

# Sequential Decentralized Detection under Noisy Channels

Yasin Yilmaz

Electrical Engineering Department  
Columbia University  
New York, NY 10027  
Email: yasin@ee.columbia.edu

George Moustakides

Dept. of Electrical & Computer Engineering  
University of Patras  
26500 Rion, Greece  
Email: moustaki@upatras.gr

Xiaodong Wang

Electrical Engineering Department  
Columbia University  
New York, NY 10027  
Email: wangx@ee.columbia.edu

**Abstract**—We consider decentralized detection through distributed sensors that perform level-triggered sampling and communicate with a fusion center via noisy channels. Each sensor computes its local log-likelihood ratio (LLR), samples it using the level-triggered sampling, and upon sampling transmits a single bit to the FC. Upon receiving a bit from a sensor, the FC updates the global LLR and performs a sequential probability ratio test (SPRT) step. We derive the fusion rules under various types of channels. We further provide non-asymptotic and asymptotic analyses on the average detection delay for the proposed channel-aware scheme, and show that the asymptotic detection delay is characterized by a KL information number. The delay analysis facilitates the choice of appropriate signaling schemes under different channel types for sending the 1-bit information from sensors to the FC.

## I. INTRODUCTION

In this paper, we consider the decentralized detection problem in which a number of sensors communicate to a fusion center (FC) under bandwidth constraints. The FC is responsible for making the final decision based on the limited information it receives from the sensors. [1] and [2], which are among the early works treating the problem, showed the optimality of likelihood ratio test as a local decision rule at sensors and a global decision rule at the FC, respectively. Most of the works on decentralized detection, including the above mentioned, followed the fixed sample size approach where the FC makes the global decision at a fixed time. On the other hand, some works, e.g., [3]–[5], considered the sequential approach where the decision time of the FC is random. It has been observed that sequential methods require, on average, approximately four times [6, Page 109] less samples (for Gaussian signals) to reach a decision than their fixed sample size counterparts, for the same level of confidence. Relaxing the strict bandwidth constraint, it is known that data fusion (multi-bit messaging) is a much more powerful technique than decision fusion (one-bit messaging), albeit it consumes higher bandwidth [7]. Moreover, the level-triggered sampling-based sequential detection schemes, recently proposed in [4] and [5], are as powerful as data fusion techniques, and at the same time they are as simple and bandwidth-efficient as decision fusion techniques.

In practice, there are two sources of uncertainty in a decentralized system, namely the noise in listening channels, cor-

rupting the sensor observations, and the noise in transmission channels, corrupting the messages received by the FC. The conventional approach to decentralized detection considers the former, but ignores the latter, e.g., [1]–[5]. In other words, the conventional approach assumes ideal transmission channels while applying a fusion rule to the messages received by the FC. Following the conventional approach, before applying a fusion rule, a communication block can be independently applied to recover the information bits transmitted from sensors. However, this two-step solution causes performance loss due to the data processing inequality [8]. For an optimal solution, the fusion rule should be channel-aware, i.e., should use the the knowledge on noisy channels [9], [10].

In this paper, we design and analyze channel-aware sequential decentralized detection schemes based on level-triggered sampling, under different types of discrete and continuous noisy channels. In particular, we first derive channel-aware sequential detection schemes based on level-triggered sampling. We then present an information theoretic framework to analyze the decision delay performance of the proposed schemes based on which we provide both non-asymptotic and asymptotic analyses on the decision delays under various types of channels. Based on the expressions of the asymptotic decision delays, we also consider appropriate signaling schemes under different continuous channels to minimize the asymptotic delays.

The remainder of the paper is organized as follows. In Section II, we describe the general structure of the decentralized detection approach based on level-triggered sampling with noisy channels between sensors and the FC. In Section III, we derive channel-aware fusion rules at the FC for various types of channels. Next, we provide analyses on the decision delay performance for ideal channel and noisy channels in Section IV and Section V, respectively. Finally, the paper is concluded in Section VI.

## II. SYSTEM DESCRIPTIONS

Consider a wireless sensor network consisting of  $K$  sensors each of which observes a Nyquist-rate sampled discrete-time signal  $\{y_t^k : t \in \mathbb{N}\}$ ,  $k = 1, \dots, K$ . Each sensor  $k$  computes the log-likelihood ratios (LLR)  $\{L_t^k\}_t$  of the signal it observes, samples the LLR sequence using the *level-triggered sampling*,

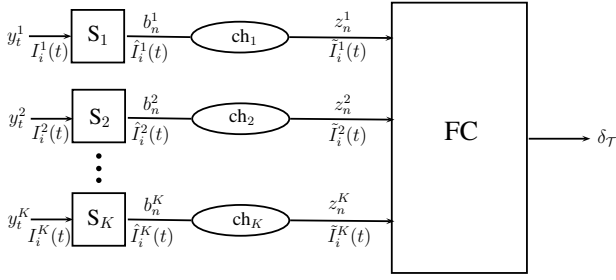


Fig. 1. A wireless sensor network with  $K$  sensors, and an FC. Sensors process their observations  $\{y_t^k\}$ , and transmits information bits  $\{b_n^k\}$ . Then, the FC, receiving  $\{z_n^k\}$  through wireless channels, makes a detection decision  $\delta_\tau$ .  $I_i^k(t)$ ,  $\hat{I}_i^k(t)$ ,  $\bar{I}_i^k(t)$  are the observed, transmitted and received information entities respectively, which will be defined in Section IV.

and then sends the LLR samples to the fusion center (FC). The FC then combines the local LLR information from all sensors, and decides between two hypotheses,  $H_0$  and  $H_1$ , in a sequential manner.

Observations collected at the same sensor,  $\{y_t^k\}_t$ , are assumed to be i.i.d., and in addition observations collected at different sensors,  $\{y_t^k\}_k$ , are assumed to be independent. Hence, the local LLR at the  $k$ -th sensor,  $L_t^k$ , and the global LLR,  $L_t$ , are computed as

$$L_t^k \triangleq \log \frac{f_1^k(y_1^k, \dots, y_t^k)}{f_0^k(y_1^k, \dots, y_t^k)} = L_{t-1}^k + l_t^k = \sum_{n=1}^t l_n^k, \quad (1)$$

and  $L_t = \sum_{k=1}^K L_t^k$ , respectively, where  $l_t^k \triangleq \log \frac{f_1^k(y_t^k)}{f_0^k(y_t^k)}$  is the LLR of the sample  $y_t^k$  received at the  $k$ -th sensor at time  $t$ ;  $f_i^k$ ,  $i = 0, 1$ , is the probability density function (pdf) of the received signal by the  $k$ -th sensor under  $H_i$ . The  $k$ -th sensor samples  $L_t^k$  via level-triggered sampling at a sequence of random sampling times  $\{t_n^k\}_n$  that are dictated by  $L_t^k$  itself. Specifically, the  $n$ -th sample is taken from  $L_t^k$  whenever the accumulated LLR  $L_t^k - L_{t_{n-1}^k}^k$ , since the last sampling time  $t_{n-1}^k$  exceeds a constant  $\Delta$  in absolute value, i.e.,

$$t_n^k \triangleq \min \left\{ t > t_{n-1}^k : L_t^k - L_{t_{n-1}^k}^k \notin (-\Delta, \Delta) \right\}, \quad (2)$$

where  $t_0^k = 0$ ,  $L_0^k = 0$ . Let  $\lambda_n^k$  denote the accumulated LLR during the  $n$ -th inter-sampling interval,  $(t_{n-1}^k, t_n^k]$ , i.e.,

$$\lambda_n^k \triangleq \sum_{t=t_{n-1}^k+1}^{t_n^k} l_t^k = L_{t_n^k}^k - L_{t_{n-1}^k}^k. \quad (3)$$

Immediately after sampling at  $t_n^k$ , as shown in Fig. 1, an information bit  $b_n^k$  indicating the threshold crossed by  $\lambda_n^k$  is transmitted to the FC, i.e.,

$$b_n^k \triangleq \text{sign}(\lambda_n^k). \quad (4)$$

Note that each sensor, in fact, performs the Wald's well known *sequential probability ratio test* (SPRT) with thresholds  $\Delta$  and  $-\Delta$ . SPRT is the optimum sequential test in the i.i.d. case in terms of minimizing the average decision delay [11].

Detailed information on SPRT can be found in [12]. At sensor  $k$  the  $n$ -th local SPRT starts at time  $t_{n-1}^k$  and ends at time  $t_n^k$  when the local test statistic  $\lambda_n^k$  exceeds either  $\Delta$  or  $-\Delta$ . This local hypothesis test produces a local decision represented by the information bit  $b_n^k$ , and induces local error probabilities  $\alpha_k$  and  $\beta_k$  which are given by

$$\alpha_k \triangleq P_0(b_n^k = 1), \quad \text{and} \quad \beta_k \triangleq P_1(b_n^k = -1), \quad (5)$$

respectively, where  $P_i(\cdot)$ ,  $i = 0, 1$ , denotes the probability under  $H_i$ .

Denote the received signal at the FC corresponding to  $b_n^k$  as  $z_n^k$  [cf. Fig. 1]. The FC then computes the LLR  $\tilde{\lambda}_n^k$  of each received signal and approximates the global LLR  $L_t$  as

$$\tilde{L}_t \triangleq \sum_{k=1}^K \sum_{n=1}^{N_t^k} \tilde{\lambda}_n^k \quad \text{with} \quad \tilde{\lambda}_n^k \triangleq \log \frac{p_1^k(z_n^k)}{p_0^k(z_n^k)}, \quad (6)$$

where  $N_t^k$  is the total number of LLR messages the  $k$ -th sensor has transmitted up to time  $t$ , and  $p_i^k(\cdot)$ ,  $i = 0, 1$ , is the pdf of  $z_n^k$  under  $H_i$ . In particular, suppose that the  $m$ -th LLR message  $\tilde{\lambda}_m$  from any sensor is received at time  $t_m$ . Then at  $t_m$ , the FC first updates the global LLR as

$$\tilde{L}_{t_m} = \tilde{L}_{t_m-1} + \tilde{\lambda}_m. \quad (7)$$

It then performs an SPRT step by comparing  $\tilde{L}_{t_m}$  with two thresholds  $\tilde{A}$  and  $-\tilde{B}$ , and applying the following decision rule

$$\delta_{t_m} \triangleq \begin{cases} H_1, & \text{if } \tilde{L}_{t_m} \geq \tilde{A}, \\ H_0, & \text{if } \tilde{L}_{t_m} \leq -\tilde{B}, \\ -, & \text{otherwise.} \end{cases} \quad (8)$$

In other words, it continues to receive LLR messages if  $\tilde{L}_{t_m} \in (-\tilde{B}, \tilde{A})$ . The thresholds  $(\tilde{A}, \tilde{B} > 0)$  are selected to satisfy the error probability constraints  $P_0(\delta_\tau = H_1) \leq \alpha$  and  $P_1(\delta_\tau = H_0) \leq \beta$  with equalities, where  $\alpha, \beta$  are target error probability bounds, and

$$\tilde{\mathcal{T}} \triangleq \min\{t > 0 : \tilde{L}_t \notin (-\tilde{B}, \tilde{A})\} \quad (9)$$

is the detection delay.

With ideal channels between sensors and the FC, we have  $z_n^k = b_n^k$ , so from (5) we can write the local LLR  $\tilde{\lambda}_n^k = \hat{\lambda}_n^k$ , where

$$\hat{\lambda}_n^k \triangleq \begin{cases} \log \frac{P_1(b_n^k=1)}{P_0(b_n^k=1)} = \log \frac{1-\beta_k}{\alpha_k}, & \text{if } b_n^k = 1, \\ \log \frac{P_1(b_n^k=-1)}{P_0(b_n^k=-1)} = \log \frac{\beta_k}{1-\alpha_k}, & \text{if } b_n^k = -1. \end{cases} \quad (10)$$

In this paper, we will assume, as in [4], that the local error probabilities  $\{\alpha_k, \beta_k\}_k$  are available at the FC in order to compute the LLR  $\hat{\lambda}_n^k$  of the received signals. For the case of ideal channels, we use the  $A$  and  $-B$  to denote the thresholds in (8), i.e.,  $\tilde{A} = A, \tilde{B} = B$ , and use  $\mathcal{T}$  to denote the decision delay in (9), i.e.,  $\tilde{\mathcal{T}} = \mathcal{T}$ .

In the case of noisy channels, the received signal  $z_n^k$  is not always identical to the transmitted bit  $b_n^k$ , and thus the LLR  $\tilde{\lambda}_n^k$  of  $z_n^k$  can be different from  $\hat{\lambda}_n^k$  of  $b_n^k$  given in (10). In the next section, we consider some popular channel models and give the corresponding expressions for  $\tilde{\lambda}_n^k$ .

### III. CHANNEL-AWARE FUSION RULES

In computing the LLR  $\tilde{\lambda}_n^k$  of the received signal  $z_n^k$ , we will make use of the local sensor error probabilities  $\alpha_k, \beta_k$ , and the channel parameters that characterize the statistical property of the channel. In this paper, we assume sampling (communication) times are reliably detected at the FC under continuous channels by using high enough signaling levels at the sensors. The unreliable detection of sampling times under continuous channels is analyzed in [13, Section VI].

#### A. Binary Erasure Channels (BEC)

Consider binary erasure channels between sensors and the FC with erasure probabilities  $\{\epsilon_k\}_k$ . Under BEC, a transmitted bit  $b_n^k$  is lost with probability  $\epsilon_k$ , and correctly received at the FC, i.e.,  $z_n^k = b_n^k$ , with probability  $1 - \epsilon_k$ . Then the LLR of  $z_n^k$  is given by

$$\tilde{\lambda}_n^k = \begin{cases} \log \frac{P_1(z_n^k=1)}{P_0(z_n^k=1)} = \log \frac{1-\beta_k}{\alpha_k}, & \text{if } z_n^k = 1, \\ \log \frac{P_1(z_n^k=-1)}{P_0(z_n^k=-1)} = \log \frac{\beta_k}{1-\alpha_k}, & \text{if } z_n^k = -1. \end{cases} \quad (11)$$

Note that under BEC the channel parameter  $\epsilon_k$  is not needed when computing the LLR  $\tilde{\lambda}_n^k$ . Note also that in this case, a received bit bears the same amount of LLR information as in the ideal channel case, although a transmitted bit is not always received. Hence, the channel-aware approach coincides with the conventional approach which relies solely on the received signal. Although the LLR updates in (10) and (11) are identical, the fusion rules under BEC and ideal channels are not since the thresholds  $\hat{A}$  and  $-\hat{B}$  of BEC, due to the information loss, are in general different from the thresholds  $A$  and  $-B$  of the ideal channel case.

#### B. Binary Symmetric Channels (BSC)

Next, we consider binary symmetric channels with crossover probabilities  $\{\epsilon_k\}_k$  between sensors and the FC. Under BSC, the transmitted bit  $b_n^k$  is flipped, i.e.,  $z_n^k = -b_n^k$ , with probability  $\epsilon_k$ , and it is correctly received, i.e.,  $z_n^k = b_n^k$ , with probability  $1 - \epsilon_k$ . The LLR of  $z_n^k$  can be computed as

$$\begin{aligned} \tilde{\lambda}_n^k(z_n^k = 1) &= \log \frac{P_1^1(z_n^k = 1)P_1^1 + P_1^{-1}(z_n^k = 1)P_1^{-1}}{P_0^1(z_n^k = 1)P_0^1 + P_0^{-1}(z_n^k = 1)P_0^{-1}} \\ &= \log \frac{(1 - \epsilon_k)(1 - \beta_k) + \epsilon_k\beta_k}{(1 - \epsilon_k)\alpha_k + \epsilon_k(1 - \alpha_k)} \\ &= \log \frac{1 - \overbrace{[(1 - 2\epsilon_k)\beta_k + \epsilon_k]}^{\hat{\beta}_k}}{\underbrace{(1 - 2\epsilon_k)\alpha_k + \epsilon_k}_{\hat{\alpha}_k}} \end{aligned} \quad (12)$$

where  $P_i^j, i = 0, 1, j = -1, 1$ , is the probability under  $H_i$  given  $b_n^k = j$ , and  $\{\hat{\alpha}_k, \hat{\beta}_k\}$  are the effective local error probabilities at the FC under BSC. Similarly we can write

$$\tilde{\lambda}_n^k(z_n^k = -1) = \log \frac{\hat{\beta}_k}{1 - \hat{\alpha}_k}. \quad (13)$$

Note that  $\hat{\alpha}_k > \alpha_k, \hat{\beta}_k > \beta_k$  if  $\alpha_k < 0.5, \beta_k < 0.5, \forall k$ , which we assume true for  $\Delta > 0$ . Thus, we have  $|\tilde{\lambda}_{n,BSC}^k| <$

$|\tilde{\lambda}_{n,BEC}^k|$  from which we expect the performance loss under BSC will be higher than that under BEC. The numerical results provided in Section V-B will illustrate this. Finally, note also that, unlike the BEC case, under BSC the FC needs to know the channel parameters  $\{\epsilon_k\}$  to operate in a channel-aware manner.

#### C. Additive White Gaussian Noise (AWGN) Channels

Now, assume that the channel between each sensor and the FC is an AWGN channel. The received signal at the FC is given by

$$z_n^k = h_n^k x_n^k + w_n^k \quad (14)$$

where  $h_n^k = h_k, \forall k, n$ , is a known constant complex channel gain;  $w_n^k \sim \mathcal{N}_c(0, \sigma_k^2)$ ;  $x_n^k$  is the transmitted signal at sampling time  $t_n^k$  given by

$$x_n^k = \begin{cases} a, & \text{if } \lambda_n^k \geq \Delta, \\ b, & \text{if } \lambda_n^k \leq -\Delta. \end{cases} \quad (15)$$

where the transmission levels  $a$  and  $b$  are complex in general.

The distribution of the received signal is then  $z_n^k \sim \mathcal{N}_c(h_k x_n^k, \sigma_k^2)$ . The LLR of  $z_n^k$  is given by

$$\begin{aligned} \tilde{\lambda}_n^k &= \log \frac{p_k^a(z_n^k)P_1^a + p_k^b(z_n^k)P_1^b}{p_k^a(z_n^k)P_0^a + p_k^b(z_n^k)P_0^b} \\ &= \log \frac{(1 - \beta_k) \exp(-c_n^k) + \beta_k \exp(-d_n^k)}{\alpha_k \exp(-c_n^k) + (1 - \alpha_k) \exp(-d_n^k)}, \end{aligned} \quad (16)$$

where the superscript  $a$  (resp.  $b$ ) denotes the conditioning on  $x_n^k = a$  (resp.  $x_n^k = b$ ),  $c_n^k \triangleq \frac{|z_n^k - h_k a|^2}{\sigma_k^2}$  and  $d_n^k \triangleq \frac{|z_n^k - h_k b|^2}{\sigma_k^2}$ .

#### D. Rayleigh Fading Channels

If a Rayleigh fading channel is assumed between each sensor and the FC, the received signal model is also given by (14)-(15), but with  $h_n^k \sim \mathcal{N}_c(0, \sigma_{h,k}^2)$ . We then have  $z_n^k \sim \mathcal{N}_c(0, |x_n^k|^2 \sigma_{h,k}^2 + \sigma_k^2)$ ; and accordingly, similar to (16),  $\tilde{\lambda}_n^k$  is written as

$$\tilde{\lambda}_n^k = \log \frac{\frac{1-\beta_k}{\sigma_{a,k}^2} \exp(-c_n^k) + \frac{\beta_k}{\sigma_{b,k}^2} \exp(-d_n^k)}{\frac{\alpha_k}{\sigma_{a,k}^2} \exp(-c_n^k) + \frac{1-\alpha_k}{\sigma_{b,k}^2} \exp(-d_n^k)} \quad (17)$$

where  $\sigma_{a,k}^2 \triangleq |a|^2 \sigma_{h,k}^2 + \sigma_k^2, \sigma_{b,k}^2 \triangleq |b|^2 \sigma_{h,k}^2 + \sigma_k^2, c_n^k \triangleq \frac{|z_n^k|^2}{\sigma_{a,k}^2}$  and  $d_n^k \triangleq \frac{|z_n^k|^2}{\sigma_{b,k}^2}$ .

#### E. Rician Fading Channels

For Rician fading channels, we have  $h_n^k \sim \mathcal{N}_c(\mu_k, \sigma_{h,k}^2)$  in (14), and hence  $z_n^k \sim \mathcal{N}_c(\mu_k x_n^k, |x_n^k|^2 \sigma_{h,k}^2 + \sigma_k^2)$ . Using  $\sigma_{a,k}^2$  and  $\sigma_{b,k}^2$  as defined in the Rayleigh fading case, and defining  $c_n^k \triangleq \frac{|z_n^k - \mu_k a|^2}{\sigma_{a,k}^2}, d_n^k \triangleq \frac{|z_n^k - \mu_k b|^2}{\sigma_{b,k}^2}$  we can write  $\tilde{\lambda}_n^k$  as in (17).

### IV. PERFORMANCE ANALYSIS FOR IDEAL CHANNELS

In this section, we derive the exact non-asymptotic expression for the average detection delay  $E_i[T]$ , and provide an asymptotic analysis on it as the error probability bounds  $\alpha, \beta \rightarrow 0$ . Before proceeding to the analysis, let us define some information entities which will be used throughout this and next sections.

### A. Information Entities

Note that the expectation of an LLR corresponds to a *Kullback-Leibler (KL) information* entity. For instance,

$$I_1^k(t) \triangleq \mathbb{E}_1 \left[ \log \frac{f_1^k(y_1^k, \dots, y_t^k)}{f_0^k(y_1^k, \dots, y_t^k)} \right] = \mathbb{E}_1[L_t^k],$$

and  $I_0^k(t) \triangleq \mathbb{E}_0 \left[ \log \frac{f_0^k(y_1^k, \dots, y_t^k)}{f_1^k(y_1^k, \dots, y_t^k)} \right] = -\mathbb{E}_0[L_t^k]$  (18)

are the KL divergences of the local LLR sequence  $\{L_t^k\}_t$  under  $H_1$  and  $H_0$ , respectively. Similarly

$$\hat{I}_1^k(t) \triangleq \mathbb{E}_1 \left[ \log \frac{p_1^k(b_1^k, \dots, b_{N_t^k}^k)}{p_0^k(b_1^k, \dots, b_{N_t^k}^k)} \right] = \mathbb{E}_1[\hat{L}_t^k], \quad \tilde{I}_1^k(t) \triangleq -\mathbb{E}_0[\hat{L}_t^k]$$

$$\tilde{I}_1^k(t) \triangleq \mathbb{E}_1 \left[ \log \frac{p_1^k(z_1^k, \dots, z_{N_t^k}^k)}{p_0^k(z_1^k, \dots, z_{N_t^k}^k)} \right] = \mathbb{E}_1[\tilde{L}_t^k], \quad \tilde{I}_0^k(t) \triangleq -\mathbb{E}_0[\tilde{L}_t^k]$$

(19)

are the KL divergences of the local LLR sequences  $\{\hat{L}_t^k\}_t$  and  $\{\tilde{L}_t^k\}_t$  respectively. Define also  $I_i(t) \triangleq \sum_{k=1}^K I_i^k(t)$ ,  $\hat{I}_i(t) \triangleq \sum_{k=1}^K \hat{I}_i^k(t)$ , and  $\tilde{I}_i(t) \triangleq \sum_{k=1}^K \tilde{I}_i^k(t)$  as the KL divergences of the global LLR sequences  $\{L_t\}$ ,  $\{\hat{L}_t\}$ , and  $\{\tilde{L}_t\}$  respectively.

In particular, we have  $I_1^k(1) = \mathbb{E}_1 \left[ \log \frac{f_1^k(y_1^k)}{f_0^k(y_1^k)} \right] = \mathbb{E}_1[l_t^k]$  and  $I_0^k(1) = \mathbb{E}_0 \left[ \log \frac{f_0^k(y_1^k)}{f_1^k(y_1^k)} \right] = -\mathbb{E}_0[l_t^k]$  as the KL information numbers of the LLR sequence  $\{l_t^k\}$ ; and  $I_i(1) \triangleq \sum_{k=1}^K I_i^k(1)$ ,  $i = 0, 1$  are those of the global LLR sequence  $\{l_t\}$ . Moreover,  $I_1^k(t_1^k) = \mathbb{E}_1 \left[ \log \frac{f_1^k(y_1^k, \dots, y_{t_1^k}^k)}{f_0^k(y_1^k, \dots, y_{t_1^k}^k)} \right] = \mathbb{E}_1[\lambda_n^k]$ ,  $\hat{I}_1^k(t_1^k) = \mathbb{E}_1 \left[ \log \frac{p_1^k(b_1^k)}{p_0^k(b_1^k)} \right] = \mathbb{E}_1[\hat{\lambda}_n^k]$ , and  $\tilde{I}_1^k(t_1^k) = \mathbb{E}_1 \left[ \log \frac{p_1^k(z_1^k)}{p_0^k(z_1^k)} \right] = \mathbb{E}_1[\tilde{\lambda}_n^k]$ , are the KL information numbers of the local LLR sequences  $\{\lambda_n^k\}$ ,  $\{\hat{\lambda}_n^k\}$ , and  $\{\tilde{\lambda}_n^k\}$ , respectively, under  $H_1$ . Likewise, we have  $I_0^k(t_1^k) = -\mathbb{E}_0[\lambda_n^k]$ ,  $\hat{I}_1^k(t_1^k) = -\mathbb{E}_0[\hat{\lambda}_n^k]$ , and  $\tilde{I}_1^k(t_1^k) = -\mathbb{E}_0[\tilde{\lambda}_n^k]$  under  $H_0$ . To summarize,  $I_i^k(t)$ ,  $\hat{I}_i^k(t)$ , and  $\tilde{I}_i^k(t)$  are respectively the observed (at sensor  $k$ ), transmitted (by sensor  $k$ ), and received (by the FC) KL information entities as illustrated in Fig. 1.

Next we define the following information ratios,  $\hat{\eta}_i^k \triangleq \frac{\hat{I}_i^k(t_1^k)}{I_i^k(t_1^k)}$  and  $\tilde{\eta}_i^k \triangleq \frac{\tilde{I}_i^k(t_1^k)}{I_i^k(t_1^k)}$ , which represent how efficiently information is transmitted from sensor  $k$  and received by the FC, respectively. Due to the data processing inequality, we have  $0 \leq \hat{\eta}_i^k, \tilde{\eta}_i^k \leq 1, \forall i, k$ . We further define  $\hat{I}_i(1) \triangleq \sum_{k=1}^K \hat{\eta}_i^k I_i^k(1) = \sum_{k=1}^K \hat{I}_i^k(1)$ , and  $\tilde{I}_i(1) \triangleq \sum_{k=1}^K \tilde{\eta}_i^k I_i^k(1) = \sum_{k=1}^K \tilde{I}_i^k(1)$ , as the effective transmitted and received values corresponding to the KL information  $I_i(1)$ , respectively. Note that  $\hat{I}_i(1)$  and  $\tilde{I}_i(1)$  are not real KL information numbers, but projections of  $I_i(1)$  onto the filtrations generated by the transmitted, (i.e.,  $\{b_n^k\}$ ), and received, (i.e.,  $\{z_n^k\}$ ), signal sequences, respectively. This is because sensors do not transmit and the FC does not receive the LLR of a single observation, but instead they transmit and it receives the LLR messages of several observations. Hence, we cannot have the KL information for single observations at the two ends of the

communication channel, but we can define hypothetical KL information to serve analysis purposes.

The KL information  $I_i^k(1)$  of a sensor whose information ratio,  $\hat{\eta}_i^k$ , is high and close to 1 is well projected to the FC. Conversely,  $I_i^k(1)$  of a sensor which undergoes high information loss is poorly projected to the FC. Note that there are two sources of information loss for sensors, namely, the overshoot effect due to having discrete-time observations and noisy transmission channels. The latter appears only in  $\tilde{\eta}_i^k$ , whereas the former appears in both  $\hat{\eta}_i^k$  and  $\tilde{\eta}_i^k$ . In general with discrete-time observations at sensors we have  $\hat{I}_i(1) \neq I_i(1)$  and  $\tilde{I}_i(1) \neq I_i(1)$ . Lastly, note that under ideal channels, since  $z_n^k = b_n^k, \forall k, n$ , we have  $\tilde{I}_i(1) = \hat{I}_i(1)$ .

### B. Analysis on Detection Delay

Let  $\{\tau_n^k : \tau_n^k = t_n^k - t_{n-1}^k\}_n$  denote the inter-arrival times of the LLR messages transmitted from the  $k$ -th sensor. Note that  $\tau_n^k$  depends on the observations  $y_{t_{n-1}^k+1}^k, \dots, y_{t_n^k}^k$ , and since  $\{y_t^k\}_t$  are i.i.d.,  $\{\tau_n^k\}_n$  are also i.i.d. random variables. Hence, the counting process  $\{N_t^k\}_t$  is a renewal process. Similarly the LLRs  $\{\lambda_n^k\}_n$  of the received signals at the FC are also i.i.d. random variables, and form a renewal-reward process. Note from (9) that the SPRT can stop in between two arrival times of sensor  $k$ , e.g.,  $t_n^k \leq \mathcal{T} < t_{n+1}^k$ . The event  $N_{\mathcal{T}}^k = n$  occurs if and only if  $t_n^k = \tau_1^k + \dots + \tau_n^k \leq \mathcal{T}$  and  $t_{n+1}^k = \tau_1^k + \dots + \tau_{n+1}^k > \mathcal{T}$ , so it depends on the first  $(n+1)$  LLR messages. From the definition of *stopping time* [14, pp. 104] we conclude that  $N_{\mathcal{T}}^k$  is not a stopping time for the processes  $\{\tau_n^k\}_n$  and  $\{\lambda_n^k\}_n$  since it depends on the  $(n+1)$ -th message. However,  $N_{\mathcal{T}}^k + 1$  is a stopping time for  $\{\tau_n^k\}_n$  and  $\{\lambda_n^k\}_n$  since we have  $N_{\mathcal{T}}^k + 1 = n \iff N_{\mathcal{T}}^k = n - 1$  which depends only on the first  $n$  LLR messages. Hence, from Wald's identity [14, pp. 105] we can directly write the following equalities

$$\mathbb{E}_i \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \tau_n^k \right] = \mathbb{E}_i[\tau_1^k] (\mathbb{E}_i[N_{\mathcal{T}}^k] + 1), \quad (20)$$

$$\text{and } \mathbb{E}_i \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \hat{\lambda}_n^k \right] = \mathbb{E}_i[\hat{\lambda}_1^k] (\mathbb{E}_i[N_{\mathcal{T}}^k] + 1). \quad (21)$$

We have the following theorem on the average detection delay under ideal channels.

**Theorem 1.** *Consider the decentralized detection scheme given in Section II, with ideal channels between sensors and the FC. Its average detection delay under  $H_i$  is given by*

$$\mathbb{E}_i[\mathcal{T}] = \frac{\hat{I}_i(\mathcal{T})}{\hat{I}_i(1)} + \frac{\sum_{k=1}^K \hat{I}_i^k(t_{N_{\mathcal{T}}^k+1}^k) - \mathbb{E}_i[\mathcal{Y}_k] \hat{I}_i^k(1)}{\hat{I}_i(1)} \quad (22)$$

where  $\mathcal{Y}_k$  is a random variable representing the time interval between the stopping time and the arrival of the first bit from the  $k$ -th sensor after the stopping time, i.e.,  $\mathcal{Y}_k \triangleq t_{N_{\mathcal{T}}^k+1}^k - \mathcal{T}$ .

*Proof:* From (20) and (21) we obtain

$$\mathbb{E}_i \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \tau_n^k \right] = \mathbb{E}_i[\tau_1^k] \frac{\mathbb{E}_i \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \hat{\lambda}_n^k \right]}{\mathbb{E}_i[\hat{\lambda}_1^k]}$$

where the left-hand side equals to  $\mathbb{E}_i[\mathcal{T}] + \mathbb{E}_i[\mathcal{Y}_k]$ . Note that  $\mathbb{E}_i[\tau_1^k]$  is the expected stopping time of the local SPRT at the  $k$ -th sensor and by Wald's identity it is given by  $\mathbb{E}_i[\tau_1^k] = \frac{\mathbb{E}_i[\hat{\lambda}_1^k]}{\mathbb{E}_i[l_1^k]}$ , provided that  $\mathbb{E}_i[l_1^k] \neq 0$ . Hence, we have

$$\begin{aligned} \mathbb{E}_i[\mathcal{T}] &= \frac{\mathbb{E}_i[\hat{\lambda}_1^k]}{\mathbb{E}_i[l_1^k]} \frac{\mathbb{E}_i \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \hat{\lambda}_n^k \right]}{\mathbb{E}_i[l_1^k]} - \mathbb{E}_i[\mathcal{Y}_k] \\ &= \frac{I_i^k(t_1^k)}{I_i^k(t_1^k)} \frac{\hat{I}_i^k(\mathcal{T}) + \hat{I}_i^k(t_{N_{\mathcal{T}}^k+1}^k)}{I_i^k(1)} - \mathbb{E}_i[\mathcal{Y}_k] \end{aligned}$$

where we used the fact that  $\mathbb{E}_1 \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \hat{\lambda}_n^k \right] = \mathbb{E}_1[\hat{L}_{\mathcal{T}}^k] + \tilde{\mathbb{E}}_1[\hat{\lambda}_{N_{\mathcal{T}}^k+1}^k] = \hat{I}_1^k(\mathcal{T}) + \hat{I}_1^k(t_{N_{\mathcal{T}}^k+1}^k)$  and similarly  $\mathbb{E}_0 \left[ \sum_{n=1}^{N_{\mathcal{T}}^k+1} \hat{\lambda}_n^k \right] = -\hat{I}_0^k(\mathcal{T}) - \hat{I}_0^k(t_{N_{\mathcal{T}}^k+1}^k)$ . Note that  $\tilde{\mathbb{E}}_i[\cdot]$  is the expectation with respect to  $\hat{\lambda}_{N_{\mathcal{T}}^k+1}^k$  and  $N_{\mathcal{T}}^k$  under  $H_i$ . By rearranging the terms and then summing over  $k$  on both sides, we obtain

$$\begin{aligned} \mathbb{E}_i[\mathcal{T}] \underbrace{\sum_{k=1}^K I_i^k(1) \frac{\hat{I}_i^k(t_1^k)}{I_i^k(t_1^k)}}_{\hat{I}_i(1)} &= \\ \hat{I}_i(\mathcal{T}) + \sum_{k=1}^K \hat{I}_i^k(t_{N_{\mathcal{T}}^k+1}^k) - \mathbb{E}_i[\mathcal{Y}_k] \underbrace{I_i^k(1) \frac{\hat{I}_i^k(t_1^k)}{I_i^k(t_1^k)}}_{\hat{I}_i^k(1)} & \end{aligned}$$

which is equivalent to (22).  $\blacksquare$

The result in (22) is in fact very intuitive. Recall that  $\hat{I}_i(\mathcal{T})$  is the KL information at the detection time at the FC. It naturally lacks some local information that has been accumulated at sensors, but has not been transmitted to the FC, i.e., the information gathered at sensors after their last sampling times. The numerator of the second term on the right hand side of (22) replaces such missing information by using the hypothetical KL information. Note that in (22)  $\hat{I}_i^k(t_{N_{\mathcal{T}}^k+1}^k) \neq \hat{I}_i^k(t_1^k)$ , i.e.,  $\tilde{\mathbb{E}}_i[\hat{\lambda}_{N_{\mathcal{T}}^k+1}^k] \neq \mathbb{E}_i[\hat{\lambda}_1^k]$ , since  $N_{\mathcal{T}}^k$  and  $\hat{\lambda}_{N_{\mathcal{T}}^k+1}^k$  are not independent.

The next result gives the asymptotic detection delay performance under ideal channels.

**Theorem 2.** *As  $\alpha, \beta \rightarrow 0$ , the average detection delay under ideal channels given by (22) satisfies*

$$\mathbb{E}_1[\mathcal{T}] = \frac{|\log \alpha|}{\hat{I}_1(1)} + O(1), \quad \text{and} \quad \mathbb{E}_0[\mathcal{T}] = \frac{|\log \beta|}{\hat{I}_0(1)} + O(1), \quad (23)$$

where  $O(1)$  represents a constant term.

*Proof:* We will prove the first equality in (23), and the proof of the second one follows similarly. Let us first prove the following lemma.

**Lemma 1.** *As  $\alpha, \beta \rightarrow 0$  we have the following KL information at the FC*

$$\hat{I}_1(\mathcal{T}) = |\log \alpha| + O(1), \quad \text{and} \quad \hat{I}_0(\mathcal{T}) = |\log \beta| + O(1). \quad (24)$$

*Proof:* We will show the first equality and the second one follows similarly. We have

$$\begin{aligned} \hat{I}_1(\mathcal{T}) &= \mathbb{P}_1(\hat{L}_{\mathcal{T}} \geq A) \mathbb{E}_1[\hat{L}_{\mathcal{T}} | \hat{L}_{\mathcal{T}} \geq A] + \\ &\quad \mathbb{P}_1(\hat{L}_{\mathcal{T}} \leq -B) \mathbb{E}_1[\hat{L}_{\mathcal{T}} | \hat{L}_{\mathcal{T}} \leq -B] \\ &= (1 - \beta)(A + \mathbb{E}_1[\theta_A]) - \beta(B + \mathbb{E}_1[\theta_B]), \quad (25) \end{aligned}$$

where  $\theta_A, \theta_B$  are overshoot and undershoot respectively given by  $\theta_A \triangleq \hat{L}_{\mathcal{T}} - A$  if  $\hat{L}_{\mathcal{T}} \geq A$  and  $\theta_B \triangleq -\hat{L}_{\mathcal{T}} - B$  if  $\hat{L}_{\mathcal{T}} \leq -B$ . From [4, Theorem 2], we have  $A \leq |\log \alpha|$  and  $B \leq |\log \beta|$ , so as  $\alpha, \beta \rightarrow 0$  (25) becomes  $\hat{I}_1(\mathcal{T}) = A + \mathbb{E}_1[\theta_A] + o(1)$ . Assuming  $0 < \Delta < \infty$  and  $|l_t^k| < \infty, \forall k, t$ , we have  $0 < \alpha_k, \beta_k < 1$ , hence from (10) we have  $|\hat{\lambda}_n^k| < \infty$ , and accordingly  $\hat{I}_i^k(t_1^k) = \mathbb{E}_i[\hat{\lambda}_1^k] < \infty$ . Since the overshoot cannot exceed the last received LLR value, we have  $\theta_A, \theta_B \leq \Theta = \max_{k,n} |\hat{\lambda}_n^k| < \infty$ . Similar to Eq. (73) in [4] we can write  $\beta \geq e^{-B-\Theta}$  and  $\alpha \geq e^{-A-\Theta}$  where  $\Theta = O(1)$  by the above argument, or equivalently,  $B \geq |\log \beta| - O(1)$  and  $A \geq |\log \alpha| - O(1)$ . Hence, we have  $A = |\log \alpha| + O(1)$  and  $B = |\log \beta| + O(1)$ .  $\blacksquare$

From the assumption of  $|l_t^k| < \infty, \forall k, t$ , we also have  $\hat{I}_i(1) \leq I_i(1) < \infty$ . Moreover, we have  $\mathbb{E}_i[\mathcal{Y}_k] \leq \mathbb{E}_i[\tau_1^k] < \infty$  since  $\mathbb{E}_i[l_1^k] \neq 0$ . Note that all of the terms on the right-hand side of (22) except  $\hat{I}_i(\mathcal{T})$  do not depend on the global error probabilities  $\alpha, \beta$ , so they are  $O(1)$  as  $\alpha, \beta \rightarrow 0$ . Finally, substituting (24) into (22) we get (23).  $\blacksquare$

It is seen from (23) that the hypothetical KL information number,  $\hat{I}_i(1)$ , plays a key role in the asymptotic detection delay expression. We can asymptotically minimize  $\mathbb{E}_i[\mathcal{T}]$  by maximizing  $\hat{I}_i(1)$ . Recalling its definition

$$\hat{I}_i(1) = \sum_{k=1}^K \frac{\hat{I}_i^k(t_1^k)}{I_i^k(t_1^k)} I_i^k(1)$$

we see that three information numbers are required to compute it. Note that  $I_i^k(1) = \mathbb{E}_i[l_1^k]$  and  $I_i^k(t_1^k) = \mathbb{E}_i[\hat{\lambda}_1^k]$ , which is given in (26) below, are computed based on local observations at sensors, thus do not depend on the channels between sensors and the FC. Specifically, we have

$$\begin{aligned} I_1^k(t_1^k) &= (1 - \beta_k)(\Delta + \mathbb{E}_1[\bar{\theta}_n^k]) - \beta_k(\Delta + \mathbb{E}_1[\underline{\theta}_n^k]), \\ \text{and} \quad I_0^k(t_1^k) &= \alpha_k(\Delta + \mathbb{E}_0[\bar{\theta}_n^k]) - (1 - \alpha_k)(\Delta + \mathbb{E}_0[\underline{\theta}_n^k]) \end{aligned} \quad (26)$$

where  $\bar{\theta}_n^k$  and  $\underline{\theta}_n^k$  are local over(under)shoots given by  $\bar{\theta}_n^k \triangleq \lambda_n^k - \Delta$  if  $\lambda_n^k \geq \Delta$  and  $\underline{\theta}_n^k \triangleq -\lambda_n^k - \Delta$  if  $\lambda_n^k \leq -\Delta$ . Due to having  $|l_t^k| < \infty, \forall k, t$  we have  $\bar{\theta}_n^k, \underline{\theta}_n^k < \infty, \forall k, n$ .

On the other hand,  $\hat{I}_i^k(t_1^k)$  represents the information received in an LLR message by the FC, so it heavily depends on the channel type. In the ideal channel case, from (10) it is

given by

$$\hat{I}_1^k(t_1^k) = (1 - \beta_k) \log \frac{1 - \beta_k}{\alpha_k} + \beta_k \log \frac{\beta_k}{1 - \alpha_k}, \quad (27)$$

$$\text{and } \hat{I}_0^k(t_1^k) = \alpha_k \log \frac{1 - \beta_k}{\alpha_k} + (1 - \alpha_k) \log \frac{\beta_k}{1 - \alpha_k}.$$

Since  $\hat{I}_i^k(t_1^k)$  is the only channel-dependent term in the asymptotic detection delay expression, in the next section we will obtain its expression for each noisy channel type considered in Section III.

## V. PERFORMANCE ANALYSIS FOR NOISY CHANNELS

In all noisy channel types that we consider in this paper, we assume that channel parameters are either constants or i.i.d. random variables across time. In other words,  $\epsilon_k, h_k$  are constant for all  $k$  (see Section III-A, III-B, III-C), and  $\{h_n^k\}_n, \{w_n^k\}_n$  are i.i.d. for all  $k$  (see Section III-C, III-D, III-E). Thus, in all noisy channel cases discussed in Section III the inter-arrival times of the LLR messages  $\{\tilde{\tau}_n^k\}_n$ , and the LLRs of the received signals  $\{\tilde{\lambda}_n^k\}_n$  are i.i.d. across time as in the ideal channel case. Accordingly the average detection delay in these noisy channels has the same expression as in (22), as given by the following proposition. The proof is similar to that of Theorem 1.

**Proposition 1.** *Under each type of noisy channel discussed in Section III, the average detection delay is given by*

$$E_i[\tilde{T}] = \frac{\tilde{I}_i(\tilde{T})}{\tilde{I}_i(1)} + \frac{\sum_{k=1}^K \tilde{I}_i^k(t_{N_{\tilde{T}}^k+1}^k) - E_i[\tilde{Y}_k] \tilde{I}_i^k(1)}{\tilde{I}_i(1)} \quad (28)$$

where  $\tilde{Y}_k \triangleq t_{N_{\tilde{T}}^k+1}^k - \tilde{T}$ .

The asymptotic performances under noisy channels can also be analyzed analogously to the ideal channel case.

**Proposition 2.** *As  $\alpha, \beta \rightarrow 0$ , the average detection delay under noisy channels given by (28) satisfies*

$$E_1[\tilde{T}] = \frac{|\log \alpha|}{\tilde{I}_1(1)} + O(1), \text{ and } E_0[\tilde{T}] = \frac{|\log \beta|}{\tilde{I}_0(1)} + O(1). \quad (29)$$

*Proof:* Note that in the noisy channel cases the FC, as discussed in Section III, computes the LLR,  $\tilde{\lambda}_n^k$ , of the signal it receives, and then performs SPRT using the LLR sum  $\tilde{L}_t$ . Hence, analogous to Lemma 1 we can show that  $\tilde{I}_1(\tilde{T}) = |\log \alpha| + O(1)$  and  $\tilde{I}_0(\tilde{T}) = |\log \beta| + O(1)$  as  $\alpha, \beta \rightarrow 0$ . Note also that due to channel uncertainties  $|\tilde{\lambda}_n^k| \leq |\hat{\lambda}_n^k|$ , so we have  $\tilde{I}_i^k(t_1^k) \leq \hat{I}_i^k(t_1^k) < \infty$  and  $\tilde{I}_i(1) \leq \hat{I}_i(1) < \infty$ . We also have  $E_i[\tilde{Y}_k] \leq E_i[\tilde{\tau}_1^k] < \infty$  as in the ideal channel case. Substituting these asymptotic values in (28) we get (29). ■

Recall that  $\tilde{I}_i(1) = \sum_{k=1}^K \frac{\tilde{I}_i^k(t_1^k)}{\tilde{I}_i^k(t_1^k)} I_i^k(1)$  in (29) where  $I_i^k(1)$  and  $I_i^k(t_1^k)$  are independent of the channel type, i.e., they are same as in the ideal channel case. In the subsequent subsections, we will compute  $\tilde{I}_i^k(t_1^k)$  for each noisy channel type. We will also consider the choices of the signaling levels  $a$  and  $b$  in (15) that maximize  $\tilde{I}_i^k(t_1^k)$ , i.e., asymptotically minimize  $E_i[\tilde{T}]$ .

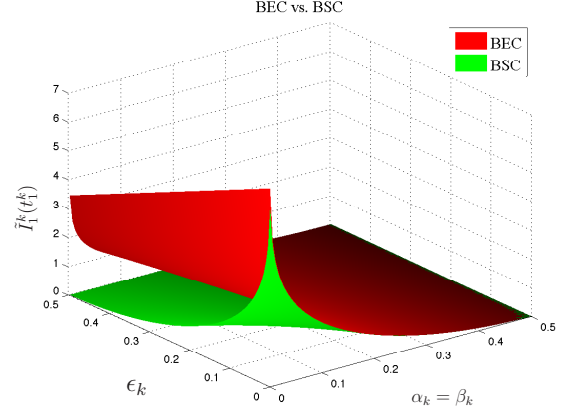


Fig. 2. The KL information,  $\tilde{I}_1^k(t_1^k)$ , under BEC and BSC, as a function of the local error probabilities  $\alpha_k = \beta_k$  and the channel error probability  $\epsilon_k$ .

### A. BEC

Under BEC, from (11) we can write the LLR of the received bits at the FC as

$$\tilde{\lambda}_n^k = \begin{cases} \hat{\lambda}_n^k, & \text{with probability } 1 - \epsilon_k, \\ 0, & \text{with probability } \epsilon_k. \end{cases} \quad (30)$$

Hence, we have

$$\tilde{I}_i^k(t_1^k) = E_i[\tilde{\lambda}_1^k] = (1 - \epsilon_k) \hat{I}_i^k(t_1^k) \quad (31)$$

where  $\hat{I}_i^k(t_1^k)$  is given in (27). As can be seen in (31) the performance degradation under BEC is only determined by the channel parameters  $\epsilon_k$ . In general, from (23), (29) and (31) this asymptotic performance loss can be quantified as  $\frac{1}{1 - \min_k \epsilon_k} \leq \frac{E_i[\tilde{T}]}{E_i[\tilde{T}]} \leq \frac{1}{1 - \max_k \epsilon_k}$ . Specifically, if  $\epsilon_k = \epsilon, \forall k$ , then we have  $\frac{E_i[\tilde{T}]}{E_i[\tilde{T}]} = \frac{1}{1 - \epsilon}$  as  $\alpha, \beta \rightarrow 0$ .

### B. BSC

Recall from (12) and (13) that under BSC local error probabilities  $\alpha_k, \beta_k$  undergo a linear transformation to yield the effective local error probabilities  $\hat{\alpha}_k, \hat{\beta}_k$  at the FC. Therefore, using (12) and (13), similar to (27),  $\tilde{I}_i^k(t_1^k)$  is written as follows

$$\tilde{I}_1^k(t_1^k) = (1 - \hat{\beta}_k) \log \frac{1 - \hat{\beta}_k}{\hat{\alpha}_k} + \hat{\beta}_k \log \frac{\hat{\beta}_k}{1 - \hat{\alpha}_k}, \quad (32)$$

$$\text{and } \tilde{I}_0^k(t_1^k) = \hat{\alpha}_k \log \frac{1 - \hat{\beta}_k}{\hat{\alpha}_k} + (1 - \hat{\alpha}_k) \log \frac{\hat{\beta}_k}{1 - \hat{\alpha}_k}$$

where  $\hat{\alpha}_k = (1 - 2\epsilon_k)\alpha_k + \epsilon_k$  and  $\hat{\beta}_k = (1 - 2\epsilon_k)\beta_k + \epsilon_k$ . Notice that the performance loss in this case also depends only on the channel parameter  $\epsilon_k$ .

In Fig. 2 we plot  $\tilde{I}_1^k(t_1^k)$  as a function of  $\alpha_k = \beta_k$  and  $\epsilon_k$ , for both BEC and BSC. It is seen that the KL information of BEC is higher than that of BSC, implying that the asymptotic average detection delay is lower for BEC, as anticipated in Section III-B.

$$\begin{aligned}
\tilde{I}_1^k(t_1^k) &= \bar{E}_1[\tilde{\lambda}_1^k] = (1 - \beta_k) \mathbb{E} \left[ \log \frac{(1 - \beta_k)e^{-u} + \beta_k e^{-v_a}}{\alpha_k e^{-u} + (1 - \alpha_k)e^{-v_a}} \right] + \beta_k \mathbb{E} \left[ \log \frac{(1 - \beta_k)e^{-v_b} + \beta_k e^{-u}}{\alpha_k e^{-v_b} + (1 - \alpha_k)e^{-u}} \right] \\
&= \underbrace{(1 - \beta_k) \log \frac{1 - \beta_k}{\alpha_k} + \beta_k \log \frac{\beta_k}{1 - \alpha_k}}_{\hat{I}_1^k(t_1^k)} + \beta_k \underbrace{\left( \frac{1 - \beta_k}{\beta_k} \mathbb{E} \left[ \log \frac{1 + \frac{\beta_k}{1 - \beta_k} e^{u - v_a}}{1 + \frac{1 - \alpha_k}{\alpha_k} e^{u - v_a}} \right] + \mathbb{E} \left[ \log \frac{1 + \frac{1 - \beta_k}{\beta_k} e^{u - v_b}}{1 + \frac{\alpha_k}{1 - \alpha_k} e^{u - v_b}} \right] \right)}_{\mathcal{C}_1^k} \quad (33)
\end{aligned}$$

### C. AWGN

In this and the following sections, we will drop the sensor index  $k$  of  $\sigma_{h,k}^2$  and  $\sigma_k^2$  for simplicity. In the AWGN case, it follows from Section III-C that if the transmitted signal is  $a$ , i.e.,  $x_n^k = a$ , then  $c_n^k = u, d_n^k = v_a$ ; and if  $x_n^k = b$ , then  $c_n^k = v_b, d_n^k = u$  where  $u \triangleq \frac{|w_n^k|^2}{\sigma^2}, v_a \triangleq \frac{|w_n^k + (a-b)h_k|^2}{\sigma^2}, v_b \triangleq \frac{|w_n^k + (b-a)h_k|^2}{\sigma^2}$ . Accordingly, from (16) we write the KL information as in (33), where  $\mathbb{E}[\cdot]$  denotes the expectation with respect to the channel noise  $w_n^k$  only, and  $\bar{E}_1[\cdot]$  denotes the expectation with respect to both  $x_n^k$  and  $w_n^k$  under  $H_1$ . Since  $w_n^k$  is independent of  $x_n^k$  under both  $H_0$  and  $H_1$ , we used the identity  $\bar{E}_1[\cdot] = \mathbb{E}[E_1[\cdot]]$  in (33).

Note from (33) that we have  $\tilde{I}_1^k(t_1^k) = \hat{I}_1^k(t_1^k) + \beta_k \mathcal{C}_1^k$  and  $\tilde{I}_0^k(t_1^k) = \hat{I}_0^k(t_1^k) + \alpha_k \mathcal{C}_0^k$ . Similar to  $\mathcal{C}_1^k$  we have  $\mathcal{C}_0^k \triangleq -\mathcal{E}_1 - \frac{1 - \alpha_k}{\alpha_k} \mathcal{E}_2$ . Since we know  $\tilde{I}_i^k(t_1^k) \leq \hat{I}_i^k(t_1^k)$ , the extra terms,  $\mathcal{C}_1^k, \mathcal{C}_0^k \leq 0$  are penalty terms that correspond to the information loss due to the channel noise. Our focus will be on this term as we want to optimize the performance under AWGN channels by choosing the transmission signal levels  $a$  and  $b$  that maximize  $\mathcal{C}_i^k$ .

Let us first consider the random variables  $\zeta_a \triangleq u - v_a$  and  $\zeta_b \triangleq u - v_b$  which are the arguments of the exponential functions in  $\mathcal{E}_1$  and  $\mathcal{E}_2$  in (33). From the definitions of  $u$  and  $v_a$ , we write  $\zeta_a = \frac{|w_n^k|^2}{\sigma^2} - \frac{|w_n^k + (a-b)h_k|^2}{\sigma^2} = -\frac{|a-b|^2|h_k|^2}{\sigma^2} - \frac{2}{\sigma^2}\gamma$  where  $\gamma \triangleq \Re\{(w_n^k)^*(a-b)h_k\}$  and  $\Re\{\cdot\}$  denotes the real part of a complex number. Similarly we have  $\zeta_b = -\frac{|a-b|^2|h_k|^2}{\sigma^2} + \frac{2}{\sigma^2}\gamma$ . Note that  $\gamma \sim \mathcal{N}(0, \frac{|a-b|^2|h_k|^2\sigma^2}{2})$  since  $w_n^k \sim \mathcal{N}_c(0, \sigma^2)$ . If we define  $\nu \triangleq \frac{\sqrt{2}}{|a-b|h_k|\sigma}\gamma$ , then we have  $\nu \sim \mathcal{N}(0, 1)$ . Upon defining  $s \triangleq \frac{|a-b|h_k|}{\sigma}$  we can then write  $\zeta_a$  and  $\zeta_b$  as

$$\zeta_a = -s^2 - \sqrt{2} s\nu \quad \text{and} \quad \zeta_b = -s^2 + \sqrt{2} s\nu.$$

If we define  $F \triangleq \frac{1 - \alpha_k}{\alpha_k}$  and  $G \triangleq \frac{1 - \beta_k}{\beta_k}$ , then we have

$$\begin{aligned}
\mathcal{C}_0^k &= \mathbb{E} \left[ \log \frac{1 + F^{-1}e^{\zeta_b}}{1 + Ge^{\zeta_b}} \right] + F^{-1} \mathbb{E} \left[ \log \frac{1 + Fe^{\zeta_a}}{1 + G^{-1}e^{\zeta_a}} \right] \\
\mathcal{C}_1^k &= \mathbb{E} \left[ \log \frac{1 + Ge^{\zeta_b}}{1 + F^{-1}e^{\zeta_b}} \right] + GE \left[ \log \frac{1 + G^{-1}e^{\zeta_a}}{1 + Fe^{\zeta_a}} \right]. \quad (34)
\end{aligned}$$

Note from (14) that the received signal,  $z_n^k$ , will have the same variance, but different means,  $ah_k$  and  $bh_k$ , if  $x_n^k = a$  and  $x_n^k = b$  are transmitted respectively. Hence, we expect that the detection performance under AWGN channels will improve if the difference between the transmission levels,  $|a - b|$ , increases. Toward that end the following result gives a

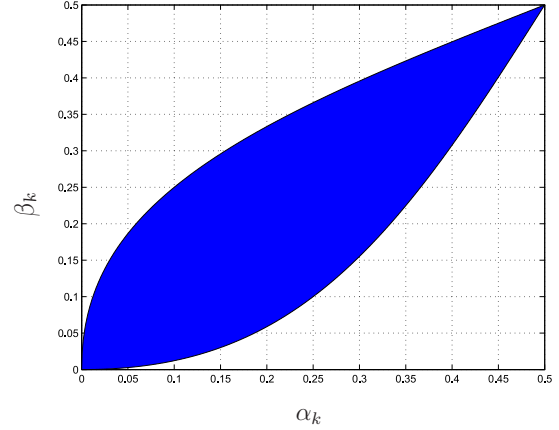


Fig. 3. The region of  $(\alpha_k, \beta_k)$  specified by Lemma 2.

sufficient condition under which the penalty term  $\mathcal{C}_i^k$  increases with  $s$ , and hence with  $|a - b|$ . The proof is given in [13, Appendix].

**Lemma 2.**  $\mathcal{C}_i^k$  is an increasing function of  $s$ ,  $i = 0, 1$ , if  $F^2 \geq G$  and  $G^2 \geq F$ .

Lemma 2 indicates that for  $\alpha_k, \beta_k$  values inside the region shown in Fig. 3,  $\mathcal{C}_i^k$  is increasing in  $|a - b|$ . Note that  $\alpha_k, \beta_k$  are local error probabilities which are directly related to the local threshold  $\Delta$ . Therefore, even if the hypotheses  $H_0$  and  $H_1$  are non-symmetric, we can ensure that we will have  $\alpha_k, \beta_k$  inside the region in Fig. 3 by employing different local thresholds,  $-\Delta_k$  and  $\Delta_k$ , in (2). In fact, even for  $\alpha_k, \beta_k$  values outside the region in Fig. 3 numerical results show that  $\mathcal{C}_i^k$  is increasing in  $s$ .

Hence, maximizing  $\mathcal{C}_i^k$  is equivalent to maximizing  $|a - b|$ . If we consider a constraint on the maximum allowed transmission power at sensors, i.e.,  $\max(|a|^2, |b|^2) \leq P^2$ , then the antipodal signaling is optimum, i.e.,  $|a| = |b| = P$  and  $a = -b$ .

### D. Rayleigh Fading

It follows from Section III-D that  $c_n^k = u_a, d_n^k = \frac{\sigma_a^2}{\sigma_b^2} u_a$  when  $x_n^k = a$ ; and  $c_n^k = \frac{\sigma_b^2}{\sigma_a^2} u_b, d_n^k = u_b$  when  $x_n^k = b$  where  $u_a \triangleq \frac{|ah_n^k + w_n^k|^2}{\sigma_a^2}, u_b \triangleq \frac{|bh_n^k + w_n^k|^2}{\sigma_b^2}$ , and  $\sigma_a^2 = |a|^2\sigma_h^2 + \sigma^2, \sigma_b^2 = |b|^2\sigma_h^2 + \sigma^2$  as defined in Section III-D. Define



$$\begin{aligned}
\tilde{I}_1^k(t_1^k) &= (1 - \beta_k) \mathbb{E} \left[ \log \frac{\frac{1-\beta_k}{\sigma_a^2} e^{-u_a} + \frac{\beta_k}{\sigma_b^2} e^{-\rho u_a}}{\frac{\alpha_k}{\sigma_a^2} e^{-u_a} + \frac{1-\alpha_k}{\sigma_b^2} e^{-\rho u_a}} \right] + \beta_k \mathbb{E} \left[ \log \frac{\frac{1-\beta_k}{\sigma_a^2} e^{-\rho^{-1} u_b} + \frac{\beta_k}{\sigma_b^2} e^{-u_b}}{\frac{\alpha_k}{\sigma_a^2} e^{-\rho^{-1} u_b} + \frac{1-\alpha_k}{\sigma_b^2} e^{-u_b}} \right] \\
&= \underbrace{(1 - \beta_k) \log \frac{1 - \beta_k}{\alpha_k} + \beta_k \log \frac{\beta_k}{1 - \alpha_k}}_{\tilde{I}_1^k(t_1^k)} + \beta_k \underbrace{\left( \mathbb{E} \left[ \log \frac{1 + G\rho^{-1} e^{\zeta_b}}{1 + F^{-1}\rho^{-1} e^{\zeta_b}} \right] + G\mathbb{E} \left[ \log \frac{1 + G^{-1}\rho e^{\zeta_a}}{1 + F\rho e^{\zeta_a}} \right] \right)}_{C_1^k} \quad (35)
\end{aligned}$$

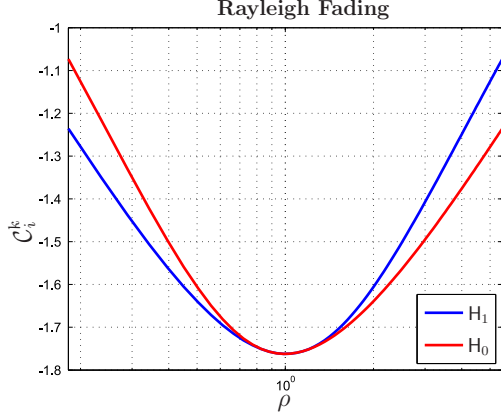


Fig. 4. The penalty term  $C_i^k$  for Rayleigh fading channels as a function of  $\rho$ .

further  $\rho \triangleq \frac{\sigma_a^2}{\sigma_b^2}$ . Hence, using (17) we write the KL information as in (35), where  $\zeta_a \triangleq u_a(1 - \rho)$  and  $\zeta_b \triangleq u_b(1 - \rho^{-1})$ . Note that when  $|a| = |b|$  which corresponds to the optimal signaling in the AWGN case, we have  $\rho = 1$ ,  $\zeta_a = \zeta_b = 0$  and therefore  $\tilde{I}_1^k(t_1^k) = 0$  in (35). This result is quite intuitive since in the Rayleigh fading case the received signals differ only in their variances. Note that  $u_a$  and  $u_b$  are chi-squared random variables with 2 degrees of freedom, i.e.,  $u_a, u_b \sim \chi_2^2$ , thus we can write the penalty term  $C_i^k$  as

$$C_0^k = \int_0^\infty \left( \log \frac{1 + F^{-1}\rho^{-1} e^{u(1-\rho^{-1})}}{1 + G\rho^{-1} e^{u(1-\rho^{-1})}} + F^{-1} \log \frac{1 + F\rho e^{u(1-\rho)}}{1 + G^{-1}\rho e^{u(1-\rho)}} \right) \frac{e^{-u}}{2} du, \quad (36)$$

$$\text{and } C_1^k = \int_0^\infty \left( \log \frac{1 + G\rho^{-1} e^{u(1-\rho^{-1})}}{1 + F^{-1}\rho^{-1} e^{u(1-\rho^{-1})}} + G \log \frac{1 + G^{-1}\rho e^{u(1-\rho)}}{1 + F\rho e^{u(1-\rho)}} \right) \frac{e^{-u}}{2} du.$$

Note that given local error probabilities  $\alpha_k, \beta_k$  the integrals in (36) is a function of  $\rho$  only. However, maximizing  $C_i^k$  in (36) with respect to  $\rho$  seems analytically intractable. As can be seen in Section III-D, the received signals at the FC will have zero mean and the variances  $\sigma_a^2$  and  $\sigma_b^2$  when  $x_n^k = a$  and  $x_n^k = b$  respectively. Therefore, in this case intuitively we should increase the difference between the two variances, i.e.,  $||a|^2 - |b|^2|$ . Consider the following

constraints:  $\max(|a|^2, |b|^2) \leq P^2$  and  $\min(|a|^2, |b|^2) \geq Q^2$ , where the first one is the peak power constraint as before, and the second is to ensure reliable detection of an incoming signal by the FC. We conjecture that the optimum signaling scheme in this case that maximizes  $C_i^k$  corresponds to  $|a| = P, |b| = Q$  or  $|a| = Q, |b| = P$ .

To numerically illustrate the behavior of  $C_i^k$  as a function of  $\rho$ , we set  $\alpha_k = \beta_k = 0.1$ ,  $\sigma_h^2 = \sigma^2 = 1$ ,  $P^2 = 10$ ,  $Q^2 = 1$ , and plot  $C_i^k$  in Fig. 4. It is seen that  $C_i^k$  has its global minimum when  $\rho = 1$ , which corresponds to the case  $|a| = |b|$  as expected. Moreover,  $C_i^k$ , validating our conjecture, monotonically grows as  $\rho$  tends to its minimum and maximum values corresponding to the cases  $|a| = Q, |b| = P$  and  $|a| = P, |b| = Q$  respectively.

Note that in Fig. 4, the curves for  $H_0$  and  $H_1$  are mirrored versions of each other around  $\rho = 1$  since we have  $\alpha_k = \beta_k$  in the example. From (36) we can say that the symmetry between  $H_0$  and  $H_1$  around  $\rho = 1$  will exist whenever  $F = G$ , i.e.,  $\alpha_k = \beta_k$ .

### E. Rician Fading

In the Rician fading case, upon defining  $\tilde{h}_n^k \triangleq h_n^k - \mu_k$  from Section III-E we have  $c_n^k = \frac{|a\tilde{h}_n^k + w_n^k|^2}{\sigma_a^2}$ ,  $d_n^k = \frac{|a\tilde{h}_n^k + w_n^k + (a-b)\mu_k|^2}{\sigma_b^2}$  when  $x_n^k = a$ ; and  $c_n^k = \frac{|b\tilde{h}_n^k + w_n^k + (b-a)\mu_k|^2}{\sigma_a^2}$ ,  $d_n^k = \frac{|b\tilde{h}_n^k + w_n^k|^2}{\sigma_b^2}$  when  $x_n^k = b$ . We will drop the subscript  $k$  in  $\mu_k$  for convenience. We further define  $\tilde{z}_a \triangleq a\tilde{h}_n^k + w_n^k$  and  $\tilde{z}_b \triangleq b\tilde{h}_n^k + w_n^k$  that are zero-mean Gaussian variables with variances  $\sigma_a^2$  and  $\sigma_b^2$ , respectively. Then from Section III-E similar to (35) we write the KL information as

$$\begin{aligned}
\tilde{I}_1^k(t_1^k) &= \tilde{I}_1^k(t_1^k) + \\
&\beta_k \underbrace{\left( \mathbb{E} \left[ \log \frac{1 + G\rho^{-1} e^{\zeta_b}}{1 + F^{-1}\rho^{-1} e^{\zeta_b}} \right] + G\mathbb{E} \left[ \log \frac{1 + G^{-1}\rho e^{\zeta_a}}{1 + F\rho e^{\zeta_a}} \right] \right)}_{C_1^k} \quad (37)
\end{aligned}$$

where  $\zeta_a \triangleq -\left( \frac{|\tilde{z}_a + (a-b)\mu|^2}{\sigma_b^2} - \frac{|\tilde{z}_a|^2}{\sigma_a^2} \right)$  and  $\zeta_b \triangleq -\left( \frac{|\tilde{z}_b + (b-a)\mu|^2}{\sigma_a^2} - \frac{|\tilde{z}_b|^2}{\sigma_b^2} \right)$ . Now we will analyze the exponents  $\zeta_a$  and  $\zeta_b$ .



$$\begin{aligned}
\mathcal{C}_1^k &= \underbrace{\int_0^\infty \log \left[ \frac{1 + G\rho^{-1}e^{u(1-\rho^{-1}) + \frac{|b-a|^2|\mu|^2}{(|b|^2-|a|^2)\sigma_h^2}}}{1 + F^{-1}\rho^{-1}e^{u(1-\rho^{-1}) + \frac{|b-a|^2|\mu|^2}{(|b|^2-|a|^2)\sigma_h^2}}} \right]}_{\mathcal{I}_1} \frac{e^{-\frac{u+\lambda_b}{2}}}{2} I_0(\sqrt{\lambda_b}u) du + \\
&G \underbrace{\int_0^\infty \log \left[ \frac{1 + G^{-1}\rho e^{u(1-\rho) + \frac{|a-b|^2|\mu|^2}{(|a|^2-|b|^2)\sigma_h^2}}}{1 + F\rho e^{u(1-\rho) + \frac{|a-b|^2|\mu|^2}{(|a|^2-|b|^2)\sigma_h^2}}} \right]}_{\mathcal{I}_2} \frac{e^{-\frac{u+\lambda_a}{2}}}{2} I_0(\sqrt{\lambda_a}u) du
\end{aligned} \tag{38}$$

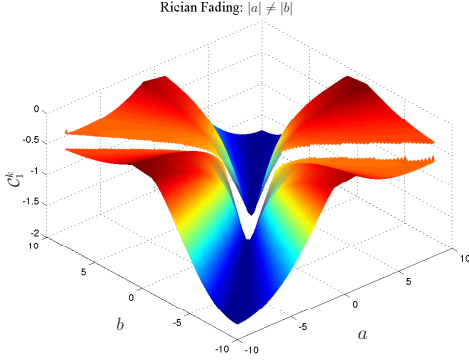


Fig. 5. The penalty term under  $\mathcal{H}_1$ , i.e.,  $\mathcal{C}_1^k$ , for Rician channels with  $|a| \neq |b|$ , as a function of the transmission levels  $a$  and  $b$ .

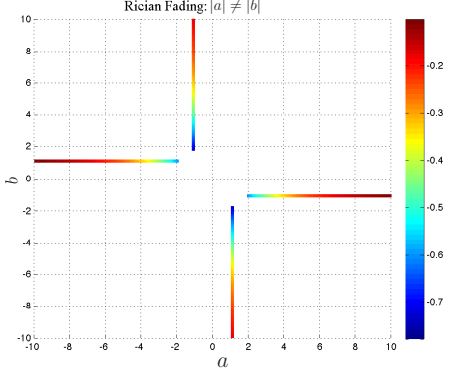


Fig. 6. The maximum contour of Fig. 5 is shown to exhibit the locus of the optimum signaling levels.  $\mathcal{C}_1^k$  is color coded according to the color bar given next to the figure.

1) *Case 1:  $|a| \neq |b|$ :* For  $\kappa \triangleq \frac{1}{\sigma_b^2} - \frac{1}{\sigma_a^2} > 0$ , i.e.,  $|a| > |b|$ , we can write  $\zeta_a$  as

$$\zeta_a = - \left[ \kappa |\tilde{z}_a|^2 + \frac{2\Re\{\tilde{z}_a^*(a-b)\mu\}}{\sigma_b^2} + \frac{|a-b|^2|\mu|^2}{\sigma_b^2} \right] \tag{39}$$

$$= - \left( \sigma_a^2 \kappa \underbrace{\left[ \frac{\tilde{z}_a}{\sigma_a} + \frac{(a-b)\mu}{\sigma_b^2 \sigma_a \kappa} \right]^2}_{\triangleq u_a} + \frac{|a-b|^2|\mu|^2}{(|b|^2-|a|^2)\sigma_h^2} \right) \tag{40}$$

$$= u_a(1-\rho) + \frac{|a-b|^2|\mu|^2}{(|a|^2-|b|^2)\sigma_h^2} \tag{41}$$

where we used  $\sigma_b^2\kappa - 1 = -\rho^{-1}$ ,  $\sigma_b^4\kappa = \rho^{-1}(\sigma_a^2 - \sigma_b^2)$  while writing (40), and  $\sigma_a^2\kappa = \rho - 1$  while writing (41). Note that  $u_a$  is a noncentral chi-squared random variable with two degrees of freedom and the noncentrality parameter  $\lambda_a \triangleq \frac{|a-b|^2|\mu|^2\sigma_a^2}{(|a|^2-|b|^2)^2\sigma_h^4}$ . Using  $\sqrt{-\kappa}$  instead of  $\sqrt{\kappa}$  it can be easily shown that (41) holds for  $\kappa < 0$ . Similarly one can obtain

$$\zeta_b = u_b(1-\rho^{-1}) + \frac{|b-a|^2|\mu|^2}{(|b|^2-|a|^2)\sigma_h^2}$$

for  $\kappa \neq 0$ , i.e.,  $|a| \neq |b|$ , where  $u_b \triangleq \left| \frac{\tilde{z}_b}{\sigma_b} + \frac{(a-b)\mu}{\sigma_a^2\sigma_b\kappa} \right|^2$  and  $u_b \sim \chi^2_2(\lambda_b)$  with  $\lambda_b \triangleq \frac{|a-b|^2|\mu|^2\sigma_b^2}{(|a|^2-|b|^2)^2\sigma_h^4}$ . Accordingly, for the non-symmetric case where  $|a| \neq |b|$  from (37) we can write  $\mathcal{C}_1^k$  as in (38). Similarly, we have  $\mathcal{C}_0^k = -\mathcal{I}_1 - F^{-1}\mathcal{I}_2$ .

The expression in (38) resembles the one in (36) for the Rayleigh fading case. And maximizing (38) analytically with respect to  $a$  and  $b$  seems even more intractable. Recall that in the Rayleigh fading case, the optimum signaling scheme was an OOK-like non-symmetric constellation, i.e.,  $|a| = P$ ,  $|b| = Q$  or  $|a| = Q$ ,  $|b| = P$ . Considering the same power constraints we conjecture that the same signaling scheme, that maximizes the difference between the variances  $\sigma_a^2$  and  $\sigma_b^2$ , is optimum in this non-symmetric case.

We provide a numerical example to illustrate the behavior of  $\mathcal{C}_i^k$  as a function of  $a$  and  $b$ . Using the same values for  $\alpha_k, \beta_k, \sigma_h^2, \sigma^2, P^2, Q^2$  as in the Rayleigh fading case, and setting  $\mu = 1 + j$  we plot  $\mathcal{C}_1^k$  in Fig. 5. The maximum contour of the three-dimensional surface in Fig. 5, which corresponds to the potential optimum signaling level pairs, is clearly shown in Fig. 6. As seen in the figure  $\mathcal{C}_1^k$  is maximized when  $|a| = P$ ,  $|b| = Q$  or  $|a| = Q$ ,  $|b| = P$  validating our conjecture.

2) *Case 2:  $|a| = |b|$ :* For  $\kappa = 0$ , we have  $\sigma_a^2 = \sigma_b^2$ , i.e.,  $|a| = |b|$ . Accordingly, from (39) we write  $\zeta_a = -s^2 - \frac{2}{\sigma_a^2}\gamma_a$ ,  $\zeta_b = -s^2 - \frac{2}{\sigma_a^2}\gamma_b$  where similar to Section V-C we define  $s \triangleq \frac{|a-b|\mu|}{\sigma_a}$ ,  $\gamma_a \triangleq \Re\{\tilde{z}_a^*(a-b)\mu\}$  and  $\gamma_b \triangleq \Re\{\tilde{z}_b^*(b-a)\mu\}$ . Defining standard Gaussian random variables  $\nu_a \triangleq \frac{\sqrt{2}}{|a-b|\mu|\sigma_a}\gamma_a$  and  $\nu_b \triangleq \frac{\sqrt{2}}{|a-b|\mu|\sigma_a}\gamma_b$ , analogous to the AWGN case, we have  $\zeta_a = -s^2 - \sqrt{2}s\nu_a$  and  $\zeta_b = -s^2 - \sqrt{2}s\nu_b$ . Therefore, from (37)  $\mathcal{C}_i^k$  is given by (34). Accordingly, Lemma 2 applies here in the case of  $|a| = |b|$  under Rician channels. This case is analogous to the AWGN case since the received signal  $z_n^k$

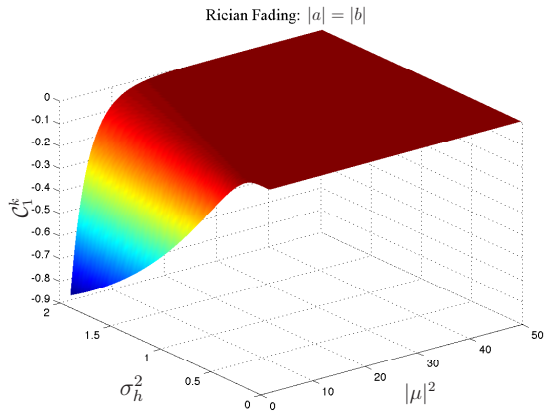


Fig. 7. The penalty term  $C_1^k$  in Rician fading channels with  $|a| = |b|$ , as a function of the mean and the variance of the channel gain.  $a = P = 10$  and  $b = -P = -10$ .

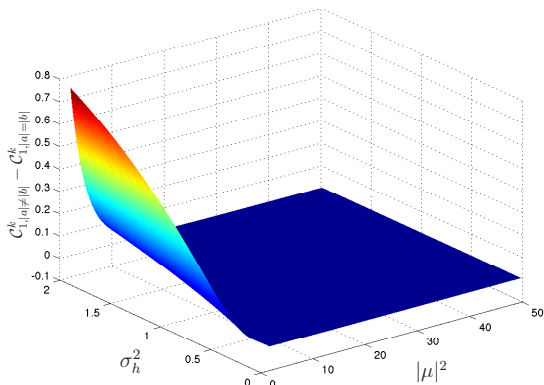


Fig. 8.  $C_{1,|a|\neq|b|}^k - C_{1,|a|=|b|}^k$  in Rician fading channels as a function of  $|\mu|^2$  and  $\sigma_h^2$ ,  $P = 10$ ,  $Q = 1$ .

has the same variance, but different means when  $x_n^k = a$  and  $x_n^k = b$ . Consequently, antipodal signaling is optimal. In Fig. 7,  $C_1^k$  is plotted as a function of the channel gain parameters  $|\mu|^2$  and  $\sigma_h^2$ . It is seen that  $C_1^k$  is increasing in  $|\mu|^2$  and decreasing in  $\sigma_h^2$  when antipodal signaling is used, which corroborates Lemma 2 since  $s$  is increasing in  $|\mu|^2$  and decreasing in  $\sigma_h^2$ .

In Fig. 8, the difference  $C_{1,|a|\neq|b|}^k - C_{1,|a|=|b|}^k$  is plotted as a function of  $|\mu|^2$  and  $\sigma_h^2$ . For  $|a| = |b|$  antipodal signaling is employed; and for  $|a| \neq |b|$ , OOK-like signaling is employed. It is seen that the OOK-like signaling is much better than antipodal signaling when the mean is low and the variance is high. Although not visible in Fig. 8, antipodal signaling is only slightly better than OOK-like signaling when the mean is high and the variance is low.

## VI. CONCLUSIONS

We have developed and analyzed channel-aware distributed detection schemes based on level-triggered sampling. The sensors form local log-likelihood ratios (LLRs) based on their

observations and sample their LLRs using the level-triggered sampling. Upon sampling each sensor sends a single bit to the fusion center (FC). The FC is equipped with the local error rates of all sensors and the statistics of the channels from all sensors. Upon receiving the bits from the sensors, the FC updates the global LLR and performs an SPRT. The fusion rules under different channel types are given. We have further provided non-asymptotic and asymptotic analyses on the average detection delay for the proposed channel-aware scheme. We have shown that the asymptotic detection delay is characterized by a KL information number, whose expressions under different channel types have been derived. Based on the delay analysis, we have also identified appropriate signaling schemes under different channels for the sensors to transmit the 1-bit information.

## REFERENCES

- [1] R.R. Tenney, and N.R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aero. Electron. Syst.*, vol. 17, no. 4, pp. 501-510, July 1981.
- [2] Z. Chair, and P.K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aero. Electron. Syst.*, vol. 22, no. 1, pp. 98-101, Jan. 1986.
- [3] A.M. Hussain, "Multisensor distributed sequential detection," *IEEE Trans. Aero. Electron. Syst.*, vol. 30, no. 3, pp. 698-708, July 1994.
- [4] G. Fellouris, and G.V. Moustakides, "Decentralized sequential hypothesis testing using asynchronous communication," *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 534-548, Jan. 2011.
- [5] Y. Yilmaz, G.V. Moustakides, and X. Wang, "Cooperative sequential spectrum sensing based on level-triggered sampling," *IEEE Trans. Sig. Proc.*, 2012, to be published.
- [6] V. Poor, "An Introduction to Signal Detection and Estimation," Springer, New York, NY, 1994.
- [7] S. Chaudhari, J. Lunden, V. Koivunen, and H.V. Poor, "Cooperative sensing with imperfect reporting channels: Hard decisions or soft decisions?," *IEEE Trans. Sig. Proc.*, vol. 60, no. 1, pp. 18-28, Jan. 2012.
- [8] B. Chen, L. Tong and P.K. Varshney, "Channel-Aware Distributed Detection in Wireless Sensor Networks," *IEEE Signal Processing Magazine*, vol. 23, issue 4, pp. 16-26, July 2006.
- [9] J.-F. Chamberland, and V.V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Sig. Proc.*, vol. 51, no. 2, pp. 407-416, Feb. 2003.
- [10] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3349-3358, July 2006.
- [11] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Stat.*, vol. 19, pp. 326-329, 1948.
- [12] Z. Govindarajulu, "Sequential Statistics," World Scientific Publishing, Hackensack, NJ, 2004.
- [13] Y. Yilmaz, G.V. Moustakides, and X. Wang, "Channel-aware Decentralized Detection via Level-triggered Sampling," <http://arxiv.org/abs/1205.5906>, May. 2012.
- [14] S. Ross, *Stochastic Processes*, Wiley, New York, NY, 1996.