. For $\pi \in [0,\bar{A}]$, $\pi \mapsto f(\pi)$ is strictly concave and $f(\pi) < 1-\pi$.*

*(iii) It is twice continuously differentiable except possibly at a finite number of points where it is continuously differentiable. Its derivative $f_\pi(\cdot)$ is bounded below by $\kappa(\ell+1,\cdot)$. Wherever it is twice continuously differentiable, it satisfies the inequality $\mathcal{L}[f](\ell+1,\pi)+c\pi \ge 0$, where the operator $\mathcal{L}$ is defined in (3.5).*

The concavity of $\pi\mapsto f(\pi)$ implies $f_\pi(\cdot) \ge -1$, which further yields $f_\pi(\cdot) \ge \max\{\kappa(\ell+1,\cdot),-1\} = U_\pi(\ell+1,\cdot)$. Moreover, we have

$$\mathcal{L}[f](\ell,\pi)+c\pi = \mathcal{L}[f](\ell+1,\pi)+c\pi - \frac{\mu^2}{2}\pi^2(1-\pi)^2 \underbrace{f_{\pi\pi}(\pi)}_{\le 0} \ge 0 \tag{4.3}$$

wherever $f(\cdot)$ is twice-differentiable; the inequality is strict for $\pi \in (0,\bar{A})$ where $f$ is strictly concave.

Using the properties given in Assumption 4.1, we will show that the continuation region of the problem in (4.2) is one-sided if $b \ge U(\ell,0) - f(0)$, and it is two-sided otherwise. The difference $U(\ell,0)-f(0)$ is positive since $f(0) \le U(\ell+1,0) < U(\ell,0)$; see Lemma 3.1. We also have $f_\pi(\cdot) \ge U_\pi(\ell+1,\cdot) \ge U_\pi(\ell,\cdot)$. Then the inequality $b \ge U(\ell,0)-f(0)$ implies that $b+f(\pi) \ge U(\ell,\pi)$ for all $\pi\in[0,1]$. If we interpret the function $f(\cdot)$ as the value function in (2.2) with $\ell+1$ sensors already in place, the inequality $b+f(\cdot) \ge U(\ell,\cdot)$ suggests that adding a new sensor would always be more costly; therefore, we should simply apply the one-sided detection policy of the static problem in Section 3.

THEOREM 4.1. Under the assumptions above, we have $f(\pi) \le \mathcal{D}[f](\ell,\pi) \le b+f(\pi)$, for all $\pi\in[0,1]$.

$T$ he inequality $\mathcal{D}[f](\ell,\pi) \le b + f(\pi)$ is obvious from the definition of $\mathcal{D}[f](\cdot,\cdot)$ in (4.2). To show the inequality $f(\pi) \le \mathcal{D}[f](\ell,\pi)$, we recall that $f_\pi(\cdot)$ is bounded in $[-1,0]$ and $\mathcal{L}[f](\ell,\pi) + c\pi \ge 0$. Therefore, when we apply the Itô rule for the process $t \mapsto f(\Pi_t^\ell)$ for a given $\mathbb{F}^\ell$-stopping time $\tau$, we obtain

$$\mathbb{E}^\pi f(\Pi_{\tau\wedge t}^\ell) - f(\pi) \ge -\mathbb{E}^\pi \int_0^{\tau\wedge t} c\Pi_s^\ell \mathrm{d}s.$$

We then let $t \to \infty$ and employ the monotone and dominated convergence theorems to obtain $f(\pi) \le \mathbb{E}^\pi \left[\int_0^\tau c\Pi_s^\ell \mathrm{d}s + f(\Pi_\tau^\ell)\right]$. Note that $f(\pi) \le 1-\pi$ by assumption (and $f(\pi) \le \min\{1-\pi, b+f(\pi)\}$). Hence we have

$$f(\pi) \le \mathbb{E}^\pi \left[\int_0^\tau c\Pi_s^\ell \mathrm{d}s + \min\left\{1-\Pi_\tau^\ell\,,\, b+f(\Pi_\tau^\ell)\right\}\right],$$

which implies that $f(\pi) \le \mathcal{D}[f](\ell,\pi)$ since $\tau$ above is arbitrary.

**4.1. Explicit solution of (4.2) for $\ell = 0$.** The process $\Pi^0$ is deterministic and solves $\mathrm{d}\Pi_t^0/\mathrm{d}t = \lambda(1-\Pi_t^0)$ with $\Pi_0^0 = \pi$. Removing the expectation operator in (4.2) the problem becomes

$$\begin{aligned}\mathcal{D}[f](0,\pi) &= \inf_t \min\left\{\int_0^t c\,\Pi_s^0 \mathrm{d}s + 1 - \Pi_t^0,\ \int_0^t c\,\Pi_s^0 \mathrm{d}s + b + f\left(\Pi_t^0\right)\right\}\\ &= \min\left\{\inf_t \int_0^t c\,\Pi_s^0 \mathrm{d}s + 1 - \Pi_t^0,\ \inf_t \int_0^t c\,\Pi_s^0 \mathrm{d}s + b + f\left(\Pi_t^0\right)\right\}.\end{aligned} \tag{4.4}$$

The first infimum above gives the function $U(0,\pi)$. For the second infimum, we note that

$$\frac{\partial}{\partial t}\left[\int_0^t c\,\Pi_s^0 \mathrm{d}s + b + f\left(\Pi_t^0\right)\right] = c\,\Pi_t^0 + a(\Pi_t^0)\, f_\pi\left(\Pi_t^0\right) \ge c\,\Pi_t^0 + a(\Pi_t^0)\,\kappa\left(1,\Pi_t^0\right)$$

since $f_\pi(\cdot) \ge \kappa(1,\cdot)$ by Assumption 4.1 (iii). Note also that $\kappa(\ell,\pi)$ solves

$$a(\pi)\,\kappa(\ell,\pi) + \frac{1}{2}\sigma^2(\ell,\pi)\,\kappa_\pi(\ell,\pi) + c\pi = 0 \tag{4.5}$$

for every $\ell \ge 0$, and therefore we obtain

$$\frac{\partial}{\partial t}\left[\int_0^t c\,\Pi_s^0 \mathrm{d}s + b + f\left(\Pi_t^0\right)\right] \ge -\frac{1}{2}\sigma^2(1,\pi)\kappa_\pi\left(1,\Pi_t^0\right) \ge 0,$$

where the last inequality follows because $\pi \mapsto \kappa(\ell,\pi)$ is (strictly) decreasing for every $\ell \ge 0$. This implies that the infimum is attained at $t = 0$ and

$$\mathcal{D}[f](0,\pi) = \min\left\{U(0,\pi)\,,\, b + f(\pi)\right\}, \tag{4.6}$$

which also shows that $\pi \mapsto \mathcal{D}[f](0,\pi)$ is strictly decreasing and concave as it is the minimum of two such functions.

Let us first assume that $b \ge U(0,0) - f(0)$. In this case, we have $b + f(\pi) \ge U(0,\pi)$ for all $\pi \in [0,1]$, and $\mathcal{D}[f](0,\cdot) = U(0,\cdot)$. Therefore, $\mathcal{D}[f](0,\cdot)$ has the explicit form given in (3.2), and it solves the variational inequalities in (3.4). It also satisfies the properties given in Assumption 4.1 with $\ell+1$ replaced by 0.

Let us next assume that the opposite inequality $b < U(0,0) - f(0)$ holds. On $(0, A_0) \equiv (0, \lambda/(\lambda + c))$, we have $U_\pi(0,\pi) = \kappa(0,\pi) < \kappa(1,\pi) \le f_\pi(\pi)$, and on $[A_0, \bar{A})$, $U_\pi(0,\pi) = -1 < f_\pi(\pi)$ because $f(\cdot)$ is strictly concave on $[0,\bar{A}]$ and $f_\pi(\bar{A}-) = -1$. Hence we observe that $\pi \mapsto b + f(\pi) - U(0,\pi)$

is strictly increasing for $\pi \in (0, \bar{A})$. Its value at $\pi = 0$ is strictly negative by assumption, and its value at $\pi = \bar{A}$ is clearly equal to $b > 0$. Therefore, there exists a point $B_0[f] < \bar{A}$, at which $b + f(\cdot)$ intersects with $U(0, \cdot)$, and

$$\begin{aligned} &\mathcal{D}[f](0,\pi) = b + f(\pi) < U(0,\pi) && \text{on } [0, B_0[f]), \\ &\mathcal{D}[f](0,\pi) = U(0,\pi) < b + f(\pi) && \text{on } (B_0[f], 1]. \end{aligned}$$

There also exists a point

$$A_0[f] := \max\{A_0, \bar{B}[f]\}, \text{ where } \bar{B}[f] := \min\{\pi \in [0,1] : b + f(\pi) \geq 1 - \pi\} < \bar{A} \tag{4.7}$$

such that $\mathcal{D}[f](0,\pi) = 1 - \pi$ to the right of the point $A_0[f]$. If $b + f(A_0) - U(0, A_0) \leq 0$, then $B_0[f] = A_0[f] = \bar{B}[f]$. Otherwise $B_0[f] < A_0[f] = A_0$ and $\mathcal{D}[f](0,\pi) = U(0,\pi) < \min\{b + f(\pi), 1 - \pi\}$ for $\pi \in (B_0[f], A_0[f])$. In either case $\mathcal{D}[f](0, \cdot)$ is strictly concave on $[0, A_0[f])$ due to strict concavity of $f(\cdot)$ and $U(0, \cdot)$ respectively on $[0, \bar{A}]$ and $[0, A_0]$.

The function $\mathcal{D}[f](0, \cdot)$ is clearly not differentiable at $\pi = B_0[f]$. At other points, it inherits its smoothness from $f(\cdot)$ and $U(0, \cdot)$; that is, it is twice continuously differentiable except possibly at finitely many points where it is still continuously differentiable. For $\pi \in (0, B_0[f]) \subseteq (0, \bar{A})$, we have $\mathcal{L}[\mathcal{D}[f]](0,\pi) + c\pi = \mathcal{L}[f](0,\pi) + c\pi > 0$ since $f(\cdot)$ is strictly concave on this region; see (4.3). Provided that $(B_0[f], A_0[f]) \neq \varnothing$ (that is, $A_0[f] = A_0$), on this interval we obviously have $\mathcal{L}[\mathcal{D}[f]](0,\pi) + c\pi = \mathcal{L}[U](0,\pi) + c\pi = 0$. Also, for $\pi \in (A_0[f], 1) \subseteq (A_0, 1) \equiv (\lambda/(\lambda + c), 1)$, $\mathcal{D}[f](0,\pi) = 1 - \pi$ and $\mathcal{L}[\mathcal{D}[f]](0,\pi) + c\pi = -\lambda(1 - \pi) + c\pi > 0$.

Finally, because $f_\pi(\cdot) \geq \kappa(1, \cdot)$ by Assumption 4.1 (iii) and $U_\pi(0, \cdot) = \max\{\kappa(0, \cdot), -1\}$, it follows that $(\mathcal{D}[f](0, \cdot))_\pi \geq \kappa(0, \cdot)$ on $(0,1) \setminus \{B[f]\}$. Hence, the function $\mathcal{D}[f](0, \cdot)$ satisfies all the properties in Assumption 4.1 above with $\ell + 1$ replaced by 0, except that it is not differentiable at $B_0[f]$. The following corollary is now immediate, and it summarizes the case for $\ell = 0$.

Proposition 4.1. The function $\mathcal{D}[f](0, \cdot)$ is concave and strictly decreasing. If $b \geq U(0,0) - f(0)$, then it equals $U(0, \cdot)$ and solves the variational inequalities in (3.4), and satisfies Assumption 4.1 above with $\ell + 1$ replaced by 0.

If $b < U(0,0) - f(0)$, then it still satisfies Assumption 4.1 with $\ell + 1$ replaced by 0, but is not differentiable at the intersection point $B_0[f]$ of the functions $b + f(\cdot)$ and $U(0, \cdot)$, and solves the variational inequalities

$$\left.\begin{cases} \mathcal{D}[f](0,\pi) = b + f(\pi) < 1 - \pi \\ \mathcal{L}[\mathcal{D}[f]](0,\pi) + c\pi > 0 \end{cases}\right\}, \quad \pi \in (0, B_0[f]),$$
$$\left.\begin{cases} \mathcal{D}[f](0,\pi) < \min\{b + f(\pi), 1 - \pi\} \\ \mathcal{L}[\mathcal{D}[f]](0,\pi) + c\pi = 0 \end{cases}\right\}, \quad \pi \in (B_0[f], A_0[f]), \tag{4.8}$$
$$\left.\begin{cases} \mathcal{D}[f](0,\pi) = 1 - \pi < b + f(\pi) \\ \mathcal{L}[\mathcal{D}[f]](0,\pi) + c\pi > 0 \end{cases}\right\}, \quad \pi \in (A_0[f], 1),$$

where $A_0[f] = \max\{A_0, \bar{B}[f]\}$.

**4.2. Solution of (4.2) for $\ell \geq 1$.** As in the case for $\ell = 0$, let us firstly assume that $b \geq U(\ell, 0) - f(0)$. Recall that this inequality implies $b + f(\pi) \geq U(\ell, \pi)$ for all $\pi \in [0,1]$, and we obtain

$$\mathcal{D}[f](\ell,\pi) \geq \inf_\tau \mathbb{E}\left[\int_0^\tau c\Pi_t^\ell \mathrm{d}t + \min\left\{1 - \Pi_\tau^\ell,\, U(\ell, \Pi_\tau^\ell)\right\}\right] = \inf_\tau \mathbb{E}\left[\int_0^\tau c\Pi_t^\ell \mathrm{d}t + U(\ell, \Pi_\tau^\ell)\right]. \tag{4.9}$$

Using the variational inequalities in (3.4) it is easy to show that the last infimum above equals the function $U(\ell, \pi)$. This implies that $\mathcal{D}[f](\ell,\pi) = U(\ell,\pi)$ thanks to Lemma 4.1. Hence the function $\mathcal{D}[f](\ell, \cdot)$ solves the variational inequalities in (3.4) and it satisfies Assumption 4.1 with $\ell + 1$ replaced with $\ell$.

In the remainder, we solve the problem in the more difficult case where $b < U(\ell,0) - f(0)$. For notational convenience, let us introduce

$$H_B[f](\ell,\pi) := e^{\frac{2\lambda}{\ell\mu^2}\alpha(B)}\left[f_\pi(B) - \kappa(\ell,B)\right] e^{-\frac{2\lambda}{\ell\mu^2}\alpha(\pi)} + \kappa(\ell,\pi), \qquad \pi \in [0,1],\ B \in (0,\bar{A}), \tag{4.10}$$

which solves the equation

$$a(\pi)\Big(H_B[f](\ell,\pi)\Big) + \frac{1}{2}\sigma^2(\ell,\pi)\,\frac{\partial\Big(H_B[f](\ell,\pi)\Big)}{\partial\pi} + c\pi = 0, \tag{4.11}$$
for $\pi \in (0,1)$, and with the condition $H_B[f](\ell,B) = f_\pi(B)$.

Note that $f_\pi(B) \geq \kappa(\ell+1,B) > \kappa(\ell,B)$, and the mappings $\pi \mapsto e^{-\frac{2\lambda}{\mu^2}\alpha(\pi)}$ and $\pi \mapsto \kappa(\ell,\pi)$ are decreasing. Hence, $\pi \mapsto H_B[f](\ell,\pi)$ is strictly decreasing. As $\pi \searrow 0$ and $\pi \nearrow 1$, $H_B[f](\ell,\pi)$ goes to $\infty$ and $-\infty$ respectively, and its value at $\pi = B$ is strictly greater than -1 since $B < \bar{A}$. Therefore, by continuity there exists a unique point, call $\Lambda_B \in (B,1)$, such that

$$H_B[f](\ell,\Lambda_B) = -1 \tag{4.12}$$

and $H_B[f](\ell,\pi)$ is strictly less (greater) than $-1$ for $\pi > \Lambda_B$ (for $\pi < \Lambda_B$).

Corollary 4.1. Recall that $A_\ell$ is the unique root of the equation $\kappa(\ell,\pi) = -1$. Since $f_\pi(B) \geq \kappa(\ell+1,B) > \kappa(\ell,B)$, it follows that $H_B[f](\ell,\pi) > \kappa(\ell,\pi)$ for $\pi < 1$, and $\Lambda_B > A_\ell$.

Theorem 4.2. The mapping $y \mapsto e^{\frac{2\lambda}{\ell\mu^2}\alpha(y)}\left[f_\pi(y) - \kappa(\ell,y)\right]$ is strictly increasing on $(0,1)$. Therefore, for $B_1 < B_2$, $H_{B_1}[f](\ell,\pi) < H_{B_2}[f](\ell,\pi)$ on $\pi \in (0,1)$ and $\Lambda_{B_1} < \Lambda_{B_2}$.

*R* ecall that $f_\pi(\cdot)$ is continuous and $f_{\pi\pi}(\cdot)$ may fail to exist at finitely many points only. Hence, it is enough to show that the derivative of the expression above with respect to $y$ is strictly positive, wherever it exists. To this end, we have $\partial\Big[e^{\frac{2\lambda}{\ell\mu^2}\alpha(y)}[f_\pi(y) - \kappa(\ell,y)]\Big]/\partial y$

$$\begin{aligned} &= e^{\frac{2\lambda}{\ell\mu^2}\alpha(y)}\left[f_{\pi\pi}(y) + \frac{2\lambda}{\ell\mu^2}\alpha'(y)f_\pi(y) - \kappa_\pi(\mu,y) - \frac{2\lambda}{\ell\mu^2}\alpha'(y)\kappa(\ell,y)\right] \\ &= e^{\frac{2\lambda}{\ell\mu^2}\alpha(y)}\left[f_{\pi\pi}(y) + \frac{2\lambda}{\ell\mu^2}\alpha'(y)f_\pi(y) + \frac{2c}{\ell\mu^2}\frac{1}{y(1-y)^2}\right]. \end{aligned} \tag{4.13}$$

For $y \geq \bar{A}$, we have $f(y) = 1 - y$ and this yields

$$\frac{\partial}{\partial y}\left[e^{\frac{2\lambda}{\ell\mu^2}\alpha(y)}\left[f_\pi(y) - \kappa(\ell,y)\right]\right] = e^{\frac{2\lambda}{\ell\mu^2}\alpha(y)}\frac{2}{\ell\mu^2}\frac{1}{y^2(1-y)^2}\left[-\lambda(1-y) + cy\right] > 0$$

since $\bar{A} \geq A_{\ell+1} > \lambda/(c+\lambda)$. For $y < \bar{A}$,

$$\frac{2\lambda}{\ell\mu^2}\alpha'(y)f_\pi(y) + f_{\pi\pi}(y) + \frac{2c}{\ell\mu^2 y(1-y)^2} = \frac{2}{\ell\mu^2 y^2(1-y)^2}\left[\mathcal{L}[f](\ell,y) + cy\right]. \tag{4.14}$$

which is again strictly positive (see (4.3) and the argument following it for strict inequality).

For fixed $B \in (0,\bar{A})$, let us now introduce the function

$$M_B[f](\ell,\pi) := \begin{cases} b + f(\pi), & \text{for } 0 \leq \pi \leq B, \\ b + f(B) + \displaystyle\int_B^\pi H_B[f](\ell,z)\mathrm{d}z, & \text{for } B < \pi \leq \Lambda_B, \\ b + f(B) + \displaystyle\int_B^{\Lambda_B} H_B[f](\ell,z)\mathrm{d}z + \int_{\Lambda_B}^\pi (-1)\mathrm{d}z, & \text{for } \Lambda_B < \pi \leq 1. \end{cases} \tag{4.15}$$

For $B < \pi \le 1$, this function can also be written as $b + f(B) + \int_B^\pi G_B[f](\ell,z)\mathrm{d}z$, where

$$G_B[f](\ell,\pi) := \max\{H_B[f](\ell,\pi)\,,\,-1\}, \quad \pi\in[0,1]. \tag{4.16}$$

By Theorem 4.2, $\pi \mapsto e^{\frac{2\lambda}{\ell\mu^2}\alpha(\pi)}\left[f_\pi(\pi)-\kappa(\ell,\pi)\right]$ is strictly increasing. Then for $\pi < B$ we have

$$e^{\frac{2\lambda}{\ell\mu^2}\alpha(\pi)}\left[f_\pi(\pi)-\kappa(\ell,\pi)\right] < e^{\frac{2\lambda}{\ell\mu^2}\alpha(B)}\left[f_\pi(B)-\kappa(\ell,B)\right]$$

and rearranging the terms yields $f_\pi(\pi) < H_B[f](\ell,\pi)$. Clearly, reversing the inequalities gives $f_\pi(\pi) > H_B[f](\ell,\pi)$ for $\pi > B$. Hence, for $\pi \le \Lambda_B$ the function $M_B[f](\ell,\cdot)$ can be written as $b+f(0)+\int_0^\pi \min\{f_\pi(z)\,,\,H_B[f](\ell,\cdot)\}\mathrm{d}z$, and for $0\le\pi\le1$ it can be represented more compactly as

$$M_B[f](\ell,\pi) = b + f(0) + \int_0^\pi \min\left\{f_\pi(z)\,,\,G_B[f](\ell,z)\right\}\mathrm{d}z. \tag{4.17}$$

LEMMA 4.2. The function $\pi\mapsto M_B[f](\ell,\pi)$ is concave, strictly decreasing and twice continuously differentiable except possibly at finitely many points where it is continuously differentiable. Moreover (except at those points) it satisfies

$$\begin{aligned} &\mathcal{L}\left[M_B[f]\right](\ell,\pi)+c\pi = 0, \quad \text{for } \pi\in(B,\Lambda_B),\\ &\mathcal{L}\left[M_B[f]\right](\ell,\pi)+c\pi > 0, \quad \text{for } \pi\in(0,B)\cup(\Lambda_B,1). \end{aligned} \tag{4.18}$$

On $[0,B)$, it equals $b+f(\pi)$ and therefore it is strictly concave. On $(B,\Lambda_B)$, it is strictly less than $b+f(\pi)$ and strictly concave. Finally, on $(\Lambda_B,1]$ it is linear with slope $-1$ and again strictly less than $b+f(\pi)$.

$B$ y construction, the function is twice-continuously differentiable except possibly at finitely many points where it is continuously differentiable. Its derivative is bounded above as $(M_B[f])_\pi(\ell,\pi) = \min\left\{f_\pi(\pi)\,,\,G_B[f](\ell,\pi)\right\} \le f_\pi(\pi) < 0$ for $\pi\in(0,1)$. Recall that $\pi\mapsto H_B[f](\ell,\pi)$ is strictly decreasing. Then $\pi\mapsto G_B[f](\ell,\pi)$ is obviously non-increasing and as the minimum of two non-increasing function so is $\pi\mapsto(M_B[f])_\pi(\ell,\pi)$. Hence, $(M_B[f])(\ell,\cdot)$ is concave and strictly decreasing.

For $\pi < B$, $M_B[f](\ell,\pi) = b + f(\pi)$, which is strictly concave since $B < \bar{A}$. We also have $\mathcal{L}\left[M_B[f]\right](\ell,\pi)+c\pi > 0$ because $f_{\pi\pi}(\cdot)<0$ on this region (wherever the second derivative exists); see (4.3).

For $\pi\in(B,\Lambda_B)$, the function $\pi\mapsto M_B[f](\ell,\pi)$ solves $\mathcal{L}\left[M_B[f]\right](\ell,\pi)+c\pi=0$ by construction since its derivative $H_B[f](\ell,\cdot)$ solves the equation (4.11). $H_B[f](\ell,\cdot)$ is also strictly decreasing and strictly less than $f_\pi(\cdot)$ on this interval. Therefore, the function $M_B[f](\ell,\cdot)$ is strictly concave and stays below the function $b+f(\cdot)$.

On $\pi\in(\Lambda_B,1]$, $M_B[f](\ell,\cdot)$ is linear with slope $-1$ (see (4.15)). Because $M_B[f](\ell,\Lambda_B) < b+f(\Lambda_B)$ and $f_\pi(\cdot)\ge -1$, it follows that $M_B[f](\ell,\cdot) < b+f(\cdot)$ on this interval as well. Moreover, we have

$$\mathcal{L}\left[M_B[f]\right](\ell,\pi)+c\pi = \lambda(1-\pi)(-1)+c\pi > 0,$$

since $\Lambda_B > A_\ell \ge A_0 \equiv \lambda/(\lambda+c)$ (see Lemma 3.1 and Corollary 4.1).

COROLLARY 4.2. The collection of functions $\left(M_B[f](\ell,\cdot)\right)_{B\in(0,\bar{A})}$ is monotone and non-decreasing in $B$ thanks to the monotonicity of $B\mapsto H_B[f](\ell,\cdot)$ and $B\mapsto G_B[f](\ell,\cdot)$ (see Theorem 4.2 and the definition of $G_B[f](\ell,\cdot)$ in (4.16)). Moreover, for $B<\pi$, we have

$$M_B[f](\ell,\pi) = b+f(B)+\int_B^\pi G_B[f](\ell,z)\mathrm{d}z,$$

which clearly shows that $M_B[f](\ell,\pi) \nearrow b+f(\pi)$ as $B\nearrow\pi$.

LEMMA 4.3. The mapping $B \mapsto M_B[f](\ell,1)$ is continuous and strictly increasing on $(0,\bar{A})$. We have

$$\lim_{B\searrow 0} M_B[f](\ell,1) = b + f(0) - U(\ell,0) < 0 \quad \text{and} \quad M_{\bar{B}[f]}[f](\ell,1) > 0,$$

where $\bar{B}[f]$ is given in (4.7). Therefore there exists a unique $B \in (0,\bar{B}[f])$ such that $M_B[f](\ell,1)$ is equal to zero. That is, $\pi \mapsto M_B[f](\ell,\pi)$ is tangent to the curve $\pi \mapsto 1-\pi$ at the point $\Lambda_B$ (given in (4.12)) and coincides with it thereafter.

$F$ or $B_1 < B_2 < \bar{A}$ we have $G_{B_1}[f](\ell,\cdot) \le G_{B_2}[f](\ell,\cdot)$, which implies that $M_{B_2}[f](\ell,\pi) - M_{B_1}[f](\ell,\pi) \ge 0$ for $\pi \in [0,1]$, and that $\pi \mapsto M_{B_2}[f](\ell,\pi) - M_{B_1}[f](\ell,\pi)$ is nondecreasing. In particular, for $\pi \in (B_1,B_2)$, we have

$$\begin{aligned}\frac{\partial\big[M_{B_2}[f](\ell,\pi) - M_{B_1}[f](\ell,\pi)\big]}{\partial \pi} &= \min\{f_\pi(\pi), G_{B_2}[f](\ell,\pi)\} - \min\{f_\pi(\pi), G_{B_1}[f](\ell,\pi)\} \\ &= f_\pi(\pi) - H_{B_1}[f](\ell,\pi) > 0.\end{aligned}$$

Therefore $M_{B_2}[f](\ell,\pi) - M_{B_1}[f](\ell,\pi) > 0$ for all $\pi > B_1$, and with $\pi = 1$ this shows that $B \mapsto M_B[f](\ell,1)$ is strictly increasing.

Observe that $(B,\pi) \mapsto \min\{f_\pi(\pi), G_B[f](\ell,\pi)\}$ is bounded and jointly continuous on $(0,\bar{A}) \times [0,1]$. Hence

$$B \mapsto M_B[f](\ell,1) = b + f(0) + \int_0^1 \min\{f_\pi(z), G_B[f](\ell,z)\}\mathrm{d}z$$

is continuous on $(0,\bar{A})$. As $B$ goes to 0, we have $H_B[f](\ell,\cdot) \searrow \kappa(\ell,\cdot)$ pointwise, and $M_B[f](\ell,1) \nearrow b + f(0) + \int_0^1 \min\{f_\pi(z), \max\{\kappa(\ell,z),-1\}\}\mathrm{d}z$ thanks to bounded convergence theorem. Because $f_\pi(\cdot) \ge \max\{\kappa(\ell+1,\cdot),-1\} \ge \max\{\kappa(\ell+1,\cdot),-1\} \equiv U_\pi(\ell,\cdot)$, we can rewrite this integral as

$$\int_0^1 \min\{f_\pi(z), \max\{\kappa(\ell,\pi),-1\}\}\mathrm{d}z = \int_0^1 U_\pi(\ell,z)\mathrm{d}z = -U(\ell,0),$$

since $U(\ell,1) = 0$. Therefore, as $B$ goes to 0, $M_B[f](\ell,1)$ decreases and goes to $b + f(0) - U(\ell,0)$, which is strictly negative by assumption.

Recall also that $M_{\bar{B}[f]}[f](\ell,\bar{B}[f]) = b + f(\bar{B}[f])$ since $\bar{B}[f] = \min\{\pi \in [0,1] : b + f(\pi) \ge 1-\pi\}$. Moreover, because $\bar{B}[f] < \bar{A}$, we have $(M_{\bar{B}[f]}[f])_\pi\big(\ell,\bar{B}[f]\big) = f_\pi(\bar{B}[f]) > -1$. Together with the lower bound $(M_{\bar{B}[f]}[f])_\pi(\ell,\pi) \ge -1$ for $\pi \in (0,1)$ this implies that the function $M_{\bar{B}[f]}[f])(\ell,\cdot)$ stays strictly above the function $1-\pi$ for $\pi > \bar{B}[f]$, and therefore $M_{\bar{B}[f]}[f])(\ell,1) > 0$.

Hence, we conclude that the mapping $B \mapsto M_B[f](\ell,1)$ is continuous on $(0,\bar{A})$ and strictly decreasing with $\lim_{B\searrow 0} M_B[f](\ell,1) = b + f(0) - U(\ell,0) < 0$ and $M_{\bar{B}[f]}[f](\ell,1) > 0$. Therefore there exists a unique $B \in (0,\bar{B}[f])$, for which $M_B[f](\ell,1) = 0$.

For notational convenience let $M^*[f](\ell,\cdot)$ denote the function $M_B[f](\ell,\pi)$ for the unique value of $B$ described in Lemma 4.3. Corollary 4.3 below summarizes the properties of the function $M^*[f](\ell,\pi)$. They follow directly from Lemmata 4.2, 4.3 and Corollary 4.2. Using these properties we show in Remark 4.2 that $\mathcal{D}[f](\ell,\cdot) = M^*[f](\ell,\cdot)$. Hence, to be consistent with the notation in Section 4.1 (case $\ell = 0$), we let $B_\ell[f]$ denote the unique $B$ given in Lemma 4.3 and $A_\ell[f]$ denote the point $\Lambda_{B_\ell[f]}$.

COROLLARY 4.3. The function $\pi \mapsto M^*[f](\ell,\pi)$ is strictly decreasing and concave. It satisfies the variational inequalities in (4.18) with $B$ and $\Lambda_{B_\ell[f]}$ replaced by $A_\ell[f]$ and $B_\ell[f]$ respectively.

Because $M^*[f](\ell,1) = 0$ and $B_\ell[f] < \bar{B}[f]$, we have the bounds $0 \le M^*[f](\ell,\pi) \le \min\{1-\pi, b + f(\pi)\}$. For $\pi \le B_\ell[f]$, it equals $b + f(\pi) < 1-\pi$, it is strictly concave, and we have $(M^*[f])_\pi(\ell,\pi) = f_\pi(\pi) \ge \kappa(\ell+1,\pi) > \kappa(\ell,\pi)$. On $(B_\ell[f], A_\ell[f])$, $M^*[f](\ell,\pi)$ is strictly concave and strictly less than $\min\{1-\pi, b + f(\pi)\}$. Moreover $(M^*[f])_\pi(\ell,\pi) = H_{B_\ell[f]}[f](\ell,\pi) > \kappa(\ell,\pi)$ . Finally, for $\pi \ge A_\ell[f]$, $M^*[f](\ell,\pi)$ equals $1-\pi < b + f(\pi)$ and we have $(M^*[f])_\pi(\ell,\pi) = -1 > \kappa(\ell,\pi)$ since $\Lambda_{B_\ell[f]} \equiv A_\ell[f] > A_\ell$; see Corollary 4.1.

Remark 4.2. The function $M^*[f](\ell,\cdot)$ coincides with the value function $\mathcal{D}[f](\ell,\cdot)$, and the exit time of $\Pi^\ell$ from the interval $(B_\ell[f],A_\ell[f])$ is an optimal stopping time for the problem in (4.2).

$T$ he function $M^*[f](\ell,\cdot)$ satisfies (4.18), it equals $\min\{b+f(\pi),1-\pi\}$ for $\pi\notin(B_\ell[f],A_\ell[f])$, and it is strictly below $\min\{b+f(\pi),1-\pi\}$ for $\pi\in(B_\ell[f],A_\ell[f])$. Then the result follows from a straightforward application of Itô Lemma and it is omitted here for conciseness.

The identity $\mathcal{D}[f](\ell,\cdot)=M^*[f](\ell,\cdot)$ in Remark 4.2, the properties of $M^*[f](\ell,\cdot)$ in Corollary 4.3, and the inequality $\mathcal{D}[f](\ell,\cdot)\le U(\ell,\cdot)$ in Lemma 4.1 imply that $\pi\mapsto\mathcal{D}[f](\ell,\pi)$ satisfies the properties in Assumption 4.1 with $\ell+1$ replaced by $\ell$. Moreover, we have

$$\begin{aligned}
&\left\{\begin{array}{l}\mathcal{D}[f](\ell,\pi)=b+f(\pi)<1-\pi\\ \mathcal{L}[\mathcal{D}[f]](\ell,\pi)+c\pi>0\end{array}\right\}, && \pi\in(0,B_\ell[f]),\\
&\left\{\begin{array}{l}\mathcal{D}[f](\ell,\pi)<\min\{b+f(\pi),1-\pi\}\\ \mathcal{L}[\mathcal{D}[f]](\ell,\pi)+c\pi=0\end{array}\right\}, && \pi\in(B_\ell[f],A_\ell[f]),\\
&\left\{\begin{array}{l}\mathcal{D}[f](\ell,\pi)=1-\pi<b+f(\pi)\\ \mathcal{L}[\mathcal{D}[f]](\ell,\pi)+c\pi>0\end{array}\right\}, && \pi\in(A_\ell[f],1).
\end{aligned}\tag{4.19}$$

Corollary 4.2 below summarizes the properties of $\mathcal{D}[f](\ell,\cdot)$ when $\ell\ge1$, and it concludes Section 4.

Proposition 4.2. For $\ell\ge1$, the function $\mathcal{D}[f](\ell,\cdot)$ satisfies the properties in Assumption 4.1 with $\ell+1$ replaced by $\ell$. If $b\ge U(\ell,0)-f(0)$, $\mathcal{D}[f](\ell,\cdot)$ equals $U(\ell,\cdot)$, and it solves the the variational inequalities in (3.4). Otherwise it coincides with the function $M^*[f](\ell,\cdot)$ and satisfies (4.19), where $B_\ell[f]$ is the unique $B$ solving $M_B[f](\ell,1)=0$ and $A_\ell[f]$ is the unique point solving $H_{B_\ell[f]}[f](\ell,A_\ell[f])=-1$ (see (4.12)).

**5. Construction of the value function.** For large values of $\ell$ we expect that it is never optimal to add a new sensor as its cost will exceed its benefit (i.e., reduction in the Bayes risk). Therefore, for large $\ell$, we expect to have $V(\ell,\cdot)=U(\ell,\cdot)=\mathcal{D}[U(\ell+1,\cdot)]$ where the last equality holds if $U(\ell,\cdot)\le U(\ell+1,\cdot)+b$; see Proposition 4.2. With this expectation in mind, let us define

$$L_b:=\max\{\ell\ge0:U(\ell,0)-U(\ell+1,0)>b\}\tag{5.1}$$

with the convention $\max\varnothing=-1$, and let us define the function $F(\cdot,\cdot)$ on $\{0,1,2,\dots\}\times[0,1]$ with

$$F(\ell,\pi)=\begin{cases}\mathcal{D}[F(\ell+1,\cdot)](\ell,\pi), & \text{for } \ell\le L_b,\\ U(\ell,\cdot), & \text{for } \ell\ge L_b+1,\end{cases}\tag{5.2}$$

as a candidate for the value function in (2.2). Observe that $L_b$ is finite and bounded above by

$$\widehat{L}_b:=\min\{\ell\ge0:b\ge U(\ell,0)\}.\tag{5.3}$$

For each $\ell\ge L_b+1$, the function $F(\ell+1,\cdot)\equiv U(\ell+1,\cdot)$ clearly satisfies Assumption 4.1, and $U(\ell,0)\le F(\ell+1,0)+b$ by the definition of $L_b$ in (5.1). Hence, we have $\mathcal{D}[F(\ell+1,\cdot)](\ell,\cdot)=U(\ell,\cdot)\equiv F(\ell,\cdot)$. Also, the results in Proposition 4.2 (and Proposition 4.1) about the variational inequalities obviously apply due to identity $F(\ell,\cdot)\equiv U(\ell,\cdot)$ as well.

For $\ell\le L_b$, $F(\ell,\cdot)=\mathcal{D}[F(\ell+1,\cdot)](\ell,\cdot)$ by construction. Assumption 4.1 holds by induction since $F(L_b+1,\cdot)$ satisfies the same assumption, and Assumption 4.1 is preserved under the dynamic programming operator. Hence, the variational inequalities described in Propositions 4.1 and 4.2 are also valid for $\ell\le L_b$.

Note that if $L_b\ge0$, we have $U(L_b,0)>U(L_b+1,0)+b\equiv F(L_b+1,0)$; see (5.1), therefore $F(L_b,\cdot)$ is different from $U(L_b,\cdot)$. However, for some $0\le\ell<L_b$, the function $F(\ell,\cdot)$ may still be equal to $U(\ell,\cdot)$ as the inequality $U(\ell,0)\le F(\ell+1,0)+b$ may hold.

PROPOSITION 5.1. We have $F(\ell,\cdot)=\mathcal{D}[F(\ell+1,\cdot)](\ell,\cdot)$ for all $\ell\geq 0$. Hence, $F(\ell+1,\cdot)\leq F(\ell,\cdot)\leq F(\ell+1,\cdot)+b$ for all $\ell\geq 0$ thanks to Theorem 4.1.

PROPOSITION 5.2. For every $\ell\geq 0$, let us define

$$\begin{aligned} A^*_\ell &:= \left\{\begin{matrix} A_\ell[F(\ell+1,\cdot)], & \text{if } F(\ell+1,0)+b<U(\ell,0) \\ A_\ell, & \text{otherwise} \end{matrix}\right\}, \\ B^*_\ell &:= \left\{\begin{matrix} B_\ell[F(\ell+1,\cdot)], & \text{if } F(\ell+1,0)+b<U(\ell,0) \\ \text{undefined}, & \text{otherwise} \end{matrix}\right\}, \end{aligned} \tag{5.4}$$

where $A_\ell$ is the boundary of the continuation region of the static problem in Section 3, and $A_\ell[F(\ell+1,\cdot)], B_\ell[F(\ell+1,\cdot)]$ are the points described in Propositions 4.1 and 4.2 with $f(\cdot)$ replaced by $F(\ell+1,\cdot)$. If $F(\ell+1,0)+b\geq U(\ell,0)$ the function $F(\ell,\cdot)$ equals $U(\ell,\cdot)$ and solves the variational inequalities

$$\begin{aligned} &\left\{\begin{matrix} F(\ell,\pi)<1-\pi \\ \mathcal{L}[F(\ell,\cdot)](\ell,\pi)+c\pi=0 \end{matrix}\right\}, & \pi\in(0,A^*_\ell), \\ &\left\{\begin{matrix} F(\ell,\pi)=1-\pi \\ \mathcal{L}[F(\ell,\cdot)](\ell,\pi)+c\pi>0 \end{matrix}\right\}, & \pi\in(A^*_\ell,1). \end{aligned} \tag{5.5}$$

Otherwise,

$$\begin{aligned} &\left\{\begin{matrix} F(\ell,\pi)=b+F(\ell+1,\pi)<1-\pi \\ \mathcal{L}[F(\ell,\cdot)](\ell,\pi)+c\pi>0 \end{matrix}\right\}, & \pi\in(0,B^*_\ell), \\ &\left\{\begin{matrix} F(\ell,\pi)<\min\{b+F(\ell+1,\pi),1-\pi\} \\ \mathcal{L}[F(\ell,\cdot)](\ell,\pi)+c\pi=0 \end{matrix}\right\}, & \pi\in(B^*_\ell,A^*_\ell), \\ &\left\{\begin{matrix} F(\ell,\pi)=1-\pi<b+F(\ell+1,\pi) \\ \mathcal{L}[F(\ell,\cdot)](\ell,\pi)+c\pi>0 \end{matrix}\right\}, & \pi\in(A^*_\ell,1). \end{aligned} \tag{5.6}$$

Using the properties of $F(\cdot,\cdot)$ given in Propositions 5.1 and 5.2 we now prove that the function $F(\cdot,\cdot)$ is the value function of the problem in (2.2). We also present an optimal sensor installation and detection policy.

PROPOSITION 5.3. We have $V(\ell,\pi)=F(\ell,\pi)$ for every $\ell\geq 0$ and $\pi\in[0,1)$. In terms of the hitting times

$$\begin{aligned} \tau^\ell_A &:= \inf\{t\geq 0;\, \Pi^\ell_t\geq A^*_\ell\}, \\ \tau^\ell_S &:= \left\{\begin{matrix} \inf\{t\geq 0;\, \Pi^\ell_t\leq B^*_\ell\}, & \text{if } F(\ell+1,0)+b<U(\ell,0) \\ \infty, & \text{otherwise} \end{matrix}\right\}, \end{aligned} \tag{5.7}$$

an optimal sensor installation policy $\delta^*=(\tau^*_1,\tau^*_2,\ldots)$ and optimal alarm time $\tau^*$ are given by $\tau^*_0\equiv 0$,

$$\begin{aligned} \tau^*_n &= \left\{\begin{matrix} \left(\tau^{\ell+n-1}_S 1_{\{\tau^{\ell+n-1}_S<\tau^{\ell+n-1}_A\}}+\infty 1_{\{\tau^{\ell+n-1}_A<\tau^{\ell+n-1}_S\}}\right)\circ\theta_{\tau^*_{n-1}} & \text{if } \tau^*_{n-1}<\infty \\ \infty & \text{if } \tau^*_{n-1}=\infty \end{matrix}\right\}, \quad n\geq 1, \\ \tau^* &= \sum_{n=0}^{\infty}\left(\tau^{\ell+n}_A 1_{\{\tau^{\ell+n}_A<\tau^{\ell+n}_S\}}\right)\circ\theta_{\tau^*_n}\cdot 1_{\{\tau^*_n<\infty\}}, \end{aligned} \tag{5.8}$$

where $\theta_t$ is the shift operator with $\Pi^\ell_s\circ\theta_t=\Pi^\ell_{t+s}$ for every $t,s\geq 0$ and $\ell\geq 0$.

L et $\delta=(\tau_1,\tau_2,\ldots)$ be any admissible sensor installation policy and $\tau$ be an $\mathbb{F}^\delta$-stopping time. Then,

$$\mathbb{E}\left[F(\delta(\tau\wedge t),\Pi^{\ell,\delta}_{\tau\wedge t})-F(\ell,\pi)\right]=\mathbb{E}\sum_{n\geq 0}1_{\{\tau_n\leq\tau\wedge t<\tau_{n+1}\}}\left[F(\ell+n,\Pi^{\ell,\delta}_{\tau\wedge t})-F(\ell,\pi)\right]. \tag{5.9}$$

First, we note that $\mathbb{E}1_{\{0\le\tau\wedge t<\tau_1\}}\left[F(\ell,\Pi^{\ell,\delta}_{\tau\wedge t})-F(\ell,\pi)\right]=$

$$\mathbb{E}1_{\{0\le\tau\wedge t<\tau_1\}}\int_0^{\tau\wedge t}\mathcal{L}[F(\ell,\cdot)](\ell,\Pi^{\ell,\delta}_s)\mathrm{d}s\ge\mathbb{E}1_{\{0\le\tau\wedge t<\tau_1\}}\int_0^{\tau\wedge t}(-c\Pi^{\ell,\delta}_s)\mathrm{d}s, \tag{5.10}$$

and this holds even when $\ell=0$ (recall that $F(0,\cdot)$ is not differentiable at $B^*_0$; see Proposition 4.1). Next, for $n\ge 1$, we have

$$\begin{aligned}
&\mathbb{E}1_{\{\tau_n\le\tau\wedge t<\tau_{n+1}\}}\left[F(\ell+n,\Pi^{\ell,\delta}_{\tau\wedge t})-F(\ell,\pi)\right]\\
&=\mathbb{E}1_{\{\tau_n\le\tau\wedge t<\tau_{n+1}\}}\Bigg[F(\ell+n,\Pi^{\ell,\delta}_{\tau\wedge t})-F(\ell+n,\Pi^{\ell,\delta}_{\tau_n})\\
&\qquad+\sum_{k=0}^{n-1}\Big(F(\ell+k,\Pi^{\ell,\delta}_{\tau_{k+1}-})-F(\ell+k,\Pi^{\ell,\delta}_{\tau_k})\Big)+\sum_{k=0}^{n-1}\Big(\underbrace{F(\ell+k+1,\Pi^{\ell,\delta}_{\tau_{k+1}})-F(\ell+k,\Pi^{\ell,\delta}_{\tau_{k+1}-})}_{\ge -b}\Big)\Bigg]\\
&\ge\mathbb{E}1_{\{\tau_n\le\tau\wedge t<\tau_{n+1}\}}\Bigg[\int_{\tau_n}^{\tau\wedge t}\underbrace{\mathcal{L}[F(\ell+n,\cdot)](\ell+n,\Pi^{\ell,\delta}_s)}_{\ge -c\Pi^{\ell,\delta}_s}\mathrm{d}s+\sum_{k=0}^{n-1}\int_{\tau_k}^{\tau_{k+1}}\underbrace{\mathcal{L}[F(\ell+k,\cdot)](\ell+k,\Pi^{\ell,\delta}_s)}_{\ge -c\Pi^{\ell,\delta}_s}\mathrm{d}s-nb\Bigg]\\
&\ge\mathbb{E}1_{\{\tau_n\le\tau\wedge t<\tau_{n+1}\}}\left[\int_0^{\tau\wedge t}(-c\Pi_s)\mathrm{d}s-nb\right].
\end{aligned} \tag{5.11}$$

Substituting the inequalities in (5.10) and (5.11) back into (5.9) yields

$$\mathbb{E}F(\delta(\tau\wedge t),\Pi^{\ell,\delta}_{\tau\wedge t})-F(\ell,\pi)\ge\mathbb{E}\left[\int_0^{\tau\wedge t}(-c\Pi^{\ell,\delta}_s)\mathrm{d}s-b\delta(\tau\wedge t)\right].$$

Next, using the inequality $F(\delta(\tau\wedge t),\Pi^{\ell,\delta}_{\tau\wedge t})\le 1-\Pi^{\ell,\delta}_{\tau\wedge t}$, we obtain

$$F(\ell,\pi)\le\mathbb{E}\left[\int_0^{\tau\wedge t}c\Pi^{\ell,\delta}_s\mathrm{d}s+b\delta(\tau\wedge t)+(1-\Pi^{\ell,\delta}_{\tau\wedge t})\right].$$

Finally, letting $t\to\infty$, we have

$$F(\ell,\pi)\le\mathbb{E}\left[\int_0^{\tau}c\Pi^{\ell,\delta}_s\mathrm{d}s+b\delta(\tau)+(1-\Pi^{\ell,\delta}_{\tau})\right].$$

thanks to monotone and bounded convergence theorems. This further implies that $F(\ell,\pi)\le V(\ell,\pi)$. For the policy $\delta^*$ and the alarm time $\tau^*$ in (5.8), the inequalities above hold with equalities.

Proposition 5.3 indicates that, for a given number of sensor $\ell\ge 0$ we first check whether $U(\ell,0)\le V(\ell+1,0)+b$. If this holds, we do not install any new sensor and we apply the detection policy in Section 3; that is, we stop the first time the process $\Pi^\ell$ exceeds the level $A^*_\ell=A_\ell$. Otherwise, we wait until the exit time of $\Pi^\ell$ from the interval $(B^*_\ell,A^*_\ell)$. If the detection boundary $A^*_\ell$ is crossed first, we stop and declare the change. If the left boundary $B^*_\ell$ is crossed instead, we instantaneously add a new sensor and proceed optimally with $\ell+1$ sensors this time. This procedure is repeated until the conditional probability process hits one of the detection boundaries $A^*_\cdot$'s.

Since $F(0,\cdot)\ge\ldots\ge F(\ell,\cdot)\ge F(\ell+1,\cdot)\ge\ldots$, we have $A^*_0\le\ldots\le A^*_\ell\le A^*_{\ell+1}\le\ldots$. However, when they are well-defined, such an ordering does not necessarily hold for sensor thresholds $B^*_\ell$'s as our numerical examples illustrate in the next section. Hence, the continuation regions $(B^*_\ell,A^*_\ell)$'s are not necessarily nested, and upon hitting a sensor threshold more than one sensor may be added.

If the unit sensor installation cost $b$ is rather small, then the upper bound $\widehat{L}_b$ in (5.3) on $L_b$ can be very large. A tighter upper bound $\check{L}_b$ on $L$ is given by

$$\check{L}_b:=\max\left\{\ell\ge 0;\int_0^1\frac{c}{\lambda}\int_0^y\frac{1}{(1-z)^2}\left(1-\exp\left\{\frac{2\lambda}{\mu}[\alpha(z)-\alpha(y)]\frac{1}{\ell(\ell+1)}\right\}\right)\mathrm{d}z\mathrm{d}y>b\right\}, \tag{5.12}$$

**Step 0:** Compute the upper bound $\check{L}_b$ in (5.12) on $L_b$ and

$$L_b = \max\{\ell \in [0, \check{L}_b); U(\ell,0) - U(\ell+1,0) > b\}$$

with the convention $\max\varnothing = -1$. Set $V(\ell,\cdot) = U(\ell,\cdot)$ for every $\ell \geq L_b + 1$. Let $A^*_\ell = A_\ell$ and $B^*_\ell$ be undefined for $\ell \geq L_b$. If $L_b = -1$, stop, otherwise set $\ell = L_b$.
**Step 1a:** If $b + V(\ell+1,0) \geq U(\ell,0)$, set $V(\ell,\cdot) = U(\ell,\cdot)$. Let $A^*_\ell = A_\ell$ and $B^*_\ell$ be undefined.
**Step 1b:** If $b + V(\ell+1,0) < U(\ell,0)$, let $B^*_\ell$ be the unique root of

$$B \mapsto b + V(\ell+1,B) + \int_B^1 G_B[V(\ell+1,\cdot)](\ell,z)dz = 0$$

for $B \in (0, \bar{B}[V(\ell+1,\cdot)])$, where $\bar{B}[V(\ell+1,\cdot)]$ is given in (4.7). Set

$$\begin{aligned} V(\ell,\pi) &= M^*[V(\ell+1,\cdot)](\ell,\pi) \\ &= \begin{cases} b + V(\ell+1,\pi), & \text{if } 0 \leq \pi \leq B^*_\ell, \\ b + V(\ell+1,B^*_\ell) + \displaystyle\int_{B^*_\ell}^{\pi} G_{B^*_\ell}[V(\ell+1,\cdot)](\ell,z)dz, & \text{if } B^*_\ell < \pi \leq 1, \end{cases} \end{aligned}$$

where $G_{B^*_\ell}[V(\ell+1,\cdot)](\ell,\cdot)$ is given in (4.16). Let $A^*_\ell$ be the smallest $A \in (A_\ell, A^*_{\ell+1})$ for which $G_{B^*_\ell}[V(\ell+1,\cdot)](\ell,A) = -1$ or unique $A \in (A_\ell, A^*_{\ell+1})$ such that $H_{B^*_\ell}[V(\ell+1,\cdot)](\ell,A) = -1$.
**Step 2:** Replace $\ell$ with $\ell - 1$. If $\ell \geq 0$, go to Step 1; otherwise, stop.

FIGURE 1. Numerical algorithm to solve the sequential sensor installation problem in (2.2).

because the double integral in (5.12) is decreasing in $\ell$, and

$$\begin{aligned} &U(\ell,0) - U(\ell+1,0) = \int_0^1 \Big(\min\{-\kappa(\ell,y),1\} - \min\{-\kappa(\ell+1,y),1\}\Big)\mathrm{d}y \\ &\leq \int_0^1 \Big(\kappa(\ell+1,y) - \kappa(\ell,y)\Big)\mathrm{d}y \\ &= \int_0^1 \frac{c}{\lambda}\int_0^y \frac{1}{(1-z)^2}\exp\left\{\frac{2\lambda}{(\ell+1)\mu^2}[\alpha(z)-\alpha(y)]\right\}\left(1 - \exp\left\{\frac{2\lambda}{\mu^2}[\alpha(z)-\alpha(y)]\frac{1}{\ell(\ell+1)}\right\}\right)\mathrm{d}z\mathrm{d}y \\ &< \int_0^1 \frac{c}{\lambda}\int_0^y \frac{1}{(1-z)^2}\left(1 - \exp\left\{\frac{2\lambda}{\mu^2}[\alpha(z)-\alpha(y)]\frac{1}{\ell(\ell+1)}\right\}\right)\mathrm{d}z\mathrm{d}y \end{aligned}$$

implies that $L_b = \max\{\ell \geq 0; U(\ell,0) - U(\ell+1,0) > b\} < \check{L}_b$.

**6. Numerical examples.** The iterative construction of the value function in Section 5 is concisely described in Figure 1. In this section, we use that algorithm to illustrate
(i) savings of optimal sequential sensor installation policies over optimal static sensor installation policies,
(ii) iterations of the method and distinct shapes of no-action spaces,
(iii) shapes of the optimal sensor addition regions as the cost parameters $c$ and $b$ jointly change.
In the basic numerical case, we set $\lambda = 0.001$, $\mu = 1$, $c = 0.1$, and $b = 0.01$. In the real applications, quickest detection of infrequent changes is often of more interest and our choice of the small $\lambda$ indicates that we expect one change in every 1000 time units. Later, each parameter is decreased and increased by several folds to measure the sensitivity of the results to the parameter changes. The range of each parameter over which nontrivial optimal sequential installation policies exist turns out to belong to some bounded intervals, which we find by trials and errors. We decrease and increase each parameter one at a time by several folds and calculate optimal static and sequential

Table 1. Sensitivity analysis.

| Case No | $\lambda$ | $\mu$ | $c$ | $b$ | Maximum Percentage Savings | Nested no-action spaces? | Number of iterations |
|---|---|---|---|---|---|---|---|
| 1 | 1e-08 | | | | 11.86 | Yes | 15 |
| 2 | 1e-06 | | | | 11.78 | Yes | 15 |
| 3 | 1e-05 | | | | 11.33 | Yes | 15 |
| 4 | 1e-04 | | | | 10.00 | Yes | 15 |
| 5 | 1e-03 | 1.00 | 1.00e-01 | 1.00e-02 | 8.04 | Yes | 14 |
| 6 | 1e-02 | | | | 6.34 | Yes | 12 |
| 7 | 1e-01 | | | | 4.58 | Yes | 10 |
| 8 | 1e+00 | | | | 1.74 | Yes | 5 |
| 9 | | 0.25 | | | 0.00 | **No** | 40 |
| 10 | | 0.30 | | | 0.70 | **No** | 37 |
| 11 | | 0.50 | | | 5.47 | Yes | 26 |
| 12 | | 0.75 | | | 7.45 | Yes | 18 |
| 13 | | 1.33 | | | 8.11 | Yes | 11 |
| | 1e-03 | 1.00 | 1.00e-01 | 1.00e-02 | 8.04 | Yes | 14 |
| 14 | | 2.00 | | | 7.67 | Yes | 7 |
| 15 | | 4.00 | | | 7.75 | Yes | 4 |
| 16 | | 8.00 | | | 7.23 | Yes | 2 |
| 17 | | 10.00 | | | 6.77 | Yes | 2 |
| 18 | | 12.00 | | | 2.80 | Yes | 1 |
| 19 | | 14.00 | | | 0.00 | Yes | 1 |
| 20 | | | 1.00e-04 | | 0.00 | Yes | 1 |
| 21 | | | 1.00e-03 | | 1.28 | Yes | 1 |
| | 1e-03 | 1.00 | 1.00e-01 | 1.00e-02 | 8.04 | Yes | 14 |
| 22 | | | 4.00e-01 | | 5.90 | Yes | 28 |
| 23 | | | 6.50e-01 | | 3.20 | Yes | 35 |
| 24 | | | 1.00e+00 | | 0.17 | **No** | 42 |
| 25 | | | | 2.00e-04 | 5.98 | Yes | 108 |
| 26 | | | | 3.00e-04 | 6.28 | Yes | 88 |
| 27 | | | | 4.00e-04 | 6.49 | Yes | 76 |
| 28 | | | | 5.00e-04 | 6.67 | Yes | 68 |
| 29 | | | | 6.00e-04 | 6.78 | Yes | 62 |
| 30 | | | | 7.00e-04 | 6.92 | Yes | 58 |
| 31 | | | | 8.00e-04 | 7.01 | Yes | 54 |
| 32 | | | | 9.00e-04 | 7.10 | Yes | 53 |
| 33 | | | | 1.00e-03 | 7.24 | Yes | 48 |
| | 1e-03 | 1.00 | 1.00e-01 | 1.00e-02 | 8.04 | Yes | 14 |
| 34 | | | | 3.00e-02 | 6.26 | Yes | 8 |
| 35 | | | | 5.00e-02 | 4.51 | Yes | 6 |
| 36 | | | | 1.00e-01 | 1.01 | Yes | 4 |
| 37 | | | | 1.50e-01 | 0.00 | Yes | 3 |

installation policies as long as optimal sequential sensor installation policy strictly beats the optimal static sensor policy for at least some values of the prior probability of change at time zero. Those parameter choices are reported in Table 1 along with maximum percentage savings of optimal sequential policy over optimal static policy in the sixth column, type of no-action spaces in the sixth column, and the number of iterations of the algorithm in the last column. The base case corresponds to Case 5 in the same table.

**(i) Savings of optimal sequential sensor installation policies over the optimal static sensor installation policies.** The sixth column of Table 1 reports the maximum percentage savings of optimal sequential sensor policies over the optimal static sensor policies. In the optimal static sensor installation problem, the observer decides on the number of sensors at time zero by solving the problem

$$\min_{\ell} b\ell + U(\ell,\pi) \quad \text{for every } \pi \in [0,1]. \tag{6.1}$$

In the optimal sequential sensor installation problem, the observer starts with no sensor ($\ell = 0$) and adds new sensors sequentially according to the optimal policy as described in (5.8). As a result, the total cost of optimal sequential sensor installation policy becomes

$$V(0,\pi) \equiv \min_{\ell} b\ell + V(\ell,\pi) \quad \text{for every } \pi \in [0,\pi]. \tag{6.2}$$

In (6.1) and (6.2), the minima are taken over the number of sensors to be installed at time zero. We define the *optimal* initial static and sequential sensor numbers as the *largest* minimizers of (6.1) and (6.2), respectively. This choice makes easier the comparisons of optimal initial static and sequential sensor numbers. Optimal initial sensor numbers depend on the prior probability $\pi$ of a change at or before time zero. These numbers are calculated and plotted on the left in Figure 2 for the base case as a function of the prior probability $\pi$ of a change. The solid and broken step functions give the optimal initial static and sequential sensor numbers, respectively. The gap between those lines is notable: on the 70% of the entire $\pi$-region, the optimal sequential installation policy starts at time zero with least 28.57% fewer numbers of sensors than the optimal static installation policy. The percentage savings

$$100 \times \frac{\min_\ell b\ell + U(\ell,\pi) - V(0,\pi)}{\min_\ell b\ell + U(\ell,\pi)} \quad \text{for every } \pi \in [0,1]$$

of optimal sequential policy over the optimal static policy also depends on the prior probability $\pi$ of a change at time zero. Its plot on the right in Figure 2 indicates that the savings can be as large as 8.04% in the base case example.

REMARK 6.1. Observe that the case $\pi = 0$ is degenerate in the sense that the optimal sequential and static policies perform equally well. In the optimal sequential policy, one installs $\ell$-many sensors for which $V(\ell,0)$ is one sided. This means, for that value of $\ell$, $V(\ell,\cdot)$ coincides with $U(\ell,\cdot)$. Hence, optimal sequential policy can be replicated by the optimal static one. As pointed out in Remark 3.1, sensors are useful when there are two separate regimes to differentiate. For higher values of $\pi$, an observer applying a sequential policy delays the sensor installations. However, when $\pi = 0$, the disorder event will occur in the future with probability one, and the investment in the sensors can be done upfront.

The sixth column of Table 1 lists the maximum over prior change probability $\pi \in [0,1]$ of percentage savings

$$\max_{\pi\in[0,1]} 100 \times \frac{\min_\ell b\ell + U(\ell,\pi) - V(0,\pi)}{\min_\ell b\ell + U(\ell,\pi)}$$

of optimal sequential policies over optimal static policies for each of 37 cases. The largest percentage savings observed in 37 cases was 11.78% and continues its steady increase as the disorder rate $\lambda$ decreases to zero; see the upper left plot in Figure 3.

Each panel of Figure 3 reports the maximum percentage savings of optimal sequential policies over optimal static policies as one of the parameters changes while others are kept the same as in Case 5 of Table 1.

In the upper left panel, the percentage savings always increase as the disorder rate $\lambda$ decreases. This is indeed intuitive; as $\lambda$ decreases, expected waiting time until a change increases. Therefore, there is more time to collect observations and a sequential policy uses the resources more effectively.

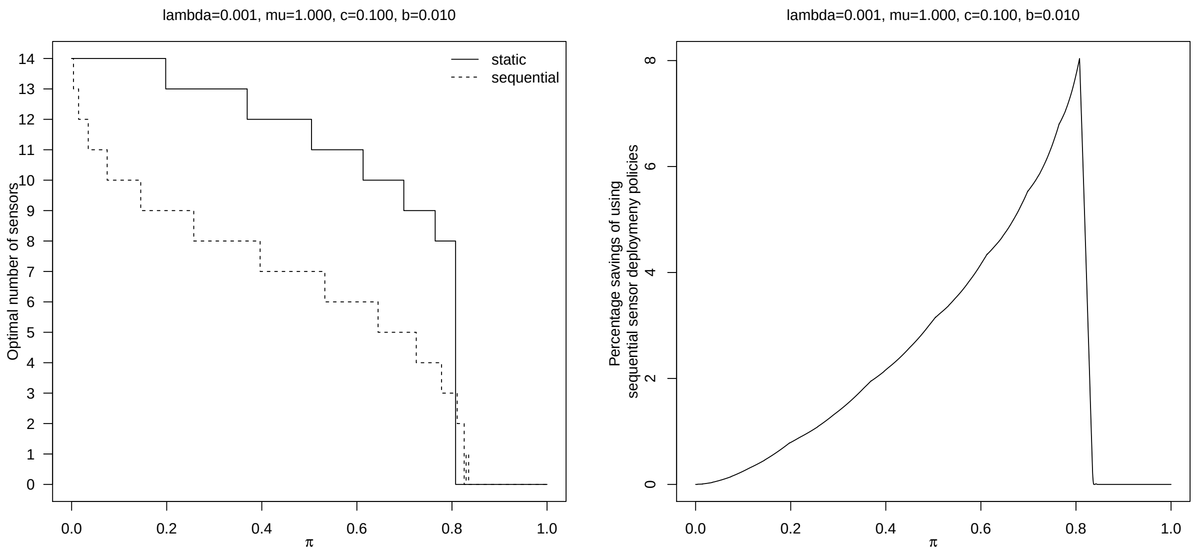



FIGURE 2. Optimal initial static and sequential sensor numbers (on the left) and percentage savings of optimal sequential policy over the optimal static policy for the base case example (Case 5 in Table 1).

As $\lambda$ decreases, the percentage savings increases to a number near 12%, which can be considered as a significant gain.

The remaining three plots of Figure 3 show that the percentage savings of optimal sequential policies first increase and then decrease as $\mu$, $c$, or $b$ increase.

In the upper right panel, the maximum percentage savings of optimal sequential policy over optimal static policy is plotted as $\mu$ changes. The savings are insignificant for very small and very large values of $\mu$. If $\mu$ is very small, then a sensor is very weak and conveys little information. Therefore, both static and sequential policies lose significant amounts of power and the difference between their performances become minuscule. On the other hand, as $\mu$ increases, even a single sensor becomes very vocal in presenting the change, and only very few sensors become adequate to detect the change time quickly. As a result, the savings of sequential policy over static policy vanish as $\mu$ increases. For the intermediate values of $\mu$, however, the percentage savings of optimal sequential policies rest on a wide plateau near 8%, which again presents a significant gain.

The maximum percentage savings of optimal sequential policy over optimal static policy as the unit detection delay cost $c$ varies is displayed in the lower left panel. Once again, the maximum percentage savings vanish if $c$ is very small or very large. If $c$ is very small, then there is sufficiently long time to collect many observations after the change happens. Hence, optimal static policy delays the alarm sufficiently long to reduce the Bayes risk and its performance approaches to that of the optimal sequential policy (in the extreme $c = 0$ case, the static policy with no sensors attains zero Bayes risk). If $c$ is very large, then the problem is forced to terminate very early and sequential sensor installation policy does not have time to flourish. Hence, the savings of optimal sequential policy is minor for very large $c$ values. For the intermediate $c$ values, however, the maximum percentage savings of optimal sequential policy reaches 8%.

Finally, the lower right panel of Figure 3 presents the maximum percentage savings of optimal sequential policy as the unit sensor cost $b$ varies. The savings vanish if $b$ is very small or very large. When the cost of a sensor is small, one can install a large number of sensors at time $t = 0$ for a richer set of observations and reduce the overall Bayes risk. In this case, a static policy will perform well, and additional savings of optimal sequential policy will be slim. If $b$ is very large, then adding

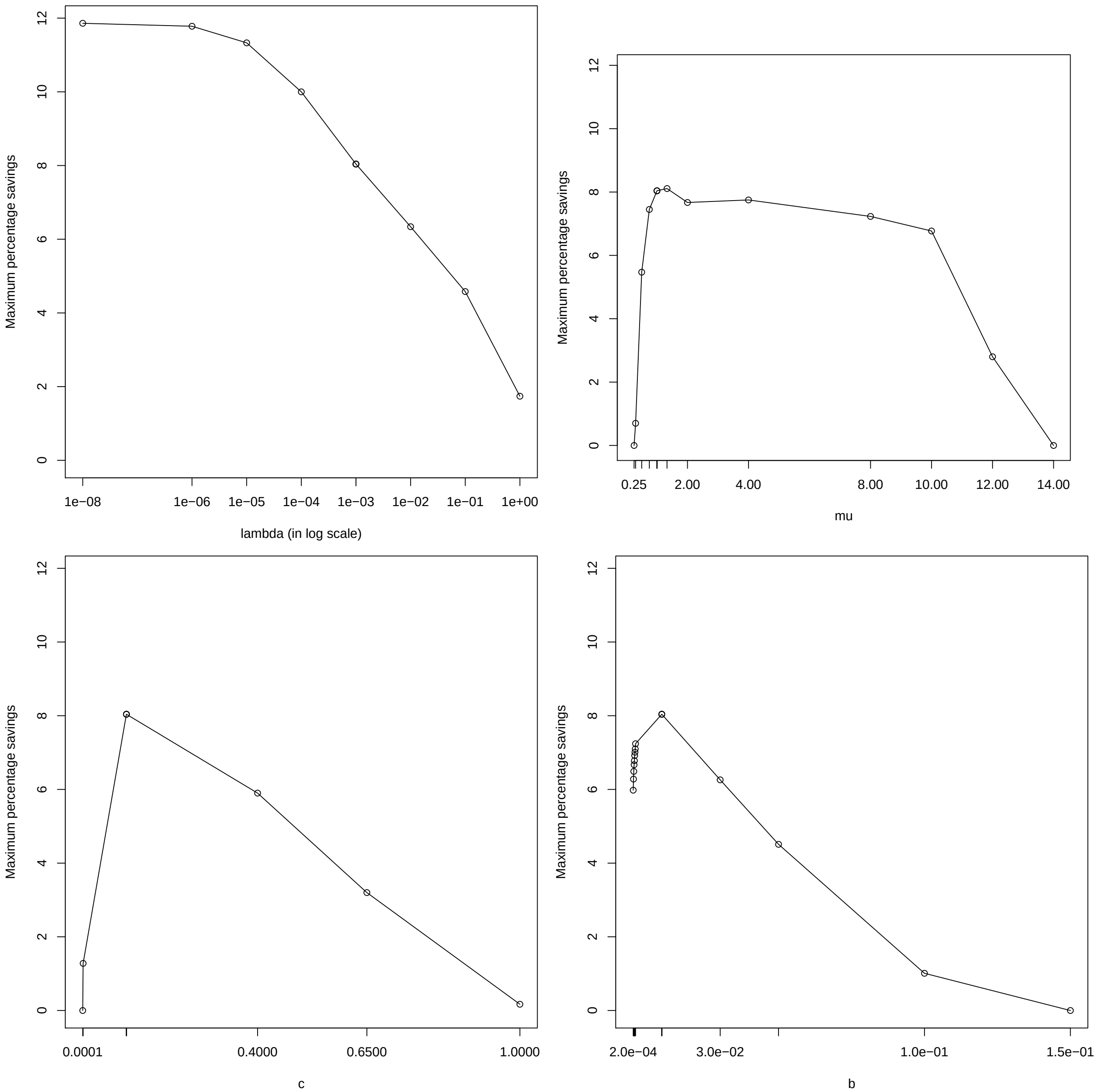



FIGURE 3. The maximum percentage savings of optimal sequential sensor installation policies over optimal static sensor installation policies as each parameter in Case 5 of Table 1 varies one at a time.

a new sensor will be prohibitively expensive. Hence, optimal sequential policy is very unlikely to install new sensors after $t=0$ either, and the savings of optimal sequential policy will again vanish.

**(ii) Iterations of the solution method and distinct shapes of no-action spaces.** Figure 4 displays the value functions for the optimal sequential (solid) and static (dashed) sensor installation problems on the left and no-action spaces for the optimal sequential sensor installation problems on the right for Cases 5 (base case), 24, and 37.

The algorithm in Figure 1 iterates $L_b = 14$, $L_b = 42$, and $L_b = 3$ times, respectively. Therefore, it calculates $V(L_b,\cdot) \equiv U(L_b,\cdot)$ as in Shiryaev's problem and iterates backward as described in Figure 1. As it iterates backwards, notice that the gap between $V(\ell,\cdot)$ and $U(\ell,\cdot)$ widens (compare the

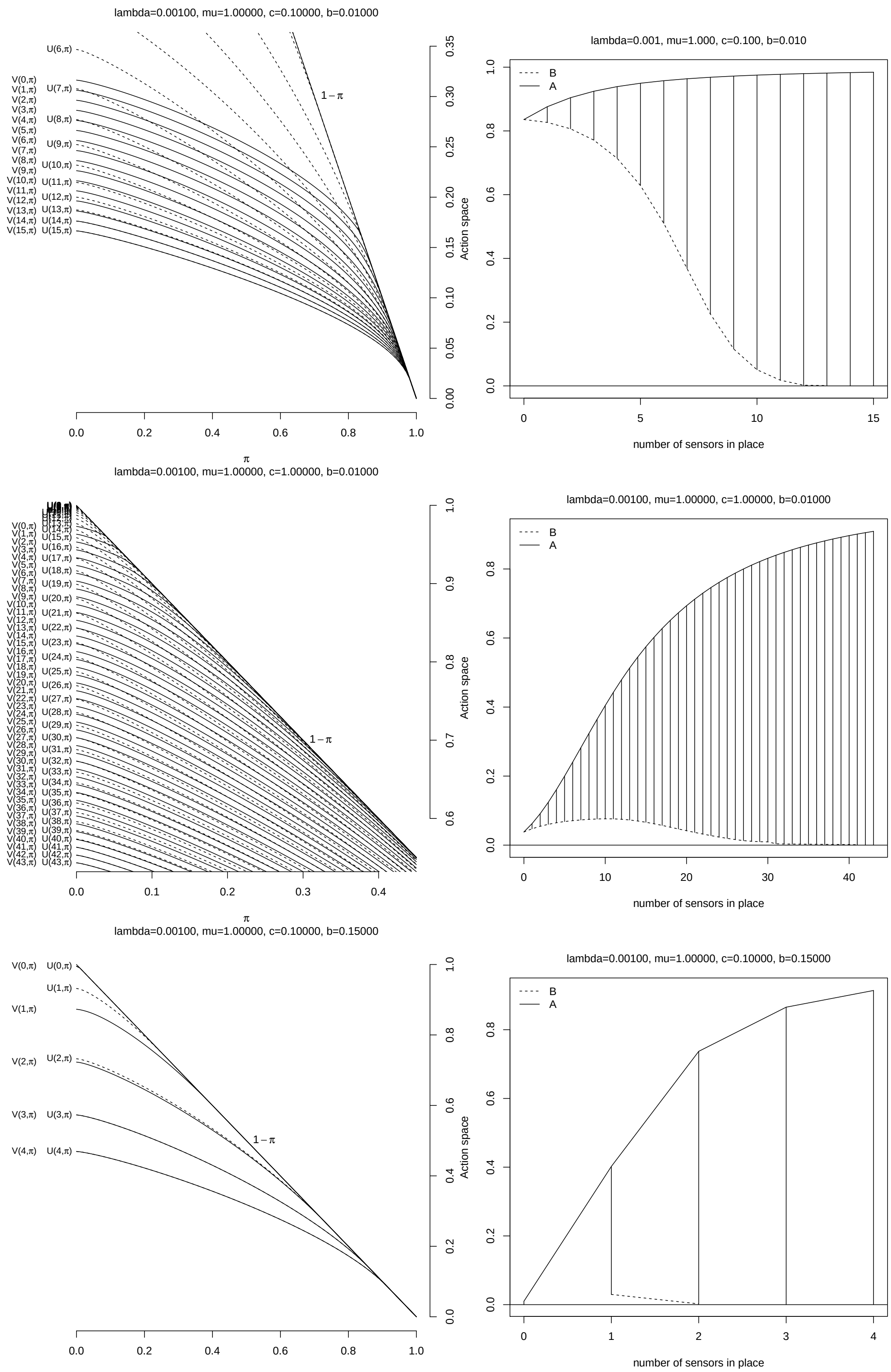
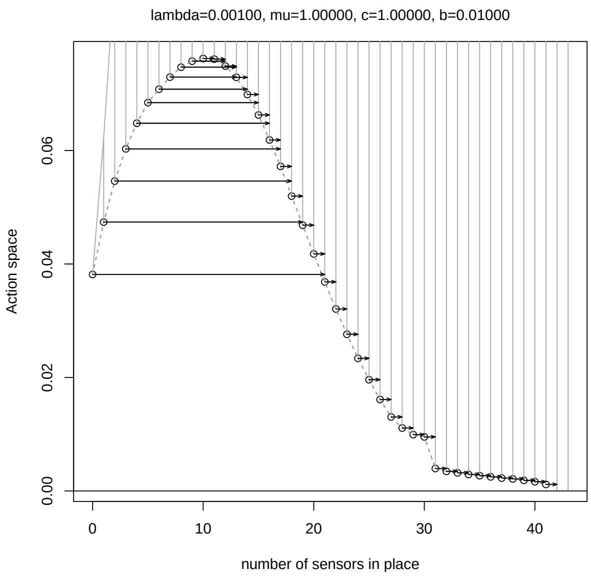



FIGURE 4. Iterations of the value functions and no-action spaces.




| Number of sensors in use | Number of sensors optimal to add |
|---|---|
| 0 | 21 |
| 1 | 18 |
| 2 | 16 |
| 3 | 14 |
| 4 | 12 |
| 5 | 10 |
| 6 | 8 |
| 7 | 6 |
| 8 | 5 |
| 9 | 3 |
| $\geq 10$ | 1 |

FIGURE 5. In Case 24 in Table 1 (also middle row in Figure 4), it is optimal to add two or more sensors instantaneously if the number of sensors in place is 0-to-9.

broken and solid curves on the left labeled, respectively, with $U(\ell,\cdot)$ and $V(\ell,0)$ as $\ell$ decreases from $L_b$ to 0).

Those three cases are especially picked to illustrate the qualitatively distinct shapes of no-action spaces, displayed on the right column of Figure 4. The vertical lines are the subintervals of $[0,1]$, the state-space of posterior probability distribution, in which it is optimal to wait according to optimal sequential sensor installation policy. Therefore, we call each of them a no-action space corresponding to the number of sensors currently in use. The righthand end-points of no-action spaces are connected by a solid curve and correspond to alarm thresholds: if the posterior probability process leaves for the first time the no-action space from the righthand end-point, then it is optimal to immediately raise a change alarm. The lefthand end-points of the no-action spaces are connected with a broken curve and correspond to add-new-sensor thresholds: if the posterior probability process leaves for the first time the no-action space from the lefthand end-point, then it is optimal to immediately add a new sensor, in which case the posterior probability process jumps instantaneously and horizontally on to the next vertical line (or its vertical extension) on the right, and this process continues.

In the upper row (Case 5), a new sensor is optimal to add if 0-to-14 sensors are in place and is not optimal to add if 15 or more sensors are already in use. The broken add-new-sensor curve is strictly decreasing: hence, the no-action spaces are nested in the sense that it is never optimal to add two or more sensors simultaneously.

The pictures of Case 24 on the second line of Figure 4 give an example for a *not-nested* no-action spaces. In this case, a new sensor is optimal to add if 0-to-42 sensors are in place and is *not* optimal to add if 43 or more sensors are already in use. More interestingly, the broken add-new-sensor curve firstly increases and then decreases. That means, the no-action spaces are *not* nested and therefore, it is sometimes optimal to add two or more sensors simultaneously. Figure 5 shows in more detail the broken add-new-sensor curve and the precise number of sensors optimal to add for each possible number of sensors in use: if currently one sensor is in place and the posterior probability processes reaches the add-new-sensor region, then the optimal action is to instantaneously install 18 new sensors. In the plot on the left of Figure 5, this action is shown by an arrow that horizontally

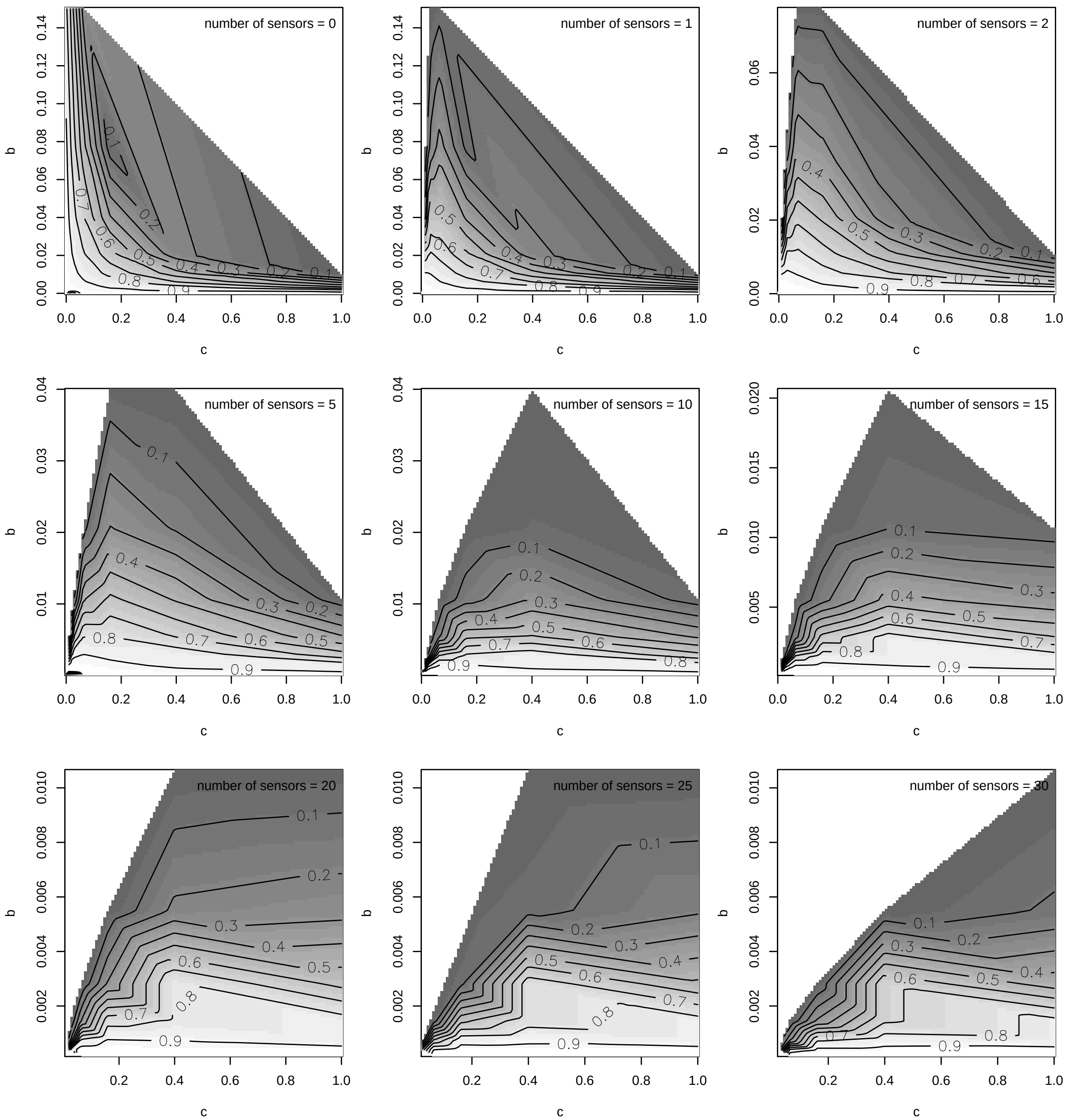



FIGURE 6. The contour lines of add-new-sensor boundary values as a function of cost parameters $c$ and $b$ for some fixed numbers of sensors in place for Case 5 in Table 1.

transfers the posterior probability process resting at the add-new-sensor boundary for one sensor in use onto the no-action space corresponding to $1+18=19$ sensors in use.

Finally, the last row of Figure 4 shows the optimal no-action space for Case 37 in Table 1. We now observe that adding a new sensor is not optimal most of the time. In particular for $\ell \geq 3$, because of the high cost of an additional sensor, we do not install any sensor and the functions $V(\ell,\cdot)$'s coincide with $U(\ell,\cdot)$'s.

**(iii) Shapes of the optimal sensor addition regions as the cost parameters $c$ and $b$ jointly change.** Figure 6 displays optimal add-new-sensor boundary values as a function of the number of sensors in use and the cost parameters $c$ and $b$. More precisely, it shows for each fixed number $\ell = 0, 1, 2, 5, 10, 15, 20, 25, 30$ of sensors in place (marked on the upper righthand corner of each panel) the contour lines of optimal add-new-sensor boundary point $B^*_\ell \equiv B^*_\ell(c, b)$ on the $(c, b)$-space for the base case example (Case 5 in Table 1). For $\ell = 0$, the contours are hyperbolic type of curves. This implies that no matter how high the value of $c$ (or $b$) is, there is always a value of $b$ (or $c$) for which new sensors are optimal to be installed. However, this structure disappears if at least one sensor is already in use. For $\ell > 0$, in order to justify additional sensor(s) installation, $b$ must be sufficiently small at every given level of $c$. In general, the contours show that, for fixed value of $c$, as $b$ increases the thresholds decrease and installation of new sensor(s) becomes less attractive.

When $b$ is fixed, in most cases, the threshold firstly increases then decreases as $c$ increases. From small to moderate values of $c$ because the problem may terminate relatively early, we may also prefer to install the sensors early. However, after some level, high values of $c$ may hurt the observer and cause her to stop the problem earlier, in which case installation of new sensors may add very little value and is increasingly discouraged. A final remark is that as the number of sensors increases, the contour lines flatten out. This suggests that the size of the region where it is optimal to add new sensors largely determined by the unit sensor cost $b$.

**7. Some extensions.** In our formulation, we assumed for simplicity that all sensors are identical devices with the same post-disorder drifts, and they yield uncorrelated signals. In practice, the observer may have different types of sensors (with different post-disorder drifts) available at different prices. This makes the problem even more difficult. In this case, it is appealing to order them based on the *unit cost per signal-to-noise* ratio. However, such an approach may not be optimal. The problem with correlated observations, on the other hand, can be reformulated in terms of non-identical sensors. If we collect observations from $\ell$-many sensors in the form

$$d\vec{X}_t = \vec{m}\, 1_{\{\theta \le t\}} dt + d\vec{W}_t, \qquad t \ge 0$$

for a column vector $\vec{m}$ of constants, and an $\ell$-dimensional correlated Brownian motion $\vec{W}$ having a positive definite correlation matrix $C$ with $S := C^{1/2}$, then, the scaling $\vec{Z}_t \triangleq S^{-1}\vec{W}_t$ yields a standard (uncorrelated) $\ell$-dimensional Brownian motion. The observations can therefore be recast as

$$d\vec{Y}_t = S^{-1}\vec{m}\, 1_{\{\theta \le t\}} dt + d\vec{Z}_t, \qquad t \ge 0,$$

as with uncorrelated sensors having possibly non-identical post-disorder drifts. It is easy to see that an observer will work with finitely many sensors at most since immediate stopping gives a uniformly bounded Bayes risk $1 - \pi \le 1$ over $\pi \in [0, 1]$. Hence it is sufficient to work with a finite dimensional matrix $S$.

We also assumed that the potential sites/locations for additional sensors are identical for the fusion center. Such an assumption allows us to obtain an optimal solution without dealing with the problem of best sensor network configuration every time a sensor is installed. Indeed, there is a separate literature on the optimal placement and deployment of static and mobile sensors in (especially wireless) sensor networks for better network coverage, reduced communication cost and energy consumption, more efficient target tracking, better data acquisition, etc. We refer the reader, for example, to Gonzalez-Banos and Latombe [17], Chakrabarty et al. [9], Zou and Chakrabarty [41], Dhillon and Chakrabarty [15], Lin and Chiu [22], Sheng et al. [29], Altinel et al. [1], Krause et al. [21], Baumgartner et al. [4] and the references therein for various models and algorithms on optimal on-site sensor deployment. Those studies also assume that the number of sensors deployed on the region of interest is determined (or given) initially at time zero. That is, the problem of sequential installation of new sensors has not been considered in that line of work, either. Therefore, we believe that the researchers in that area can also benefit from the novel formulation of the current paper.

The framework presented in the paper can also extended to the mixed Wiener and compound Poisson observation case as in Dayanik et al. [14]. This corresponds to a setting where we have two types of sensors; one yielding Brownian observations (as in the current paper), and the other providing compound Poisson observations whose compensator changes at the disorder time, say, from $\lambda_0 ds \times \nu_0(dy)$ to $\lambda_1 ds \times \nu_1(dy)$ for some $\lambda_0, \lambda_1 > 0$ and measures $\nu_0$ and $\nu_1$ on $\mathcal{B}(\mathbb{R}^d)$.

When we collect observations from $\ell_W$-many Wiener processes $X^{(1)}, \ldots, X^{(\ell_W)}$ and $\ell_P$-many point processes $(T_n^{(j)}, Y_n^{(j)})_{n\geq 1}$, $j = 1, \ldots, \ell_P$, each forming a compound Poisson processes, the conditional probability process $\Pi^\delta$ follows the dynamics

$$d\Pi_t^\delta = \lambda(1-\Pi_{t-}^\delta)\,dt + \Pi_{t-}^\delta(1-\Pi_{t-}^\delta)\Bigg\{\sum_{i=1}^{\ell_W} \mu\big[dX_{t-}^{(i)} - \mu\Pi_{t-}^\delta dt\big]$$
$$\sum_{j=1}^{\ell_P}\int_{R^d} \frac{\big[\frac{\lambda_1}{\lambda_0}\frac{d\nu_1}{d\nu_0}(z) - 1\big]}{\big[(1-\Pi_{t-}^\delta) + \Pi_{t-}^\delta \frac{\lambda_1}{\lambda_0}\frac{d\nu_1}{d\nu_0}(z)\big]}\Big[p^{(j)}(dt\times dz) - (1-\Pi_{t-}^\delta)\lambda_0 dt\,\nu_0(dz) - \Pi_{t-}^\delta \lambda_1 dt\,\nu_1(dz)\Big]\Bigg\} \tag{7.1}$$

until the next sensor installation, where $p^{(j)}$ is the counting measure

$$p^{(j)}((0,t]\times A) := \sum_{n=1}^\infty 1_{(0,t]\times A}(T_n^{(j)}, Y_n^{(i)}), \qquad t\geq 0,\ A\in\mathcal{B}(\mathbb{R}^d), \qquad j = 1,\ldots,\ell_P.$$

The dynamics in (7.1) indicate that the effect of an additional Brownian observation is on the volatility term whereas a new Poisson process alters the drift term and introduces an additional source of discontinuity. Therefore, it is not clear what type of sensor one should install when it is optimal to do so. One expects to have a solution structure in which the answer depends on the current value of the conditional probability process at this time and the number of sensors in place of each type.

The problem is three dimensional; at any time the current number of Poisson observations/sensors should also be stored. For the minimization of the Bayes risk in (1.3) with different cost of sensors $b_W$ and $b_P$ respectively for Brownian and Poisson observations, the dynamic programming operator takes the form

$$\begin{aligned} V(\ell_W, \ell_P, \pi) := \inf_\tau \mathbb{E}^\pi \Big[ \int_0^\tau c\,\Pi_t^{\delta_0} \mathrm{d}t \\ + \min\big\{1-\Pi_\tau^{\delta_0},\, b_W + V(\ell_W+1, \ell_P, \Pi_\tau^{\delta_0}),\, b_P + V(\ell_W, \ell_P+1, \Pi_\tau^{\delta_0})\big\}\Big]. \end{aligned}$$

This problem is difficult not only because the DP operator is an optimal stopping problem for a jump diffusion but also because the value function needs to be constructed on the three dimensional space.

Another direction of extension that one can explore is obtained by introducing a running cost term for sensors in the Bayes risk. If there is a cost of $\rho\geq 0$ per unit time for running each sensor, then the static problem with $\ell$-many sensors in place has the form

$$U^\rho(\ell,\pi) = \inf_{\tau\in\mathbb{F}^{\ell,0}} \mathbb{E}^\pi\left[\int_0^\tau \big(c\,\Pi_t^{\ell,\delta} + \rho\ell\big)\mathrm{d}t + 1 - \Pi_\tau^{\ell,\delta}\right],$$

which is the Lagrangian form of the constrained problem of minimizing the expected delay in detection subject to false alarm and budget constraints

$$\mathbb{P}\{\tau < \Theta\} \leq \alpha \qquad \text{and} \qquad \rho\ell\,\mathbb{E}[\tau] \leq B,$$

respectively. The latter constraint is generally omitted in the earlier work on Bayesian change detection as the focus is primarily on the timely detection of the change (i.e., finding the right balance between late detection and an early false alarm). This term rather appears in the hypothesis testing formulation of Wald and Wolfowitz [38], in which there is emphasis in the cost of collecting information. If the cost of running the sensors are significant, then it may also be included in the formulation of a change-detection problem.

Using standard standard verification arguments, one can show that

$$U^\rho(\ell,\pi) = \int_\pi^1 \min\{-\kappa^\rho(\ell,z), 1\} dz,$$

where

$$\kappa^\rho(\ell,\pi) = \left(1+\frac{\ell\rho}{c}\right)\kappa(\ell,\pi) - \frac{\ell\rho}{\lambda}.$$

We already know that $\pi \mapsto \kappa(\ell,\pi)$ is strictly decreasing with boundary conditions $\kappa(\ell,0+) = 0$ and $\kappa(\ell,1-) = -\infty$. Hence, $\pi \mapsto \kappa^\rho(\ell,\pi)$ is also strictly decreasing with boundary conditions $\kappa^\rho(\ell,0+) = -\frac{\ell\rho}{\lambda}$ and $\kappa^\rho(\ell,1-) = -\infty$.

If $\ell\rho < \lambda$ (small observation cost), the unique root $A^\rho_\ell$ of the equation $\kappa^\rho(\ell,\pi) = -1$ always exists and lies in the interval $(\frac{\lambda-\ell\rho}{\lambda+c}, 1)$. If, however, $\ell\rho \geq \lambda$ (higher observation cost), $U^\rho(\ell,\pi) = 1-\pi$ and immediate stopping is always an optimal action for all $\pi$ values. Note that, under the classical formulation with no proportional running cost, we do not have any case for which immediate stopping is optimal for all $\pi$ values.

In this new setting, the dynamic programming operator has the form

$$\mathcal{D}^\rho[f](\ell,\pi) := \inf_{\tau\in\mathbb{F}^\ell} \mathbb{E}^\pi\left[\int_0^\tau \left[c\,\Pi^\ell_t + \rho\ell\right]\mathrm{d}t + \min\left\{1-\Pi^\ell_\tau\,,\, b + f(\Pi^\ell_\tau)\right\}\right]$$

for a concave function $f$ bounded from above by $1-\pi$. The mappings $\ell \mapsto \kappa^\rho(\ell,\cdot)$ and $\ell \mapsto U^\rho(\ell,\cdot)$ are not monotone and our analysis in Section 4 does not hold as it is. Understanding what kind of solution structure this operator yields requires further analysis, which we leave for future research.

**Acknowledgments.** Savas Dayanik was supported in part by the TÜBİTAK Research Grants 109M714 and 110M610. The authors thank the anonymous referees, associate editor, and area editor for their constructive comments and suggestions, which greatly improved the presentation in the paper. We dedicate this paper to Professor Erhan Çınlar on the occasion of his 75th birthday.

## References

[1] Altinel, K., N. Aras, E. Guney, C. Ersoy. 2006. Effective coverage in sensor networks: Binary integer programming formulations and heuristics. *Proc. IEEE International Conference on Communication* **9** 4014–4019.

[2] Baron, M., A.G. Tartakovsky. 2006. Asymptotic optimality of change-point detection schemes in general continuous-time models. *Sequential Analysis* **25**(3) 257–296.

[3] Basseville, M., I. V. Nikiforov. 1993. *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, Information and System Sciences Series. Prentice Hall Inc., Englewood Cliffs, NJ.

[4] Baumgartner, K., S. Ferrari, T.A. Wettergren. 2009. Robust deployment of dynamic sensor networks for cooperative track detection. *IEEE Sensors Journal* **9**(9) 1029–1048.

[5] Beibel, M. 1996. A note on Ritov's Bayes approach to the minimax property of the CUSUM procedure. *Ann. Statist.* **24** 1804–1812.

[6] Beibel, M. 2000. A note on sequential detection with exponential penalty for the delay. *Ann. Statist.* **28** 1696–1701.

[7] Blum, R.S., S.A. Kassam, H.V. Poor. 1997. Distributed detection with multiple sensors: Part II-Advanced topics. *Proceedings of the IEEE* **85**(1) 64–799.

[8] Carmona, R., S. Dayanik. 2008. Optimal multiple stopping of linear diffusions. *Mathematics of Operations Research* **33**(2) 446–460.

[9] Chakrabarty, K., S.S. Iyengar, H. Qi, E. Cho. 2002. Grid coverage for surveillance and target location in distributed sensor networks. *IEEE Transactions on Computers* **51**(12) 1448–1453.

[10] Chamberland, J.F., V.V. Veeravalli. 2003. Decentralized detection in sensor networks. *IEEE Transactions on Signal Processing* **51**(2) 407–416.

[11] Crow, R.W., S.C. Schwartz. 1996. Quickest detection for sequential decentralized decision systems. *IEEE Transactions on Aerospace and Electronic Systems* **32**(1) 267–283.

[12] Dayanik, S. 2010. Wiener disorder problem with observations at fixed discrete time epochs. *Mathematics of Operations Research* **35**(4) 756–785.

[13] Dayanik, S., M. Ludkovski. 2007. Filling the gap between American and Russian options: adjustable regret. *Stochastics* **79**(1-2) 61–83.

[14] Dayanik, S., H.V. Poor, S.O. Sezer. 2008. Multisource Bayesian sequential change detection. *Annals of Applied Probability* **18**(2) 552–590.

[15] Dhillon, S.S., K. Chakrabarty. 2003. Sensor placement for effective coverage and surveillance in distributed sensor networks. *Proc. IEEE Wireless Communications and Networking* **3** 1609–1614.

[16] Gapeev, P., G. Peskir. 2006. The Wiener disorder problem with finite horizon. *Stochastic Processes and their Applications* **116** 1770–1791.

[17] Gonzalez-Banos, H., J-C. Latombe. 2001. A randomized art-gallery algorithm for sensor placement. *Proc. 17th Annual Symposium on Computational Geometry* 232–240.

[18] Hadjiliadis, O. 2005. Optimality of the 2-CUSUM drift equalizer rules for detecting two-sided alternatives in the Brownian motion model. *J. Appl. Probab.* **42**(4) 1183–1193.

[19] Hadjiliadis, O., G.V. Moustakides. 2006. Optimal and asymptotically optimal CUSUM rules for change point detection in the Brownian motion model with multiple alternatives. *Theory Probab. Appl.* **50**(1) 75–85.

[20] Hadjiliadis, O., H. Zhang, H.V. Poor. 2009. One-shot schemes for decentralized quickest detection. *IEEE Transactions on Information Theory* **55**(7) 3346–3359.

[21] Krause, A., A. Singh, C. Guestrin. 2008. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research* **8** 235–284.

[22] Lin, F.Y.S., P.L. Chiu. 2005. A near-optimal sensor placement algorithm to achieve complete coverage-discrimination in sensor networks. *IEEE Communication Letters* **9**(1) 43–45.

[23] Mei, Y. 2005. Information bounds and quickest change detection in decentralized decision systems. *IEEE Transactions on Information Theory* **51**(7) 2669–2681.

[24] Moustakides, G.V. 2006. Decentralized CUSUM change detection. *Proceedings of the 9th International Conference on Information Fusion, Florence, Italy* **1** 878–885.

[25] Peskir, G., A. N. Shiryaev. 2006. *Optimal Stopping and Free-boundary problems*. Birkhauser-Verlag, Lectures in Mathematics, ETH Zurich.

[26] Poor, H. V., O. Hadjiliadis. 2009. *Quickest detection*. Cambridge University Press, Cambridge.

[27] Raghavan, V., V.V. Veeravalli. 2010. Quickest change detection of a Markov process across a sensor array. *IEEE Transactions on Information Theory* **56**(4) 1961–1981.

[28] Sezer, S.O. 2010. On the Wiener disorder problem. *Annals of Applied Probability* **20**(4) 1537–1566.

[29] Sheng, S., L. Zhang, R.X. Gao. 2006. A systematic sensor-placement strategy for enhanced defect detection in rolling bearings. *IEEE Sensors Journal* **6**(5) 1346–1354.

[30] Shiryaev, A.N. 1963. On optimum methods in quickest detection problems. *Theory of Probability and its Applications* **8**(1) 22–46.

[31] Shiryaev, A.N. 1978. *Optimal stopping rules*. Springer-Verlag, Berlin.

[32] Shiryaev, A.N. 1996. Minimax optimality of the method of cumulative sums (CUSUM) in the continuous time case. *Uspekhi Mat. Nauk.* **51**(4(310)) 173–174.

[33] Tartakovsky, A., A.S. Polunchenko. 2008. Quickest changepoint detection in distributed multisensor systems under unknown parameters. *Proceedings of the 11th International Conference on Information Fusion, Cologne, Germany* **1** 878–885.

[34] Tartakovsky, A., V.V. Veeravalli. 2002. An efficient sequential procedure for detecting changes in multichannel and distributed systems. *Proceedings of the 5th International Conference on Information Fusion, Annapolis, MD* **1** 41–48.

[35] Tartakovsky, A., V.V. Veeravalli. 2004. Change-point detection in multichannel and distributed systems with applications. *In "Applications of Sequential Methodologies" (N. Mukhopadhyay, S. Datta and S.Chattopadhyay, Eds), Marcel Dekker, Inc., New York, pp. 487-521* .

[36] Tartakovsky, A., V.V. Veeravalli. 2008. Asymptotically optimal quickest change detection in distributed sensor systems. *Sequential Analysis* **27**(4) 441–475.

[37] Veeravalli, V.V. 2001. Decentralized quickest change detection. *IEEE Transactions on Information Theory* **47** 1657–1665.

[38] Wald, A., J. Wolfowitz. 1950. Bayes solutions of sequential decision problems. *Ann. Math. Statistics* **21** 82–99.

[39] Zhang, H., O. Hadjiliadis, T. Schäfer, H.V. Poor. 2014. Quickest detection in coupled system. *SIAM Journal on Control and Optimization* **52**(3) 1567–1596.

[40] Zhang, H., N. Rodosthenous, O. Hadjiliadis. 2014. Robustness of the N -CUSUM stopping rule in a wiener disorder problem. *Annals of Applied Probability* **forthcoming**. URL http://arxiv.org/abs/1410.8765.

[41] Zou, Y., K. Chakrabarty. 2003. Sensor deployment and target localization based on virtual forces. *Proc. IEEE Infocom Conference* **2** 1293–1303.