Parameter Estimation Over Noisy Communication Channels in Distributed Sensor Networks

Thakshila Wimalajeewa[1], Sudharman K. Jayaweera[1] and Carlos Mosquera[2]
[1]*Dept. of Electrical and Computer Engineering, University of New Mexico,*
[2]*Dept. de Teoría de la Señal y Comunicaciones, Universidad de Vigo,*
[1]*USA*
[2]*Spain*

1. Introduction

With the recent development of low-cost, low-power, multi-functional sensor nodes, sensor networks have become an attractive emerging technology in a wide variety of applications including, but not limited to, military surveillance, civilian, industrial and environmental monitoring [1]–[5]. In most of these new applications sensor nodes are capable not only of sensing but also of data processing, wireless communications and networking. It can be argued that it is their ability of ad hoc wireless networking that has attracted much interest to wireless sensor networks in recent years.

A typical sensor network may consist of a large number of spatially distributed nodes to make a decision on a Parameter of Interest (PoI). This can be detection, estimation or tracking of a target or multiple targets. Once the network is deployed, the network resources, such as node power and communication bandwidth, are limited in many situations. This is due to the fact that reinstalling and recharging the batteries might be difficult, or even impossible, once the network is deployed. A common question arising in such networks is how to effectively combine the information from all the nodes in the network to arrive at a final decision while consuming the resources in an optimum way. In a distributed sensor network, the distributed nodes make observations of PoI and process them locally to make a summary of their observations. The final decision is usually made by combining these locally processed data.

Once local decisions are made at each individual sensor node, the natural questions are how to combine the local decisions and where the final decision is made. When there is a possibility that the sensor network can have a central node (generally called as the fusion center) with relatively high processing power compared to distributed nodes, the summary of the local observations can be sent to the fusion center. The fusion center combines all local decisions in an optimum way to arrive at a final decision in what is known as the centralized architecture. The disadvantage of such a system is that if there is a failure in the fusion center, the whole network fails. On the other hand, in some applications, it might be of interest that nodes communicate with each other to reach at a final decision without

Source: Sensor and Data Fusion, Book edited by: Dr. ir. Nada Milisavljević,
ISBN 978-3-902613-52-3, pp. 490, February 2009, I-Tech, Vienna, Austria

depending on a central node. In this set-up the node that makes the final decision may change over time due to the dynamic nature of the sensor network and the PoI. This may lead to a more robust architecture compared to that with centralized architecture.

Irrespective of the data fusion architecture, the local information from sensor nodes needs to be shared over a communication channel that, in general, can undergo both path loss attenuation and multipath fading. As a result, the received signal at a destination node, be it another distributed node or a central fusion center, is corrupted by both multiplicative and additive noise. The performance of the final decision will depend not only on measurement noise at the distributed nodes but also on channel quality of communication links.

The performance of resource-constrained wireless sensor networks with communication and measurement noise has been addressed by many authors in different contexts. For example, performance of the sensor networks under power and bandwidth constraints are analyzed in [6]–[22] and [10], [11], [23]–[36], respectively. Collaborative signal processing, including sequential communication, is addressed in [15], [37]–[42].

In this chapter we address the problem of muti-sensor data fusion over noisy communication channels. The objective of the sensor network is to estimate a deterministic parameter. Distributed nodes make noisy observations of the PoI. Each node generates either an amplified version of its own observation or a quantized message based on its own observation, and shares it with other nodes over a wireless channel. The final decision can be made either at a central node (fusion center) or fully distributively. In the case of centralized architecture, the locally processed messages can be sent to the fusion center over a set of orthogonal channels or a multiple-access channel in which nodes share a common communication channel. In the fully distributive architecture, there is no explicit central fusion center and nodes communicate with each other to arrive at a final decision. There are several variations of this architecture: in one setting, nodes may communicate sequentially with neighbors to sequentially update an estimator (or a sufficient statistic for the parameter). The final decision can be declared by any node during this sequential updating process. On the other hand, it might sometimes be of interest for all nodes in the network to arrive at a common final decision. This leads to a distributed consensus estimation problem. Note that, here all nodes communicate with each other in contrast to the sequential communication architecture above until they reach an agreement.

The rest of this chapter is organized as follows: Section 2 formulates the problem of parameter estimation over noisy communication channels in a distributed sensor network. Section 2-A presents the assumed observation model. The ideal centralized estimation is reviewed in Section 2-B. Ideal estimation is, of course, not possible in a wireless sensor network since communication is over a noisy channel and the network is constrained by available communication resources. By sharing only a summary of the observations with each other, the scarce communication resources can be efficiently utilized. Local processing schemes to achieve this goal are discussed in Section 2-C.

Section 3 focuses on centralized estimation architecture with noisy communication channels between distributed nodes and the fusion center. Estimation performance with orthogonal channels is discussed in Section 3-A and that with non-orthogonal communication channels is discussed in Section 3-B.

Sections 4 and 5 discuss the distributed estimation performance in a sensor network with collaborative information processing. Section 4 considers the distributed sensor network architecture with sequential communication where inter-sensor communication links are assumed to be noisy. In Section 5 collaborative estimation with distributed consensus is addressed. Here, the nodes are allowed to communicate with a set of other nodes that are considered as their neighbors. Sections 5-A and 5-B address static parameters whereas section 5-C considers, time-varying parameters. Finally chapter summary is given in section 6.

2. Data fusion problem in a wireless sensor network

Throughout, we consider a spatially distributed sensor network that is deployed to estimate a PoI. It is natural to expect that a final decision be obtained by combining the information from different nodes. In a distributed sensor network, nodes share summaries of their observations over noisy communication channels. Since network resources, in particular the node power and the communication bandwidth, are scarce it is important that the observations at each node are locally processed to reduce the observation to a concise summary. The final decision can then be made based on these local outputs that nodes share with each others and/or with a fusion center.

A. Multi-sensor observation model

We consider a situation in which multiple sensors observe a PoI. When these nodes form a sensor network, the final decision can be made in either a centralized or distributed way. In the centralized architecture, each node sends a summary of its observations to a central node called a fusion center. There is no inter-node communication. The fusion center combines all received information in an effective way to arrive at a final decision. In the distributed decision-making architecture, on the other hand, the nodes collaborate with each other to arrive at a final decision distributively, without the aid of a central fusion node.

Irrespective of the architecture, communication between sensors and the fusion center, or among sensors, is over a noisy channel. Thus, the information sent sees distortion due to both additive as well as multiplicative noise. The multiplicative noise is due to path loss attenuation and multipath fading encountered, for example, in a wireless channel. In this section, we first consider the centralized architecture as shown in Fig. 1. The distributed architecture is covered in Sections 4 and 5.

Consider a spatially distributed network of n sensors. Let us assume that the network is to estimate, in general, a vector parameter $\mathbf{\Theta}$ where $\mathbf{\Theta}$ is a p-vector. The observation at each node is related to the parameter $\mathbf{\Theta}$ that we wish to estimate via the following observation model;

$$z_k(t) = f_k(\mathbf{\Theta}) + v_k(t),$$

where $z_k(t)$ is the observation at the k-th node at time t, $f_k : \mathbb{R}^p \to \mathbb{R}$ is a function of the parameter vector $\mathbf{\Theta}$ (in general, non-linear) and $v_k(t)$ is the additive observation noise at node k. In the special case of linear observation model, the joint observation vector at n nodes at time t can be written as

$$\mathbf{z}(t) = \mathbf{B}\mathbf{\Theta} + \mathbf{v}(t), \tag{1}$$

where $\mathbf{B}$ is an $n \times p$ (known) matrix and $\mathbf{v}$ is the observation noise vector having a zero mean and a covariance matrix of $\mathbf{\Sigma_v}$. In this chapter we focus mainly on scalar parameter estimation (where we assume p = 1) although the techniques developed and the results can easily be extended to vector parameter estimation. For a scalar parameter, the observation vector (1) formed by observations at all n nodes reduces to,

$$\mathbf{z} = \theta \mathbf{e} + \mathbf{v}, \tag{2}$$

where we have suppressed the timing index t and **e** is the n-vector of all ones.

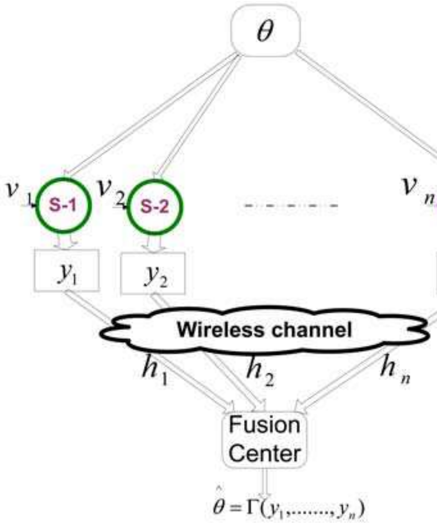

Fig. 1. Distributed estimation with a central fusion center

B. Ideal centralized data fusion

When local observation vector **z** is directly available at the fusion center, the problem is termed the ideal centralized data fusion. The optimal final estimator and its mean-squared error performance are summarized in the following lemma:

Lemma 1: [43] *When the observation vector (2) is available at the fusion center, the best linear unbiased estimator (BLUE) for the scalar parameter θ is given by*

$$\hat{\theta}(\mathbf{z}) = \frac{\mathbf{e}^T \mathbf{\Sigma_v}^{-1} \mathbf{z}}{\mathbf{e}^T \mathbf{\Sigma_v}^{-1} \mathbf{e}}, \tag{3}$$

where x^T denotes the transpose of x. The corresponding mean squared error (MSE) achieved by (3) is

$$MSE = \mathbb{E}\{|\hat{\theta} - \theta|^2\} = \left(\mathbf{e}^\mathbf{T} \mathbf{\Sigma}_\mathbf{v}^{-1} \mathbf{e}\right)^{-1}, \tag{4}$$

where $\mathbb{E}\{.\}$ denotes the mathematical expectation. Further, if the local observations are i.i.d., so that $\mathbf{\Sigma_v} = \sigma_v^2 \mathbf{I}$, where $\mathbf{I}$ is the $n \times n$ identity matrix, the estimator in (3) simplifies to the sample mean of the observations,

$$\hat{\theta}(\mathbf{z}) = \frac{1}{n} \sum_{k=1}^{n} z_k,$$

with the corresponding MSE in (4) simplified to

$$MSE = \frac{\sigma_v^2}{n}.$$

Since communication from distributed nodes to the fusion center is over noisy channels, in practice signals transmitted by the distributed nodes undergo distortion. Hence, direct access from a distant fusion center to the exact observations at distributed nodes may not be possible. However, the lemma 1 will serve as a benchmark for other schemes that we will discuss in this chapter.

C. Local processing at sensor nodes

To facilitate efficient utilization of node and network resources, each node in a sensor network locally processes its observation to generate a useful summary. The transmitted signal at the k-th node is then given by $y_k = g_k(z_k)$. In the following we consider two local processing schemes:

1) Amplify-and-Forward (AF) local processing: In many practical situations where sensor observations are corrupted by additive noise, the amplify and forward strategy has been shown to perform well. In this method, each node directly amplifies its observation and sends it to the fusion center. The transmitted signal from node k is

$$y_k = g_k z_k,$$

where g_k is the amplifying gain at the k-th node. In order to save the node power, it is important to select the amplification gain g_k for $k = 1,..., n$ appropriately depending on the other network parameters such as channel quality and observation quality, etc.. If nodes are operated at the same power level, sometimes it may lead to an unnecessary usage of the network power especially when observation qualities of nodes and channel qualities are not the same for all nodes. Therefore, choosing g_k's in a meaningful way is an important issue to be addressed in designing resource-constrained sensor networks. This problem is discussed in section 3.

With AF local processing, the received signal vector at the fusion center with noiseless communication is given by

$$\mathbf{r} = \mathbf{G}\mathbf{z}, \tag{5}$$

where $\mathbf{G}$ = diag(g_1,..., g_n) is the channel gain matrix. Then the Best Linear Unbiased Estimator and its corresponding mean squared error is given by the following lemma:

Lemma 2: [34], [43] *If the received signal at the fusion center is as given in (5), then the BLUE estimator based on the received signal vector is given by*

$$\hat{\theta}(\mathbf{r}) = \frac{\mathbf{e^T}\boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{G}^{-1}\mathbf{r}}{\mathbf{e^T}\boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{e}},$$

and the corresponding MSE is

$$MSE = \mathbb{E}\{|\hat{\theta} - \theta|^2\} = \left(\mathbf{e^T}\boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{e}\right)^{-1}.$$

Further, if the local observations are i.i.d., so that $\boldsymbol{\Sigma}_\mathbf{v} = \sigma_v^2\mathbf{I}$, the MSE simplifies to

$$MSE = \frac{\sigma_v^2}{n}.$$

Fig. 2. Probabilistic quantization

2) Quantized local processing: To save node energy and communication bandwidth, sensors can compress their observations before transmitting to the fusion center. In this set up, local nodes quantize their observations to generate finite-range messages $m_k(z_k)$ where each m_k is represented by L_k number of bits [9]. Based on the quantized messages received from nodes, the fusion center computes the final estimator

$$\hat{\theta} = \Gamma(\hat{m}_1, \hat{m}_2, \cdots, \hat{m}_n)$$

where $\hat{m}_k$'s are the corrupted versions of quantized messages m_k's received at the fusion center and Γ(.) is the final estimator mapping.

There are several quantization schemes proposed in the literature each having its own advantages and disadvantages [9], [22], [21]. For simplicity, throughout this chapter we concentrate on the universal decentralized quantization scheme given in [9]. According to this scheme, each node locally quantizes its own observation z_k into a discrete message $m_k(z_k,L_k)$ of L_k bits. Due to the lack of knowledge of probability density function (pdf) of noise, the quantizer $Q_k : z_k \rightarrow m_k(z_k,L_k)$ at local nodes is designed to be a uniform randomized quantizer [9]. To that end, suppose the observation range of each sensor is $[-W,W]$ where W is a known parameter determined by the physical properties of the sensor nodes. At each node the range $[-W, W]$ is divided into 2^{L_k}-1 intervals of length $\Delta_k = 2W/(2^{L_k} - 1)$ each as shown in Fig. 2. The quantizer Q_k rounds-off z_k to the nearest endpoint of one of these intervals in a probabilistic manner. For example, suppose, $i\Delta_k \leq z_k < (i + 1)\Delta_k$ where $-2^{L_k-1} \leq i \leq 2^{L_k-1}$. Then Q_k will quantize z_k into $m_k(z_k,L_k)$ so that

$$P\{m_k(z_k, L_k) = i\Delta_k\} = 1 - r,$$

and

$$P\{m_k(z_k, L_k) = (i+1)\Delta_k\} = r,$$

where $r \equiv (z_k - i\Delta_k)/\Delta_k \in [0, 1]$. Note that the quantizer noise $q_k(z_k,L_k) = m_k(z_k,L_k)-z_k$ is then a Bernoulli random variable taking values of $q_k(z_k,L_k) = -r\Delta_k$ and $q_k(z_k,L_k) = (1 - r)\Delta_k$ with probabilities

$$P\{q_k(z_k, L_k) = -r\Delta_k\} = 1 - r,$$

and

$$P\{q_k(z_k, L_k) = (1-r)\Delta_k\} = r.$$

With this local processing scheme the quantized message at node k can be expressed as

$$m_k(z_k, L_k) = z_k + q_k = \theta + v_k + q_k; \text{ for } k = 1, \cdots, n, \tag{6}$$

where we have made use of (2). Note that the quantization noise q_k and the observation noise v_k in (6) will be assumed to be independent. Moreover, q_k is independent across sensors since quantization is performed locally at each sensor.

It can be easily shown that $m_k(z_k, L_k)$ is an unbiased estimator of θ so that $\mathbb{E}\{m_k\} = \theta$ with the MSE (which is, in fact, the variance of the estimator) upper bounded as

$$V_k(m_k) \leq \frac{W^2}{(2^{L_k}-1)^2} + \sigma_v^2 = \delta_k^2 + \sigma_v^2 \quad \text{for} \quad k = 1, \cdots, n, \tag{7}$$

where $\delta_k^2 = \frac{W^2}{(2^{L_k}-1)^2}$. Hereafter we use the short-hand notation m_k to denote $m_k(z_k, L_k)$, so that the transmitted signal y_k at node k is $y_k = m_k$ for $k = 1,..., n$.

With quantized processing, the received signal vector at the fusion center, with noiseless communication is

$$\mathbf{r} = \mathbf{m} = \theta\mathbf{e} + \mathbf{v} + \mathbf{q} \tag{8}$$

where $\mathbf{q} = [q_1,..., q_n]^T$ is the quantization noise vector and $\mathbf{m} = [m_1,..., m_n]^T$. The BLUE estimator at the fusion center and its performance are characterized in lemma 3 below.

Lemma 3: [9] *The BLUE estimator based on the received signal in (8) is given by*

$$\hat{\theta}(\mathbf{m}) = \frac{\mathbf{e^T}(\mathbf{\Sigma_v} + \mathbf{\Sigma_q})^{-1}\mathbf{m}}{\mathbf{e^T}(\mathbf{\Sigma_v} + \mathbf{\Sigma_q})^{-1}\mathbf{e}},$$

where $\Sigma_q = diag(\delta_1^2, ..., \delta_n^2)$. *An upper bound for the MSE of above estimator can be found to be (using(7))*

$$MSE = \mathbb{E}\{|\hat{\theta} - \theta|^2\} \leq \left(\mathbf{e^T}(\mathbf{\Sigma_v} + \mathbf{\Sigma_q})^{-1}\mathbf{e}\right)^{-1}. \tag{9}$$

When local observations are i.i.d. the MSE upper bound (9) can be further simplified as

$$MSE \leq \left(\sum_{k=1}^{n} \frac{1}{\sigma_v^2 + \delta_k^2}\right)^{-1}. \tag{10}$$

Of course, in practice the above ideal estimators cannot be realized due to imperfect communications between distributed nodes and the fusion center. These imperfections can be due to multiplicative noise (caused by channel fading and path loss attenuation) and additive noise at the receiver. When the sensor system has to conform with resource constraints on node power and communication bandwidth, it is important to consider the minimum achievable error performance taking into account these channel imperfections. Parameter estimation under imperfect communications in a distributed sensor network is discussed in the next section.

3. Optimal decision fusion over noisy communication channels

The performance of a final estimator when locally processed data are transmitted to the destination over a noiseless channel was discussed in the latter part of Section 2. In this section we discuss the final estimator performance at a fusion center in the presence of noisy communication channels from distributed nodes to the fusion center. In the following we consider two communication schemes where sensors transmit data over orthogonal or non-orthogonal channels.

A. Communication over orthogonal channels

When locally processed sensor data are transmitted through orthogonal channels (either TDMA, FDMA or CDMA), the received signal vector at the fusion center can be written as

$$\mathbf{r} = \mathbf{H_c u} + \mathbf{w}, \tag{11}$$

where $\mathbf{H_c}$ = diag(h_1,..., h_n) are the fading coefficients of each channel and **w** is the receiver noise vector with mean zero and the covariance matrix $\mathbf{\Sigma_w}$. Note that in (11) we have assumed flat fading channels between sensors and the fusion center which can be a reasonable assumption in certain WSN's but not all. When the channels are selective one can modify (11) by using a tapped-delay line model. The statistics of h_k is determined by the type of fading distributions. Throughout this chapter we will assume that h_k's are Rayleigh distributed.

1) AF local processing: With AF local processing and orthogonal communication channels, the received signal vector at the fusion center is given by

$$\begin{aligned} \mathbf{r} &= \mathbf{H_c G z} + \mathbf{w} \\ &= \mathbf{H_c G e}\theta + \mathbf{n}, \end{aligned} \tag{12}$$

where $\mathbf{n} = \mathbf{H_c G v} + \mathbf{w}$ is the effective noise vector at the fusion center with mean zero and covariance matrix $\Sigma_n = \mathbf{H_c G \Sigma_v G H_c} + \mathbf{\Sigma_w}$, assuming that the receiver noise and the node observation noise are independent. In the following lemma we summarize the optimal estimator at the fusion center based on the received signal (12) and its performance:

Lemma 4: [34] *If the fusion center has the knowledge of channel fading coefficients, the BLUE estimator and the MSE based on the received signal (12) is given by*

$$\hat{\theta}(\mathbf{r}) = \frac{\mathbf{e^T G H_c \Sigma_n^{-1} r}}{\mathbf{e^T G H_c \Sigma_n^{-1} H_c G e}}, \tag{13}$$

and

$$MSE = \left(\mathbf{e^T G H_c \Sigma_n^{-1} H_c G e}\right)^{-1}. \tag{14}$$

In the special case when local observations and the receiver noise are both i.i.d. such that $\mathbf{\Sigma_v} = \sigma_v^2 \mathbf{I}$ *and* $\mathbf{\Sigma_w} = \sigma_w^2 \mathbf{I}$, *the BLUE estimator (13) and the MSE (14) further simplify to*

$$\hat{\theta}(\mathbf{r}) = \frac{\sum_{k=1}^{n} \frac{g_k h_k r_k}{\sigma_v^2 h_k^2 g_k^2 + \sigma_w^2}}{\sum_{k=1}^{n} \frac{h_k^2 g_k^2}{\sigma_v^2 h_k^2 g_k^2 + \sigma_w^2}}, \tag{15}$$

and

$$MSE = \left(\sum_{k=1}^{n} \frac{h_k^2 g_k^2}{\sigma_v^2 h_k^2 g_k^2 + \sigma_w^2}\right)^{-1}, \tag{16}$$

where σ_w^2 *is the receiver noise power.*

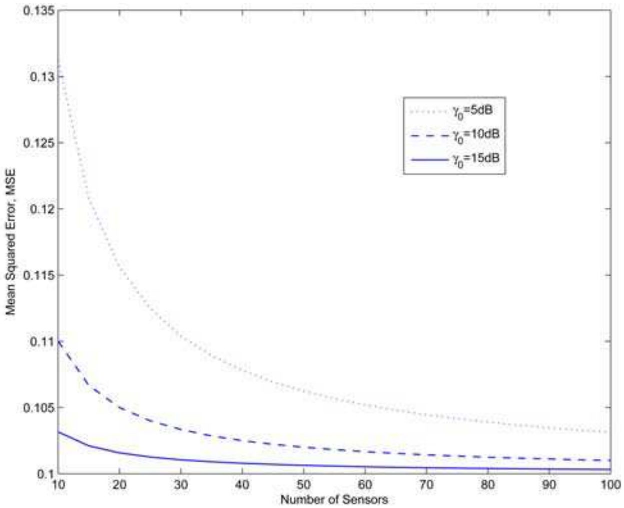

Fig. 3. Mean squared error performance vs. number of nodes: The total network power is constant.

The performance of the BLUE estimator (15) is shown in Figs. 3 and 4 given that the total power in the network is constant. Note that in the both Figs. 3 and 4 the node power is the same at each sensor, so that $g_k = g$ for $k = 1,..., n$ and each channel gain is unity (i.e. h_k=1 for all k). Hence, if total network power is P_T then the individual node power is given by $g^2 = P_T/n$. In this case, the MSE in (16) is further simplified to $MSE = \frac{\sigma_v^2}{n} + \frac{\sigma_w^2}{P_T}$. The local SNR, γ_0 is defined as P_s/σ_v^2 where P_s is the average power at local nodes. In the simulations we have let $P_s = 1$. It can be seen that when either the number of sensors or the total network power is increased, the performance of the BLUE estimator is floored: i.e. $\lim_{n\to\infty} MSE \approx \frac{\sigma_w^2}{P_T}$ and $\lim_{P_T\to\infty} MSE \approx \frac{\sigma_v^2}{n}$. The first of these limits is illustrated in Fig. 3 for a constant total network power, as parameterized by the local observation SNR γ_0. It is seen from Figs. 3 and 4 that when local SNR is high the system shows better performance which intuitively makes sense. From Fig. 4, it can be seen that in the region of low local SNR, the performance of the system can be improved by increasing the number of nodes. But in high local SNR region, increasing the number of nodes may not affect the final performance much since ultimately the performance is limited by the channel quality between nodes and the fusion center.

Allocating equal power for all nodes may not result in the best performance since all nodes may not have the same quality observations or communication channels. This is particularly true when one considers channel fading. Let us consider the power allocation among nodes such that the network consumes the minimum possible energy to achieve a desired performance. The optimization problem can be formulated as

$$\min_{g_k \geq 0, k=1,\cdots,n} \sum_{k=1}^{n} g_k^2 \text{ such that } MSE \leq \epsilon_1,$$

where ϵ_1 is the required MSE threshold at the fusion center. If we assume that the local observations are independent, the optimization problem can be rewritten as

$$\min_{g_k \geq 0, k=1,\cdots,n} \sum_{k=1}^{n} g_k^2 \text{ such that } \epsilon_1' - \sum_{k=1}^{n} \frac{h_k^2 g_k^2}{\sigma_v^2 h_k^2 g_k^2 + \sigma_w^2} \leq 0, \quad (17)$$

where we have defined $\epsilon_1' = \frac{1}{\epsilon_1}$. The optimal power allocation strategy is stated in the following lemma assuming that the channel state information (CSI) is available at the distributed nodes.

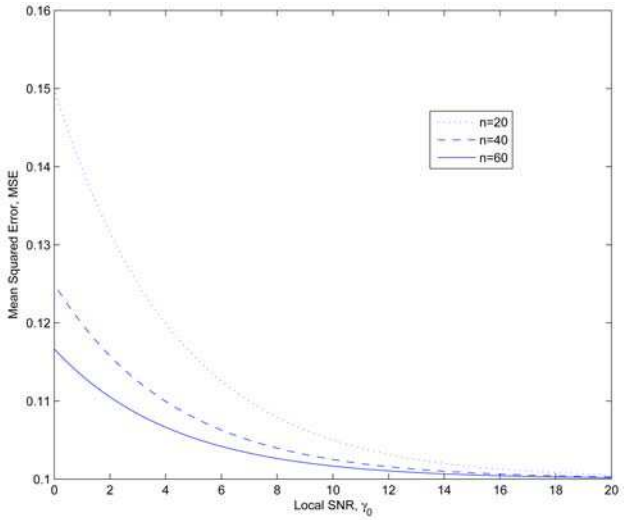

Fig. 4. Performance of mean squared error vs. local SNR, γ_0; The total network power is constant.

Lemma 5: [34] *When local observations are i.i.d., the optimal power allocation solution to (17) is given by*

$$g_k^{*2} = \begin{cases} \frac{\sigma_w^2}{h_k^2 \sigma_v^2} \left[\frac{h_k \sum_{j=1}^{K_1} \frac{1}{h_j}}{(K_1 - \epsilon_1' \sigma_v^2)} - 1 \right] & ; \textit{if } k < K_1 \textit{ and } n > \epsilon_1' \sigma_v^2 \\ 0 & ; \textit{if } k > K_1 \textit{ and } n > \epsilon_1' \sigma_v^2 \\ \textit{infeasible} & ; \textit{if } n < \epsilon_1' \sigma_v^2 \end{cases} \quad (18)$$

where assuming, without loss of generality, $h_1 \geq h_2 \geq \ldots \geq h_n$, K_1 is found such that $s_1(K_1) < 1$ and $s_1(K_1 + 1) \geq 1$ for $1 \leq K_1 \leq n$ where $s_1(k) = \frac{(k - \epsilon_1' \sigma_v^2)}{h_k \sum_{j=1}^{k} \frac{1}{h_j}}$ for $1 \leq k \leq n$.

Lemma 5 says that the optimal power at each node depends on its observation quality, channel quality and the required MSE threshold at the fusion center. Note that letting $\sqrt{\lambda_0} = \sigma_w \frac{\sum_{k=1}^{K_1} \frac{1}{h_k}}{K_1 - \epsilon_1' \sigma_v^2}$, for $s_1(k)$ - 1 < 0 and $n > \epsilon_1' \sigma_v^2$, the optimal g_k^{*2} can be written as $g_k^{*2} = \frac{\sigma_w^2}{h_k^2 \sigma_v^2} \left(\frac{h_k \sqrt{\lambda_0}}{\sigma_w} - 1 \right)$. Hence, when CSI is available at distributed nodes, each node can determine its power using $\sqrt{\lambda_0}$ as a side information that is provided by the fusion center.

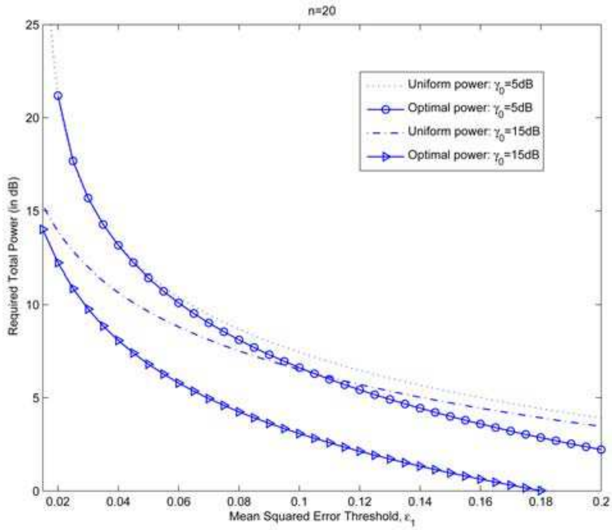

Fig. 5. Optimal power allocation scheme vs. uniform power allocation scheme: The required optimal power to achieve a given MSE of ϵ_1 as given in (18) is shown in the figure parameterized by local SNR γ_0 for n=20 is shown. The comparison of the required uniform power to achieve the same MSE threshold is illustrated.

Figure 5 shows the performance of the optimal power allocation scheme (18) in achieving a desired MSE performance at the fusion center. Figure 5 assumes that fading coefficients are drawn from a Rayleigh distribution with unity mean. It is seen that allocating power optimally as in (18) gives a significant power saving over the uniform power allocation only when either the local observation SNR is high or when the required MSE at the fusion center is not significantly low. This is not surprising since if local observation SNR is high node estimators are good enough on their own and thus perhaps collecting the local estimators from only those nodes with very good fading coefficients can save total power while also meeting the MSE requirement at the fusion center. Moreover, when the MSE required at the fusion center is not very demanding, we may meet it by only collecting local estimators of few nodes (and turning others off), so that the optimal power allocation, may lead to better power savings over the uniform power allocation scheme.

2) Quantized local processing: Recall that with the quantization scheme presented in Section 2-C2, an upper bound for the MSE at the fusion center is given by (9) when the communication between the sensors and the fusion center is noiseless. When discrete messages m_k's are transmitted over noisy communication channels, however, bit errors may occur in a resource constrained network with a finite power.

Let us assume that the discrete messages are transmitted over a noisy channel where bit errors occur due to imperfect communication. Let $\hat{m}_k$ and p_b^k be the decoded message at the fusion center corresponding to the transmitted message m_k from the k-th node and the associate bit error probability, respectively. To compute the resulting MSE of the estimator $\hat{\theta}$ at the fusion center based on the decoded messages $\{\hat{m}_1,\dots, \hat{m}_n\}$, the bit errors caused by the channel should be taken into account. For the quantization scheme presented in Section

2-C2, a complete analysis of the resulting MSE at the fusion center with noisy channels is given in [9]. According to [9], for i.i.d. local observations an upper bound for the MSE, when the messages are transmitted over a memoryless binary symmetric channel is given by the following lemma:

Lemma 6: [9] *If the bit error rates from node k is* p_b^k*, then the MSE achieved by the fusion center based on the decoded messages* $\{\hat{m}_1, ..., \hat{m}_n\}$ *is upper bounded by*

$$MSE \leq (1+p_0)^2 \left(\sum_{k=1}^{n} \frac{1}{\sigma_v^2 + \delta_k^2} \right)^{-1}, \tag{19}$$

where $\delta_k^2 = \frac{W^2}{(2^{L_k}-1)^2}$ *and* $p_0 > 0$ *satisfies the following condition:*

$$p_0 \geq \frac{8W}{\sigma_v} \sqrt{\frac{np_b^k}{3}} \text{ for all } k = 1, \cdots, n.$$

By comparing (19) with (10) it is observed that the achievable MSE with imperfect communication deviates by that with noiseless communication by a constant factor.

Let the communication channel between node k and the fusion center undergo path loss attenuation a_k proportional to d_k^α where d_k is the transmission distance from node k to fusion center and α is the path loss attenuation index. Assuming that node k sends L_k bits using quadrature amplitude modulation (QAM) with constellation size 2^{L_k}, at a bit error probability of p_b^k the transmission power spent by node k is $P_k = B_s E_k$, where B_s is the transmission symbol rate and E_k is the transmission energy per symbol, given by,

$$E_k = \varsigma_k a_k \left(\ln \frac{2}{p_b^k} \right) (2^{L_k} - 1), \tag{20}$$

with $\varsigma_k = 2N_f N_0 G_d$ where N_f is the receiver noise figure, N_0 is the single sided thermal noise spectral density and G_d is a system constant [9].

It can be easily seen from (20) that $L_k = \log \left(1 + \frac{P_k}{B_s \varsigma_k a_k \left(\ln \frac{2}{p_b^k} \right)} \right)$. Thus, to determine the optimal number of bits L_k to be allocated to node k in order to meet a desired MSE performance at the fusion center while minimizing the total network power, [9] solves the following optimization problem (assuming ς_k, B_s and p_b^k are the same for all nodes):

$$\min \|\mathbf{P}\|_2 \text{ such that } MSE \leq \epsilon_2, \tag{21}$$

where $\|\mathbf{P}\|_2 = \sqrt{\sum_{k=1}^{n} P_k}$ is the L^2-norm of the power vector $\mathbf{P} = [P_1, ..., P_n]^T$, ϵ_2 is the desired MSE threshold at the fusion center and the MSE is as given by (19). The optimal number of bits L_k^* to quantize the observations at node k, that is given by the solution to (21), are characterized in the following lemma.

Lemma 7: [9] *The optimal number of bits used to quantize the observations at the k-th node found by solving (21) is*

$$L_k^* = \begin{cases} \log\left(1 + \frac{W}{\sigma_v}\sqrt{\frac{\eta_0}{a_k} - 1}\right) & ;if\ k \leq K_2\ and\ n > \frac{\sigma_v^2}{\epsilon_2'} \\ 0 & ;if\ k \geq K_2 + 1\ and\ n > \frac{\sigma_v^2}{\epsilon_2'} \\ infeasible & ;if\ n < \frac{\sigma_v^2}{\epsilon_2'} \end{cases}$$

where $\eta_0 = \left(\frac{K_2}{\sigma_v^2} - \frac{1}{\epsilon_2'}\right)^{-1} \sum_{k=1}^{K_2} \frac{a_k}{\sigma_v^2}$, $\epsilon_2' = \frac{\epsilon_2}{(1+p_0)^2}$ *and assuming, without loss of generality,* $a_1 \leq a_2 \leq \ldots \leq a_n$, K_2 *is found such that* $s_2(K_2) < 1$ *and* $s_2(K_2 + 1) \geq 1$ *for* $1 \leq K_2 \leq n$ *where* $s_2(k) = a_k\left(\frac{k}{\sigma_v^2} - \frac{1}{\epsilon_2'}\right)\left(\sum_{j=1}^{k} \frac{a_k}{\sigma_v^2}\right)^{-1}$. *Then the optimal transmission power at the k-th node is given by*

$$P_k^* = \varsigma_k B_s \left(\ln \frac{2}{p_b}\right) \frac{W h_k}{\sigma_v} \sqrt{\left(\frac{\eta_0}{a_k} - 1\right)^+},$$

where $(x)^+$ *equals to zero if* $x < 0$ *and equals to x otherwise.*

Note that again the optimal power at node k is determined by the observation quality, channel quality and the required MSE threshold as was the case with AF local processing we saw in lemma 5. Figure 6 shows the number of sensors that are active in the network to achieve a desired MSE threshold at the fusion center. In Fig. 6, the network size $n = 1000$ and $\alpha = 2$. The distance from node k to fusion center, d_k, is drawn from a uniform distribution on [1, 2]. It is observed that when the required MSE threshold increases the number of active sensors decreases greatly. That is, the network discards the observations at nodes with poor observation and channel quality. This is similar to what we observed in Fig. 5 earlier.

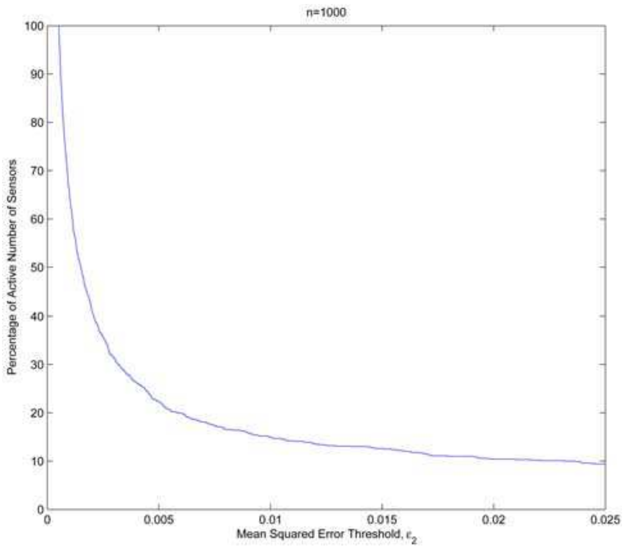

Fig. 6. Number of active sensors in the network according to the optimal power allocation scheme given by lemma 7. The number of total sensors in the network is $n = 1000$ and $\alpha = 2$

Figure 7 shows the energy saving due to the optimal power allocation scheme given in lemma 7 compared to the uniform power allocation scheme. Clearly Fig. 7 shows that significant energy savings are possible by optimally selecting number of bits, especially at moderate levels of desired MSE at the fusion center.

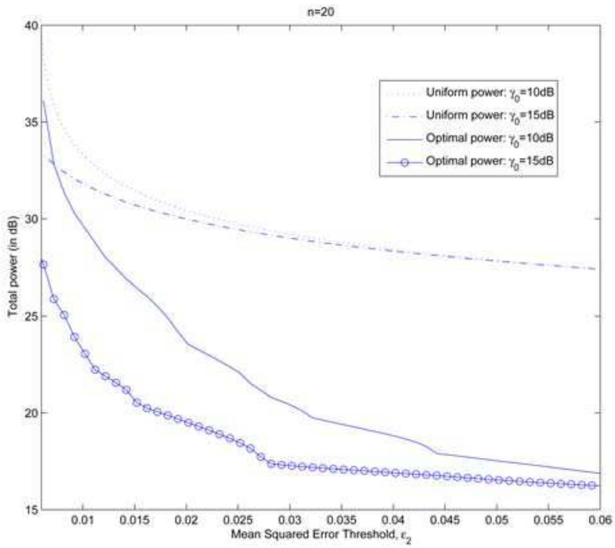

Fig. 7. Performance of optimal power allocation scheme given in lemma 7 vs. uniform power allocation scheme: network size n = 20

B. Communication over multiple access channels

One of the disadvantages of using orthogonal channels to transmit local decisions is the large bandwidth consumption as the number of distributed nodes n increases. An alternative is to allow multiple sensor nodes to share a common channel. Such multiple access communication (MAC) in bandwidth constrained wireless sensor networks has been investigated in, among others, [10], [11], [17], [24], [25], [34], [36], [44]. For example, in [24], [25], [44] the authors proposed a type based multiple-access communication in which sensors transmit according to the type of their observation in a shared channel where the *type* is as defined in [45]. An analysis of both orthogonal and MAC channels for distributed detection in a sensor network was presented in [44]. MAC with correlated observations was considered in [34] and [46]. The use of CDMA signaling in distributed detection of deterministic and Gaussian signals under strict power constraints was presented in [10] and [11], respectively. When all sensor nodes communicate with the fusion center coherently, with amplify-and-forward local processing the estimator performance can be improved compared to that of orthogonal communication due to the coherent beam-forming gain [47], [46]. Performance of MAC communication with asynchronous transmissions was discussed in [48].

In the following we consider the form and performance of the final estimator at the fusion center when communications from distributed nodes to the fusion center is over noisy multiple-access channels. Assuming perfect synchronization among sensor transmissions, the received signal at the fusion center over a MAC can be written as

$$r = \sum_{k=1}^{n} h_k y_k + w,$$

where w is the receiver noise with zero mean and variance of σ_w^2 and h_k is the channel fading coefficient from node k to the fusion center, as defined earlier. For the AF local processing, substituting $y_k = g_k z_k$, the resulting received signal is given by

$$r = \sum_{k=1}^{n} h_k g_k z_k + w. \tag{22}$$

Fusion center computes the final estimator based on the received coherent signal r. The resulting BLUE estimator and its performance is given by the following lemma.
Lemma 8: [34] *The BLUE estimator and the resulting MSE based on the received signal (22) can be shown to be*

$$\hat{\theta}(r) = \frac{r}{\sum_{k=1}^{n} h_k g_k},$$

and

$$MSE = \frac{\mathbf{e^T G H_c \Sigma_v H_c G e} + \sigma_w^2}{(\mathbf{e^T H_c G e})^2}.$$

With i.i.d. local observations the MSE simplifies to

$$MSE = \frac{\sigma_v^2 \sum_{k=1}^{n} h_k^2 g_k^2 + \sigma_w^2}{(\sum_{k=1}^{n} h_k g_k)^2}.$$

The MSE performance of the BLUE estimator under both orthogonal and multiple-access channels, with i.i.d. observations, is depicted in Fig. 8 as a function of total network power. Figure 8 assumes equal node powers and unity channel gains. Moreover, MAC communication is assumed to be perfectly synchronized among nodes. As seen from Fig. 8, when total network power is small, the MAC communication leads to a better MSE performance compared to that with orthogonal communication. But as total network power increases both schemes converge to the same performance level. This is because when the network can afford a large transmission power, irrespective of the communication scheme the overall estimator performance is only limited by the local observation quality and the effects of additive/multiplicative channel noise is mitigated by the large gain in the transmission. However, when a practical sensor network is power-constrained the MAC communication may be able to provide a much better performance over that of the orthogonal transmissions when nodes are perfectly synchronized.
Figure 8 assumes equal transmission powers at all nodes. However, when the fusion center needs to achieve only a target estimator quality, say an MSE of ϵ_3, one can consider non-uniform power allocations such that,

$$\min_{g_k \geq 0, k=1,\cdots,n} \sum_{k=1}^{n} g_k^2 \text{ such that } MSE \leq \epsilon_3, \tag{23}$$

where MSE is as given in lemma 8. When the observations are i.i.d., a tractable analytical solution for the above optimization problem was given in [34] that is stated in the following lemma.

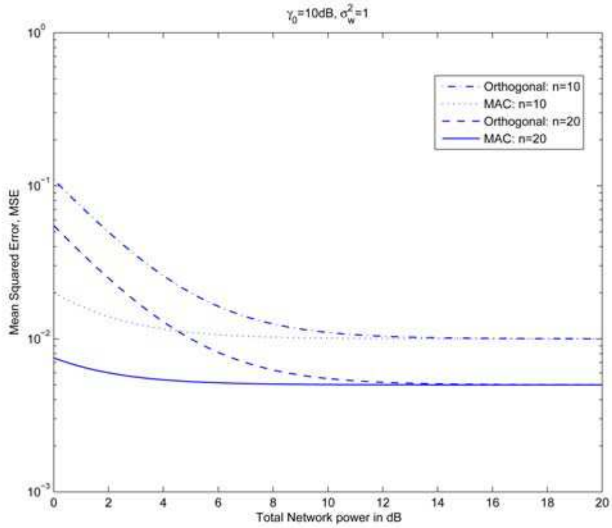

Fig. 8. Mean squared error performance vs. total network power for different network sizes

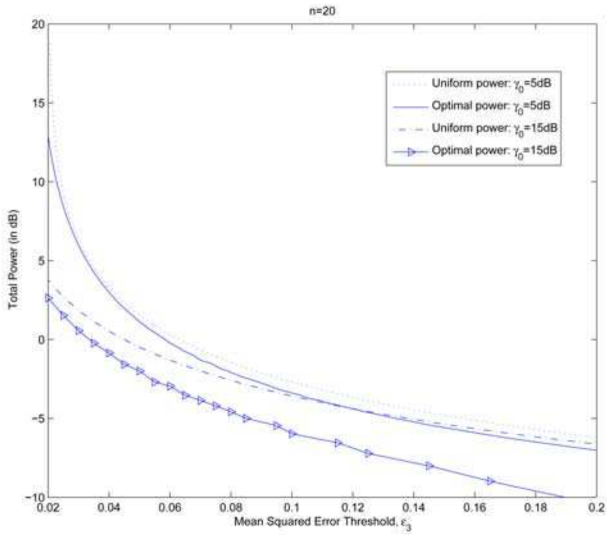

Fig. 9. Performance of optimal power allocation scheme vs. uniform power allocation scheme

Lemma 9: With i.i.d. local observations, the optimal power at node k, g_k^{*2} *that solves the optimization problem in (23) with the MSE as given in lemma 8 is* $g_k^{*2} = \frac{\mu^2}{4}\frac{h_k^2}{(1+\eta_0 h_k^2)^2}$ *for* k = 1, 2,..., *n*, where η_0 *and* μ *can be found numerically by solving the equations* $\sum_{k=1}^{n}\frac{\eta_0 h_k^2}{(1+\eta_0 h_k^2)} = \frac{1}{\epsilon_3^{'}}$ *and* $\mu = 2\frac{\sigma_w}{\sigma_v}\left(\frac{1}{\eta_0^2\epsilon_3^{'}} - \sum_{k=1}^{n}\frac{h_k^4}{(1+\eta_0 h_k^2)^2}\right)^{-\frac{1}{2}}$ *where* $\epsilon_3^{'} = \frac{\epsilon_3}{\sigma_v^2}$.

It is observed that η_0 has a feasible solution only when $n > \frac{\sigma_v^2}{\epsilon_3}$ [34]. The total power spent by the network with the above optimal power allocation scheme is given by $P_{total} = \sum_{k=1}^{n} g_k^{*2} = \frac{\sigma_w^2}{\sigma_v^2}\eta_0$.

Figure 9 shows the performance of the optimal power allocation scheme compared to that of uniform power allocation scheme for a network size of $n = 20$. Again, the optimal power scheduling scheme has a significant performance gain over the uniform power allocation scheme especially when local SNR is high or the required MSE threshold at the fusion center is moderate, similar to what was observed in Section 3-A in the case of orthogonal communication.

C. Effects of synchronization errors on MAC

To achieve coherent gain with MAC transmissions, it is important that the sensor transmissions are synchronized. For this discussion on node synchronization, we assume, i.i.d observations and AF local processing. Analysis would remain essentially the same for other network models as well.

To achieve synchronization in node transmissions, one may assume that there is a master-node (that can be taken as the fusion center itself, for simplicity) that broadcasts the carrier and timing signals to the distributed nodes [47]. Suppose that the k-th node is located at a distance of $d_k + \delta_k$ from the fusion center, for $k = 1, 2,..., n$, where d_k and δ_k are the nominal distance and the sensor placement error of the k-th node, respectively. The fusion center broadcasts a carrier signal $\cos(2\pi f_0 t)$ where f_0 is the carrier frequency. The received carrier signal at the k-th node is a noisy version of $\cos(2\pi f_0 t + \psi_k + \psi_{ek})$ where $\psi_k = \frac{2\pi f_0 d_k}{c}$ and $\psi_{ek} = \frac{2\pi f_0 \delta_k}{c}$. Each node employs a Phase Locked Loop (PLL) to lock onto the carrier. If each node precompensates for the difference in their nominal distances d_k, by transmitting its locally processed and modulated observation with a proper delay and phase shift ψ_k, then the received signal at the fusion center is corrupted only by the phase shift due to the sensor placement error δ_k. Considering only the phase shift due to this sensor placement error, the received signal at the fusion center is given by $r = \sum_{k=1}^{n} h_k g_k z_k \cos(\psi_{ek}) + w$, assuming AF local processing at sensor nodes. In the following lemma we assume that the placement error δ_k is Gaussian with zero mean and variance σ_δ^2.

Lemma 10: [34] *Assuming that* $\sigma_\delta^2 << \iota_0$ *where* $\iota_0 = \frac{c}{f_0}$, *so that phase error* $\psi_{ek} \sim \mathcal{N}(0, \sigma_\psi^2)$ *where* σ_ψ^2 *is small, the BLUE estimator at the fusion center when local observations are i.i.d. is* $\hat{\theta}(r) = r\left(e^{-\frac{\sigma_\psi^2}{2}} \sum_{k=1}^{n} h_k g_k\right)^{-1}$. *The resulting MSE with the phase synchronization errors is*

$$MSE = \frac{\sigma_v^2 \sum_{k=1}^{n} h_k^2 g_k^2 + e^{\sigma_\psi^2}\sigma_w^2}{\left(\sum_{k=1}^{n} h_k g_k\right)^2}. \tag{24}$$

Figure 10 shows the MSE performance (24) of a sensor system in the presence of phase synchronization errors. It can be seen that the performance is robust against synchronization errors as long as the variance of the phase error σ_ψ^2 is sufficiently small.