Unstructured sequential testing in sensor networks

Georgios Fellouris and Alexander Tartakovsky

Abstract—The problem of sequentially detecting a signal in a sensor network is considered, when the set of sensors in which signal may be present is not known in advance. We formulate this problem as a sequential hypothesis testing problem with a simple null hypothesis (signal is absent in all sensors) and a composite alternative (signal is present in some sensors). We propose a sequential test that controls explicitly the type-I and type-II error probabilities and minimizes asymptotically the expected sample size under the alternative hypothesis for any possible set of affected sensors. Moreover, we propose modifications of this procedure that require minimal transmission activity from the sensors to the fusion center and, at the same time, preserve this asymptotic optimality property.

I. INTRODUCTION

Consider $K$ sources of observations (sensors) which transmit their data to a global decision maker (fusion center), as it is suggested by Figure 1. Each sensor $k$ observes a sequence $(X^k_t)_{t \in \mathbb{N}}$ of independent and identically distributed (i.i.d.) random variables with common density $f^k$ with respect to a $\sigma$-finite measure $\nu(dx)$, $1 \leq k \leq K$. There are two possibilities for $f^k$, $f^0_k$ and $f_1^k$, which are assumed to have a common support, and we set:

$$\Lambda^k_t := \prod_{s=1}^{t} \frac{f_1^k(X^k_s)}{f_0^k(X^k_s)},$$

$$Z^k_t = \log \Lambda^k_t = \sum_{s=1}^{t} \log \frac{f_1^k(X^k_s)}{f_0^k(X^k_s)},$$

i.e., $\Lambda^k_t$ (resp. $Z^k_t$) is the likelihood (resp. log-likelihood) ratio at sensor $k$ at time $t$. The goal at the fusion center is to distinguish between the following two hypotheses:

$$H_0 : f^k = f_0^k, \quad 1 \leq k \leq K$$

$$H_1 : f^k = f_1^k, \quad k \notin \mathcal{A} \quad \text{and} \quad f^k = f_1^k, \quad k \in \mathcal{A},$$

(1)

where $\mathcal{A} \subset \{1, \ldots, K\}$ is a subset of sensors. It is natural to think that the observations at sensor $k$ are distributed according to $f^k_f$ (resp. $f_0^k$) when signal is present (resp. absent) at sensor $k$. With this interpretation in mind, the null hypothesis $H_0$ represents the case that all sensors observe “noise”, whereas the alternative hypothesis $H_1$ represents the situation in which signal is present in some sensors.

Assuming that the observations at the sensors and the fusion center are acquired sequentially, the goal is to find a sequential test $(T, d_T)$ that consists of an $\mathcal{F}_T$-stopping time $T$ and an $\mathcal{F}_T$-measurable random variable $d_T$ that takes values in $\{0, 1\}$, so that $H_j$ is selected on $\{d_T = j\}, j = 0, 1$, where

$$\mathcal{F}_t := \sigma(X^k_s; 1 \leq s \leq t, 1 \leq k \leq K), \quad t \in \mathbb{N}.$$ 

An ideal sequential test should have small detection delay under both hypotheses, while controlling its error probabilities below prescribed levels. More specifically, if $C_{\alpha, \beta}$ is the class of sequential tests with probabilities of type-I and type-II error bounded below $\alpha$ and $\beta$, respectively, i.e.,

$$C_{\alpha, \beta} := \{(T, d_T) : P_0(d_T = 1) \leq \alpha \quad \text{and} \quad P_0(d_T = 0) \leq \beta\},$$

the goal is to find a sequential test $(S, d_S)$ that attains

$$\inf_{(T, d_T) \in C_{\alpha, \beta}} E_0[T] \quad \text{and} \quad \inf_{(T, d_T) \in C_{\alpha, \beta}} E_1[T],$$

(2)

where here, and in what follows, $P_j$ and $E_j$ represent probability and expectation, respectively, under $H_j, j = 0, 1$. Clearly, $P_1$ and $H_1$ depend on the set $\mathcal{A}$ and this is one of the main difficulties of this setup, since we will assume that the set $\mathcal{A}$ is not known in advance. However, we chose to suppress this dependence in order to lighten the notation.

Wald and Wolfowitz [1] proved that, for any $\alpha, \beta$ so that $\alpha + \beta < 1$, both infima are attained by Wald’s [2] Sequential Probability Ratio Test (SPRT):

$$S := \inf \{t : Z_t^A \notin (-A, B)\}, \quad d_S := 1_{\{Z_t^A \geq B\}}$$

where $A, B > 0$ are positive thresholds chosen so that $P_0(d_S = 1) = \alpha$ and $P_1(d_S = 0) = \beta$, whereas $Z_t^A$ is the log-likelihood ratio between the two hypotheses at time $t$. Assuming independence across sensors, which will be our
standing assumption from now on, we have:

\[ Z_t^A := \sum_{k \in A} Z_t^k. \]

Moreover, it is well-known that the asymptomatic perfor-
mance of the SPRT as \( \alpha, \beta \to 0 \) (to a first order) is:

\[
E_1[|S|] = \frac{\log \alpha}{I_1} \left( 1 + o(1) \right)
\]

\[
E_0[|S|] = \frac{\log \beta}{I_0} \left( 1 + o(1) \right),
\]

where \( I_1 := \sum_{k \in A} I_k^1, I_0 := \sum_{k \in A} I_k^0 \) and \( I_k^1, I_k^0 \) are the Kullback-Leibler information numbers that measure the “distance” between \( f_k^1 \) and \( f_k^0 \) and vice-versa, i.e.,

\[
I_k^1 := \int \log \left( \frac{f_k^1(x)}{f_k^0(x)} \right) f_0(x) \nu(dx),
\]

\[
I_k^0 := \int \log \left( \frac{f_k^0(x)}{f_k^1(x)} \right) f_0(x) \nu(dx).
\]

Based on these results, we will say that a sequential test \( (\hat{S}, d_s) \) is asymptotically optimal under \( P_j \) as \( \alpha, \beta \to 0 \), if

\[
E_j[\hat{S}] = \inf_{(T,d_T) \in C_{\alpha,\beta}} E_j[T] \to 1
\]

\[
J = 0, 1, \text{ no matter what the set } A \text{ is. Thus, } (\hat{S}, d_s) \text{ will be asymptotically optimal under } P_1 \text{ as } \alpha, \beta \to 0, \text{ if for any set } A \text{ we have}
\]

\[
E_1[\hat{S}] = \frac{\log \alpha}{I_1} \left( 1 + o(1) \right)
\]

(and similarly under \( P_0 \)).

Of course, the SPRT is not implementable when the set \( \mathcal{A} \) is not known. In the case that signal may be present at most one source, i.e., the cardinality of set \( \mathcal{A} \) is equal to 1 (\( |\mathcal{A}| = 1 \)), it has been suggested [4] to replace \( Z^A \) with either mixture-based statistics of the form \( \log(\sum_{k=1}^{K} p_k \Lambda_k) \), where \( p = (p_1, \ldots, p_K) \) is an arbitrary \( K \)-dimensional vector of positive numbers, or with weighted generalized likelihood ratio statistics (GLR) of the form \( \log(\max_{1 \leq k \leq K} p_k \Lambda_k) \). The latter class generalized the approach in [4], where the GLR statistic was considered, i.e., \( p_1 = \ldots = p_K = 1 \). In this context, where \( |\mathcal{A}| = 1 \), it was shown (see [4], [3]) that for any choice of \( p \), the resulting sequential tests attain asymptotically, as \( \alpha, \beta \to 0 \), both infima in (2) when \( \mathcal{A} = \{k\} \) for every \( 1 \leq k \leq K \), i.e., no matter what sensor is affected.

The first contribution of the present work is that we propose a sequential test that belongs in \( C_{\alpha,\beta} \) and attains asymptotically \( \inf_{(T,d_T) \in C_{\alpha,\beta}} E_1[T] \), i.e., the optimal expected sample size under the alternative hypothesis, for any possible subset \( \mathcal{A} \). More specifically, we assume that the number of sensors in which signal may be present is at least \( K \) and at most \( K \), i.e., \( K \leq |\mathcal{A}| \leq K \), known in advance. When, in particular, \( K = 1 \) and \( K = K \), this corresponds to a completely unstructured problem, i.e., there is absolutely no prior information regarding \( \mathcal{A} \). When \( K = K \), this corresponds to the problem considered in [4], [3].

The implementation of the SPRT in sensor networks is also prohibited by the fact that it is typically not possible for the sensors to transmit their complete observations to the fusion center, due to practical constraints. More specifically, in many applications that are characterized by limited communication bandwidth and energy constraints, as it is often the case with wireless sensor networks, it is very important to design schemes with small communication load (see, e.g., [5], [6]). This is achieved by having the sensors transmitting quantized versions of their observations at a much lower rate, if possible, than their local sampling rate. The sequential testing problem has been considered extensively under such constraints when signal is either absent or present in all sensors, i.e., \( \mathcal{A} = \{1, \ldots, K\} \) (see, e.g., [7]-[15]). In this context, asymptotically optimal sequential tests have been proposed in [13] and [14] (see also [15]).

The second goal of the present work is construct sequential tests that attain asymptotically \( \inf_{(T,d_T) \in C_{\alpha,\beta}} E_1[T] \) for any set \( \mathcal{A} \), but at the same time require infrequent transmission of (one-bit) messages to the fusion center. In order to achieve this, for each sensor \( k \), we sample the log-likelihood ratio process \( Z^k \) using level-triggered sampling, as in [14], but at two different sequences of sampling instances.

The remaining paper is organized as follows: in Section 2, we described the suggested approach and we show how to designed the proposed sequential test in order to control the probabilities of error and achieve the aforementioned uniform asymptotic optimality. In Section 3, we modify the proposed scheme to account for communication constraints. In Section 4, we conclude and we discuss potential extensions of this work in progress.

II. THE PROPOSED SEQUENTIAL TEST

Let \( \hat{Z}^k \) and \( \hat{Z}^k \) be arbitrary \( \{\mathcal{F}_t^k\} \)-adapted statistics so that for every \( t \in \mathbb{N} \):

\[
m_t^k := \inf_{0 \leq s \leq t} Z_s^k \leq \hat{Z}_t^k \leq 0 \leq \hat{Z}_t^k \leq M_t^k := \sup_{0 \leq s \leq t} Z_s^k. \tag{3}
\]

We set

\[
\hat{Z}_t := \max_{B : |B| = K} \hat{Z}_t^B, \quad \text{where } \hat{Z}_t^B := \sum_{k \in B} \hat{Z}_t^k,
\]

\[
\hat{Z}_t := \max_{B : |B| = K} \hat{Z}_t^B, \quad \text{where } \hat{Z}_t^B := \sum_{k \in B} \hat{Z}_t^k.
\]

Alternatively, the statistics \( \hat{Z} \) and \( \hat{Z} \) can be defined as follows:

\[
\hat{Z}_t = \sum_{k=1}^{K} \hat{Z}_t^k, \quad \hat{Z}_t = \sum_{k=1}^{K} \hat{Z}_t^k,
\]

where \( \hat{Z}_t^{(1)} \geq \ldots \geq \hat{Z}_t^{(K)} \) and \( \hat{Z}_t^{(1)} \geq \ldots \geq \hat{Z}_t^{(K)} \).
form:
\[ \hat{Z}_t = \sum_{k=1}^{K} \hat{Z}_t^k, \quad \hat{Z}_t = \max_{1 \leq k \leq K} \hat{Z}_t^{(k)}, \]

The proposed sequential test stops at
\[ \hat{T}_B \land \hat{T}_A := \min\{\hat{T}_B, \hat{T}_A\} \]
and selects \( H_1 \) (resp. \( H_0 \)) when \( \hat{T}_B \leq \hat{T}_A \) (resp. \( \hat{T}_A < \hat{T}_B \)), where
\[ \hat{T}_B := \inf\{t: \hat{Z}_t \geq B\}, \quad \hat{T}_A := \inf\{t: -\hat{Z}_t \geq A\} \]
and \( A, B \) are positive thresholds selected so that the resulting type-I and type-II error probabilities are upper bounded by \( \alpha \) and \( \beta \), respectively.

### A. Controlling the error probabilities

For any \( \alpha, \beta \in (0, 1) \), let \( B_\alpha \) and \( A_\beta \) be the unique roots of the equations \( F(x; K) = \alpha \) and \( F(x; K) = \beta \), respectively, i.e.,
\[ F(B_{\alpha}; K) = \alpha, \quad F(A_{\beta}; K) = \beta, \quad (4) \]
where for any \( m \in \mathbb{N} \) such that \( m \geq 2 \) we denote by \( F(x; m) \) the survival function of the Erlang distribution with parameters 1 and \( m \), i.e.,
\[ F(x; m) = e^{-x} \sum_{j=0}^{m-1} \frac{x^j}{j!}, \quad x > 0. \quad (5) \]

The following lemma describes the asymptotic relationship between \( B_\alpha \) and \( A_\beta \) as \( \alpha \to 0 \), as well as \( A_\beta \) and \( \beta \) as \( \beta \to 0 \).

**Lemma 2.1:** As \( \alpha, \beta \to 0 \),
\[ B_\alpha = |\log \alpha| + (K - 1) \log |\log \alpha| - \log (K - 1)! + o(1) \]
\[ A_\beta = |\log \beta| + (K - 1) \log |\log \beta| - \log (K - 1)! + o(1). \]

**Proof:** Taking logarithms in the definition of \( B_\alpha \) in (4) we obtain
\[ B_\alpha = |\log \alpha| + \log \left( \sum_{j=0}^{K-1} \frac{B_{\alpha}^j}{j!} \right) \]
\[ \quad = |\log \alpha| + (K - 1) \log B_\alpha - \log (K - 1)! + o(1) \]
as \( \alpha \to 0 \), where the second equality follows from the fact that
\[ \sum_{j=0}^{K-1} \frac{B_{\alpha}^j}{j!} = \frac{B_{\alpha}^{K-1}}{(K - 1)!} \left( 1 + o(1) \right). \]
The definition of \( B_\alpha \) also implies that \( B_\alpha = |\log \alpha|(1 + o(1)) \) and consequently \( \log B_\alpha = |\log \alpha| + o(1) \) as \( \alpha \to 0 \), which leads to the first relationship in (2.1). The second one can be shown in a similar manner. \( \square \)

**Theorem 2.1:** If \( A = A_\beta \) and \( B = B_\alpha \), then the proposed test belongs in \( C_{\alpha, \beta} \).

**Proof:** First of all, we observe that
\[ \hat{T}_B \geq \hat{R}_B := \inf\{t: \sum_{k=1}^{K} M_t^k \geq B\}, \]
since from (3) and the definition of \( \hat{Z} \) it follows that
\[ \hat{Z}_t \leq \sum_{k=1}^{K} \hat{Z}_t^k \leq \sum_{k=1}^{K} M_t^k. \]
Therefore, for any \( A, B > 0 \) we have
\[ P_0(\hat{T}_B < \hat{T}_A) \leq P_0(\hat{R}_B < \infty) \leq P_0(\hat{R}_B < t) \leq \lim_{t \to \infty} P_0(\hat{R}_B \leq t) \]
\[ = \lim_{t \to \infty} P_0(\sum_{k=1}^{K} M_t^k \geq B). \]
For every \( k \) and \( t \), the random variable \( M_t^k \) is stochastically dominated by an exponential random variable with rate 1. Indeed, setting \( T_B^k := \inf\{t: Z_t^k \geq B\} \) we have:
\[ P_0(M_t^k \geq B) = P_0(T_B^k \leq t) \leq P_0(T_B^k < \infty) = E_1^k \left[ e^{-Z_t^k B} \right] \leq e^{-B}, \]
where \( E_1^k \) is expectation under \( P_1^k \), i.e., the probability measure under which \( f^k = f_1^k \) (and \( f_j^j = f_0^j \), for \( j \neq k \)). Then, due to the assumed independence across sensors, \( M_1^1, \ldots, M_1^K \) are independent and \( \sum_{k=1}^{K} M_t^k \) is stochastically dominated by an Erlang random variable with parameters 1 and \( K \), i.e.,
\[ P_0 \left( \sum_{k=1}^{K} M_t^k \geq B \right) \leq F(B; K), \]
where \( F(x; K) \) is defined in (5). From the latter observation and (2.2) it follows that for any \( A > 0 \):
\[ P_0(\hat{T}_{B_\alpha} < \hat{T}_A) \leq F(B_\alpha; K) = \alpha. \]
From (3) and the definition of \( \hat{Z} \) it also follows that
\[ \hat{Z}_t = \max_{B: |B| = K} \hat{Z}_t^B \geq \max_{B \leq A: |B| = K} \hat{Z}_t^B \geq \max_{B \leq A: |B| = K} \hat{Z}_t^B \geq \sum_{k \in A} \hat{Z}_t^k \geq \sum_{k \in A} m_k^k, \]
and, as a result, we have:
\[ \hat{T}_A \geq \hat{R}_A := \inf\{t: -\sum_{k \in A} m_k^k \geq A\}. \]
Therefore, using the same reasoning as before, we can show that for any \( A, B > 0 \):
\[ P_1(\hat{T}_A \leq \hat{T}_B) \leq \lim_{t \to \infty} P_1 \left( \sum_{k \in A} -m_k^k \geq A \right), \]
and, consequently,
\[ P_1 \left( \sum_{k \in A} -m_k^k \geq A \right) \leq F(A; |A|) \leq F(A; K), \]
where the second inequality is correct because \( F(x; m) \) is increasing in \( m \) for any \( x \) and and \( |A| \leq K \). From the latter
observation and the previous relationships it follows that for any $B > 0$:
\[ P_1(\hat{T}_{A_\alpha} < \bar{T}_B) \leq F(A_\alpha; \mathcal{K}) = \beta, \]
which completes the proof.

B. Asymptotic optimality

In the next theorem we show that if $\hat{Z}_k$ and $\hat{Z}^k$ are chosen so that
\[ \hat{Z}_t^k = \max\{Z_t^k, 0\} \quad \text{and} \quad \hat{Z}_t = \min\{Z_t^k, 0\}, \quad (7) \]
a choice that clearly satisfies (3), then the proposed test attains $\inf_{(T, I, \tau) \in \mathcal{C}_{n, \alpha}} E[I(T)]$ asymptotically as $\alpha \to 0$ for any set $\mathcal{A}$. In order to prove this result, we will additionally need to assume that the increments of each random walk $Z^k$ have a finite second moment, i.e.,
\[ E_1[(Z_t^k)^2] < \infty, \quad 1 \leq k \leq K. \quad (8) \]

**Theorem 2.2:** Suppose that $\hat{Z}_k$ and $\hat{Z}^k$ are selected according to (7) and condition (8) holds. Then, as $\alpha \to 0$,
\[ E_1[\hat{T}_{B_n} \wedge \hat{T}_{A_\alpha}] \leq \frac{|\log \alpha| + (K - 1)|\log \alpha| + O(1)}{I_1}. \]

**Proof:** For any $A, B > 0$ we have:
\[ I_1 E_1[\hat{T}_B \wedge \hat{T}_A] \leq I_1 E_1[\hat{T}_B] = E_1[Z^k_{T_B}] \leq E_1[Z^k_{\mathcal{A}}] = E_1[\hat{T}_{B_n}]. \]
The first inequality holds because $\hat{T}_{B_n} \wedge \hat{T}_{A_\alpha} \leq \hat{T}_{B_n}$, whereas the equality is due to Wald’s identity. Moreover, since $\hat{Z}_k = \max\{Z_t^k, 0\}$, the second inequality holds because $\hat{Z}_k \geq Z^k$ for every $k$ and, consequently, $\hat{Z}^A \geq Z^A$, whereas the third inequality holds because $\hat{Z}_k \geq 0$ for every $k$, therefore $\hat{Z}^A \leq Z^B$ for every $A \subseteq B$ and, consequently, $\hat{Z}^A \leq \hat{Z}$.

Moreover, from the second moment assumption (8) we have as $B \to \infty$:
\[ E_1[\hat{Z}_{T_B}] = B + E_1[\hat{Z}_{\mathcal{T}_B} - B] = B + O(1). \quad (9) \]
Selecting now $B = B_\alpha$ and $A = A_\alpha$ and applying Lemma (2.1) we obtain the desired result.

III. COMMUNICATION CONSTRAINTS

Selecting the statistics $\hat{Z}_k$ and $\hat{Z}^k$ according to (7) requires that sensor $k$ transmits to the fusion center an infinite-bit message at any given time $t$. As we discussed in the Introduction, this may not be a realistic assumption in many applications that involve sensor networks, in which the sensors should ideally transmit, infrequently, low-bit messages to the fusion center. Our goal in this section is to suggest alternative specifications for $\tilde{Z}_k$ and $\tilde{Z}^k$ that satisfy (3) and preserve the asymptotic optimality property of Theorem 2.2, while at the same time require low transmission activity.

In order to achieve this, we assume that each sensor $k$ communicates with the fusion center at the following two sequences of stopping times:
\[ \tau^k_n := \inf\{t \geq \tau^k_{n-1} : Z^k_t - Z^k_{\tau^k_{n-1}} \geq \Delta^k\}, \quad n \in \mathbb{N}, \]
\[ \tau^k_n := \inf\{t \geq \tau^k_{n-1} : Z^k_t - Z^k_{\tau^k_{n-1}} \leq -\Delta^k\}, \quad n \in \mathbb{N}. \quad (10) \]

where $\tau^k_0 := 0 =: \tau^k_0$ and $\Delta^k$ and $\Delta^k$ are arbitrary positive thresholds, for any $1 \leq k \leq K$. It is clear that by selecting the thresholds $\Delta^k$ and $\Delta^k$ to be “large”, this communication scheme induces low-rate transmission from sensor $k$ to the fusion center. We also emphasize the difference with the communication scheme
\[ \tau^k_n := \inf\{t \geq \tau^k_{n-1} : Z^k_t - Z^k_{\tau^k_{n-1}} \notin (-\Delta^k, \Delta^k)\}, \quad n \in \mathbb{N} \]
which was considered in [3].

Moreover, we define the counting processes that count the number of transmitted messages, i.e.,
\[ N^k_t := \max\{n \in \mathbb{N} : \tau^k_n \leq t\}, \quad t \in \mathbb{N} \]
\[ N^k_t := \max\{n \in \mathbb{N} : \tau^k_n \leq t\}, \quad t \in \mathbb{N} \quad (11) \]
and we set
\[ \hat{\ell}^k_n := Z^k_{\tau^k_n} - Z^k_{\tau^k_{n-1}}, \quad n \in \mathbb{N} \]
\[ \hat{\ell}^k_n := Z^k_{\tau^k_n} - Z^k_{\tau^k_{n-1}}, \quad n \in \mathbb{N}, \quad (12) \]
i.e., $\hat{\ell}^k_n$ (resp. $\hat{\ell}^k_n$) is the realized log-likelihood ratio at sensor $k$ between $\tau^k_{n-1}$ and $\tau^k_n$ (resp. $\tau^k_{n-1}$ and $\tau^k_n$).

A. No quantization constraints

If at each time $\tau^k_n$ (resp. $\tau^k_n$), sensor $k$ can transmit the exact value of $\hat{\ell}^k_n$ (resp. $\hat{\ell}^k_n$) to the fusion center, a transmission that requires an infinite bit message, then we can define $\tilde{Z}^k$ (resp. $\tilde{Z}^k$) as the value of $Z^k$ at the last communication instance, i.e., $\tilde{Z}^k_k$ (resp. $\tilde{Z}^k_k$). In other words, we can select $\hat{Z}_k$ and $\hat{Z}^k$ as follows:
\[ \tilde{Z}_t^k = \sum_{n=1}^{N^k_t} \hat{\ell}^k_n \quad \text{and} \quad \tilde{Z}_t^k = \sum_{n=1}^{N^k_t} \hat{\ell}^k_n. \quad (13) \]

With this selection, it is clear that $\tilde{Z}_k$ starts from 0 and has piecewise constant and increasing paths. Moreover, since it coincides with $Z^k$ at the communication times $\tau^k_n$ and stays flat in between, it is clear that $\tilde{Z}_t^k$ cannot be larger than the maximum value of $Z^k$ up to time $t$. In a similar way we can argue that $\tilde{Z}_t^k$ takes non-positive values and is smaller than the minimum value of $Z^k$ up to time $t$. Therefore, selecting $\tilde{Z}_k$ and $\tilde{Z}^k$ according to (13) satisfies (3). In the following theorem, we will show that it also preserves the asymptotic optimality of Theorem 2.2.

**Theorem 3.1:** Suppose that $\tilde{Z}_k$ and $\tilde{Z}^k$ are selected according to (13) and (8) holds. Then, as $\alpha \to 0$, we have
\[ E_1[\tilde{T}_{B_n} \wedge \tilde{T}_{A_\alpha}] \leq \frac{|\log \alpha| + (K - 1)|\log \alpha| + O(\Delta_{\max})}{I_1}, \]
where $\Delta_{\max} := \max_{1 \leq k \leq K} \Delta^k$.

**Proof:** We observe that for any $A, B > 0$:
\[ I_1 E_1[\tilde{T}_B \wedge \tilde{T}_A] \leq I_1 E_1[\tilde{T}_B] = E_1[Z^k_{\tilde{T}_B}] = E_1[\tilde{Z}^k_{\tilde{T}_B}] + E_1[Z^A_{\tilde{T}_B}], \quad (14) \]
thus we need to find appropriate upper bounds for each of the two terms in the right-hand side. For the first term we have

\[ E_1[(Z^A - \hat{Z}^A)_{\hat{T}_{B,a}}] = \sum_{k \in A} E_1[(Z^k - \hat{Z}^k)_{\hat{T}_{B,a}}] \leq \sum_{k \in A} \Delta^k \leq |A| \Delta_{\text{max}} \leq K \Delta_{\text{max}}. \]

The first inequality follows by the definition of \( \hat{Z}^k \) in (13) and the communication scheme defined by (10), according to which \( Z^k \) cannot exceed \( \hat{Z}^k \) by more than \( \Delta^k \), i.e., \( Z^k \leq \hat{Z}^k + \Delta^k \), for any \( t, k \).

For the second term, we have

\[ E_1[\hat{Z}^A_{\hat{T}_{B,a}}] \leq E_1[\hat{Z}_{T_{B,a}}] = B + E_1[\hat{Z}_{T_{B,a}} - B] = B + O(1), \]

where \( O(1) \) is a bounded term as \( B, \Delta_{\text{max}} \to \infty \), due the second moment assumption (8). Selecting now \( B = B_\alpha \) and applying Lemma 2.1 we obtain the desired result.

**Corollary 3.1:** Under the conditions of Theorem 3.1, the proposed test is asymptotically optimal under \( P_1 \) as \( \alpha \to 0 \) for any fixed \( \{\bar{\Delta}^k, \Delta^k\}_{1 \leq k \leq K} \), as well as when \( \Delta^k \to \infty \), as long as \( \Delta_{\text{max}} = o(|\log \alpha|) \).

**B. Quantization constraints**

In the previous subsection we assumed that each sensor is able to transmit the exact value of \( \{\hat{t}^k_n\} \) (resp. \( \{\tilde{t}^k_n\} \)) at \( \hat{t}^k_n \) (resp. \( \tilde{t}^k_n \)). However, this may not be possible in applications with severe bandwidth constraints, where the sensors should transmit only low-bit messages to the fusion center. In order to take such quantization constraints into account, we set

\[ \hat{Z}^k_t := \hat{N}^k_t \Delta^k, \quad \tilde{Z}^k_t := \tilde{N}^k_t \Delta^k. \]  

(15)

In this way, sensor \( k \) simply needs to inform the fusion center regarding the evolution of the counting processes \( \hat{N}^k_t \) and \( \tilde{N}^k_t \), which can be done with the transmission of only one bit at the communication times \( (\hat{t}^k_n)_n \) and \( (\tilde{t}^k_n)_n \). In other words, in comparison to (13), sensor \( k \) no longer transmits to the fusion center the values of the overshoots:

\[ \hat{n}^k_n := \hat{t}^k_n - \Delta^k \quad \text{and} \quad \tilde{n}^k_n := - (\tilde{t}^k_n + \Delta^k) \]

(16)

that are associated with the transmissions by sensor \( k \) at \( \hat{t}^k_n \) and \( \tilde{t}^k_n \), respectively.

For notational simplicity, and without loss of generality, let us assume that \( \hat{N}^k_t = m \), i.e. \( t \in [\hat{t}^k_m, \hat{t}^k_{m+1}) \). Then:

\[ Z^k_t - \hat{Z}^k_t = Z^k_t - Z^k_{\hat{T}_{B,a}} + \sum_{n=1}^{m} \hat{n}^k_n \]

(17)

and since \( Z^k_t \) cannot exceed \( Z^k_{\hat{T}_{B,a}} \) by more than \( \Delta^k \), we obtain

\[ Z^k_t - \hat{Z}^k_t \leq \Delta^k + \sum_{n=1}^{m} \hat{n}^k_n \]

(18)

In order to see that selecting \( \hat{Z}^k_t \) and \( \tilde{Z}^k_t \) according to (15) satisfies (3), we now set \( t = \hat{t}^k_n \) in (17) and we obtain

\[ Z^k_{\hat{T}_{B,a}} - \hat{Z}^k_{\hat{T}_{B,a}} = \sum_{n=1}^{m} \hat{n}^k_n \geq 0, \]

which means that \( \hat{Z}^k_t \) cannot be larger than \( Z^k_t \) at any communication time \( \hat{t}^k_n \). But since \( \hat{Z}^k_{\hat{T}_{B,a}} \) remains constant between communication times, it is clear that \( \hat{Z}^k_t \) cannot be larger than the maximum value of \( Z^k_t \) up to time \( t \). This implies the last inequality in (3) and we can verify the validity of the first inequality in (3) in a similar way.

**Theorem 3.2:** Suppose that \( \hat{Z}^k_t \) and \( \tilde{Z}^k_t \) are selected according to (15) and condition (8) holds. Then, as \( \alpha \to 0 \),

\[ E_1[\hat{T}_{B,a} \land \hat{T}_{A}] \leq \frac{\log \alpha + (K-1) \log |\log \alpha|}{I_1} + O(\Delta_{\text{max}}) + \frac{O(|\log \alpha|)}{\Delta_{\text{min}}}, \]

where \( \Delta_{\text{min}} := \min_{1 \leq k \leq K} \bar{\Delta}^k \).

**Proof:** Working as in (14) we have for any \( A, B > 0 \):

\[ J_1 E_1[\hat{T}_{B} \land \hat{T}_{A}] \leq \sum_{k \in A} E_1[(Z^k_t - \hat{Z}^k_t)_{\hat{T}_{B,a}}] + E_1[\hat{Z}_{T_{B,a}}]. \]

From (18) we obtain

\[ E_1[(Z^k_t - \hat{Z}^k_t)_{\hat{T}_{B,a}}] \leq \Delta^k + \sum_{n=1}^{\hat{N}^{\hat{T}_{B,a}} + 1} \hat{n}^k_n \]

\[ = \Delta^k + E_1[\hat{N}^{\hat{T}_{B,a}} + 1] + E_1[\hat{n}^k_n], \]

where the equality follows from Wald’s identity, which can be applied since \( (\hat{n}^k_n)_{n \in N} \) is a sequence of \( \{\hat{t}^k_n\} \)-adapted, i.i.d. random variables and \( \hat{N}^{\hat{T}_{B,a}} + 1 \) is an \( \{\hat{t}^k_n\} \)-adapted stopping time, where \( \hat{t}^k_n := \hat{T}_{B,a} \hat{t}^k_n \) for every \( n \in N \).

If we now set \( C := \max_{1 \leq k \leq K} E[\hat{n}^k_n] \), we can write

\[ E_1[(Z^k_t - \hat{Z}^k_t)_{\hat{T}_{B,a}}] \leq \Delta^k + C + C E_1[\hat{N}^{\hat{T}_{B,a}}] \]

\[ = O(\Delta_{\text{max}}) + C E_1[\hat{N}^{\hat{T}_{B,a}}], \]

since, due to the second moment assumption (8), \( C \) is clearly an \( O(1) \) term as \( \Delta_{\text{max}} \to \infty \). Summing now over \( k \in A \) we obtain

\[ \sum_{k \in A} E_1[(Z^k_t - \hat{Z}^k_t)_{\hat{T}_{B,a}}] \leq O(\Delta_{\text{max}}) + C \sum_{k \in A} E_1[\hat{N}^{\hat{T}_{B,a}}] \]

\[ \leq O(\Delta_{\text{max}}) + C E_0[\hat{N}^{\hat{T}_{B,a}}] \]

\[ \leq O(\Delta_{\text{max}}) + C \frac{E_0[\hat{Z}_{T_{B,a}}]}{\Delta_{\text{min}}}, \]

where \( \hat{N}_t := \sum_{k=1}^{K} \hat{N}^k_t \) and the last inequality holds because \( \Delta_{\text{min}} \hat{N}_t \leq \hat{Z}_t \) for every \( t \geq 0 \), which follows from the definition of \( \hat{Z}_t \) in (15). Combining the previous relationships we have:

\[ J_1 E_1[\hat{T}_{B} \land \hat{T}_{A}] \leq O(\Delta_{\text{max}}) + \left(1 + \frac{C}{\Delta_{\text{min}}} \right) E_1[\hat{Z}_{T_{B,a}}]. \]
Since the overshoot of $\hat{Z}$ over the boundary $B$ cannot exceed $K\Delta_{\text{max}}$, we have $\hat{Z}_{T_B} \leq B + K\Delta_{\text{max}}$ and, consequently, we can write

$$I_1 E[I_B \wedge \hat{T}_A] = O(\Delta_{\text{max}}) + C \frac{B}{\Delta_{\text{min}}} + B.$$ 

Setting now $B = B_\alpha$ and applying Lemma 2.1 we obtain the desired result.

**Corollary 3.2:** Under the conditions of Theorem 3.2, the proposed test is asymptotically optimal under $P_1$ as $\alpha \to 0$, as long as $\Delta_{\text{min}} \to \infty$, i.e., $\Delta_k \to \infty$ for every $1 \leq k \leq K$, so that $\Delta_{\text{max}} = O(\log\alpha)$.

Note that when $\hat{Z}^k$ and $\hat{Z}^b$ are selected according to (15), instead of (13), it is now required that $\Delta_k \to \infty$ for every $1 \leq k \leq K$, i.e., $\Delta_{\text{min}} \to \infty$, for the proposed test to be asymptotically optimal. In other words, we can simply ignore the overshoots $(\hat{Z}_n)$, as long as the thresholds $\Delta^k$, $1 \leq k \leq K$ are sufficiently large. Thus, in this context, a very high communication rate will not be beneficial, due to the fast accumulation of unobserved overshoots, a phenomenon that was also observed in [14], where a similar communication scheme was employed.

**IV. CONCLUSIONS AND EXTENSIONS**

In this work, we proposed a class of sequential tests that are asymptotically optimal under the alternative hypothesis of problem 1, with and without communication constraints. By definition, the proposed sequential test is aggressive towards stopping under $H_1$ and conservative towards stopping under $H_0$. This is desirable in our context, since it is more important to detect quickly the presence of signal (and take any necessary action) than stopping to declare the absence of signal. However, a similarly aggressive approach under $H_0$, i.e., setting $\hat{Z} = \min_{0 \leq |A| \leq R} Z_i^A$ would make it possible for $|Z|$ to take large values under $P_1$, especially when $|A| \ll R$, leading to early stopping associated with falsely accepting the null hypothesis, destroying the ability of the test to detect a signal reliably.

We plan to extend the results presented in this work for more general statistical models. Indeed, it is easy to see that the proof of the key Theorem 2.1 does not make any use of the i.i.d. structure of the sensor observations. Such extensions can be carried out to a great degree of generality using the Kullback-Leibler modification of the criterion suggested in [18] (see also [19]).

Finally, we plan to compare the proposed approach numerically with other sequential tests, such as the SPRT bank, described in [4], Section IV.

**ACKNOWLEDGEMENTS**

This work was supported by the U.S. Air Force Office of Scientific Research under MURI grant FA9550-10-1-0569, by the U.S. Defense Threat Reduction Agency under grant HDTRA1-10-1-0086, by the U.S. Defense Advanced Research Projects Agency under grant W911NF-12-1-0034 and by the U.S. National Science Foundation under grants CCF-0830419, EFRI-1025043, and DMS-1221888 at the University of Southern California, Department of Mathematics.

**REFERENCES**