Decentralized CUSUM Change Detection

George V. Moustakides
Department of Computer & Communication Engineering
University of Thessaly
Volos, Greece
moustaki@inf.uth.gr

Abstract -*We consider the problem of decentralized change detection using the CUSUM test. More than one sensors acquire independent signals and send their quantized version to a fusion center that uses this information to detect a simultaneous change in all sensors. By introducing a recurrence relation that defines the optimum performance of the CUSUM test for given quantization, we further optimize this measure with respect to the quantization scheme. We compare the resulting optimum test with a simple, asynchronous one shot strategy, where each sensor performs a local CUSUM test and communicates with the fusion center only once to signal its detection.*

Keywords: Change detection, CUSUM, Decentralized detection.

1   Introduction

We consider the problem of sequential change detection under a distributed setup. As it is illustrated in Fig. 1, K sensors acquire sequentially discrete time signals $\{x_n^i\}_{n=1}^\infty$, $i = 1, \ldots, K$, and send the corresponding samples to a fusion center for further processing. In order to reduce communication load the samples x_n^i before transmission are quantized to a finite alphabet with the help of a quantizer $Q^i(x)$. Consequently the fusion center receives $z_n^i = Q^i(x_n^i)$ instead of x_n^i.
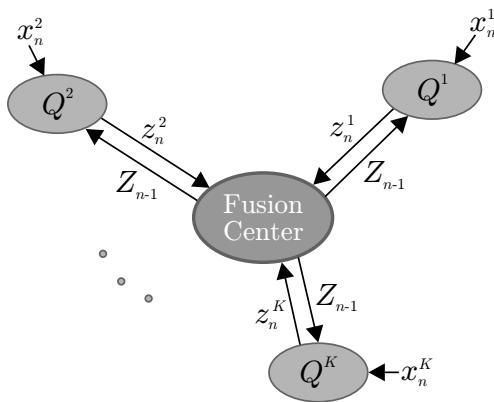

Figure 1: Topology of a distributed change detection scheme.

The problem we would like to solve is the distributed sequential change detection. Specifically we assume that the collection of K signals $\{x_n^i\}$ follows some nominal statistics $\mathbb{P}_\infty$ up to and including some time τ, while after time τ

these statistics change to $\mathbb{P}_0$. Time τ is considered unknown and we would like to detect it as soon as possible. Our goal is to develop a detection scheme that can be implemented at the fusion center. This suggests that the corresponding test can use only the available quantized information $z_n^i = Q^i(x_n^i)$ as well as any existing prior information. Unlike conventional change detection, in the distributed version, in addition to the detection scheme, we also need to specify all quantizers $Q^i(x)$.

This problem has been considered in the past in [8, 9, 6] following a Bayesian approach and using notions introduced in [7]. A very interesting alternative idea is presented in [2] where instead of quantizing the observations, each sensor performs locally its own change detection test and sends its decisions to the fusion center; with the latter being responsible for the final decision.

In this work we consider the change detection problem using Lorden's criterion as our performance measure. With the help of an integral equation we relate the performance of the CUSUM test to the quantizers in the sensors and propose a means to obtain combined quantization/detection strategies that minimize Lorden's criterion. This latter task can only be performed numerically.

In the last part of our work, we follow a similar line of reasoning as in [2] and consider sensors that perform a local change detection test. The sensors are allowed to communicate with the fusion center only *once*, namely whenever they need to signal a detection. The communication with the fusion center is clearly asynchronous allowing for alternative final decision strategies for the fusion center. The performance of different "one shot" schemes is compared against the performance of the optimum quantizer/detector structure of the first part.

2   Background

The general change detection model presented in the Introduction will now be specified and drastically simplified. We recall that even under the conventional change detection setup, a nonasymptotic solution exists only for the i.i.d. case.

Let the signals $\{x_n^i\}$, $i = 1, \ldots, K$, be independent and for each sensor i assume that the acquired samples are i.i.d. before and after the change. More specifically for the density functions before and after time τ we assume that

$$x_n^i \sim \begin{cases} f_\infty^i(x), & \text{for } 1 \le n \le \tau \\ f_0^i(x), & \text{for } \tau < n, \end{cases}$$

where the density functions $f_\infty^i(x), f_0^i(x), i = 1, \ldots, K$, are assumed known, and the time τ is deterministic but unknown.

2.1 Quantization models

As far as the quantizers $Q^i(x), i = 1, \ldots, K$, is concerned, there are different possibilities.

Memoryless Quantizers: The simplest model consists in applying a memoryless nonlinearity of the form

$$z_n^i = Q^i(x_n^i), i = 1, \ldots, K,$$

where $Q^i(x)$ is a function only of x.

Quantizers with Memory: An alternative possibility is to consider quantizers that take into account the data history. More specifically let us denote with $Z_n = [z_n^1, \ldots, z_n^K]$ the data vector arriving at the fusion center at time n after we applied the quantization process. The collection $\mathcal{F}_n = \{Z_n, \ldots, Z_1\}$ constitutes the available information at the fusion center at time n, that can be used to make a decision. It is possible this information to become also available to each sensor, if the fusion center transmits to all sensors (see Fig.1) the vector Z_{n-1} of the previous time instant. Information $\mathcal{F}_{n-1}$ can subsequently be used by each sensor to quantize the next sample x_n^i. In other words we are proposing the following quantization scheme at every sensor i

$$z_n^i = Q^i(x_n^i, \mathcal{F}_{n-1}).$$

As an intermediate step between the memoryless quantizer and the previous one where the whole fusion history $\mathcal{F}_n$ is used, we can consider a scheme where $\mathcal{F}_n$ is summarized with the help of a *finite length vector* $\mathcal{Z}_n$. Such quantizers are particularly attractive in cases where sensor memory is limited. The vector $\mathcal{Z}_{n-1}$ can then be used for the quantization of the next data samples, that is,

$$z_n^i = Q^i(x_n^i, \mathcal{Z}_{n-1}).$$

This can be compactly written as

$$Z_n = Q(X_n, \mathcal{Z}_{n-1}); \qquad (1)$$

where $Q = [Q^1, \ldots, Q^K]^t$ is the collection of all quantizers in a vector function form and $X_n = [x_n^1, \ldots, x_n^K]^t$ is the vector of acquired signal samples at the sensors. The vector of quantized values Z_n will in turn be used to update the vector $\mathcal{Z}_n$ that summarizes the history, in the following way

$$\mathcal{Z}_n = \Phi(Z_n, \mathcal{Z}_{n-1}) \qquad (2)$$

where $\Phi(\cdot, \cdot)$ is a vector function. The two quantization schemes, namely the one that uses the whole history $\mathcal{F}_n$ and the one that uses the summary $\mathcal{Z}_n$ produce the same effect if $\mathcal{Z}_n$ is a *sufficient statistics* for the quantization problem.

2.2 Lorden's criterion

A distributed sequential detection scheme consists of a set of quantizers $Q = [Q^1, \ldots, Q^K]$ and a stopping time (s.t.) T

which is adapted to the filtration $\mathcal{F}$ generated by the quantized data arriving at the fusion center. For each combination $[Q, T]$ we can then define Lorden's performance measure [1]

$$J(Q, T) = \sup_{0 \leq \tau < \infty} \text{essup}\, \mathbb{E}_\tau \left[(T - \tau) \mathbb{1}_{\{T > \tau\}} | \mathcal{F}_\tau \right],$$

where $\mathbb{P}_\tau, \mathbb{E}_\tau$ denote the probability measure and the corresponding expectation induced by the data when the change time is τ. In other words Lorden proposes the use of the worst possible average detection delay with respect to the history before (and including) time τ as a measure of performance of a detection scheme. Although Lorden's performance measure does not explicitly reveal any dependence on the quantization strategies Q, this dependence is present because the s.t. T uses the quantized sequence to decide whether to stop or continue sampling.

The goal is to minimize Lorden's performance measure by suitably selecting the quantities we can control, namely the quantization strategies Q and the s.t. T. Of course, at the same time we also need to assure a minimal false alarm rate. Extending Lorden's original optimality setup [1], we propose the following optimization problem that is consistent with the goal and the requirement we just stated

$$\inf_{Q,T} J(Q, T); \text{ under the constraint: } \mathbb{E}_\infty[T] \geq \gamma; \quad (3)$$

where γ a given constant. In other words we would like to select the quantization scheme and the s.t. that minimize Lorden's criterion assuring at the same time that false alarms will occur at an average period which is no less than a prescribed value γ.

3 The CUSUM test

Lorden's criterion is very closely related to the CUSUM test [1, 3, 4] and our goal is to demonstrate that the optimality properties of CUSUM continue to hold in a distributed setup as well.

Let $\{Z_n\}$ be a sequence of random data that are available sequentially and consider two alternative conditional probability densities for Z_n: $f_\infty(Z|\mathcal{F}_{n-1})$ for nominal condition and $f_0(Z|\mathcal{F}_{n-1})$ for changed. We can then define the CUSUM statistics $\mathcal{S}_n$ by the following recursion

$$\mathcal{S}_n = \max \left\{ \mathcal{S}_{n-1} + \log \frac{f_0(Z_n|\mathcal{F}_{n-1})}{f_\infty(Z_n|\mathcal{F}_{n-1})}, 0 \right\}, \; \mathcal{S}_0 = 0,$$

and the corresponding CUSUM s.t. by

$$T = \inf_n \{n : \mathcal{S}_n \geq \nu\}$$

where ν is a constant threshold. Threshold ν is selected so that the CUSUM test satisfies the false alarm constraint with equality (i.e. $\mathbb{E}_\infty[T] = \gamma$).

If in particular a process $\{\mathcal{Z}_n\}$ summarizes the fusion history and is used for quantization as described in (1),(2), then the CUSUM update becomes

$$\mathcal{S}_n = \max \left\{ \mathcal{S}_{n-1} + \log \frac{f_0(Z_n|\mathcal{Z}_{n-1})}{f_\infty(Z_n|\mathcal{Z}_{n-1})}, 0 \right\}, \; \mathcal{S}_0 = 0, \quad (4)$$

with the whole history $\mathcal{F}_{n-1}$ being replaced by its summary vector $\mathcal{Z}_{n-1}$. It is interesting to note that the pair $(\mathcal{S}_n, \mathcal{Z}_n)$ is Markovian, which allows for application of optimal stopping theory.

3.1 Lower bound for Lorden's criterion

A crucial step in proving optimality of CUSUM with respect to Lorden's criterion is the introduction of a suitable, and simpler to work with, lower bound. We would like this lower bound to *coincide* with Lorden's performance measure in the case of the CUSUM s.t. This assures that, if CUSUM optimizes the lower bound, then it will also optimize the original criterion. The lower bound we are interested in, is given in the following lemma.

Lemma 1 *For every combination* (Q, T) *we have that*

$$J(Q, T) \geq \frac{\mathbb{E}_\infty\left[\sum_{n=0}^{T-1} e^{\mathcal{S}_n}\right]}{\mathbb{E}_\infty[e^{\mathcal{S}_T}]} = \mathcal{J}(Q, T).$$

We have equality whenever T *is equal to the CUSUM s.t.* $\mathcal{T}$.

Proof: The proof follows the same steps as in [3]. ∎

Based on Lemma 1, instead of solving (3) we are going to consider the following alternative constrained optimization problem

$$\inf_{Q,T} \mathcal{J}(Q, T); \text{ under the constraint: } \mathbb{E}_\infty[T] \geq \gamma; \quad (5)$$

where we have replaced the original Lorden measure with the lower bound. If the CUSUM test is optimum in (5) it will clearly be optimum in (3) as well.

4 Optimum schemes

In this section we are going to define the optimum combination quantizer/stopping time. Our work will heavily rely on existing results proposed for alternative performance measures in [6, 8, 9]. We first consider optimum schemes for quantizers with (infinite) memory.

4.1 Quantizers with memory

The first step consists in limiting the search for the optimum s.t. to within the class that satisfies the constraint with equality.

Lemma 2 *For any s.t. with* $\mathbb{E}_\infty[T] > \gamma$ *we can always find an alternative s.t.* T' *with* $\mathbb{E}_\infty[T'] = \gamma$ *and such that* $\mathcal{J}(T, Q) \geq \mathcal{J}(T', Q)$.

Proof: We use the same randomization technique presented in [3] ∎

If c is the optimum value of the problem in (5) and since, due to Lemma 2, without loss of generality we can limit ourselves to s.t. that satisfy the constraint with equality, we can then write for any such s.t. T that

$$\begin{aligned}
\mathcal{U}(T, Q) &= \mathbb{E}_\infty\left[\sum_{n=0}^{T-1} e^{\mathcal{S}_n}\right] - c\mathbb{E}_\infty[e^{\mathcal{S}_T}] - d\mathbb{E}_\infty[T] \\
&= \mathbb{E}_\infty\left[\sum_{n=0}^{T-1}(e^{\mathcal{S}_n} - d) - ce^{\mathcal{S}_T}\right] \geq -d\gamma
\end{aligned}$$

where $d > 0$ a Lagrange multiplier. We can now attempt to solve the *unconstrained* problem

$$\inf_{T,Q} \mathcal{U}(T, Q). \quad (6)$$

It turns out that the solution to the unconstrained problem, for a suitable selection of the parameters c, d can also solve the constrained problem (5). This special parameter combination can be obtained as follows. For any given c, d denote the optimum pair in the sense of (6) as $Q_{c,d}, T_{c,d}$. For every value of $c > 0$ we compute a value $d(c)$ such that the optimum s.t. satisfies $\mathbb{E}_\infty[T_{c,d(c)}] = \gamma$. Among the combinations $c, d(c)$ we look for the specific selection $c = c_\star$ for which $\mathcal{U}(T_{c_\star,d(c_\star)}, Q_{c_\star,d(c_\star)}) = -d(c_\star)\gamma$. This specific parameter selection yields also the optimum for the constrained problem in (5). This is true because for any s.t. T that satisfies the constraint with equality, we have

$$\mathbb{E}_\infty\left[\sum_{n=0}^{T-1}(e^{\mathcal{S}_n} - d(c_\star)) - c_\star e^{\mathcal{S}_T}\right] \geq -d(c_\star)\gamma$$

$$\Leftrightarrow \quad \mathbb{E}_\infty\left[\sum_{n=0}^{T-1} e^{\mathcal{S}_n} - c_\star e^{\mathcal{S}_T}\right] \geq 0$$

$$\Leftrightarrow \quad \frac{\mathbb{E}_\infty\left[\sum_{n=0}^{T-1} e^{\mathcal{S}_n}\right]}{\mathbb{E}_\infty[e^{\mathcal{S}_T}]} \geq c_\star.$$

From the preceding discussion, it is clear that the key for finding the optimum scheme is to be able to obtain the solution to the unconstrained problem in (6). For this, as we said in the beginning of this section, we are going to rely on methodology already developed in [6, 8, 9] for other performance measures. We start by considering the problem of finding the optimum quantizer strategy Q for given s.t. Q. We have the following theorem.

Theorem 1 *For given s.t.* T *the optimum quantizer strategy* Q *that minimizes* $\mathcal{U}(T, Q)$ *is a monotone likelihood ratio quantizer (MLRQ), with thresholds that depend on the history* $\mathcal{F}_{n-1}$.

Proof: We use exactly similar arguments as in [6, 8, 9]. ∎

In view of Theorem 1 we can limit ourselves to MLRQ and continue with the final optimization of $\mathcal{U}(T, Q)$ over all s.t. and the thresholds of the MLRQ.

Theorem 2 *The optimum s.t. that solves (6) is the CUSUM and the thresholds of the corresponding optimum MLRQ depend only on the CUSUM statistics* $\mathcal{S}_{n-1}$.

Proof: As in [6, 8, 9], our claim can be proved by considering the finite horizon problem first and showing, by induction, that in every step n the optimum stopping is a threshold rule on $\mathcal{S}_n$ while the optimum MLRQ involves thresholds that depend on $\mathcal{S}_{n-1}$. Letting then the horizon tend to ∞ transforms all thresholds to stationary (not depending on n). ∎

As a consequence of Theorem 2, if $\mathcal{U}(\mathcal{S}) = \inf_{T,Q} \mathcal{U}(T, Q)$ denotes the optimum cost to go when the initial value of the CUSUM statistics is $\mathcal{S}_0 = \mathcal{S}$, we can write the following Bellman (dynamic programming) equation

$$\mathcal{U}(\mathcal{S}) = \min\left\{ e^{\mathcal{S}} - d + \min_Q \mathbb{E}_\infty[\mathcal{U}(\mathcal{S}_1)|\mathcal{S}_0 = \mathcal{S}], -ce^{\mathcal{S}} \right\} \quad (7)$$

where the minimization over Q is with respect to the thresholds of the MLRQ. Notice that the thresholds will be functions of S since the minimization is performed point-wise for every S. The minimization $\min_Q \mathbb{E}_\infty[\mathcal{U}(S_1)|S_0 = S]$ results in a function $\tilde{\mathcal{U}}(S)$ which yields $\mathcal{U}(S) = \min\{e^S - d + \tilde{\mathcal{U}}(S), -ce^S\}$. We stop sampling whenever the right term is smaller than the left. According to Theorem 2, this happens when $S \geq \nu$. Equ. (7) can serve as a basis for finding the optimum quantization/detection pair, iteratively.

4.2 Memoryless quantizers

In the memoryless case, as was indicated in Subsection 2.1, the quantizers do not depend on the history. Although we do not know what the optimum form of these quantizers is, it seems only natural to consider again MLRQ but with constant thresholds. Proceeding along the same lines of the previous case, we end up with a similar Bellman equation

$$\mathcal{U}(S) = \min\left\{e^S - d + \min_{Q, S=0} \mathbb{E}_\infty[\mathcal{U}(S_1)|S_0 = S], -ce^S\right\} \tag{8}$$

only now the minimization over Q (the thresholds) is performed not for every value of S, but only for $S = 0$.

4.2.1 Example

Let us present in more detail the two-sensor and two-level quantization case. We assume that the two sensor signals before the change are both Gaussian $\mathcal{N}(0, \sigma^2)$ and after the change $\mathcal{N}(\rho, \sigma^2)$. Call $a(\lambda) = \log(\Phi(\frac{\lambda-\rho}{\sigma})/\Phi(\frac{\lambda}{\sigma}))$ and $b(\lambda) = \log([1 - \Phi(\frac{\lambda-\rho}{\sigma})]/[1 - \Phi(\frac{\lambda}{\sigma})])$ where $\Phi(\cdot)$ denotes the Gaussian cdf and λ the common quantization threshold. We then have

$$\mathbb{E}_\infty[\mathcal{U}(S_1)|S] = \Phi^2\left(\frac{\lambda}{\sigma}\right)\mathcal{U}([S+2a]^+)$$
$$+ 2\Phi\left(\frac{\lambda}{\sigma}\right)\left[1 - \Phi\left(\frac{\lambda}{\sigma}\right)\right]\mathcal{U}([S+a+b]^+)$$
$$+ \left[1 - \Phi\left(\frac{\lambda}{\sigma}\right)\right]^2\mathcal{U}([S+2b]^+). \tag{9}$$

Notice that this function depends only on the normalized parameters $\tilde{\lambda} = \lambda/\sigma$ and $\tilde{\rho} = \rho/\sigma$.

For the optimum quantizer with memory, we perform the minimization

$$\tilde{\mathcal{U}}(S) = \min_\lambda \mathbb{E}_\infty[\mathcal{U}(S_1)|S]$$

and obtain a threshold $\lambda(S)$ as a function of S. If on the other hand we consider

$$\tilde{\mathcal{U}}(S) = \min_{\lambda, S=0} \mathbb{E}_\infty[\mathcal{U}(S_1)|S]$$

where the minimization over λ is performed *only* for $S = 0$, this yields the optimum threshold for the memoryless case.

As was indicated above, the Bellman equation can give rise to an iterative computation of the optimum scheme. The iteration we refer to, is the following

$$\mathcal{U}_n(S) =$$
$$\min\left\{e^S - d + \min_{\lambda, (S=0)} \mathbb{E}_\infty[\mathcal{U}_{n-1}(S_1)|S_0 = S], -ce^S\right\},$$

where we initialize with $\mathcal{U}_0(S) = 0$.

5 One shot schemes

In this section we consider schemes in which each sensor performs, locally, a CUSUM test using its own observations. The sensors communicate with the fusion center only *once*, when they must signal a detection. The communication is clearly asynchronous and the fusion center can use the acquired information in different ways. Here we consider the following two strategies

a) *Minimal:* An alarm is issued when any of the sensors reports a detection.

b) *Maximal:* An alarm is issued when all sensors have reported a detection.

For simplicity we consider *continuous time* processes $\{x_t^i\}$ in order to make use of the formulas that are available for CUSUM, for the Brownian Motion case [5]. Again we limit ourselves to the two-sensor case and signals with identical statistics. Specifically we consider the detection of a change from the process $dx_t^i = dw_t^i$ to the $dx_t^i = \mu dt + dw_t^i$, where w_t^i is a standard Wiener processes. In other words in each sensor, before the change, the signal is zero mean Gaussian whereas after the change it has a nonzero mean.

If $\mathcal{T}_i$ is the CUSUM stopping time in each sensor, then the minimal strategy stops at $\mathcal{T}_m = \min\{\mathcal{T}_1, \mathcal{T}_2\}$, whereas the maximal at $\mathcal{T}_M = \max\{\mathcal{T}_1, \mathcal{T}_2\}$. To measure the performance of these two schemes we need to compute the average run length under both hypotheses, that is, $\mathbb{E}_i[\mathcal{T}_m]$, $\mathbb{E}_i[\mathcal{T}_M]$, $i = 0, \infty$. It is easy to see that the two s.t. are related to each other as follows

$$\mathcal{T}_M = \mathcal{T}_1 + \mathcal{T}_2 - \mathcal{T}_m,$$

therefore it suffices to compute the performance of $\mathcal{T}_m$. From [5] we have the following formula for the expectation of the CUSUM s.t. $\mathcal{T}_1$

$$\mathbb{E}_i[\mathcal{T}_1] = \frac{2}{\mu^2}\{e^{-\epsilon_i\nu} + \epsilon_i\nu - 1\} \tag{10}$$

where $\epsilon_0 = 1$, $\epsilon_\infty = -1$ and ν denotes the threshold. Finding however the corresponding expression for the expectation of $\mathcal{T}_m$ (the minimum between two CUSUM s.t.) turns out to be a difficult task. We were able to obtain only a series expansion for the desired expectation. In order to present this result we need the moment generating function of the CUSUM s.t. which will be given in terms of the following functions

$$\Phi_i(\omega) = \frac{\omega e^{\epsilon_i\nu/2}}{\omega\cosh(\omega) + \epsilon_i(\nu/2)\sinh(\omega)}, \quad i = 0, \infty.$$

According to [5], the CUSUM moment generating function can then be written as

$$\mathbb{E}_i[e^{-s\mathcal{T}_1}] = \Phi_i\left(\frac{\nu}{2}\sqrt{1 + \frac{8s}{\mu^2}}\right) = F_i(s), \quad i = 0, \infty.$$

We are interested in the poles of the function $F_i(s)$ and for this we first need to identify the poles of $\Phi_i(\omega)$. Let z_0 be the *positive real* root of the equation $\tanh(z) + (\epsilon_i\nu/2)z = 0$ (there exists at *most* one) and $\{z_n\}, n = 1, 2, \ldots$ the *positive real* roots of the equation $\tan(z) +$

$(\epsilon_i \nu / 2)z = 0$ (there exists an infinite number). If we define $\omega_0 = z_0$ and $\omega_n = j z_n$, then the $\{\omega_n\}$ constitute poles for $\Phi_i(\omega)$. If now we define

$$s_n = \frac{\mu^2}{8}\left(\frac{4\omega_n^2}{\nu^2} - 1\right)$$

then $\{s_n\}$ are the poles of $F_i(s)$. With the help of $\{\omega_n\}$ and $\{s_n\}$ we are now in a position to compute the desired average delay.

Theorem 3 *Let* $\{\omega_n\}, \{s_n\}$ *be the poles defined above, then*

$$\mathbb{E}_i[\mathcal{T}_m] = \mathbb{E}_i[\mathcal{T}_1]$$

$$-\frac{64 e^{\epsilon_i \nu / 2}}{\mu^2 \nu^2} \sum_{n=0(1)}^{\infty} \frac{\Phi_i\left(\frac{\nu}{2}\sqrt{2 - \frac{4\omega_n^2}{\nu^2}}\right)\omega_n^3}{\left(\frac{4\omega_n^2}{\nu^2} - 1\right)^2 \left[\frac{\nu^2}{4} + \frac{\epsilon_i \nu}{2} + \omega_n^2\right]\sinh(\omega_n)}.$$

The sum includes the pole ω_0 *whenever this pole exists. We recall that* $\epsilon_0 = 1$, $\epsilon_\infty = -1$.

Proof: Since $\mathcal{T}_m$ is the minimum between the two random variables $\mathcal{T}_1, \mathcal{T}_2$, we have that $\mathbb{E}_i[\mathcal{T}_m] = \int_0^\infty \mathbb{P}_i^2[\mathcal{T}_1 > t]\,dt$. The Laplace transform of $\mathbb{P}_i[\mathcal{T}_1 > t]$ is equal to $[1 - F_i(s)]/s$ which suggests that the Laplace of $\mathbb{P}_i^2[\mathcal{T}_1 > t]$ is the following convolution in the transform domain

$$\int e^{-st}\mathbb{P}_i^2[\mathcal{T}_1 > t]\,dt = \frac{1}{2\pi j}\int \frac{1 - F_i(w)}{w}\frac{1 - F_i(s-w)}{s - w}dw.$$

By substituting $s = 0$, we obtain

$$\int \mathbb{P}_i^2[\mathcal{T}_1 > t]\,dt = \frac{1}{2\pi j}\int \frac{1 - F_i(w)}{w}\frac{1 - F_i(-w)}{-w}dw$$

$$= \sum_{n=0(1)}^{\infty} \frac{1 - F_i(-s_n)}{s_n^2}\lim_{s \to s_n}(s - s_n)F_i(s).$$

The last equality is a consequence of the residue theorem applied to the poles $\{s_n\}$ of $F_i(s)$. Substituting s_n in terms of ω_n and computing the limit (with the help of Hospital's rule), yields the desired series expansion. ∎

It is interesting to note that the expression for the average delay is of the form $C(\nu)/\mu^2$ where $C(\nu)$ a function of ν only. Notice that this form is also enjoyed by the average delay of the CUSUM s.t., as we can see from (10).
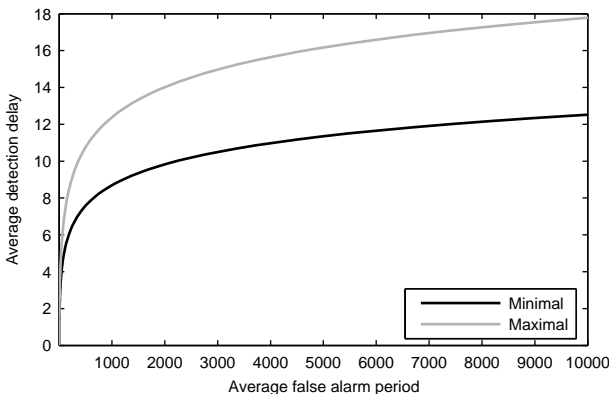

Figure 2: Detection delay as a function of false alarm.

Using the result of Theorem 3 we can now compute the worst detection and false alarm delays and plot the first as a function of the second. Fig. 2 depicts exactly this curve for the minimal and maximal strategy. We can clearly deduce from this figure that the minimal strategy is preferable to the maximal one. In other words, when locally we use the CUSUM tests, it is preferable for the fusion center to issue an alarm whenever it receives the first detection, than to wait until all sensors have acknowledged a detection.

We need to point out that, in the minimal strategy, using the CUSUM s.t. as the local test at the sensors, is by no means optimum. Although CUSUM minimizes the average detection delay at each sensor, this does not necessarily guarantee that it will also minimize the average detection delay of the *minimum* of the local s.t.

6 Comparisons

Let us now compare the minimal, one shot scheme, against the optimum detector of Section 4. We must proceed with caution since the first is a continuous time detection structure whereas the second a discrete time. To be able to apply the discrete time scheme, we are going to assume that we apply uniform sampling with period P. This means that the samples x_n^i are going to be i.i.d. and Gaussian $\mathcal{N}(0, P)$ before the change and $\mathcal{N}(\mu P, P)$ after. Since there is specific sampling involved, we can relate discrete time n to continuous time t using the relation $t = n \times P$. This suggests that the average delay of the scheme in Section 4 must be multiplied by P in order to find its corresponding continuous time performance.

Notice from the Example 4.2.1 and specifically Equ. (9) that the optimum performance in discrete time is a function of the parameter $A = \frac{\rho}{\sigma} = \mu\sqrt{P}$ and the false alarm constraint γ. If therefore we call $\mathcal{C}(A, \gamma)$ the optimum performance of the discrete time scheme, its continuous time analog is equal to $P \times \mathcal{C}(A, \gamma) = \frac{A^2 \mathcal{C}(A, \gamma)}{\mu^2}$, which is again of the form C/μ^2, as in the minimal strategy.

The ultimate point of reference is of course the centralized CUSUM s.t. $\mathcal{T}_o$ performed at the fusion center that has available both continuous time signals. To compute its performance, for the two-sensor and identically distributed signals, we need the log-likelihood ratio u_t, which for this scenario satisfies the following sde

$$\begin{aligned} du_t &= (-\frac{\mu^2}{2}dt + \mu\,dx_t^1) + (-\frac{\mu^2}{2}dt + \mu\,dx_t^2) \\ &= \epsilon_i\mu^2\,dt + \mu(dw_t^1 + dw_t^2) = \epsilon_i\mu^2\,dt + \sqrt{2}\mu\,dw_t, \end{aligned}$$

where w_t^1, w_t^2, w_t are standard Wiener processes. If we use u_t to form the CUSUM s.t. $\mathcal{T}_o$, then the corresponding optimum performance is

$$\mathbb{E}_i[\mathcal{T}_o] = \frac{1}{\mu^2}\{e^{-\epsilon_i\nu} + \epsilon_i\nu - 1\}.$$

Comparing this to the single sensor performance case in (10), we realize that there is a gain factor equal to 2. Again we recover the same form C/μ^2 for the optimum performance, as in the previous two cases.
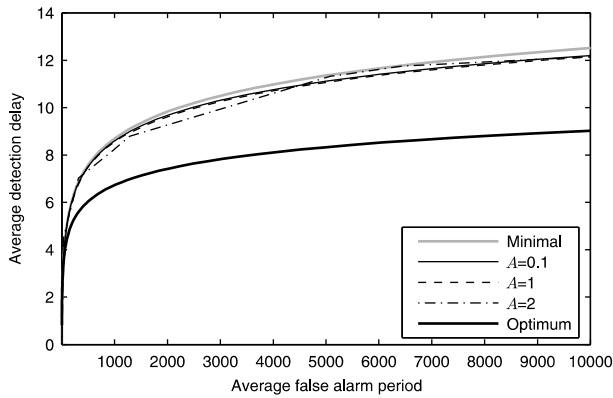
Figure 3: Relative performance of detection schemes.

Since all competing schemes have performance that can be written as C/μ^2, without loss of generality we can assume that the drift parameter is normalized, i.e. $\mu = 1$. In Fig. 3 we plot the average detection delay of the minimal strategy $\mathcal{T}_m$, the optimum $\mathcal{T}_o$ and the optimum quantized scheme for different values of the parameter A. We can see that the simple minimal strategy follows remarkably close the optimum quantized detector. The performance of the latter is in fact relatively insensitive to the parameter A that controls the sampling period. This of course raises serious questions about the practical usefulness of the detection schemes proposed in Section 4, since with a much simpler strategy like $\mathcal{T}_m$, we can obtain comparable results.

Finally we observe that the proposed classical distributed test, as well as, the simple one shot scheme, fall significantly behind the optimum centralized CUSUM in performance. It is clearly new ideas, as the one presented in [2], that can close this important gap.

References

[1] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, pp. 1897-1908, 1971.

[2] Y. Mei, "Information Bounds and Quickest Change Detection in Decentralized Decision Systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2669-2681, July 2005.

[3] G.V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Statist.*, vol. 14, pp. 1379-1387, 1986.

[4] G.V. Moustakides, "Optimality of the CUSUM procedure in continuous time," *Ann. Statist.*, vol. 32, pp. 302-315, 2004.

[5] H.M. Taylor, "A stopped Brownian Motion formula," *Ann. Probab.*, vol. 3, no. 2, pp. 234-246, 1975.

[6] A.G. Tartakovski and V.V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory Probab. Appl.*, vol. 49, no. 3, pp. 458-497, 2005.

[7] J.N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H.V. Poor and J.B. Thomas, Eds. Greenwich, CT: JAI Press, 1990.

[8] V.V. Veeravalli, "Sequential decision fusion: theory and applications," *J. Frank. Inst.*, no. 336, pp. 301-322, 1999.

[9] V.V. Veeravalli, "Decentralized Quickest Change Detection," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1657-1665, May 2001.