On the Impact of Correlation on Distributed Detection in Wireless Sensor Networks with Relays Deployment

Mohammed W. Baidas[1], Ahmed S. Ibrahim[1], Karim G. Seddik[2], and K. J. Ray Liu[1]

[1] Department of Electrical and Computer Engineering, and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA
(email: baidas, asalah, kjrliu@umd.edu)
[2] Department of Electrical Engineering, Alexandria University, Alexandria 21544, Egypt (email: kseddik@ieee.org)

Abstract—In this paper, a binary hypothesis distributed detection problem in correlated wireless sensor networks with cooperative relays deployment is considered. In particular, the effect of correlation between sensor nodes is modeled and analyzed in Rayleigh flat fading channels in order to explore the natural tradeoffs between the number of sensor/relay nodes and the detection error performance in the network. Specifically, two communication protocols are utilized; in Protocol I, each sensor node communicates its observation directly to the fusion center while in Protocol II, Amplify-and-Forward (AF) cooperative relays are deployed and a fewer number of sensors is used. Based on the theoretical analysis and simulations, it is revealed that employing less sensor nodes and instead deploying relay nodes results in significant performance gains under strict network power constraint. It is concluded that with cooperative distributed detection and exploitation of spatial diversity, better detection error performance is achieved as well as reduction in the required number of sensor nodes.

Index Terms—Amplify-and-Forward (AF), cooperation, correlation, modeling, Rayleigh, relays, sensor networks

I. INTRODUCTION

COOPERATIVE relay networks have recently attracted much attention in the wireless research literature due to their ability to combat fading effects and increase system capacity. Motivated by applications in wireless sensor networks (WSNs) [1]; the idea of cooperation among spatially distributed sensor nodes has gained an increasing interest [2][3]. However, due to the limited power, memory and processing capabilities of such nodes, the design of cooperative communication protocols taking into account different network parameters is particularly attractive in order to achieve efficiency in power consumption and bandwidth utilization and also for accurate detection [4]. On the other hand, due to the spatial correlation of sensor nodes, sending all the local observations from sensors can be inefficient in terms of network resources utilization. This in turn suggests that correlation among sensors can be exploited such that correlated sensors can cooperate and share transmission channels via relays deployment.

In research literature, several variants of the distributed detection problem of correlated observations have been studied. For instance, in [5], the optimal constellation size and number of MIMO cooperating nodes have been analytically quantified with respect to the correlation coefficient where it was shown that optimal constellation size is an increasing function of the correlation coefficient while the optimal number of MIMO cooperating nodes is a decreasing function of the correlation coefficient. The authors in [6] considered the problem of determining whether it is better to use a few high-cost, high power nodes or to use many low-cost, lower power nodes in a system, where the correlated observations are obtained from the nodes distributed uniformly on a straight line. In [7], it was shown that the distributed detection of deterministic signals in

additive Gaussian noise with a set of identical binary sensors is asymptotically optimal as the number of observations per sensor approaches infinity.

In this paper, the problem of distributed detection of correlated observations with relays deployment in the wireless sensor networks is considered. The wireless sensor network performance is studied in Rayleigh fading channels in order to explore the natural tradeoffs between the total number of sensor/relay nodes, degree of correlation and detection error performance. In particular, the main question to be addressed in this paper is: under strict network power constraint, is it better to employ many spatially correlated sensors or to have a fewer number of correlated sensors paired with cooperating relays?

Contributions of this paper include proposing a model suitable for the analysis of wireless sensor networks with correlated observations, along with the effect of path-loss and inter-sensor separation on the detection error performance for arbitrary N number of sensor/relay nodes. Furthermore, the effect of correlation is analyzed in Rayleigh fading channels through two protocols where it will be shown that using fewer sensors with relays can improve the network's detection performance significantly depending on the spatial separation/correlation and also reduces the number of sensor nodes required for the optimal detection performance under strict network power constraint.

In the remainder of this paper, the system model and network protocols are presented in Sections II and III, respectively. In Section IV, the proposed wireless sensor network model is discussed; while in Section V, the theoretical error performance for each network protocol in Rayleigh fading channels is derived. Simulation results and performance evaluation of the network protocols are presented in Section VI. Finally, conclusions are drawn in Section VII.

II. SYSTEM MODEL

In this section, the distributed detection with correlated observations problem is described. A wireless sensor network with N sensor nodes are used to sense a particular phenomenon after which their observations are communicated to the fusion center for decision-making. In this work, dumb sensor nodes are assumed (i.e. they do not perform any local decisions due to their limited capabilities). In addition, the specific case of correlated spatial data $\mathbf{X}$ gathering is modeled using the *Gaussian Random Field* model [8] which is given by the *N-dimensional* multi-variate normal distribution

$$f(\mathbf{X}) = \frac{1}{(2\pi)^{N/2}|\Phi_N|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{X} - \mathbf{m})^H \Phi_N^{-1}(\mathbf{X} - \mathbf{m})\right), \quad (1)$$

where $\mathbf{m} \equiv E[\mathbf{X}]$ is the means vector and $\Phi_N \triangleq E[(\mathbf{X} - \mathbf{m})(\mathbf{X} - \mathbf{m})^H]$ is the covariance matrix of $\mathbf{X}$. The diagonal elements of Φ_N (symmetric and positive-definite) are the measurement noise
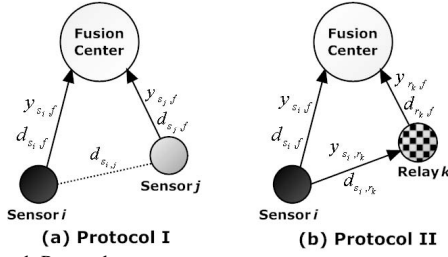
Fig. 1.   Network Protocols

variances of each sensor node $\Phi_{N,ii} = \sigma^2$. The rest of the coefficients in Φ_N (for $i \neq j$) are expressed as $\Phi_{N,i,j} = \sigma^2 \rho_{i,j}$, where $\rho_{i,j} = \exp(-\lambda d_{S_{i,j}}^2)$ is the inter-sensor distance $d_{S_{i,j}}$ dependent correlation coefficient, and λ is a medium-dependent correlation decay factor.

The work under consideration is formulated as a binary hypothesis testing problem where the measured data from the sensor nodes are based on two hypotheses, namely $\mathcal{H}_1$ and $\mathcal{H}_0$, in which the former hypothesis indicates the presence of the phenomenon within the sensor network while the latter indicates its absence. Thus, the vector of measured data $\mathbf{X}$ under each hypotheses is given by

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{X} \sim \mathcal{CN}(0, \Phi_N) \\ \mathcal{H}_1 &: \mathbf{X} \sim \mathcal{CN}(\mathbf{m}, \Phi_N), \end{aligned} \qquad (2)$$

where $\mathbf{X} \triangleq \mathcal{CN}(\mathbf{m}, \Phi_N)$ denotes the complex *correlated* Gaussian random variables.

III. NETWORK PROTOCOLS

In this section, the network transmission protocols are presented. In Protocol I, each sensor transmits its sensed observation directly to the fusion center without employing any intermediate relay nodes; while in Protocol II, sensor nodes are paired with amplify-and-forward (AF) relay nodes deployed closer to the fusion center. The network protocols assume a simple geometric path-loss that is proportional to $1/d^\alpha$ where α is the path-loss exponent and d is the distance between the sensor/relay nodes and the fusion center.

A. Protocol I

In Protocol I (shown in Fig. 1a), each sensor directly transmits its observations to the fusion center where uncoded transmission at each sensor node is assumed. The received signals at the fusion center from two different sensors S_i and S_j are expressed as

$$y_{s_i,f} = \sqrt{P_{s_i} d_{s_i,f}^{-\alpha}} h_{s_i,f} x_i + n_{s_i,f}, \qquad (3)$$

and

$$y_{s_j,f} = \sqrt{P_{s_j} d_{s_j,f}^{-\alpha}} h_{s_j,f} x_j + n_{s_j,f}, \qquad (4)$$

respectively; where $h_{s_i,f}$ and $h_{s_j,f}$ are the channel coefficients from the S_i and S_j sensor nodes to the fusion center and are modeled as zero-mean complex Gaussian random variables with variance $1/2$ per dimension. In addition, P_{s_i} and P_{s_j} are the transmission power assigned to sensors S_i and S_j respectively and are selected to satisfy a total power constraint P_T while $n_{s_i,f}$ and $n_{s_i,f}$ are the additive white Gaussian noise at the fusion center. In this protocol, all the N nodes are utilized as sensors for the transmission of observations (i.e. $|\mathcal{S}| = N$) (where $\mathcal{S}$ is the set of all the sensor nodes in the network) and it is assumed that sensor node S_j is closer to the fusion center than sensor node S_i (i.e. $d_{s_j,f} < d_{s_i,f}$).

B. Protocol II

In Protocol II (Fig. 1b), the farther sensor node S_i is paired with a known neighboring cooperative relay node R_k located at $d_{r_k,f}$ from the fusion center. In the AF relaying, signal transmissions are separated into two phases through orthogonal channels by using either TDMA or FDMA [9]. In Phase 1, sensor node S_i transmits its observation x_i with power P_{s_i} where the received signals at the fusion center $y_{s_i,F}$ and at the relay node y_{s_i,r_k} are

$$y_{s_i,f} = \sqrt{P_{s_i} d_{s_i,f}^{-\alpha}} h_{s_i,f} x_i + n_{s_i,f}, \qquad (5)$$

and

$$y_{s_i,r_k} = \sqrt{P_{s_i} d_{s_i,r_k}^{-\alpha}} h_{s_i,r_k} x_i + n_{s_i,r_k}, \qquad (6)$$

respectively; where $h_{s_i,f}$ and h_{s_i,f_k} are the fading channel coefficients from sensor S_i to the fusion center and to the R_k relay node, respectively; whereas, $d_{s_i,f}$ and d_{s_i,r_k} are the distances from S_i to the fusion center and to R_k relay node, respectively and $n_{s_i,f}$ and n_{s_i,r_k} are the white Gaussian noise. In Phase 2, the relay node amplifies the received signal and forwards it to the destination with transmit power P_{r_k}. The received signal at the fusion center can be written as

$$\begin{aligned} y_{r_k,f} &= \sqrt{\tilde{P}_{r_k} d_{r_k,f}^{-\alpha}} h_{r_k,f} y_{s_i,r_k} + n_{r_k,f} \\ &= \sqrt{\tilde{P}_{r_k} P_{s_i} d_{s_i,r_k}^{-\alpha} d_{r_k,f}^{-\alpha}} h_{r_k,f} h_{s_i,r_k} x_i + \tilde{n}_{r_k,f}, \end{aligned} \qquad (7)$$

where $\tilde{n}_{r_k,f} = \sqrt{\tilde{P}_{r_k} d_{r_k,f}^{-\alpha}} h_{r_k,f} n_{s_i,r_k} + n_{r_k,f}$. In addition, $\tilde{P}_{r_k}$ represents the normalized transmit power and is chosen to ensure a transmit power at the relay node of P_{r_k} and thus is specified as $\tilde{P}_{r_k} = (P_{r_k})/(P_{s_i} d_{s_i,r_k}^{-\alpha} |h_{s_i,r_k}|^2 \sigma^2 + N_o)$. Furthermore, $d_{r_k,f}$ is the distance from the relay to the fusion center, $h_{r_k,f}$ is the channel coefficient between the kth relay node to the fusion center and $n_{r_k,f}$ is the additive Gaussian noise. It should be noted that the noise $n_{s_i,f}$, n_{s_i,r_k} and $n_{r_k,f}$ are modeled as independent complex Gaussian random variables with zero means and variance N_o. In this protocol the N nodes in the network are split into two sets $\mathcal{S}$ and $\mathcal{R}$ representing sensor and relay nodes respectively with equal cardinality (i.e. $|\mathcal{S}| = |\mathcal{R}| = N/2$) since each sensor node is paired with a cooperating relay node.

IV. WIRELESS SENSOR NETWORK MODEL

In this section, a wireless sensor network model suitable for correlation analysis amongst sensor nodes is discussed. The sensor/relay nodes are uniformly distributed over a circular area with three concentric circles with the fusion center at the center as illustrated in Fig. 2 for $N = 4$ for both protocols. The two larger circles of radii $d_{so,f}$ and $d_{se,f}$ contain two sets ($\mathcal{S_O}$ and $\mathcal{S_E}$) of uniformly distributed *o*dd-numbered ($S_1, S_3, ..., S_{N-1}$) and *e*ven-numbered ($S_2, S_4, ..., S_N$) sensor nodes respectively. On the other hand, the smallest circle contains the set of relay nodes $\mathcal{R}$ with radius of $d_{r,f}$ and uniformly deployed ($R_2, R_4, ..., R_N$) with even-numbered indices such that ($d_{so,f} > d_{se,f} > d_{r,f}$). That is, the relays are deployed closer to the fusion center than the two sets of sensor nodes. Also, every sensor (in the outer circle - set $\mathcal{S_O}$) is separated from its cooperating nodes by a distance $d_{so,r}$. This circular model as will be seen in later sections simplifies the analysis greatly as all the nodes in $\mathcal{S_O}$ are of equal distance $d_{so,f}$ from the fusion center and so are the nodes in $\mathcal{S_E}$ and $\mathcal{R}$ with distances $d_{se,f}$ and $d_{r,f}$ respectively. The proposed model also introduces symmetry that allows matrix decompositions of covariance matrices and makes the analysis of a network of N nodes mathematically tractable.
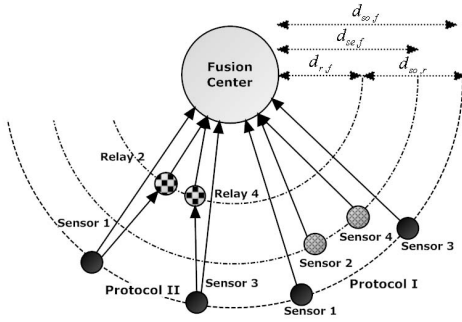
Fig. 2. Wireless Sensor Network Model

In this model, it is assumed that the sensors, relays and the fusion center are stationary and each node (sensor/relay) knows its own and its neighbors distances from the fusion center (on the concentric circles). In addition, the sensor nodes in each of the sets $\mathcal{S}_\mathcal{O}$ and $\mathcal{S}_\mathcal{E}$ are allocated a total power of $P_\mathcal{O}$ and $P_\mathcal{E}$, respectively under Protocol I such that $P_T = P_\mathcal{O} + P_\mathcal{E}$. In Protocol II, the relay nodes in the set $\mathcal{R}$ are allocated a total power of $P_\mathcal{R}$ such that the network power constraint is $P_T = P_\mathcal{O} + P_\mathcal{R}$.

V. PERFORMANCE ANALYSIS

In this section, the probability of detection error P_e expressions of the two network protocols are derived, which are then compared as function of the correlation coefficient between the sensor nodes to determine when Protocol I or Protocol II should be used. To simplify notation and the analysis, it is assumed (without loss of generality) for both protocols, that all the nodes in each set have equal power, i.e. $P_{so} = P_{so_i}$ for $\forall i \in \mathcal{S}_\mathcal{O}$, $P_{se} = P_{se_j}$ for $\forall j \in \mathcal{S}_\mathcal{E}$ and $P_r = P_{r_k}$ for $\forall k \in \mathcal{R}$. Also, let $d_{so,f} = d_{so_i,f}$, $d_{se,f} = d_{se_j,f}$ and $d_{r,f} = d_{r_k,f}$ in each set. Furthermore, it is assumed that all the measured data have the same mean (i.e. under $\mathcal{H}_1$ all sensor nodes measure the same phenomenon with the same *real* mean $m_i = m$ for $i = 1, ..., N$, within an area $\mathcal{A}$).

Given the foregoing system model and protocols, the optimal likelihood decision rule at fusion center of the received observations over Rayleigh fading channels can be derived as follows. Let π_0 and π_1 denote the prior probabilities for $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively where it is assumed for simplicity that $\pi_0 = \pi_1 = 1/2$. The *N-dimensional* received observations vector $\mathbf{Y_N}$ (assuming perfect channel state information (CSI) at the fusion center) under each hypothesis is given by

$$\begin{aligned} \mathcal{H}_0 : \mathbf{Y_N} &\sim \mathcal{CN}(\mathbf{0}, \Sigma_N) \\ \mathcal{H}_1 : \mathbf{Y_N} &\sim \mathcal{CN}(\mu_N, \Sigma_N), \end{aligned} \tag{8}$$

where $\mu_N \triangleq E[\mathbf{Y_N}]$ is the means vector and $\Sigma_N \triangleq E[(\mathbf{Y_N} - \mu_N)(\mathbf{Y_N} - \mu_N)^H]$ is the covariance matrix of the received observations. The optimal test for minimizing the probability of detection error is the likelihood ratio test (LRT) given by [10]

$$\frac{\frac{1}{(2\pi)^{N/2}|\Sigma_N|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{Y_N} - \mu_N)^H \Sigma_N^{-1}(\mathbf{Y_N} - \mu_N)\}}{\frac{1}{(2\pi)^{N/2}|\Sigma_N|^{1/2}} \exp\{-\frac{1}{2}\mathbf{Y_N}^H \Sigma_N^{-1} \mathbf{Y_N}\}} \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} 1. \tag{9}$$

The decision rule can be further simplified and expressed in terms of the sufficient statistic to $T(\mathbf{Y_N}) = \frac{1}{2}\mu_N^H \Sigma_N^{-1} \mathbf{Y_N} + \frac{1}{2}\mathbf{Y_N}^H \Sigma_N^{-1} \mu_N \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \frac{1}{2}\mu_N^H \Sigma_N^{-1} \mu_N$ where it should be noted that $T(\mathbf{Y_N})$ is linear transformation of the Gaussian random vector $\mathbf{Y_N}$ and it characterizes its distributions under $\mathcal{H}_0$ and $\mathcal{H}_1$ by

finding its means and variances under each hypothesis as $P_{\mathcal{H}_0}(T) \sim \mathcal{CN}(0, \mu_N^H \Sigma_N^{-1} \mu_N)$ and $P_{\mathcal{H}_1}(T) \sim \mathcal{CN}(\mu_N^H \Sigma_N^{-1} \mu_N, \mu_N^H \Sigma_N^{-1} \mu_N)$ [11]. By defining probabilities of detection error conditioned on hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ as $P_{\varepsilon|\mathcal{H}_0}$ and $P_{\varepsilon|\mathcal{H}_1}$, respectively; the conditional probability of detection error is expressed as

$$P_e = \pi_0 P_{\varepsilon|\mathcal{H}_0} + \pi_1 P_{\varepsilon|\mathcal{H}_1} = E\left\{Q\left(\frac{1}{2}\sqrt{\mu_N^H \Sigma_N^{-1} \mu_N}\right)\right\}, \tag{10}$$

where $Q(u) \triangleq \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp\left(-\frac{t^2}{2}\right) dt$ is the Gaussian Q-function and the expectation is taken with respect to the channel statistics. It should be noted that finding a closed-form expression for the expectation in (10) is in general very difficult; thus, it will be numerically evaluated in the performance evaluation section.

A. Protocol I - A Two Sensors Network

In this section, the probability of error detection of the $N = 2$ sensors network (shown in Fig. 1a) under Protocol I is derived. The received observation vector is defined as $\mathbf{Y}_{I,2} = [y_{s_1,f} \quad y_{s_2,f}]^T$ (see (3) and (4)) which is employed by the fusion center to decide between the two hypothesis. It can be shown that the data model under hypothesis $\mathcal{H}_0$ is written as $\mathbf{Y}_{I,2} = \mathcal{CN}(0, \Sigma_{I,2})$ while under $\mathcal{H}_1$ is given by $\mathbf{Y}_{I,2} = \mathcal{CN}(\mu_{I,(1,2)}, \Sigma_{I,2})$. The vector of means can be shown to be $\mu_{I,(1,2)} = [\sqrt{P_{so} d_{so,f}^{-\alpha}} h_{s_1,f} m \quad \sqrt{P_{se} d_{se,f}^{-\alpha}} h_{s_2,f} m]^T$, where $h_{s_1,f}$ and $h_{s_2,f}$ are the channel coefficients from sensors S_1 and S_2, respectively. Furthermore, $\Sigma_{I,2}$ is the covariance matrix of the received observations vector $\mathbf{Y}_{I,2}$ defined as shown in (11) (next page), which can be expressed in the matrix form as $\Sigma_{I,2} = (\Gamma_{(1,2)})^H \Omega_2 \Gamma_{(1,2)} + N_o \mathbf{I}$ where $(.)^H$ is the hermitian of the parameter matrix, $\mathbf{I}$ is the identity matrix of appropriate dimensions and matrices $\Gamma_{(1,2)}$ and Ω_2 are defined as

$$\Gamma_{(1,2)} = \begin{pmatrix} \sqrt{P_{so} d_{so,f}^{-\alpha} \sigma^2} h_{s_1,f}^* & 0 \\ 0 & \sqrt{P_{se} d_{se,f}^{-\alpha} \sigma^2} h_{s_2,f}^* \end{pmatrix}, \tag{12}$$

and

$$\Omega_2 = \begin{pmatrix} 1 & \rho_{1,2} \\ \rho_{1,2} & 1 \end{pmatrix}, \tag{13}$$

respectively. Hence, the probability of detection error can be expressed as $P_{e,I} = E\left\{Q\left(\frac{1}{2}\sqrt{(\mu_{I,(1,2)})^H (\Sigma_2)^{-1} \mu_{I,(1,2)}}\right)\right\}$ which after some manipulation can be written as shown in (14) (next page). For the case where sensor nodes S_1 and S_2 have independent observations (i.e. $\rho_{1,2} = 0$), the probability of detection error can be expressed as

$$P_{e,I} = E\left\{Q\left(\frac{1}{2}\sqrt{\frac{P_{so}|h_{s_1,f}|^2 d_{so,f}^{-\alpha} m^2}{P_{so}|h_{s_1,f}|^2 d_{so,f}^{-\alpha} \sigma^2 + N_o} + \frac{P_{se}|h_{s_2,f}|^2 d_{se,f}^{-\alpha} m^2}{P_{se}|h_{s_2,f}|^2 d_{se,f}^{-\alpha} \sigma^2 + N_o}}\right)\right\}. \tag{15}$$

B. Protocol I - N Sensors Network

The analysis is extended to the case of arbitrary large N number of sensors in the network (where N is even). The N nodes are split into two sets $\mathcal{S}_\mathcal{O}$ and $\mathcal{S}_\mathcal{E}$ of equal cardinality. Due to the symmetry of the model, the $N - dimensional$ vector of received observations is $\mathbf{Y}_{I,N} = [y_{s_1,f} \quad y_{s_2,f} \quad \cdots \quad y_{s_{N-1},f} \quad y_{s_N,f}]^T$ for which the covariance matrix $\Sigma_{I,N}$ can be expressed as $\Sigma_{I,N} = (\Gamma_N)^H \Omega_N \Gamma_N + N_o \mathbf{I}$ and matrices Ω_N and Γ_N are given by

$$\Omega_N = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,N} \\ \rho_{1,2} & 1 & \cdots & \rho_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,N} & \rho_{2,N} & \cdots & 1 \end{pmatrix}, \tag{16}$$

$$\Sigma_{I,2} = \begin{pmatrix} P_{so}d_{so,f}^{-\alpha}|h_{s_1,f}|^2\sigma^2 + N_o & \sqrt{P_{so}P_{se}d_{so,f}^{-\alpha}d_{se,f}^{-\alpha}}h_{s_1,f}h_{s_2,f}^*\sigma^2\rho_{1,2} \\ \sqrt{P_{so}P_{se}d_{so,f}^{-\alpha}d_{se,f}^{-\alpha}}h_{s_1,f}^*h_{s_2,f}\sigma^2\rho_{1,2} & P_{se}d_{se,f}^{-\alpha}|h_{s_2,f}|^2\sigma^2 + N_o \end{pmatrix} \tag{11}$$

$$P_{e,I} = E\left\{ Q\left(\frac{1}{2}\sqrt{\frac{2P_{so}P_{se}|h_{s_1,f}|^2|h_{s_2,f}|^2(d_{so,f}^{-\alpha}d_{se,f}^{-\alpha})m^2\sigma^2(1-\rho_{1,2}) + N_om^2(P_{so}|h_{s_1,f}|^2d_{so,f}^{-\alpha} + P_{se}|h_{s_2,f}|^2d_{se,f}^{-\alpha})}{P_{so}P_{se}|h_{s_1,f}|^2|h_{s_2,f}|^2(d_{so,f}^{-\alpha}d_{se,f}^{-\alpha})\sigma^4(1-\rho_{1,2}^2) + N_o\sigma^2(P_{so}|h_{s_1,f}|^2d_{so,f}^{-\alpha} + P_{se}|h_{s_2,f}|^2d_{se,f}^{-\alpha}) + N_o^2}} \right) \right\} \tag{14}$$

$$P_{e,I} = E\left\{ Q\left(\frac{1}{2}\sqrt{\sum_{i\in\mathcal{S}_\mathcal{O}}\frac{P_{so}|h_{s_i,f}|^2d_{so,f}^{-\alpha}m^2}{P_{so}|h_{s_i,f}|^2d_{so,f}^{-\alpha}\sigma^2 + N_o} + \sum_{j\in\mathcal{S}_\mathcal{E}}\frac{P_{se}|h_{s_j,f}|^2d_{se,f}^{-\alpha}m^2}{P_{se}|h_{s_j,f}|^2d_{se,f}^{-\alpha}\sigma^2 + N_o}} \right) \right\} \tag{21}$$

$$\Sigma_{II,2} = \begin{pmatrix} P_{so}d_{so,f}^{-\alpha}|h_{s_1,f}|^2\sigma^2 + N_o & \sqrt{\tilde{P}_{(1,2)}P_{so}P_rd_{so,f}^{-\alpha}d_{r,f}^{-\alpha}}h_{s_1,f}h_{r_2,f}^*h_{s_1,r_2}^*\sigma^2 \\ \sqrt{\tilde{P}_{(1,2)}P_{so}P_rd_{so,f}^{-\alpha}d_{r,f}^{-\alpha}}h_{s_1,f}^*h_{r_2,f}h_{s_1,r_2}\sigma^2 & P_rd_{r,f}^{-\alpha}|h_{r_2,f}|^2 + N_o \end{pmatrix} \tag{22}$$

$$P_{e,II} = E\left\{ Q\left(\frac{1}{2}\sqrt{\frac{m^2N_o(P_{so}d_{so,f}^{-\alpha}|h_{s_1,f}|^2 + \tilde{P}_{(1,2)}P_rd_{r,f}^{-\alpha}|h_{r_2,f}|^2|h_{s_1,r_2}|^2) + P_{so}P_r|h_{s_1,f}|^2|h_{r_2,f}|^2d_{so,f}^{-\alpha}d_{r,f}^{-\alpha}m^2(1-\tilde{P}_{(1,2)}\sigma^2)}{P_{so}P_r|h_{s_1,f}|^2|h_{r_2,f}|^2d_{so,f}^{-\alpha}d_{r,f}^{-\alpha}\sigma^2(1-\tilde{P}_{(1,2)}|h_{s_1,r_2}|^2\sigma^2) + N_o(P_{so}d_{so,f}^{-\alpha}|h_{s_1,f}|^2\sigma^2 + P_rd_{r,f}^{-\alpha}|h_{r_2,f}|^2) + N_o^2}} \right) \right\} \tag{25}$$

and

$$\Gamma_N = \begin{pmatrix} \Gamma_{(1,2)} & 0 & \cdots & 0 \\ 0 & \Gamma_{(3,4)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Gamma_{(N-1,N)} \end{pmatrix}, \tag{17}$$

respectively; where in general

$$\Gamma_{(i,j)} = \begin{pmatrix} \sqrt{P_{so}d_{so,f}^{-\alpha}\sigma^2}h_{s_i,f}^* & 0 \\ 0 & \sqrt{P_{se}d_{se,f}^{-\alpha}\sigma^2}h_{s_j,f}^* \end{pmatrix}. \tag{18}$$

Furthermore, the vector of means $\mu_{I,N}$ can expressed as

$$\mu_{I,N} = [\mu_{I,(1,2)} \quad \mu_{I,(3,4)} \quad \cdots \quad \mu_{I,(N-1,N)}]^T, \tag{19}$$

where; $\mu_{I,(i,j)} = [\sqrt{P_{so}d_{so,f}^{-\alpha}}h_{s_i,f}m \quad \sqrt{P_{se}d_{se,f}^{-\alpha}}h_{s_j,f}m]^T$. Therefore, the probability of detection error is given by

$$P_{e,I} = E\left\{ Q\left(\frac{1}{2}\sqrt{(\mu_{I,N})^H(\Sigma_{I,N})^{-1}\mu_{I,N}} \right) \right\}. \tag{20}$$

Since it is difficult to specify an arbitrary correlation matrix Ω_N that allows fair comparison between the protocols and due to the fact that the aim of this investigation is to characterize the effect "degree" of correlation and the optimum error performance; two *extreme* scenarios are considered, namely the Equally-Correlated Scenario (ECS) and the Mutually-Independent Scenario (MIS) [12].

In the Equally-Correlated Scenario, it is assumed that the N sensors are distributed in the wireless network such that every pair of sensor nodes are *equi-correlated* (i.e. $\rho_{i,j} = \rho$ for $i \neq j$ in Ω_N). In this case, the probability of detection error is evaluated by (20).

In the Mutually-Independent Scenario, the N sensors in the two sets $\mathcal{S}_\mathcal{E}$ and $\mathcal{S}_\mathcal{O}$ are spatially separated far enough from each other such that $\rho_{i,j} \approx 0$ for $i \neq j$. Therefore, due to the symmetry of the model, it can be easily shown that the probability of detection error is expressed as shown in (21). It should also be noted that under MIS, the optimal *lower bound* error performance is achieved.

C. Protocol II - A Sensor with a Relay Network

The probability of detection error of Protocol II (shown in Fig. 1b) is analyzed in this section. It is noteworthy that in this case, there is no correlation between the sensor and the relay nodes since the former node takes measurements of the phenomenon while the latter node only amplifies and forwards the received signals from the source sensor node (see (5-7)). Under protocol II, the received signal vector $\mathbf{Y}_{II,2} = [y_{s_1,f} \quad y_{r_2,f}]^T$ and the data model under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by $\mathbf{Y}_{II,2} = \mathcal{CN}(\mathbf{0}, \Sigma_{II,2})$ and $\mathbf{Y}_{II,2} = \mathcal{CN}(\mu_{II,(1,2)}, \Sigma_{II,2})$, respectively. The covariance matrix $\Sigma_{II,2}$ is shown in (22) which can be expressed in matrix form as $\Sigma_{II,2} = (\Upsilon_{(1,2)})^H\Omega_{(1,2)}\Upsilon_{(1,2)} + N_o\mathbf{I}$ where

$$\Upsilon_{(1,2)} = \begin{pmatrix} \sqrt{P_{so}d_{so,f}^{-\alpha}\sigma^2}h_{s_1,f}^* & 0 \\ 0 & \sqrt{P_rd_{r,f}^{-\alpha}}h_{r_2,f}^* \end{pmatrix}, \tag{23}$$

$$\Omega_{(1,2)} = \begin{pmatrix} 1 & \sqrt{\tilde{P}_{(1,2)}\sigma^2}h_{s_1,r_2}^* \\ \sqrt{\tilde{P}_{(1,2)}\sigma^2}h_{s_1,r_2} & 1 \end{pmatrix}, \tag{24}$$

and $\tilde{P}_{(1,2)} = (P_{so}d_{so,r}^{-\alpha})/(P_{so}d_{so,r}^{-\alpha}|h_{s_1,r_2}|^2\sigma^2 + N_o)$. The vector of means can be expressed as, $\mu_{II,(1,2)} = [\sqrt{P_{so}d_{so,f}^{-\alpha}}h_{s_1,f}m \quad \sqrt{\tilde{P}_{(1,2)}P_rd_{r,f}^{-\alpha}}h_{r_2,f}h_{s_1,r_2}m]^T$. In this case, the probability of detection error can be written as shown in (25).

D. Protocol II - N Nodes Network

In this case, the N nodes in the network are split into two sets; $\mathcal{S}_\mathcal{O}$ and $\mathcal{R}$. However, it should be noted that in this case, the measurements obtained from each of the sensors in the set $\mathcal{S}_\mathcal{O}$ can be correlated. For example, for $(N = |\mathcal{R}| + |\mathcal{S}_\mathcal{O}| = 4)$, the observations from sensors S_1 and S_3 (paired with relays R_2 and R_4, respectively) are correlated by $\rho_{1,3}$ that is a function of the distance $d_{s_{1,3}}$ between them. For a network with arbitrary N nodes, the $N - dimensional$ vector of received observations is $\mathbf{Y}_{II,N} = [y_{s_1,f} \quad y_{r_2,f} \quad \cdots \quad y_{s_{N-1},f} \quad y_{r_N,f}]^T$ where the covariance matrix can be shown to have the form $\Sigma_{II,N} = (\Upsilon_N)^H\Omega_N\Upsilon_N + N_o\mathbf{I}$. In addition, every pair of sensors in $\mathcal{S}_\mathcal{O}$ can be correlated depending on their inter-sensor separation. In general, it can be easily shown that for $(N \geq 4)$, the matrix Υ_N can be expressed as

$$\Upsilon_N = \begin{pmatrix} \Upsilon_{(1,2)} & 0 & \cdots & 0 \\ 0 & \Upsilon_{(3,4)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Upsilon_{(N-1,N)} \end{pmatrix} \tag{26}$$

$$\Omega_N = \begin{pmatrix} \Omega_{(1,2)} & \rho_{1,3}\Theta_{(1,2),(3,4)} & \cdots & \rho_{1,N/2-1}\Theta_{(1,2),(N-1,N)} \\ \rho_{1,3}\Theta_{(3,4),(1,2)} & \Omega_{(3,4)} & \cdots & \rho_{3,N/2-1}\Theta_{(3,4),(N-1,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,N/2-1}\Theta_{(N-1,N),(1,2)} & \rho_{3,N/2-1}\Theta_{(N-1,N),(3,4)} & \cdots & \Omega_{(N-1,N)} \end{pmatrix} \quad (28)$$

$$P_{e,II} = E\left\{ Q\left(\frac{1}{2}\sqrt{\sum_{(i,k)\in(\mathcal{S}_O,\mathcal{R})} \frac{m^2 N_o(P_{so}d_{so,f}^{-\alpha}|h_{s_i,f}|^2 + \tilde{P}_{(i,k)}P_r d_{r,f}^{-\alpha}|h_{r_k,f}|^2|h_{s_i,r_k}|^2) + P_{so}P_r|h_{s_i,f}|^2|h_{r_k,f}|^2 d_{so,f}^{-\alpha}d_{r,f}^{-\alpha}m^2(1-\tilde{P}_{(i,k)}\sigma^2)}{P_{so}P_r|h_{s_i,f}|^2|h_{r_k,f}|^2 d_{so,f}^{-\alpha}d_{r,f}^{-\alpha}\sigma^2(1-\tilde{P}_{(i,k)}|h_{s_i,r_k}|^2\sigma^2) + N_o(P_{so}d_{so,f}^{-\alpha}|h_{s_i,f}|^2\sigma^2 + P_r d_{r,f}^{-\alpha}|h_{r_k,f}|^2) + N_o^2}} \right) \right\} \quad (32)$$

where in general $\Upsilon_{(i,k)}$ is defined as

$$\Upsilon_{(i,k)} = \begin{pmatrix} \sqrt{P_{so}d_{so,f}^{-\alpha}\sigma^2}h_{s_i,f}^* & 0 \\ 0 & \sqrt{P_r d_{r,f}^{-\alpha}}h_{r_k,f}^* \end{pmatrix}. \quad (27)$$

The matrix Ω_N has the structure shown in (28) where

$$\Omega_{(i,k)} = \begin{pmatrix} 1 & \sqrt{\tilde{P}_{(i,k)}\sigma^2}h_{s_i,r_k}^* \\ \sqrt{\tilde{P}_{(i,k)}\sigma^2}h_{s_i,r_k} & 1 \end{pmatrix}, \quad (29)$$

for $\tilde{P}_{(i,k)} = (P_{so}d_{so,r}^{-\alpha})/(P_{so}d_{so,r}^{-\alpha}|h_{s_i,r_k}|^2\sigma^2 + N_o)$ and $\Theta_{(i,j),(k,l)}$ can be expressed as

$$\Theta_{(i,j),(k,l)} = \begin{pmatrix} 1 & \sqrt{\tilde{P}_{(k,l)}\sigma^2}h_{s_k,r_l}^* \\ \sqrt{\tilde{P}_{(i,j)}\sigma^2}h_{s_i,r_j} & \sqrt{\tilde{P}_{(i,j)}\tilde{P}_{(k,l)}\sigma^2}h_{s_i,r_j}h_{s_k,r_l}^* \end{pmatrix}. \quad (30)$$

In addition, the vector of means is given by

$$\mu_{II,N} = \begin{bmatrix} \mu_{II,(1,2)} & \mu_{II,(3,4)} & \cdots & \mu_{II,(N-1,N)} \end{bmatrix}^T, \quad (31)$$

where; $\mu_{II,(i,k)} = [\sqrt{P_{so}d_{so,f}^{-\alpha}}h_{s_i,f}m \quad \sqrt{\tilde{P}_{(i,k)}P_r d_{r,f}^{-\alpha}}h_{r_k,f}h_{s_i,r_k}m]^T$. Under the Equally-Correlated Scenario, the probability of detection error can be determined using, $P_{e,II} = E\left\{ Q\left(\frac{1}{2}\sqrt{(\mu_{II,N})^H(\Sigma_{II,N})^{-1}\mu_{II,N}} \right) \right\}$ while for the mutually independent scenario, the probability of detection error is given by (32).

VI. PERFORMANCE EVALUATION OF NETWORK PROTOCOLS

In this section, performance evaluation of both protocols under consideration in Rayleigh fading channels are compared in terms of the derived probability of detection error P_e. To allow a fair comparison between the performance of the network protocols, it is assumed that within a particular area $\mathcal{A}$, the total network power constraint P_T of N nodes is equally distributed amongst all nodes. That is, under protocol I, $P_{so} = P_{se} = P_T/N$ for all the sensor nodes while under protocol II, $P_{so} = P_r = P_T/N$ for all sensor and relay nodes. This in turn implies that the power allocation per each node decreases linearly with the increase in the total number of nodes in the network area $\mathcal{A}$. The simulation parameters of the network protocols are as follows. The measurement noise variance is $\sigma^2 = 0.01$ while the mean of measurements is $m = 1$. In addition, the path-loss exponent is $\alpha = 3$ and the correlation decay factor is $\lambda = 0.05$. The distance from the sensors in $\mathcal{S}_O$ and $\mathcal{S}_\mathcal{E}$ and the relays in $\mathcal{R}$ are $d_{so,f} = 4.5$, $d_{se,f} = 3.5$ and $d_{r,f} = 2.5$, respectively.

Fig. 3 illustrates the P_e for Protocol I with $N = 2$. As explained in the proposed model, different inter-sensor separations result in different correlation coefficients since $\rho_{1,2} = \exp(-\lambda d_{s_{1,2}}^2)$. Therefore, it is clear that as the inter-sensor separation increases, the correlation coefficient decreases and the error performance improves. This is to be expected since the lower the correlation, the more statistically independent the observations become; hence more information is fed
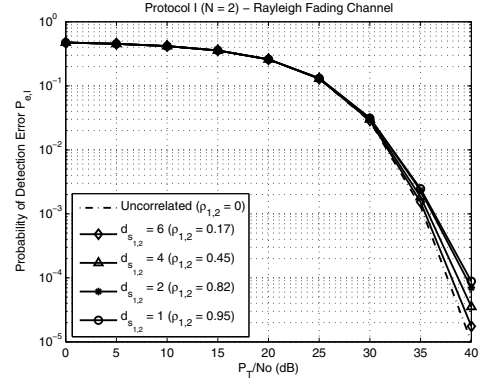

Fig. 3. Protocol I - Probability of Detection Error for $N = 2$ in Rayleigh Fading Channel for different ρ values (as a function of $d_{s_{1,2}}$)
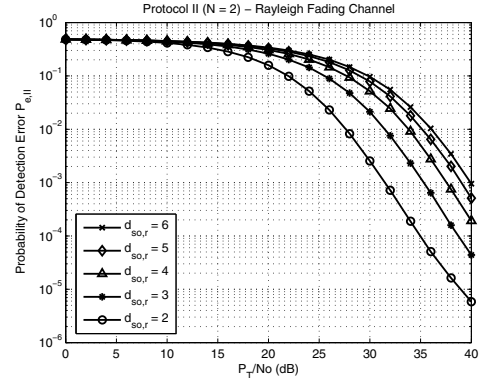

Fig. 4. Protocol II - Probability of Detection Error for $N = 2$ in Rayleigh Fading Channel for different values of inter-sensor/relay separation $d_{so,r}$

to the fusion center which in turn results in better error detection performance. Also shown, the case where both sensors S_1 and S_2 are far away from each other since that their correlation coefficient $\rho_{1,2} = 0$ (i.e. uncorrelated) and this achieves the optimal *(lower bound)* error performance of Protocol I as a function of $\rho_{1,2}$.

Under Protocol II for $N = 2$ (Fig. 4), it can also be seen that as the inter-sensor/relay distance $d_{so,r}$ increases, the error performance degrades. This is due to the fact that farther apart the sensor and the relay become, the weaker the received signal and hence the worse the performance. Hence, the closer the sensor to the cooperating relay, the better the performance, and in this case, the closest the sensor to the cooperating relay node is when $d_{so,r} = 2$ (since $d_{so,f} = 4.5$ and $d_{r,f} = 2.5$ and hence the relay node is on a straight line and in the middle between the sensor and the fusion center).

In order to determine when it is better to use Protocol I (as a function of $d_{s_{1,2}}$) or Protocol II (as a function of $d_{so,r}$) for the simple $N = 2$ network, the error performance at $P_T/N_o = 40dB$ is plotted in Fig. 5. It can be seen that for inter-sensor/relay distances
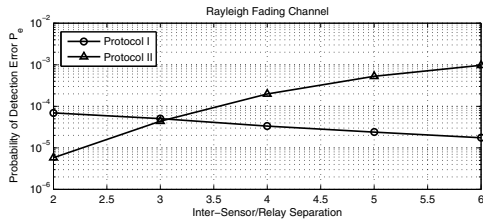
Fig. 5. Protocol I vs. Protocol II as a Function of Inter-Sensor/Relay Separation for $N = 2$ in Rayleigh Fading Channels for $P_T/No = 40dB$
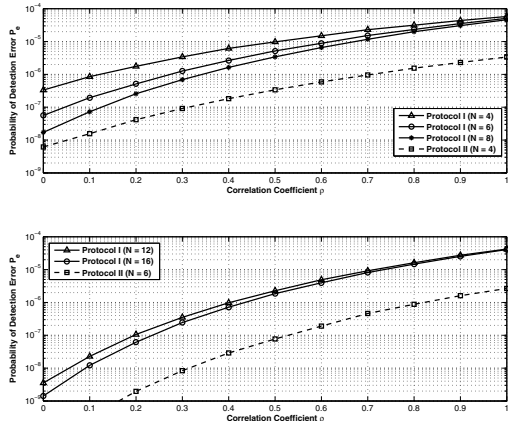

Fig. 6. Protocol I vs. Protocol II - Equally-Correlated Scenario in the Rayleigh Fading Channel for $P_T/N_o = 40dB$

less than 3, Protocol II outperforms Protocol I. In other words, if two sensors S_1 and S_2 have inter-sensor separation $d_{s_{1,2}} < 3$, then it would be better to not use S_2 for transmission but instead pair sensor node S_1 with the cooperating relay R_2 (closer to the fusion center) to amplify-and-forward the transmitted observations of S_1 with inter-sensor/relay separation $d_{so,r} < 3$. It is also evident from Figs. (4 and 5) that the sensor/relay separation when $d_{so,r} = 2$ gives the best performance for Protocol II. Thus, it will be assumed in the subsequent simulation results.

In Fig. 6, the performance of Protocol I and Protocol II for $N \geq 4$ (under the ECS) as a function of ρ is evaluated. As noticed earlier, when the correlation amongst the observations increases, the performance of the network degrades. It is also clear that the
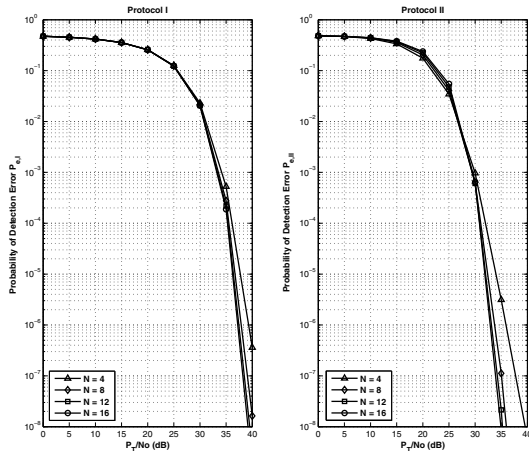

Fig. 7. Protocol I vs. Protocol II - Mutually Independent Scenario in the Rayleigh Fading Channel

performance of Protocol II with $N = 4$ outperforms that of Protocol I with $N = 4$, $N = 6$ and $N = 8$. Also, the performance of Protocol II with $N = 6$ is significantly better than Protocol I with $N = 12$ and $N = 16$. This implies that with relays deployment, the network detection error performance is improved significantly due to the spatial diversity achieved by cooperation. Therefore, sending many observations from sensors nodes does not necessary improve the detection performance significantly. However, by cooperatively relaying a fewer number of observations (i.e. with a fewer number of sensor/relay nodes), greater detection reliability is achieved at the fusion center.

In Fig. 7, Protocols I and II are evaluated under the MIS. It can be seen that Protocol II outperforms Protocol I for different numbers of nodes. It is also clear that increasing the number of sensor/relay nodes beyond $N = 12$ does not improve the error performance significantly given the power constraint P_T on the network. In addition, the performance obtained from Protocol II with $N = 8$ sensor/relay nodes is better than the performance obtained from Protocol I with $(N \geq 12)$ sensor nodes as was shown earlier.

VII. Conclusions

In this paper, a wireless sensor network model suitable for the analysis of direct transmission of observations from sensor nodes (Protocol I) as well as cooperative (AF) relay nodes (Protocol II) is proposed. It was demonstrated that the proposed model facilitates the analysis of the theoretical performance for possibly correlated observations from sensors with relays deployment. It has also been shown that for a simple network with only two nodes, Protocol II outperforms Protocol I when the observations are highly correlated (sensor nodes are in close proximity of each other) and thus it is better to employ a single sensor with a cooperating relay node. On a larger scale (for $N \geq 4$), it was shown that employing many sensor nodes does not necessarily achieve the best network detection error performance. However, by employing relays, significant detection error performance gains are achieved from diversity and also a reduction in the total number of nodes required for optimal detection.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Comunications Magazine*, pp. 102–114, 2002.
[2] B. Wang, J. Zhang, and L. Zheng, "Achievable rates and scaling laws of power-constrained wireless sensory relay networks," *IEEE Transactions on Information Theory*, pp. 4084–4104, 2006.
[3] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of mimo and cooperative mimo techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1089–1098, 2004.
[4] Y. Sung, S. Misra, L. Tong, and A. Ephremides, "Cooperative routing for distributed detection in large sensor networks," *IEEE Journal on Selected Areas in Communications*, pp. 471–483, 2007.
[5] L. S. Pilluta and V. Krishnamurthy, "Minimum energy data gathering in correlated sensor networks with cooperative transmission," *IEEE International Conference on Communcations*, pp. 3623–3628, 2007.
[6] J.-F. Chamberland and V. V. Veeravalli, "How dense should a sensor nework be for detection with correlated observations," *IEEE Transactions on Information Theory*, vol. 52, pp. 5099–5106, 2006.
[7] ——, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, pp. 407–416, 2003.
[8] R. Critescu, B. Beferull-Lozano, and M. Vetterli, "On network correlated data gathering," *IEEE INFOCOM 2004*, vol. 4, pp. 2571–2582, 2004.
[9] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, pp. 3062–3080, 2004.
[10] H. L. V. Trees, *Detection, Estimation and Modulation Theory*. Wiley, 1968.
[11] H. V. Poor, *An Introduction to Signal Detection and Estimation - Second Edition*. Springer-Verlag, 1994.
[12] E. Drakopoulos and C. C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. on Aero. and Elec. Sys.*, pp. 593–606, 1991.