Biao Chen, Lang Tong, and Pramod K. Varshney

Distributed Signal Processing in Sensor Networks

© IMAGESTATE

Channel-Aware Distributed Detection in Wireless Sensor Networks

The integration of wireless channel conditions in algorithm design

In a distributed detection (DD) system, multiple sensors/detectors work collaboratively to distinguish between two or more hypotheses, e.g., the absence or presence of a target. While DD can be traced back to the advent of democracy and associated voting schemes, one of its earliest formal treatments can be found in the work of Radner in the early 1960s, [1] who considered the problem of decision making by a team of multiple persons. Each person has access to different information and independently makes his/her decision. The coupling (i.e., the concept of a "team") lies in the fact that the payoff function of the decision-making process depends on all the decisions and the state of situation in an inseparable way.

1053-5888/06/$20.00©2006IEEE

A more prevailing model for DD, as studied extensively in the DD literature and arguably more relevant to engineering applications, is a system involving both distributed sensors *and* a fusion center. The fusion center is responsible for the final decision-making task based on information gathered from local sensors. A distinct feature that makes DD both challenging and meaningful is that local sensor observations need to be compressed before they are jointly processed by a fusion center. This need may arise due to the large data volume observed at local sensors as well as the limited channel capacity between sensors and the fusion center. If, on the other hand, the raw data observed at local sensors are accessible in their entirety at the fusion center, the problem is reduced to the classical hypothesis testing (HT) problem [2] at the fusion center with multiple data samples supplied by multiple sensors.

The recent emergence of wireless sensor networks (WSNs) has added a new dimension to DD system design. While the need for distributed data compression is still the dominant issue, during system design and, in particular, the development of signal processing algorithms, one is confronted with both the stringent resource constraints imposed by a typical WSN as well as the unreliable transmission channels among the sensors and between the sensors and the fusion center. The focus of this tutorial is to revisit the classical decentralized detection theory in the light of these new constraints and requirements. The central theme that transcends various aspects of signal processing design is that an integrated channel-aware approach needs to be taken for optimal detection performance given the available resources.

A BRIEF REVIEW OF CLASSICAL DD

Closely related to DD is the so-called multiterminal inference problem first introduced by Berger in 1979 [3]. There, the standard HT problem was put in the information theoretic framework where multiple remote terminals are subject to rate constraints in communicating their information to the decision maker. The primary goal is to establish existence theorems and bounds on the error exponent for multiterminal HT as a function of the rate constraint [4]–[6]. The results, however, are exclusively established in the asymptotic regime and do not lend themselves easily to practical design procedures to achieve the desired error exponent. (Here the asymptotics are in the temporal domain for a fixed number of sensors. This differs from the "large system" asymptotics where the number of sensors grows to infinity, as in [7].) For a more comprehensive survey in this area, the readers are referred to [8] and the references therein.

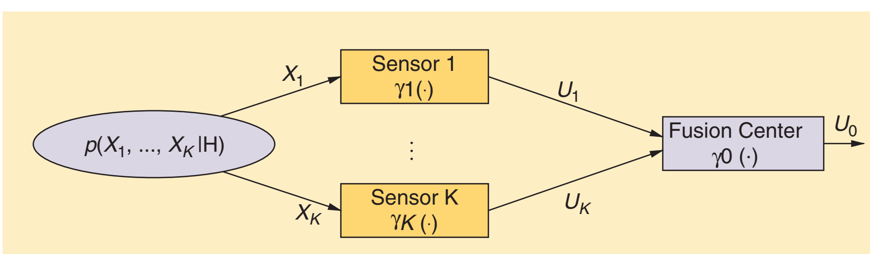
The inception of DD, as we are familiar with today in the signal processing community, can be largely attributed to the seminal work of Tenney and Sandell in 1981 [9]. Since then, DD has blossomed into a research discipline with broad appeal to researchers in the signal processing community [10]–[13]. We consider in this tutorial primarily the so-called canonical DD system, consisting of a fusion center and parallel distributed sensors communicating directly with the fusion center. This is illustrated in Figure 1.

From the signal processing perspective, two inherently different problems need to be considered for the canonical DD system: the design of the decision rule at the fusion center (often referred to as the fusion rule), which strives for an optimal system performance using compressed input from distributed sensors, and the design of local sensor signal processing algorithms. These two problems are intertwined with each other; they need to be jointly designed to optimize a prescribed performance criterion.

The design of fusion rules is conceptually straightforward—with perfect knowledge of system parameters, the optimum fusion rule amounts to a likelihood ratio test (LRT) and has been obtained for both binary and multibit (soft) local sensor outputs [11], [14]. Obtaining local sensor decision rules is considerably more complicated because of its distributed nature. Many examples can be found in the literature that highlight both the difficulty of the problem and the counterintuitiveness of some results. For example, under the Bayesian criterion, it was demonstrated that even with identical sensor observations, the optimal local sensor decision rules need not be identical [15]. For the Neyman-Pearson (NP) problem, nonconcavity of the receiver operating characteristics curves has been observed [10], [16] for certain distributions of sensor observations.

Nonetheless, thanks to the collective effort of many researchers, tremendous progress has been made over the years. For example, under various problem settings and different criteria, the optimality of LRT at the local sensors has been established under the conditional independence assumption [9], [17]–[19]. Relaxing the binary local sensor output assumption to multilevel quantization, the optimality of LR quantization has also been established [10], [20], [21].

Establishing the optimality of LRT at local sensors does not completely solve the problem. Unlike an isolated detection system, the LRT thresholds at the sensors are coupled with each other; they affect the system performance in an interdependent manner. This is true regardless of the detection paradigm employed, i.e., Bayesian, NP, or other more heuristic criteria (e.g., maximizing the Bhattacharyya distance [22]). Almost



[FIG1] A canonical DD system. The kth local sensor observes X_k and sends its decision $U_k = \gamma_k(X_k)$ to the fusion center. The fusion center makes a final decision U_0 regarding H using the fusion rule $\gamma_0(U_1, \ldots, U_K)$. The classical DD system assumes that local sensor outputs $U_1, \ldots, U_K$ are reliably received at the fusion center.

invariably used for finding the local sensor thresholds is the so-called person-by-person optimization (PBPO) approach, where each sensor's threshold is optimized assuming fixed decision rules at all other sensors and the fusion center [23]. The convergence of PBPO is guaranteed as the error probability is lower bounded by zero.

To circumvent the need for the PBPO approach, which becomes intractable when the size of the sensor network becomes large, Tsitsiklis has pioneered the use of error exponents in a decentralized detection system [7]. Using large deviation theory, it was shown that in the asymptotic regime (i.e., the number of sensors grows large), an identical decision rule at the sensors can be used that achieves the optimal error exponent for binary HT, provided that the local sensor observations are conditionally independent and identically distributed. This work, and in particular, the use of error exponent, has also had profound impact on some recent work in the area of detection for WSN [24]–[35]. For example, the optimality of identical sensor-level decision rules in the asymptotic regime was later generalized in the context of WSN [25]–[27]. Other works that resort to the large deviation theory for large-scale sensor networks include the development of universal detection scheme in the absence of knowledge of noise statistics [28], [29], the impact of bandwidth constraint in large-scale sensor networks [24]–[30], and type-based detection schemes that have been shown to attain the best achievable error exponent [32]–[35].

WHY CHANNEL-AWARE SYSTEM DESIGN?

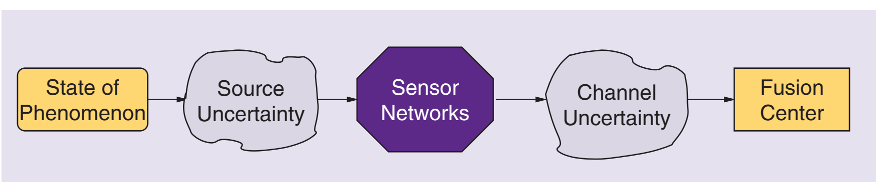
In WSN, the design of a DD system faces the challenge of dealing with an interference-rich transmission environment and stringent resource constraints, most notably the energy constraint with in-situ unattended sensor nodes operating on irreplaceable battery supply, and the delay constraint for a system engaged in a situation awareness mission.

From a system perspective, decision making in an inference-centric WSN is affected by two levels of uncertainty: the first level accounts for the disturbance and noise in the sensor observations, i.e., the uncertainty associated with the source phenomenon as observed by the sensors; the second level of uncertainty is due to the transmission channels between the sensors and the fusion center that are typically affected by receiver noise, channel fading, and interference. This concept is illustrated in Figure 2.

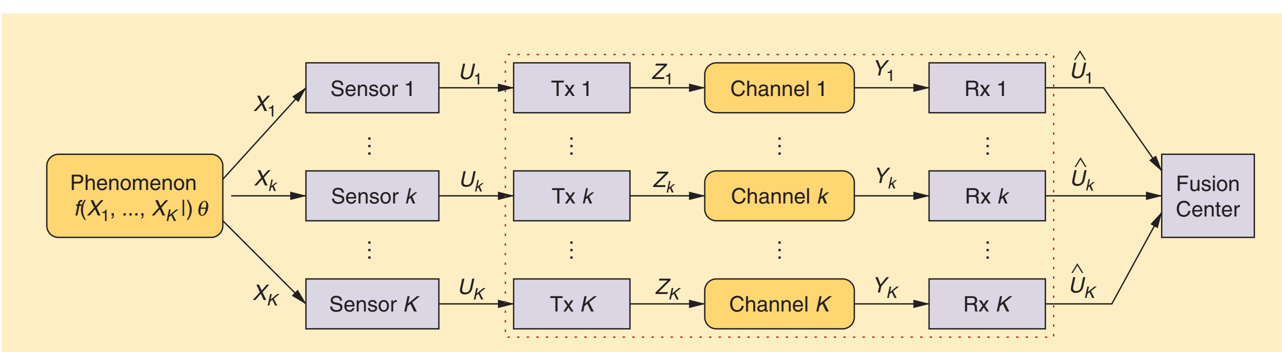
Classical DD addresses only the first layer of uncertainty, that is, sensor signal processing algorithms are designed under the premise that the transmission between the sensors and the fusion center is always reliable. Directly applying classical DD theory to sensor networks with nonideal transmission channels leads to a separation approach: the designs of communication schemes between the sensors and the fusion center are disconnected from the signal processing algorithms for the underlying detection problem.

To elaborate further, consider Figure 3 that depicts a canonical parallel fusion system with θ being the underlying parameter of inference interest. We point out here the difference between data-centric and inference-driven approaches. In the former, the objective is to recover X_1 through X_K while the latter attempts to draw inference regarding θ. This distinction will significantly affect the signal processing design at all levels of the network. Accentuated using the dotted box is the communication block. In resource and delay constrained applications, an important question is this: considering all the constraints imposed by the practical application (e.g., power, bandwidth, delay), is it *good enough* to design the wireless communication system under these constraints, *irrespective* of the sensing/processing at both the local sensor and fusion center level (or in general, the inference task at hand)?

The answer to this question is no, which can be illuminated by examining the decision fusion problem. If the communication block is designed independently of signal processing,



[FIG2] An illustration of a sensor network with two layers of uncertainty, one contributed by the observation uncertainty and the other introduced due to the nonideal transmission channels.



[FIG3] A block diagram for a WSN tasked with inference problems regarding θ. Treating the communication block (dotted box) as an independent component leads to the separation approach.

then each receiver, Rx k in Figure 3, is tasked with the inference problem regarding U_k, and the fusion center would simply process the receiver output $\hat{U}_k$. However, for a fusion network, the ultimate goal is to evaluate the underlying parameter θ instead of recovering $U_1, \ldots, U_K$. The data processing inequality [36] mandates that one should jointly process the output of the channels, namely $Y_1, \ldots, Y_K$, rather than processing the output of the receivers to avoid any potential information loss. This illustrates the importance of integrating the transceiver design with the design of the fusion algorithms.

A similar argument can be made for the signal processing algorithms designed at local sensors. For concreteness, we use Example 1 to demonstrate the need for channel-aware distributed signal processing at local sensors for optimal detection performance.

EXAMPLE 1

Consider the detection of a known signal in additive Gaussian noise with two sensors

$$\begin{array}{ll} H_0 & X_k = N_k \\ H_1 & X_k = S + N_k \end{array} \qquad (1)$$

for $k = 1, 2$, where N_1 and N_2 are independent and identically distributed according to $\mathcal{N}(0, \sigma^2)$. Without loss of generality, we assume $\pi_0 = Pr[H_0] = 0.8$, $S = 1$, and $\sigma^2 = 1$. Each sensor makes a binary decision based on its observation X_k

$$U_k = \gamma_k(X_k).$$

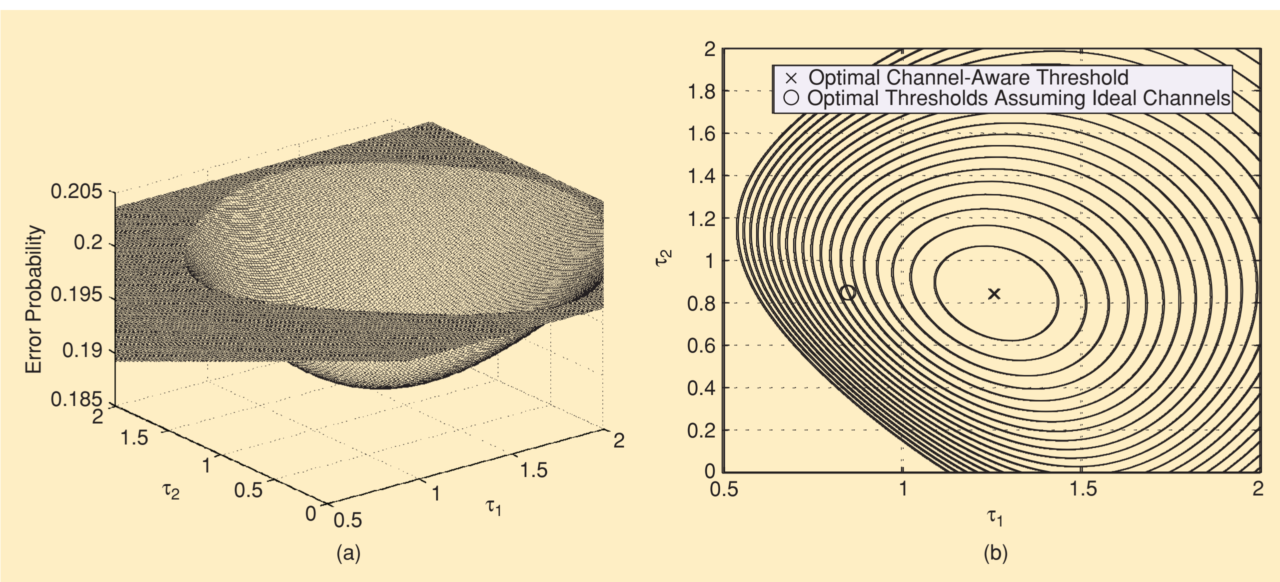
Each U_k is then transmitted through a binary symmetric channel (BSC) with crossover probability α_k, and let us assume $\alpha_1 = 0.05$ and $\alpha_2 = 0.15$. We denote by $\tau_1$and τ_2 the thresholds on X_1 and X_2. For these parameters, the channel-aware design yields a threshold pair $(\tau_1, \tau_2) = (0.8426, 1.2570)$ [37]. The validity of the result is confirmed by Figure 4, where the mesh and contour plots of minimum achievable error probability as a function of (τ_1, τ_2) are given. Figure 4(b) shows that the obtained threshold pair, marked with an "x," achieves the minimum error probability $P_e = 0.1889$.

If one assumes ideal transmission, a pair of identical thresholds $\tau_1 = \tau_2 = 0.8474$ [marked with an "o" in Figure 4(b)] can be obtained using classical DD theory [10], with a corresponding error probability $P_e = 0.1928$. While the degradation from the channel-aware design in terms of error probability is nominal in this case, the penalty of assuming ideal transmission will become more severe when the number of sensors and/or the quantization levels increase [39].

WHEN IS CHANNEL-AWARE PROCESSING NECESSARY?

We conclude this section by looking at a detection example where optimal sensor signal processing does *not* need to be channel aware. Consider the special case of an isolated sensor attempting to distinguish between two hypotheses and sending its decision to the final decision maker via a noisy channel. Assume that the sensor can use a total of m bits for transmission, and it has perfect knowledge of the transmission channel in terms of its conditional probability $p(Y|U)$. This is depicted in Figure 5. Intuitively, if the channel is ideal, only a single bit is needed; the sensor simply implements the optimum detector and sends its decision to the fusion center. Surprisingly, we show that the same holds even in the presence of nonideal channels regardless of the exact channel state $p(Y|U)$.

For this single remote sensor system, the probability of error at the fusion center can be written as, using the fact that $H \to X \to U \to Y \to U_0$ form a Markov chain



[FIG4] The mesh and contour plots of minimum achievable error probability as a function of (τ_1, τ_2) for Example 1. The parameter settings are $\pi_0 = 0.8$, $\alpha_1 = 0.05$, and $\alpha_2 = 0.15$.

$$P_{e0} = \pi_0 P(U_0 = 1|\mathrm{H}_0) + \pi_1 P(U_0 = 0|\mathrm{H}_1) = \int_X \left[\sum_{i=0}^{M-1} P(U = i|X) D_i(X) \right] dX,$$

where $M = 2^m$ is the total number of quantization levels and

$$D_i(X) = \pi_0 p(X|\mathrm{H}_0) P(U_0 = 1|U = i) + \pi_1 p(X|\mathrm{H}_1) P(U_0 = 0|U = i), \quad (2)$$

where $P(U_0 = j|U = i)$ can be computed by

$$P(U_0 = j|U = i) = \int_Y P(U_0 = j|Y) p(Y|U = i) dY.$$

The optimal sensor decision rule is to set

$$P(U = i^*|X) = 1$$

for

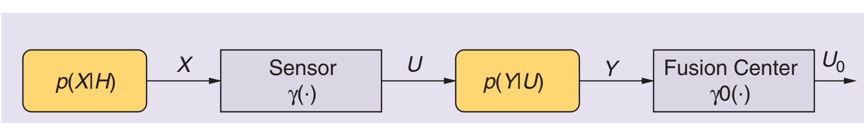
$$i^* = \gamma(X) = \arg \min_{i \in \{0,1,\ldots,M-1\}} D_i(X),$$

i.e., index i^* is what the sensor will transmit. Comparing $D_i(X)$ and $D_j(X)$ with $i \neq j$, we get, from (2),

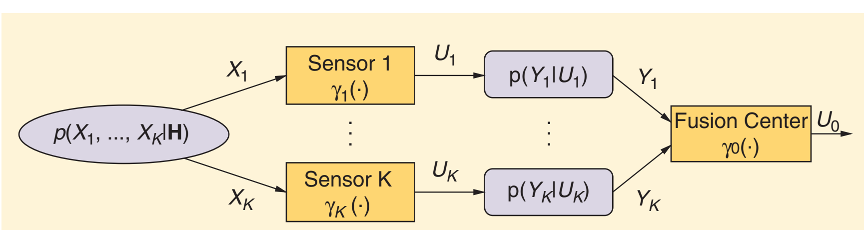
$$D_i(X) < D_j(X) \Longleftrightarrow \pi_1 (P(U_0 = 0|U = i) - P(U_0 = 0|U = j)) \left(\frac{p(X|\mathrm{H}_1)}{p(X|\mathrm{H}_0)} - \frac{\pi_0}{\pi_1} \right) < 0.$$

Notice that the term $\pi_1(P(U_0 = 0|U = i) - P(U_0 = 0|U = j))$ is independent of the sensor observation X. Therefore, even with m-bit signaling, the local sensor signaling is an LR quantizer with only a single threshold π_0/π_1, which corresponds to the optimum Bayesian detection at the remote sensor and is independent of the channel state. It follows from [38] and [39] that only two output indices out of a total of $M = 2^m$ will be used by the quantizer. Given that there are a total of m bits to be used to transmit the quantized output, the optimal encoder structure at the remote sensor is therefore a binary quantizer followed by an $(m, 1)$ code. An intuitive choice is an $(m, 1)$ repetition code, which is optimal, for example, for the Gaussian channel. This structure, as well as the LRT threshold, are independent of the channel state $P(Y|U)$. We remark here that even with the optimal sensor quantization rule being channel blind, the optimal fusion rule $\gamma_0(\cdot)$ is still channel aware.

Unfortunately, this is the only known instance where separation of "source quantization" and "channel encoding" is optimal for HT problems involving remote terminals. In all other cases, including the single-sensor M-ary ($M > 2$) HT problem, or the multiple sensor case, optimal sensor signaling is invariably channel-aware. (This result parallels in some sense that of [40], where it was shown that, for a single remote sensor case, estimation and quantization decouple for optimal quantizer design for a noisy source under a quadratic distortion measure. With multiple sensors, decoupled structure is no longer optimal for the distributed estimation problem with noisy sources [41], [42].) In particular, in a typical WSN where two or more sensors are engaged in the detection problem, channel-aware design always leads to performance improvement under given resource constraints [38]. The need for channel-aware design is also attractive from the energy efficiency viewpoint. To the extent that an improved detection performance may result from an integrated design approach, the transmitter power can be conserved by signaling at a reduced power level.



[FIG5] Detection with a single remote sensor.



[FIG6] A canonical DD system involving the channel layer.

CHANNEL-AWARE SIGNAL PROCESSING FOR DD

At the core of the channel-aware processing concept is the fact that any signal processing design for the underlying detection problem, whether at the fusion center or distributed sensors, needs to include the effect of transmission channels. In a strict sense, the channel-aware approach refers to the integration of the knowledge of the channel state information (CSI) into the design of signal processing algorithms. However, the concept of channel-aware processing discussed in the sequel can be broader than that of simply incorporating CSI into the signal processing design. It also encompasses the design approach that addresses the effect of transmission channels even in the absence of the precise knowledge of CSI. In the following, we categorize various channel-aware signal processing design approaches for detection in sensor networks, each designed specifically under a particular assumption regarding the knowledge of transmission channels.

The basic model to be considered in this section is depicted in Figure 6. As with the classical DD problem, two different problems need to be addressed: the design of a channel-aware fusion rule at the fusion center and the design of distributed signal processing algorithms at the local sensors. The former is conceptually simple: given the local sensor decision rules, the fusion rule design is a centralized detection problem and classical detection theory applies [43]–[46]. The major issue is that the degree of knowledge regarding the channels will affect the transmission schemes and the ensuing fusion rule. For example, without channel phase information, incoherent transmissions have to be used which result in totally different fusion statistics [45], [47], [48] compared with those for coherent systems [43], [44].

[FROM A SYSTEM PERSPECTIVE, DECISION MAKING IN AN INFERENCE-CENTRIC WIRELESS SENSOR NETWORK IS AFFECTED BY TWO LEVELS OF UNCERTAINTY.]

We will focus in the following on the design of *distributed* channel-aware signal processing algorithms. For concreteness, in all the examples presented below, we restrict ourselves to the case of binary HT with $K = 2$ or 3 sensors. Most of the ensuing discussions directly apply to the general K sensor case, although for large K, one is confronted with the issue of increased computational complexity.

WITH COMPLETE CHANNEL KNOWLEDGE

The ideal case for channel-aware design is the genie-aided approach, where the global CSI (i.e., the knowledge about $p(Y_k|U_k)$ for all k) is available to the designer. Particularizing to wireless fading channels, we denote by $\mathbf{h}$ the global CSI. The design criterion with the knowledge of the global CSI can be succinctly summarized as

$$\min_{\gamma_0(\cdot),\ldots,\gamma_K(\cdot)} P_{e0}(\gamma_0, \ldots, \gamma_K; \mathbf{h}) \tag{3}$$

i.e., the decision rules $\gamma_k(\cdot)$, $k = 0, \ldots, K$, are optimized for a given channel realization $\mathbf{h}$. The designer decides the optimal signaling scheme at the local sensors and the optimal decision rule at the fusion center. These schemes need to be broadcast to the sensors and the fusion center through a reliable channel. If the following two assumptions hold:

1) the observations at the local sensors are independent conditioned on any given hypothesis
2) the sensors communicate with the fusion center through independent channels

it is straightforward to show that the set of LRs

$$\left(\frac{p(X_1|\mathrm{H}_1)}{p(X_1|\mathrm{H}_0)}, \ldots, \frac{p(X_K|\mathrm{H}_1)}{p(X_K|\mathrm{H}_0)}\right)$$

form a sufficient statistic for the HT problem at the local sensor level. (Without either of the above two assumptions, channel-aware design becomes computationally challenging. This is especially the case when conditionally dependent observations are observed at local sensors. In this case, the set of local LRs is no longer sufficient for the underlying detection problem and the problem has not been completely understood even under the ideal channel assumption [49].) It follows directly that the optimal local decision rules amount to quantizing the LR and the channel awareness manifests itself in the channel-dependent LR quantization thresholds [37]–[39], [50]. A similar argument can be made for M-ary hypotheses testing though each local sensor quantizer needs to operate on an $M-1$ dimensional sufficient statistic instead of a scalar one [10], [51].

The advantage of the channel-aware approach lies in its inherent adaptivity [38], [39]: the sensor quantizer is not only driven by the observation distribution but is also channel adaptive. This has already been demonstrated in Example 1, where the quantization thresholds using the channel-aware approach are channel dependent, which results in improved detection performance compared with the classical approach. In Example 2, we provide a more thorough comparison to illustrate the inherent adaptivity of the integrated channel-aware approach.

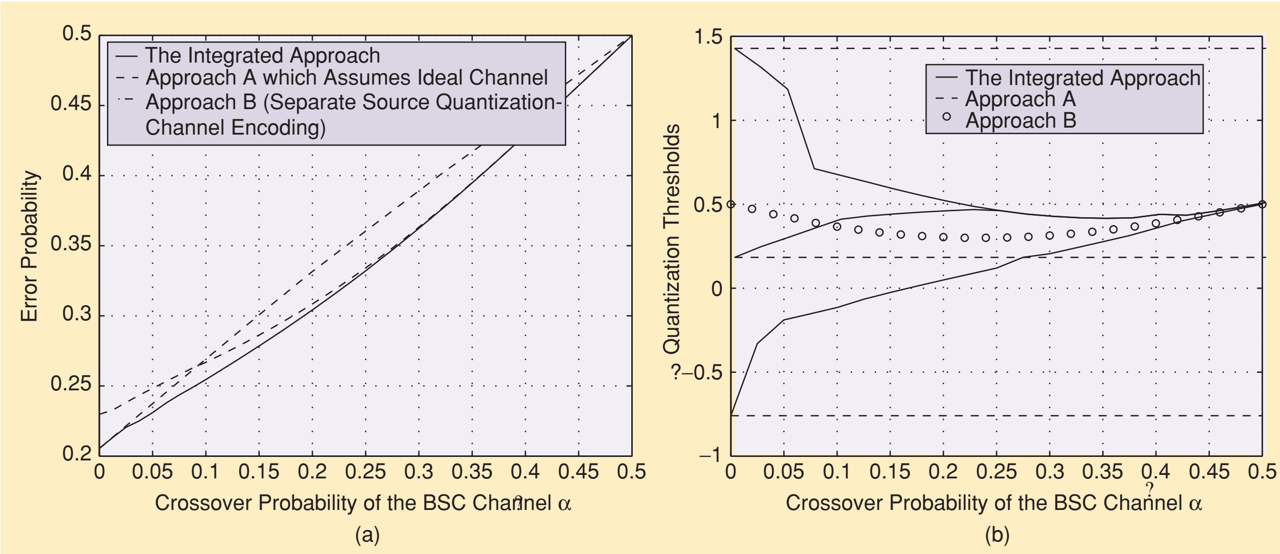
EXAMPLE 2

We again consider the detection of a known signal in independent Gaussian noise as in (1), but with $K = 3$ sensors. Each local sensor employs $m = 2$ bit quantization. We mention here that the performance advantage of the channel-aware approach becomes even more pronounced as the number of sensors and or bits/sensor increase. Limiting to the case of $m = 2$, however, makes our presentation much easier as only three quantization thresholds are needed [c.f. Figure 7(b)]. Each sensor transmits its m bits through m uses of a single BSC with crossover probability α, and α is assumed to be identical for all channels.

The channel-aware distributed quantizer design is compared to the following two alternative approaches:

- Approach A: optimum quantizer design assuming ideal transmission channels; i.e., the m bits are presumed to be fully accessible at the fusion center.
- Approach B: separate source-quantization channel-coding approach. With a total of m bits for each U_k, one can use an n-bit ($n < m$) quantizer followed by a block-channel encoder mapping each n-bit quantizer output to an m-bit code word. Essentially, each $\gamma_k(\cdot)$ in Figure 6 is divided into two parts: a quantizer followed by a channel encoder. In the following example, we set $n = 1$; i.e., a binary quantizer. This is followed by an $(m, 1)$ repetition code which is optimal among all $(m, 1)$ block codes for the BSC as it maximizes the Hamming distance. The binary quantizer threshold is optimized by taking into account the aggregate transmission channel; i.e., using the approach in [37] and [52].

Figure 7(a) plots, for $\pi_0 = 0.5$, the error probability as a function of α for the three different approaches. The integrated approach provides uniformly the best performance among the three approaches. Figure 7(b) gives the thresholds obtained by



[FIG7] Error probability and thresholds plots as a function of channel crossover probability for the three different approaches. The parameters are $\pi_0 = 0.5$ with $K = 3$ and $m = 2$. Hence, three thresholds are needed for the integrated approach and Approach A while Approach B only uses a single threshold.

the three different approaches; these obtained thresholds turn out to be identical for all the three sensors. Notice that the thresholds for Approach A remain constant (i.e., they are channel blind) while the channel-aware approach adapts its thresholds according to the channel parameter. For Approach B, a single threshold is used due to the binary quantizer restriction. An interesting observation is that as α increases, the three thresholds corresponding to the integrated approach will start to merge and eventually become identical to that of Approach B. This is not a coincidence: as the channels become noisier, the integrated approach adaptively puts more emphasis on combating channel impairment and eventually reduces to Approach B—a binary quantizer followed by a repetition code—to provide maximum protection against the noisy channel.

WITH PARTIAL CHANNEL KNOWLEDGE

The clairvoyant case described above has theoretical significance as it provides the best achievable detection performance to which any suboptimal approach needs to be compared. On the other hand, it lacks practical significance due to the requirement of exact knowledge of global CSI. This is further exacerbated by the potential mobility of sensors that leads to fast fading channels: decision rules for all sensors need to be synchronously updated for different channel realizations.

To make the channel-aware design more practical, the reliance on the global CSI in obtaining the local decision rules needs to be relaxed. Instead, partial channel knowledge can be assumed. In the context of WSN, a more reasonable assumption is the availability of channel fading statistics, which may remain stationary during sufficiently long periods of time. As such, decision rules can be updated at a more realistic rate.

Given the knowledge of channel fading statistics, a sensible criterion is to use the *average* error probability at the fusion center where the averaging is performed with respect to the channel state. One can summarize this design criterion as

$$\min_{\gamma_0(\cdot),\ldots,\gamma_K(\cdot)} \int_{\mathbf{h}} P_{e0}(\gamma_0, \ldots, \gamma_K; \mathbf{h}) p(\mathbf{h}) d\mathbf{h}, \tag{4}$$

where $P_{e0}(\gamma_0, \ldots, \gamma_K; \mathbf{h})$ is the error probability at the fusion center for the channel state $\mathbf{h}$ and the set of decision rules $\gamma_k(\cdot)$, $k = 1, \ldots, K$. Here, the sensor decision rule γ_k has only channel fading statistics, instead of the instantaneous channel state $\mathbf{h}$, as its side information.

The above integration is, however, not numerically amenable as $P_{e0}(\gamma_0, \ldots, \gamma_K; \mathbf{h})$ is a highly nonlinear function of $\gamma_0, \ldots, \gamma_K$ and does not yield a closed-form expression. The only possible way of finding the optimal $(\gamma_1, \ldots, \gamma_K)$ appears to be an exhaustive search, which becomes intractable when either K or m becomes large.

An alternative approach has been proposed in [53]. Instead of directly minimizing the average error probability as in (4), one can first average the channel transition probability with respect to the fading channel. That is, the channel probability $p(Y|U)$ in Figure 6 is computed by marginalizing the channel fading:

$$p(Y_k|U_k) = \int_h p(Y_k|U_k, h_k) f(h_k) dh_k. \tag{5}$$

With this marginalization, one can use the channel-aware design approach that tends to the "averaged" transmission channel.

Denote by P_{e0}^0 the average error probability of the channel-aware approach with global CSI, P_{e0}^1 the average error probability of the approach using (4), and P_{e0}^2 the average error probability of the channel-aware approach using the marginalized channel transition probability, the following inequality holds:

$$P_{e0}^0 \leq P_{e0}^1 \leq P_{e0}^2. \quad (6)$$

The first half of (6) arises from the fact that P_{e0}^0 can be written as

$$P_{e0}^0 = \int_{\mathrm{h}} \left[\min_{\gamma_0(\cdot),\ldots,\gamma_K(\cdot)} P_{e0}(\gamma_0, \ldots, \gamma_K; \mathrm{h}) \right] p(\mathrm{h}) d\mathrm{h},$$

and the inequality follows as minimization is taken inside the integration for P_{e0}^0. For the second half of (6), we note that

$$P_{e0}^2 = \int_{\mathrm{h}} P_{e0}(\gamma_0', \ldots, \gamma_K'; \mathrm{h}) p(\mathrm{h}) d\mathrm{h}, \quad (7)$$

where γ_k', $k = 0, \ldots, K$, are a fixed set of decision rules optimized for $p(Y_k|U_k)$ defined in (5). Inequality follows immediately by comparing (4) with (7).

The performance difference among the three different approaches depends on various parameters of the system, including the observation and channel signal-to-noise ratio as well as the fading channel model. An extreme example is the case with no channel fading (e.g., an additive white Gaussian channel between each sensor and the fusion center). In that case, all three approaches are equivalent to each other.

WITHOUT CHANNEL KNOWLEDGE

There are applications where even partial channel knowledge may be untenable or highly unreliable. For instance, high mobility may induce nonstationary fading, i.e., the fading statistics are continuously changing. This renders the previous approach assuming stationary fading channel statistics impractical. Applying channel-aware design to such cases may appear paradoxical: how can one apply channel-aware approach when there is no tangible information regarding the transmission channels available?

The approach presented in [54] is motivated by the multiple description code (MDC) principle that has been widely used to combat transmission loss [55], [56]. As illustrated in Figure 8(a) with two encoders and three decoders, the encoders are so designed that in the case of loss of one of the two channels, the two side decoders (Decoders 1 and 2) are guaranteed to attain certain acceptable performance; while in the case of successful transmission of both sources, the center decoder (Decoder 0) will have enhanced performance. While this principle may be carried over to WSN applications, the distributed nature of the problem makes it drastically different from the conventional MDC. In conventional MDC, the two encoders encode a common source, while in the context of WSN, each encoder encodes its own observations without access to the other's input source. This distributed MD problem is illustrated in Figure 8(b).

Applying the MD principle to the DD problem, one can pose the following constrained minimization problem:

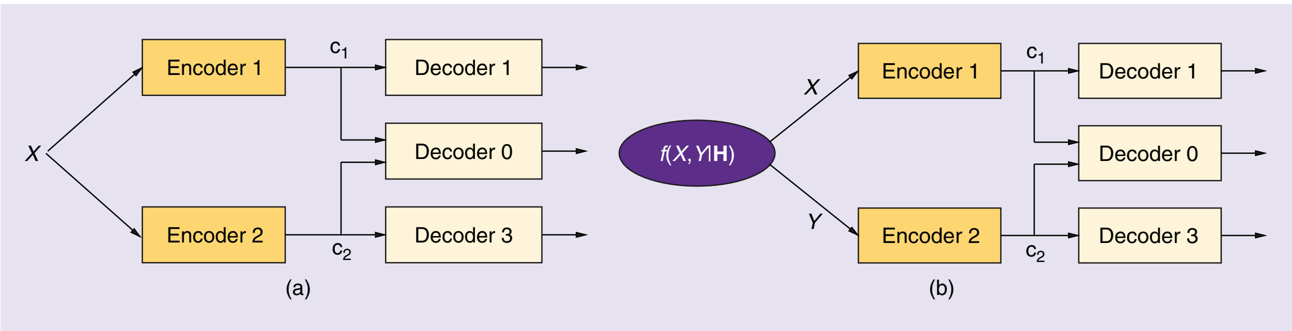
$$\begin{aligned} \min \quad & \pi_0 P(U_0 = 1|\mathrm{H}_0) + \pi_1 P(U_0 = 0|\mathrm{H}_1) \\ \text{subject to} \quad & \pi_0 P(U_i = 1|\mathrm{H}_0) + \pi_1 P(U_i = 0|\mathrm{H}_1) \leq \eta \\ & \text{for } i = 1, 2, \end{aligned} \quad (8)$$

where U_i is the output of Decoder i and the minimization is with respect to $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$, the local sensor decision rules, and $\gamma_0(\cdot)$, the fusion rule at Decoder 0 in Figure 8(b) when both transmissions are successful. Thus we want to minimize the error probability when both transmissions are successful, under the constraint that if only one of the two transmissions is successful, the error probability is no greater than η. The constraints ensure robust detection performance in the presence of potential channel outages. Example 3 gives a "toy" example to motivate the use of the MD principle for DD.

EXAMPLE 3

Assume a binary HT problem with a two-sensor parallel fusion system where each sensor employs a binary quantizer. The two hypotheses under test, H_0 and H_1, are a priori equally likely. The local sensor observations at the two sensors, X_1 and X_2, are conditionally independent and identically distributed ternary random variables with

$$\begin{cases} P(X_k = 0|\mathrm{H}_0) &= 0.95 \\ P(X_k = 1|\mathrm{H}_0) &= 0.05 \\ P(X_k = 2|\mathrm{H}_0) &= 0 \end{cases} \quad \begin{cases} P(X_k = 0|\mathrm{H}_1) &= 0.05 \\ P(X_k = 1|\mathrm{H}_1) &= 0.9 \\ P(X_k = 2|\mathrm{H}_1) &= 0.05 \end{cases}$$



[FIG8] (a) Conventional MD problem and (b) distributed MD for decentralized detection. In the conventional MD problem, the two encoders have access to the common source *X*, which is to be recovered at the decoders; whereas for the decentralized detection problem each encoder encodes its own observations without access to the other, and the ultimate goal is to infer about H instead of recovering the observation *X* and *Y*.

for $k = 1, 2$. By the monotonicity of the LR in the sensor observations (i.e., the local sensor LR values are monotone in X_k), we need to consider only the two binary local decision rules at each sensor [15]:

$$\text{(A)} \quad U_k = \begin{cases} 0 & X_k = 0 \\ 1 & X_k = 1 \text{ or } 2 \end{cases}$$

$$\text{(B)} \quad U_k = \begin{cases} 0 & X_k = 0 \text{ or } 1 \\ 1 & X_k = 2. \end{cases}$$

Adopting the classical DD approach, it is straightforward to show that the two sensors should employ different decision rules to achieve the minimum error probability of 0.04875 at the fusion center. Without loss of generality, we assume that sensor 1 uses (A) while sensor 2 uses (B). If the decision of sensor 1 does not reach the fusion center due to a channel outage, the actual minimum error probability by using the decision from sensor 2 alone becomes 0.475. This is a significant degradation from the case when both sensor outputs are available. This error probability essentially renders the detection system useless as it is close to 0.5. A more robust design is to use decision rule (A) at both the sensors. In this case, both the fusion center and each local sensor have an identical error probability of 0.05, thus there is no degradation in the event of a lost transmission. Compared with the classical DD approach for which error probabilities at the fusion center and the local sensors are 0.04875 and 0.475 respectively, the alternative approach provides a more robust performance in the presence of a transmission loss. This robust solution can be obtained via the constrained minimization formulation for the constraint $0.05 \leq \eta < 0.475$. □

The phenomenon that the optimal sensor decision rules across the sensors may be different even if the sensor observations have identical distributions has been observed for classical DD (see, e.g., [15]). Here we provide an intuitive explanation that will help understand the robustness issue in the event of a transmission loss. Assuming ideal transmission channels, the design of local decision rules can be considered as a special case of distributed source coding with correlated sources, albeit with the ultimate goal of distinguishing the underlying hypothesis instead of recovering the source at the fusion center. Notice that even in the case of conditional independence assumption, the observations at different sensors are still marginally correlated under the Bayesian framework (i.e., the underlying hypothesis is a random variable). For the above case where sensor 1 using decision rule (A) while sensor 2 using (B) constitutes an optimal distributed quantization configuration, one can treat the decision output sent from sensor 2 as side information. This interpretation is cognizant of the fact that threshold rule (A) gives the minimum error probability at a single sensor. Thus in choosing decision rule (B) at sensor 2, the premise is that the information sent from sensor 1 is always available and the detection performance can be maximally enhanced by incorporating the side information sent from sensor 2. In the event of a transmission loss for sensor 1, however, merely sending the side information to the fusion center will result in severely degraded detection performance.

The constrained minimization problem in (8) can be solved using the Lagrange multiplier method. This approach, however, becomes highly intractable for a WSN with more than two sensors. As the number of sensors increases, the number of constraints increases exponentially as one has to take into account all combinations of channel successes/failures. Alternatively, one may treat this as a multiobjective optimization problem with the number of objective functions increasing exponentially. An alternative approach is to impose an erasure channel (EC) model for each channel between the sensor and the fusion center [54], [57]. This EC model was first used in [56] for MD quantizer design for point-to-point communication. For the detection problem, imposing the EC model allows one to collapse the multiple objective functions in the original MD problem into a single error probability criterion. An added advantage is that the EC model enables the direct application of the channel-aware approach [37], [38] in obtaining the sensor thresholds.

SUMMARY

The most challenging issue in decentralized detection for WSNs is the need for distributed data compression at the local sensors. This same challenge is also driving the research efforts for other inference problems in sensor networks. For example, a comprehensive coverage on distributed compression for estimation in sensor networks can be found in [58] where a canonical fusion model, similar to the one considered here, is adopted. Going beyond the parallel fusion model, a graphical model perspective is presented in [59] that provides a unifying framework in treating many of the prototypical applications in sensor networks.

The problem addressed herein and the associated sensor network structure are also similar to that of distributed source coding [60]–[62] or distributed joint source channel coding problems [63], [64]. The distinction lies in that our objective is the inference of the underlying hypothesis as opposed to data recovery. Also related to this work is the so-called CEO (chief executive officer) problem [65]–[68], where multiple agents, each of them subject to a rate constraint in communicating to the CEO, are deployed to observe independently corrupted versions of a common random process. Theoretically achievable fidelity, either in the form of expected Hamming distance for discrete processes or mean square error when extended later to the Gaussian case, as a function of source coding rate was established. While the CEO problem is also concerned with the underlying inference problem, the focus of the present work is on practical signal processing algorithm design *with a strict delay constraint* instead of obtaining performance bounds or achievable error exponents, which implicitly requires long or infinite delays. Various channel-aware distributed signal processing designs were presented under progressively relaxed CSI assumptions. These include the clairvoyant case where global CSI is available for which a channel optimized quantizer design was developed. The other extreme is when no tangible CSI is available. A proactive design approach using the MD principle can be applied for robust detection performance.

While, in principle, all of these techniques can be generalized to large-scale sensor networks, this extension is typically hindered by the prohibitive complexity when the size of the network becomes large. Compromises can be made in terms of performance/complexity tradeoffs by imposing extra conditions on the optimal solutions (e.g., all sensor nodes use an identical decision rule). However, other practical constraints in large-scale sensor networks that can otherwise be neglected in small-scale networks may call for completely different approaches. For example, implicit in the canonical fusion model is the orthogonal transmission assumption where each sensor communicates with the fusion center through an independent channel. This assumption becomes unrealistic for large-scale sensor networks in which medium access control needs to be achieved in a distributed fashion. An alternative to the orthogonal transmission is the general multiple access channel (MAC), which was first considered in [50]. This MAC model was also used in some recent work in decentralized detection and estimation for sensor networks [33]–[35], [69]–[71]. The interaction between the underlying inference problem and various networking issues is examined in [72] through a cross-layer design perspective.

ACKNOWLEDGMENTS

The authors would like to thank the editors and the anonymous reviewers for their comments and suggestions that helped improve this article. This work was supported in part by NSF under Grants ECS-0501534 and NETS-NOSS CNS-0435190, by DoD Multidisciplinary University Research Initiative (MURI) program administered by the Army Research Office under Grant DAAD19-00-1-0352, and by the Air Force Office of Scientific Research under Grant FA9550-06-1-0051.

AUTHORS

Biao Chen (bichen@ecs.syr.edu) received his B.E. and E.E. in electrical engineering from Tsinghua University, Beijing, China, in 1992 and 1994, respectively. From 1994 to 1995, he was with AT&T (China) Inc., Beijing. He received his M.S. in statistics and Ph.D. in electrical engineering in 1998 and 1999, respectively, from the University of Connecticut, Storrs. From 1999 to 2000 he was with Cornell University as a post-doctoral research associate. Since 2000, he has been with Syracuse University, as an assistant professor with the Department of Electrical Engineering and Computer Science. He is on the editorial board of *EURASIP Journal on Wireless Communications and Networking* and a guest editor for a special issue on wireless sensor networks. He is the recipient of an NSF CAREER Award. His area of interest mainly focuses on signal processing for wireless sensor and ad hoc networks and in multiuser MIMO systems.

Lang Tong (ltong@ece.cornell.edu) joined Cornell University in 1998 where he is now a professor. Prior to joining Cornell University, he was on faculty at the West Virginia University and the University of Connecticut. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology. He received the B.E. degree from Tsinghua University, Beijing, P.R. China in 1985 and the Ph.D. degree in electrical engineering from the University of Notre Dame, Indiana, in 1991. He was a postdoctoral research affiliate at the Information Systems Laboratory, Stanford University, in 1991. He is a Fellow of IEEE. He received the Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 best paper award (with M. Dong) from IEEE Signal Processing Society, and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with P. Venkitasubramaniam and S. Adireddy). He also received the Young Investigator Award from the Office of Naval Research. His research is in the general area of statistical signal processing, communication systems and networks, and information theory.

Pramod K. Varshney received the B.S. degree in electrical engineering and computer science and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign in 1972, 1974, and 1976, respectively. Since 1976 he has been with Syracuse University, where he is currently a professor of electrical engineering and computer science and the research director of the New York State Center for Advanced Technology in Computer Applications and Software Engineering. His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing and remote sensing. He is a Fellow of the IEEE and has received numerous awards. He serves as a distinguished lecturer for the IEEE Aerospace and Electronic Systems Society. He was the 2001 president of International Society of Information Fusion.

REFERENCES

[1] R. Radner, "Team decision problems," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 857–881, 1962.

[2] E.L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Wiley, 1986.

[3] T. Berger, "Decentralized estimation and decision theory," presented at the IEEE 7th Spring Workshop Information Theory, Mt. Kisco, NY, Sept. 1979.

[4] R. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE Trans. Inform. Theory*, vol. 32, pp. 533–542, July 1986.

[5] T.S. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. Inform. Theory*, vol. 33, pp. 759–772, Nov. 1987.

[6] H.M.H. Shalaby and A. Papamarcou, "Multiterminal detection with zero-rate data compression," *IEEE Trans. Inform. Theory*, vol. 38, pp. 254–267, Mar. 1992.

[7] J.N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control, Signals, Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[8] T.S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2300–2324, Oct. 1998.

[9] R.R. Tenney and N.R. Sandell Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, pp. 501–510, July 1981.

[10] J.N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H.V. Poor and J.B. Thomas, Eds. Greenwich, CT: JAI Press, 1993.

[11] P.K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer, 1997.

[12] 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.

[13] R.S. Blum, S.A. Kassam, and H.V. Poor, "Distributed detection with multiple sensors: Part II—Advanced topics," *Proc. IEEE*, vol. 85, pp. 64–79, Jan. 1997.

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

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

[16] Q. Yan and R. Blum, "Distributed signal detection under the Neyman-Pearson criterion," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1368–1377, May 2001.

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

[18] I.Y. Hoballah and P.K. Varshney, "Distributed Bayesian signal detection," *IEEE*

Trans. Inform. Theory, vol. 35, pp. 995–1000, Sept. 1989.

[19] S.C.A. Thomopoulos, R. Viswanathan, and D.K. Bougoulias, "Optimal distributed decision fusion," *IEEE Trans Aerospace Elec. Syst.*, vol. 25, pp. 761–765, Sept. 1989.

[20] D.J. Warren and P.K. Willett, "Optimal decentralized detection for conditionally independent sensors," in *Proc. Amer. Control Conf.*, Pittsburgh, PA, June 1989, vol. 2, pp. 1326–1329.

[21] D.J. Warren and P.K. Willett, "Optimum Quantization for Detector Fusion: Some Proofs, Examples, and Pathology," *J. Franklin Inst.*, vol. 336, pp. 323–359, Mar. 1999.

[22] M. Longo, T.D. Lookabaugh, and R.M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inform. Theory*, vol. 36, no. 2, pp. 241–255, 1990.

[23] Z.-B. Tang, K.R. Pattipati, and D.L. Kleinman, "An algorithm for determining the decision thresholds in a distributed detection problem," *IEEE Trans. Syst., Man Cybern.*, vol. 21, pp. 231–237, Jan./Feb. 1991.

[24] J. Chamberland and V.V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Processing*, vol. 51, pp. 407–416, Feb. 2003.

[25] J. Chamberland and V.V. Veeravalli, "Asymptotic results for detection in power constrained wireless sensor networks," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT'2003)*, Yokohama, Japan, June–July 2003, p. 230.

[26] J. Chamberland and V.V. Veeravalli, "The impact of fading on decentralized detection in power constrained wireless sensor networks," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing (ICASSP'2004)*, Montreal, Canada, May 2004, vol. 3, pp. 837-840.

[27] J. Chamberland and V.V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1007–1015, Aug. 2004.

[28] J. Xiao and Z-Q. Luo, "Decentralized detection in a bandwidth constrained sensor network," in *Proc. IEEE GLOBECOM 2004*, Dallas, TX, Nov. 2004, vol. 1, pp. 123-128.

[29] J. Xiao and Z-Q. Luo, "Universal decentralized detection in a bandwidth constrained sensor network," *IEEE Trans. Signal Processing*, vol. 53, pp. 2617–2624, Aug. 2005.

[30] S.A. Aldosari and J.M.F. Moura, "Detection in decentralized sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Montreal, Canada, May 2004, vol. 2, pp. 277-280.

[31] Q. Cheng, B. Chen, and P.K. Varshney, "Detection performance limit of channel impaired distributed sensor networks," in *Proc. 2004 Annu. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2004, vol. 2, pp. 2213-2217.

[32] K. Liu and A.M. Sayeed, "Asymptotically optimal decentralized type-based detection in wireless sensor networks," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, Montreal, Canada, May 2004, vol. 3, pp. 873-876.

[33] K. Liu and A.M. Sayeed, "Optimal distributed detection strategies for wireless sensor networks," in *Proc. Annu. Allerton Conf. Communication, Control, Computing*, Monticello, IL, Oct. 2004.

[34] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access in sensor networks," in *Proc. SPAWC 2005*, New York, June 2005, pp. 1018-1022.

[35] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *IEEE Trans. Signal Processing*, to be published.

[36] T.M. Cover and J.A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[37] B. Chen and P.K. Willett, "On the optimality of likelihood ratio test for local sensor decisions in the presence of non-ideal channels," *IEEE Trans. Inform. Theory,* vol. 51, pp. 693–699, Feb. 2005.

[38] B. Liu and B. Chen, "Joint source-channel coding for distributed sensor networks," in *Proc. 2004 Annu. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2004, vol. 2, pp. 1397-1401.

[39] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *IEEE Trans. Inform. Theory*, to be published.

[40] Y. Ephraim and R.M. Gray, "A unified approach for encoding clean and noisy sources by means of waveform and autoregressive model," *IEEE Trans. Inform. Theory*, vol. 34, pp. 826–835, July 1988.

[41] W.M. Lam and A.R. Reibman, "Quantizer design for decentralized estimation systems with communication constraints," in *Proc. Conf. Information Sciences Systems*, Baltimore, MD, Mar. 1989, pp. 489–494.

[42] W.M. Lam and A. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. Commun.*, vol. 41, pp. 1602–1605, Nov. 1993.

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

[44] R. Niu, B. Chen, and P.K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 54, pp. 1018–1027, Mar. 2006.

[45] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2668–2673, Nov. 2005.

[46] Y. Lin, B. Chen, and P.K. Varshney, "Decision fusion rules in multi-hop wireless sensor networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, pp. 475–488, Apr. 2005.

[47] R. Jiang and B. Chen, "Decision fusion with censored sensors," in *Proc. IEEE Int. Conf. Acoustic Speech, Signal Processing (ICASSP'2004)*, Montreal, Canada, May 2004, vol. 2, pp. 289-292.

[48] R. Jiang, S. Misra, B. Chen, and A. Swami, "Robust supoptimal decision fusion in wireless sensor networks," in *Proc. IEEE MILCOM'05*, Atlantic City, NJ, Oct. 2005.

[49] P.K. Willett, P.F. Swaszek, and R.S. Blum, "The good, bad, and ugly: distributed detection of a known signal in dependent Gaussian noise," *IEEE Trans. Signal Processing*, vol. 48, pp. 3266–3279, Dec. 2000.

[50] T.M. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 469–476, Apr. 1998.

[51] B. Liu, B. Chen, and R.S. Blum, "Exploiting the finite-alphabet property for cooperative relays," in *Proc. IEEE Int. Conf. Acoustic Speech, Signal Processing (ICASSP'2005)*, Philadelphia, PA, Mar. 2005, vol. 3, pp. 357-360.

[52] B. Chen and P.K. Willett, "Channel optimized binary quantizers for distributed sensor networks," in *Proc. IEEE Int. Conf. Acoustic Speech, Signal Processing (ICASSP'2004)*, Montreal, Canada, May 2004, vol. 3, pp. 845-848.

[53] B. Liu and B. Chen, "Distributed detection in sensor networks with channel fading statistics," in *Proc. 2006 Int. Conf. Acoustic Speech Signal Processing (ICASSP'06)*, Toulouse, France, May 2006.

[54] Y. Lin, B. Chen, and B. Suter, "Multiple description quantizer design for distributed sensor networks," in *Proc. Conf. Information Sciences Systems*, Baltimore, MD, Mar. 2005.

[55] V.K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Mag.*, vol. 19, no. 5, pp. 74–93, Sept. 2002.

[56] V. Vaishampayan, "Design of multiple description scalar quantizer," *IEEE Trans. Inform. Theory*, vol. 39, pp. 821–834, May 1993.

[57] Y. Lin, B. Chen, and B. Suter, "Robust binary quantizers for distributed detection," *IEEE Trans. Wireless Commun.*, to be published.

[58] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G.B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Mag.*, vol. 23, no. 4,pp. 27–41, 2006.

[59] M. Çetin, L. Chen, J.W. Fisher III, A.T. Ihler, R.L. Moses, M.J. Wainwright, and A.S. Willsky, "Distributed fusion in sensor networks," *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 42–55, 2006.

[60] S.S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Mag.*, vol. 19, no. 3, pp. 40–50, Mar. 2002.

[61] Z. Xiong, A.D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag.*, vol. 21, no. 5, pp. 80–94, Sept. 2004.

[62] J. Chen and T. Berger, "Robust coding schemes for distributed sensor networks with unreliable sensors," in *Proc. 2002 IEEE Int. Symp. Information Theory*, Chicago, IL, July 2004, p. 116.

[63] S.D. Servetto, "The reachback channel for wireless sensor networks," in *Proc. DIMACS Workshop Network Information Theory*, Piscataway, NJ, Mar. 2003.

[64] J. Garcia-Frias and Z. Xiong, "Distributed source and joint source-channel coding: From theory to practice," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Philadelphia, PA, Mar. 2005, vol. 5, pp. 1093-1096.

[65] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inform. Theory*, vol. 42, pp. 887–902, May 1996.

[66] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1549–1559, Sept. 1997.

[67] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1057–1070, May 1998.

[68] P. Ishwar, R. Puri, K. Ramchandran, and S.S. Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 765–775, Apr. 2005.

[69] Y. Sung, L. Tong, and A. Swami, "Asymptotic locally optimal detector in large scale sensor networks under the Poisson regime," in *Proc. 2004 IEEE Int. Conf. Acoust. Speech Signal Processing*, Montreal, Canada, May 2004, vol. 2, pp. 1077-1080.

[70] Y. Sung, L. Tong, and A. Swami, "Asymptotic locally optimal detector for large-scale sensor networks under the Poisson regime," *IEEE Trans. Signal Processing*, vol. 53, pp. 2005–2017, June 2005.

[71] G. Mergen and L. Tong, "Type based estimation over multiple access channels," *IEEE Trans. Signal Processing*, vol. 54, pp. 613–426, Feb. 2006.

[72] Q. Zhao, A. Swami, and L. Tong, "The interplay between signal processing and networking in sensor networks," *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 84–93, 2006.