Quickest detection in coupled systems

Olympia Hadjiliadis, Tobias Schaefer and H. Vincent Poor

Abstract— **This work considers the problem of quickest detection in a coupled system of N sensors, which receive continuous sequential observations from the environment. The signals received by each sensor are assumed to have the same distribution which allows for a general dependence structure in the observations (e.g. continuous-time analogs of Gaussian autoregressive processes). In particular, the signals are modeled through general Itô processes. Linear state space systems commonly used for structural health -monitoring are special cases of the models assumed. Moreover, the observations received in one sensor, after the onset of a signal (time of a change point in the law of the observations), are coupled by those received in another. Yet, the change points in each sensor can be different because whatever causes the signal to start may arrive at different sensors of the system at different times and thus may be felt at only one sensor. The objective in this paper is the optimal detection of the first time in which the system receives a signal. The problem is formulated as a stochastic optimization problem in which is used as a performance index of detection delay an extended Kullback-Leibler divergence criterion. We consider the case in which the sensors employ cumulative sum (CUSUM) strategies is considered and it is proved that the minimum of N CUSUMs is asymptotically optimal as the mean time between false alarms increases without bound.**

Keywords: Kullback-Leibler divergence, CUSUM, quickest detection

I. INTRODUCTION

We are interested in the problem of quickest detection in a system of N sensors. We consider the situation in which, although the observations in one sensor can affect the observations in another, the actual onset of a signal can occur at different times in each of the N sensors; that is the change points can be different for each of the N sensors. As an example of this case consider a system of sensors monitoring the health of a given structure. Before the change in a given sensor, we have only noise. Then, after the change, the system is vibrating and thus the signal received in any location reflects a vibrating system. That is, observations at any given sensor are coupled with those received in other locations mainly. The change points can be at different times because whatever causes the vibrations to start (i.e. the excitation) may arrive at different structural elements at different times and thus may be felt at only one sensor. Relevant literature of such examples includes, but is not restricted to [1]–[4], [6], [7], [14].

We assume that the probability law of the observations is the same across sensors. This assumption although seemingly restrictive, is realistic in view of the fact the system of sensors is coupled. We model the signals through continuous-time Itô processes. The advantage of such models is the fact that they can capture complex dependencies in the observations. For example, an autoregressive process is a special case of the discrete-time equivalent of an Itô process, namely an Ornstein-Uhlenbeck process. Continuous models are good approximations for observations received at a high rate. Other special cases of this model include Markovian models, and linear state-space systems commonly used in vibration-based structural analysis and health monitoring problems [1]–[4], [6], [7], [14]. It is important to stress that the fact that the system of N sensors is coupled makes the probabilistic treatment of the problem equivalent to the one in which all observations become available in one location.

Our objective is to detect the first onset of a signal in such a system. So far in the literature of this type of problem (see [11], [18]–[21]) it has been assumed that the change points are the same across sensors. Recently the case was also considered of change points that propagate in a sensor array [16]. However, in this configuration the propagation of the change points depends on the <u>unknown</u> 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. The objective is to detect the minimum of the change points. We demonstrate that, in the situation described above, at least asymptotically, the minimum of N CUSUMs is asymptotically optimal in detecting the minimum of the N different change points, as the mean time between false alarms tends to ∞, with respect to an appropriately extended Kullback-Leibler divergence criterion criterion [10] that incorporates the possibility of N different change points.

In the next section we formulate the problem, discuss special cases of our Itô models and demonstrate asymptotic optimality (as the mean time between false alarms tends to ∞), in an extended min-max Kullback-Leibler sense, of the minimum of N CUSUM stopping times. We finally discuss extensions of these results to the case of different structures of observations in each sensor.

This work was not supported by any organization

Olympia Hadjiliadis is with the Department of Mathematics, Brooklyn College, City University of New York, and with the Department of Computer Science, Graduate Center, City University of New York ohadjiliadis@brooklyn.cuny.edu

Tobias Schaefer is with the Department of Mathematics, College of Staten Island, City University of New York, and with the Department of Physics, Graduate Center, City University of New York tobias@math.csi.cuny.edu

H. Vincent Poor is with the Department of Electrical Engineering, Princeton University, poor@princeton.edu

II. FORMULATIONS & RESULTS

We sequentially observe the processes $\{Z_t^{(i)}; t \geq 0\}$ for all $i = 1, \ldots, N$. In order to formalize this problem we consider the measurable space $(\Omega, \mathcal{F})$, where $\Omega = C[0,\infty]^N$ and $\mathcal{F} = \cup_{t>0}\mathcal{F}_t$ with $\mathcal{F}_t = \sigma\{s \leq t; Z_s^{(1)}, \ldots, Z_s^{(N)}\}$.

The processes $\{Z_t^{(i)}; t \geq 0\}$ for all $i = 1, \ldots, N$ are assumed to have the following dynamics:

$$dZ_t^{(i)} = \begin{cases} dw_t^{(i)} & t \leq \tau_i \\ \alpha_t^{(i)}\, dt + dw_t^{(i)} & t > \tau_i, \end{cases} \tag{1}$$

where $\{\alpha_t^{(i)}; t \geq 0\}$ is a process on the same probability space adapted to the filtration $\{\mathcal{F}_t\}$ and $\{w_t^{(i)}; t \geq 0\}$ are independent standard Brownian motion. The case considered in this paper that in which $\alpha_t^{(i)}$ is the same for all i. This can be described as a signal symmetry across sensors.

We notice that $\{\mathcal{F}_t\}$ is the filtration generated by the observations received by all sensors. Thus by requiring that $\alpha_t^{(i)}$ be $\mathcal{F}_t$-measurable for all i , we have managed to capture the coupled nature of the system. In particular, in the special case in which, say, $\alpha_t^{(1)} = -r\sum_{i=1}^N Z_t^{(i)}$, (1) describes a process which displays an autoregressive (or its continuous equivalent [12]) behavior in $\{Z_t^{(1)}; t \geq 0\}$, while still being coupled with the observations received by the other sensors. More specifically, the magnitude of each increment of the process $\{Z_t^{(1)}; t \geq 0\}$ at each instant t is not only affected by $Z_t^{(1)}$ but also by $Z_t^{(i)}$, $i = 2, \ldots, N$ the observations at sensor $2, \ldots, N$. This couples the observations received in sensor 1 with those received in sensors $2, \ldots, N$ at each instant t and results in a system of interdependent sensors. We notice that the special case described above can also be written in the form of a linear state-space model as follows:

$$\begin{aligned} d\begin{pmatrix} Z_t^{(1)} \\ \ldots \\ Z_t^{(N)} \end{pmatrix} &= -r\begin{pmatrix} 1 & \ldots & 1 \\ \ldots & \ldots & \ldots \\ 1 & \ldots & 1 \end{pmatrix}\begin{pmatrix} Z_t^{(1)} \\ \ldots \\ Z_t^{(N)} \end{pmatrix} dt \\ &+ I\begin{pmatrix} dW_t^{(1)} \\ \ldots \\ dW_t^{(N)} \end{pmatrix}. \end{aligned}$$

On the space Ω, we have the following family of probability measures $\{P_{\tau_1,\ldots,\tau_N}\}$, where $P_{\tau_1,\ldots,\tau_N}$ corresponds to the measure generated on Ω by the processes $(Z_t^{(1)}, \ldots, Z_t^{(N)})$ when the change in the N-tuple process occurs at time point τ_i, $i = 1, \ldots, N$. Notice that the measure $P_{\infty,\ldots,\infty}$ corresponds to the measure generated on Ω by N independent standard Brownian motions.

Our objective is to find a stopping rule T that balances the trade-off between a small detection delay subject to a lower bound on the mean-time between false alarms and will ultimately detect $\min\{\tau_1, \ldots, \tau_N\}$. In what follows we will use $\tilde{\tau}$ to denote $\min\{\tau_1, \ldots, \tau_N\}$.

To this effect we propose a generalization of the J_{KL} of [10], namely:

$$J_{KL}^{(N)}(T) = \sup_{\tau_1,\ldots,\tau_N} \text{essup } \mathrm{E}_{\tau_1,\ldots,\tau_N}\left\{\frac{1}{2}\left(\sum_{\mathrm{i}=1}^{\mathrm{N}}\int_{\tau_\mathrm{i}}^{\mathrm{T}}(\alpha_\mathrm{s}^{(\mathrm{i})})^2 \mathrm{ds}\right) \mid \mathcal{F}_{\tilde{\tau}}\right\}, \tag{2}$$

where the supremum over $\tau_1, \ldots, \tau_N$ is taken over the set in which $\min\{\tau_1, \ldots, \tau_N\} < \infty$. That is, we consider the worst detection delay over all possible realizations of paths of the N-tuple of stochastic processes $(Z_t^{(1)}, \ldots, Z_t^{(N)})$ up to $\min\{\tau_1, \ldots, \tau_N\}$ and then consider the worst detection delay over all possible N-tuples $\{\tau_1, \ldots, \tau_N\}$ over a set in which at least one of them is forced to take 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. This gives rise to the following stochastic optimization problem:

$$\begin{gathered} \inf_T J_{KL}^{(N)}(T), \quad \text{subject to} \\ E_{\infty,\ldots,\infty}\left\{\tfrac{1}{2}\int_0^T \sum_{i=1}^N (\alpha_s^{(i)})^2 ds\right\} \geq \gamma. \end{gathered} \tag{3}$$

The criterion in (2) can be similarly motivated by considering the Kullback-Leibler divergence:

$$\begin{aligned} &E_{\tau_1,\ldots,\tau_N}\left\{\log \frac{dP_{\tau_1,\ldots,\tau_N}}{dP_{\infty,\ldots,\infty}}\Big|_{\mathcal{F}_t}\Big| \mathcal{F}_{\tilde{\tau}}\right\} \\ &= E_{\tau_1,\ldots,\tau_N}\left\{\frac{1}{2}\sum_{i=1}^N \int_{\tau_i}^t (\alpha_r^{(i)})^2 dr | \mathcal{F}_{\tilde{\tau}}\right\} \end{aligned} \tag{4}$$

where the last equality follows as long as

$$E_{\tau_1,\ldots,\tau_N}\left\{\int_{\tilde{\tau}}^t (\alpha_r^{(i)})^2 dr \Big| \mathcal{F}_{\tilde{\tau}}\right\} < \infty \quad a.s. \tag{5}$$

for all $i = 1, \ldots, N$ and all $t < \infty$.

Using an argument similar to the randomization argument of [5], it is also possible to show that the optimal stopping rule T^* must be an equalizer rule in that it would react at exactly the same time regardless of which change takes place first. In order to demonstrate this fact we begin by noting that minimization of (2) is equivalent to minimizing

$$\sup_{\tau_1,\ldots,\tau_N} \text{essup } \mathrm{E}_{\tau_1,\ldots,\tau_N}\left\{\frac{1}{2}\left(\sum_{\mathrm{i}=1}^{\mathrm{N}}\int_{\tilde{\tau}}^{\mathrm{T}}(\alpha_\mathrm{s}^{(\mathrm{i})})^2 \mathrm{ds}\right) \mid \mathcal{F}_{\tilde{\tau}}\right\} \tag{6}$$

Now define

$$J_i^{(N)}(T) = \sup_{\tau_i \leq \tau_j, j \neq i} \text{essup} E_{\tau_1,\ldots,\tau_N}\left\{\left(\frac{1}{2}\sum_{i=1}^N \int_{\tau_i}^T (\alpha_s^{(i)})^2 ds\right)\Big| \mathcal{F}_t\right\},$$

for $i = 1, \ldots, N$. That is, $J_i^{(N)}(T)$ is the detection delay of the stopping rule T when $\tau_i \leq \min_{j\neq i}\{\tau_j\}$. Then

$$J_{KL}^{(N)}(T) = \max\left\{J_1^{(N)}(T), J_2^{(N)}(T), \ldots, J_N^{(N)}(T)\right\}. \tag{7}$$

The optimal solution to (3), T^*, satisfies

$$(8)\quad J_1^{(N)}(T^*) = J_2^{(N)}(T^*) = \ldots = J_N^{(N)}(T^*).$$

To see this, let us consider the case when $N = 2$. Let T be a stopping rule such that $J_1^{(2)}(T) < J_2^{(2)}(T)$. Consider another stopping rule S, which stops as T does, but observes $Z_t^{(2)}$ in place of $Z_t^{(1)}$ and $Z_t^{(1)}$ in place of $Z_t^{(2)}$. It follows that

$$J_1^{(2)}(S) = J_2^{(2)}(T) \quad \text{and} \quad J_2^{(2)}(S) = J_1^{(2)}(T).$$

We trivially also have that

$$E_{\infty,\infty}\{S\} = E_{\infty,\infty}\{T\}.$$

Now let us use a binary random variable $X \in \{0, 1\}$, which is independent of $\{\mathcal{F}_t\}$, to construct a randomized stopping rule adapted to $\hat{\mathcal{F}}_t = \mathcal{F}_t \vee \sigma(X)$,

$$(9)\quad \hat{T} = XT + (1 - X)S.$$

It is easy to observe that

$$E_{\infty,\infty}\{\hat{T}\} = E_{\infty,\infty}\{T\},$$

and

$$J_1^{(2)}(\hat{T}) = J_2^{(2)}(\hat{T}) = \frac{1}{2}\left[J_1^{(2)}(T) + J_2^{(2)}(T)\right] < J_2^{(2)}(T),$$

which implies

$$J^{(2)}(\hat{T}) < J^{(2)}(T),$$

by (7). Therefore the optimal solution to (3) must satisfy (8)[1].

For a fixed i, and the dynamics of (1) the CUSUM stopping rule is

$$(10)\quad T = \inf\{t \geq 0; y_t^{(i)} = \nu\},$$

where

$$(11)\quad y_t^{(i)} = u_t^{(i)} - m_t^{(i)},\ i = 1, \ldots, N$$

with $m_t^{(i)} = \inf_{s \leq t} u_s^{(i)}$, $i = 1, \ldots, N$ and

$$(12)\quad u_t^{(i)} = \int_0^t \alpha_s^{(i)} dZ_s^{(i)} - \frac{1}{2}\int_0^t (\alpha_s^{(i)})^2 ds.$$

In the case that $N = 1$, in which the drift denoted by α_t is measurable with respect to the filtration generated by only one process, say $\{Z_t; t \geq 0\}$ the CUSUM stopping rule (10) is optimal in minimizing the Kullback-Leibler divergence criterion of [10] subject to the false alarm constraint $E_\infty\{T\} \geq \gamma$. The ν in (10) is chosen so that $E_\infty\left\{\int_0^{T_\nu} \alpha_t^2 dt\right\} \equiv f(\nu) = \gamma$, with $f(\nu) = e^\nu - \nu - 1$ (see [10]) and

$$(13)\quad J_{KL}^{(1)}(T_\nu) \equiv E_0\left\{\int_0^{T_\nu} \alpha_t^2 dt\right\} = f(-\nu).$$

[1]Although $\hat{T}$ of equation (9) is measurable with respect to the enlarged filtration $\left\{\hat{\mathcal{F}}_t\right\}$, the optimal solution to (3) must be adapted to the original filtration $\{\mathcal{F}_t\}$.

The fact that the worst detection delay is the same as that incurred in the case in which the change point is exactly 0 is a consequence of the non-negativity of the CUSUM process, from which it follows that the worst detection delay occurs when the CUSUM process at the time of the change is at 0 [10].

The CUSUM stopping rule (10) is an optimal solution to one-dimensional problem of detecting one change-point in the one-dimensional equivalent of (3). The details can be found in [10] and [15]. It is important however to point out that a vital assumption necessary for the optimality of the CUSUM (10) in [10] is

$$\begin{aligned} P_{\tau_i}\left(\int_0^\infty {\alpha_s}^2 ds = \infty\right) &= P_\infty\left(\int_0^\infty {\alpha_s}^2 ds = \infty\right) \\ (14)\qquad &= 1. \end{aligned}$$

This assumption ensures the a.s. finiteness of the CUSUM stopping time (see [8]), whose physical interpretation is that the signal received after the change point has sufficient energy. We will thus assume that conditions (14) are satisfied for all processes $\{\alpha_s^{(i)}\}$.

We remark here that if the N change points were the same then the problem (3) is equivalent to observing only one stochastic process which is now N-dimensional. Thus, in this case, the detection delay and mean time between false alarms are given by the formulas in the above paragraph.

Returning to problem (3), it is easily seen that in seeking solutions to this problem, we can restrict our attention to stopping times that achieve the false alarm constraint with equality [9]. The optimality of the CUSUM stopping rule in the presence of only one observation process suggests that a CUSUM type of stopping rule might display similar optimality properties in the case of multiple observation processes. In particular, an intuitively appealing rule, when the detection of $\min\{\tau_1, \ldots, \tau_N\}$ is of interest, is $T_h = T_h^1 \wedge \ldots \wedge T_h^N$, where T_h^i is the CUSUM stopping rule for the process $\{Z_t^{(i)}; t \geq 0\}$ for $i = 1, \ldots, N$. That is, we use what is known as a multi-chart CUSUM stopping time [17], which can be written as

$$(15)\quad T_h = \inf\left\{t \geq 0; \max\{y_t^{(1)}, \ldots, y_t^{(N)}\} \geq h\right\},$$

where

$$y_t^{(i)} = \sup_{0 \leq \tau_i \leq t} \log \left.\frac{dP_{\tau_i}}{dP_\infty}\right|_{\mathcal{F}_t},$$

and the P_{τ_i} are the restrictions of the measure $P_{\tau_1,\ldots,\tau_N}$ to $C[0, \infty)$.

It is easy to see that (15) is an equalizer rule. That is, it satisfies (8). This follows from the assumption that $\{\alpha_t^{(i)}\}$ are the same for all i.

Moreover,

$$\begin{aligned} J_{KL}^{(N)}(T_h) = E_{0,\infty,\ldots,\infty}\{T_h\} &= E_{\infty,0,\infty,\ldots,\infty}\{T_h\} \\ &= \ldots \\ (16)\qquad &= E_{\infty,\ldots,\infty,0}\{T_h\}. \end{aligned}$$

This is because the worst detection delay occurs when at least one of the N processes does not change regime. Thus, the worst detection delay will occur when none of the other processes changes regime and due to the non-negativity of the CUSUM process the worst detection delay will occur when the remaining one processes is exactly at 0.

Notice that the threshold h is used for the multi-chart CUSUM stopping rule (15) in order to distinguish it from ν the threshold used for the one sided CUSUM stopping rule (10).

In what follows we will demonstrate asymptotic optimality of (15) as $\gamma \to \infty$. In view of the discussion in the previous paragraph, in order to assess the optimality properties of the multi-chart CUSUM rule (15) we will thus need to begin by evaluating $E_{0,\infty,\ldots,\infty}\left\{\int_0^{T_h}\sum_{i=1}^N(\alpha_t^{(i)})^2dt\right\}$ and $E_{\infty,\ldots,\infty}\left\{\int_0^{T_h}\sum_{i=1}^N(\alpha_t^{(i)})^2dt\right\}$.

In order to demonstrate asymptotic optimality of (15) we bound the detection delay $J_{KL}^{(N)}$ of the unknown optimal stopping rule T^* by

$$E_{0,\infty,\ldots,\infty}\left\{\int_0^{T_h}\sum_{i=1}^N(\alpha_t^{(i)})^2dt\right\} \quad > \quad J_{KL}^{(N)}(T^*), \tag{17}$$

where h is chosen so that

$$E_{\infty,\ldots,\infty}\left\{\int_0^{T_h}\sum_{i=1}^N(\alpha_t^{(i)})^2dt\right\} \quad = \quad \gamma. \tag{18}$$

It is also obvious that $J_{KL}^{(N)}(T^*)$ is bounded from below by the detection delay of the one CUSUM when there is only one observation process, in view of the fact that

$$\begin{aligned}&\sup_{\tau_1,\ldots,\tau_N}\mathrm{essup}E_{\tau_1,\ldots,\tau_N}\left\{\int_{\tilde{\tau}}^T\sum_{i=1}^N(\alpha_t^{(i)})^2dt|\mathcal{F}_{\tilde{\tau}}\right\}\geq\\ &\geq\sup_{\tau_1}\mathrm{essup}\,E_{\tau_1}\left\{\int_{\tau_1}^T\alpha_t^2dt|\mathcal{F}_{\tau_1}^{(1)}\right\},\end{aligned} \tag{19}$$

where α_t is measurable w.r.t. the filtration generated by the 1-dimensional process $\{Z_t^{(1)}\}$, denoted by $\{\mathcal{F}_t^{(1)}\}$, and is the projection of $\{\alpha_t^{(i)}\}$ on the filtration $\{\mathcal{F}_t^{(1)}\}$. It is important to stress here that inequality (19) is valid as long as T is measurable with respect to the filtration $\{\mathcal{F}_t^{(1)}\}$.

The stopping time that minimizes the right hand side is the CUSUM stopping rule T_ν of (10), with ν chosen so as to satisfy

$$E_\infty\left\{\int_0^{T_\nu}\alpha_t^2dt\right\} \quad = \quad \gamma. \tag{20}$$

We will demonstrate that the difference between the upper and lower bounds

$$\begin{aligned}E_{0,\infty,\ldots,\infty}\left\{\int_0^{T_h}\sum_{i=1}^N(\alpha_t^{(i)})^2dt\right\} \quad &> \quad J_{KL}^{(N)}(T^*)\\ &> \quad E_0\left\{\int_0^{T_\nu}\alpha_t^2dt\right\},\end{aligned} \tag{21}$$

is bounded by a constant as $\gamma \to \infty$, with h and ν satisfying (18) and (20), respectively.

Lemma 1: Suppose that $\{\alpha_t^{(i)}\}$ are the same for all i. We have

$$E_{0,\infty,\ldots,\infty}\left\{\int_0^{T_h}\sum_{i=1}^N(\alpha_t^{(i)})^2dt\right\} = [\log\gamma+\log N-1+o(1)], \tag{22}$$

as $\gamma \to \infty$

Proof: Please refer to the Appendix for a sketch of the proof. Moreover, it is easily seen from (13) that

$$E_0\left\{\int_0^{T_\nu}\alpha_t^2dt\right\} \quad = \quad [\log\gamma-1+o(1)]. \tag{23}$$

Thus we have the following result.

Theorem 1: Suppose that $\{\alpha_t^{(i)}\}$ are the same for all i. Then the difference in detection delay $J_{KL}^{(N)}$ of the unknown optimal stopping rule T^* and the detection delay of T_h of (15) with h satisfying (18) is bounded above by

$$\log N,$$

as $\gamma \to \infty$.

Proof: The proof follows from Lemma 1 and (23).

Remark: Since $J_{KL}^{(N)}(T_h)$ increases without bound as $\gamma \to \infty$, Theorem 1 asserts the asymptotic optimality of T_h.

III. CONCLUSIONS AND FUTURE WORKS

In this paper we have demonstrated the asymptotic optimality of the minimum of N CUSUMs for detecting the minimum of N different change points in a coupled system of N sensors which receive sequential observations from the environment. We have allowed for a general dependence structure in the observations and we have shown that the N-CUSUM stopping rule is asymptotically optimal, as the mean time to the first false alarm increases without bound, in detecting the minimum of N different change-points in the sense that it minimizes a worst Kullback-Leibler divergence criterion. This has been seen by the fact that the difference in detection delay of the proposed N-CUSUM stopping rule and the unknown optimal stopping rule is bounded above by the constant $\log N$. An interesting extension of this work would incorporate the fact that the distributions of the signals received in different sensors may be different. In this case the fact that the optimal stopping rule has to be an equalizer rule (i.e. satisfy (8)) would determine the optimal selection of thresholds in each sensor which in the general case should be different.

IV. ACKNOWLEDGMENTS

The authors are grateful to Professor Jay Rosen for promoting this collaboration.

REFERENCES

[1] M. Basseville, M. Abdelghani, A. Benveniste, Subspace-based fault detection algorithms for vibration monitoring, *Automatica*, Vol. 36, No.1, 2000, pp. 101- 109.

[2] M. Basseville, A. Benveniste, M. Goursat, L. Mevel, 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*, Vol. 2007, No. 1, 2007, pp. 200-213.

[3] M. Basseville, L. Mevel, M. Goursat, Statistical model-based damage detection and localization, subspace-based residuals and damage-to-noise sensitivity ratios, *Journal of Sound and Vibration*, Vol. 275, No. 3-5, 2004, pp. 769-794.

[4] D. J. Ewins, *Modal Testing: Theory, Practice and Applications*, Research Studies Press, Letchworth, Hertfordshire UK, 2nd edition, 2000.

[5] O. Hadjiliadis, H. Zhang and H.V. Poor, One shot schemes for decentralized quickest change detection, *IEEE Transactions of Information Theory*, to appear.

[6] W. Heylen, S. Lammens, and P. Sas, *Modal Analysis Theory and Testing*, Department of Mechanical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, 1995.

[7] J. N. Juang, *Applied System Identification*, Prentice Hall, Englewood Cliffs, NJ, USA, 1994.

[8] R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes II*, Springer-Verlag, Berlin, 1979.

[9] G. V. Moustakides, Optimal stopping times for detecting changes in distributions, *Annals of Statistics*, Vol. 14, No. 4, pp. 1379-1387, 1986.

[10] G. V. Moustakides, Optimality of the CUSUM procedure in continuous time, *Annals of Statistics*, Vol. 32, No. 1, 2004, pp. 302-315.

[11] G. V. Moustakides, Decentralized CUSUM change detection, *Proceedings of the 9th International Conference on Information Fusion (ICIF)*, Florence, Italy, 2006, pp. 1 - 6.

[12] A. A. Novikov, On the first passage time of an autoregressive process over a level and an application to a "disorder" problem, *Theory of Probability and its Applications*, Vol. 35, No. 2, 1987.

[13] B. Oksendal, *Stochastic Differential Equations*, Springer-Verlag, New York, 2002.

[14] B. Peeters and G. De Roeck, Reference-based stochastic subspace identification for output-only modal analysis, *Mechanical Systems and Signal Processing*, Vol. 13, No. 6, 1999, pp. 855-877.

[15] H. V. Poor and O. Hadjiliadis, *Quickest Detection*, Cambridge University Press, Cambridge UK, 2008.

[16] V. Raghavan and V. V. Veeravalli, Quickest detection of a change process across a sensor array, *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, July 1-3 2008.

[17] A. G. Tartakovsky, Asymptotic performance of a multichart CUSUM test under false alarm probability constraint, *Proceedings of the 44th IEEE Conference on Decision and Control*, Seville, Spain, December 12 - 15 2005, pp. 320 - 325.

[18] A. G. Tartakovsky and H. Kim, Performance of certain decentralized distributed change detection procedures, *9th International Conference on Information Fusion (ICIF)*, Florence, Italy, July 10-13 2006, pp. 1-8.

[19] A. G. Tartakovsky and V. V. Veeravalli, Quickest change detection in distributed sensor systems, *Proceedings of the 6th Conference on Information Fusion*, Cairns, Australia, July 8-11 2003.

[20] A. G. Tartakovsky and V. V. Veeravalli, Change-point detection in multichannel and distributed systems with applications, in *Applications of Sequential Methodologies*, pp. 331-363, (N. Mukhopadhay, S.Datta and S. Chattopadhay, Eds), Marcel Dekker, New York, 2004.

[21] A. G. Tartakovsky and V. V. Veeravalli, Asymptotically optimum quickest change detection in distributed sensor systems, *Sequential Analysis*, to appear.

V. APPENDIX

As an illustration for the general case, let us prove the result for $N = 2$.

We begin by deriving the Partial Differential equations satisfied by the functions

- $\tilde{S}(\tilde{x},\tilde{y}) = E_{0,\infty}^{(\tilde{x},\tilde{y})}\left\{\int_0^{T_h}\sum_{i=1}^2(\alpha_t^{(i)})^2dt\right\}$,
- $\tilde{T}(\tilde{x},\tilde{y}) = E_{\infty,\infty}^{(\tilde{x},\tilde{y})}\left\{\int_0^{T_h}\sum_{i=1}^2(\alpha_t^{(i)})^2dt\right\}$,

where the subscript $(\tilde{x},\tilde{y})$ indicates the indicates the initial value of the pair of CUSUM processes $(y_t^{(1)}, y_t^{(2)})$. With this representation, it is easy to see that $E_{0,\infty}\left\{\int_0^{T_h}\sum_{i=1}^2(\alpha_t^{(i)})^2dt\right\} = \tilde{S}(0,0)$ and $E_{\infty,\infty}\left\{\int_0^{T_h}\sum_{i=1}^2(\alpha_t^{(i)})^2dt\right\} = \tilde{T}(0,0)$. In the sequel we will denote by $\tilde{T}_{\tilde{x}}$, $\tilde{T}_{\tilde{y}}$, $\tilde{S}_{\tilde{x}}$, $\tilde{S}_{\tilde{y}}$, the first partial derivatives of $\tilde{T}$ and $\tilde{S}$ with respect to $\tilde{x}$ and $\tilde{y}$ respectively. Similarly, we will denote by $\tilde{T}_{\tilde{x}\tilde{x}}$, $\tilde{T}_{\tilde{y}\tilde{y}}$, $\tilde{S}_{\tilde{x}\tilde{x}}$, $\tilde{S}_{\tilde{y}\tilde{y}}$ the second partials.

Using Itô's rule [13], we have

(24)
$$\tilde{T}(y_t^{(1)}, y_t^{(2)}) - \tilde{T}(\tilde{x},\tilde{y}) = \int_0^t (\alpha_s^{(1)})^2\tilde{T}_{\tilde{x}}dw_s^{(1)} + (\alpha_s^{(2)})^2\tilde{T}_{\tilde{y}}dw_s^{(2)}$$
$$-\int_0^t (\alpha_s^{(1)})^2\tilde{T}_{\tilde{x}}(y_s^{(1)}, y_s^{(2)})dm_s^{(1)} + (\alpha_s^{(2)})^2\tilde{T}_{\tilde{y}}(y_s^{(1)}, y_s^{(2)})dm_s^{(2)}$$
$$+\int_0^t \alpha_s^{(i)}(\tilde{T}_{\tilde{x}\tilde{x}} + \tilde{T}_{\tilde{y}\tilde{y}} - \tilde{T}_{\tilde{x}} - \tilde{T}_{\tilde{y}})ds,$$

where the arguments of each of the above functions are $(\tilde{y}_s^{(1)}, \tilde{y}_s^{(2)})$ when omitted and where in the last line we use the fact that $\alpha_s^{(i)}$ are of the same form for all i. Evaluating the above equation at T_h and taking expectations under the $P_{\infty,\infty}$ measure, while using conditions (5), (14), we obtain that $\tilde{T}$ has to satisfy

$$\tilde{T}_{\tilde{x}\tilde{x}} + \tilde{T}_{\tilde{y}\tilde{y}} - \tilde{T}_{\tilde{x}} - \tilde{T}_{\tilde{y}} = -1,\ (\tilde{x},\tilde{y}) \in \tilde{\mathcal{D}} = [0,h]^2, \tag{25}$$

with the Dirichlet boundary conditions

$$\tilde{T}(\tilde{x},\tilde{y})|_{\tilde{x}=h} = \tilde{T}(\tilde{x},\tilde{y})|_{\tilde{y}=h} = 0 \tag{26}$$

and the Neumann boundary conditions

$$\left.\frac{\partial\tilde{T}}{\partial\tilde{x}}\right|_{\tilde{x}=0} = \left.\frac{\partial\tilde{T}}{\partial\tilde{y}}\right|_{\tilde{y}=0} = 0. \tag{27}$$

Notice that the Neumann boundary conditions ensure that the terms in the second line of (24) vanish. Similarly, $\tilde{S}$ satisfies

$$\tilde{S}_{\tilde{x}\tilde{x}} + \tilde{S}_{\tilde{y}\tilde{y}} + \tilde{S}_{\tilde{x}} - \tilde{S}_{\tilde{y}} = -1,\ (\tilde{x},\tilde{y}) \in \tilde{\mathcal{D}} = [0,h]^2, \tag{28}$$

with the same boundary conditions as $\tilde{T}$.

We can now introduce a change of variable $x = \frac{\tilde{x}}{h}$ and $y = \frac{\tilde{y}}{h}$. By setting $\epsilon = \frac{1}{h}$, we can rewrite (25) as

(29)
$$\epsilon^2\tilde{T}_{xx} + \epsilon^2\tilde{T}_{yy} - \epsilon\tilde{T}_x - \epsilon\tilde{T}_y = -1,\ (x,y) \in \mathcal{D} = [0,1]^2,$$

with the Dirichlet boundary conditions

$$\tilde{T}(x,y)|_{x=1} = \tilde{T}(x,y)|_{y=1} = 0 \tag{30}$$

and the Neumann boundary conditions (27). By letting $\epsilon\tilde{T} = T$, we now obtain

(31)
$$\epsilon T_{xx} + \epsilon T_{yy} - T_x - T_y = -1,\ (x,y) \in \mathcal{D} = [0,1]^2,$$

with T satisfying the Dirichlet boundary conditions of (30) and the Neumann condition of (27). We are interested in the asymptotics of $T(0,0)$ for small values of ϵ (or equivalently large values of h). $T(0,0)$ can be interpreted as the mean exit time of a particle that is placed initially at the origin, with reflecting boundaries along the axes and absorbing boundaries on the top and the right side of the rectangular domain $\mathcal{D}$. In order to solve the above problem, we note, that we can write the solution T as

$$T(x,y) = \int_0^\infty G(x,y,t)\,dt \tag{32}$$

where G denotes the probability that the particle, initially placed at a point (x,y) in $\mathcal{D}$ leaves the domain $\mathcal{D}$ at a time $\tau > t$. The evolution of G is then governed by the backward Fokker-Planck equation:

$$\frac{\partial G}{\partial t} = \epsilon \Delta G - \frac{\partial G}{\partial x} - \frac{\partial G}{\partial y}\,. \tag{33}$$

Boundary conditions for G correspond to boundary conditions of T and the initial condition of G is given by the fact that, at $t = 0$, G has the value 1 in $\mathcal{D}$.

In the case of the particular geometry under consideration, we can find an approximate solution to (33) and use this to find T. This is due to the fact that, for a rectangular domain under the assumptions given, the solution of (33) can be found by simple separation of variables, hence we find G as a product of the form

$$G(x,y,t) = G_1(x,t)G_2(y,t), \tag{34}$$

where G_1 satisfies the equation

$$\frac{\partial G_1}{\partial t} = \epsilon \frac{\partial^2 G_1}{\partial x^2} - \frac{\partial G_1}{\partial x} \tag{35}$$

on $[0,1]$ with reflecting boundary at 0 and absorbing boundary at 1. The same holds for G_2 with respect to the variable y.

In order to solve (35), we apply a Laplace transform in t and obtain for $\tilde{G}_1 = \tilde{G}_1(s,x)$ the ordinary differential equation

$$s\tilde{G}_1 - 1 = \epsilon \tilde{G}_1'' - \tilde{G}_1'. \tag{36}$$

Making use of the fact that ϵ is small, we find as leading order approximation to the solution of (36):

$$\tilde{G}_1(0,s) \approx \frac{\epsilon\,\mathrm{e}^{1/\epsilon}}{\epsilon s\,\mathrm{e}^{1/\epsilon} + 1}. \tag{37}$$

For this approximation it is simple to find the inverse Laplace transform to obtain

$$G_1(0,t) \approx \exp\left(-\frac{1}{\epsilon}\mathrm{e}^{-1/\epsilon} t\right). \tag{38}$$

Using this formula for both $G_1(0,t)$ and $G_2(0,t)$ we obtain immediately for $T(0,0)$ in (31) the asymptotic formula

$$T(0,0) \approx \frac{1}{2}\epsilon\,\mathrm{e}^{1/\epsilon} \tag{39}$$

from which it follows that $\tilde{T}(0,0) \approx \frac{1}{2}\,\mathrm{e}^{1/\epsilon}$. Setting $\tilde{T}(0,0) = \gamma$, and using $h = \frac{1}{\epsilon}$, we further obtain that as $\gamma \to \infty$, $h \approx \log\gamma + \log 2$.

For the asymptotic formula of $\tilde{S}(0,0)$ of (28), we also let $S = \epsilon\tilde{S}$ and use the same change of variable as in the previous case. The only difference is that we have to solve for $\tilde{G}_1$ the different problem

$$s\tilde{G}_1 - 1 = \epsilon \tilde{G}_1'' + \tilde{G}_1'. \tag{40}$$

In this case, the approximate solution takes the form

$$\tilde{G}_1(0,s) \approx \frac{1}{s}\left(1 - \frac{\mathrm{e}^{-s}}{1+\epsilon\,s}\right) - 2\epsilon\,\mathrm{e}^{-s}. \tag{41}$$

From here we obtain after inverse Laplace transform

$$G_1(0,t) \approx \mathcal{H}(1-t) + \mathcal{H}(t-1)\,\mathrm{e}^{-(t-1)/\epsilon} - 2\epsilon\,\delta(t-1) \tag{42}$$

where $\mathcal{H}$ denotes the Heaviside function and δ denotes the Dirac delta distribution. Combining the formulas (42) for G_1 and (38) for G_2 we find as approximation of $S(0,0)$ for the problem (28)

$$S(0,0) = \int_0^\infty G_1(0,t)G_2(0,t)\,dt \approx 1 - \epsilon, \tag{43}$$

from which we obtain $\tilde{S}(0,0) \approx \frac{1}{\epsilon} - 1 = h - 1$, from which it follows that $\tilde{S}(0,0) \approx \log\gamma + \log 2 - 1$ as $\gamma \to \infty$.

Using the same derivational steps it is possible to generalize to N sensors. In particular, in this case the integrand for $T(x_1,\ldots,x_N)$ in (32) becomes the product (see (34)) of N functions, $G_1(x_1,t),\ldots,G_N(x_N,t)$ each of which satisfies equation (35) with the same boundary conditions with respect to their respective variables. Their respective Laplace transforms satisfy (36). This leads to

$$T(0,\ldots,0) \approx \frac{1}{N}\epsilon\,\mathrm{e}^{1/\epsilon}. \tag{44}$$

Similarly, $S(0,\ldots,0)$ takes the form (43), with integrand consisting of the product of N functions, the Laplace transform of the first of which satisfies (40) and the Laplace transforms of the others satisfy (36). Following the same steps as before, this leads to the asymptotic formula

$$S(0,\ldots,0) \approx 1 - \epsilon. \tag{45}$$

Using (44) and (45), we derive $\tilde{S}(0,0) \approx \log\gamma + \log N - 1$ as $\gamma \to \infty$. This completes the proof of Lemma 1.