Decentralized Detection with Correlated Observations

Jayakrishnan Unnikrishnan and Venugopal V. Veeravalli
Dept. of Electrical and Computer Engineering
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801
Email: {junnikr2, vvv}@uiuc.edu

***Abstract*—We study the problem of decentralized detection when the observations are dependent conditioned on the hypothesis. The optimal rule for fusing the quantized versions of the observations received from the individual sensors requires complete statistical information about the quantized observations. Since such a fusion rule may not be easy to implement in practice, we propose a suboptimal detector that optimizes a deflection metric in the class of linear-quadratic detectors. Using simulations it is shown that when the observations are correlated, the deflection-optimal detector outperforms the detector that ignores the correlation information completely.**

I. Introduction

A typical decentralized detection setup consists of a group of sensors jointly trying to identify the state of the environment based on their observations. The statistics of the signals observed at the sensors are assumed to be dependent in some way on the state of the environment. The sensors communicate some function of their observations to a fusion center which makes the final decision.

In this paper we study the Neyman-Pearson version of a decentralized binary hypothesis testing problem, in which the state of the environment can be one of two possibilities, H_0 and H_1. The observations at the sensors have distinct distributions under the two hypotheses. The sensors transmit a quantized version of their observations to a fusion center over an error-free channel where the final decision is made. The objective of the test is to maximize the probability of correct detection under H_1 subject to a constraint on the false alarm rate under H_0.

There has been a lot of research in the area of decentralized detection over the past twenty years. However, most of the significant works have focused on problems where the observations at the sensors are assumed to be conditionally independent conditioned on the hypothesis (see for example [1], [2], [3] and [4] for an overview of these results). The correlated case has also been studied ([5], [6], [7]) but the results are often not very easy to implement in practice.

The conditional independence assumption which is well justified in some cases makes the decentralized detection problem a lot more tractable than the one where the observations are conditionally dependent. However, scenarios with dependent observations also do occur in practice, for instance in distributed wireless reception where the signals at the wireless receivers may be correlated due to shadowing.

A complete solution to a decentralized detection problem should address two issues, one of quantizer design at the sensors and one of the decision rule to be used at the fusion center. There have been many works on criteria for optimal quantizer design [7] [8]. It is known that in the conditionally *i.i.d.* case, using identical monotone likelihood ratio quantizers is asymptotically optimal when the number of sensors is large [8]. For the correlated case, however, the results of [7] have suggested that design of optimal quantizers is not a straightforward problem even when there are just two sensors. It is known that [9] the optimal quantizer design problem is, in general, NP-complete.

The problem of designing the optimal fusion rule has also been studied in many works [5] [6]. It is known that the optimal fusion rule is to compute the joint likelihood ratio of the bits and compare it with a threshold chosen so as to meet the false alarm rate requirement. But this solution typically requires the knowledge of the joint statistics of the quantized observations under both hypotheses. In many practical cases, however, obtaining complete statistical information of correlated quantized observations may not be feasible. This is true even for the popular Gaussian observation model.

In this paper, we attempt a suboptimal solution to the decentralized detection problem with conditionally dependent observations when the fusion center has access to only partial statistical information about the quantized observations in the form of some lower order moments. We do not address the issue of quantizer-design and focus on obtaining a good fusion rule that makes use of the available partial statistical information. Using deflection as a performance criterion, we solve for the optimal linear-quadratic detector that maximizes the deflection between the two distributions. This solution requires only the knowledge of up to the fourth order moments under one hypothesis and up to the second order moments under the other.

In the next section we introduce the problem statement and our proposed solutions. In the subsequent sections we present our simulation results and conclusions.

This research was supported by the NSF under award CCF 0049089, through the University of Illinois at Urbana-Champaign and by a Vodafone Foundation Graduate Fellowship.

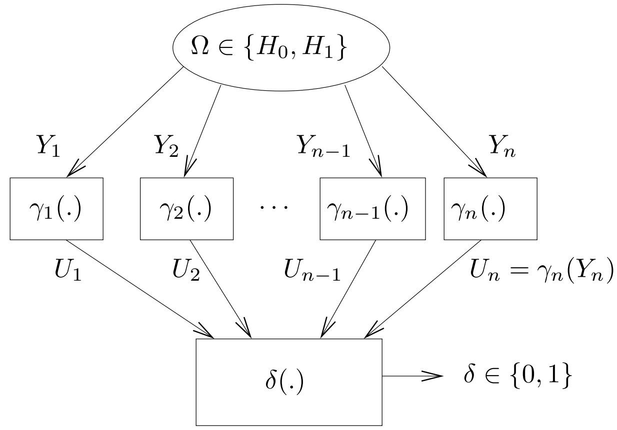

Fig. 1. Decentralized detection setup

II. Problem statement and proposed solutions

The state of the environment could be one of two hypotheses: H_0 and H_1. There are n sensors whose observations are denoted by random variables Y_i, $i = 1, 2, \ldots n$. The sensors quantize their observations to obtain quantized observations $\{U_i\}_1^n$ which are transmitted over an error-free channel to a fusion center which then makes the decision $\delta(\underline{U})$ by computing some function of the received vector of quantized observations $\underline{U}$ and comparing it with some threshold τ_f. The objective is to maximize the probability of making a correct decision under H_1 subject to a bound on the probability of making an error under H_0. The decentralized detection setup is illustrated in Fig. 1.

As mentioned earlier, in this work we do not address the issue of quantizer design and try to come up with a good solution to the fusion problem at the fusion center. For ease of illustration, in the rest of the paper, we stick to the special case where the observations at the sensors are identically distributed, though not necessarily independent. Furthermore, we assume that all sensors employ identical binary likelihood ratio quantizers. We use p_j to denote the distributions of the Y_i's, $i = 1, 2, \ldots, n$ and q_j for the distributions of the U_i's, $i = 1, 2, \ldots, n$, under hypothesis H_j, $j = 0, 1$. We also use P_j to denote the probability measure under hypothesis H_j and E_j to denote the expectation operator under hypothesis H_j, $j = 0, 1$. All sensors use the same thresholds τ_s to quantize the log-likelihood ratios of their observations Y_i to obtain bits $U_i = I_{\{\log(L(Y_i)) > \tau_s\}}$ where $L(Y_i) = \frac{p_1(Y_i)}{p_0(Y_i)}$ represents the likelihood ratio of the observations at sensor i and $I_{\{\}}$ represents the indicator function, which takes on value 1 when its argument is true and 0 otherwise. The sensor thresholds τ_s could be set by using some local criterion chosen appropriately for the application of interest. Although the results in this paper are derived under these special scenarios of identical distributions and identical quantizers, they are easily generalizable to arbitrary distributions of the observations and arbitrary quantizers at the sensors.

The joint statistics of the vector of observations $\underline{Y}$ are assumed known. However, the joint statistics of the vector of bits $\underline{U}$ are unknown except for the first few moments. The objective is to come up with a criterion for designing the fusion center decision rule making use of only the knowledge of some lower order moments of the distributions.

A. Counting Rule

One of the simplest suboptimal solutions to the data fusion problem, the Counting Rule [10], is to just count the number of sensor nodes that vote in favor of H_1 and compare it with a threshold. Equivalently, the decision is based solely on the *type* [11] of the received vector of bits. The threshold value is chosen so as to obtain equality in the false alarm constraint under H_0. It has to be set using simulations since the joint statistics of the U_i's under H_0 are not available. It is easy to see that under the special scenario where the quantized observations are $i.i.d$ under both hypotheses, this is the optimal rule since the joint likelihood ratio of the bits is a function of only the *type* of the received bit vector. Thus this would be a reasonable thing to do even when nothing is known about the correlation structure. How well a rule designed for a decentralized hypothesis test with correlated observations makes use of the correlation information could thus be quantified by comparing its performance with that of the Counting Rule for the same observations. Moreover, the fact that the thresholds for even a simple detector like the Counting Rule detector need to be set using simulations suggest that the same is to be expected for more sophisticated detectors.

B. Linear Quadratic detector

In this section we present the main contribution of this paper - a general suboptimal solution to the fusion problem that uses partial statistical knowledge and gives better performance than the one obtained by ignoring the correlation information completely. This solution makes use of the second order statistics of the quantized observations $\{U_i\}_1^n$ under H_1 and the fourth order statistics under H_0 in the form of moments. The second order moments under H_1 can be calculated by calculating or estimating just the pairwise statistics under H_1. The calculation of the fourth order moments under H_0 is also particularly simple if the observations are independent under this hypothesis. In any case, we note that obtaining information about these moments is in general significantly easier than obtaining the entire joint statistics of the quantized observations especially when there are a large number of cooperating nodes.

We consider detectors in the class of linear-quadratic (LQ) detectors, i.e. detectors that compare a linear-quadratic function of $\{U_i\}_1^n$ with a threshold. Since we are including quadratic terms as well while computing our decision metric, we expect to see improved performance over the Counting Rule that was purely linear. Moreover, since we are using only moment information about $\{U_i\}_1^n$, this detector is quite general and can be used for all classes of distributions of the signals. Since the exact joint statistics are unknown, it is not possible to optimize over the class of LQ detectors using the probability of error criterion. Instead, we optimize the detector using the *generalized signal-to-noise ratio* or *deflection* criterion [12].

The deflection of a detection rule that compares any function $T(\underline{X})$ of the observations $\underline{X}$ with a threshold is defined as:

$$D_T = \frac{[E_1(T(\underline{X})) - E_0(T(\underline{X}))]^2}{\text{Var}_0(T(\underline{X}))} \quad (1)$$

where H_0 is typically the noise-only hypothesis. Although deflection cannot be related directly to the error-probability for non-Gaussian observations, a detector with a higher value of deflection is expected to have better error-probability performance than one with a lower value of deflection. We show using simulations that the optimal deflection-based LQ detector that we derive in this paper gives improved error performance over the Counting Rule in correlated environments.

Following [13], we solve for the optimal LQ detector. Since we now have quadratic terms as well, the values that we assign to the bits become significant. Values of 1 or 0 assigned to the quantized random variables U_i do not make much sense simply because they are not representative of the actual values taken by the signal Y_i. The question then becomes: what values should be assigned to the decision variables?

For a binary hypothesis test involving two Gaussian vector distributions with equal variances and arbitrary covariances, it can be shown that [14] the optimal detection rule is, in general, to compare a linear quadratic function of the observations with a threshold. In this case, the decision metric can be viewed as a linear quadratic function of the log-likelihood ratios of the individual random variables. This observation suggests that an intelligent choice of values to be assigned to the quantized observations in our problem would be the log-likelihood ratios of the bits themselves. Hence we express our decision metric as:

$$T(\underline{X}) = \underline{h}^\top \underline{X} + \underline{X}^\top M \underline{X} \quad (2)$$

where $\underline{X}$ is the vector of log-likelihood ratios of the quantized observations with means under H_0 subtracted, given by

$$X_i = \log\left(\frac{q_1(U_i)}{q_0(U_i)}\right) - E_0\left[\log\left(\frac{q_1(U_i)}{q_0(U_i)}\right)\right]$$

while $\underline{h}$ is a vector of length n and M is an $n \times n$ square matrix. We need to find the optimal LQ metric of the form (2) that maximizes the deflection given by (1). Clearly, this optimization will require the knowledge of up to the second order moments of the quantized observations under H_1 and up to the fourth order moments under H_0 since these terms explicitly appear in the expression for the deflection (1).

Define matrix $C = E_0[\underline{X}\underline{X}^\top]$. Since adding a constant to the decision metric leaves the deflection unchanged, (2) can be replaced by a new decision metric given by:

$$S(\underline{Z}) = \underline{p}^\top \underline{Z} \quad (3)$$

where $\underline{p}$ is now an $(n^2+n)\times 1$ vector and $\underline{Z}$ is an $(n^2+n)\times 1$ vector given by:

$$\begin{aligned}\underline{Z} = [\; X_1 \dots \quad X_n \qquad & X_1^2 - C_{11} \dots \quad X_1X_n - C_{1n} \\ & X_2X_1 - C_{21} \dots \quad X_2X_n - C_{2n} \\ & \qquad \dots \\ & X_nX_1 - C_{n1} \dots \quad X_n^2 - C_{nn} \;]^\top\end{aligned}$$

In other words, we form $\underline{Z}$ by appending $\underline{X}$ with the raster-scanned form of $\underline{X}\underline{X}^\top - C$. So the first n elements of $\underline{Z}$ are the elements of $\underline{X}$, the next n are the elements of the first row of $\underline{X}\underline{X}^\top - C$, followed by the elements of the second row of $\underline{X}\underline{X}^\top - C$ and so on. Similarly, $\underline{p}$ can be viewed as a vector formed by appending vector $\underline{h}$ with matrix M in raster-scanned form. So the problem of finding optimal $\underline{h}$ and M reduces to solving for the optimal $\underline{p}$ that maximizes the deflection for this decision metric.

From the construction of $\underline{Z}$ it is easy to see that $\underline{Z}$ has zero mean under H_0. Hence, applying (1) to (3), we have deflection for $S(\underline{Z})$ given by:

$$D_S = \frac{(\underline{p}^\top \underline{\mu})^2}{\underline{p}^\top K \underline{p}} \quad (4)$$

where $\underline{\mu} = E_1(\underline{Z})$ and $K = E_0(\underline{Z}\underline{Z}^\top)$. Matrix K is a function of the second, third and fourth order moments of X_i's under H_0. In general, matrix K is positive semi-definite and not strictly positive definite. But it can be shown that any vector $\underline{p}$ that drives the denominator of (4) to zero drives the numerator also to zero. This is because if $\underline{p}^\top K \underline{p} = 0$, we have:

$$E_0(\underline{p}^\top \underline{Z}\underline{Z}^\top \underline{p}) = E_0((\underline{p}^\top \underline{Z})^2) = 0$$

whence $\underline{p}^\top \underline{Z} = 0 \; w.p. \; 1$ and therefore,

$$E_1(\underline{p}^\top \underline{Z}) = \underline{p}^\top \underline{\mu} = 0$$

which would mean that the two distributions of $S(\underline{Z})$ have the same mean. Since we do not desire this, it is sufficient to perform the optimization of (4) over those $\underline{p}$ vectors that do not lie in the singular space of matrix K.

Since K is positive semi-definite, it can be diagonalized as $K = V\Lambda V^\top$ where V is a unitary matrix and Λ is a diagonal matrix with non-negative entries. Therefore, (4) can be written as:

$$\frac{(\underline{p}^\top \underline{\mu})^2}{\underline{p}^\top K \underline{p}} = \frac{(\underline{\tilde{p}}^\top \underline{\tilde{\mu}})^2}{\underline{\tilde{p}}^\top \Lambda \underline{\tilde{p}}} \quad (5)$$

which can be reduced to

$$D_S = \frac{(\underline{\widetilde{p_a}}^\top \underline{\widetilde{\mu_a}})^2}{\underline{\widetilde{p_a}}^\top \Lambda_a \underline{\widetilde{p_a}}} \quad (6)$$

where $\underline{\tilde{\mu}} = V^\top \underline{\mu}$ and $\underline{\tilde{p}} = V^\top \underline{p}$. Equation (6) follows by defining Λ_a as the diagonal matrix containing only the non-zero diagonal elements of Λ and defining $\underline{\widetilde{p_a}}$ and $\underline{\widetilde{\mu_a}}$ as the vectors composed of the corresponding elements of $\underline{\tilde{p}}$ and $\underline{\tilde{\mu}}$ respectively. Hence we just have to optimize over $\underline{\widetilde{p_a}}$.

Now (6) can be written as:

$$\begin{aligned} D_S &= \frac{(\underline{\widetilde{p_a}}^\top \Lambda_a^{1/2}\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}})^2}{(\underline{\widetilde{p_a}}^\top \Lambda_a^{1/2}\Lambda_a^{1/2}\underline{\widetilde{p_a}})} \\ &= \frac{((\Lambda_a^{1/2}\underline{\widetilde{p_a}})^\top(\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}}))^2}{(\Lambda_a^{1/2}\underline{\widetilde{p_a}})^\top(\Lambda_a^{1/2}\underline{\widetilde{p_a}})} \\ &= \left(\frac{(\Lambda_a^{1/2}\underline{\widetilde{p_a}})^\top(\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}})}{|(\Lambda_a^{1/2}\underline{\widetilde{p_a}})|}\right)^2 \end{aligned} \quad (7)$$

The right hand side of (7) is of the form $(\frac{x^\top y}{|x|})^2$ which is maximized when the vectors x and y are collinear. It follows that the optimal $\underline{\widetilde{p_a}}$ satisfies:

$$(\Lambda_a^{1/2}\underline{\widetilde{p_a}}) = k\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}} \tag{8}$$

where k in (8) is any constant which is chosen to be unity to give:

$$(\underline{\widetilde{p_a}})_{opt} = \Lambda_a^{-1}\underline{\widetilde{\mu_a}} \tag{9}$$

Since Λ_a contains only the positive eigenvalues of Λ as diagonal elements it is non-singular and (9) is well-defined. Hence the optimal decision metric from (3) has the form:

$$\begin{aligned} S_{opt}(\underline{Z}) &= \underline{p_{opt}}^\top \underline{Z} = \underline{\widetilde{p_{opt}}}^\top \underline{\tilde{Z}} \\ &= (\underline{\widetilde{p_a}})_{opt}^\top \underline{\widetilde{Z_a}} \\ &= \underline{\widetilde{\mu_a}}^\top \Lambda_a^{-1} \underline{\widetilde{Z_a}} \end{aligned} \tag{10}$$

where $\underline{\tilde{Z}} = V^\top \underline{Z}$ and $\underline{\widetilde{Z_a}}$ is obtained by keeping only the terms of $\underline{\tilde{Z}}$ corresponding to those of $\underline{\tilde{\mu}}$ that appear in $\underline{\widetilde{\mu_a}}$, and the optimal deflection obtained is thus:

$$(D_S)_{opt} = \underline{\widetilde{\mu_a}}^\top \Lambda_a^{-1}\underline{\widetilde{\mu_a}} \tag{11}$$

Hence the deflection-optimal LQ detector compares the metric given by (10) to a threshold. Clearly, the computation of $\underline{\tilde{\mu}}$ and K requires only the knowledge of up to the fourth order statistics of the decision variables under H_0 and up to the second order statistics under H_1. Hence the detector based on (10) can be used for all distributions of the original observations as long as these lower order statistics can be calculated.

III. Simulation Results

We compared the performances of the deflection-optimal LQ detector and the Counting Rule detector for a sensor network comprising 9 cooperating sensors placed uniformly in a unit square with separation between nearest neighbors kept at one-half. The sensors employ energy detectors on their received signals and quantize their observations into bits which are then communicated to the fusion center. We assume a log-normal shadowing model and an exponentially decaying correlation function under the signal-present hypothesis. Under the signal absent-hypothesis, the received power is sum of noise and interference powers which we model as being *i.i.d.* across the sensors. After subtracting the mean noise powers from the received signals the resultant signal powers at the sensors in dB have the following distributions under the two Gaussian hypotheses:

$$\begin{aligned} H_0 : \underline{Y} &\sim \mathcal{N}(\underline{0}, \sigma_0^2 I) \\ H_1 : \underline{Y} &\sim \mathcal{N}(\mu\underline{1}, \Sigma) \end{aligned} \tag{12}$$

where $\underline{1}$ and $\underline{0}$ are the vectors with all elements set to 1 and 0 respectively. Parameter μ is the mean received SNR in dB under H_1, σ_0 captures the uncertainty of the noise power, and Σ represents the covariance matrix under H_1 with $\Sigma_{ij} = \sigma_1^2\rho^{d_{ij}}$, where d_{ij} is the separation between sensors indexed by i and j and $\rho < 1$ is some constant parameter that quantifies the amount of correlation present in the environment. Such a model is useful in cooperative sensing of licensed spectrum for cognitive radio applications [15] where the cooperating radios try to detect vacancies in licensed spectrum for unlicensed access.

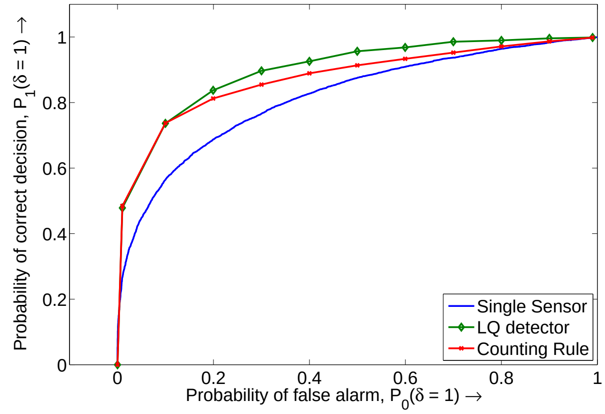

Fig. 2. ROCs for parameters: $\mu = 2.25$, $\sigma_1 = 2$, $\sigma_0 = 1.5$, $\rho = 0.6$

For this problem, we design the the binary quantizers employed at the sensors such that the decisions made by each sensor independently satisfy the false alarm constraint with equality. Since the decision metric at the fusion center is discrete-valued, we allow *randomization* to be used at the fusion center. Since analytical expressions for the error probabilities of the various fusion rules cannot be obtained, we need to resort to simulations for estimating their performances. The performance of a detector can be illustrated using its *receiver operating characteristic* (ROC) [14] which is the plot of the detection probability under H_1 against the false alarm probability under H_0 obtained with the detector. If the ROC of one detector lies above that of another at a particular value of false alarm probability, it means that the former detector performs better than the latter for a Neyman-Pearson test at that level. For the detection problem (12), the detection probability under H_1 is given by $P_1(\delta(\underline{U}) = H_1)$ and the probability of false alarm under H_0 is given by $P_0(\delta(\underline{U}) = H_1)$ where δ represents the final decision about the hypothesis made at the fusion center.

Clearly, we can see from ROCs in Fig. 2 that the deflection-optimal detector provides substantial gains in detection probability over the Counting Rule for a wide range of false alarm rates. In the same figure we have also included the performance obtained with the single sensor detector to illustrate the performance gain obtained by using multiple sensors.

In Fig. 3 we plot the detection performances obtained with various detectors as a function of the correlation parameter ρ for a fixed false alarm probability of 0.2. We keep all the other parameters and sensor configuration the same as in the previous simulation. As we see in the figure, the LQ detector outperforms the Counting Rule detector for all values of ρ greater than 0.4. For low values of ρ the Counting Rule does better than the LQ detector, though not by much. This is because the Counting Rule is the optimal detector when the observations are *i.i.d.* under both hypotheses while the LQ

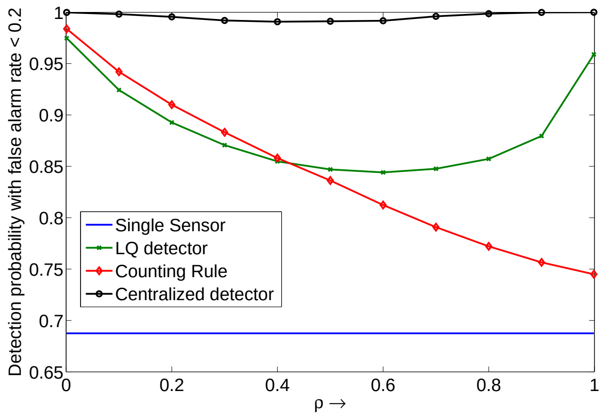

Fig. 3. Performance as a function of the correlation parameter ρ

detector is only a sub-optimal detector. However, for high values of ρ, the LQ detector significantly outperforms the Counting Rule detector thus demonstrating the usefulness of this detector. The figure also shows the performance of the optimal centralized detector that uses the entire unquantized observations $\{Y_i\}_1^n$. It can also be seen that the behavior of the LQ detector mirrors that of the optimal centralized detector in that it shows a similar non-monotonicity in performance as a function of ρ. The Counting Rule, however, fails to capture the improvement in performance for high values of ρ and gives a monotonically decreasing performance with ρ.

IV. Conclusions

We have presented a suboptimal solution to the fusion problem for decentralized detection with correlated observations. Our results clearly indicate that it is not advisable to ignore the correlation information completely. Detection performance can be improved substantially by using a deflection based linear-quadratic detector that requires statistical knowledge of only up to the fourth order under H_0 and second order under H_1. Moreover, since this detector uses only moment information, it can be used for all statistics of the observations, as long as these moments can be computed even when obtaining exact expressions for the joint statistics may be impossible. Hence this detector gives a practical solution to the decentralized detection problem with correlated observations.

Although the results were shown for the special case where all sensors have identically distributed observations and employ identical binary quantizers, this approach is applicable to more general distributions of the signal and general quantizers. Higher level quantizers would, however, increase the complexity of computing the required moment information.

The LQ detector proposed in this paper can also be used for a Bayesian decentralized detection problem with conditionally dependent observations. For such an application the fusion center threshold would have to be chosen to minimize the average probability of error under both hypotheses averaged over their prior probabilities.

References

[1] J. N. Tsitsiklis, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, pp. 297-344, 1993.
[2] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems*, vol. 1, no. 2, pp. 167-182, 1988.
[3] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407-416, 2003.
[4] V. V. Veeravalli and J.-F. Chamberland, "Detection in sensor networks," to appear in *Wireless Sensor Networks. Signal Processing and Communications Perspectives*, A. Swami et al (Eds.), Wiley, 2006.
[5] E. Drakopoulos and C.-C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, no. 4, pp. 5-14, July 1991.
[6] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 916-920, 1992.
[7] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: Distributed detection of a known signal in dependent Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 48, pp. 3266-3279, 2000.
[8] J. N. Tsitsiklis, "Extremal properties of likelihood-ratio quantizers," *IEEE Transactions on Communications*, vol. 41, pp. 550-558, April 1993.
[9] J. N. Tsitsiklis, and M. Athans, "On the complexity of decentralized detection making and detection problems," *IEEE Transactions on Automatic Control*, vol. 30, no. 5, pp. 440-446, 1985.
[10] V. Aalo and R. Viswanathan, "Asymptotic performance of a distributed detection system in correlated Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 40, pp. 211-213, 1992.
[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[12] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 3, pp. 1072-1081, 1995.
[13] B. Picinbono and P. Duvaut, "Optimal linear-quadratic systems for detection and estimation," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 304-311, 1988.
[14] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.
[15] J. Unnikrishnan, and V. V. Veeravalli, "Cooperative sensing and detection for cognitive radio," to be presented in *IEEE Globecom*, Washington D.C., Nov. 2007.