Universal Decentralized Detection in a Bandwidth-Constrained Sensor Network

Jin-Jun Xiao, *Student Member, IEEE,* and Zhi-Quan (Tom) Luo, *Senior Member, IEEE*

***Abstract*—Consider the problem of decentralized detection with a distributed sensor network where the communication channels between sensors and the fusion center are bandwidth constrained. Previous approaches to this problem typically rely on quantization of either the sensor observations or the local likelihood ratios, with quantization levels optimally designed using the knowledge of noise distribution. In this paper, we assume that each sensor is restricted to send a 1-bit message to the fusion center and that the sensor noises are additive, zero mean, and spatially independent but otherwise unknown and with possibly different distributions across sensors. We construct a universal decentralized detector using a recently proposed isotropic decentralized estimation scheme [10], [11] that requires only the knowledge of either the noise range or its second-order moment. We show that the error probability of this detector decays exponentially at a rate that is lower bounded either in terms of the noise range for bounded noise or the signal-to-noise ratio for noise with unbounded range.**

***Index Terms*—Distributed detection, large deviation, sensor networks.**

I. Introduction

WE CONSIDER the problem of decentralized detection with a wireless sensor network consisting of a fusion center and a large number of geographically distributed sensors. We assume that sensor nodes can communicate with the fusion center but not with each other, and there is no feedback from the fusion center to local sensors. Because of their low-power budget, sensors may have limited dynamic range, resolution, or communication capability. As a result, local quantization/compression of sensor observations is of great importance. In general, optimal local compression schemes are dependent on sensor noise distributions that, in many practical situations, can be hard to characterize due to large network size and unpredictable environment change. Consequently, we are motivated to design decentralized detection algorithms that work universally for any unknown noise and have low bandwidth requirements.

There has been a long history of research on decentralized detection (see [17] for an excellent survey of the early work and [1] and [20] for a more recent tutorial on this subject). Typically, the sensor observations are assumed to be conditionally independent with known distributions [16], since, in the absence of conditional independence assumption, the problem of determining the optimal decentralized detection strategy is NP-hard [18]. Other work [6], [9] focused on exploiting the known joint observation probability distribution function (pdf) to design local quantization schemes. The decentralized detection problem with communication constraint was studied in [15], whereby the sensors employ a "send/no send" strategy depending on whether the likelihood ratios fall in a certain range. Also, the quantization or thresholding of local likelihood ratios is studied in [2], [3], [17], and [19], where the goal is to reduce the communication requirement from sensors to the fusion center. Without the knowledge of observation distribution, the work of [14] proposed to use a training sequence to aid the design of local data quantization strategies.

Recently, several universal decentralized estimation schemes were proposed [10], [12], [13] for distributed parameter estimation in the presence of unknown, additive sensor noises that are bounded and identically distributed. These universal decentralized estimation schemes have a low bandwidth requirement: Each observation is compressed to exactly one binary bit per sensor. In particular, half of the sensors will send to the fusion center the first most significant bit (MSB) of their observations, one fourth of the sensors will send the second MSB of their observations, and so on. When properly combined at the fusion center, these bits can be used to estimate the unknown parameter, resulting in a mean-squared error (MSE) that is within a constant factor of four to the minimum achievable. The isotropic universal estimation in a bandwidth-constrained *ad hoc* sensor network was considered in [11], where sensors adopt an identical probabilistic quantization strategy: Each sensor quantizes its observation to the first MSB with probability $1/2$, quantizes to the second MSB with probability $1/4$, and so on.

In this paper, we construct a universal decentralized scheme for the detection of a deterministic signal corrupted by additive noise with unknown distribution by using the recently proposed isotropic decentralized estimation scheme [10], [11]. Operationally, each sensor collects an observation, performs a local probabilistic compression, and sends a 1-bit message to the fusion center, while the latter performs the final detection by appropriately combining the received bits from all sensors. The resulting universal distributed detection scheme requires only the knowledge of either the noise range or its second-order moment and has an error probability that decays exponentially. Moreover, the decaying rate (called error exponent) is shown to be lower bounded either in terms of the noise range when noise is

Manuscript received March 17, 2004; revised July 29, 2004. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant OPG0090391, by the Canada Research Chair Program, and by the National Science Foundation under Grant DMS-0312416. This work was presented in part at the IEEE Global Telecommunications Conference (Globecom), Dallas, TX, November 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zidong Wang.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: xiao@ece.umn.edu; luozq@ece.umn.edu).

Digital Object Identifier 10.1109/TSP.2005.850334

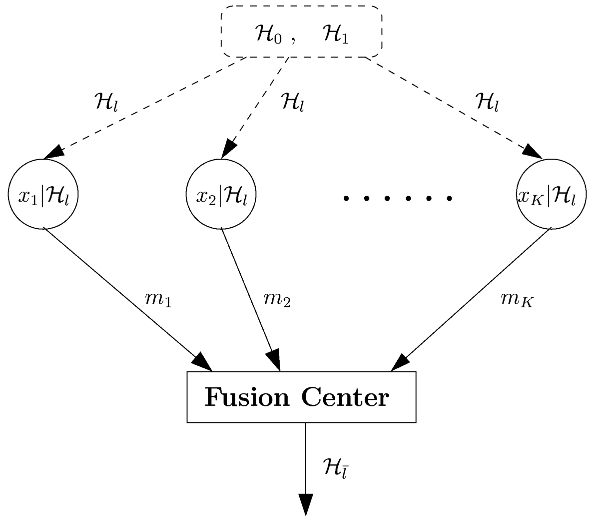


Fig. 1. Decentralized detection scheme.

bounded or the signal-to-noise ratio (SNR) for noise with unbounded range.

Our paper is organized as follows. In Section II, the decentralized detection problem is formulated. In Section III, we propose a decentralized detection scheme for the bounded noise case. Its performance is analyzed and compared with that of a detection scheme based on real-valued observations. The extension of this decentralized detection scheme to the case of unbounded noises is given in Section IV. Section V contains some concluding remarks.

II. Problem Formulation

We consider the decentralized binary hypothesis detection problem structured in Fig. 1. Suppose that the two prior hypotheses are $\mathcal{H}_0, \mathcal{H}_1$, and there are K distributed sensor nodes. To accomplish the detection task, each sensor collects a noisy observation x_k, performs a (probabilistic) local quantization, and then transmits the quantization result $m_k(x_k)$ (which is also called local message function) to the fusion center. Upon receiving the messages $\{m_1, m_2, \ldots, m_K\}$, the fusion center selects one of the two hypotheses as the final decision by combining the received messages according to a detection rule (to be designed). Our goal is to design the decentralized detection scheme (DDS), which consists of the local message functions $\{m_1, m_2, \ldots, m_K\}$, and the final fusion strategy at the fusion center, so that the detection error probability $\mathsf{P}_e(K) = \mathsf{P}(\mathcal{H}_{\bar{l}} \neq \mathcal{H}_l)$ is minimized. In contrast to DDS, a centralized detection scheme (CDS) does not perform local decisions; it generates the final detection decision on the basis of real-valued sensor observations.

In what follows, we consider the problem of detecting a known deterministic signal corrupted by additive noises. For symmetry, we assume that the sensor observations are described by

$$\begin{aligned} \mathcal{H}_0&: x_k = -d + n_k, \quad k = 1, 2, \ldots, K \\ \mathcal{H}_1&: x_k = d + n_k, \quad k = 1, 2, \ldots, K \end{aligned} \tag{1}$$

where d and $-d$ are the deterministic signals under hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively, and $\{n_1, n_2, \ldots, n_K\}$ are mutually independent noise random variables with zero mean. The distributions of sensor noises are unknown and may be different across sensors. The variances of n_k are assumed to be upper bounded by a known constant σ^2.

The error probability is given by

$$\mathsf{P}_e(K) = \mathsf{P}(\text{decide}\, \mathcal{H}_1 \,|\, \mathcal{H}_0)\mathsf{P}(\mathcal{H}_0) + \mathsf{P}(\text{decide}\, \mathcal{H}_0 \,|\, \mathcal{H}_1)\mathsf{P}(\mathcal{H}_1).$$

For any reasonable detection strategy, the associated detection error $\mathsf{P}_e(K)$ typically decays exponentially as $K \to \infty$. Thus, we can define the error exponent as

$$R = \liminf_{K \to \infty} R(K), \quad \text{where } R(K) = -\frac{1}{K} \log \mathsf{P}_e(K). \tag{2}$$

When the noise pdfs $\{f_k(x) : k = 1, 2, \ldots, K\}$ and the prior probabilities of the hypotheses $\mathsf{P}(\mathcal{H}_0)$ and $\mathsf{P}(\mathcal{H}_1)$ are known, the optimal decision rule that minimizes the detection error probability is given by the maximum *a posteriori* detector (see, e.g., [8]). The decision rule of this detector is the likelihood ratio test (LRT)

$$\frac{1}{K} \sum_{k=1}^{K} \log \frac{f_k(x_k - d)}{f_k(x_k + d)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \frac{1}{K} \log \frac{\mathsf{P}(\mathcal{H}_0)}{\mathsf{P}(\mathcal{H}_1)}. \tag{3}$$

When $\{n_1, n_2, \ldots, n_K\}$ are independent and identically distributed (i.i.d.) with pdf $f(x)$, the error exponent attained by the LRT in (3) is the Chernoff distance between $g_0(x) = f(x + d)$ and $g_1(x) = f(x - d)$ (if $g_0(x)$ and $g_1(x)$ are mutually absolute continuous; see, e.g., [5])

$$\mathcal{C}(g_0, g_1) = -\min_{0 \le s \le 1} \log \int g_0(x)^s g_1^{1-s}(x)\, dx.$$

Notice that $g_0(x) = f(x + d)$ and $g_1(x) = f(x - d)$ are the conditional sensor observation distributions under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Chernoff distance characterizes the largest possible exponential rate at which the error probability of any detection scheme decreases to zero as $K \to \infty$. Therefore, it can serve as a useful benchmark for evaluating the asymptotic error performance of various detection schemes.

Implementing the above likelihood test in a sensor network requires both the knowledge of the noise pdfs and the transmission of real-valued messages. This makes distributed likelihood-testing impractical for a large–scale sensor network. To ensure low bandwidth requirement and universality to noise distributions, we will constrain the message function $m_k(x_k)$ to be 0-1 valued and be independent of noise pdf. A main objective of this paper is to understand the impact of these constraints on the decentralized detection performance.

As an example, consider the simple choice of taking each m_k as a sign detector

$$m_k(x_k) = \begin{cases} 1, & \text{if } x_k \ge 0 \\ 0, & \text{if } x_k < 0. \end{cases}$$

It is easy to see that m_k has the conditional distribution

$$\mathsf{P}(m_k = 0 \mid \mathcal{H}_0) = \int_{-\infty}^{0} f_k(x+d)dx = F_k(d)$$
$$\mathsf{P}(m_k = 1 \mid \mathcal{H}_0) = \int_{0}^{\infty} f_k(x+d)dx = 1 - F_k(d)$$

where $F_k(x)$ is the cumulative noise density function of n_k. Similarly

$$\mathsf{P}(m_k = 0 \mid \mathcal{H}_1) = F_k(-d)$$
$$\mathsf{P}(m_k = 1 \mid \mathcal{H}_1) = 1 - F_k(-d).$$

However, if $f_k(x) = 0$ for all $x \in [-d, d]$ and $1 \le k \le K$, then $F_k(-d) = F_k(d)$, which implies that the distributions of m_k are the same under both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. Hence, $\mathcal{H}_0$ and $\mathcal{H}_1$ are not distinguishable from the message functions $\{m_1, m_2, \ldots, m_K\}$. Thus, local sign detectors cannot work *universally* for all noise pdfs. In the ensuing sections, we describe universal DDSs for both bounded and unbounded noises.

III. Detection in Bounded Noise Case

In this section, the sensor noises are assumed to be bounded in an interval $[-U, U]$. In other words, if we let

$$\mathcal{M}_U = \left\{ f(x)\colon \int_{-U}^{U} f(x)dx = 1, \quad \int_{-U}^{U} xf(x)dx = 0 \right.$$
$$\left. f(x) \ge 0, \quad \mathrm{supp}(f) \subseteq [-U, U] \right\}$$

then the noise pdfs $f_k(x) \in \mathcal{M}_U$ for all $k = 1, 2, \ldots, K$. Let us now recall two large deviation results.

Proposition 3.1 (Hoeffding [7]): Let $\{\eta_1, \eta_2, \ldots, \eta_K\}$ be mutually independent random variables with $a \le \eta_k \le b$ for all $k = 1, 2, \ldots, K$. Suppose $\bar{\eta}_K = (1/K)(\eta_1 + \eta_2 + \cdots + \eta_K)$; then for any $\epsilon \ge 0$

$$\left. \begin{array}{l} \mathsf{P}(\bar{\eta}_K - \mathsf{E}(\bar{\eta}_K) \ge \epsilon) \\ \mathsf{P}(\mathsf{E}(\bar{\eta}_K) - \bar{\eta}_K \ge \epsilon) \end{array} \right\} \le \exp\left(-\frac{2K\epsilon^2}{(b-a)^2} \right).$$

Proposition 3.2 (Chernoff [4]): Let $\{\zeta_1, \zeta_2, \ldots, \zeta_K\}$ be i.i.d. Bernoulli random variables with $\mathsf{P}(\zeta_k = 0) = q, \mathsf{P}(\zeta_k = 1) = p$. Suppose $\bar{\zeta}_K = (1/K)(\zeta_1 + \zeta_2 + \cdots + \zeta_K)$; then for any $s \ge 0$, it holds that

$$\begin{array}{ll} \mathsf{P}(\bar{\zeta}_K \ge s) \le \exp(-K\phi_p(s)), & \text{if } s \ge p \\ \mathsf{P}(\bar{\zeta}_K \le s) \le \exp(-K\phi_p(s)), & \text{if } s \le p \end{array} \quad (4)$$

where

$$\phi_p(s) = s \ln \frac{s}{p} + (1-s) \ln \frac{1-s}{1-p}.$$

Moreover, $\phi_p(s)$ is the best achievable exponent in (4) in the sense that, for any $\phi > \phi_p(s)$, there holds

$$\liminf_{K \to \infty} \mathsf{P}(\bar{\zeta}_K \ge s) \exp(K\phi) = \infty, \quad \text{if } s \ge p$$
$$\liminf_{K \to \infty} \mathsf{P}(\bar{\zeta}_K \le s) \exp(K\phi) = \infty, \quad \text{if } s \le p.$$

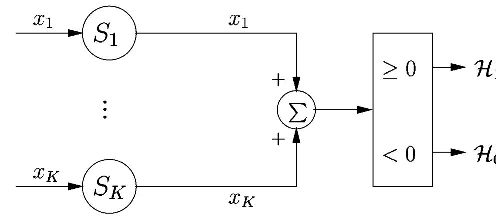


Fig. 2. Centralized detection scheme.

A. Centralized Detection With Real-Valued Observations

We first consider the centralized detection scheme (CDS) where the detector has direct access to the sensor observations $\{x_1, x_2, \ldots, x_K\}$, but the noise pdfs are unknown. Due to the insufficient statistic, it is natural to perform the following decision strategy:

$$\bar{x}_K := \frac{1}{K} \sum_{k=1}^{K} x_k \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lesseqgtr}} \gamma \quad (5)$$

where the threshold γ is to be chosen. By applying Proposition 3.1 to the detection scheme (5), we have the following result.

Theorem 3.3: For the detection problem (1), if the noises $\{n_1, n_2, \ldots, n_K\}$ are bounded in $[-U, U]$, then the detection scheme (5) with threshold $\gamma = 0$ (see Fig. 2) achieves detection error probability that decays exponentially with rate

$$R(K) \ge \frac{d^2}{2U^2}, \quad \text{for all } K > 0$$

where $R(K)$ is defined in (2). Thus, the error exponent

$$R = \liminf_{K \to \infty} R(K) \ge \frac{d^2}{2U^2}. \quad (6)$$

.

Proof: Notice that $\mathsf{E}(x_k \mid \mathcal{H}_0) = -d, \mathsf{E}(x_k \mid \mathcal{H}_1) = d$, so

$$\mathsf{E}(\bar{x}_K \mid \mathcal{H}_0) = -d, \quad \mathsf{E}(\bar{x}_K \mid \mathcal{H}_1) = d.$$

The sensor observations are bounded as

$$\begin{cases} -U - d \le x_k \le U - d, & \text{under } \mathcal{H}_0 \\ -U + d \le x_k \le U + d, & \text{under } \mathcal{H}_1. \end{cases}$$

Applying Proposition 3.1, we obtain the error probability of (5)

$$\begin{aligned} \mathsf{P}_e(K) &= \mathsf{P}(\bar{x}_K \ge \gamma \mid \mathcal{H}_0)\mathsf{P}(\mathcal{H}_0) + \mathsf{P}(\bar{x}_K < \gamma \mid \mathcal{H}_1)\mathsf{P}(\mathcal{H}_1) \\ &= \mathsf{P}(\bar{x}_K - \mathsf{E}(\bar{x}_K) \ge d + \gamma \mid \mathcal{H}_0)\mathsf{P}(\mathcal{H}_0) \\ &\quad + \mathsf{P}(\mathsf{E}(\bar{x}_K) - \bar{x}_K > d - \gamma \mid \mathcal{H}_1)\mathsf{P}(\mathcal{H}_1) \\ &\le \exp\left(-\frac{K(d+\gamma)^2}{2U^2} \right) \mathsf{P}(\mathcal{H}_0) \\ &\quad + \exp\left(-\frac{K(d-\gamma)^2}{2U^2} \right) \mathsf{P}(\mathcal{H}_1) \end{aligned} \quad (7)$$

where the last step applied the fact that $\bar{x}_k$ takes value in an interval of length $2U$ under either hypothesis. Since the error exponent is determined by the least of the two exponents in (7), we take $\gamma = 0$ and obtain

$$\mathsf{P}_e(K) \le \exp\left(-\frac{Kd^2}{2U^2} \right)$$

and

$$R(K) \ge \frac{d^2}{2U^2}, \quad \text{for all } K > 0.$$

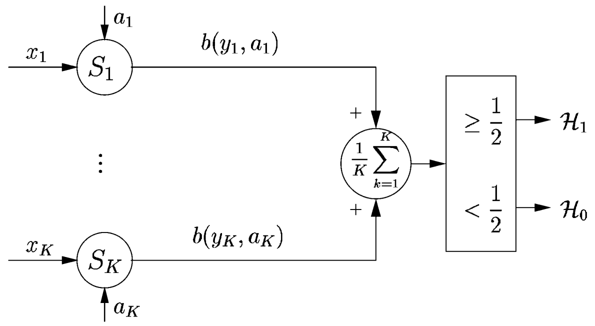


Fig. 3. Universal decentralized detection in bounded noise case.

Thus, letting $K \to \infty$ shows that the error exponent R is also lower bounded by $d^2/2U^2$. □

The lower bound (6) on the error exponent is universal for all noise distributions $f_k \in \mathcal{M}_U$. When U is much larger than d, this bound is actually tight since (6) holds with equality when the noises are i.i.d. with pdf $f(x) = 1/2\delta(x+U)+1/2\delta(x-U)$, where $\delta(\cdot)$ denotes the usual point mass delta function. In particular, we can, in this case, linearly transform each observation x_k to a 0–1 Bernoulli random variable and apply Proposition 3.2 to obtain the error exponent

$$R = \phi_{\frac{U+d}{2U}}\left(\frac{1}{2}\right) = \frac{1}{2}\ln\left(1-\frac{d^2}{U^2}\right) + \frac{d}{2U}\log\frac{U+d}{U-d} \approx \frac{d^2}{2U^2}, \quad \text{when } U \gg d.$$

This matches the lower bound (6).

Notice that the detection strategy (5) is universal since it does not require the knowledge of noise pdf. However, when applied in sensor networks, it is not bandwidth efficient since real-valued messages need to be transmitted from local sensors to the fusion center. In what follows, we impose a 1-bit per sample bandwidth constraint at each sensor node and develop a universal DDS under this condition.

B. Decentralized Detection Under a 1-Bit Bandwidth Constraint

Throughout this section, we assume that each local message is a 0-1 binary function. The basic idea is to introduce, at each sensor, an auxiliary random variable a_k with which each sensor can randomize its message function. Specifically, sensor k generates a binary local message $m_k(x_k, a_k)$ based on not only its observation x_k but also the outcome of the auxiliary random variable a_k. The DDS can be briefly described as follows (also see Fig. 3).

- First, the auxiliary random variable a_k at each sensor is specified to have the following distribution:

$$\mathsf{P}(a_k = i) = 2^{-i}, \quad i = 1, 2, 3, \ldots \tag{8}$$

  Notice that all sensors in the network use identical but independent auxiliary random variables.
- Second, let $W = U + d$; then, $x_k \in [-W, W]$ under both hypotheses. Suppose that W is known to local sensors, and a_k is the local auxiliary random variable outcome, and the notation $b(z; i)$ denotes the ith bit of a real number z. Then, the local message function is given by

$$m_k = b(y_k; a_k), \quad \text{where } y_k = \frac{W + x_k}{2W} \in [0, 1]. \tag{9}$$

  In other words, the message m_k is simply the a_kth bit of the real number y_k, which is either 0 or 1.
- Finally, upon receiving the 1-bit binary messages $\{m_1, m_2, \ldots, m_K\} = \{b(y_1; a_1), b(y_2; a_2), \ldots, b(y_K; a_K)\}$, the fusion center performs the hypothesis testing according to the following rule:

$$\bar{m}_K := \frac{1}{K}\sum_{k=1}^{K} m_k = \frac{1}{K}\sum_{k=1}^{K} b(y_k; a_k) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \frac{1}{2}. \tag{10}$$

We now analyze the detection performance of DDS (8)–(10).

Theorem 3.4: Consider the decentralized detection problem (1) where each sensor is restricted to send a 1-bit message to the fusion center. If the noises $\{n_1, n_2, \ldots, n_K\}$ are bounded in $[-U, U]$, with the knowledge of $W = U+d$, the DDS defined by the local message functions $m_k(x_k, a_k)$ (8) and (9), and the final fusion rule (10) has a detection error that decays exponentially at a rate

$$R(K) \geq \frac{d^2}{2W^2}, \quad \text{for all } K > 0.$$

Proof: It is easy to see that $\{b(y_1; a_1), b(y_2; a_2), \ldots, b(y_K; a_K)\}$ are conditionally independent. In addition, it follows from the distribution of a_k in (8) that $b(y_k; a_k)$ is an unbiased quantization of y_k, i.e.,

$$\mathsf{E}_{a_k}(b(y_k; a_k)) = \sum_{i=1}^{\infty} b(y_k; i)2^{-i} = y_k.$$

Therefore, we obtain

$$\mathsf{E}(b(y_k; a_k)) = \mathsf{E}_{n_k}(y_k) = \frac{W + \mathsf{E}(x_k)}{2W} \tag{11}$$

where the expectation of $b(y_k; a_k)$ is taken with respect to both the noise n_k and the auxiliary random variable a_k. Furthermore, we can use (11) to compute the conditional expectations

$$\begin{cases} \mathsf{E}(b(y_k; a_k) \,|\, \mathcal{H}_0) = \frac{W+\mathsf{E}(x_k \,|\, \mathcal{H}_0)}{2W} = \frac{1}{2} - \frac{d}{2W} \\ \mathsf{E}(b(y_k; a_k) \,|\, \mathcal{H}_1) = \frac{W+\mathsf{E}(x_k \,|\, \mathcal{H}_1)}{2W} = \frac{1}{2} + \frac{d}{2W}. \end{cases} \tag{12}$$

The detection error probability is given by

$$\mathsf{P}_e(K) = \mathsf{P}(\bar{m}_K \geq 1/2 \,|\, \mathcal{H}_0)\mathsf{P}(\mathcal{H}_0) + \mathsf{P}(\bar{m}_K < 1/2 \,|\, \mathcal{H}_1)\mathsf{P}(\mathcal{H}_1). \tag{13}$$

Applying Proposition 3.2 with $p = 1/2 - d/2W, q = 1/2 + d/2W$, and $s = 1/2$, we obtain

$$\mathsf{P}(\bar{m}_K \geq 1/2 \,|\, \mathcal{H}_0) \leq \exp(-K\phi_p(1/2))$$

and

$$\mathsf{P}(\bar{m}_K < 1/2 \,|\, \mathcal{H}_1) \leq \exp(-K\phi_q(1/2)).$$

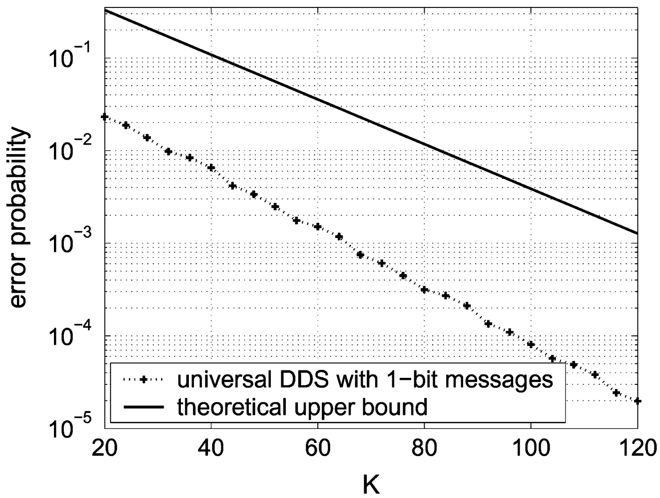


Fig. 4. Error exponents of DDS in bounded noise case.

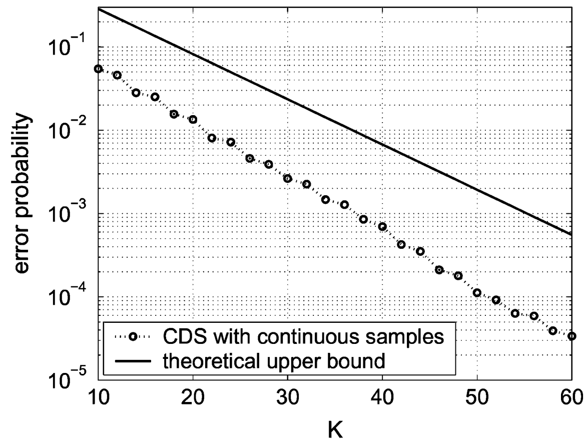


Fig. 5. Error exponents of CDS in bounded noise case.

Since $\phi_p(1/2) = \phi_q(1/2) = -1/2\ln(4pq)$, it follows from (13) that

$$\begin{aligned}\mathsf{P}_e(K) &\le \exp(-K\phi_p(1/2))\mathsf{P}(\mathcal{H}_0) + \exp(-K\phi_q(1/2))\mathsf{P}(\mathcal{H}_1)\\ &= \exp(-K\phi_p(1/2)).\end{aligned}$$

Thus, we have

$$R(K) \ge \phi_p(1/2) = \frac{1}{2}\ln\frac{1}{4pq} = \frac{1}{2}\ln\frac{1}{1-d^2/W^2} \ge \frac{d^2}{2W^2}.$$

The proof is complete. □

The error exponent bounds in Theorems 3.3 and 3.4 differ by a factor of $W^2/U^2 = (U+d)^2/U^2$, where d is the amplitude of the deterministic signal, and U is the noise range. This suggests that to ensure the same universal performance, the DDS only needs a multiplicative constant factor of $(U+d)^2/U^2$ more sensors than that needed by the CDS to achieve the same detection performance, providing that local sensors have the knowledge the observation range W in DDS. In other words, comparing these two universal detectors, the low bandwidth constraint (one bit per sample per node) only results in a constant factor increase in the sensor network size.

C. Simulations

In the simulations, we choose $d = 1$. For simplicity, the noises are selected to be spatially i.i.d. in the range $[-2, 2]$ with pdf $0.5\delta(x+2) + 0.5\delta(x-2)$. For each K (the total number of sensors), we obtain the probability of detection error by repeating the experiments 10^6 times. Fig. 4 shows the detection performance of the universal DDS in (8)–(10) with 1-bit message functions, while Fig. 5 shows the detection performance of the CDS in (5) with real-valued observations. The two theoretical upper bounds of detection errors are $\exp(-(Kd^2)/(2U^2))$ and $\exp(-(Kd^2)/(2W^2))$, respectively, where $W = U + d$ is the range of the sensor observations.

In Figs. 4 and 5, we can see that the DDS achieves the same error probability as the CDS with about twice as many sensors. This agrees well with the theoretical upper bounds for the respective error exponents, which are related by a factor of $W^2/U^2 = 9/4 = 2.25$. The experimental error probability curves are parallel to the theoretical error probability curves, suggesting that the two are related by a bounded multiplicative factor.

IV. Universal Decentralized Detection in the Unbounded Noise Case

In this section, the DDS is extended to the case when sensor noises have infinite range. In the extended scheme, the unbounded observations are first truncated to an appropriate finite range; then, the DDS developed in the previous sections is applied to these truncated observations.

Throughout this section, we assume that sensor noises have finite variances upper bounded by σ^2. In other words, if we let

$$\mathcal{N}_\sigma = \left\{f(x) : \int_{-\infty}^{\infty} f(x)dx = 1, \quad \int_{-\infty}^{\infty} xf(x)dx = 0 \right. \\ \left. \int_{-\infty}^{\infty} x^2 f(x)dx = \sigma^2, \quad f(x) \ge 0\right\}$$

then $f_k(x) \in \mathcal{N}_\sigma$ for all $1 \le k \le K$.

A. Universal DDS

For any $W > 0$ and sensor observation x_k, we say x_k^* is the truncation of x_k to the range $[-W, W]$ if

$$x_k^* = \begin{cases} W, & \text{if } x_k \ge W \\ x_k, & \text{if } -W < x_k < W \\ -W, & \text{if } x_k \le -W. \end{cases}$$

First, we have the following lemma.

Lemma 4.1: For any random variable x with finite second-moment Ω_2, if x^* is the truncation of x to the range $[-W, W]$, where $W > 0$, then

$$|\mathsf{E}(x) - \mathsf{E}(x^*)| \le \frac{\Omega_2}{4W}.$$

Proof: Let $g(x)$ denote the pdf of random variable x. Then

$$\begin{aligned}|\mathsf{E}(x) - \mathsf{E}(x^*)| &= \left|\int_W^{\infty}(x-W)g(x)\,dx + \int_{-\infty}^{-W}(x+W)g(x)\,dx\right|\\ &= \left|\int_W^{\infty}(x-W)g(x)\,dx - \int_{-\infty}^{-W}(-x-W)g(x)\,dx\right|\\ &\le \max\{T_1(W), T_2(W)\} \qquad (14)\end{aligned}$$

where we let

$$T_1(W) = \int_W^{\infty} (x - W) g(x) dx \geq 0$$
$$T_2(W) = \int_{-\infty}^{-W} (-x - W) g(x) dx \geq 0.$$

It is easy to see that

$$\begin{aligned}\Omega_2 &= \int_{-\infty}^{\infty} x^2 f(x) dx \geq \int_W^{\infty} x^2 f(x)\, dx \\ &= \int_W^{\infty} (x - W + W)^2 f(x)\, dx \\ &\geq \int_W^{\infty} 4W(x - W) f(x)\, dx \\ &= 4W \int_W^{\infty} (x - W) f(x)\, dx.\end{aligned}$$

Therefore, we obtain

$$T_1(W) = \int_W^{\infty} (x - W) f(x)\, dx \leq \frac{\Omega_2}{4W}.$$

Similarly, we have $T_2(W) \leq \Omega_2/4W$. Combining these bounds on $T_1(W)$ and $T_2(W)$ with (14) proves the lemma. □

By the assumption that $f_k(x) \in \mathcal{N}_\sigma$ for all $1 \leq k \leq K$, it follows that the second moment of any sensor observation x_k is at most $\sigma^2 + d^2$ under either $\mathcal{H}_0$ or $\mathcal{H}_1$. Let us take

$$W = \frac{\sigma^2 + d^2}{2d} \tag{15}$$

and truncate x_k to $[-W, W]$ to get x_k^*. By Lemma 4.1, we obtain

$$\begin{cases} |\mathsf{E}(x_k^* \,|\, \mathcal{H}_0) - \mathsf{E}(x_k \,|\, \mathcal{H}_0)| \leq \dfrac{\Omega_2}{4W} \leq \dfrac{\sigma^2 + d^2}{4W} = \dfrac{d}{2} \\ |\mathsf{E}(x_k^* \,|\, \mathcal{H}_1) - \mathsf{E}(x_k \,|\, \mathcal{H}_1)| \leq \dfrac{d}{2}. \end{cases}$$

Also, it follows from (1) that $\mathsf{E}(x_k \,|\, \mathcal{H}_0) = -d, \mathsf{E}(x_k \,|\, \mathcal{H}_1) = d$. Hence, we have

$$\mathsf{E}(x_k^* \,|\, \mathcal{H}_0) \leq -\frac{d}{2}, \quad \mathsf{E}(x_k^* \,|\, \mathcal{H}_1) \geq \frac{d}{2}. \tag{16}$$

It is easy to see that $\{x_1^*, x_2^*, \ldots, x_K^*\}$ are conditionally independent and bounded in $[-W, W]$. This motivates us to construct a DDS as follows (see Fig. 6).

- After collecting an observation x_k, each sensor first truncates x_k to the range $[-W, W]$ to obtain x_k^*, where W is specified in (15).
- Then, similar to the DDS described in Section III for the case when observations are bounded, x_k^* is quantized to a 1-bit probabilistic message

$$m_k^* = b(y_k^*; a_k), \quad \text{where } y_k^* = \frac{W + x_k^*}{2W} \in [0, 1] \tag{17}$$

where a_k is defined in (8), and the fusion center performs hypothesis testing according to the following rule:

$$\bar{m}_K^* := \frac{1}{K} \sum_{k=1}^{K} m_k^* = \frac{1}{K} \sum_{k=1}^{K} b(y_k^*, a_k) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \frac{1}{2}.$$

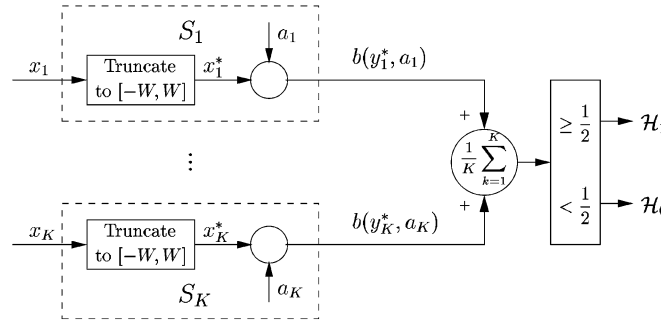


Fig. 6. Universal decentralized detection in unbounded noise case.

Let us now analyze the detection performance of the above 1-bit DDS.

Theorem 4.2: Consider the decentralized detection problem (1), where each sensor is restricted to send a 1-bit message to the fusion center,, and the sensor noise variances are bounded by σ^2. Then, the universal DDS specified above achieves an error probability that decays exponentially at a rate

$$R(K) \geq \frac{1}{2} \ln\left(1 + \frac{r^2}{2r + 1}\right) \tag{18}$$

where $r = d^2/\sigma^2$ is the SNR.

Proof: Similar to the derivation of (12), with (15)–(17), we have

$$\begin{cases} \mathsf{E}(b(y_k^*, a_k) \,|\, \mathcal{H}_0) = \dfrac{W + \mathsf{E}(x_k^* \,|\, \mathcal{H}_0)}{2W} \leq \dfrac{1}{2} - \dfrac{d^2}{2(d^2 + \sigma^2)} \\ \mathsf{E}(b(y_k^*, a_k) \,|\, \mathcal{H}_1) = \dfrac{W + \mathsf{E}(x_k^* \,|\, \mathcal{H}_1)}{2W} \geq \dfrac{1}{2} + \dfrac{d^2}{2(d^2 + \sigma^2)}. \end{cases}$$

Applying Proposition 3.2 with

$$p = \frac{1}{2} - \frac{d^2}{2(d^2 + \sigma^2)} = \frac{1}{2} - \frac{r}{2(1 + r)}$$
$$q = \frac{1}{2} + \frac{d^2}{2(d^2 + \sigma^2)} = \frac{1}{2} + \frac{r}{2(1 + r)}$$

and $s = 1/2$ and using an argument similar to the proof of Theorem 3.4, we obtain that the detection error probability

$$\begin{aligned}\mathsf{P}_e(K) &= \mathsf{P}(\bar{m}_K \geq 1/2 \,|\, \mathcal{H}_0) \mathsf{P}(\mathcal{H}_0) \\ &\quad + \mathsf{P}(\bar{m}_K < 1/2 \,|\, \mathcal{H}_1) \mathsf{P}(\mathcal{H}_1) \\ &\leq \exp(-K \phi_p(1/2)).\end{aligned}$$

Thus, we have

$$\begin{aligned}R(K) \geq \phi_p(1/2) &= \frac{1}{2} \ln \frac{1}{4pq} = \frac{1}{2} \ln \frac{1}{1 - \left(\frac{r}{r+1}\right)^2} \\ &= \frac{1}{2} \ln\left(1 + \frac{r^2}{2r + 1}\right).\end{aligned}$$

It follows that $R \geq (1/2) \ln(1 + (r^2)/(2r + 1))$ when we let $K \to \infty$. □

B. Simulation

Again, we let $d = 1$. We choose sensor noises as Gaussian random variables with standard deviation $\sigma = 1$. By (15), the

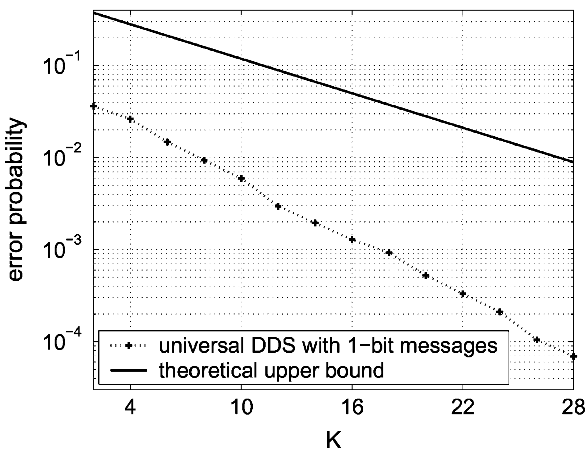


Fig. 7. Error exponent of DDS in the unbounded noise case.

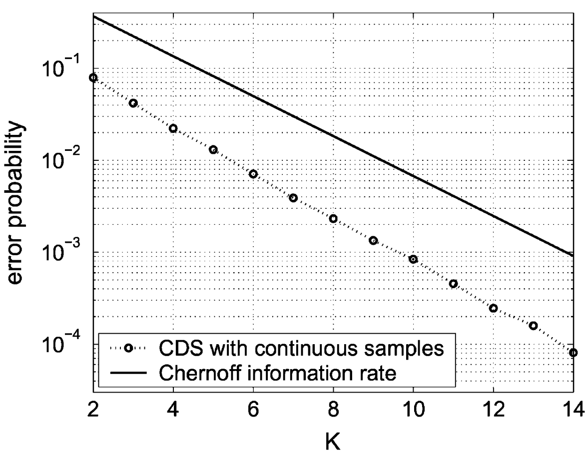


Fig. 8. Error exponents of CDS in the unbounded noise case.

truncation range $W = (\sigma^2 + d^2)/2d = 1$. The detection error upper bound for the DDS is $\mathsf{P}_e(K) \leq \exp(-KR(K))$, where $R(K)$ is specified in (18). From Fig. 7, we see that the simulation performance is better than the theoretical prediction. This is because the theoretical upper bound is universal for any noise with zero mean and variance bounded by σ^2, while the simulation is for Gaussian noise only.

The best achievable error exponent in the case of Gaussian noise is $d^2/2\sigma^2$, which is the Chernoff distance between the two noise pdfs under $\mathcal{H}_0$ and $\mathcal{H}_1$. Fig. 8 plots this Chernoff error bound curve $\exp(-Kd^2/2\sigma^2)$ as well as the simulation results for the CDS (5) with $\gamma = 0$. The two simulation curves in Figs. 7 and 8 suggest that, with about twice as many sensors, the DDS can achieve comparable performance as the optimal likelihood ratio test.

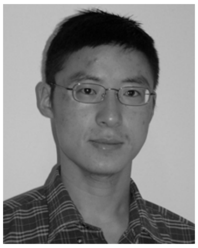
V. Conclusion

In this paper, we have considered the problem of universal decentralized detection of a deterministic signal in a bandwidth-constrained sensor network, where each sensor is restricted to send a 1-bit message to the fusion center. We have proposed a universal DDS whose performance is compared to a classical centralized detector for bounded sensor noises. We have also extended this DDS to the unbounded noise case and analyzed the two universal DDSs using the theory of large deviations. Simulation results agree well with the theoretical bounds on the corresponding error exponents.

The proposed universal DDSs have several attractive properties: 1) *isotropic*: All sensors in the network use the same local message decision rule, 2) *bandwidth efficient*: Each sensor only needs to send a 1-bit message to the fusion center, 3) *universal*: They work universally for any sensor noises with either bounded noise range or finite noise variance, and 4) *good performance*: The probability error of this universal detector decays exponentially at a provably nondiminishing rate, which is lower bounded either in terms of the noise range when nose is bounded or the SNR for noise with unbounded range. These properties make these DDSs attractive candidates for implementation in bandwidth-constrained sensor networks.

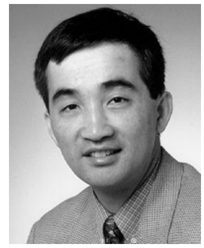
References

[1] R. S. Blum, S. A. Kassam, and H. V. Poor, “Distributed detection with multiple sensors: Part II—Advanced topics,” *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.
[2] Z. Chair and P. K. Varshney, “Distributed Bayesian hypothesis testing with distributed data fusion,” *IEEE Trans. Syst., Man, Cybern.*, vol. 18, no. 5, pp. 695–699, Sep./Oct. 1988.
[3] J. F. Chamberland and V. V. Veeravalli, “Decentralized detection in sensor networks,” *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
[4] H. Chernoff, “A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations,” *Ann. Math. Statist.*, vol. 23, pp. 493–507, 1952.
[5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer-Verlag, 1998.
[6] I. Y. Hoballah and P. K. Varshney, “An information theoretic approach to the distributed detection problem,” *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 988–994, Sep. 1989.
[7] W. Hoeffding, “Probability inequalities for sums of bounded random variables,” *J. Amer. Statist. Assoc.*, vol. 58, pp. 13–30, 1963.
[8] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
[9] M. Longo, T. D. Lookabaugh, and R. M. Gray, “Quantization for decentralized hypothesis testing under communication constraints,” *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 241–255, Mar. 1990.
[10] Z.-Q. Luo, “Universal decentralized estimation in a bandwidth constrained sensor network,” *IEEE Trans. Inf. Theory*, to be published.
[11] ——, “An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 735–744, Apr. 2005.
[12] Z.-Q. Luo and J.-J. Xiao, “Decentralized estimation in an inhomogeneous environment,” in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, p. 517.
[13] ——, “Decentralized estimation in an inhomogeneous sensing environment,” *IEEE Trans. Inf. Theory*, to be published.
[14] V. Megalooikonomou and Y. Yesha, “Quantizer design for distributed estimation with communication constraints and unknown observation statistics,” *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 181–184, Feb. 2000.
[15] C. Rago, P. Willett, and Y. Bar-Shalom, “Censoring sensors: A low-communication-rate scheme for distributed detection,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.
[16] R. R. Tenny and N. R. Sandell Jr., “Detection with distributed sensors,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, no. 4, pp. 501–510, Jul. 1981.
[17] J. N. Tstsiklis, “Decentralized detection,” in *Advances in Statistical Signal Processing*. Greenwich, CT: JAI Press, 1993, vol. 2, pp. 297–344.
[18] J. N. Tsitsiklis and M. Athans, “On the complexity of decentralized decision making and detection problems,” *IEEE Trans. Autom. Control*, vol. 30, no. 5, pp. 440–446, May 1985.
[19] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1997.
[20] R. Viswanathan and P. K. Varshney, “Distributed detection with multiple sensors: Part I—Fundamentals,” *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.



Jin-Jun Xiao (S'04) received the B.Sc. degree in applied mathematics from Jilin University, Jilin, China, in 1997 and the M.Sc. degree in mathematics from the University of Minnesota, Minneapolis, in 2003. He is currently working toward the Ph.D. degree in electrical engineering at the University of Minnesota.

His research interests are in wireless sensor networks, information theory, and optimization.



Zhi-Quan (Tom) Luo (SM'03) received the B.Sc. degree in mathematics from Peking University, Beijing, China, in 1984. During the academic year of 1984 to 1985, he was with Nankai Institute of Mathematics, Tianjin, China. From 1985 to 1989, he studied at the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, where he received the Ph.D. degree in operations research.

In 1989, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he became a Professor in 1998 and held the Canada Research Chair in Information Processing since 2001. Since April 2003, he has been a Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, and holds an ADC Chair in Digital Technology. He is presently serving as an associate editor for several international journals, including the *SIAM Journal on Optimization*, *Mathematical Programming*, *Mathematics of Computation*, and *Mathematics of Operations Research*. His research interests lie in the union of large-scale optimization, signal processing, data communications, and information theory.

Prof. Luo is a member of SIAM and MPS.