SRC TR 87-153

Robust Distributed Block and Sequential Continuous-Time Detection

by

E. Geraniotis

ROBUST DISTRIBUTED BLOCK AND SEQUENTIAL CONTINUOUS-TIME DETECTION

Evaggelos Geraniotis

Department of Electrical Engineering
and Systems Research Center
University of Maryland
College Park, MD 20742

ABSTRACT

Two detectors making independent observations which are the outputs of stochastic dynamical systems driven by colored Gaussian noise must decide which one of two hypotheses is true. Detection with a fixed observation interval (block detection) and sequential detection are considered. The decisions are coupled through a common cost function which for tests with a fixed observation interval consists of the sum of the error probabilities while for sequential tests it comprises the sum of the error probabilities and the expected stopping times. For the case of block detection the time-varying parameters of the dynamical system belong to uncertainty classes determined by 2-alternating capacities or to classes with minimal and maximal elements. For the case of sequential detection the time-invariant parameters of the dynamical system belong to classes with minimal and maximal elements.

A minimax robust (worst-case) design is pursued according to which the two detectors employ tests with a fixed observation interval or sequential probability ratio tests whose likelihood ratios and thresholds depend on the least-favorable parameters over the uncertainty class. For the aforementioned cost function the optimal thresholds of the two detectors turn out to be coupled. It is shown that, despite the uncertainty, the two detectors are thus guaranteed a minimum level of acceptable performance.

This research was supported in part by the Systems Research Center at the University of Maryland, College Park, through National Science Foundation CDR-85-00108 and in part by the Office of Naval Research under contract N00014-86-K-0013.

I. INTRODUCTION

In [1] and [2] distributed discrete-time fixed-sample-size (block) detection and sequential detection problems, respectively, were formulated and solved. Continuous-time distributed detection problems with known statistics are considered in [3]. The two detectors collect independent observations and make decisions which are coupled through a common cost function. Then, the optimal decisions are characterized by thresholds which are coupled. The hypothesis testing models considered in [1], [2], and [3] assume perfect knowledge of the statistics of the observations.

In this paper we investigate distributed detection with a fixed observation interval and sequential detection in continuous time. The observations are the output of a stochastic dynamical system driven by colored Gaussian noise. Due to modeling uncertainties or partially known noise characteristics, the parameters of the stochastic differential equation are only known to belong to uncertainty classes. Similar distributed discrete-time problems with statistical uncertainty are treated in [4], the companion to this paper.

Two distinct types of uncertainty classes are considered: (i) classes determined by two alternating capacities and (ii) classes with minimal or maximal elements. These uncertainty models have been very popular among the robust statisticians [5]-[8] because they include several useful models of uncertainty and result in closed form expressions for the least-favorable elements in the class.

The design philosophy that we pursue is that of minimax robustness. According to it, the worst-case operational conditions are identified with regard to the error probabilities of decision designs and the optimal such design for these conditions is derived. Subsequently this decision design is employed independently of actual conditions (which are not known, except for the fact that they belong to structured uncertainty classes of the types (i) and (ii) above) and it is shown that it achieves desirable performance despite the uncertainty. Minimax robust signal processing techniques have received considerable attention in the past fifteen years (see the tutorial in [9]).

In this paper, we first consider distributed detection schemes with fixed observation interval (block detection) and observations which are the output of two stochastic dynamical systems (driven by mutually independent Wiener processes) whose time-varying parameters (means and/or variances) belong to either 2-alternating capacity classes or to classes with maximal or minimal elements. After the system and uncertainty models are introduced, the least favorable operating conditions, as well as the minimax robust likelihood ratio tests which are coupled through their thresholds, are derived. We also examine the behavior of the Chernoff upper bounds on the joint cost function of the test in the presence of uncertainty and show that the exponential convergence to zero (as the duration of the observation interval increases) is guaranteed for statistical uncertainty within the aforementioned classes. These problems are treated in Section II of the paper.

In Section III sequential tests are employed. In this case we derive the least favorable operational conditions and the minimax robust sequential probability ratio test for two stochastic dynamical systems, whose parameters belong to classes with minimal or maximal elements. The joint performance criterion now includes--besides the error probabilities of the tests--the expected values of the stopping time of each decision maker under the two hypotheses. The sum of the asymptotic speeds of the two sequential tests which are inversly proportional to the informational divergence (Kullback-Leibler distance) is also robustified.

II. ROBUST DISTRIBUTED CONTINUOUS-TIME BLOCK DETECTION

II.A Problem Formulation and Models of Uncertainty

Consider the following hypothesis testing problem of two simple hypotheses H_0 and H_1 with two decision-makers. Decision-maker i $(i = 1, 2)$ is equipped with a sensor and is faced with testing the hypotheses H_1 versus H_0:

$$\begin{aligned} H_1 &: \ dY_{t,i} = M_{1,i}(\ dt\ ) + dW_{t,i} \\ H_0 &: \ dY_{t,i} = M_{0,i}(\ dt\ ) + dW_{t,i} \quad , 0 \le t \le T \end{aligned} \tag{1}$$

In (1) $M_{j,i}$ for $j = 0, 1$ and $i = 1, 2$ are measures--defined on the sample space $\Omega = [0,T]$ and the associated σ-field B--which belong to classes of the form

$$\mathbf{M}_{j,i} = \left\{ M_{j,i} \in M \ \mid \ M_{j,i}\ (\ A\ ) \le v_{j,i}\ (\ A\ ),\ \forall A \in B,\ M_{j,i}\ (\Omega) = v_{j,i}\ (\Omega) \right\}, \tag{2}$$

where $\boldsymbol{M}$ is the class of measures on $(\Omega,\boldsymbol{B})$. The non-standard Wiener process $W_{t,i}$ are mutually independent for $i = 1, 2$, have zero means and variances

$$E\left\{ \left[\int_A dW_{t,i} \right]^2 \right\} = \Sigma_i(A),\ \ \forall A \in \boldsymbol{B}\ . \tag{3}$$

The measures Σ_i belong to uncertainty classes of the form

$$\Sigma_i = \left\{ \ \Sigma_i \in M \ \mid \ \Sigma_i\ (A) \le v_i(A),\ \forall A \in B,\ \Sigma_i\ (\Omega) = v_i(\Omega) \ \right\} \tag{4}$$

The quantities $v_{j,i}$ in (2) and v_i in (4) are 2-alternating capacities and will be defined below.

The decision making of detectors 1 and 2 is coupled through the following cost structure:

$$C(d_1,d_2;h) = \begin{cases} 0 \text{ for } & d_1 = d_2 = h \\ e \text{ for } & d_1 \neq d_2 \\ f \text{ for } & d_1 = d_2 \neq h \end{cases}, \tag{5}$$

where $d_1, d_2, h \in \{0,1\}$, e and f are non-negative constants, and we assume that $f > 2e$. Since the cost $[C(1,1;0) = C(0,0;1)]$ of wrong decisions by both detectors is expected to be considerably larger than the cost $[C(0,1;0) = C(1,0;0) = C(0,1;1) = C(1,0;1)]$ of a wrong decision by one of the detectors, this assumption does not impose a serious restriction on the generality of our problem formulation.

Next we define the 2-alternating capacities:

Definition: A positive set function v on a sample space Ω and assosciated σ-field B is called a **2-alternating capacity** if it is increasing, continuous from below, continuous from above on closed sets, and satisfies the conditions $v(\phi) = 0$, $v(A \cup B) + v(A \cap B) \leq v(A) + v(B)$. Suppose now that M is the class of measures on (Ω, B) and $m \in M$ is any such measure. Consider the uncertainty class which is determined by the 2-alternating capacity v as follows [compare with (2) and (4)]:

$$M_v = \left\{ m \in M \mid m(A) \leq v(A), \ VA \in B, m(\Omega) = v(\Omega) \right\}. \tag{6}$$

When Ω is compact several popular uncertainty models like ϵ-contaminated neighborhoods [5], total variation neighborhoods [5], band classes [6] and p-point classes [7] are special cases of this model.

Example: The ϵ-contaminated model [5]

$$M_\epsilon = \left\{ m \in M \mid m(A) = (1-\epsilon)\, m^0(A) + \epsilon \tilde{m}(A),\ \forall A \in B,\ m^0(\Omega) = \tilde{m}(\Omega) \right\}, \quad (7)$$

for $\epsilon \in [0, 1]$. Then $v(A) = (1-\epsilon)\, m^0(A) + \epsilon\, m^0(\Omega)$

Fundamental properties of these uncertainty models have been studied by Huber and Strassen [8]. We will state the relevant properties as a Lemma.

Lemma 1: Suppose v_0 and v_1 are 2-alternating capacities on (Ω, B) and M_0 and M_1 are the uncertainty classes determined by them as in (1). Then there exists a Lebesgue-measurable function $\pi_v : \Omega \to [0, \infty]$ such that

$$\theta\, v_0(\{\pi_v > \theta\}) + v_1(\{\pi_v \leq \theta\}) \leq \theta\, v_0(A) + v_1(A^c) \quad (8)$$

for all $A \in B$ and all $\theta \geq 0$. Furthermore there exist measures $(\hat{m}_0, \hat{m}_1)$ in $M_0 \times M_1$ such that

$$\hat{m}_0(\{\pi_v > \theta\}) = v_0(\{\pi_v > \theta\}) \quad (9)$$

$$\hat{m}_1(\{\pi_v \leq \theta\}) = v_1(\{\pi_v \leq \theta\}) \quad (10)$$

(that is, π_v is stochastically largest over M_0 under $\hat{m}_0$ and stochastically smallest over M_1 under $\hat{m}_1$) and π_v is a version of $d\hat{m}_1/d\hat{m}_0$ and is unique a.e. $[\hat{m}_0]$. The measures $(\hat{m}_0, \hat{m}_1)$ are termed the **least-favorable** measures over $M_0 \times M_1$.

Example: The ϵ-contaminated mixture uncertainty classes described by

$$M_j = \left\{ m_j \in M \mid m_j = (1-\epsilon_j)\, m_j^0 + \epsilon_j \tilde{m}_j,\ \tilde{m}_j(\Omega) = m^0(\Omega) \right\},\ j = 0, 1 \quad (11)$$

associated with the 2-alternating capacities

$$v_j(A) = \begin{cases} (1-\epsilon_j)\, m_j^0(A) + \epsilon_j, & A \neq \phi \\ 0, & A = \phi \end{cases}, \quad (12)$$

have the least-favorable distributions

$$d\hat{m}_0/d\lambda = \begin{cases} (1-\epsilon_0)\, dm_0^0/d\lambda \ , & dm_1^0/dm_0^0 \le c_2 \\ \dfrac{1-\epsilon_0}{c_2}\, dm_1^0/d\lambda \ , & c_2 < dm_1^0/dm_0^0 \end{cases} \tag{13}$$

$$d\hat{m}_1/d\lambda = \begin{cases} (1-\epsilon_1)\, dm_1^0/d\lambda \ , & c_1 < dm_1^0/dm_0^0 \\ c_1(1-\epsilon_1)\, dm_0^0/d\lambda \ , & dm_1^0/dm_0^0 \le c_1 \end{cases} \tag{14}$$

and the Huber-Strassen derivative π_v

$$\pi_v = d\hat{m}_1/d\hat{m}_0 = \frac{1-\epsilon_1}{1-\epsilon_0} \min \left\{ c_2 \ , \max(c_1 \ , dm_1^0/dm_0^0) \right\} \tag{15}$$

where $0 \le c_1 \le c_2 < \infty$ are such that $\hat{m}_1(\Omega) = \hat{m}_0(\Omega) = 1$.

For the uncertainty classes in (2) and (4) we assume that the nominal measures satisfy $M_{j,i}^0 << \lambda$ and $\Sigma_i^0 << \lambda$ (i.e., they are absolutely continuous with respect to λ, the Lebesgue measure on $\Omega = [0,T]$) so that the least-favorable measures $\hat{M}_{j,i} << \lambda$ and $\hat{\Sigma}_i << \lambda$ and we define $\hat{\mu}_{j,i}(t) = (d\hat{M}_{j,i}/d\lambda)(t)$ and $[\hat{\sigma}_i(t)]^2 = (d\hat{\Sigma}_i/d\lambda)(t)$.

We also consider the hypothesis testing problem

$$\begin{aligned} H_1 &: \quad dY_{t,i} = \mu_{1,i}(t)\, dt + \sigma_i(t)\, dW_{t,i} \\ H_0 &; \quad dY_{t,i} = \mu_{0,i}(t)\, dt + \sigma_i(t)\, dW_{t,i} \quad , 0 \le t \le T \end{aligned} \tag{16}$$

where $W_{t,i}$ are mutually idependent standard Wiener processes $(i = 1, 2)$, $\mu_{j,i}(t)$ $(j = 0, 1 \ , i = 0, 1)$ belong to the following classes with **minimal or maximal elements**.

$$M_{1,i} = \left\{ \mu_{1,i}(t) \mid \mu_{1,i}(t) \ge \hat{\mu}_{1,i}(t), 0 \le t \le T \right\} \tag{17}$$

$$M_{0,i} = \left\{ \mu_{0,i}(t) \mid \mu_{0,i}(t) \le \hat{\mu}_{0,i}(t), 0 \le t \le T \right\}, \tag{18}$$

and $\sigma_i(t)$ ($i=1,2$) belong to the classes

$$\Sigma_i = \left\{ \sigma_i(t) \mid \sigma_i(t) \le \hat{\sigma}_i(t) \ , \ 0 \le t \le T \right\} \tag{19}$$

For either the hypothesis testing problem (1) or that of (16) we assume that the a priori probabilities for the hypotheses H_0 and H_1 are λ and $1-\lambda$, respectively, and that likelihood ratio tests are employed, the average cost is

$$\begin{aligned} J(L_1,L_2,\tilde{\eta}_1,\tilde{\eta}_2) = \lambda\{ & e\,[m_{0,1}(\{L_1>\tilde{\eta}_1\}) + m_{0,2}(\{L_2>\tilde{\eta}_2\})] \\ & + (f-2e)m_{0,1}(\{L_1>\tilde{\eta}_1\})\cdot m_{0,2}(\{L_2>\tilde{\eta}_2\})\} \\ & + (1-\lambda)\{e\,[m_{1,1}(\{L_1\le\tilde{\eta}_1\}) + m_{1,2}(\{L_2\le\tilde{\eta}_2\})] \\ & + (f-2e)m_{1,1}(\{L_1^{(n)}\le\tilde{\eta}_1\})\cdot m_{1,2}(\{L_2\le\tilde{\eta}_2\})\} \end{aligned} \tag{20}$$

In (20) $m_{j,i}$ are the probability measures induced by the Wiener processes of (1) [or (16)] of the i-th decision-maker ($i=1,2$) under hypothesis H_j ($j=0,1$). By $L_i = dm_{1,i}/dm_{0,i}$ we denote the likelihood ratio based on observations of the i-th decision-maker over the interval $[0,T]$ and by $\tilde{\eta}_i$ its threshold.

The optimal thresholds for (20) are the pair (η_1,η_2) which minimizes the average cost function $J(L_1,L_2,\tilde{\eta}_1,\tilde{\eta}_2)$, that is

$$(\eta_1,\eta_2) = arg \min_{\tilde{\eta}_1,\tilde{\eta}_2} J(L_1,L_2,\tilde{\eta}_1,\tilde{\eta}_2) \tag{21}$$

Actually the likelihood ratio tests (LRTs) are the optimal policies for the two-decision-maker problem formulated above as stated in the following proposition

Proposition 1: Likelihood ratio tests (LRTs) with thresholds which minimize $J(L_1,L_2,\tilde{\eta}_1,\tilde{\eta}_2)$ of (20) are optimal over all tests for the aforementioned common cost structure

Proof: The proof follows closely the corresponding proof of [1] about the optimality of the one-detector strategy (i.e., the likelihood ratio test) in this case of decision makers with independent observations, and will be ommitted.

II.B Robust Distributed Block Detection

The expression for the average cost function in (20) is valid for the case that there is no uncertainty in the statistics of the observations of the two decision makers. In the presence of uncertainty within the 2-alternating classes $M_{j,i}$ and Σ_i of (2) and (4) [or the classes with maximal or minimal elements of (17)-(19)], the likelihood ratios $\hat{L}_i$ and the thresholds $\hat{\eta}_i$, $i = 1, 2$, which are matched to the least-favorable mean and variance measures $\hat{M}_{j,i}$ and $\hat{\Sigma}_i$ (singled out by Lemma 1) of the classes $M_{j,i}$ of (2) and Σ_i of (4), respectively, are employed. Similarly in the case of the uncertainty classes of (17)-(19). In these cases the **average cost function under mismatch**--that is, when the statistics of the observations are actually governed by the probability measures $m_{j,i}$ induced by the Wiener processes with the means and variances above--is given by $J(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2)$ which is obtained from (20), if we replace L_i by $\hat{L}_i$ and η_i by $\hat{\eta}_i$, for $i = 1, 2$, and these thresholds are the solution to the minimization problem:

$$(\hat{\eta}_1,\hat{\eta}_2) = arg \min_{\tilde{\eta}_1,\tilde{\eta}_2} \hat{J}(\hat{L}_1,\hat{L}_2,\tilde{\eta}_1,\tilde{\eta}_2) , \tag{22}$$

where $\hat{J}(\hat{L}_1,\hat{L}_2,\tilde{\eta}_1,\tilde{\eta}_2)$ is the average cost when the likelihood ratios $\hat{L}_i$ ($i = 1, 2$ for the two detectors) are employed and the observations are distributed according to the probability measures $\hat{m}_{j,i}$ induced by the Wiener process with the aforementioned least-favorable means and variances.

From the results presented in [10] we cite the following Lemma for a single detector (say the i-th detector $i = 1$, or 2)

Lemma 2: For either the hypothesis testing problem (1) [for the i-th detector] and the uncertainty models (2) and (4), or the problem of (16) and the uncertainty models of (17)-(19), the likelihood ratio test based on

$$ln\hat{L}_i = \int_\Omega \frac{\hat{\mu}_{1,i}(t) - \hat{\mu}_{0,i}(t)}{[\hat{\sigma}_i(t)]^2}\, dY_{t,i} - \frac{1}{2}\int_\Omega \frac{[\hat{\mu}_{1,i}(t)]^2 - [\hat{\mu}_{0,i}(t)]^2}{[\hat{\sigma}_i(t)]^2}\, dt\,, \tag{23}$$

is a minimax robust test; i.e., the error probabilities for this test under the two hypotheses satisfy the inequalities

$$m_{0,i}(\{\hat{L}_i > \eta_i\}) < \hat{m}_{0,i}(\{\hat{L}_i > \eta_i\}) \leq \hat{m}_{0,i}(\{G_i \geq \eta_i\}) \tag{24}$$

$$m_{1,i}(\{\hat{L}_i \leq \eta_i\}) \leq \hat{m}_{1,i}(\{\hat{L}_i \leq \eta_i\}) \leq \hat{m}_{1,i}(\{G_i < \eta_i\}) \tag{25}$$

for all thresholds η_i satisfying

$$|\, ln\, \eta_i \,| \leq \frac{1}{2}\int_\Omega \frac{[\hat{\mu}_{1,i}(t) - \hat{\mu}_{0,i}(t)]^2}{[\hat{\sigma}_i(t)]^2} dt.$$

In (24)-(25) $\hat{m}_{j,i}$ are the measures induced by the Wiener processes of (1) or (16) when the parameters involved in these equations are the least-favorable elements of the corresponding uncertainty classes.

Proof: See Propositions 3 and 6 of [10].

The main result of this section now follows:

Proposition 2: The LRTs of the two detectors based on the log-likelihood-ratio of (23) for $i = 1, 2$ are minimax robust with respect to the average cost function defined in (20), for the problems (1) and (16) with parameters in the uncertainty classes (2) and (4) or (17)-(19), respectively, that is

$$J(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2) \leq \hat{J}(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2) \leq \hat{J}(G_1,G_2,\eta_1,\eta_2) \tag{26}$$

where G_i $(i = 1, 2)$ are any decision statistics operating on the outputs of the dynamical systems of (1) or (16).

Proof: The right-hand-side inequality in (26) is a straightforward application of Proposition 1 to the case characterized by $\hat{m}_{j,i}$ $(j = 0,1$ and $i = 1, 2)$ for which the likelihood ratios are $\hat{L}_i$ and the optimal thresholds are $\hat{\eta}_i$.

The left-hand-side inequality in (26) is a consequence of Lemma 2. Specifically we apply the left-hand-side inequalities in (24) and (25) to the probability measures $(m_{0,i}, \hat{m}_{0,i})$ and $(m_{1,i}, \hat{m}_{1,i})$, respectively, of the two detectors $(i = 1, 2)$, and then use the definitons of the mismatch average cost function J and the average cost function $\hat{J}$ matched to the pair of probability measures $(\hat{m}_{0,i}, \hat{m}_{1,i})$ induced by the least-favorable parameters of the classes (2) and (4) or (17)-(19).

Note: The optimal thresholds $(\hat{\eta}_1,\hat{\eta}_2)$ can be determined from the error probabilities $\hat{\alpha}_i,\hat{\beta}_i$ $(i=1,2)$ for the least-favorable case of problem (1) by minimizing

$$\min\left\{\lambda\left[e(\hat{\alpha}_1+\hat{\alpha}_2)+(f-2e)\hat{\alpha}_1\hat{\alpha}_2\right]+(1-\lambda)\left[e(\hat{\beta}_1+\hat{\beta}_2)+(f-2e)\hat{\beta}_1\hat{\beta}_2\right]\right\}$$

under the constraints $\hat{\beta}_i = f_i(\hat{\alpha}_i)$ [operating receiver characteristic (ROC) for detector i], $0 \leq \hat{\alpha}_i \leq 1$, $0 \leq \hat{\beta}_i \leq 1$, and $\hat{\alpha}_i + \hat{\beta}_i \leq 1$ for $i=1,2$.

II.C Asymptotic Performance

It is of interest to examine the behaviour of the joint cost function under the two hypotheses as T, the length of the observation interval, increases. In particular, it is desirable that the robust likelihood ratio tests described in Propositions 2 behave asymptotically (for large T) in an optimal way, For continuous-time problems with uncertain statistics the asymptotic performance of the Chernoff bounds on the error probabilities of a single detector (say the i-th detector) with observations obeying (1) or (16), and uncertainty models (2) and (4) or (17)-(19)-when the likelihood ratio test of (23) is employed-has been examined in [10]. The next two Lemmas contain the results of [10] in a condensed form.

Lemma 3: The error probabilities for the hypothesis testing problems of (1) and (16) [for the i-th detector] can be upperbounded as

$$m_{0,i}(\{\hat{L}_i > \gamma_i T\}) \leq \exp\{-T[s\gamma_i + C_{0,i}(s,\hat{L}_i)]\} \tag{27}$$

and

$$m_{1,i}(\{\hat{L}_i \leq \gamma_i T\}) \leq \exp\{-T[-s\gamma_i + C_{1,i}(s,\hat{L}_i)]\} \tag{28}$$

for all s in (0,1).

In (27)-(28) the **Chernoff distances** $C_{j,i}(s,\hat{L}_i)$ for ($j = 0, 1$) and $i = 1, 2$)under mismatch are defined as

$$C_{0,i}(s,\hat{L}_i) = -\frac{1}{T}\ln[E_{0,i}\{\hat{L}_i^s\}] \tag{29}$$

and

$$C_{1,i}(s,\hat{L}_i) = -\frac{1}{T}\ln[E_{1,i}\{\hat{L}_i^{-s}\}] \tag{30}$$

and the threshold is $\eta_i = \gamma_i T$. For uncertainties in (2) and (4) the Chernoff distances actually take the form

$$C_{0,i}(s,\hat{L}_i) = \frac{1}{T}\left[-\frac{s^2}{2}\int_\Omega\left\{\frac{\hat{\mu}_{1,i}(t)-\hat{\mu}_{0,i}(t)}{[\hat{\sigma}_i(t)]^2}\right\}^2\Sigma_i(dt) - s\int_\Omega\frac{\hat{\mu}_{1,i}(t)-\hat{\mu}_{0,i}(t)}{[\hat{\sigma}_i(t)]^2}M_{0,i}(dt) + \frac{s}{2}c_i\right] \tag{31}$$

$$C_{1,i}(s,\hat{L}_i) = \frac{1}{T}\left[-\frac{s^2}{2}\int_\Omega \left\{ \frac{\hat{\mu}_{1,i}(t)-\hat{\mu}_{0,i}(t)}{[\hat{\sigma}_i(t)]^2} \right\}^2 \Sigma_i(dt) + s\int_\Omega \frac{\hat{\mu}_{1,i}(t)-\hat{\mu}_{0,i}(t)}{[\hat{\sigma}_i(t)]^2} M_{1,i}(dt) - \frac{s}{2}c_i \right] \tag{32}$$

where

$$c_i = \int_0^T \frac{[\hat{\mu}_{1,i}(t)]^2-[\hat{\mu}_{0,i}(t)]^2}{[\hat{\sigma}_i(t)]^2} dt. \tag{33}$$

For uncertainties in (17)-(19) we need to replace $M_{j,i}(dt)$ by $\mu_{j,i}(t)dt$ and $\Sigma_i(dt)$ by $[\sigma_i(t)]^2 dt$, respectively, for $j = 0, 1$ and $i = 1, 2$.

Proof: See Proposition 4 of [10].

Lemma 4: The Chernoff upper bounds on the error probabilities of the hypothesis-testing problems (1) and (16) with the uncertainty models of (2) and (4) and (17)-(19) aproach zero exponentially with increasing T--despite the uncertainty--when the likelihood ratio test of (23) is employed.

Proof: See Proposition 5 of [10].

The following proposition provides the desired asymptotic result for the mismatch average cost function $J(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2)$ as the length of the observations interval T increases:

Proposition 3: Under the assumptions of Proposition 2, the average cost function under mismatch converges to zero exponentially as the length of observations interval T increases, despite the uncertainty; that is, $J(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2)\rightarrow 0$, as $T \rightarrow \infty$ for all values of the parameters in the uncertainty class (2) and (4) or (17)-(19).

Proof: By applying Lemma 3 to the error probabilities of the hypothesis testing problem of each of the two detectors and using the definition of $J(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2)$ we derive an upper bound on the average cost under mismatch in terms of the Chernoff bounds. This takes the form

$$J(\hat{L}_1,\hat{L}_2,\hat{\eta}_1,\hat{\eta}_2) \leq \lambda\{e\,[\exp\{-T\,[s\,\hat{\gamma}_1+C_{0,1}(s\,,\hat{L}_1)]\} + \exp\{-T\,[s\,\hat{\gamma}_2+C_{0,2}(s\,,\hat{L}_2)]\}]$$
$$+ (f\,-2e\,)\exp\{-T\,[s\,\hat{\gamma}_1+C_{0,1}(s\,,\hat{L}_1)]\}\exp\{-T\,[s\,\hat{\gamma}_2+C_{0,2}(s\,,\hat{L}_2)]\}\}$$
$$+ (1-\lambda)\{e\,[\exp\{-T\,[-s\,\hat{\gamma}_1+C_{1,1}(s\,,\hat{L}_1)]\} + \exp\{-T\,[-s\,\hat{\gamma}_2+C_{1,2}(s\,,\hat{L}_2)]\}]$$
$$+ (f\,-2e\,)\exp\{-T\,[-s\,\hat{\gamma}_1+C_{1,1}(s\,,\hat{L}_1)]\}\exp\{-T\,[-s\,\hat{\gamma}_2+C_{1,2}(s\,,\hat{L}_2)]\}\} \tag{34}$$

where $\hat{\eta}_i = \hat{\gamma}_i\,T$ is the threshold for the i-th detector ($i = 1, 2$). Finally we apply Lemma 4 to (34) to complete the proof of Proposition 3.

III. ROBUST DISTRIBUTED CONTINUOUS-TIME SEQUENTIAL DETECTION

III.A Problem Formulation and Model of Uncertainty

The distributed sequential detection problem that we consider in this section has a lot of sismilarities with the problem considered in the previous section. The two decision makers are faced with the hypothesis testing problem

$$\begin{aligned} H_0 &: \quad dY_{t,i} = \mu_{0,i}\ dt + \sigma_i\ dW_{t,i} \ , \\ H_1 &: \quad dY_{t,i} = \mu_{1,i}\ dt + \sigma_i\ dW_{t,i} \ , \quad 0 \le t \le \tau_i \end{aligned} \tag{35}$$

where τ_i is the stopping time [11] of the i-th detector and $\mu_{1,i}$, $\mu_{0,i}$, σ_i are time-invariant parameters which belong to the following uncertainty classes characterized by their minimal or maximal elements (for which $\hat{\mu}_{1,i} \ge \hat{\mu}_{0,i}$).

$$M_{1,i} = \left\{ \mu_{1,i} \mid \mu_{1,i} \ge \hat{\mu}_{1,i} \right\}, \tag{36}$$

$$M_{0,i} = \left\{ \mu_{0,i} \mid \mu_{0,i} \le \hat{\mu}_{0,i} \right\}, \tag{37}$$

$$\Sigma_i = \left\{ \sigma_i \mid \sigma_i \le \hat{\sigma}_i \right\}. \tag{38}$$

The cost function $C(\cdot,\cdot;\cdot)$ of (5) remains the same as in section II. However, now there is also a cost for collecting data, which for the i-th decision maker ($i = 1, 2$) is defined by:

$$k_i\left[\lambda E_{0,i}\{\tau_i\} + (1-\lambda)E_{1,i}\{\tau_i\}\right], \tag{39}$$

where k_i ($i = 1, 2$) are nonnegative constants, $E_{j,i}$ denotes expectation with respect to the probability measure $m_{j,i}$ (under the hypothesis H_j, $j = 0, 1$, and for the i-th detector, $i = 1, 2$) induced by the Wiener process $W_{t,i}$ of (35), the a priori probabilities for the hypotheses H_0 and H_1 are λ and $1-\lambda$, respectively, and the random variable τ_i is the stopping time of the i-th detector; i.e., the necessary length of the observation interval in order to reach a decision in favor of one of the two hypotheses.

Recall [11]-[12] that for a single sequential detector, the optimal test, termed the sequential probability ratio test (SPRT), consists of keep collecting observations till the likelihood ratio L_τ based on the observation interval $[0,\tau]$ exceeds B or falls below A -- the two thresholds ($0<A<1<B$)--in which case a decision is made in favor of H_1 or H_0, respectively.

Assuming that SPRTs are employed by both detectors, we can write the average cost as

$$J(L_1,L_2,\tilde{A}_1,\tilde{B}_1,\tilde{A}_2,\tilde{B}_2) = \lambda\Big\{ k_1 E_{0,1}\{\tau_1 \mid L_1\} + k_2\, E_{0,2}\{\tau_2 \mid L_2\}$$

$$+e\,[m_{0,1}^{(*)}(\{L_{1,\tau_1}\geq\tilde{B}_1\}) + m_{0,2}^{(*)}(\{L_{2,\tau_2}\geq\tilde{B}_2\})]$$

$$+(f-2e)\,m_{0,1}^{(*)}(\{L_{1,\tau_1}\geq\tilde{B}_1\})\cdot m_{0,2}^{(*)}(\{L_{2,\tau_2}\geq\tilde{B}_2\})\Big\}$$

$$+(1-\lambda)\Big\{ k_1 E_{1,1}\{\tau_1 \mid L_1\} + k_2\, E_{1,2}\{\tau_2 \mid L_2\}$$

$$+e\,[m_{1,1}^{(*)}(\{L_{1,\tau_1}\leq\tilde{A}_1\}) + m_{1,2}^{(*)}(\{L_{2,\tau_2}\leq\tilde{A}_2\})]$$

$$+(f-2e)\,m_{1,1}^{(*)}(\{L_{1,\tau_1}\leq\tilde{A}_1\})\cdot m_{1,2}^{(*)}(\{L_{2,\tau_2}\leq\tilde{A}_2\})\Big\} \tag{40}$$

where τ_i $(i = 1, 2)$ are stopping times for the two detectors, that is, if the likelihood ratio $L_{i,\tau}$, which is based on observations in the interval $[0,\tau]$, is larger than or equal to $\tilde{B}_i$, it is decided that H_1 is true, the test terminates and $\tau_i = \tau$; if it is smaller than or equal to $\tilde{A}_i$, it is decided that H_0 is true, the test again terminates and $\tau_i = \tau$; otherwise, more observations are collected ($[0,\tau]$ increases) and the procedure continues. $m_{j,i}^{(*)}$ is the probability measure which governs the observations of the i-th detector under hypothesis H_j ($j = 0, 1$) when the SPRT terminates after the interval $[0,\tau_i]$ has been processed. The notation $E_{j,i}\{\tau_i \mid L_i\}$ has been preferred over the notation $E_{j,i}\{\tau_i\}$ for the expected value of τ_i under the probability measure $m_{j,i}$ and an SPRT employing the likelihood ratio $L_i = dm_{1,i}/dm_{0,i}$, because it allows us to consider situations of mismatch, that is, when the likelihood ratio employed is not the one corresponding to the operating probability measures.

The optimal thresholds for (40) are the quadruple (A_1,B_1,A_2,B_2) which minimizes the average cost function $J(L_1,L_2,\tilde{A}_1,\tilde{B}_1,\tilde{A}_2,\tilde{B}_2)$, that is

$$(A_1,B_1,A_2,B_2) = arg \min_{\tilde{A}_1,\tilde{B}_1,\tilde{A}_2,\tilde{B}_2} J(L_1,L_2,\tilde{A}_1,\tilde{B}_1,\tilde{A}_2,\tilde{B}_2) \tag{41}$$

Actually the sequential probability ratio tests (SPRTs) are the optimal policies for the two-decision-maker problem formulated above as stated in the following proposition

Proposition 4: SPRTs with thresholds which minimize $J(L_1,L_2,\tilde{A}_1,\tilde{B}_1,\tilde{A}_2,\tilde{B}_2)$ of (40) are optimal over all tests for the aforementioned common cost structure.

Proof: The proof is provided in [2] for discrete-time sequential detection and in [3] for continuous-time sequential detection and it establishes the optimality of the one-detector strategy (i.e., the SPRT) in this case of decision makers with independent observations. It will be ommitted.

III.B Robust Distributed Sequential Detection

The expression for the average cost function in (40) is valid for the case that there is no uncertainty in the statistics of the observations of the two decision makers. In the presence of uncertainty within the classes of (36)-(38), the likelihood ratios $\hat{L}_i$ and the thresholds $(\hat{A}_i, \hat{B}_i)$, $i = 1, 2$, which are matched to the measures $\hat{m}_{j,i}$ induced by the least-favorable parameters in the above classes are employed. In this case the **average cost function under mismatch**--that is, when the statistics of the observations are actually governed by $m_{j,i}$ induced by any parameters in (36)-(38)--is given by $J(\hat{L}_1, \hat{L}_2, \hat{A}_1, \hat{B}_1, \hat{A}_2, \hat{B}_2)$ which is obtained from (40), if we replace L_i by $\hat{L}_i$ and (A_i, B_i) by $(\hat{A}_i, \hat{B}_i)$, for $i = 1, 2$. These thresholds are the solution to the minimization problem:

$$(\hat{A}_1, \hat{B}_1, \hat{A}_2, \hat{B}_2) = arg \min_{\tilde{A}_1, \tilde{B}_1, \tilde{A}_2, \tilde{B}_2} \hat{J}(\hat{L}_1, \hat{L}_2, \tilde{A}_1, \tilde{B}_1, \tilde{A}_2, \tilde{B}_2) \ , \tag{42}$$

where $\hat{J}(\hat{L}_1, \hat{L}_2, \tilde{A}_1, \tilde{B}_1, \tilde{A}_2, \tilde{B}_2)$ is the average cost when SPRTs based on the likelihood ratios $\hat{L}_i$ and the thresholds $(\tilde{A}_i, \tilde{B}_i)$ ($i = 1, 2$ for the two detectors) are employed and the observations are governed by the probability measure $\hat{m}_{j,i}$ which is induced by the least-favorable parameters ($j = 0, 1$ for the two hypotheses).

From [10] we cite the following two results for a single sequential detector:

Lemma 5: Suppose τ_i is the stopping time of the SPRT associated with (35) and the SPRT employs a likelihood ratio

$$ln\hat{L}_{i,\tau} = \frac{\hat{\mu}_{1,i} - \hat{\mu}_{0,i}}{\hat{\sigma}_i^2} Y_{i,\tau} - \frac{\hat{\mu}_{1,i}^2 - \hat{\mu}_{0,i}^2}{2\hat{\sigma}_i^2} \tau \tag{43}$$

based on observations over $[0,\tau]$ and thresholds $(\hat{A}, \hat{B})$ which are matched to the case $(\hat{\mu}_{0,i}, \hat{\mu}_{1,i}, \hat{\sigma}_i)$ [the least-favorable parameters in the uncertainty classes (36)-(38)] while the actual operating conditions are determined [through (35)] by $(\mu_{0,i}, \mu_{1,i}, \sigma_i)$. Then the following identities are true for the error probabilities under mismatch:

$$\alpha_i = m_{0,i}(\{\hat{L}_{i,\tau_i} \geq \hat{B}_i\}) = \frac{\hat{A}_i^{r_{0,i}} - 1}{(\hat{B}_i^{-1}\hat{A}_i)^{r_{0,i}} - 1} \tag{44}$$

where

$$r_{0,i} = \frac{\hat{\sigma}_i^2 \, (2\,\mu_{0,i} - \hat{\mu}_{1,i} - \hat{\mu}_{0,i})}{\sigma_i^2 \, (\hat{\mu}_{1,i} - \hat{\mu}_{0,i})} \tag{45}$$

and

$$\beta_i = m_{1,i}(\{\hat{L}_{i,\tau_i} \leq \hat{A}_i\}), = \frac{\hat{B}_i^{r_{1,i}} - 1}{(\hat{A}_i^{-1}\hat{B}_i)^{r_{1,i}} - 1} \tag{46}$$

where

$$r_{1,i} = \frac{\hat{\sigma}_i^2 \, (2\,\mu_{1,i} - \hat{\mu}_{1,i} - \hat{\mu}_{0,i})}{\sigma_i^2 \, (\hat{\mu}_{1,i} - \hat{\mu}_{0,i})} \,. \tag{47}$$

For the expected stopping times under mismatch the following identities hold:

$$E_{0,i} \, [\, \tau_i \mid \hat{L}_i \,] = \frac{-2\,\hat{\sigma}_i^2 \, \omega \, (\hat{\alpha}_i , \hat{\beta}_i , \alpha_i)}{(\hat{\mu}_{1,i} - \hat{\mu}_{0,i}) \, (2\,\mu_{0,i} - \hat{\mu}_{1,i} - \hat{\mu}_{0,i})} \tag{48}$$

and

$$E_{1,i} \, [\, \tau_i \mid \hat{L}_i \,] = \frac{2\,\hat{\sigma}_i^2 \, \omega \, (\hat{\beta}_i , \hat{\alpha}_i , \beta_i)}{(\hat{\mu}_{1,i} - \hat{\mu}_{0,i}) \, (2\,\mu_{1,i} - \hat{\mu}_{1,i} - \hat{\mu}_{0,i})} \tag{49}$$

where

$$\omega \, (\hat{x} , \hat{y} , x) = (1 - x) \ln \frac{1 - \hat{x}}{\hat{y}} + x \ln \frac{\hat{x}}{1 - \hat{y}}, \tag{50}$$

and

$$\hat{\alpha}_i = \hat{m}_{0,i}(\{\hat{L}_{i,\tau_i} \geq \hat{B}_i\}) = \frac{1 - \hat{A}_i}{\hat{B}_i - \hat{A}_i}, \tag{51}$$

and

$$\hat{\beta}_i = \hat{m}_{1,i}(\{\hat{L}_{i,\tau_i} \leq \hat{A}_i\}) = \frac{\hat{A}_i(\hat{B}_i - 1)}{\hat{B}_i - \hat{A}_i}, \tag{52}$$

are the error probabilities in the matched case. The thresholds $\hat{A}_i$ and $\hat{B}_i$ are given by

[11]

$$\hat{A}_i = \frac{\hat{\beta}_i}{1-\hat{\alpha}_i} \tag{53}$$

$$\hat{B}_i = \frac{1-\hat{\beta}_i}{\hat{\alpha}_i} \tag{54}$$

Proof: See Lemmas 3 and 2 of [10].

lemma 6: Consider the hypothesis testing problem of (35) [for the i-th detector] with parameters in the sets (36)-(38). Suppose that the following assumptions are satisfied:

$$(i)\ \ \hat{\mu}_{1,i} \geq \hat{\mu}_{0,i} \tag{55}$$

$$(ii)\ \ \ln \frac{1-\hat{\beta}_i}{\hat{\alpha}_i} >> -\hat{\beta}_i \ln \frac{\hat{\alpha}_i\ \hat{\beta}_i}{(1-\hat{\alpha}_i)(1-\hat{\beta}_i)} . \tag{56}$$

Then the following inequalities are true

$$\alpha_i = m_{0,i}(\{\hat{L}_{i,\tau_i} \geq \hat{B}_i\}) \leq \hat{m}_{0,i}(\{\hat{L}_{i,\tau_i} \geq \hat{B}_i\}) = \hat{\alpha}_i\ , \tag{57}$$

$$\beta_i = m_{1,i}(\{\hat{L}_{i,\tau_i} \leq \hat{A}_i\}), \leq \hat{m}_{1,i}(\{\hat{L}_{i,\tau_i} \leq \hat{A}_i\}) = \hat{\beta}_i\ , \tag{58}$$

and

$$E_{j,i}\ [\ \tau_i\ \mid \hat{L}_i\ ] \leq \hat{E}_{j,i}\ [\ \tau \mid \hat{L}_i\ ]\ ,\quad j = 0, 1\ . \tag{59}$$

Proof: See Proposition 8 of [10].

The main result of this section now follows:

Proposition 5: The SPRTs which employ thresholds $(\hat{A}_i, \hat{B}_i)$ and likelihood ratios $\hat{L}_{i,\tau}$ defined as in Lemma 5 [equation (43)] (for the two detectors $i = 1, 2$) are minimax robust with respect to the average cost function defined in (40), that is

$$J(\hat{L}_1,\hat{L}_2,\hat{A}_1,\hat{B}_1,\hat{A}_2,\hat{B}_2) \leq \hat{J}(\hat{L}_1,\hat{L}_2,\hat{A}_1,\hat{B}_1,\hat{A}_2,\hat{B}_2) \leq \hat{J}(G_1,G_2,A_1,B_1,A_2,B_2) \tag{60}$$

where G_i $(i = 1, 2)$ is any decision statistic operating on the observations over the interval $[0,\tau_i]$, if for $i = 1, 2$ $\hat{\alpha}_i$ and $\hat{\beta}_i$ of (53)-(54) satisfy the condition (56).

Note: The optimal thresholds $(\hat{A}_i, \hat{B}_i)$ $(i = 1, 2)$ can be determined from equations (53)-(54) where the error probabilities $\hat{\alpha}_i$ and $\hat{\beta}_i$ are solutions to the minimization problem:

$$\min \left\{ \lambda \left[k_1 \frac{\omega(\hat{\alpha}_1,\hat{\beta}_1,\hat{\alpha}_1)}{\hat{E}_{0,1}\{-ln\hat{L}_{1,\tau_1}\}} + k_2 \frac{\omega(\hat{\alpha}_2,\hat{\beta}_2,\hat{\alpha}_2)}{\hat{E}_{0,2}\{-ln\hat{L}_{2,\tau_2}\}} + e(\hat{\alpha}_1+\hat{\alpha}_2) + (f-2e)\hat{\alpha}_1\hat{\alpha}_2 \right] \right.$$

$$\left. + (1-\lambda) \left[k_1 \frac{\omega(\hat{\beta}_1,\hat{\alpha}_1,\hat{\beta}_1)}{\hat{E}_{1,1}\{\ln\hat{L}_{1,\tau_1}\}} + k_2 \frac{\omega(\hat{\beta}_2,\hat{\alpha}_2,\hat{\beta}_2)}{\hat{E}_{1,2}\{\ln\hat{L}_{2,\tau_2}\}} + e(\hat{\beta}_1+\hat{\beta}_2) + (f-2e)\hat{\beta}_1\hat{\beta}_2 \right] \right\} \quad (61)$$

under the constraints $0 \leq \hat{\alpha}_i \leq 1$, $0 \leq \hat{\beta}_i \leq 1$, and $\hat{\alpha}_i + \hat{\beta}_i \leq 1$ for $i = 1, 2$. In (61) the quantities in the denominators satisfy the inequalities:

$$E_{0,i}\{-ln\hat{L}_{i,\tau_i}\} = \frac{-(\hat{\mu}_{1,i} - \hat{\mu}_{0,i})(2\mu_{0,i} - \hat{\mu}_{1,i} - \hat{\mu}_{0,i})}{2\hat{\sigma}_i^2} \geq \frac{(\hat{\mu}_{1,i} - \hat{\mu}_{0,i})^2}{2\hat{\sigma}_i^2} = \hat{E}_{0,i}\{-ln\hat{L}_{i,\tau_i}\} \quad (62)$$

$$E_{1,i}\{ln\hat{L}_{i,\tau_i}\} = \frac{(\hat{\mu}_{1,i} - \hat{\mu}_{0,i})(2\mu_{1,i} - \hat{\mu}_{1,i} - \hat{\mu}_{0,i})}{2\hat{\sigma}_i^2} \geq \frac{(\hat{\mu}_{1,i} - \hat{\mu}_{0,i})^2}{2\hat{\sigma}_i^2} = \hat{E}_{1,i}\{ln\hat{L}_{i,\tau_i}\} \quad (63)$$

which are valid because of the definition of the log-likelihood ratios $ln\hat{L}_{i,\tau_i}$ and of the inequalities that define the uncertainty classes (36)-(38).

III.C Asymptotic Performance

The following proposition provides a result on the common asymptotic speed--which is defined as the sum of the asymptotic (for small error probabilities) stopping times of the two detectors--of the robust sequential test.

Proposition 6: Suppose that for the problem (35) with the uncertainty classes (36)-(38) and under the mismatch conditions of Lemmas 5, 6 and Proposition 5 above, the error probabilities $\hat{\alpha}_i$ and $\hat{\beta}_i$ (for $i=1,2$) approach zero. Then the sum of the asymptotic expected stoppong times--under mismatch and for the least-favorable case--satisfy the inequalities

$$\lambda\left[k_1\frac{-ln\hat{A}_1}{E_{0,1}\{-ln\hat{L}_1\}}+k_2\frac{-ln\hat{A}_2}{E_{0,2}\{-ln\hat{L}_2\}}\right]+(1-\lambda)\left[k_1\frac{ln\hat{B}_1}{E_{1,1}\{ln\hat{L}_1\}}+k_2\frac{ln\hat{B}_2}{E_{1,2}\{ln\hat{L}_2\}}\right]$$

$$\leq\lambda\left[k_1\frac{-ln\hat{A}_1}{\hat{E}_{0,1}\{-ln\hat{L}_1\}}+k_2\frac{-ln\hat{A}_2}{\hat{E}_{0,2}\{-ln\hat{L}_2\}}\right]+(1-\lambda)\left[k_1\frac{ln\hat{B}_1}{\hat{E}_{1,1}\{ln\hat{L}_1\}}+k_2\frac{ln\hat{B}_2}{\hat{E}_{1,2}\{ln\hat{L}_2\}}\right]$$

$$\leq\lambda\left[k_1\underline{\hat{E}}_{0,1}\{\tau_1\mid G_1\}+k_2\underline{\hat{E}}_{0,2}\{\tau_2\mid G_2\}\right]+(1-\lambda)\left[k_1\underline{\hat{E}}_{1,1}\{\tau_1\mid G_1\}+k_2\underline{\hat{E}}_{1,2}\{\tau_2\mid G_2\}\right] \tag{64}$$

where G_1 and G_2 are any other sequential tests different from the SPRT; $\underline{\hat{E}}_{j,i}$ denotes the limit of the expectation $\hat{E}_{j,i}$ as $\hat{\alpha}_i\to 0$ and $\hat{\beta}_i\to 0$.

Proof: As $\hat{\alpha}_i\to 0$ and $\hat{\beta}_i\to 0$, then α_i and β_i approach zero as well, since $\alpha_i\leq\hat{\alpha}_i$ and $\beta_i\leq\hat{\beta}_i$. Thus $J(\hat{L}_1,\hat{L}_2,\hat{A}_1,\hat{B}_1,\hat{A}_2,\hat{B}_2)$ (the joint cost function under mismatch) reduces to the first sum in (64), whereas $\hat{J}(\hat{L}_1,\hat{L}_2,\hat{A}_1,\hat{B}_1,\hat{A}_2,\hat{B}_2)$ (the joint cost function matched to the least-favorable case) reduces to the second sum in (64). The first sum is smaller than the second sum because of the inequalities (62)-(63). The second inequality in (64) holds because of a theorem by Wald [12] (for the matched single detector case) which states that the SPRT has the minimum asymptotic speed (expected stopping time) among all sequential tests.

REFERENCES

[1] R. R. Tenney, and N. R. Sandell, "Detection with distributed sensors," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-17, pp. 501-510, July 1981.

[2] D. Teneketzis and Y.-C. Ho, "The decentralized Wald problem," 1985, preprint.

[3] A. LaVigna, A. M. Makowski and J. S. Baras, "A continuous-time distributed version of Wald' s sequential hypothesis testing problem," 1986, preprint.

[4] E. Geraniotis, "Robust distributed block and sequential discrete-time detection in uncertain environments", accepted for presentation at the 1987 Conference on Information Sciences and Systems, Johns Hopkins University, March 1987; to appear in the conference proceedings.

[5] P. J. Huber, "A robust version of the probability ratio test," *Ann. Math. Stat.,* Vol. 36, pp. 1753-1758, 1965.

[6] S. A. Kassam, "Robust hypothesis testing for bounded classes of probability densities," *IEEE Transactions on Information Theory,* vol. IT-27, pp. 242-247, 1981.

[7] K. S. Vastola and H. V. Poor, "On the p-point uncertainty class," *IEEE Transactions on Information Theory,* Vol. IT-30, pp.374-376, 1984.

[8] P. J. Huber and V. Strassen, "Minimax tests and the Neyman-Pearson lemma for capacities," *Ann. Statist.,* Vol. 1, pp. 251-265, 1973.

[9] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing," *Proceedings of the IEEE,* Vol. 73, pp. 433-481, March 1985.

[10] E. Geraniotis and Y. A. Chau, "On robust continuous-time discrimination," submitted for publication to the *IEEE Transactions on Information Theory,* March 1987.

[11] N. Shiryayev, *Optimal Stopping Ruels,* New York, Springer-Verlag, 1977.

[12] A. Wald, *Sequential Analysis,* New York, Wiley, 1947.