ROBUSTNESS OF THE N-CUSUM STOPPING RULE IN A WIENER DISORDER PROBLEM

By Hongzhong Zhang[‡,*] , Neofytos Rodosthenous[§]
and Olympia Hadjiliadis[¶,†]

*Columbia University[‡], Queen Mary University of London[§],
and City University of New York[¶]*

We study a Wiener disorder problem of detecting the minimum of N change-points in N observation channels coupled by correlated noises. It is assumed that the observations in each dimension can have different strengths and that the change-points may differ from channel to channel. The objective is the quickest detection of the minimum of the N change-points. We adopt a min-max approach and consider an extended Lorden's criterion, which is minimized subject to a constraint on the mean time to the first false alarm. It is seen that, under partial information of the post-change drifts and a general nonsingular stochastic correlation structure in the noises, the minimum of N cumulative sums (CUSUM) stopping rules is asymptotically optimal as the mean time to the first false alarm increases without bound. We further discuss applications of this result with emphasis on its implications to the efficiency of the decentralized versus the centralized systems of observations which arise in engineering.

1. Introduction. The problem of quickest detection has been known in the engineering literature since the 1930s. Since then there have been various analytical considerations of the quickest detection problem in a variety of models and setups (see [30] for an overview). The quickest detection problem, also known as the disorder problem, concerns the detection of a change point in the statistical behavior of a stream of sequential observations. The objective is to blanace the trade off between a small detection delay and small frequency of false alarms. Of this problem there are two main formulations, the Bayesian and the min-max. In the former the change point or disorder time is assumed to have an a priori distribution usually independent of the observation process while in the latter it is assumed to be an unknown constant. An interesting variation of the Bayesian problem in which the change point is assumed to depend on the observations is discussed in [26] and treated under Poisson dynamics in [32].

Yet in all formulations considered thus far, it is assumed that there is either one stream of observations in which there is one [7, 14, 20, 23, 24, 33] or multiple alternatives regarding the law of the post change distribution of the observations [5, 8, 9], or alternatively, multiple streams of observations of various models all undergoing a disorder at the same time [11, 25, 35, 36, 37]. In our work we assume that there are N sources of observations coupled by correlated noise. The observations are assumed to be continuous and thus a Wiener model is used. The problem considered in this work is that in which the N different streams of observations coupled by correlated noise may undergo a change at N distinct change points. The objective is then to detect the minimum of the change points or disorder times. Of this type of problem there has thus far been a

---

Bayesian formulation in independent streams of Poisson observations [6]. Recently the case was also considered of change points that propagate in a sensor array [31]. However, in this configuration the propagation of the change points depends on the *unknown* identity of the first sensor affected and considers a restricted Markovian mechanism of propagation of the change.

In this paper we consider the case in which the change points can be different and do not propagate in any specific configuration. In fact in our formulation the change points or disorder times are assumed to be unknown constants and a min-max approach to their estimation is taken. In particular we consider an extended Lorden criterion to measure the worst detection delay over all observation paths and change points. The objective is then to find a stopping rule that minimizes the detection delay subject to a lower bound constraint on the mean time to the first false alarm. The N streams of observations are coupled through correlated noise. In particular, correlations are modeled through a stochastic correlation matrix that is assumed to be non-singular and predictable. This work is a continuation of the problem considered in [18] in which the case is considered of independent observations received at each sensor. In that work, it is seen that the decentralized system of sensors in which each sensor employs its own cumulative sum (CUSUM) [30] strategy and communicates its detection through a binary asynchronous message to a central fusion center, which in turn decides at the first onset of a signal based on the first communication performs asymptotically just as well as the centralized system. In other words, the minimum of N CUSUMs is asymptotically optimal in detecting the minimum of N distinct change points in the case of independent observations as the mean time to the first false alarm increases without bound. The mean time to the first false alarm can be used as a benchmark in actual applications in which the engineer or scientist may make several runs of the system while it is in control in order to uniquely identify, the appropriate parameter that would lead to a tolerable rate of false detection. The problem of optimal detection then boils down to minimizing the detection delay subject to a tolerable rate of false alarms. Asymptotic optimality is then proven by comparing the rate of increase in detection delay to the rate of false alarms as the threshold parameter varies. A series of more recent related work includes the case in which the system of sensors is coupled through the drift parameter as opposed to the noise [17, 39]. In that work it is once again seen that the minimum of N CUSUMs is also asymptotically optimal in detecting the minimum of N distinct change points with respect to a generalized Kullback-Leibler distance criterion inspired by [24].

Yet, in none of the above cases is the case of correlated noise considered even though it is very important in practical applications. In fact there are multiple applications of this problem especially in the area of communications where sensor networks are widely used and multiple correlated streams of observations are present. The change points, usually representing the onset of a signal in a specific sensor, may well be distinct. The minimum of the change points then represents the onset of a signal in the system. The presence of correlations is due to the fact that, although sensors are placed typically at different locations, they are subject to the same physical environment. For example, in the case of sensors monitoring traffic in opposite (same) directions may have negative (positive) correlations due to environmental factors such as the direction of the wind [10]. Moreover, the appearance of a signal at one location may or may not cause interference of the signal at another location, thus causing correlations whose structure may even be time or observations dependent. This happens when the sensors are closely spaced relative to the curvature of the field being sensed. For example, temperature sensors or humidity sensors that are in a similar geographic region will produce readings that are correlated. A stochastic correlation matrix would best describe such a situation. Some of the relevant literature that includes such examples can be found in [1, 2, 3, 12, 19, 21, 28].

In an earlier work the authors in [38] treat the problem of quickest detection of the minimum of two change points in the special case of two streams of sequential observations when the correlation

in the noise of the observations is constant and negative and the same drifts are assumed after each of the disorder times. This work treats the general case of N correlated streams of observations in the presence of partial information regarding the post-change drifts which can as such be different. Moreover, we consider a general stochastic correlation matrix allowing for both positive and negative time and state dependent correlations in the system. The results found in this work are in fact rather surprising. It is seen that the minimum of N CUSUM stopping rules maintains its asymptotically optimal character as the mean time to the first false alarm increases without bound even in the case of partially known drifts and a stochastic correlation matrix coupling the noise of N streams of observations. In particular, it is proved that the N-CUSUM stopping rule (defined in Algorithm 2.1) is second order asymptotically optimal[1] in the case the post-disorder drift parameters assumed across the N streams of observations are the same, and is third order asymptotically optimal when the post-disorder drift parameters are different for an appropriately chosen set of threshold parameters whose form is explicitly given.

The method used to prove the asymptotic optimality of the N-CUSUM stopping rule is to bound the optimal detection delay from above and from below. Then we examine the rate at which the difference between the upper and the lower bounds approach each other as the mean time to the first false alarm increases without bound. This method is similar to [13, 16, 17, 18, 24, 29, 38, 39]. However, the methodology developed in this work for establishing the upper and lower bounds is more efficient and robust in that it is based on probabilistic arguments. In contrast the existing work in continuous-time, which is either relied on brute computation of the asymptotic behaviors of maximum drawdown densities [18] or on the derivation of sharp solutions to Dirichlet problems with Neumann conditions [17, 24, 38, 39], is very difficult in high-dimension and highly sensitive to the model parameters. The methodology developed in this paper is universal and can thus handle a non-Markovian, predictable correlation matrix process for the noises, which is very useful in practical applications. Finally, our methodology can be applied to other detection problems not covered in this paper, for example quickest detection with multiple alternatives [15, 16]. In establishing the lower bound, we give a non-trivial generalization of a measure change technique developed in [24] to N-dimensions. Although we don't get the exact optimality as in one dimension [24], we do prove that the optimal detection delay in N-dimensions is bounded from below by that obtained in one dimension, under any predictable, nonsingular correlation matrix.

In the next section we formulate the problem mathematically, review the existing results in one dimension, and introduce the N-CUSUM stopping rule. In Section 3, we establish a robust upper bound and a robust lower bound for both the optimal detection delay and the detection delay of the N-CUSUM stopping rule. These bounds are then used in Section 4 to show the main result of the paper - the asymptotic optimality of the N-CUSUM stopping rule under complete or partial information of the drifts and a stochastic cross-correlated noise structure in the observations. Applications of these results are discussed in Section 5. We conclude with some closing remarks in Section 6. The proof of the lemma that is omitted can be found in the Appendix.

Throughout the paper, we denote by $s \wedge t = \min\{s, t\}$, $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_+ = [0, \infty)$ and $\bar{\mathbb{R}}_+ = [0, \infty]$.

2. Formulation of the problem. Consider a filtered probability space $(\Omega, \overline{\mathcal{F}}, \overline{\mathbb{F}}, P)$ with filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t \geq 0}$, and the processes $\xi^{(i)} := \{\xi_t^{(i)}\}_{t \geq 0}$, $i = 1, \ldots, N$, are assumed to satisfy the following stochastic differential equations:

$$(2.1) \qquad d\xi_t^{(i)} = \mu_i \mathbf{1}_{\{t \geq \tau_i\}} dt + dw_t^{(i)}.$$

---

[1]See Definition 2.1 below.

Here $\{\tau_i\}_{1\leq i\leq N}$ are deterministic but *unknown* positive constants or ∞, $\{\mu_i\}_{1\leq i\leq N}$ are positive constants[2] that are either completely known or partially known. In the latter case, we assume that $\mu_1 > 0$ is a known constant, and for $i = 2,\ldots,N$, there are known positive constants $\mu_1 \leq \underline{\mu}_i \leq \overline{\mu}_i$ such that $\mu_i \in [\underline{\mu}_i, \overline{\mu}_i]$ holds.[3] The processes $\{w^{(i)}\}_{1\leq i\leq N}$ for $w^{(i)} := \{w_t^{(i)}\}_{t\geq 0}$ are N correlated standard Brownian motions with a predictable, non-singular, stochastic instantaneous correlation matrix $\Sigma_t = (\rho_t^{i,j})$. That is, $\rho_t^{i,j}$ is the instantaneous correlation between Brownian motions $w^{(i)}$ and $w^{(j)}$ (see also [34, page 227]).

An example covered by the above assumptions is one in which $\rho_t^{i,j} = \rho e^{-t}$ for $i \neq j$ and some $\rho \in (0,1)$. In other words, there is a deterministic exponential decay in the instantaneous correlation of the two sensors i and j. Such a situation may arise by the sudden arrival of a passing rainstorm at sensors i and j, which are customarily placed in the same geographical region and are therefore also subject to the same climet conditions. Yet our formulation is even more general in that it is also able to capture state dependent correlations which is a very realistic scenario since observations of higher intensity are typically more likely to cause higher correlations in the noise, for instance $N = 2$ and $\rho_t^{i,j} = \frac{\xi_t^{(i)}\xi_t^{(j)}}{1+|\xi_t^{(i)}\xi_t^{(j)}|}e^{-t}$ for $i \neq j$. Another example of a correlated non-stationary white noise structure arises in the problem of monitoring the vibration of a mechanical system and is discussed in full detail in subsection 11.1.4.1 of [4].

To facilitate our analysis, we introduce a family of probability measures on the canonical space $(C(\mathbb{R}_+^N), \mathbb{F})$: $\{\mathbb{P}_{s_1,\ldots,s_N}\}_{(s_1,\ldots,s_N)\in(\bar{\mathbb{R}}_+)^N}$. Here $\mathbb{P}_{s_1,\ldots,s_N}$ corresponds to the measure generated on $C(\mathbb{R}_+^N)$ by the processes $(\xi^{(1)},\ldots,\xi^{(N)})$ when the change in the N-tuple process occurs at the time points $\tau_i = s_i$, $1 \leq i \leq N$, respectively. In particular, the measure $\mathbb{P}_{\infty,\ldots,\infty}$ characterizes the law of N correlated standard Brownian motions $\{w^{(i)}\}_{1\leq i\leq N}$. For other s_i's, the measure $\mathbb{P}_{s_1,\ldots,s_N}$ can be defined through the Radon-Nikodym derivative process $\frac{d\mathbb{P}_{s_1,\ldots,s_N}}{d\mathbb{P}_{\infty,\ldots,\infty}}\big|_{\mathcal{F}_t}$. To this end, we assume that the correlation matrix Σ_t fulfills the Novikov condition:

$$(2.2) \qquad \mathbb{E}_{\infty,\ldots,\infty}\left\{\exp\left(\frac{1}{2}\langle\log(\frac{d\mathbb{P}_{s_1,\ldots,s_N}}{d\mathbb{P}_{\infty,\ldots,\infty}}\big|_{\mathcal{F}_t})\rangle\right)\right\} < \infty, \forall t \geq 0, \forall(s_1,\ldots,s_N) \in (\bar{\mathbb{R}}_+)^N.$$

We comment that the "reality" measure $\mathbb{P}_{\tau_1,\ldots,\tau_N}$ is one unknown element in $\{\mathbb{P}_{s_1,\ldots,s_N}\}_{(s_1,\ldots,s_N)\in(\bar{\mathbb{R}}_+)^N}$.

To describe the "marginal" law of the i-th component of the N-tuple process $(\xi^{(1)},\ldots,\xi^{(N)})$, we also introduce the measure $\{\mathbb{P}_{s_i}^i\}$, which is the probability measure generated by the process $\xi^{(i)}$ on the space $(C(\mathbb{R}), \mathbb{G}^{(i)})$, where $\mathbb{G}^{(i)} = \{\mathcal{G}_t^{(i)}\}_{t\geq 0}$ for $\mathcal{G}_t^{(i)} = \sigma\{(\xi_s^{(i)}); s \leq t\}$, is the natural filtration of $\xi^{(i)}$, and $\tau_i = s_i$ is the value of the change-point for process $\xi^{(i)}$.

Our objective is to find a stopping rule T, which is adapted to the natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$: $\mathcal{F}_t = \sigma(\xi_s^{(1)},\ldots,\xi_s^{(N)}; s \leq t)$,[4] to balance the trade-off between a small detection delay subject to a lower bound on the mean-time to the first false alarm and will ultimately detect $\tau_1 \wedge \tau_2 \wedge \ldots \wedge \tau_N$, which will be denoted by $\tilde{\tau}$ in what follows. As a performance measure we consider

$$(2.3) \qquad J^{(N)}(T) = \sup_{\substack{(s_1,\ldots,s_N)\in\bar{\mathbb{R}}_+^N \\ \tilde{s}<\infty}} \operatorname{essup} \mathbb{E}_{s_1,\ldots,s_N}\left\{(T-\tilde{s})^+ \mid \mathcal{F}_{\tilde{s}}\right\}$$

---

[2]The condition can be relaxed. For example, if we know a priori that $\mu_i < 0$ (but not necessarily the value of it), then we can take $-\xi^{(i)}$ as the i-th observation process so that the post-change drift is $-\mu_i > 0$. We do not treat in this paper, however, the case in which we don't know the sign of the post-change drift.

[3]If μ_i is known, we can conveniently take $\underline{\mu}_i = \overline{\mu}_i = \mu_i$.

[4]Note that Σ needs not to be adapted to $\mathbb{F}$. For example, Σ can be driven by a N-dimensional Brownian motion which is independent of $w^{(i)}$'s.

where $\tilde{s} = s_1 \wedge s_2 \wedge \ldots \wedge s_N$, $\mathbb{E}_{s_1,\ldots,s_N}$ denotes the expectation under the probability measure $\mathbb{P}_{s_1,\ldots,s_N}$, and the supremum over $s_1,\ldots,s_N$ is taken over the set in which $\tilde{s} < \infty$. In other words, we consider the worst detection delay over all possible realizations of paths of the N-tuple of the stochastic process $\{(\xi_t^{(1)},\ldots,\xi_t^{(N)})\}_{t\geq 0}$ up to time $\tilde{s}$, and then consider the worst detection delay over all possible N-tuples $(s_1,\ldots,s_N)$ over a set in which at least one of the components takes a finite value. This is because T is a stopping rule meant to detect the minimum of the N change points and therefore if one of the N processes undergoes a regime change, any unit of time by which T delays in reacting, should be counted towards the detection delay. Although it seems to be a quite pessimistic measure for detection delay, this framework has the merit that one do not need to impose any prior knowledge of the distribution of the change-points τ_i's, as is discussed in [26]. In all, this gives rise to the following stochastic optimization problem:

$$
(2.4) \qquad
\begin{aligned}
&\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) \\
&\text{with } \mathcal{T}_\gamma = \{\mathbb{F}\text{-stopping rule } T \; : \; \mathbb{E}_{\infty,\ldots,\infty}\{T\} \geq \gamma\}
\end{aligned}
$$

where $\mathbb{E}_{\infty,\ldots,\infty}\{T\}$ captures the mean time to the first false alarm and as such the above constraint describes the tolerance on the false alarms. In particular, the constant $\gamma > 0$ is the lowest acceptable value of the mean time to the first false alarm. In other words, the reciprocal of γ, namely $\frac{1}{\gamma}$, captures the highest tolerance to the frequency of false alarms of the family of stopping times considered in this problem.

When detecting τ_i is our only concern, and that μ_i is a known constant, the problem reduces to an one-dimensional problem of detecting a one-sided change in a sequence of Brownian motion observations, whose optimality was found in [7] and [33]. It is shown that the optimal stopping rule under Lorden's criterion is the continuous time version of Page's CUSUM stopping rule, namely the first passage time of the process

$$
(2.5) \qquad
\begin{aligned}
\tilde{y}_t^{(i)} &:= \sup_{0\leq s\leq t} \left.\frac{\mathrm{d}\mathbb{P}_s^i}{\mathrm{d}\mathbb{P}_\infty^i}\right|_{\mathcal{G}_t^{(i)}} = \tilde{u}_t^{(i)} - \tilde{m}_t^{(i)}, \\
&\text{for} \quad \tilde{u}_t^{(i)} := \mu_i \xi_t^{(i)} - \tfrac{1}{2}\mu_i^2 t \quad \text{and} \quad \tilde{m}_t^{(i)} := \inf_{0\leq s\leq t} \tilde{u}_s^{(i)},
\end{aligned}
$$

and the CUSUM stopping rule with the threshold $\nu_i^\star > 0$ is given by

$$
(2.6) \qquad \tilde{T}_{\nu_i^\star}^i = \inf\{t \geq 0 \; : \; \tilde{y}_t^{(i)} \geq \nu_i^\star\}.
$$

The optimal threshold $\nu_i^\star$ is chosen so that,

$$
(2.7) \qquad \mathbb{E}_\infty^i\{\tilde{T}_{\nu_i^\star}^i\} = (2/\mu_i^2)\, g(\nu_i^\star) = \gamma, \text{ where } g(\nu) := \mathrm{e}^\nu - \nu - 1, \quad \forall \nu > 0.
$$

The corresponding optimal detection delay achieved by the CUSUM stopping rule $\tilde{T}_{\nu_i^\star}^i$ is then given by

$$
(2.8) \qquad J^{(1)}(\tilde{T}_{\nu_i^\star}^i) = \mathbb{E}_0^i\{\tilde{T}_{\nu_i^\star}^i\} = \frac{2}{\mu_i^2}\, g(-\nu_i^\star).
$$

The fact that the worst detection delay in the one-dimensional problem is the same as that incurred in the case that the change point is exactly at 0 is a consequence of the non-negativity and strong Markov property of the CUSUM process, from which it follows that the worst detection delay occurs when the CUSUM process is at 0 at the time of the change (see also [23]).

The optimality of the CUSUM stopping rule in the presence of only one observation process with a known drift suggests that a CUSUM type of stopping rule might display similar optimality

properties in the case of multiple observation processes for the problem (2.4). In particular, an intuitively appealing rule, when the detection of $\tilde{\tau}$ is of interest, is to take the minimum of N CUSUM-like stopping rules (see, e.g. [15]), which we formalize in the following algorithm.

ALGORITHM 2.1.   *The N-CUSUM stopping rule with a threshold vector $\hbar = (h_1, \ldots, h_N) \in (\mathbb{R}_+)^N$ is given by $T_\hbar = T_{h_1}^1 \wedge T_{h_2}^2 \wedge \ldots \wedge T_{h_N}^N$, where for each $i = 1, \ldots, N$,*

$$(2.9) \qquad \begin{aligned} T_{h_i}^i &= \inf\{t \geq 0 \,:\, y_t^{(i)} \geq h_i\}, \text{ with } y_t^{(i)} = u_t^{(i)} - m_t^{(i)}, \\ &\text{for } u_t^{(i)} := \underline{\mu}_i \xi_t^{(i)} - \tfrac{1}{2}\underline{\mu}_i^2 t \text{ and } m_t^{(i)} := \inf_{0 \leq s \leq t} u_t^{(i)}. \end{aligned}$$

That is, we use what is known as a multi-chart CUSUM stopping rule [25], which can be written as

$$T_\hbar \;=\; \inf\left\{ t \geq 0 \,:\, \max\left\{ \frac{y_t^{(1)}}{h_1}, \ldots, \frac{y_t^{(N)}}{h_N} \right\} \geq 1 \right\},$$

where $\{y_t^{(i)}\}_{t \geq 0}$ is the semi-martingale defined in (2.9), for $i = 1, \ldots, N$. We notice that each of the $T_{h_i}^i$, for $i = 1, \ldots, N$, are stopping rules also with respect to each of the smaller filtrations $\mathbb{G}^{(i)}$, and thus they can be employed by each one of the sensors S_i, for each i independently. Each of the sensors can then subsequently communicate an alarm to a central fusion center once its threshold, say h_i, is reached by its own CUSUM statistic process $y^{(i)}$. The resulting rule, namely Algorithm 2.1, can then be devised by the central fusion center in that it will declare a detection at the first instance one of the N sensors communicates.

REMARK 2.1.   *From (2.5) and (2.9), it is easily seen that $y^{(i)} \equiv \tilde{y}^{(i)}$ and $T_{h_i}^i = \tilde{T}_{h_i}^i$, a.s., provided that $\mu_i = \underline{\mu}_i$ is known. In particular, we always have $y^{(1)} \equiv \tilde{y}^{(1)}$ and $T_{h_1}^1 = \tilde{T}_{h_1}^1$.*

While it seems prohibitively difficult to devise a stopping rule that achieves the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ under a general nonsingular correlation matrix $(\Sigma_t)_{t \geq 0}$, the above N-CUSUM stopping rule $T_\hbar$ provides a low-complexity candidate detection rule for detecting $\tilde{\tau}$.

In particular, we will show that the N-CUSUM stopping rule is asymptotically optimal. To this effect we give the following definitions of asymptotic optimality as in [13].

DEFINITION 2.1.   *Given $\gamma > 0$ and a stopping time $T' \in \mathcal{T}_\gamma$, we say that,*

1. T' has the first order asymptotic optimality for problem (2.4) if and only if

$$\lim_{\gamma \to \infty} \frac{J^{(N)}(T')}{\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)} = 1 \text{ and } \lim_{\gamma \to \infty} \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) = \infty.$$

2. T' has the second order asymptotic optimality for problem (2.4) if and only if

$$\lim_{\gamma \to \infty}[J^{(N)}(T') - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)] < \infty \text{ and } \lim_{\gamma \to \infty} \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) = \infty.$$

3. T' has the third order asymptotic optimality for problem (2.4) if and only if

$$\lim_{\gamma \to \infty}[J^{(N)}(T') - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)] = 0 \text{ and } \lim_{\gamma \to \infty} \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) = \infty.$$

Below we will investigate the performance of $T_\hbar$ by contrasting it with the optimal detection delay.

3. Robust bounds for the optimal detection delay. In this section, we examine the performance of the N-CUSUM stopping rule by presenting an upper bound and a lower bound for both the detection delay of the N-CUSUM stopping rule and the optimal detection delay defined in (2.4). To this end, we derive a robust upper bound for the detection delay of a particular N-CUSUM stopping rule $T_\hbar$ in $\mathcal{T}_\gamma$. Because $T_\hbar$ cannot beat the unknown optimal stopping rule (if it ever exists), this upper bound will also bound the optimal detection delay from above. We then demonstrate that the optimal detection delay in the N-dimensional system is bounded from below by the optimal delay in 1-dimensional systems.

3.1. *The upper bound.* In this subsection, we derive a robust upper bound for the detection delay of a N-CUSUM stopping rule $T_\hbar$, whose thresholds set $\hbar$ is chosen so that $T_\hbar \in \mathcal{T}_\gamma$ for any $\gamma > 0$. The upper bound, that we obtain, also dominates the optimal detection delay, due to the fact that $J^{(N)}(T_\hbar) \geq \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ holds.

Now let us introduce

$$(3.1) \qquad J_j^{(N)}(T) \;=\; \sup_{\substack{(s_1,\ldots,s_N)\in\bar{\mathbb{R}}_+^N \\ s_j = \tilde{s} < \infty}} \operatorname*{essup} \mathbb{E}_{s_1,\ldots,s_N}\big\{(T - s_j)^+ \,\big|\, \mathcal{F}_{s_j}\big\},$$

for $j = 1, \ldots, N$, where $J_j^{(N)}(T)$ is the detection delay of the stopping rule T when $s_j \leq \min_{i\neq j}\{s_i\}$, implying that the performance measure defined in (2.3) is given by $J^{(N)}(T) = \max_{1\leq j\leq N} J_j^{(N)}(T)$. We now consider the case when all drifts μ_i's are known constants. In this case, we select $\hbar$ such that,

$$(3.2) \qquad \mathbb{E}_0^1\{T_{h_1}^1\} = \mathbb{E}_0^2\{T_{h_2}^2\} = \ldots = \mathbb{E}_0^N\{T_{h_N}^N\},$$

or equivalently (by (2.8)),

$$\frac{1}{\mu_1^2}g(-h_1) = \frac{1}{\mu_2^2}g(-h_2) = \ldots = \frac{1}{\mu_N^2}g(-h_N).$$

Due to the monotonicity of function g, h_i's are uniquely determined once $h_1 > 0$ is given. In general, if we only have partial information about μ_i's for $i = 2, \ldots, N$, we instead consider

$$(3.3) \qquad \frac{1}{\mu_1^2}g(-h_1) = \frac{1}{\underline{\mu}_2^2}g(-h_2) = \ldots = \frac{1}{\underline{\mu}_N^2}g(-h_N).$$

By choosing the N-CUSUM stopping rule $T_\hbar$ in this way, we are able to get an easily computable upper bound for the worst detection delay $J^{(N)}(T_\hbar)$. The assertion is proved in the following proposition.

PROPOSITION 3.1. *Suppose that $\hbar \in \mathbb{R}_+^N$ satisfies the equations in* (3.2) *or* (3.3), *then we have for the N-CUSUM stopping rule $T_\hbar$, that*

$$(3.4) \qquad J^{(N)}(T_\hbar) \leq \mathbb{E}_0^1\{T_{h_1}^1\} = \frac{2}{\mu_1^2}\,g(-h_1),$$

where the function g is defined in (2.7).

PROOF. For any $(s_1, \ldots, s_N) \in \bar{\mathbb{R}}_+^N$ such that $s_j \leq \min_{i \neq j}\{s_i\}$ and $s_j < \infty$, we have

$$(3.5) \quad \mathbb{E}_{s_1,\ldots,s_N}\big\{(T_\hbar - s_j)^+ \,\big|\, \mathcal{F}_{s_j}\big\} = \mathbb{E}_{s_1,\ldots,s_N}\big\{(T_{h_1}^1 \wedge \ldots \wedge T_{h_N}^N - s_j)^+ \,\big|\, \mathcal{F}_{s_j}\big\}$$

$$\leq \mathbb{E}_{s_1,\ldots,s_N}\big\{(T_{h_j}^j - s_j)^+ \,\big|\, \mathcal{F}_{s_j}\big\} = \mathbb{E}_{s_j}^j\big\{(T_{h_j}^j - s_j)^+ \,\big|\, \mathcal{G}_{s_j}^j\big\}$$

where the last equality follows from the fact that the CUSUM stopping rule $T_{h_j}^j$ is $\mathbb{G}^{(j)}$-measurable and we can thus use the "marginal" law of the j-th component of the N-tuple process $(\xi^{(1)}, \ldots, \xi^{(N)})$, given in this case by the measure $\mathbb{P}_{s_j}^j$. By taking the essential supremum and then the supremum over $s_1, \ldots, s_N$ such that $s_j \leq \min_{i \neq j}\{s_i\}$ on both sides of (3.5), and using the definitions in (2.3) and (3.1) for $N = 1$, we get that

$$(3.6) \qquad\qquad J_j^{(N)}(T_\hbar) \leq J^{(1)}(T_{h_j}^j) = \sup_{s_j < \infty} \operatorname{essup}\mathbb{E}_{s_j}^j\big\{(T_{h_j}^j - s_j)^+|\mathcal{G}_{s_j}^j\big\}.$$

To get the conditional expectation in the above expression, we use the strong Markov property of the processes $y_t^{(j)}$ (see, e.g. [27, Theorem 7.2.4]) and apply Itô's formula to $\{g(-y_t^{(j)})\}_{s_j \leq t < T_{h_j}^j}$ (see, e.g. [27, Theorem 4.1.2]) for the function g given in (2.7) (see also Shiryaev [33] and Moustakides [24]), we obtain that (by the monotonicity of g)

$$(3.7) \quad g(-h_j) \geq \mathbf{1}_{\{T_{h_j}^j > s_j\}}[g(-y_{T_{h_j}^j}^{(j)}) - g(-y_{s_j}^{(j)})]$$

$$= \mathbf{1}_{\{T_{h_j}^j > s_j\}}\bigg[\int_{s_j}^{T_{h_j}^j} \underline{\mu}_j\Big(\mu_j - \frac{1}{2}\underline{\mu}_j\Big)\mathrm{d}s - \int_{s_j}^{T_{h_j}^j} g'(-y_s^{(j)})\mathrm{d}m_s^{(j)} + M_{T_{h_j}^j} - M_{s_j}\bigg],$$

where the process $m^{(j)}$ is given by (2.9) and the continuous square integrable martingale $M = \{M_t\}_{t \geq 0}$ is given by

$$M_t = \int_0^{t \wedge T_{h_j}^j} \underline{\mu}_j g'(-y_s^{(j)})\mathrm{d}w_s^{(j)}.$$

Taking into account that the process $m^{(j)}$ decreases only on the random set $\{t \geq 0 : y_t^{(j)} = 0\}$ and the measure $\mathrm{d}m_t^{(j)} = 0$ off this set, together with the fact that $g'(0) = 0$, we conclude that the integral in (3.7) can be set equal to zero. We then take the conditional expectations with respect to the probability measure $\mathbb{P}_{s_j}^j$ given $\mathcal{G}_{s_j}^j$ in (3.7) and by means of the Doob's optional sampling theorem (see, e.g. [22, Chapter 1, Theorem 3.22]), we have

$$(3.8) \qquad g(-h_j) \geq \mathbf{1}_{\{T_{h_j}^j > s_j\}}\mathbb{E}_{s_j}^j\bigg\{ \int_{s_j}^{T_{h_j}^j} \underline{\mu}_j\Big(\mu_j - \frac{1}{2}\underline{\mu}_j\Big)\mathrm{d}s \,\bigg|\, \mathcal{G}_{s_j}^j\bigg\} \geq \frac{\underline{\mu}_j^2}{2}\mathbb{E}_{s_j}^j\big\{(T_{h_j}^j - s_j)^+|\mathcal{G}_{s_j}^j\big\}.$$

Therefore, by (3.6) and (3.8) we have $J_j^{(N)}(T_\hbar) \leq \frac{2}{\underline{\mu}_j^2}g(-h_j)$. By the arbitrariness of j, we have

$$(3.9) \qquad J^{(N)}(T_\hbar) = \max_{1 \leq j \leq N} J_j^{(N)}(T_\hbar) \leq \max\Big\{\frac{2}{\underline{\mu}_1^2}g(-h_1), \ldots, \frac{2}{\underline{\mu}_N^2}g(-h_N)\Big\} = \frac{2}{\underline{\mu}_1^2}g(-h_1),$$

where the last equality is a consequence of the equations in (3.3).   □

The condition (3.3) reduces the thresholds' selection problem from N dimension to one dimension. In order to bound the optimal detection delay in (2.4) using the result in Proposition 3.1, we will choose h_1 so that the resulting N-CUSUM stopping rule $T_\hbar \in \mathcal{T}_\gamma$. That is,

$$(3.10) \qquad\qquad\qquad\qquad \mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} \geq \gamma.$$

To this end, we derive a lower bound for the mean time to the first false alarm $\mathbb{E}_{\infty,\ldots,\infty}\{T_h\}$, which is robust with respect to the covariance matrix $(\Sigma_t)_{t\geq 0}$. In the sequel we first study the case of equal drift size with complete information, where $0 < \mu_1 = \mu_2 = \ldots = \mu_N = \mu$ are known constant, and then treat the case of unequal drift size with complete information, such that μ_i's are all known and $0 < \mu_1 = \mu_2 = \ldots = \mu_k < \min_{i>k} \mu_i$ holds for some $k \in \{1,\ldots,N-1\}$. Finally, we study the general case with partial information, where we only know μ_1 and the intervals $[\underline{\mu}_i, \overline{\mu}_i] \ni \mu_i$ for all $i = 2,\ldots,N$.

3.1.1. *Equal drift case - complete information about μ_i's.* In this case, it is assumed that all μ_i' are known and $\mu_i = \mu_1 = \mu > 0$ for all $i = 1,\ldots,N$. Then the monotonicity of function g and (3.2) imply that $h_1 = h_2 = \ldots = h_N = h$ for the N-CUSUM stopping rule. Hence, with a slight abuse of notation, we denote by $T_h = T_\hbar$. Below we derive a lower bound for the mean time of the first false alarm of the latter.

PROPOSITION 3.2. *Suppose that all thresholds of the N-CUSUM stopping rule are chosen to be equal to $h > 0$. Then the first false alarm for the N-CUSUM stopping rule T_h satisfies*

$$(3.11) \qquad \mathbb{E}_{\infty,\ldots,\infty}\{T_h\} \geq \frac{1}{N}\mathbb{E}_\infty^1\{T_h^1\} = \frac{2}{N\mu^2}\, g(h),$$

where the function g is defined in (2.7).

PROOF. For any $i = 1,\ldots,N$, we have

$$(3.12) \qquad \begin{aligned} \mathbb{E}_\infty^i\{T_h^i\} &= \mathbb{E}_{\infty,\ldots,\infty}\{T_h^i\} = \mathbb{E}_{\infty,\ldots,\infty}\{T_h\} + \mathbb{E}_{\infty,\ldots,\infty}\big\{(T_h^i - T_h)\,\mathbf{1}_{\{T_h^i \neq T_h\}}\big\} \\ &= \mathbb{E}_{\infty,\ldots,\infty}\{T_h\} + \mathbb{E}_{\infty,\ldots,\infty}\big\{\mathbb{E}_{\infty,\ldots,\infty}\{T_h^i - T_h \,|\, \mathcal{F}_{T_h}\}\,\mathbf{1}_{\{T_h^i \neq T_h\}}\big\}, \end{aligned}$$

where the third equality follows from the tower property of the conditional expectation and the finiteness of $T_\hbar$.

As in the proof of Proposition 3.1, we apply Itô's formula to $\{g(y_t^{(i)})\}_{T_h \leq t < T_h^i}$ (see, e.g. [27, Theorem 4.1.2]) to obtain that

$$(3.13) \qquad g\big(y_{T_h^i}^{(i)}\big) - g\big(y_{T_h}^{(i)}\big) = \frac{\mu^2}{2}(T_h - T_h^i) - \int_{T_h}^{T_h^i} g'(y_s^{(i)})\mathrm{d}m_s^{(i)} + M_{T_h^i} - M_{T_h}$$

where the process $m^{(i)}$ is given by (2.5) and the continuous square integrable martingale $M = \{M_t\}_{t\geq 0}$ (with respect to $\mathbb{P}_{\infty,\ldots,\infty}$) is given by

$$(3.14) \qquad M_t = \mu \int_0^{t \wedge T_h^i} g'(y_s^{(i)})\mathrm{d}w_s^{(i)}$$

Taking into account that the process $m^{(i)}$ decreases only on the random set $\{t \geq 0 \,:\, y_t^{(i)} = 0\}$ and the measure $\mathrm{d}m_t^{(i)} = 0$ off this set, together with the fact that $g'(0) = 0$, we conclude that the integral in (3.13) can be set equal to zero. We then take the conditional expectations with respect to the probability measure $\mathbb{P}_{\infty,\ldots,\infty}$ in (3.13) and by means of the Doob's optional sampling theorem (see, e.g. [22, Chapter 1, Theorem 3.22]), we have

$$(3.15) \qquad \mathbb{E}_{\infty,\ldots,\infty}\{g\big(y_{T_h^i}^{(i)}\big) - g\big(y_{T_h}^{(i)}\big) \,|\, \mathcal{F}_{T_h}\} = \frac{\mu^2}{2}\mathbb{E}_{\infty,\ldots,\infty}\{T_h^i - T_h \,|\, \mathcal{F}_{T_h}\}$$

Therefore, using equation (3.15) in the expression of (3.12) we have that

$$(3.16) \qquad \frac{2}{\mu^2} g(h) = \mathbb{E}_\infty^i \{T_h^i\}$$

$$= \mathbb{E}_{\infty,\dots,\infty} \{T_h\} + \mathbb{E}_{\infty,\dots,\infty} \left\{ \frac{2}{\mu^2} \left(g(y_{T_h^i}^{(i)}) - g(y_{T_h}^{(i)}) \right) \mathbf{1}_{\{T_h^i \neq T_h\}} \right\}$$

$$= \mathbb{E}_{\infty,\dots,\infty} \{T_h\} + \mathbb{E}_{\infty,\dots,\infty} \left\{ \frac{2}{\mu^2} \left(g(h) - g(y_{T_h}^{(i)}) \right) \mathbf{1}_{\{T_h^i \neq T_h\}} \right\}$$

$$\leq \mathbb{E}_{\infty,\dots,\infty} \{T_h\} + \frac{2}{\mu^2} g(h) \, \mathbb{P}_{\infty,\dots,\infty} \big(T_h^i \neq T_h \big),$$

where the first equality and the function g are given by (2.7) and the third equality follows from the definition of the one-dimensional CUSUM stopping rule in (2.9). It follows that,

$$\frac{\mu^2}{2} \mathbb{E}_{\infty,\dots,\infty} \{T_h\} \geq g(h) \, \mathbb{P}_{\infty,\dots,\infty} (T_h = T_h^i).$$

Hence by summing both sides over all $i = 1, \dots, N$, we get

$$\frac{N\mu^2}{2} \mathbb{E}_{\infty,\dots,\infty} \{T_h\} \geq g(h) \sum_{i=1}^N \mathbb{P}_{\infty,\dots,\infty} (T_h = T_h^i)$$

$$\geq g(h) \mathbb{P}_{\infty,\dots,\infty} (T_h = T_h^i \text{ for some } i \in \{1,\dots,N\}) = g(h),$$

which completes the proof of (3.11). $\qquad\square$

As a result of Proposition 3.2, when $\mu_i = \mu$ for all $i = 1, \dots, N$, for any $\gamma > 0$ and any N-dimensional, predictable, non-singular, stochastic instantaneous correlation matrix Σ_t, we can choose the threshold h using

$$(3.17) \qquad \mathbb{E}_\infty^1 \{T_h^1\} \equiv \frac{2}{\mu^2} (e^h - h - 1) = N\gamma.$$

Then we will have $T_\hbar \in \mathcal{T}_\gamma$. Moreover, Proposition 3.1 implies that, both the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ and the detection delay of this N-CUSUM stopping rule $J^{(N)}(T_h)$, are bounded above by $\frac{2}{\mu^2} g(-h)$.

3.1.2. *Unequal drift case - complete information about μ_i's.* In this case, it is assumed that all μ_i' are known and $0 < \mu_1 = \mu_2 = \dots = \mu_k < \min_{i>k} \mu_i$ holds for some $k \in \{1, \dots, N-1\}$. Then the monotonicity of function g and (3.2) imply that $h_1 = h_2 = \dots = h_k$, for the N-CUSUM stopping rule $T_\hbar$. When h_i's are all big, the condition (3.2) is approximately a linear constraint on h_i's, and hence $h_1 < \min_{i>k} h_i$. Intuitively, with high chances, $T_{h_i}^i$ for $1 \leq i \leq k$ will proceed $T_{h_j}^j$ for $k+1 \leq j \leq N$ due to their smaller thresholds. Hence, it is expected that $\mathbb{E}_{\infty,\dots,\infty} \{T_\hbar\} \approx \mathbb{E}_{\infty,\dots,\infty} \{T_{h_1}^1 \wedge \dots \wedge T_{h_k}^k\} \geq \frac{2}{k\mu_1^2} g(h_1)$, where the inequality follows from (3.11). Below we rigorously show the validness of this heuristic argument.

PROPOSITION 3.3. *Suppose that the drifts μ_i of the observation processes $\xi_t^{(i)}$, $i = 1, \dots, N$ are such that $0 < \mu_1 = \mu_2 = \dots = \mu_k < \min_{i>k} \mu_i$ holds. Suppose also that the thresholds $\hbar$ satisfy (3.2). Then the mean time to the first false alarm for the N-CUSUM stopping rule $T_\hbar$ satisfies*

$$(3.18) \qquad \mathbb{E}_{\infty,\dots,\infty} \{T_\hbar\} \geq \left(1 - \sum_{j=k+1}^N \frac{\mathbb{E}_\infty^1 \{T_{h_1}^1\}}{\mathbb{E}_\infty^j \{T_{h_j}^j\}} \right) \frac{1}{k} \mathbb{E}_\infty^1 \{T_{h_1}^1\} = \left(1 - \sum_{j=k+1}^N \frac{\mu_j^2}{\mu_1^2} \frac{g(h_1)}{g(h_j)} \right) \frac{2}{k\mu_1^2} g(h_1),$$

where the function g is defined in (2.7).

PROOF. Let us denote by $R_{h_1} := T_{h_1}^1 \wedge \ldots \wedge T_{h_1}^k$. For any $k + 1 \leq j \leq N$, following similar arguments to the ones in (3.12) through (3.16), we have in this case that

$$(3.19) \qquad \frac{2}{\mu_j^2} g(h_j) = \mathbb{E}_{\infty,\ldots,\infty}\{T_{h_j}^j\}$$

$$= \mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} + \mathbb{E}_{\infty,\ldots,\infty}\{(T_{h_j}^j - T_\hbar)\,\mathbf{1}_{\{T_\hbar \neq T_{h_j}^j\}}\}$$

$$= \mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} + \mathbb{E}_{\infty,\ldots,\infty}\{\mathbb{E}_{\infty,\ldots,\infty}\{T_{h_j}^j - T_\hbar \mid \mathcal{F}_{T_\hbar}\}\,\mathbf{1}_{\{T_\hbar \neq T_{h_j}^j\}}\}$$

$$\leq \mathbb{E}_{\infty,\ldots,\infty}\{T_{h_1}^1\} + \frac{2}{\mu_j^2}\,\mathbb{E}_{\infty,\ldots,\infty}\{\big(g(y_{T_{h_j}^j}^{(j)}) - g(y_{T_\hbar}^{(j)})\big)\,\mathbf{1}_{\{T_\hbar \neq T_{h_j}^j\}}\}$$

$$\leq \frac{2}{\mu_1^2}\,g(h_1) + \frac{2}{\mu_j^2}\,g(h_j)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq T_{h_j}^j),$$

which implies that

$$(3.20) \qquad \mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = T_{h_j}^j) \leq \frac{\mu_j^2}{\mu_1^2}\,\frac{g(h_1)}{g(h_j)}.$$

On the other hand, for any $1 \leq i \leq k$, we similarly have

$$\frac{2}{\mu_1^2}\,g(h_1) = \mathbb{E}_{\infty,\ldots,\infty}\{T_{h_1}^1\} \quad \leq \quad \mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} + \frac{2}{\mu_1^2}\,g(h_1)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq T_{h_1}^i),$$

which implies that

$$\mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} \geq \frac{2}{\mu_1^2}\,g(h_1)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = T_{h_1}^i).$$

Summing up both sides of the above inequality for all $1 \leq i \leq k$, we obtain that

$$(3.21) \qquad k\,\mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} \geq \frac{2}{\mu_1^2}\,g(h_1)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = R_{h_1}) = \frac{2}{\mu_1^2}\,g(h_1)\,[1 - \mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq R_{h_1})].$$

However, we also have

$$(3.22) \qquad \mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq R_{h_1}) \leq \sum_{j=k+1}^N \mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = T_{h_j}^j) \leq \sum_{j=k+1}^N \frac{\mu_j^2}{\mu_1^2}\,\frac{g(h_1)}{g(h_j)},$$

where we used (3.20) in the above inequality. It follows from (3.21) and (3.22) that,

$$(3.23) \qquad \mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} \geq \left(1 - \sum_{j=k+1}^N \frac{\mu_j^2}{\mu_1^2}\,\frac{g(h_1)}{g(h_j)}\right)\frac{2}{k\mu_1^2}g(h_1).$$

which completes the proof. $\qquad\qquad\square$

As a result of Proposition 3.3, when $\mu_1 = \ldots = \mu_k < \min_{i>k} \mu_i$, then for any $\gamma > 0$ and any N-dimensional, predictable, non-singular, stochastic instantaneous correlation matrix Σ_t, we can choose the set of thresholds $\hbar$ using (3.2) and the transcendental equation

$$(3.24) \quad \left(1 - \sum_{j=k+1}^N \frac{\mathbb{E}_\infty^1\{T_{h_1}^1\}}{\mathbb{E}_\infty^j\{T_{h_j}^j\}}\right)\frac{1}{k}\,\mathbb{E}_\infty^1\{T_{h_1}^1\} = \left(1 - \sum_{j=k+1}^N \frac{\mu_j^2}{\mu_1^2}\,\frac{e^{h_1} - h_1 - 1}{e^{h_j} - h_j - 1}\right)\frac{2}{k\mu_1^2}(e^{h_1} - h_1 - 1) = \gamma,$$

then the resulting N-CUSUM stopping rule $T_\hbar \in \mathcal{T}_\gamma$. Again, Proposition 3.1 then implies that, both the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ and the detection delay of this N-CUSUM stopping rule $J^{(N)}(T_\hbar)$, are bounded above by $\frac{2}{\mu_1^2}g(-h_1)$.

3.1.3. *The general case - partial information about μ_i's.* In this case, it is assumed that only $\mu_1, \underline{\mu}_i, \overline{\mu}_i$, $i = 2, \ldots, N$, are known, and that $0 < \mu_1 \leq \underline{\mu}_i \leq \mu_i \leq \overline{\mu}_i$. Without loss of generality, we assume that $0 < \mu_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_{k'} < \min_{i>k'} \underline{\mu}_i$ holds for some $k' = \{1, \ldots, N-1\}$.

PROPOSITION 3.4. *Suppose that the drifts μ_i of the observation processes $\xi_t^{(i)}$, $i = 1, \ldots, N$ are such that $0 < \mu_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_{k'} < \min_{i>k'} \underline{\mu}_i$ holds and $\mu_i \in [\underline{\mu}_i, \overline{\mu}_i]$ for all $i = 2, \ldots, N$, Suppose also that the thresholds $\hbar$ satisfy* (3.3). *Then the mean time to the first false alarm for the N-CUSUM stopping rule $T_\hbar$ satisfies*

$$(3.25) \qquad \mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} \geq \frac{2}{\sum_{1 \leq i \leq k'} \mu_1(2\overline{\mu}_i - \mu_1)}\left(1 - \sum_{k'+1 \leq j \leq N} \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{\mu_1^2} \frac{g(h_1)}{g(h_j)}\right)g(h_1),$$

where the function g is defined in (2.7).

PROOF. According to (3.3), we have $h_1 = h_2 = \ldots = h_{k'}$. Let us denote by $R_{h_1} := T_{h_1}^1 \wedge \ldots \wedge T_{h_1}^{k'}$. Similar as (3.20) in the proof of Proposition 3.4, for any $k'+1 \leq j \leq N$, we have

$$g(h_j) = \mathbb{E}_{\infty,\ldots,\infty}\left\{\int_0^{T_{h_j}^j} \underline{\mu}_j\left(\mu_j - \frac{1}{2}\underline{\mu}_j\right)ds\right\}$$

$$= \mathbb{E}_{\infty,\ldots,\infty}\left\{\int_0^{T_\hbar} \underline{\mu}_j\left(\mu_j - \frac{1}{2}\underline{\mu}_j\right)ds\right\}$$

$$+ \mathbb{E}_{\infty,\ldots,\infty}\left\{\mathbb{E}_{\infty,\ldots,\infty}\left\{\int_{T_\hbar}^{T_{h_j}^j} \underline{\mu}_j\left(\mu_j - \frac{1}{2}\underline{\mu}_j\right)ds \,\Big|\, \mathcal{F}_{T_\hbar}\right\}\mathbf{1}_{\{T_\hbar \neq T_{h_j}^j\}}\right\}$$

$$\leq \mathbb{E}_{\infty,\ldots,\infty}\left\{\int_0^{T_{h_1}^1} \underline{\mu}_j\left(\mu_j - \frac{1}{2}\underline{\mu}_j\right)ds\right\} + \mathbb{E}_{\infty,\ldots,\infty}\left\{\left(g\big(y_{T_{h_j}^j}^{(j)}\big) - g\big(y_{T_\hbar}^{(j)}\big)\right)\mathbf{1}_{\{T_\hbar \neq T_{h_j}^j\}}\right\}$$

$$\leq \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{2}\mathbb{E}_{\infty,\ldots,\infty}\{T_{h_1}^1\} + g(h_j)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq T_{h_j}^j)$$

$$= \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{\mu_1^2}g(h_1) + g(h_j)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq T_{h_j}^j),$$

It follows that

$$\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = T_{h_j}^j) \leq \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{\mu_1^2} \frac{g(h_1)}{g(h_j)}.$$

which implies that

$$\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq R_{h_1}) = \mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = T_{h_j}^j, \text{for some } j \in \{k'+1, \ldots, N\})$$

$$(3.26) \qquad\qquad \leq \sum_{k'+1 \leq j \leq N} \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{\mu_1^2} \frac{g(h_1)}{g(h_j)}.$$

On the other hand, for any $1 \leq i \leq k'$, by Itô's formula (see, e.g. [27, Theorem 4.1.2]) we have

$$
\begin{aligned}
g(h_1) &= \mathbb{E}_{\infty,\ldots,\infty}\left\{ \int_0^{T_{h_1}^i} \mu_1\left(\mu_i - \frac{1}{2}\mu_1\right)\mathrm{d}s \right\} \\
&= \mathbb{E}_{\infty,\ldots,\infty}\left\{ \int_0^{T_\hbar} \mu_1\left(\mu_i - \frac{1}{2}\mu_1\right)\mathrm{d}s \right\} \\
&\quad + \mathbb{E}_{\infty,\ldots,\infty}\left\{ \mathbb{E}_{\infty,\ldots,\infty}\left\{ \int_{T_\hbar}^{T_{h_1}^i} \mu_1\left(\mu_i - \frac{1}{2}\mu_1\right)\mathrm{d}s \,\Big|\, \mathcal{F}_{T_\hbar} \right\} \mathbf{1}_{\{T_\hbar \neq T_{h_1}^i\}} \right\} \\
&= \mathbb{E}_{\infty,\ldots,\infty}\left\{ \int_0^{T_\hbar} \mu_1\left(\mu_i - \frac{1}{2}\mu_1\right)\mathrm{d}s \right\} + \mathbb{E}_{\infty,\ldots,\infty}\left\{ \left(g\big(y_{T_{h_1}^i}^{(i)}\big) - g\big(y_{T_\hbar}^{(i)}\big)\right)\mathbf{1}_{\{T_\hbar \neq T_{h_1}^i\}} \right\} \\
&\leq \frac{\mu_1(2\overline{\mu}_i - \mu_1)}{2}\mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} + g(h_1)\,\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq T_{h_1}^i),
\end{aligned}
$$

which implies that

$$
\frac{\mu_1(2\overline{\mu}_i - \mu_1)}{2}\mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} \geq g(h_1)\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = T_{h_1}^i).
$$

Summing up both sides of the above inequality for all $1 \leq i \leq k'$, we obtain that

$$
\begin{aligned}
(3.27) \qquad \frac{\sum_{1 \leq i \leq k'} \mu_1(2\overline{\mu}_i - \mu_1)}{2}\mathbb{E}_{\infty,\ldots,\infty}\{T_\hbar\} &\geq g(h_1)\mathbb{P}_{\infty,\ldots,\infty}(T_\hbar = R_{h_1}) \\
&= g(h_1)[1 - \mathbb{P}_{\infty,\ldots,\infty}(T_\hbar \neq R_{h_1})] \\
&\geq g(h_1)\left(1 - \sum_{k'+1 \leq j \leq N} \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{\mu_1^2}\frac{g(h_1)}{g(h_j)}\right),
\end{aligned}
$$

where we used (3.26) in the last step. The conclusion of the proposition follows immediately. $\qquad\square$

As a result of Proposition 3.4, when we only known μ_1 and possible ranges for other drift μ_i's, given any $\gamma > 0$ and any N-dimensional, predictable, non-singular, stochastic instantaneous correlation matrix Σ_t, we can choose the set of thresholds $\hbar$ using (3.3) and the transcendental equation

$$
(3.28) \qquad \left(1 - \sum_{j=k'+1}^{N} \frac{\underline{\mu}_j(2\overline{\mu}_j - \underline{\mu}_j)}{\mu_1^2}\frac{\mathrm{e}^{h_1} - h_1 - 1}{\mathrm{e}^{h_j} - h_j - 1}\right)\frac{2(\mathrm{e}^{h_1} - h_1 - 1)}{\sum_{1 \leq i \leq k'} \mu_1(2\overline{\mu}_i - \mu_1)} = \gamma,
$$

then the resulting N-CUSUM stopping rule $T_\hbar \in \mathcal{T}_\gamma$. Again, Proposition 3.1 then implies that, both the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ and the detection delay of this N-CUSUM stopping rule $J^{(N)}(T_\hbar)$, are bounded above by $\frac{2}{\mu_1^2}g(-h_1)$.

3.2. *The lower bound.* In this subsection, we present a robust lower bound for the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$. In fact, we can prove a stronger statement: for any stopping rule $T \in \mathcal{T}_\gamma$, its detection delay $J^{(N)}(T)$, is bounded below by the optimal detection delay in one dimension. The proof is accomplished by a change of measure argument as in [24] plus a decomposition formula for the Radon-Nikodym derivative in N dimensions.

LEMMA 3.5. *Let $\mathbb{Q} = \mathbb{P}_{\infty,\ldots,\infty}$ be the law of the N-tuple process $W_t := (w_t^{(1)}, w_t^{(2)}, \ldots, w_t^{(N)})$ for the Brownian motions defined in* (2.1). *And let $\mathbb{Q}_1$ be the law of the N-tuple process $(\mu_1 t + w_t^{(1)}, w_t^{(2)}, \ldots, w_t^{(N)})$. Then for all $t > 0$,*

$$(3.29) \qquad \left.\frac{d\mathbb{Q}_1}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = \mathrm{e}^{u_t^{(1)}} \cdot \mathcal{E}(B^{(1)})_t,$$

where $u_{\cdot}^{(1)}$ is defined in (2.5) *and $\mathcal{E}(B^{(1)})_{\cdot}$ is the stochastic exponential of the local martingale $B_{\cdot}^{(1)}$ defined in* (A.1)-(A.2). *Moreover the standard Brownian motions driving $B_{\cdot}^{(1)}$ are independent of $w^{(1)}$.*

PROOF. The proof can be found in the Appendix. $\qquad\qquad\square$

PROPOSITION 3.6. *For any stopping rule $T \in \mathcal{T}_\gamma$, we have $J^{(N)}(T) \geq (2/\mu_1^2)\, g(-\nu_1^\star)$, where $\nu_i^\star$ satisfies $g(\nu_1^\star) = (\mu_1^2/2)\,\gamma$ for the function g defined in* (2.7).

PROOF. Let T be an arbitrary $\mathbb{F}$-stopping rule such that $\mathbb{E}_{\infty,\ldots,\infty}\{T\} \geq \gamma$ holds and observe that

$$(3.30) \qquad J^{(N)}(T) \geq \tilde{J}_1^{(N)}(T) := \sup_{s \in \mathbb{R}_+} \ \text{essup} \ \mathbb{E}_{s,\infty,\ldots,\infty}\{(T-s)^+ \mid \mathcal{F}_s\} \geq \tilde{J}_1^{(N)}(T_\nu)$$

where $T_\nu := T \wedge T_\nu^1 \leq T$, a.s., and T_ν^1 is the CUSUM stopping rule given in (2.6) for some threshold ν which will be determined later. Clearly, T_ν is a finite stopping rule. In what follows we will demonstrate that for any given $\epsilon > 0$, there exists a $\nu > 0$ such that

$$(3.31) \qquad \tilde{J}_1^{(N)}(T_\nu) \geq \frac{2}{\mu_1^2}\, g(-\nu_1^*) - \epsilon.$$

where $\nu_1^\star$ is chosen so that $g(\nu_1^\star) = (\mu_1^2/2)\gamma$ and the function g is given by (2.7). Because ϵ in (3.31) can be arbitrarily small, (3.30) and (3.31) will imply the assertion in the Proposition for $i = 1$. This is in a similar light as in [24].

By applying Itô's formula (see, e.g. [27, Theorem 4.1.2] and [24]) to $\{g(-y_t^{(1)})\}_{s \wedge T_\nu \leq t < T_\nu}$ and proceed by using similar arguments as in (3.13)-(3.15) in Proposition 3.2, we obtain that, for any fixed $s \in \mathbb{R}_+$,

$$(3.32) \qquad \mathbb{E}_{s,\infty,\ldots,\infty}\big\{(T_\nu - s)^+ \mid \mathcal{F}_s\big\} = \frac{2}{\mu_1^2}\, \mathbb{E}_{s,\infty,\ldots,\infty}\{g(-y_{T_\nu}^{(1)}) - g(-y_s^{(1)}) \mid \mathcal{F}_s\}\mathbf{1}_{\{T_\nu \geq s\}}.$$

Using Girsanov's theorem (see, e.g. [22, Chapter 3, Theorem 5.1]) and Lemma 3.5 at the finite stopping rule $T_\nu \wedge n$ for a fixed $n > 0$, we have that

$$\mathbb{E}_{s,\infty,\ldots,\infty}\big\{g(-y_{T_\nu \wedge n}^{(1)}) - g(-y_s^{(1)}) \mid \mathcal{F}_s\big\}\mathbf{1}_{\{T_\nu \wedge n > s\}}$$
$$= \mathbb{E}_{\infty,\ldots,\infty}\Big\{\mathrm{e}^{u_{T_\nu \wedge n}^{(1)} - u_s^{(1)}} \cdot \frac{\mathcal{E}(B^{(1)})_{T_\nu \wedge n}}{\mathcal{E}(B^{(1)})_s}\, [g(-y_{T_\nu \wedge n}^{(1)}) - g(-y_s^{(1)})]\Big| \mathcal{F}_s\Big\}\mathbf{1}_{\{T_\nu \wedge n > s\}}.$$

Consider the enlargement of filtration $\mathcal{F}_t^a = \mathcal{F}_t \vee \mathcal{G}_{T_\nu^1}^{(1)}$. Then clearly, $\mathcal{F}_t^a \equiv \mathcal{F}_t$ for all $t \geq T_\nu^1$, but on the event $\{T_\nu \wedge n > s\}$, for all $t \in [s, T_\nu \wedge n] \subset [0, T_\nu^1]$, we have $\mathcal{F}_t \subsetneq \mathcal{F}_t^a$. By the tower property of

conditional expectation, on the event that $\{T_\nu \wedge n > s\}$,

$$(3.33) \qquad \mathbb{E}_{\infty,\ldots,\infty}\left\{\mathrm{e}^{u^{(1)}_{T_\nu \wedge n} - u^{(1)}_s} \cdot \frac{\mathcal{E}(B^{(1)})_{T_\nu \wedge n}}{\mathcal{E}(B^{(1)})_s}\left[g(-y^{(1)}_{T_\nu \wedge n}) - g(-y^{(1)}_s)\right]\Big|\mathcal{F}_s\right\}$$

$$=\mathbb{E}_{\infty,\ldots,\infty}\left\{\mathrm{e}^{u^{(1)}_{T_\nu \wedge n} - u^{(1)}_s}\left[g(-y^{(1)}_{T_\nu \wedge n}) - g(-y^{(1)}_s)\right] \cdot \mathbb{E}_{\infty,\ldots,\infty}\left\{\frac{\mathcal{E}(B^{(1)})_{T_\nu \wedge n}}{\mathcal{E}(B^{(1)})_s}\Big|\mathcal{F}^a_s\right\}\Big|\mathcal{F}_s\right\}$$

$$=\mathbb{E}_{\infty,\ldots,\infty}\left\{\mathrm{e}^{u^{(1)}_{T_\nu \wedge n} - u^{(1)}_s}\left[g(-y^{(1)}_{T_\nu \wedge n}) - g(-y^{(1)}_s)\right]\Big|\mathcal{F}_s\right\},$$

where the last equality is due to the fact that $\mathcal{E}(B^{(1)})_t$ is a $\overline{\mathbb{F}}$-exponential martingale (under assumption equation (2.2)) driven by Brownian motions that are independent of $w^{(1)}$ (see Lemma 3.5). Similarly, it can be shown that,

$$(3.34) \qquad \mathbf{1}_{\{T_\nu \wedge n > s\}} = \mathbf{1}_{\{T_\nu \wedge n > s\}}\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u^{(1)}_{T_\nu \wedge n} - u^{(1)}_s} \,|\, \mathcal{F}_s\},$$

We now let $n \uparrow \infty$ in (3.33) and (3.34). From the fact that $u^{(1)}_{T_\nu \wedge n} - u^{(1)}_s \leq u^{(1)}_{T_\nu \wedge n} - m^{(1)}_{T_\nu \wedge n} = y^{(1)}_{T_\nu \wedge n} \leq \nu$, and the monotonicity of function $g(-h)$, we have that

$$0 < \mathrm{e}^{u^{(1)}_{T_\nu \wedge n} - u^{(1)}_s} \leq \mathrm{e}^\nu < \infty \quad \text{and} \quad |g(-y^{(1)}_{T_\nu \wedge n}) - g(-y^{(1)}_s)| \leq 2\,g(-\nu) < \infty, \ \ \forall n > 0$$

and thus the bounded convergence theorem implies that, on the event $\{T_\nu > s\}$,

$$(3.35) \qquad 1 = \mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u^{(1)}_{T_\nu} - u^{(1)}_s} \,|\, \mathcal{F}_s\}$$

$$(3.36) \qquad \mathbb{E}_{s,\infty,\ldots,\infty}\{g(-y^{(1)}_{T_\nu}) - g(-y^{(1)}_s) \,|\, \mathcal{F}_s\} = \mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u^{(1)}_{T_\nu} - u^{(1)}_s}\left[g(-y^{(1)}_{T_\nu}) - g(-y^{(1)}_s)\right] \,|\, \mathcal{F}_s\}.$$

It follows from (3.30), (3.32), (3.35) and (3.36), that

$$(3.37) \qquad \tilde{J}^{(N)}_1(T_\nu)\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u^{(1)}_{T_\nu} - u^{(1)}_s} \,|\, \mathcal{F}_s\}\mathbf{1}_{\{T_\nu > s\}} = \tilde{J}^{(N)}_1(T_\nu)\mathbf{1}_{\{T_\nu > s\}}$$

$$\geq \mathbb{E}_{s,\infty,\ldots,\infty}\{(T_\nu - s)^+ \,|\, \mathcal{F}_s\}\mathbf{1}_{\{T_\nu > s\}}$$

$$= \frac{2}{\mu^2_1}\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u^{(1)}_{T_\nu} - u^{(1)}_s}\left[g(-y^{(1)}_{T_\nu}) - g(-y^{(1)}_s)\right] \,|\, \mathcal{F}_s\}\mathbf{1}_{\{T_\nu > s\}}.$$

Following the same arguments as in Theorem 2 of [24], we integrate both sides of the above inequality with respect to $(-\mathrm{d}m^{(1)}_s)$ for all $s \in [0, T_\nu]$ and then take the expectation under $\mathbb{P}_{\infty,\ldots,\infty}$. We therefore obtain that

$$\tilde{J}^{(N)}_1(T_\nu)\,\mathbb{E}_{\infty,\ldots,\infty}\left\{\mathrm{e}^{u^{(1)}_{T_\nu}}\int_0^{T_\nu}\mathrm{e}^{-u^{(1)}_s}(-\mathrm{d}m^{(1)}_s)\right\}$$

$$\geq \frac{2}{\mu^2_1}\,\mathbb{E}_{\infty,\ldots,\infty}\left\{\mathrm{e}^{u^{(1)}_{T_\nu}}\int_0^{T_\nu}\mathrm{e}^{-u^{(1)}_s}\left[g(-y^{(1)}_{T_\nu}) - g(-y^{(1)}_s)\right](-\mathrm{d}m^{(1)}_s)\right\}.$$

Notice that the measure $\mathrm{d}m^{(1)}_s$ is supported on the random set $\{s \mid y^{(1)}_s = 0\} = \{s \mid u^{(1)}_s = m^{(1)}_s\}$, and that $g(0) = 0$, thus we obtain that

$$\tilde{J}^{(N)}_1(T_\nu)\,\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y^{(1)}_{T_\nu}} - \mathrm{e}^{u^{(1)}_{T_\nu}}\} \geq \frac{2}{\mu^2_1}\,\mathbb{E}_{\infty,\ldots,\infty}\{[\mathrm{e}^{y^{(1)}_{T_\nu}} - \mathrm{e}^{u^{(1)}_{T_\nu}}]\,g(-y^{(1)}_{T_\nu})\}.$$

On the other hand, by letting $s = 0$ in (3.37) we have that

$$\tilde{J}_1^{(N)}(T_\nu)\,\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u_{T_\nu}^{(1)}}\} \geq \frac{2}{\mu_1^2}\,\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{u_{T_\nu}^{(1)}}\,g(-y_{T_\nu}^{(1)})\}.$$

In all, we have that

(3.38) $$\tilde{J}_1^{(N)}(T_\nu)\,\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\} \geq \frac{2}{\mu_1^2}\,\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\,g(-y_{T_\nu}^{(1)})\}$$

holds.

To relate the detection delay in (3.38) to the first false alarm constraint γ, we use similar arguments as in (3.13)-(3.15) in Proposition 3.2, to obtain that

(3.39) $$\frac{2}{\mu_1^2}\,\mathbb{E}_{\infty,\ldots,\infty}\{g(y_{T_\nu}^{(1)})\} = \mathbb{E}_{\infty,\ldots,\infty}\{T_\nu\}.$$

By taking the limit as $\nu \uparrow \infty$ and using monotone convergence theorem, we have that $T_\nu \uparrow T$, and $\lim_{\nu\uparrow\infty} \mathbb{E}_{\infty,\ldots,\infty}\{T_\nu\} = \mathbb{E}_{\infty,\ldots,\infty}\{T\} \geq \gamma$, which implies that there exists a large enough ν, such that

$$\frac{2}{\mu_1^2}\,\mathbb{E}_{\infty,\ldots,\infty}\{g(y_{T_\nu}^{(1)})\} = \mathbb{E}_{\infty,\ldots,\infty}\{T_\nu\} \geq \gamma - \epsilon$$

holds for any pre-specified $\epsilon > 0$. Now consider the non-negative function $p(y) := \mathrm{e}^y\,[g(-y) - g(-\nu_1^\star)] - g(y) + g(\nu_1^\star)$, using which we trivially have $\mathbb{E}_{\infty,\ldots,\infty}\{p(y_{T_\nu}^{(1)})\} \geq 0$, implying that

(3.40) $$\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\,g(-y_{T_\nu}^{(1)})\} \geq \mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\}\,g(-\nu_1^\star) + \mathbb{E}_{\infty,\ldots,\infty}\{g(y_{T_\nu}^{(1)})\} - g(\nu_1^\star)$$

$$= \mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\}\,g(-\nu_1^\star) + \mathbb{E}_{\infty,\ldots,\infty}\{g(y_{T_\nu}^{(1)})\} - \frac{\mu_1^2}{2}\,\gamma$$

$$\geq \mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\}\,g(-\nu_1^\star) - \frac{\mu_1^2}{2}\,\epsilon$$

$$\geq \mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\}\,[g(-\nu_1^\star) - \frac{\mu_1^2}{2}\,\epsilon],$$

since $\mathbb{E}_{\infty,\ldots,\infty}\{\mathrm{e}^{y_{T_\nu}^{(1)}}\} \geq 1$. The above inequality in (3.40) together with (3.38) yields (3.31), which completes the proof. $\qquad\square$

4. Asymptotic optimality of the N-CUSUM stopping rule.

In this section we demonstrate the asymptotic optimality of the N-CUSUM stopping rule $T_\hbar$ for $\hbar$ chosen such that (3.2) and either (3.17) or (3.24) hold, or (3.3) and (3.28) hold. To this end, we examine the asymptotic behavior of the robust upper and low bounds established in Section 3. We show that the additional detection delay of $T_\hbar$ over the optimal detection delay remains bounded as the mean time of the first false alarm γ increases without bound.

Let any sufficiently large $\gamma > 0$ and recall from Section 3 (in particular Propositions 3.1 and 3.6) that the optimal detection delay in (2.4) is bounded from below and above as:

(4.1) $$\frac{2}{\mu_1^2}g(-\nu_1^\star) \leq \inf_{T\in\mathcal{T}_\gamma} J^{(N)}(T) \leq J^{(N)}(T_\hbar) \leq \frac{2}{\mu_1^2}g(-h_1),$$

where the set of thresholds $\nu_1^\star$ and $\hbar$ is respectively determined using $(\mu_1^2/2)g(\nu_1^\star) = \gamma$ and either (3.2) together with (3.17) or with (3.24), or (3.3) together with (3.28), when the drifts' sizes μ_i are all known and equal or unequal,[5] or partially known, respectively.

---

[5] Note that when the drifts μ_i are different, we do not necessarily require the uniqueness of $\hbar$ that solves (3.2) and (3.24).

It is easily seen from Result 3 in the Appendix of [18] that, as $\gamma \to \infty$,

$$(4.2) \qquad \nu_1^\star = \log \frac{\mu_1^2}{2} + \log \gamma + o(1),$$

$$(4.3) \qquad \frac{2}{\mu_1^2} g(-\nu_1^\star) = \frac{2}{\mu_1^2} \Big(\log \frac{\mu_1^2}{2} + \log \gamma - 1 + o(1) \Big).$$

Moreover, when all the drifts are known and $\mu_i = \mu > 0$ for all $i = 1, \ldots, N$, the thresholds $h_i = h > 0$ for all $i = 1, \ldots, N$. Using (3.17) and Result 3 in the Appendix of [18] we have that, as $\gamma \to \infty$,

$$(4.4) \qquad h_1 = \log \frac{N\mu_1^2}{2} + \log \gamma + o(1),$$

$$(4.5) \qquad \frac{2}{\mu_1^2} g(-h_1) = \frac{2}{\mu_1^2} \Big(\log \frac{N\mu_1^2}{2} + \log \gamma - 1 + o(1) \Big).$$

As a result, we have the following optimality result.

THEOREM 4.1. *Assume that the drift sizes are all known and $\mu_i = \mu > 0$ for all $i = 1, \ldots, N$. Then for any predictable, non-singular, stochastic instantaneous correlation matrix covariance matrix Σ_t, the N-CUSUM stopping rule $T_\hbar$ defined in Algorithm 2.1, where the set of thresholds $\hbar$ is chosen using (3.2) and (3.17), is asymptotically optimal to the problem (2.4). More specifically, the difference between the detection delay of the N-CUSUM stopping rule, $J^{(N)}(T_\hbar)$, and the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$, is bounded above by $\frac{2}{\mu_1^2} \log N$, as $\gamma \to \infty$.*

PROOF. The result follows from (4.1), (4.3) and (4.5):

$$
\begin{aligned}
0 \leq J^{(N)}(T_\hbar) - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) &\leq \frac{2}{\mu_1^2} g(-h_1) - \frac{2}{\mu_1^2} g(-\nu_1^\star) \\
&\leq \frac{2}{\mu_1^2} \Big(\log \frac{N\mu_1^2}{2} + \log \gamma - 1 + o(1) \Big) - \frac{2}{\mu_1^2} \Big(\log \frac{\mu_1^2}{2} + \log \gamma - 1 + o(1) \Big) \\
&= \frac{2}{\mu_1^2} \log N + o(1),
\end{aligned}
$$

as $\gamma \to \infty$. □

On the other hand, in the more general case that the drifts are all known and $\mu_1 = \ldots = \mu_k < \min_{i>k} \mu_i$, using (3.2), (3.24), and Result 3 in the Appendix of [18], we obtain that,

$$(4.6) \qquad h_1 = \log \frac{k\mu_1^2}{2} + \log \gamma + o(1),$$

$$(4.7) \qquad \frac{2}{\mu_1^2} g(-h_1) = \frac{2}{\mu_1^2} \Big(\log \frac{k\mu_1^2}{2} + \log \gamma - 1 + o(1) \Big).$$

It follows that we have the following optimality result.

THEOREM 4.2. *Assume that the drift sizes are known and $0 < \mu_1 = \ldots = \mu_k < \min_{i>k} \mu_i$. Then for any predictable, non-singular, stochastic instantaneous correlation matrix Σ_t, the N-CUSUM stopping rule $T_\hbar$ defined in Algorithm 2.1, where the set of thresholds $\hbar$ is chosen using (3.2) and (3.24), is asymptotically optimal to the problem (2.4). More specifically, the difference between the detection delay of the N-CUSUM stopping rule, $J^{(N)}(T_\hbar)$, and the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$, is bounded above by $\frac{2}{\mu_1^2} \log k$, as $\gamma \to \infty$. In particular, if $k = 1$, then $T_\hbar$ is equivalent to the optimal solution to (2.4) asymptotically.*

PROOF. The result follows from (4.1), (4.3) and (4.9):

$$
\begin{aligned}
0 \leq J^{(N)}(T_\hbar) - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) &\leq \frac{2}{\mu_1^2} g(-h_1) - \frac{2}{\mu_1^2} g(-\nu_1^\star) \\
&\leq \frac{2}{\mu_1^2} \Big(\log \frac{k\mu_1^2}{2} + \log\gamma - 1 + o(1) \Big) - \frac{2}{\mu_1^2} \Big(\log \frac{\mu_1^2}{2} + \log\gamma - 1 + o(1) \Big) \\
&= \frac{2}{\mu_1^2} \log k + o(1),
\end{aligned}
$$

as $\gamma \to \infty$. If $k = 1$, the above upper bound for $J^{(N)}(T_\hbar) - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ is $o(1)$, and hence, the N-CUSUM stopping rule is equivalent to the optimal solution to (2.4) asymptotically. $\square$

Finally, if we only know μ_1 and have partial information about other drifts, i.e. $\mu_i \in [\underline{\mu}_i, \overline{\mu}_i]$ for all $i = 2, \ldots, N$ and $0 < \mu_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_{k'} < \min_{i>k'} \underline{\mu}_i$ for some $k' \in \{1, 2, \ldots, N-1\}$. using (3.3), (3.28), and Result 3 in the Appendix of [18], we obtain that,

$$
\tag{4.8} h_1 = \log \frac{\sum_{1 \leq i \leq k'} \mu_1(2\overline{\mu}_i - \mu_1)}{2} + \log\gamma + o(1),
$$

$$
\tag{4.9} \frac{2}{\mu_1^2} g(-h_1) = \frac{2}{\mu_1^2} \Big(\log \frac{\sum_{1 \leq i \leq k'} \mu_1(2\overline{\mu}_i - \mu_1)}{2} + \log\gamma - 1 + o(1) \Big).
$$

It follows that we have the following optimality result.

THEOREM 4.3. *Assume that the μ_1 is known, $\mu_i \in [\underline{\mu}_i, \overline{\mu}_i]$ for all $i = 2, \ldots, N$ and that $0 < \mu_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_{k'} < \min_{i>k'} \underline{\mu}_i$ for some $k' \in \{1, \ldots, N-1\}$. Then for any predictable, non-singular, stochastic instantaneous correlation matrix Σ_t, the N-CUSUM stopping rule $T_\hbar$ defined in Algorithm 2.1, where the set of thresholds $\hbar$ is chosen using (3.3) and (3.28), is asymptotically optimal to the problem (2.4). More specifically, the difference between the detection delay of the N-CUSUM stopping rule, $J^{(N)}(T_\hbar)$, and the optimal detection delay $\inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$, is bounded above by*

$$
\frac{2}{\mu_1^2} \log \Big(\sum_{1 \leq i \leq k'} (2\overline{\mu}_i/\mu_1 - 1) \Big),
$$

as $\gamma \to \infty$. In particular, if $k' = 1$, then $T_\hbar$ is equivalent to the optimal solution to (2.4) asymptotically.

PROOF. The result follows from (4.1), (4.3) and (4.9):

$$
\begin{aligned}
0 \leq J^{(N)}(T_\hbar) - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T) &\leq \frac{2}{\mu_1^2} g(-h_1) - \frac{2}{\mu_1^2} g(-\nu_1^\star) \\
&\leq \frac{2}{\mu_1^2} \Big(\log \frac{\sum_{1 \leq i \leq k'} \mu_1(2\overline{\mu}_i - \mu_1)}{2} + \log\gamma - 1 + o(1) \Big) - \frac{2}{\mu_1^2} \Big(\log \frac{\mu_1^2}{2} + \log\gamma - 1 + o(1) \Big) \\
&= \frac{2}{\mu_1^2} \log \frac{\sum_{1 \leq i \leq k'} (2\overline{\mu}_i - \mu_1)}{\mu_1} + o(1),
\end{aligned}
$$

as $\gamma \to \infty$. If $k' = 1$, the above upper bound for $J^{(N)}(T_\hbar) - \inf_{T \in \mathcal{T}_\gamma} J^{(N)}(T)$ is $o(1)$, and hence, the N-CUSUM stopping rule is equivalent to the optimal solution to (2.4) asymptotically. $\square$

REMARK 4.1.   *From Definition 2.1, we know that the order of the asymptotic optimality achieved in Theorems 4.1, 4.2, 4.3 is of the second order. If μ_1 is strictly smaller than all the other drifts ($k = 1$ in Theorem 4.2 and $k' = 1$ in Theorem 4.3), then the N-CUSUM stopping rule given by Algorithm 2.1, and either (3.2) and (3.24), or (3.3) and (3.28) exhibits third order asymptotic optimality. Moreover, it can be seen after a perusal of the proofs that the order of the asymptotic optimality of the N-CUSUM does not change if we model μ_i's as $\overline{\overline{\mathbb{F}}}$-adapted processes bounded by known constants $\underline{\mu}_i$ and $\overline{\mu}_i$, for all $i = 2, \ldots, N$.*

5. Applications.   In this section we discuss one of the applications of the results in decentralized communication systems. Let us now suppose that each of the observation processes $\{\xi_t^{(i)}\}$ become sequentially available at a particular location monitored by sensor S_i, which then employs an asynchronous communication scheme to a central fusion center. In particular, sensor S_i communicates to the central fusion center only when it wants to signal an alarm, which is elicited according to a CUSUM stopping rule $T_{h_i}^i$ as in (2.9) adapted to the small filtration $\{\mathcal{G}_t^i\}$. The observations received at the N sensors can change dynamics at distinct unknown points τ_i. An example of such a case is described in [2] where the motivation suggested arises in the health-monitoring of mechanical, civil and aeronautic structures. The fusion center, whose objective is to detect the first time when there is a change in at least one of the sensors devises a minimal strategy; that is, it declares that a change has occurred at the first instance when one of the sensors communicates an alarm. The implication of the main Theorems in Section 4 is that in fact this strategy is the best, at least asymptotically, in that there is no loss in performance, between the case in which the fusion center receives the raw data $\{\xi_t^{(1)}, \ldots \xi_t^{(N)}\}$ summarized in the large filtration $\{\mathcal{F}_t\}$ directly and the case in which the communication that takes place is limited to the decentralized setup. In other words, the CUSUM stopping rule $T_\hbar$ is a sufficient statistic (at least asymptotically) of the minimum N possibly distinct change points. That is, the stopping rule $T_\hbar$ is an asymptotically optimal solution to the problems of quickest detection presented in (2.4). In practice sensors are cheap and easy to replace devices whereas central fusion centers or central processing units are not. Transferring most of the processing work to the sensors while incurring no loss in the efficiency of the system is thus valuable and can render cost and speed effective communication systems.

6. Summary.   In this paper we study the problem of detecting the minimum of N different change points in a N-dimensional Brownian system with partial information of the drifts and an arbitrary, predictable, non-singular, stochastic instantaneous correlation matrix Σ_t. It is shown that, under an extended Lorden's minmax criterion, the N-CUSUM stopping rule exhibits asymptotic optimality in the tradeoff between detection delay and false alarms, as the mean time to the first false alarm increases without bound. Moreover, the performance of the N-CUSUM stopping rule under dependence is *no worse than* that under independence [18]. This optimality result is obtained by establishing a robust upper bound and a robust lower bound for the optimal detection delay. The contribution of this work can be seen in two folds. First, we designed a low complexity, efficeint stopping rule without using the explicit information of the covariance matrix Σ_t. This stopping rule is guaranteed to have a comparable performance or identical performance as the optimal stopping rule, even with cross-correlated observations - a non-trivial extension to the existing literature and the first formal treatment of correlated noise in change-point detection. Secondly, the robust bounds obtained in this work provide a unified robust probabilistic (rather than analytical) approach to treat detection problems with multiple change-points or multiple alternatives [15, 16, 38]. This is especially useful when the analytical characteristic such as joint density or Green functions are not explicitly available.

APPENDIX A

Let us denote by $\Sigma_t^1 = (\rho_t^{i,j})_{i,j \neq 1}$ the $(N-1) \times (N-1)$ matrix obtained from Σ_t by removing its first column and first row, and by $\tilde{\Sigma}_t^1$ the $(N-1) \times (N-1)$ matrix $(\rho_t^{1,i} \cdot \rho_t^{1,j})_{i,j \neq 1}$. We further introduce a local martingale

$$(A.1) \quad B_t^{(1)} = -\mu_1 \int_0^t (\rho_s^{2,1}, \ldots, \rho_s^{N,1})(\Sigma_s^1 - \tilde{\Sigma}_s^1)^{-1} \left(\sqrt{1 - (\rho_s^{2,1})^2} d\tilde{w}_s^{(2)}, \ldots, \sqrt{1 - (\rho_s^{N,1})^2} d\tilde{w}_s^{(N)} \right)',$$

$$(A.2) \quad \tilde{w}_t^{(i)} = \int_0^t \frac{dw_s^{(i)} - \rho_s^{i,1} dw_s^{(1)}}{\sqrt{1 - (\rho_s^{i,1})^2}}, \quad 2 \leq i \leq N.$$

The local martingale $B^{(1)}$ will naturally appear in (A.8) of the following proof. We are now ready to prove the assertion in Lemma 3.5.

PROOF OF LEMMA 3.5. Since Σ_t is non-singular at all time a.s., we can use a Cholesky decomposition to obtain a lower triangular, non-singular matrix-valued process $L_t = (L_t^{i,j})_{1 \leq i,j \leq N}$, and a N-dimensional standard Brownian motion $Z = (z^{(1)}, \ldots, z^{(N)})'$, such that

$$(A.3) \qquad\qquad dW_t' = L_t \, dZ_t$$

holds. In particular, we have $L_t^{1,1} \equiv 1$, $z_t^{(1)} \equiv w_t^{(1)}$, and $L_t^{i,1} = \rho_t^{i,1}$ for all $2 \leq i \leq N$. Using Girsanov's theorem (see, e.g. [22, Chapter 3, Theorem 5.1]) and the condition in (2.2), the measure change from $\mathbb{Q}$ to $\mathbb{Q}_1$ is given by the exponential martingale

$$(A.4) \qquad\qquad \frac{d\mathbb{Q}_1}{d\mathbb{Q}}\bigg|_{\mathcal{F}_t} = \mathcal{E}(\int_0^{\cdot} \nu_s dZ_s)_t,$$

where $\nu = (\nu^{(1)}, \ldots, \nu^{(N)})$ is a N-tuple process, such that

$$(A.5) \qquad\qquad (\mu_1, 0, \ldots, 0)' = L_t \nu_t'.$$

It is easily seen that $\nu_t^{(1)} \equiv \mu_1$. Moreover, from $L_t^{i,1} = \rho_t^{i,1}$ for any $2 \leq i \leq N$, we know that,

$$(A.6) \qquad\qquad \tilde{L}_t \left(\nu_t^{(2)}, \ldots, \nu_t^{(N)} \right)' = -\mu_1 \left(\rho_t^{2,1}, \ldots, \rho_t^{N,1} \right)',$$

where $\tilde{L}_t = (L_t^{i,j})_{2 \leq i,j \leq N}$ is a $(N-1) \times (N-1)$ non-singular matrix-valued process . On the other hand, notice that

$$(A.7) \qquad d(z_t^{(2)}, \ldots, z_t^{(N)})' = (\tilde{L}_t)^{-1} \left(dw_t^{(2)} - \rho_t^{2,1} dw_t^{(1)}, \ldots, dw_t^{(N)} - \rho_t^{N,1} dw_t^{(1)} \right)'.$$

Using the equations in (A.6) and (A.7), we conclude that

(A.8)
$$\int_0^t \nu_s \mathrm{d}Z_s$$
$$= \int_0^t \nu_s^{(1)} \mathrm{d}z_s^{(1)} + \int_0^t \left(\nu_s^{(2)}, \dots, \nu_s^{(N)} \right) (\mathrm{d}z_s^{(2)}, \dots, \mathrm{d}z_s^{(N)})'$$
$$= \mu_1 w_t^{(1)} - \mu_1 \int_0^t (\rho_s^{2,1}, \dots, \rho_s^{N,1})(\tilde{L}_s')^{-1}(\tilde{L}_s)^{-1} \left(\mathrm{d}w_s^{(2)} - \rho_s^{2,1}\mathrm{d}w_s^{(1)}, \dots, \mathrm{d}w_s^{(N)} - \rho_s^{N,1}\mathrm{d}w_s^{(1)} \right)'$$
$$= \mu_1 w_t^{(1)} - \mu_1 \int_0^t (\rho_s^{2,1}, \dots, \rho_s^{N,1})(\Sigma_s^1 - \tilde{\Sigma}_s^1)^{-1}\left(\sqrt{1 - (\rho_s^{2,1})^2}\mathrm{d}\tilde{w}_s^{(2)}, \dots, \sqrt{1 - (\rho_s^{N,1})^2}\mathrm{d}\tilde{w}_s^{(N)} \right)'$$
$$= \mu_1 w_t^{(1)} + B_t^{(1)}.$$

where the third equality follows from the fact that $\tilde{L}_t(\tilde{L}_t)' + \tilde{\Sigma}_t^1 = \Sigma_t^1$ holds (see accompanying internet supplement). Finally, by the way we construct $B^{(1)}$, we know that the Brownian motions that drive $B^{(1)}$ and $w^{(1)}$ are independent, and this completes the proof. $\qquad\square$

REFERENCES

[1] BASSEVILLE, M., ABDELGHANI, M. and BENVENISTE, A. (2000). Subspace-based fault detection algorithms for vibration monitoring. *Automatica* **36** 101-109.

[2] BASSEVILLE, M., BENVENISTE, A., GOURSAT, M. and MEVEL, L. (2007). Subspace-based algorithms for structural identification, damage detection, and sensor data fusion. *Journal of Applied Signal Processing, Special Issue on Advances in Subspace-Based Techniques for Signal Processing and Communications* **2007** 200-213.

[3] BASSEVILLE, M., MEVEL, L. and GOURSAT, M. (2004). Statistical model-based damage detection and localization, subspace-based residuals and damage-to-noise sensitivity ratios. *Journal of Sound and Vibration* **275** 769-794.

[4] BASSEVILLE, M. and NIKIFOROV, I. (1993). *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, Englewood Cliffs NJ.

[5] BAYRAKTAR, E., DAYANIK, S. and KARATZAS, I. (2006). Adaptive poisson disorder problem. *The Annals of Applied Probability* **16** 1190-1261.

[6] BAYRAKTAR, E. and POOR, H. V. (2007). Quickest detection of a minimum of two Poisson disorder times. *SIAM Journal on Control and Optimization* **46** 308-331.

[7] BEIBEL, M. (1996). A note on Ritov's Bayes approach to the minimax property of the CUSUM procedure. *Annals of Statistics* **24** 1804-1812.

[8] BEIBEL, M. (1997). Sequential change-point detection in continuous time when the post-change drift is unknown. *Bernoulli* **3** 457-478.

[9] BEIBEL, M. and LERCHE, H. R. (2003). Sequential Bayes detection of trend changes. In *Foundations of Statistical Inference* (Y. Haitovsky, H. R. Lerche and Y. Ritov, eds.) 117-130. hysica, Heidelberg.

[10] CUCHIARA, R., PICCARDI, M. and MELLO, P. (2000). Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Intelligent Transportation Systems Society* **1** 119-130.

[11] DAYANIK, S., POOR, H. V. and SEZER, S. O. (2008). Multisource Bayesian sequential change detection. *The Annals of Applied Probability* **18** 552-590.

[12] EWINS, D. J. (2000). *Modal Testing: Theory, Practice and Applications*, 2nd ed. Research Studies Press, Letchworth, Hertfordshire UK.

[13] FELLOURIS, G. and MOUSTAKIDES, G. V. (2011). Decentralized sequential hypothesis testing using asynchronous communication. *IEEE Transactions on Information Theory* **57** 534-548.

[14] GAPEEV, P. V. (2005). The disorder problem for compound Poisson processes with exponential jumps. *The Annals of Applied Probability* **15** 487-499.

[15] HADJILIADIS, O., HERNANDEZ-DEL-VALLE, G. and STAMOS, I. (2009). A comparison of 2-CUSUM stopping rules for quickest detection of two-sided alternatives in a Brownian motion model. *Sequentially Analysis* **28** 92-114.

[16] HADJILIADIS, O. and MOUSTAKIDES, G. V. (2006). Optimal and asymptotically optimal CUSUM rules for change point detection in the Brownian motion model with multiple alternatives. *Theory of Probability and its Applications* **50** 131-144.

[17] HADJILIADIS, O., SCHÄFER, T. and POOR, H. V. (2009). Quickest detection in coupled systems. In *Proceedings of the 48th IEEE Conference on Decision and Control* 4723 - 4728.

[18] HADJILIADIS, O., ZHANG, H. and POOR, H. V. (2009). One-shot schemes for decentralized quickest detection. *IEEE Transactions on Information Theory* **55** 3346-3359.

[19] HEYLEN, W., LAMMENS, S. and SAS, P. (1995). Modal Analysis Theory and Testing Technical Report, Department of Mechanical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium.

[20] IVANOFF, B. G. and MERZBACH, E. (2010). Optimal detection of a space set in a spatial Poisson process. *The Annals of Applied Probability* **20** 640-659.

[21] JUANG, J. N. (1994). *Applied System Identification*. Prentice Hall, Englewood Cliffs, NJ, USA.

[22] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer.

[23] MOUSTAKIDES, G. V. (1986). Optimal stopping rules for detecting changes in distributions. *Annals of Statistics* **14** 1379-1387.

[24] MOUSTAKIDES, G. V. (2004). Optimality of the CUSUM procedure in continuous time. *Annals of Statistics* **32** 302-315.

[25] MOUSTAKIDES, G. V. (2006). Decentralized CUSUM change detection. In *Proceedings of the 9th International Conference on Information Fusion (ICIF)* 1-6.

[26] MOUSTAKIDES, G. V. (2008). Sequential change detection revisited. *Annals of Statistics* **36** 787-807.

[27] ØKSENDAL, B. (2002). *Stochastic Differential Equations*. Springer, New York.

[28] PEETERS, B. and ROECK, G. D. (1999). Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical Systems and Signal Processing* **13** 855-877.

[29] POLUNCHENKO, A. S. and TARTAKOVSKY, A. G. (2010). On optimality of the Shiryaev-Roberts procedure for detecting a change in distribution. *Annals of Statistics* **38** 3445-3457.

[30] POOR, H. V. and HADJILIADIS, O. (2008). *Quickest Detection*. Cambridge University Press, Cambridge, UK.

[31] RAGHAVAN, V. and VEERAVALLI, V. V. (July 1-3, 2008). Quickest detection of a change process across a sensor array. In *Proceedings of the 11th International Conference on Information Fusion* 1-8.

[32] SEZER, S. (2010). On the Wiener disorder problem. *The Annals of Applied Probability* **20** 1537-1566.

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

[34] SHREVE, S. E. (2004). *Stochastic Calculus for Finance* **2**. Springer, New York.

[35] TARTAKOVSKY, A. G. and KIM, H. (2006). Performance of certain decentralized distributed change detection procedures. In *Proceedings of the 9th International Conference on Information Fusion (ICIF)* 1-8.

[36] TARTAKOVSKY, A. G. and VEERAVALLI, V. V. (2004). Change-point detection in multichannel and distributed systems with applications. In *Applications of Sequential Methodologies* (N. Mukhopadhay, S. Datta and S. Chattopadhay, eds.) 331-363. Marcel Dekke, New York.

[37] TARTAKOVSKY, A. G. and VEERAVALLI, V. V. (2008). Asymptotically optimum quickest change detection in distributed sensor systems. *Sequential Analysis* **27** 441-475.

[38] ZHANG, H. and HADJILIADIS, O. (2012). Quickest detection in a system with correlated noise. In *Proceedings of the 51st IEEE Conference on Decision and Control* 4757-4763.

[39] ZHANG, H., HADJILIADIS, O., SCHÄFER, T. and POOR, H. V. (2014). Quickest detection in coupled systems. *SIAM Journal on Control and Optimization* **52** 1567-1596.

HONGZHONG ZHANG
STATISTICS DEPARTMENT
COLUMBIA UNIVERSITY
1255 AMSTERDAM AVE
NEW YORK, NY 10027, USA
E-MAIL: hzhang@stat.columbia.edu

NEOFYTOS RODOSTHENOUS
SCHOOL OF MATHEMATICAL SCIENCES
QUEEN MARY UNIVERSITY OF LONDON
LONDON E1 4NS, UNITED KINGDOM
E-MAIL: N.Rodosthenous@qmul.ac.uk

OLYMPIA HADJILIADIS
DEPARTMENT OF MATHEMATICS
BROOKLYN COLLEGE OF THE CITY UNIVERSITY OF NEW YORK
E-MAIL: ohadjiliadis@brooklyn.cuny.edu