Detection Performance Limits for Distributed Sensor Networks in the Presence of Nonideal Channels

Qi Cheng, Biao Chen and Pramod K. Varshney

***Abstract*— Existing studies on the classical distributed detection problem typically assume idealized transmissions between local sensors and a fusion center. This is not guaranteed in the emerging wireless sensor networks with low-cost sensors and stringent power/delay constraints. By focusing on discrete transmission channels, we study the performance limits, in both asymptotic and non-asymptotic regimes, of a distributed detection system as a function of channel characteristics. For asymptotic analysis, we compute the error exponents of the underlying hypothesis testing problem; while for cases with a finite number of sensors, we determine channel conditions under which the distributed detection systems become useless – observing the channel outputs cannot help reduce the error probability at the fusion center. We demonstrate that as the number of sensors or the quantization levels at local sensors increase, the requirements on channel quality can be relaxed.**

***Index Terms*— Distributed detection, wireless sensor networks, nonideal transmission channels, Chernoff Information, Kullback-Leibler Distance, detection performance limits.**

The authors are with Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244. Email: qcheng{bichen,varshney} @ecs.syr.edu. Part of this work has been presented at the 38th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 2004. This work was supported in part by NASA-Langley under Order No. NNL04A127P and by NSF under Grant ECS-0501534.

I. INTRODUCTION

Distributed detection has been intensively studied in the past few decades (see [1]–[3] and references therein). Optimal local quantizer and fusion rule design can be found in the vast literature under either the Bayesian criterion [4]–[6] or the Neyman-Pearson criterion [7]. When the number of sensors is allowed to go to infinity, asymptotic performance analysis as well as asymptotically optimum detection structures have been addressed in [8], [9].

There is an implicit assumption in the majority of existing results: the transmissions between local sensors and the fusion center are error-free. Although by proper encoding and decoding, any noisy channel can be made arbitrarily reliable, provided the information rate is less than that prescribed by the channel capacity, this is difficult to guarantee in realistic wireless sensor networks with stringent power and delay constraints. If we are limited to zero-memory encoding, many new challenges and research opportunities arise in this area of channel-aware distributed detection. There are several recent papers that have addressed related problems [10]–[15]. In [13], [14], the impact of transmission channels on the decentralized detection system was studied. Under both the additive Gaussian noise and fading channel assumptions, the authors established the optimality of identical local sensor decision rules for the binary hypothesis testing problem using the asymptotic error exponent (Chernoff information) as the performance criterion. The fusion rule design in the presence of nonideal transmission channels between the sensors and the fusion center was studied in [11], [12]. The dual problem of designing local decision rules was considered in [15]. The optimality of a likelihood ratio test at local sensors was established, where the optimality is in the sense of minimizing the error probability at the fusion center. A more general case with soft (multi-bit) local sensor output has been treated in [16].

In this correspondence, we aim to quantify the detection performance limits on a distributed sensor network imposed by the discrete channels both in the asymptotic and non-asymptotic regimes. The discrete channels may arise as approximations of the fading channels. More importantly, with this channel model, a connection between the error exponents and the channel capacity can be readily established. Specifically, in the asymptotic regime, there is always performance gain for the detection problem if the channel itself has non-zero capacity. However, in the non-asymptotic regime, we

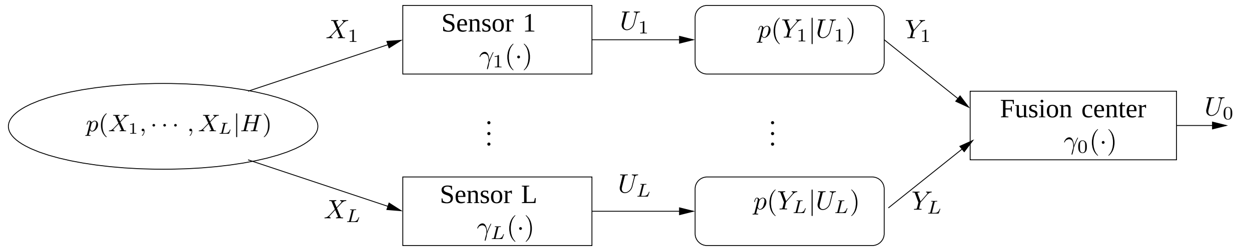


Fig. 1. A canonical distributed detection system involving the channel layer.

establish conditions under which observing the channel output does not reduce the error probability at the fusion center, regardless of channel-aware optimal processing.

This correspondence is organized as follows. Problem is formulated in Section 2. In Sections 3 and 4, asymptotic and non-asymptotic cases are treated respectively. Numerical examples are given in Section 5 followed by conclusions in Section 6.

II. Problem Formulation

Consider the parallel decentralized detection system given in Fig. 1, where L sensors collect observations that are assumed to be independent and identically distributed conditioned on a given hypothesis (H_0 or H_1). Upon observing x_l, the lth sensor maps it to a finite alphabet message u_l and sends it to the fusion center through a discrete channel. This discrete channel is characterized by a channel transition probability matrix $\mathbf{M}$, which, for simplicity, is assumed to be identical across sensors. The local sensor output $u_l = \gamma_l(x_l), l = 1, \ldots, L$, takes one of D_u possible values where $\gamma_l(\cdot)$ is the local quantizer to be optimized. The signal received at the fusion center from the lth sensor, y_l, is assumed to have D_y levels. As such, $\mathbf{M}$ is a $D_y \times D_u$ matrix with the ijth entry defined as

$$M_{ij} = P(y_l = i|u_l = j), i = 1, \ldots, D_y, j = 1, \ldots, D_u$$

where $\sum_{i=1}^{D_y} M_{ij} = 1$ for any $j \in \{1, ..., D_u\}$. Our study assumes zero-memory quantizer at local sensors and consequently, we do not consider complex channel codes that incur significant delays and rely on long input sequences.

Notice that for fixed local quantizers at the sensors, the optimal maximum *a posteriori* probability (MAP) fusion rule at the fusion center can be derived in a straightforward manner. For this reason, the optimization is carried out only with respect to $\gamma_l(\cdot)$ and the obtained optimal encoding scheme will be useful in the non-asymptotic analysis.

While optimal channel-aware local decision rules and the fusion rule are available [11], [15], the performance of the detection system is fundamentally limited by the transmission channel characteristics. A simple (and trivial) example when this may happen is the situation in which the channels connecting the sensors and the fusion center have zero capacity (e.g., a binary symmetric channel (BSC) with crossover probability 0.5). In this correspondence, we show that as long as the channels have non-zero capacity, the error probability will eventually decay to zero for the asymptotic case ($L \to \infty$). For the non-asymptotic case (finite L), we demonstrate that even with non-zero capacity channels, the distributed detection system may still be useless in that observing the channel output cannot help reduce the error probability at the fusion center. By assuming a BSC, we establish condition under which sensors become useless.

III. Asymptotic Analysis

Under the ideal channel assumption, when the number of sensors $L \to \infty$, for any sensible detection system, the probability of error goes to zero eventually [8]. This was established by using error exponents which describe how fast the error probability goes to zero [17]. Thus, a system is considered useful if and only if the error exponents are greater than zero. In this section, we will show how the observation statistics and channel parameters affect the error exponents under both Bayesian and Neyman-Pearson criteria.

A. Bayesian criterion

Using the Bayesian criterion where one wants to minimize the probability of error, we have the following theorem.

Theorem 1: Assume that $X_1, \ldots, X_L$ are independent and identically distributed (i.i.d.) with distribution Q and the form of Q depends on the hypothesis: $Q = P_0$ with prior probability π_0 and $Q = P_1$ with prior probability π_1. The best achievable exponent (Chernoff Information)[1]

$$\mathbf{CI}_c = -\min_{0\le s\le 1}\min_{\gamma}\log_2 \left[\sum_{y=1}^{D_y}\left(\sum_{u=1}^{D_u} P(y|u)P_0(u)\right)^s \cdot \left(\sum_{u=1}^{D_u} P(y|u)P_1(u)\right)^{1-s}\right] \quad (1)$$

is 0 if and only if $rank(\mathbf{M}) = \mathbf{1}$, i.e., all the columns of the transition matrix $\mathbf{M}$ are identical ($P(y|u) = P(y)$).

The proof is provided in the Appendix. The matrix $\mathbf{M}$ is rank one if and only if it has identical columns. Hence $P(y|u) = P(y)$, and the corresponding discrete channel has zero capacity. On the other hand, it is also easy to show that for the perfect channel, i.e., when $\mathbf{M}$ is an identity matrix, $\mathbf{CI}_c$ reduces to the conventional CI for the parallel distributed detection system as obtained in [8]:

$$\mathbf{CI} = -\min_{0\le s\le 1}\min_{\gamma}\log_2\left[\sum_{u=1}^{D_u} P_0^s(u)P_1^{1-s}(u)\right] \quad (2)$$

[1]The subscript c is to indicate the fact that the transmission channels are considered in deriving the $\mathbf{CI}$, hence it can be distinguished from that of the ideal channel case (cf. Eq. (2)).

B. Neyman-Pearson criterion

Analogous to the CI, the best achievable error exponent for the Neyman-Pearson test is the Kullback-Leibler Distance (KLD). The main result is summarized in the following theorem.

Theorem 2: The Kullback Leibler Distance

$$\mathbf{KLD}_c = \max_\gamma \sum_{y=1}^{D_y} \left(\sum_{u=1}^{D_u} P(y|u)P_0(u)\right) \log_2 \frac{\sum_{u=1}^{D_u} P(y|u)P_0(u)}{\sum_{u=1}^{D_u} P(y|u)P_1(u)} \tag{3}$$

is 0 if and only if $rank(\mathbf{M}) = \mathbf{1}$.

It is also easy to show that for the perfect channel, $\mathbf{KLD}_c = \max_\gamma \sum_{u=1}^{D_u} P_0(u) \log_2 \frac{P_0(u)}{P_1(u)}$, the expression without channel being considered. In the other extreme, $\mathbf{KLD}_c = 0$ if and only if the channel transition matrix $\mathbf{M}$ is rank one, thus all the columns are identical.

The above results can be interpreted as saying that the error exponents are zero if and only if the channel between each sensor and the fusion center has zero capacity. In Fig. 2, we present the error exponents under the Bayesian and the NP criteria for a binary hypothesis testing problem. The two hypotheses are $\mathrm{H}_0 : X_k \sim \mathcal{N}(-1,1)$ and $\mathrm{H}_1 : X_k \sim \mathcal{N}(1,1)$. A BSC is used between each sensor and the fusion center with crossover probability α (X-axis). Four-level quantizers are used to process the local observations and the two bits are assumed to be transmitted by using the same BSC twice. Note that the error exponents are positive for $\alpha < 0.5$, which is consistent with our analysis.

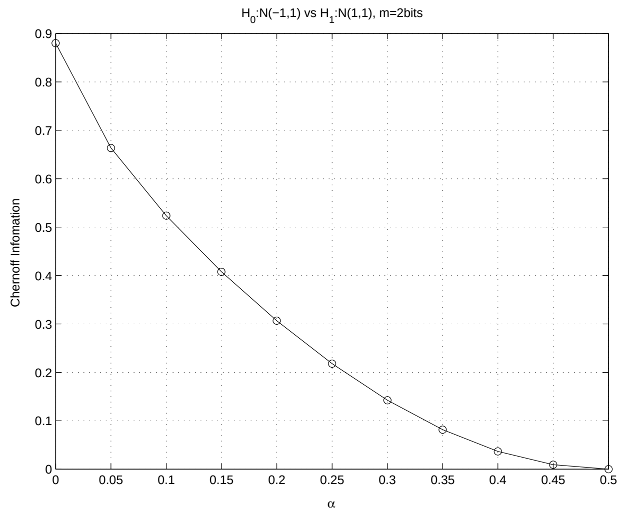
IV. Non-asymptotic Analysis

In contrast to the asymptotic results, with finite L, there exist situations in which even though the channel capacity is greater than zero, the overall system does not lead to any reduction in error probability compared to a detector using only prior information. Because of the simplicity of the BSC (single parameter characterization), we restrict the non-asymptotic analysis to the case where each sensor is connected to the fusion center by a BSC (multiple bits are sent through repeated use of the same BSC). The parameter characterizing a BSC is the crossover probability α and we only consider $0 \leq \alpha \leq 0.5$.[2]

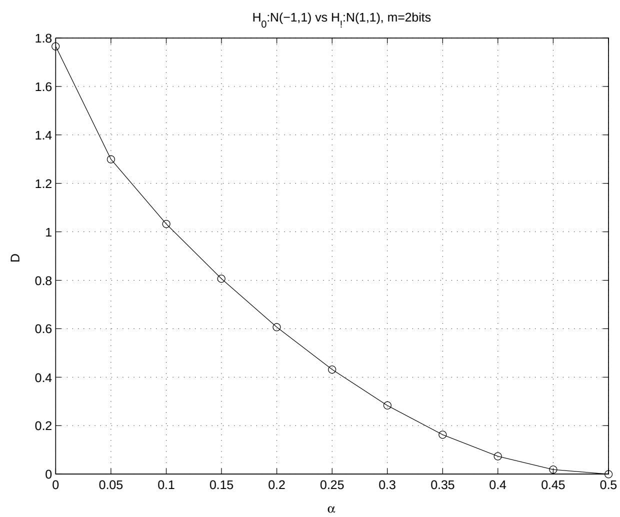
Under this formulation, we find that there exists a threshold α_s that divides all possible BSCs into two groups. If $\alpha < \alpha_s$, the distributed detection system is useful and is able to reduce the error probability from that obtained by utilizing only the prior information; otherwise, the system is useless and one should simply rely on the prior probability in making a final decision. This threshold is a function of prior probabilities, observation statistics, number of sensors in the system and quantization levels at the local sensors.

Without loss of generality, we assume $\pi_0 \leq \pi_1$. One should achieve a performance no worse than $Pe = \min(\pi_0, \pi_1) = \pi_0$, which is obtained when the decision maker declares $U_0 = 1$ by totally ignoring any observations. We assume that the L sensors employ identical decision rules. It is asymptotically

[2]Any BSC with $\alpha > 0.5$ is equivalent to a BSC with crossover probability $1 - \alpha$



(a)



(b)

Fig. 2. Error exponents as a function of α. Each sensor uses 2-bit quantization. (a)Chernoff Information (b)Kullback-Leibler Distance.

optimal to apply identical decision rules to identical local sensor observations when the number of sensors is large [8], [13], [14]. Even in the non-asymptotic case, identical local decision rules are optimal in most situations, with an exception only for discrete local observations with carefully selected probability mass functions [18].

The optimal MAP rule for final decision making at the fusion center is

$$\log \frac{P(y_1, \ldots, y_L|H_1)}{P(y_1, \ldots, y_L|H_0)} \underset{H_1}{\overset{H_0}{\lessgtr}} \log \frac{\pi_0}{\pi_1} \tag{4}$$

If the log likelihood ratio (LLR) of $\mathbf{Y} = [y_1, \ldots, y_L]$ is always greater than the threshold $\log \frac{\pi_0}{\pi_1}$, then the distributed detection system is useless in that it cannot improve the performance[3]. Transmission of m-bit messages from local sensors to the

[3]The situation that LLR of $\mathbf{Y}$ is always less than $\log \frac{\pi_0}{\pi_1}$ is not possible under the assumption $\pi_0 \leq \pi_1$. If $\pi_0 \geq \pi_1$ is assumed, we will consider LLR to be always less than instead of greater than the threshold.

fusion center is carried out by m uses of a BSC. Then the LLR is a function of γ (local m bit quantizer) and α (BSC crossover probability), hence the notation $LLR(\gamma, \alpha)$. The threshold α_s is the minimum value of α which satisfies the inequality

$$\min_{\gamma} \min_{\mathbf{Y}} LLR(\gamma, \alpha) \geq \log \frac{\pi_0}{\pi_1} \tag{5}$$

That is, for any $\alpha > \alpha_s$, H_1 is declared regardless of $\mathbf{Y}$, hence the observations become useless. It can be shown that if $\min_\gamma \min_{\mathbf{Y}} LLR(\gamma, \alpha) < \log \frac{\pi_0}{\pi_1}$, then Pe is always less than π_0. For L sensors using identical m-bit local quantizers, there are a total of $N = \begin{pmatrix} L + 2^m - 1 \\ 2^m - 1 \end{pmatrix}$ potentially distinct LLR values. Sort these LLRs in an ascending order $LLR_{(1)} \leq \cdots LLR_{(n)} \leq LLR_{(n+1)} \leq \ldots \leq LLR_{(N)}$. Let $P_{1(n)}, P_{0(n)}$ denote the likelihood functions under H_1 and H_0 respectively, with each pair corresponding to $LLR_{(n)}$. If $LLR_{(n)} < \log \frac{\pi_0}{\pi_1} \leq LLR_{(n+1)}$, then we have $Pe = \pi_1(P_{1(1)} + \ldots + P_{1(n)}) + \pi_0(P_{0(n+1)} + \ldots + P_{0(N)}) < \pi_0$.

When $m = 1$, it reduces to the binary quantizer case. Since, by the virture of the BSC channel, y can take only two possible values, 0 or 1, we can rewrite (4) as

$$\sum_{l=1}^{L} y_l \quad \underset{H_1}{\overset{H_0}{\lessgtr}} \quad \frac{\log \frac{\pi_0}{\pi_1} - L \log \frac{1-p_d}{1-p_f}}{\log \frac{p_d(1-p_f)}{p_f(1-p_d)}} \triangleq t \tag{6}$$

The condition in (5) becomes the inequality $t \leq 0$. For the binary quantizer case, we can find the closed form solution for the threshold α_s and we summarize it in the following theorem.

Theorem 3: For an L sensor system, the threshold α_s is given by

$$\alpha_s = \frac{\max_s f(s) - 1}{2 \max_s f(s) - 1} \tag{7}$$

where $f(s) = \frac{p'_d(s) - (\frac{\pi_0}{\pi_1})^{\frac{1}{L}} p'_f(s)}{1 - (\frac{\pi_0}{\pi_1})^{\frac{1}{L}}}$, $p'_d(s) = P(U = 1|H_1)$ and $p'_f(s) = P(U = 1|H_0)$. s is the likelihood ratio quantization threshold at the local sensor.

Proof:

$$\begin{aligned} t \leq 0 \quad & \Longrightarrow \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}} \leq \frac{1-p_d}{1-p_f} \\ & \Longrightarrow p_d \leq \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}} p_f + \left(1 - \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}}\right) \end{aligned} \tag{8}$$

Since $p_d = p'_d(1-\alpha) + (1-p'_d)\alpha$ and $p_f = p'_f(1-\alpha) + (1-p'_f)\alpha$, Inequality (8) is equivalent to

$$p'_d \leq \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}} p'_f + \left(1 - \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}}\right) \frac{1-\alpha}{1-2\alpha} \tag{9}$$

It is obvious that if $\alpha \geq \pi_0 (\alpha \leq 0.5)$, (9) is satisfied. Inequality (9) can also be written as

$$\frac{1-\alpha}{1-2\alpha} \geq \frac{p'_d - \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}} p'_f}{1 - \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{L}}} = f(s) \tag{10}$$

Threshold α_s should satisfy $\alpha_s = \frac{\max_s f(s) - 1}{2 \max_s f(s) - 1}$, where $\alpha_s < \pi_0$. Q.E.D.

V. Numerical Examples

In this section, we present some numerical examples to show the impact of channels on the detection performance limits. In particular, we strive to understand how various system parameters, including prior probabilities of the hypotheses, observation statistics, the number of sensors used, and quantization levels at the local sensors, determine when the system becomes useless. Consider the case where the local observations are identically distributed according to $\mathrm{H}_0 : X \sim \mathcal{N}(-\mu, \sigma^2)$ and $\mathrm{H}_1 : X \sim \mathcal{N}(\mu, \sigma^2)$.

Figs. 3 and 4 show the threshold α_s as a function of the number of sensors used by varying the prior probability and observation noise variance respectively. Binary quantizers are used at the local sensors. The threshold α_s curve divides the distributed detection systems into 'useful' and 'useless' categories, which represents the tradeoff between prior and likelihood functions – a system becomes useless when the prior dominates the likelihood functions and renders the channel outputs irrelevant to the detection problem. For example, extreme values of the prior probability (π_0 is very large or small) yields a very small α_s, indicating that only high quality channels can potentially improve the detection performance with an informative prior.

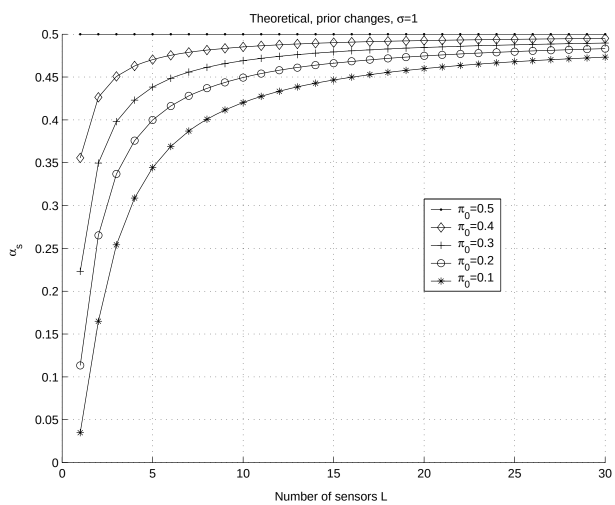


Fig. 3. Theoretical results for threshold α_s as a function of the number of sensors L. The prior probability of H_0 varies from $\pi_0 = 0.1$ to 0.5 for the set of curves.

Fig. 5 shows α_s as a function of the number of sensors L with different prior probabilities and $m = 2$. The threshold α_s increases compared with that of $m = 1$. By using multibit quantization, the range of useful α values increases, thus making previously useless system useful again.

For the non-asymptotic analysis, the same BSC is used for all the local sensors. However, if the environment is non-homogeneous, the channel parameters may be different for different sensors. Fig. 6 is an example that two types of BSC exist with crossover probabilities α_1 and α_2 respectively. Instead of separate thresholds for each paramter, there is a curve dividing the whole $\alpha_1 - \alpha_2$ plane into two parts, with the lower left region indicating the usefulness of the system.

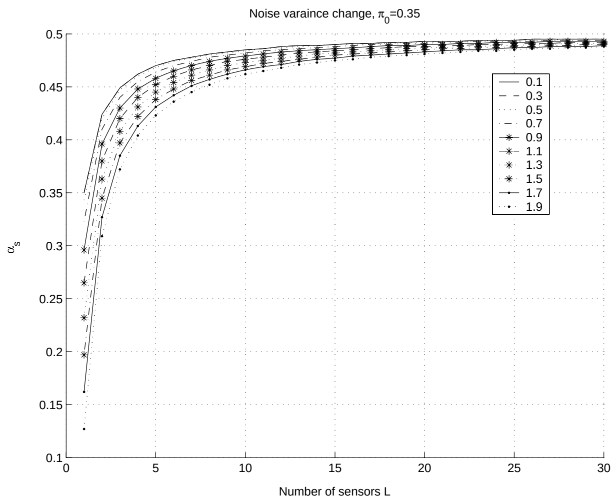


Fig. 4. Theoretical results for threshold α_s as a function of the number of sensors L. The prior probability of H_0 is $\pi_0 = 0.35$, and the noise standard deviation varies from 0.1 to 1.9.

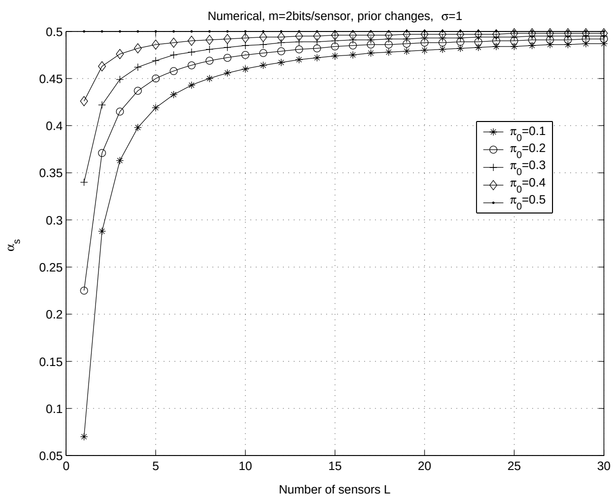


Fig. 5. Numerical results for threshold α_s as a function of the number of sensors L. The prior probability of H_0 varies from 0.1 to 0.5, the quantization bits is $m = 2$bits.

Compared with the results for the identical BSC case, this example again shows that individual sensors and channels might be useless, however, diversity can help make the system useful again.

VI. Conclusions

In this correspondence, distributed detection systems with imperfect channels from the local sensors to a fusion center are considered. It is found that, in the asymptotic regime, there is always performance gain from the output of the system if the channel itself has non-zero capacity. However, it was shown that with a finite number of sensors, observing the channel output cannot help in reducing the error probability performance at the fusion center under certain conditions, regardless of channel-aware optimal system design. Increasing the number of sensors or the quantization levels at local

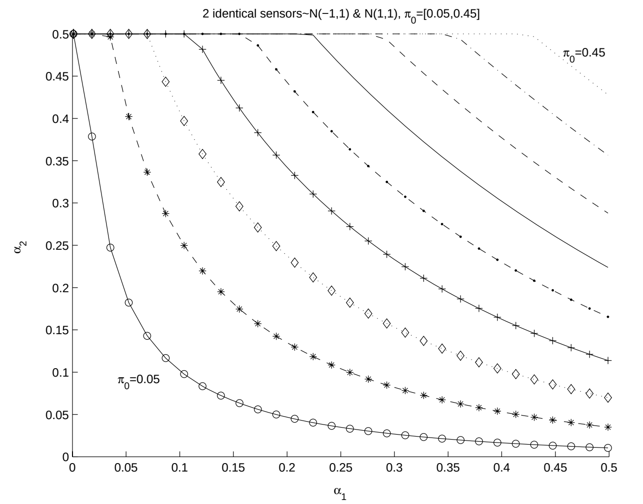


Fig. 6. Two identical sensors are connected with the fusion center by two different BSCs with respective crossover probability α_1 and α_2. The prior probability of H_0 varies from 0.05 to 0.45. The lower left region against thresholds is the useful parameter region.

sensors can help relax the requirements on channel quality thus providing a meaningful way for enhanced channel usage.

Appendix
Proof of *Theorem 1*

The CI is given by $\mathbf{CI}_c = -\min_{0\le s\le 1}\min_\gamma \log_2\left[\sum_{y=1}^{D_y} P_0^s(y)P_1^{1-s}(y)\right]$ [17]. Since $P_k(y) = \sum_{u=1}^{D_u} P(y|u)P_k(u), k = 0, 1$, Eq. (1) is easy to be derived.

To show $rank(\mathbf{M}) = \mathbf{1} \Longrightarrow \mathbf{CI}_c = 0$, notice that $rank(\mathbf{M}) = \mathbf{1}$ implies that all columns of M are the same, i.e., $P(y|u) = P(y)$, which, when plugged into Eq. (1), can readily show $\mathbf{CI}_c = -\min_{0\le s\le 1}\min_\gamma \log_2\left(\sum_{y=1}^{D_y} P(y)\right) = 0$. To show the *only if* part, i.e., $\mathbf{CI}_c = 0 \Longrightarrow P(y|u) = P(y)$ for any (y, u) pair, we note

$$\begin{aligned}\mathbf{CI}_c = 0 \quad &\Longrightarrow \quad \sum_{u=1}^{D_u} P(y|u)P_0(u) = \sum_{u=1}^{D_u} P(y|u)P_1(u) \\ &\Longrightarrow \quad \sum_{u=1}^{D_u} P(y|u)(P_0(u) - P_1(u)) = 0 \qquad (11)\end{aligned}$$

for $y = 1, \ldots, D_y$. The above equation is true for *any* vector

$$\mathbf{p}_{0-1} = [P_0(1) - P_1(1), \ldots, P_0(D_u) - P_1(D_u)]^T \qquad (12)$$

which satisfies $\sum_{u=1}^{D_u}(P_0(u) - P_1(u)) = 0$. Thus, there exists a rank $D_u - 1$ matrix $\mathbf{P_{0-1}}$ that spans the whole space of these vectors, i.e., for any $\mathbf{p}_{0-1}$ in (12), there exists a vector $\mathbf{a}$ such that $\mathbf{p}_{0-1} = \mathbf{P}_{0-1}\mathbf{a}$. Therefore, from Eq. (11), we have $\mathbf{MP}_{0-1} = 0$. Since for any two matrices $\mathbf{A}$ and $\mathbf{B}$, $rank(\mathbf{A})+rank(\mathbf{B})-n \le rank(\mathbf{AB})$, where n is the number of rows in $\mathbf{B}$, we have $rank(\mathbf{M}) + rank(\mathbf{P}_{0-1}) - D_u \le rank(\mathbf{0}) = 0$. Therefore $rank(\mathbf{M}) = 1$. Q.E.D.

REFERENCES

[1] P. K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, 1997.

[2] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors: Part I–Fundamentals," *Proceedings of the IEEE*, vol. 85, pp. 54–63, Jan. 1997.

[3] R. S. Blum, S. A. Kassam and H. V. Poor, "Distributed detection with multiple sensors: part II–Advanced topics," *Proceedings of the IEEE*, vol. 85, pp. 64–79, Jan. 1997.

[4] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 98-101, Jan. 1986.

[5] A. Reibman and L. Nolte, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 24-30, Jan. 1987.

[6] I. Hoballah and P. Varshney, "Distributed Bayesian signal detection," *IEEE Trans. Inform. Theory*, vol. 35, pp. 995-1000, Sept. 1989.

[7] S. Thomopoulos, R. Viswanathan, and D. Bougoulias, "Optimal distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 761-765, Sept. 1989.

[8] J.N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems*, vol. 1, pp. 167-182, 1988.

[9] P. Chen and A. Papamarcou, "New Asymptotic Results in Parallel Distributed Detection," *IEEE Trans. Information Theory*, vol. 39, no. 6, pp. 1847–1863, Nov. 1993.

[10] S.C.A. Thomopoulos and L. Zhang, "Distributed decision fusion with networking delays and channel errors," *Information Science*, vol. 66, pp. 91-118, Dec. 1992.

[11] B. Chen, R. Jiang, T. Kasetkasem, and P.K. Varshney, "Channel aware decision fusion for wireless sensor networks," in *IEEE Trans. Signal Processing*, vol. 52, pp. 3454–3458, December 2004.

[12] R. Niu, B. Chen, and P.K. Varshney, "Decision fusion rules in wireless sensor networks using fading statistics," in *Proc. 37th Annual Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 2003.

[13] J. Chamberland and V.V. Veeravalli, "Asymptotic results for detection in power constrained wireless sensor networks," in *Proc. IEEE Intl. Symp. Inform. Theory (ISIT2003)*, Yokohama, Japan, June-July 2003.

[14] J. Chamberland and V.V. Veeravalli, "The impact of fading on decentralized detection in power constrained wireless sensor networks," in *Proc. IEEE Intl. Conf. Acoust. Speech, Sig. Proc. (ICASSP2004)*, Montreal, Canada, May 2004.

[15] B. Chen and P.K. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," in *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 693–699, Feb. 2005.

[16] B. Liu and B. Chen, "Joint source channel coding for distributed sensor networks," *Proc. 38th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2004.

[17] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, John Wiley, 1991.

[18] J. N. Tsitsiklis, "On threshold rules in decentralized detection", in *Proc. 25th IEEE Conf. Decision and Control*, vol. 1, pp. 232-236, Athens, Greece, 1986.