52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

Sequential Decision Making in Two-Dimensional Hypothesis Testing

Michael Carlisle, Olympia Hadjiliadis

Abstract— **In this work, we consider the problem of sequential decision making on the state of a two-sensor system with correlated noise. Each of the sensors is either receiving or not receiving a signal obstructed by noise, which gives rise to four possibilities: (noise, noise), (signal, noise), (noise, signal), (signal, signal). We set up the problem as a min-max optimization in which we devise a decision rule that minimizes the length of continuous observation time required to make a decision about the state of the system subject to error probabilities.**

We first assume that the noise in the two sources of observations is uncorrelated, and propose running in parallel two sequential probability ratio tests, each involving two thresholds. We compute these thresholds in terms of the error probabilities of the system. We demonstrate asymptotic optimality of the proposed rule as the error probabilities decrease without bound. We then analyze the performance of the proposed rule in the presence of correlation and discuss the degenerate cases of perfect positive or negative correlation. Finally, we purport the benefits of our proposed rule in a decentralized sensor system versus one in constant communication with a fusion center.

I. INTRODUCTION

Sequential hypothesis testing is one of the most classical problems arising in statistical sequential analysis and dating as far back as the work of Wald [19]. In the classical setup the problem is that of receiving a stream of sequential observations whose law follows one of two hypotheses, the null H_0 or the alternative H_1. The objective is then to minimize the number of observations required to make a decision subject to a pre-set error tolerance described as Type I and Type II errors (see, for instance, [17]). There are two main approaches to this problem: the Bayesian, in which each hypothesis is assigned an a priori probability of being true, and the min-max, in which no such assumption is made. The optimal solution to this problem is known to be given by the sequential probability ratio test (see, for instance, Chow, Robbins, and Siegmund [3]; Shiryaev [18]).

In the later years, the optimality of the sequential probability ratio test was examined and extended to various models (see, for instance, Liptser and Shiryaev [14], Irle [12]). Moreover, much work was done along the lines of extending two hypotheses regarding a stream of observations to the problem of distinguishing between many hypotheses both in the Bayesian (Baum and Veeravalli [1] , Dayanik, Poor, and Sezer [4], Dayanik and Goulding [5], Dragalin, Tartakovsky, and Veeravalli [6], Golubev and Khasminski [8]) and the min-max set up (Brodsky and Darkhovsky [2]). However, none of the above works consider the case of making an inference along multiple streams of observations.

This work was supported by the PSC-CUNY grant 65625-00 43, the NSA-MSP grant 081103, the NSF-DMS-ATD grant 1222526, and the NSF-CCF-MSC grant 0916452.

M. Carlisle is with the Department of Mathematics, Baruch College, City University of New York, 1 Bernard Baruch Way, New York, NY 10010 USA michael.carlisle@baruch.cuny.edu

O. Hadjiliadis is with the Department of Mathematics, Brooklyn College, City University of New York, 2900 Bedford Avenue, Brooklyn, NY 11210 USA ohadjiliadis@brooklyn.cuny.edu

In this paper, we examine the problem of testing four hypotheses on two streams of observations. Each of the hypotheses represents a physical state of the presence versus the absence of a signal obstructed by noise. In particular, we capture data and attempt to distinguish the presence of a signal, represented by a drift, in each of the two sensors, from Brownian noise (see, for example, [16], [18]). In other words, our problem is that of sequentially distinguishing a standard vs a drifted two-dimensional Brownian motion, and the objective is to minimize sampling time subject to error probabilities.

To address this problem we devise a sequential decision rule consisting of a stopping rule that declares the optimal time to stop sampling and a decision variable that declares a decision of the state of our system. Our proposed rule is the maximum of two sequential probability ratio tests, each with distinct thresholds, and our decision variable is determined from the exit location of the two sequential probability ratio statistics. Our first contribution is the explicit computation of the thresholds of the proposed rule in terms of the error probabilities. The novelty of our approach is in the construction of a test using a purely two-dimensional structure, allowing detection of a signal in each coordinate rather than merely the detection of a signal somewhere in the system (i.e. we are able to distinguish in which coordinate(s) a signal exists), subject to error probabilities for every possible case. We then proceed to show asymptotic optimality of the proposed rule in the absence of across-sensor correlation (i.e. the independent case) as the error probabilities decrease without bound (see [7]). We then proceed to examine the relationship of the error probabilities and sampling times in the presence of correlation to the error probabilities and sampling times of the proposed rule in the independent case.

A very important property of our proposed rule is that it can be implemented in a decentralized setup and still enjoy the same asymptotic optimality properties. In other words, each of the individual sequential probability ratio tests can be devised by each of the sensors separately, which can then communicate a binary bit of information to a central fusion center consisting of the alarm by the sequential probability ratio test and its exit side. The central fusion center can then make a decision once it receives a communication from both sensors and is thus not required to have access to the two-dimensional stream of sequential observations. This property

978-1-4673-5716-6/13/$31.00 ©2013 IEEE

makes the system very powerful because an optimal outcome can be achieved with a rather limited level of communication.

In Section II we set up the problem mathematically and present our proposed rule. In Section III, we investigate the properties of the rule in the absence of correlation (i.e. the independent case). In the first part of the section we show the relationship between the threshold selection and the error probabilities, while in the second part of the section we demonstrate asymptotic optimality of the proposed rule. In the third part of this section, we conduct a performance analysis of the expected sampling time in terms of the signal-to-noise ratio. In Section IV, we consider the case of non-zero correlation, establish inequalities relating the error probabilities and expected sampling times to those of the independent case, and discuss the perfect-correlation cases. In Section V we discuss the ability of our proposed rule to be implemented in a decentralized sensor system. We conclude with a discussion on future work.

II. Mathematical Setup

Consider a two-dimensional stochastic process,

$$dZ_t^{(k)} = \sigma_k dW_t^{(k)} + \mu_k dt, \;\; k = 1, 2$$

where $W_t^{(k)}$ are correlated Brownian motions with correlation $-1 < \rho < 1$, *i.e.*, $dW_t^{(1)} dW_t^{(2)} = \rho\, dt$. (We will also consider the limiting cases $\rho = \pm 1$ in Section IV.) We sequentially observe the 2-dimensional process $(Z_t^{(1)}, Z_t^{(2)})$, $t \geq 0$, and wish to test the following hypotheses:

$$\begin{aligned} &H_{00} : \mu_1 = 0, \, \mu_2 = 0 \qquad H_{10} : \mu_1 = m_1, \, \mu_2 = 0 \\ &H_{01} : \mu_1 = 0, \, \mu_2 = m_2 \quad H_{11} : \mu_1 = m_1, \, \mu_2 = m_2 \end{aligned} \tag{1}$$

where $m_1, m_2 \neq 0$ and are not necessarily equal.

The filtration generated by the vector $(Z_t^{(1)}, Z_t^{(2)})$ is denoted $\mathcal{F}_t = \sigma((Z_s^{(1)}, Z_s^{(2)}) : 0 \leq s \leq t)$. We distinguish this filtration from the marginal filtrations $\mathcal{F}_t^{(k)} = \sigma(Z_s^{(k)} : 0 \leq s \leq t)$, $k \in \{1, 2\}$. The hypotheses H_{ij} and the correlation ρ induce the joint normal probability measures $P_{ij,\rho}$. We distinguish $P_{ij,\rho}$ from their marginal probability measures $P_i^{(k)}$, which are the measures generated by the one-dimensional processes $Z_t^{(k)}$ on coordinate $k \in \{1, 2\}$; $P_{ij,\rho}$ has as its marginal measures $P_i^{(1)}$ and $P_j^{(2)}$, under the associated one-dimensional hypotheses

$$H_0^{(k)} : \mu_k = 0 \qquad\qquad H_1^{(k)} : \mu_k = m_k \tag{2}$$

for $k \in \{1, 2\}$. We are interested in developing a decision rule (τ, δ_τ) for this problem, where τ is a stopping rule with respect to $\mathcal{F}_t$, and δ_τ is an $\mathcal{F}_\tau$-measurable decision variable taking values in the index set $\{00, 01, 10, 11\}$. Optimality of our rule will be assessed in terms of minimizing the expected sampling time under each measure $P_{ij,\rho}$ subject to probabilities of error

$$\alpha_{ij,\rho} := P_{ij,\rho}(\delta_\tau \neq ij) \tag{3}$$

where $\alpha_{ij,\rho} < 0.5$. (The reason for these upper bounds will be given in Section III-A.) We will test the hypotheses (1) by tracking the two-dimensional process $(u_t^{(1)}, u_t^{(2)})$, where $u_t^{(k)}$ is the log-likelihood ratio of the marginal density in coordinate $k = 1, 2$. That is,

$$\begin{aligned} u_t^{(k)} &= x_k + \log \frac{dP_1^{(k)}}{dP_0^{(k)}} = \frac{m_k}{\sigma_k^2} \left(Z_t^{(k)} - \frac{1}{2} m_k t \right) \\ &= x_k + \frac{m_k}{\sigma_k^2} \left[\sigma_k W_t^{(k)} + \left(\mu_k - \frac{m_k}{2} \right) t \right] \end{aligned} \tag{4}$$

where $u_0^{(k)} = x_k$, $\mu_k = 0$ on $P_0^{(k)}$, and $\mu_k = m_k$ on $P_1^{(k)}$.

Finally, as a notational convenience, we define $P_{ij,\rho}^{(x,y)}$ and $E_{ij,\rho}^{(x,y)}$ as the probability and associated expectation given the starting point $(u_0^{(1)}, u_0^{(2)}) = (x_1, x_2) = (x, y)$, and $P_i^{(x_k)(k)}$ and $E_i^{(x_k)(k)}$ as the one-dimensional probability and expectation for measure $P_i^{(k)}$ given the starting point x_k in coordinate k. Wherever this notation is omitted, it is assumed that $(x_1, x_2) = (x, y) = (0, 0)$.

We wish to find an optimal decision rule of the threshold type

$$\tau = \inf\{t \geq 0 : (u_t^{(1)}, u_t^{(2)}) \notin A\} \tag{5}$$

where A is an appropriate set. The optimal decision rule $(\tilde{\tau}, \delta_{\tilde{\tau}})$ has the property

$$E_{ij,\rho}(\tilde{\tau}) \leq E_{ij,\rho}(\tau), \, i, j \in \{0, 1\}. \tag{6}$$

As in the one-dimensional case, we begin at

$$(u_0^{(1)}, u_0^{(2)}) = (0, 0) \in [a_1, b_1] \times [a_2, b_2] = A \tag{7}$$

and determine the values of $a_1 < 0 < b_1, a_2 < 0 < b_2$ based on the error probabilities $\alpha_{ij,\rho}$ defined in (3).

Our decision rule (τ, δ_τ) in one dimension (see, for example, [18]) is based on the first escape time of the interval $[a, b]$ where a and b are determined by the error probabilities under question. For the two-dimensional case, we are escaping a rectangle. When the noises are uncorrelated, our proposed rule runs two independent one-dimensional decision rule $(\tau_k, \delta_{\tau_k}^{(k)})$, where $\delta_{\tau_k}^{(k)} \in \{0, 1\}$, and

$$\begin{aligned} \tau_k &= \inf\{t \geq 0 : u_t^{(k)} \notin [a_k, b_k]\} \\ \delta_{\tau_k}^{(k)} &= 0 \text{ if } u_{\tau_k}^{(k)} = a_k, \, \delta_{\tau_k}^{(k)} = 1 \text{ if } u_{\tau_k}^{(k)} = b_k. \end{aligned} \tag{8}$$

We thus devise the two-dimensional decision rule (τ, δ_τ), where

$$\tau = \tau_1 \vee \tau_2, \;\; \delta_\tau = \delta_{\tau_1}^{(1)} \delta_{\tau_2}^{(2)}. \tag{9}$$

Next, we calculate the generator $\mathcal{L}_{ij,\rho}$ of the two-dimensional process $(u_t^{(1)}, u_t^{(2)})$ on the measure $P_{ij,\rho}$:

$$\begin{aligned} \mathcal{L}_{ij,\rho} := S_1 \left(\partial_{xx} + (-1)^{i+1} \partial_x \right) + S_2 \left(\partial_{yy} + (-1)^{j+1} \partial_y \right) \\ + 2\rho \sqrt{S_1 S_2} \partial_{xy} \end{aligned} \tag{10}$$

where $S_k := \frac{m_k^2}{2\sigma_k^2}, \, k \in \{1, 2\}$. Note that, for correlations ρ, ρ',

$$\mathcal{L}_{ij,\rho} = \mathcal{L}_{ij,\rho'} + 2(\rho - \rho') \sqrt{S_1 S_2} \partial_{xy}. \tag{11}$$

III. Independence: $\rho = 0$

In this section, we consider the case $\rho = 0$. We relate the thresholds of our proposed rule (9) to the error probabilities $\alpha_{ij,0}$ and prove asymptotic optimality of (9) as $\alpha_{ij,0}$ decrease without bound.

A. Error Probabilities and Thresholds

In this subsection we calculate the error probabilities and relate them to the thresholds of our two-dimensional proposed decision rule (9). Define

$$\gamma_{ij,\rho}(x,y) := P_{ij,\rho}^{(x,y)}(\delta_\tau = ij) \tag{12}$$

as the probability of correct decision in world ij. Then

$$\gamma_{ij,\rho}(0,0) = P_{ij,\rho}(\delta_\tau = ij) = 1 - \alpha_{ij,\rho}. \tag{13}$$

Moreover, $\gamma_{ij,\rho}(x,y)$ satisfies

$$\mathcal{L}_{ij,\rho}\gamma_{ij,\rho}(x,y) = 0 \tag{14}$$

subject to $0 \le \gamma_{ij,\rho}(x,y) \le 1$ for all $(x,y) \in A$, and the boundary conditions

$$\begin{aligned}
&\gamma_{00,\rho}(a_1,a_2) = 1 && \gamma_{01,\rho}(a_1,b_2) = 1 \\
&\gamma_{00,\rho}(b_1,y) = 0 \ \forall y \in [a_2,b_2] && \gamma_{01,\rho}(b_1,y) = 0 \ \forall y \in [a_2,b_2] \\
&\gamma_{00,\rho}(x,b_2) = 0 \ \forall x \in [a_1,b_1] && \gamma_{01,\rho}(x,a_2) = 0 \ \forall x \in [a_1,b_1] \\
&\gamma_{10,\rho}(b_1,a_2) = 1 && \gamma_{11,\rho}(b_1,b_2) = 1 \\
&\gamma_{10,\rho}(a_1,y) = 0 \ \forall y \in [a_2,b_2] && \gamma_{11,\rho}(a_1,y) = 0 \ \forall y \in [a_2,b_2] \\
&\gamma_{10,\rho}(x,b_2) = 0 \ \forall x \in [a_1,b_1] && \gamma_{11,\rho}(x,a_2) = 0 \ \forall x \in [a_1,b_1].
\end{aligned} \tag{15}$$

For $\rho = 0$, (14) reduces by separation of variables to

$$(\gamma_{ij,0})_{xx}(x,y) + (-1)^{i+1}(\gamma_{ij,0})_x(x,y) = \frac{\lambda}{S_1} \tag{16}$$

$$(\gamma_{ij,0})_{yy}(x,y) + (-1)^{j+1}(\gamma_{ij,0})_y(x,y) = -\frac{\lambda}{S_2} \tag{17}$$

with λ a constant. Solving the differential equations (16) and (17), we obtain $\lambda = 0$, and have the solutions

$$\gamma_{ij,0}(x,y) = f_i(x)g_j(y), \ (x,y) \in [a_1,b_1] \times [a_2,b_2] \tag{18}$$

where $f_i(x)$ and $g_j(y)$ are

$$\begin{aligned}
f_0(x) &= \frac{e^{b_1} - e^x}{e^{b_1} - e^{a_1}} \\
f_1(x) &= \frac{e^{-a_1} - e^{-x}}{e^{-a_1} - e^{-b_1}} = \frac{e^{b_1} - e^{a_1+b_1-x}}{e^{b_1} - e^{a_1}} = f_0(a_1+b_1-x) \\
g_0(y) &= \frac{e^{b_2} - e^y}{e^{b_2} - e^{a_2}} \\
g_1(y) &= \frac{e^{-a_2} - e^{-y}}{e^{-a_2} - e^{-b_2}} = \frac{e^{b_2} - e^{a_2+b_2-y}}{e^{b_2} - e^{a_2}} = g_0(a_2+b_2-y).
\end{aligned} \tag{19}$$

The fact that the differential equations above result in $\lambda = 0$ means that the choices of m_k and σ_k (and hence S_k) do not affect these probabilities in the independent case. The only parameters determining these thresholds are the error probabilities $\alpha_{ij,0}$ in (3).

We next give the $\alpha_{ij,0}$ in terms of the four thresholds a_1, a_2, b_1, b_2. Set $X = e^x$, $Y = e^y$, and $A_k = e^{a_k}$, $B_k = e^{b_k}$ for $k = 1,2$. Then

$$\begin{aligned}
\gamma_{00,0}(x,y) &= f_0(x)g_0(y) = \frac{X - B_1}{A_1 - B_1} \cdot \frac{Y - B_2}{A_2 - B_2} \\
\gamma_{01,0}(x,y) &= f_0(x)g_1(y) = \frac{X - B_1}{A_1 - B_1} \cdot \frac{B_2(A_2 - Y)}{Y(A_2 - B_2)} \\
\gamma_{10,0}(x,y) &= f_1(x)g_0(y) = \frac{B_1(A_1 - X)}{X(A_1 - B_1)} \cdot \frac{Y - B_2}{A_2 - B_2} \\
\gamma_{11,0}(x,y) &= f_1(x)g_1(y) = \frac{B_1(A_1 - X)}{X(A_1 - B_1)} \cdot \frac{B_2(A_2 - Y)}{Y(A_2 - B_2)}.
\end{aligned} \tag{20}$$

Setting $(x,y) = (0,0)$ and applying (13) yields

$$\begin{aligned}
\alpha_{00,0} &= \frac{(A_1 - B_1)(A_2 - B_2) - (1 - B_1)(1 - B_2)}{(A_1 - B_1)(A_2 - B_2)} \\
\alpha_{01,0} &= \frac{(A_1 - B_1)(A_2 - B_2) - (1 - B_1)B_2(A_2 - 1)}{(A_1 - B_1)(A_2 - B_2)} \\
\alpha_{10,0} &= \frac{(A_1 - B_1)(A_2 - B_2) - B_1(A_1 - 1)(1 - B_2)}{(A_1 - B_1)(A_2 - B_2)} \\
\alpha_{11,0} &= \frac{(A_1 - B_1)(A_2 - B_2) - B_1(A_1 - 1)B_2(A_2 - 1)}{(A_1 - B_1)(A_2 - B_2)}.
\end{aligned} \tag{21}$$

We can relate all four $\alpha_{ij,0}$. From the equalities

$$\begin{aligned}
C_1 := \frac{\gamma_{10,0}(0,0)}{\gamma_{00,0}(0,0)} &= \frac{1 - \alpha_{10,0}}{1 - \alpha_{00,0}} = \frac{f_1(0)}{f_0(0)} \\
= \frac{\gamma_{11,0}(0,0)}{\gamma_{01,0}(0,0)} &= \frac{1 - \alpha_{11,0}}{1 - \alpha_{01,0}}
\end{aligned} \tag{22}$$

$$\begin{aligned}
C_2 := \frac{\gamma_{01,0}(0,0)}{\gamma_{00,0}(0,0)} &= \frac{1 - \alpha_{01,0}}{1 - \alpha_{00,0}} = \frac{g_1(0)}{g_0(0)} \\
= \frac{\gamma_{11,0}(0,0)}{\gamma_{10,0}(0,0)} &= \frac{1 - \alpha_{11,0}}{1 - \alpha_{10,0}}
\end{aligned} \tag{23}$$

we obtain

$$(1 - \alpha_{00,0})(1 - \alpha_{11,0}) = (1 - \alpha_{01,0})(1 - \alpha_{10,0}) \tag{24}$$

which restricts the choice of error probabilities. In fact, we can use (21), (22), and (23) to relate A_k to B_k, $k = 1,2$:

$$B_k = \frac{C_k}{C_k + A_k - 1}, \ A_k = 1 + \frac{C_k(1 - B_k)}{B_k}. \tag{25}$$

Therefore,

$$b_k = -\ln\left(1 - \frac{1 - e^{a_k}}{C_k}\right), \ a_k = \ln\left(1 + C_k(e^{-b_k} - 1)\right). \tag{26}$$

We can relate our thresholds further if we have more information about the system:

Proposition 3.1: If $\rho = 0$, then any two of the following equalities imply the third: $a_1 = a_2$, $b_1 = b_2$, $\alpha_{01,0} = \alpha_{10,0}$.
Proof These follow directly from (26). ■

Furthermore, under $\rho = 0$, we can relate B_1 to B_2. From (25),

$$A_k - B_k = \frac{(C_k + B_k)(1 - B_k)}{B_k}. \tag{27}$$

From (21), we obtain

$$(A_1 - B_1)(A_2 - B_2)(1 - \alpha_{00,0}) = (1 - B_1)(1 - B_2). \tag{28}$$

Combining (27) and (28) yields

$$B_1 = \frac{C_1}{\frac{B_2}{(1-\alpha_{00,0})(C_2+B_2)} - 1} \tag{29}$$

which after taking logarithms results in

$$b_1 = \ln\left(\frac{C_1}{\frac{e^{b_2}}{(1-\alpha_{00,0})(C_2+e^{b_2})} - 1}\right). \tag{30}$$

Switching coordinates 1 and 2 in (29) and (30) holds by the symmetry of (28). We can now determine a lower bound for b_k: since $b_k > 0$, then using (29) and (30), we have

$$b_1 > \ln\left(\frac{1-\alpha_{10,0}}{\alpha_{00,0}}\right), \ \ b_2 > \ln\left(\frac{1-\alpha_{01,0}}{\alpha_{00,0}}\right) \tag{31}$$

which gives the reasoning for the error probability bounds given in Section II.

Applying (31) to (26) yields upper bounds for a_k. The fact that $B_k > 1$ implies that the denominator in the left equation of (25) is positive. This yields lower bounds for the a_k. Combined, these are, if $0 < C_k < 1$,

$$\ln(1 - C_k) < a_k < \ln\left(1 - C_k + \frac{\alpha_{00,0}}{1-\alpha_{00,0}}\right) \tag{32}$$

which can be written as

$$\ln\left(\frac{\alpha_{10,0}-\alpha_{00,0}}{1-\alpha_{00,0}}\right) < a_1 < \ln\left(\frac{\alpha_{10,0}}{1-\alpha_{00,0}}\right)$$
$$\ln\left(\frac{\alpha_{01,0}-\alpha_{00,0}}{1-\alpha_{00,0}}\right) < a_2 < \ln\left(\frac{\alpha_{01,0}}{1-\alpha_{00,0}}\right) \tag{33}$$

In the case that $C_k \geq 1$, the lower bound on a_k does not apply. In fact, the case $C_k > 1$ induces upper bounds on the b_k: by applying the fact that $0 < A_k < 1$ to the right equation of (25), we have

$$b_k < \ln\left(\frac{C_k}{C_k - 1}\right). \tag{34}$$

Equations (31) and (34) can be combined, if $C_k > 1$, as

$$\ln\left(\frac{1-\alpha_{10,0}}{\alpha_{00,0}}\right) < b_1 < \ln\left(\frac{1-\alpha_{10,0}}{\alpha_{00,0}-\alpha_{10,0}}\right)$$
$$\ln\left(\frac{1-\alpha_{01,0}}{\alpha_{00,0}}\right) < b_2 < \ln\left(\frac{1-\alpha_{01,0}}{\alpha_{00,0}-\alpha_{01,0}}\right) \tag{35}$$

A user setting three of the four error probabilities $\alpha_{ij,0}$ will automatically determine the fourth by (24). Then, setting a_1 (subject to (33)) will determine b_1 by (26), b_2 by switching coordinates in (30), and finally a_2 by (26) again.

B. Asymptotic Optimality

In this subsection we demonstrate asymptotic optimality of $\tau = \tau_1 \vee \tau_2$ of order-3, as per Fellouris and Moustakides [7]. That is, we prove, using the stopping times (8) with the error probabilities (21),

$$E_{ij,0}(\tau_1 \vee \tau_2) - E_{ij,0}(\tau_1) = o(1) \tag{36}$$

as the error probabilities $\alpha_{ij,0} \downarrow 0$. In establishing our result, we use the exponential killing trick: for any nonnegative random variable Y with no point mass at zero, its Laplace transform $E(e^{-\lambda Y})$ is its cumulative distribution function (CDF) $F_Y(t)$ killed at an exponential random variable with parameter λ. For $X_\lambda \sim \exp(\lambda)$ independent of Y, we have

$$\begin{aligned} E(e^{-\lambda Y}) &= \int_0^\infty e^{-\lambda t} dF_Y(t) \\ &= [e^{-\lambda t} F_Y(t)]_0^\infty + \int_0^\infty \lambda e^{-\lambda t} F_Y(t) dt \\ &= E(F_Y(X_\lambda)). \end{aligned} \tag{37}$$

We use this approach on τ_1 and $\tau_1 \vee \tau_2$, noting that, in the case that $\rho = 0$, τ_1 and τ_2 are independent. Hence, implicitly under measure $P_{ij,0}$ for our CDFs, $F_{\tau_1 \vee \tau_2}(t) = F_{\tau_1}(t) F_{\tau_2}(t)$, yielding

$$E_{ij,0}(e^{-\lambda \tau_1}) = E_{ij,0}(F_{\tau_1}(X_\lambda)) \tag{38}$$
$$E_{ij,0}(e^{-\lambda(\tau_1 \vee \tau_2)}) = E_{ij,0}(F_{\tau_1}(X_\lambda) F_{\tau_2}(X_\lambda)). \tag{39}$$

Since the stopping time τ_1 is an escape time, the tail probability $P_{ij,0}(\tau_1 > t) = 1 - F_{\tau_1}(t)$ is the probability that our process $u_t^{(1)}$ has not yet escaped the interval $[a_1, b_1]$. Using Laplace transforms, we rewrite (36) as

$$\begin{aligned} &E_{ij,0}(\tau_1 \vee \tau_2) - E_{ij,0}(\tau_1) \\ &\quad = \lim_{\lambda \to 0} \left[\frac{1 - E_{ij,0}(e^{-\lambda(\tau_1 \vee \tau_2)})}{\lambda} - \frac{1 - E_{ij,0}(e^{-\lambda \tau_1})}{\lambda}\right] \\ &\quad = \lim_{\lambda \to 0} \frac{E_{ij,0}(e^{-\lambda \tau_1}) - E_{ij,0}(e^{-\lambda(\tau_1 \vee \tau_2)})}{\lambda}. \end{aligned} \tag{40}$$

Our numerator is, using exponential killing,

$$\begin{aligned} &E_{ij,0}(e^{-\lambda \tau_1}) - E_{ij,0}(e^{-\lambda(\tau_1 \vee \tau_2)}) \\ &\quad = E_{ij,0}[F_{\tau_1}(X_\lambda)(1 - F_{\tau_2}(X_\lambda))] \end{aligned} \tag{41}$$

turning (40) into

$$\begin{aligned} &E_{ij,0}(\tau_1 \vee \tau_2) - E_{ij,0}(\tau_1) \\ &\quad = \lim_{\lambda \to 0} \frac{E_{ij,0}[F_{\tau_1}(X_\lambda)(1 - F_{\tau_2}(X_\lambda))]}{\lambda}. \end{aligned} \tag{42}$$

By the standard literature on Brownian motion with drift (for example, [15]), the difference on the left hand side of (42) is finite. Note that τ_k's boundaries a_k, b_k are functions of the $\alpha_{ij,0}$ (once the initial a_1 is selected within the bounds (33)). Sending the $\alpha_{ij,0} \to 0$ at the same rate, and noting

$$0 < A_k = e^{a_k} < 1 < B_k = e^{b_k} < \infty \tag{43}$$

then by (26) we see that we get

$$\lim_{\alpha_{ij,0} \to 0} C_k = 1 \implies \lim_{\alpha_{ij,0} \to 0} a_k = -b_k. \tag{44}$$

Rewriting (29) gives us

$$\frac{C_1 + B_1}{B_1} \cdot \frac{C_2 + B_2}{B_2} = \frac{1}{1 - \alpha_{00,0}}$$

which, by sending $\alpha_{ij,0} \to 0$ at the same rate on both sides, yields via $C_k \to 1$

$$\lim_{\alpha_{ij,0} \to 0} \frac{C_1 + B_1}{B_1} \cdot \frac{C_2 + B_2}{B_2} = \frac{1 + B_1}{B_1} \cdot \frac{1 + B_2}{B_2} = 1. \quad (45)$$

Since $B_k > 1$, (45) implies that $B_1 = B_2 = \infty$, and so $b_1 = b_2 = \infty$ in the limit. Thus, by (44), in the limit, $a_k \to -\infty$. Hence, as $\alpha_{ij,0} \to 0$, $\tau_k \to \infty$ a.s. and so $F_{\tau_k}(\cdot) \to 0$. Since CDFs are bounded by 1, we can use the dominated convergence theorem to obtain

$$\lim_{\alpha_{ij,0} \to 0} \frac{E_{ij,0}[F_{\tau_1}(X_\lambda)(1 - F_{\tau_2}(X_\lambda))]}{\lambda} = 0. \quad (46)$$

By continuity of the thresholds in the $\alpha_{ij,0}$, the finiteness of the left-hand side of (42), and again by the dominated convergence theorem (this time under λ),

$$\begin{aligned} &\lim_{\alpha_{ij,0} \to 0} \lim_{\lambda \to 0} \frac{E_{ij,0}[F_{\tau_1}(X_\lambda)(1 - F_{\tau_2}(X_\lambda))]}{\lambda} \\ &= \lim_{\lambda \to 0} \lim_{\alpha_{ij,0} \to 0} \frac{E_{ij,0}[F_{\tau_1}(X_\lambda)(1 - F_{\tau_2}(X_\lambda))]}{\lambda} = 0 \quad (47) \end{aligned}$$

which yields the asymptotic optimality limit

$$\lim_{\alpha_{ij,0} \to 0} E_{ij,0}(\tau_1 \vee \tau_2) - E_{ij,0}(\tau_1) = 0. \quad (48)$$

C. Performance Analysis

In this section we conduct a performance analysis of the expected sampling time of the proposed rule in terms of the signal-to-noise ratio of the first component, $S_k = \frac{m_k^2}{2\sigma_k^2}$, while fixing all other parameters.

In conducting this performance analysis, we set the parameters of the system as follows: the error probabilities $\alpha_{ij,0} = 0.01$ for $i, j = 0, 1$, and setting the free threshold $a_1 = -5.29078$ yields the remaining thresholds, by (26) and (30), of $a_2 = -5.29078$, $b_1 = b_2 = 5.29078$. This is a symmetric system, in which our starting point $(u_0^{(1)}, u_0^{(2)}) = (0, 0)$ is in the center of a square.

We use Monte Carlo simulation to approximate, with small $\lambda > 0$, $E_{ij,0}(\tau_1 \vee \tau_2)$, using exponential killing as seen in (39) and (40); that is,

$$\begin{aligned} T(m_1, m_2) := E_{ij,0}(\tau_1 \vee \tau_2) &\approx \frac{1 - E_{ij,0}(e^{-\lambda(\tau_1 \vee \tau_2)})}{\lambda} \\ &= \frac{1 - E_{ij,0}[F_{\tau_1}(X_\lambda) F_{\tau_2}(X_\lambda)]}{\lambda}. \quad (49) \end{aligned}$$

Figure 1 plots $T(m_1, m_2)$ for $(m_1, m_2) \in (0.1, 1.3)^2$ with the above thresholds, and $\sigma_1 = \sigma_2 = 1$. Note that, with symmetric signal strengths, error probabilities, and thresholds, Figure 1 is symmetric on $m_1 = m_2$, and predictably decreases as signal strength increases. Note that the symmetric example used results in the same values for $E_{ij,0}(\tau_1 \vee \tau_2)$ for any $i, j \in \{0, 1\}$.

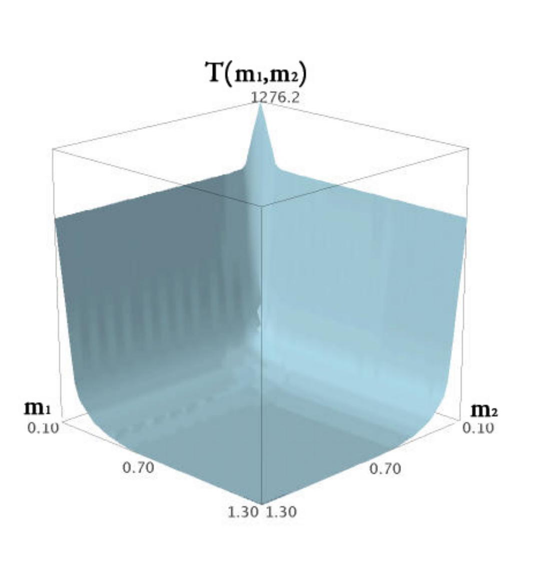

Fig. 1. Expected sampling time, Monte Carlo

IV. Correlation: $\rho \neq 0$

In this section we generalize to the correlated case $\rho \neq 0$ by relating it to the independent case.

A. Error Probabilities and Expected Sampling Times

Recall that

$$\gamma_{ij,\rho}(x, y) = P_{ij,\rho}^{(x,y)}(\delta_\tau = ij)$$

is the probability of making the correct decision under measure $P_{ij,\rho}$ with the ρ-correlation regime. We now relate these ρ-correlated probabilities to the independent case via the generator $\mathcal{L}_{ij,\rho}$.

Define $\Delta\gamma_{ij,\rho}(x, y) := \gamma_{ij,\rho}(x, y) - \gamma_{ij,0}(x, y)$. By (11),

$$\mathcal{L}_{ij,\rho}(\Delta\gamma_{ij,\rho}(0, 0)) = -2\rho\sqrt{S_1 S_2}\,\partial_{xy}\gamma_{ij,0}(0, 0). \quad (50)$$

The signs of $\rho \neq 0$ and $\partial_{xy}\gamma_{ij,0}$ determine whether $\Delta\gamma_{ij,\rho}(u_{t\wedge\tau_1}^{(1)}, u_{t\wedge\tau_2}^{(2)})$ is a supermartingale (if $\mathcal{L}_{ij,\rho}(\Delta\gamma_{ij,\rho}(x, y)) \leq 0$) or a submartingale (if $\mathcal{L}_{ij,\rho}(\Delta\gamma_{ij,\rho}(x, y)) \geq 0$) (see, for example, [15]). If $\Delta\gamma_{ij,\rho}(u_{t\wedge\tau_1}^{(1)}, u_{t\wedge\tau_2}^{(2)})$ is a supermartingale, then by the optional stopping theorem at the absorption time $\tau = \tau_1 \vee \tau_2$,

$$\Delta\gamma_{ij,\rho}(x, y) \geq E_{ij,\rho}^{(x,y)}(\Delta\gamma_{ij,\rho}(u_{\tau\wedge\tau_1}^{(1)}, u_{\tau\wedge\tau_2}^{(2)})) = 0 \quad (51)$$

which implies

$$\gamma_{ij,\rho}(x, y) \geq \gamma_{ij,0}(x, y). \quad (52)$$

The inequalities in (51) and (52) flip in the case that $\Delta\gamma_{ij,\rho}(u_{t\wedge\tau_1}^{(1)}, u_{t\wedge\tau_2}^{(2)})$ is a submartingale.

By (18) and (19),

$$\partial_{xy}\gamma_{ij,0}(x, y) = (\partial_x f_i(x))(\partial_y g_j(y)) \begin{cases} \geq 0 & ij \in \{00, 11\} \\ \leq 0 & ij \in \{01, 10\}. \end{cases}$$

Hence, by (50), for the cases $\rho > 0, ij \in \{00, 11\}$ and $\rho < 0, ij \in \{01, 10\}$, $\Delta\gamma_{ij,\rho}(u_{t\wedge\tau_1}^{(1)}, u_{t\wedge\tau_2}^{(2)})$ is a supermartingale, and for the cases $\rho < 0, ij \in \{00, 11\}$ and $\rho > 0, ij \in$

$\{01, 10\}$, $\Delta\gamma_{ij,\rho}(u^{(1)}_{t\wedge\tau_1}, u^{(2)}_{t\wedge\tau_2})$ is a submartingale. Therefore, by (50), (51), and (52), for $\rho < 0 < \rho'$,

$$\begin{aligned}\gamma_{ij,\rho}(0,0) \le \gamma_{ij,0}(0,0) \le \gamma_{ij,\rho'}(0,0),\ ij \in \{00, 11\}\\ \gamma_{ij,\rho}(0,0) \ge \gamma_{ij,0}(0,0) \ge \gamma_{ij,\rho'}(0,0),\ ij \in \{01, 10\}.\end{aligned} \quad (53)$$

As we send $|\rho|, |\rho'| \to 1$, we have upper bounds on these probabilities. This will be explained in Section IV-B.

Next, we establish inequalities relating $E_{ij,\rho}(\tau)$ to $E_{ij,0}(\tau)$ for $\rho \ne 0$. Set, on coordinate $k = 1, 2$,

$$\begin{aligned}G^{(k)}_i(x,t) &:= P^{(x)(k)}_i(\tau_k > t)\\ G_{ij,\rho}(x,y,t) &:= P^{(x,y)}_{ij,\rho}(\tau > t)\end{aligned} \quad (54)$$

and define the functions $g_{ij,\rho}$ by

$$g_{ij,\rho}(x,y) := E^{(x,y)}_{ij,\rho}(\tau) = \int_0^\infty G_{ij,\rho}(x,y,t)\,dt. \quad (55)$$

The function $g_{ij,\rho}(x,y)$ satisfies (see, for example, [15]),

$$\mathcal{L}_{ij,\rho} g_{ij,\rho}(x,y) = -1. \quad (56)$$

Our boundary conditions on $[a_1, a_2] \times [b_1, b_2]$ state that starting at any corner of the rectangle means no motion, and starting on a wall (where $x = a_k$ or $y = b_k$) reduces to the associated one-dimensional problem:

$$\begin{aligned}g_{ij,\rho}(a_1,a_2) &= g_{ij,\rho}(a_1,b_2) \quad (57)\\ &= g_{ij,\rho}(b_1,a_2) = g_{ij,\rho}(b_1,b_2) = 0\\ g_{ij,\rho}(x,a_2) &= g_{ij,\rho}(x,b_2) = E^{(x)(1)}_0(\tau_1)\ \forall x \in [a_1,b_1]\\ g_{ij,\rho}(a_1,y) &= g_{ij,\rho}(b_1,y) = E^{(y)(2)}_0(\tau_2)\ \forall y \in [a_2,b_2].\end{aligned}$$

From [18, (4.94)-(4.95)], the one-dimensional cases are, on coordinate $k = 1, 2$,

$$\begin{aligned}E^{(x)(k)}_0(\tau_k) &= \int_0^\infty P^{(x)(k)}_0(\tau_k > t)\,dt\\ &= \frac{2\sigma_k^2}{m_k^2}\left(\frac{e^{b_k} - e^x}{e^{b_k} - e^{a_k}}(b_k - a_k) - (b_k - x)\right)\\ &= \frac{1}{S_k}\left(f_0(x)(b_k - a_k) - (b_k - x)\right) \quad (58)\\ E^{(x)(k)}_1(\tau_k) &= \int_0^\infty P^{(x)(k)}_1(\tau_1 > t)\,dt\\ &= \frac{2\sigma_k^2}{m_k^2}\left(\frac{e^{b_k} - e^{a_k + b_k - x}}{e^{b_k} - e^{a_k}}(b_k - a_k) - (x - a_k)\right)\\ &= \frac{1}{S_k}\left(f_1(x)(b_k - a_k) - (x - a_k)\right). \quad (59)\end{aligned}$$

Define $\Delta g_{ij,\rho}(x,y) := g_{ij,\rho}(x,y) - g_{ij,0}(x,y)$. Then, as in (50),

$$\mathcal{L}_{ij,\rho}(\Delta g_{ij,\rho}(x,y)) = -2\rho\sqrt{S_1 S_2}\,\partial_{xy} g_{ij,0}(x,y). \quad (60)$$

As in the argument in (50)-(52), we need the signs of $\rho \ne 0$ and $\partial_{xy} g_{ij,0}(x,y)$ to determine inequalities about $g_{ij,\rho}(x,y)$. Note that, since

$$\tau = \tau_1 \vee \tau_2 = \tau_1 + \tau_2 - \tau_1 \wedge \tau_2 \quad (61)$$

we can rewrite $\partial_{xy} g_{ij,0}(x,y)$ in terms of $\tau_1 \wedge \tau_2$:

$$\begin{aligned}\partial_{xy} g_{ij,0}(x,y) &= \partial_{xy} E^{(x,y)}_{ij,0}(\tau)\\ &= \partial_{xy} E^{(x,y)}_{ij,0}(\tau_1) + \partial_{xy} E^{(x,y)}_{ij,0}(\tau_2) - \partial_{xy} E^{(x,y)}_{ij,0}(\tau_1 \wedge \tau_2)\\ &= \partial_y \partial_x E^x_i(\tau_1) + \partial_x \partial_y E^y_j(\tau_2) - \partial_{xy} E^{(x,y)}_{ij,0}(\tau_1 \wedge \tau_2)\\ &= -\partial_{xy} E^{(x,y)}_{ij,0}(\tau_1 \wedge \tau_2).\end{aligned} \quad (62)$$

We further decompose this into its integral form:

$$\begin{aligned}\partial_{xy} E^{(x,y)}_{ij,0}(\tau_1 \wedge \tau_2) &= \partial_{xy} \int_0^\infty P_{ij,0}(\tau_1 \wedge \tau_2 > t)\,dt\\ &= \int_0^\infty \partial_x P^{(x)(1)}_i(\tau_1 > t)\,\partial_y P^{(y)(2)}_j(\tau_2 > t)\,dt\\ &= \int_0^\infty \partial_x G^{(1)}_i(x,t)\,\partial_y G^{(2)}_j(y,t)\,dt.\end{aligned} \quad (63)$$

It is easily shown that the one-dimensional hitting times are concave for $x \in [a_k, b_k]$; the maxima are achieved at

$$\begin{aligned}\partial_x E^{(x)(k)}_0(\tau_k) = 0 &\iff x = x^{*(k)}_0 := \ln\left(\frac{e^{b_k} - e^{a_k}}{b_k - a_k}\right)\\ \partial_x E^{(x)(k)}_1(\tau_k) = 0 &\iff x = x^{*(k)}_1 := a_k + b_k - x^{*(k)}_0.\end{aligned} \quad (64)$$

$G^{(k)}_i(x,t)$ is a strictly decreasing function in t. Since $x^{*(k)}_i$ yields the maximum of

$$E^{(x)(k)}_i(\tau_k) = \int_0^\infty G^{(k)}_i(x,t)\,dt,\ k = 1, 2;\ i = 0, 1 \quad (65)$$

it therefore also yields the maximum of $G^{(k)}_i(x,t)$ in x for every t. Hence, we have that $\partial_x G^{(k)}_i(x^{*(k)}_i, t) = 0$ for every t.

This allows us to calculate the sign of $\partial_x G^{(1)}_i(x,t)\,\partial_y G^{(2)}_i(y,t)$, and hence the sign of $\partial_{xy} E^{(x,y)}_{ij,0}(\tau_1 \wedge \tau_2)$, at $(x,y) = (0,0)$. By the above argument, the point $(x^{*(1)}_i, y^{*(2)}_j)$ yields the maximum of $G^{(1)}_i(x,t)\,G^{(2)}_j(y,t)$ in A, and so $\partial_x G^{(1)}_i(x,t)\,\partial_y G^{(2)}_j(y,t) = 0$ for the line segments $x = x^{*(1)}_i$ and $y = y^{*(2)}_j$. The position of $(x^{*(1)}_i, y^{*(2)}_j)$ relative to $(0,0)$ determines the sign of each first partial, which determines the sign of the product, and therefore the sign of the integral (63). For $x' < x^{*(1)}_i < x''$,

$$\begin{aligned}\partial_x G^{(1)}_i(x',t) \ge \partial_x G^{(1)}_i(x^{*(k)}_i, t) = 0\\ \ge \partial_x G^{(1)}_i(x'',t)\end{aligned} \quad (66)$$

and likewise for $y' < y^{*(2)}_j < y''$ and $\partial_y G^{(2)}_j(y,t)$. We have

$$x^{*(1)}_1 < 0 < x^{*(1)}_0,\ y^{*(2)}_1 < 0 < y^{*(2)}_0. \quad (67)$$

Thus, by the same optional stopping theorem argument as in (51)-(52) with starting point $(x,y) = (0,0)$, for the cases $\rho > 0$, $ij \in \{00, 11\}$ and $\rho < 0$, $ij \in \{01, 10\}$, $\Delta g_{ij,\rho}(u^{(1)}_{t\wedge\tau_1}, u^{(2)}_{t\wedge\tau_2})$ is a submartingale, and for the cases $\rho < 0, ij \in \{00, 11\}$ and $\rho > 0, ij \in \{01, 10\}$,

$\Delta g_{ij,\rho}(u^{(1)}_{t\wedge\tau_1}, u^{(2)}_{t\wedge\tau_2})$ is a supermartingale. Therefore, by (60), (51), and (52), for $\rho < 0 < \rho'$,

$$\begin{aligned} &g_{ij,\rho}(0,0) \geq g_{ij,0}(0,0) \geq g_{ij,\rho'}(0,0),\ ij \in \{00, 11\} \\ &g_{ij,\rho}(0,0) \leq g_{ij,0}(0,0) \leq g_{ij,\rho'}(0,0),\ ij \in \{01, 10\}. \quad (68) \end{aligned}$$

As we send $|\rho|, |\rho'| \to 1$, we have bounds on these sampling times. This will be explained in Section IV-B.

B. The Degenerate Case: $\rho = \pm 1$

In the extreme cases $\rho = \pm 1$, complete correlation between the Brownian components degenerates the problem into a one-dimensional test. However, the choice of hypotheses "in play" depends on the sign of ρ, and the issue of decentralization vs. centralization becomes a consideration. This will be discussed in Section V.

We first examine $\rho = 1$: under perfect positive noise correlation, *i.e.*, $W^{(1)}_t = W^{(2)}_t$, our one-dimensional case exhibits an intriguing level of immediacy. If, in this case, $S_1 = S_2$, then the hypotheses H_{00} and H_{11} effectively run the one-dimensional test against each other, while the hypotheses H_{01} and H_{10} are immediately decidable. This is explained by examining $u^{(2)}_t$ as a linear combination of $u^{(1)}_t$ and drift:

$$\begin{aligned} &W^{(1)}_t = W^{(2)}_t \implies u^{(2)}_t = m^* u^{(1)}_t + y^*_{ij} t, \text{ where} \\ &m^* = \frac{m_2 \sigma_1}{m_1 \sigma_2} = \sqrt{\frac{S_2}{S_1}} \text{ and} \\ &y^*_{ij} = \frac{m_2}{\sigma_2}\left(\frac{\mu_2 - \frac{m_2}{2}}{\sigma_2} - \frac{\mu_1 - \frac{m_1}{2}}{\sigma_1}\right) \\ &\quad = 2\sqrt{S_2}\left[(-1)^i \sqrt{S_1} + (-1)^{j+1}\sqrt{S_2}\right]. \end{aligned}$$

In the case $S_1 = S_2$, $y^*_{00} = y^*_{11} = 0$ and so these two hypotheses are tested on the one-dimensional case for the two-dimensional point $(u^{(1)}_t, m^* u^{(1)}_t)$ moving along the line segment $y = m^* x$ between $(a_1, m^* a_1)$ and $(b_1, m^* b_1)$. If, for any $t > 0$, it is found that $m^* u^{(1)}_t > u^{(2)}_t$, then it is clear that $(-1)^i\sqrt{S_1} > (-1)^{j+1}\sqrt{S_2}$, *i.e.*, $ij = 01$; if $m^* u^{(1)}_t < u^{(2)}_t$, then $(-1)^i\sqrt{S_1} < (-1)^{j+1}\sqrt{S_2}$ implies $ij = 10$. For $S_1 \neq S_2$, all four hypotheses are immediately decidable, as the four possible values for y^*_{ij} are distinct and nonzero.

The case $\rho = -1$ operates similarly: if $W^{(2)}_t = -W^{(1)}_t$, then $u^{(2)}_t$ is a linear combination of $u^{(1)}_t$ and drift. That is,

$$\begin{aligned} &W^{(1)}_t = -W^{(2)}_t \implies u^{(2)}_t = -m^* u^{(1)}_t + y^{**}_{ij} t, \text{ where} \\ &y^{**}_{ij} = \frac{m_2}{\sigma_2}\left(\frac{\mu_2 - \frac{m_2}{2}}{\sigma_2} + \frac{\mu_1 - \frac{m_1}{2}}{\sigma_1}\right) \\ &\quad = 2\sqrt{S_2}\left[(-1)^{i+1}\sqrt{S_1} + (-1)^{j+1}\sqrt{S_2}\right], \end{aligned}$$

which results in $y^{**}_{01} = y^{**}_{10}$. Therefore, the hypotheses H_{01} and H_{10} play the one-dimensional test in the setting $S_1 = S_2$ along the line segment $y = -m^* x$ between $(a_1, -m^* a_1)$ and $(b_1, -m^* b_1)$. If, for any $t > 0$, it is found that $-m^* u^{(1)}_t > u^{(2)}_t$, then $ij = 00$, and $-m^* u^{(1)}_t < u^{(2)}_t$ implies $ij = 11$. In the case $S_1 \neq S_2$, again all four hypotheses are immediately decidable at any time $t > 0$.

As our decisions on certain cases are perfect, we have the following results on the correct-decision probabilities $\gamma_{ij,\rho}$ for $\rho = \pm 1$: upper bounds on (53) are

$$\gamma_{00,1}(0,0) = \gamma_{11,1}(0,0) = \gamma_{01,-1}(0,0) = \gamma_{10,-1}(0,0) = 1.$$

It immediately follows that we cannot specify certain error probabilities under $\rho = \pm 1$:

$$\alpha_{00,1}(0,0) = \alpha_{11,1}(0,0) = \alpha_{01,-1}(0,0) = \alpha_{10,-1}(0,0) = 0.$$

In addition, we have lower bounds on (68) (decisions are instant):

$$g_{00,-1}(0,0) = g_{11,-1}(0,0) = g_{01,1}(0,0) = g_{10,1}(0,0) = 0.$$

Upper bounds under perfect correlation are one-dimensional sampling times: by

$$\begin{aligned} g_{00,1}(0,0) = g_{01,-1}(0,0) = E^{(x=0)(1)}_0(\tau_1), \\ g_{11,1}(0,0) = g_{10,-1}(0,0) = E^{(x=0)(1)}_1(\tau_1). \end{aligned}$$

There is a geometric intuition that goes nicely with this problem: under perfect positive noise correlation, one expects data to fall on a line with positive drift. In our hypothesis test, the endpoints of this positive drift line are precisely the decision boundaries for hypotheses H_{00} and H_{11}. Likewise for perfect negative correlation; the endpoints of the line with negative slope are the decision boundaries for H_{01} and H_{10}. This hints that there may be a formulation for this problem which deforms the shape of the boundary in exchange for fixing the expected sampling time.

V. Decentralization

Our proposed rule (9) is implementable in a decentralized setup (see [7],[10]): τ_k, $k = 1, 2$, are stopping times of the log-likelihood ratios $u^{(k)}_t$ of the marginal probability measures $P^{(k)}_i$ instead of the joint measures $P_{ij,\rho}$. Thus, the stopping time τ_k only requires access to the information from the filtration $\mathcal{F}^{(k)}_t = \sigma(Z^{(k)}_s : 0 \leq s \leq t)$ instead of the much larger filtration of the joint process $\mathcal{F}_t = \sigma((Z^{(1)}_s, Z^{(2)}_s) : 0 \leq s \leq t)$. We can thus build a system of sensors in the following fashion: sensor k receives signal $Z^{(k)}_t$ and makes a decision on (2) with rule (8). It then communicates this decision to a fusion center with the single bit $\delta^{(k)}_{\tau_k}$. At time $\tau = \tau_1 \vee \tau_2$, the fusion center makes the decision $\delta_\tau = \delta^{(1)}_{\tau_1}\delta^{(2)}_{\tau_2}$.

The asymptotic optimality result (36) implies that there is no loss in performance between the centralized case, in which the fusion center receives the continuous raw data $\{(Z^{(1)}_t, Z^{(2)}_t)\}_{t \geq 0}$, and the decentralized case described above. This greatly increases the usefulness of this model in sensor network analysis.

The results of Section IV-B imply that the level of correlation gives a measure to how useful decentralization is for such a sensor setup: the more closely correlated the noise, the less useful decentralization is in reducing the sampling time to decision. In fact, a single communication to the fusion center can make a decision under perfect correlation

($\rho = \pm 1$) if the signal-to-noise ratios of the sensors do not match ($S_1 \neq S_2$), and even in certain cases where they do.

VI. Conclusions, Future Work

We have shown that, for $-1 < \rho < 1$, the threshold boundaries of this test are the corners of a rectangle whose one-dimensional absorption sides reduce the two-dimensional problem to the related one-dimensional problem. A related question worthy of investigation is, given $\rho \neq 0$ and fixed error probabilities $\alpha_{ij,\rho}$, how would we change the shape of the boundary of A (call it A_ρ) and the rule (9) to achieve asymptotic optimality of this test? In general, this boundary need not be rectangular.

For example, in the range $0 < \rho < 1$, should the domain "squeeze shut" continuously from the rectangle to the diagonal $\{(x, y) \in [a_1, b_1] \times [m^* a_1, m^* b_1]\}$ at $\rho = 1$? This would seem to imply that the time it takes to reach the boundary would reduce, *i.e.*, $E_{ij,\rho}(\tau) \leq E_{ij,0}(\tau)$ for $\rho > 0$, $ij \in \{01, 10\}$. Likewise, should it "flip and squeeze shut" for $-1 < \rho < 0$, ending in the diagonal $\{(x, y) \in [a_1, b_1] \times [-m^* a_1, -m^* b_1]\}$ at $\rho = -1$? Finally, the case of nonconstant correlation, i.e. $\rho = \rho(t)$, should be examined.

The analysis here is done under a continuous-monitoring scenario. While this is a good approximation to random walk / discrete models of the same type, a more thorough analysis can be undertaken to examine the issue of discretized data collection, such as in batch or sliding-window processing. Just as in the one-dimensional discrete time paradigm of Shiryaev [18], we expect that the results are valid in discretized data collection.

VII. Acknowledgements

The authors gratefully acknowledge Jay Rosen for introducing the authors.

References

[1] C. Baum and V. Veeravalli, A sequential procedure for multihypothesis testing, *IEEE Transactions on Information Theory*, Vol. 40, No. 6, pp. 1994-1997, 1994.

[2] B. Brodsky and B. Darkhovsky, Minimax methods for multihypothesis sequential testing and change-point detection problems, *Sequential Analysis*, Vol. 27, No. 2, pp. 141-173, 2008.

[3] Y. S. Chow, H. Robbins, and D. Siegmund, *The theory of optimal stopping*, (Dover Publications), 1991.

[4] S. Dayanik, H, V. Poor, and S. Sezer, Sequential multi-hypothetis testing for compound Poisson processes, *Stochastics*, Vol. 80, No. 1, pp. 19-50, 2008.

[5] S. Dayanik and C. Goulding, Sequential detection and identification of a change in the distribution of a Markov-modulated random sequence, *IEEE Transactions on Information Theory*, Vol. 55, No. 7, pp. 3323-3345, 2009.

[6] V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli, Multihypothesis sequential probability ratio tests-Part I: Asymptotic Optimality, *IEEE Transactions on Information Theory*, Vol. 45, No. 7, pp. 2448-2461, 1999.

[7] G. Fellouris and G.V. Moustakides, Decentralized sequential hypothesis testing using asynchronous communication, *IEEE Transactions on Information Theory*, Vol. 57, No. 1, pp. 534-548, 2011.

[8] G. K. Golubev and R.Z. Khas'minski, Sequential testing for several signals in Gaussian white noise, *Theory of Probability and its applications*, Vol. 28, pp. 573-584, 1983.

[9] O. Hadjiliadis, T. Schäfer, and H. V. Poor, Quickest detection in coupled systems, *Proceedings of the 48th IEEE Conference on Decision and Control*, Shanghai, China, December 16-18, 2009.

[10] O. Hadjiliadis, H. Zhang, and H. V. Poor, One shot schemes for decentralized quickest change detection, *IEEE Transactions on Information Theory*, Vol. 55, No. 7, pp. 3346-3359, 2009.

[11] H. He, W. P. Kierstead, and J. Rebholz, Double lookbacks. *Mathematical Finance*, Vol. 8, pp. 201-228, 1998.

[12] A. Irle, Transitivity in problems of optimal stopping, *Annals of Probability*, Vol. 9, pp. 642-647, 1981.

[13] T. L. Lai, Asymptotic optimality of invariant sequential probability ratio test, *Annals of Statistics*, Vol. 9, pp. 318-333, 1981.

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

[15] B. Øksendal, *Stochastic Differential Equations.* (Springer-Verlag, New York), 2000.

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

[17] D. Ruppert, *Statistics and Data Analysis for Financial Engineering*. (Springer Texts in Statistics), 2011.

[18] A. N. Shiryayev, *Optimal Stopping Rules*. (Springer-Verlag, New York), 1978.

[19] A. Wald, *Sequential Analysis.* (Wiley, New York), 1947.

[20] Y. Wang and Y. Mei, Asymptotic Optimality Theory for Decentralized Sequential Multihypothesis Testing Problems, *IEEE Transactions on Information Theory*, Vol. 57, No. 10, pp. 7068-7083, 2011.

[21] H. Zhang and O. Hadjiliadis. Quickest detection in a system with correlated noise, *Proceedings of the 51st IEEE Conference on Decision and Control*, Maui, Hawai, Deecember 10-13, 2012.