

# Distributed parameter estimation in wireless sensor networks using fused local observations

Mohammad Fanaei, Matthew C. Valenti, Natalia A. Schmid, and Marwan M. Alkhweldi

Lane Department of Computer Science and Electrical Engineering  
West Virginia University, Morgantown, WV, USA

## ABSTRACT

The goal of this paper is to reliably estimate a vector of unknown deterministic parameters associated with an underlying function at a fusion center of a wireless sensor network based on its noisy samples made at distributed local sensors. A set of noisy samples of a deterministic function characterized by a finite set of unknown parameters to be estimated is observed by distributed sensors. The parameters to be estimated can be some attributes associated with the underlying function, such as its height, its center, its variances in different directions, or even the weights of its specific components over a predefined basis set. Each local sensor processes its observation and sends its processed sample to a fusion center through parallel impaired communication channels. Two local processing schemes, namely analog and digital, are considered. In the analog local processing scheme, each sensor transmits an amplified version of its local analog noisy observation to the fusion center, acting like a relay in a wireless network. In the digital local processing scheme, each sensor quantizes its noisy observation before transmitting it to the fusion center. A flat-fading channel model is considered between the local sensors and fusion center. The fusion center combines all of the received locally-processed observations and estimates the vector of unknown parameters of the underlying function. Two different well-known estimation techniques, namely maximum-likelihood (ML), for both analog and digital local processing schemes, and expectation maximization (EM), for digital local processing scheme, are considered at the fusion center. The performance of the proposed distributed parameter estimation system is investigated through simulation of practical scenarios for a sample underlying function.

**Keywords:** Distributed parameter estimation, maximum likelihood (ML), expectation maximization (EM), analog/digital local processing, local quantization, integrated mean-squared error, flat-fading channel, fusion center, wireless sensor networks.

## 1. INTRODUCTION

Wireless sensor networks (WSNs) are typically formed by a large number of densely-deployed geographically-distributed sensors with limited capabilities that cooperate with each other to achieve a common goal. One of the most important applications of such networks is distributed parameter estimation, which is the first step in a wider range of applications such as event detection and classification. In a WSN performing distributed parameter estimation, distributed local sensors observe the conditions of their surrounding environment, process their local observations, and send their processed data to a fusion center, which then combines all of the locally-processed samples to perform the ultimate global estimation. This paper studies the problem of estimating a vector of unknown parameters associated with a deterministic function at the fusion center of a WSN from its distributed noisy samples observed by local sensors and communicated through parallel flat-fading channels. Two local processing schemes, namely *analog* and *digital*, will be considered. In the analog local processing scheme, each sensor acts as a pure relay and transmits an amplified version of its raw analog local observation to the fusion center. In the digital local processing method, each sensor quantizes its local noisy observation and sends its quantized sample to the fusion center using a digital modulation format. A major application of this work could be in cases, where the estimated parameters can then be used in detection and classification of the underlying object that has created the observed influence-field intensity function. The main contribution of this paper is a generalized formulation of distributed parameter estimation in the context of wireless sensor

networks, where local observations are not linearly dependent on the underlying parameters to be estimated and no specific observation model has been considered in the analysis.

Xiao et al.<sup>1</sup> have considered the *linear* coherent decentralized mean-squared error (MSE) estimation of an unknown vector under stringent bandwidth and power constraints, where the local observation models, compression functions at local sensors, and the fusion rule at the fusion center are all linear. As a result of the bandwidth constraint, each sensor in their proposed framework transmits to the fusion center a fixed number of real-valued messages per observation. The power constraint in their model limits the strength of the transmitted signals. Based on their proposed algorithm, each sensor linearly encodes its observations using a linear amplify-and-forward coding strategy, and the fusion center applies a linear mapping to estimate the unknown vector based on the received messages from local sensors through a coherent multiple-access channel (MAC) assuming a perfect synchronization between sensors and the fusion center so that the transmitted messages from local sensors can be coherently combined at the fusion center. It is shown in<sup>2</sup> that if the sensor statistics are Gaussian, i.e., the parameter to be estimated and the independent observation noises are Gaussian, and the communication channels between local sensors and the fusion center are standard Gaussian multiple-access channels, a simple uncoded analog-and-forward scheme, in which each sensor's channel input is merely a scaled version of its noisy observation, drastically outperforms the separate source-channel coding approach in the sense of mean-squared error. Therefore, the proposed decentralized estimation algorithm in<sup>1</sup> can perform optimally in applications that satisfy the requirements summarized in<sup>2</sup>. In this paper, a general non-linear model for the distributed local observations is considered and analyzed.

Ishvar et al.<sup>3</sup> have considered the problem of estimation with unreliable communication links and have derived an information-theoretic achievable rate-distortion region characterizing the per-sample sensor bit-rate  $R$  versus the estimation error  $D$ . In particular, they have considered two cases of fully-distributed estimation schemes with no inter-sensor collaboration and localized or collaborative estimation schemes in which it is assumed that the network is divided into a number of sensor clusters, where collaboration is allowed among sensors within the same cluster but not across clusters. Nowak et al.<sup>4</sup> have studied the theoretical upper and lower bounds and tradeoffs between the estimation error and energy consumption in WSNs as functions of local sensor density. They have proposed practical in-network processing approaches based on hierarchical data handling and multi-scale partitioning methods for the estimation of two-dimensional inhomogeneous fields composed of two or more homogeneous smoothly-varying regions separated by smooth boundaries.

Ribeiro and Giannakis<sup>5</sup> have proposed a distributed bandwidth-constrained maximum-likelihood (ML) estimation method for estimating a scalar deterministic parameter in the presence of zero-mean additive white Gaussian observation noise using only quantized versions of the original local observations perfectly available at the fusion center. In a sequel work, they have considered more realistic signal models such as known univariate but generally non-Gaussian noise probability density functions (pdfs), known noise pdfs with a finite number of unknown parameters, completely unknown noise pdfs, and practical generalizations to vector parameters and multivariate and possibly correlated noise pdfs<sup>6</sup>. The observation model in this work could in general be a non-linear function of the vector parameter to be estimated, but it still ignores the effects of imperfect wireless channels between local sensors and the fusion center. It is shown in these works that transmitting a few bits (or even a single bit) per sensor can approach the performance of the estimator based on unquantized data under realistic conditions. In this paper, a general observation model will be considered in which local observations are not necessarily linear functions of the vector of parameters to be estimated. Furthermore, impaired flat-fading channels between local sensors and the fusion center will be considered in the analysis.

Niu and Varshney<sup>7</sup> have proposed a signal intensity- (or energy-) based ML target location estimation scheme to estimate the coordinates of an energy emitting source using quantized local observations, which are assumed to be perfectly received by the fusion center. They have considered an isotropic signal intensity attenuation model in which the intensity of the signal is assumed to be inversely proportional to the square of the distance from its source. An optimal infeasible as well as two heuristic practical design methods for optimal local quantization thresholds are proposed by them that minimize the summation of estimation error variances for the target's two coordinates. In a sequel work, Ozdemir et al.<sup>8</sup> have added the effects of imperfect fading and noisy wireless communication channels between local sensors and the fusion center to their problem formulation and have developed similar results. Maşazade et al.<sup>9</sup> have considered a similar problem in which the quantized version

of sensor measurements are transmitted to the fusion center over error-free channels. They have proposed an energy-efficient iterative source localization scheme in which the algorithm begins with a coarse location estimate obtained from measurement data from a set of anchor sensors. Based on the accumulated information at each iteration, the posterior pdf of the source location is approximated using an importance-sampling-based Monte Carlo method, which is then used to activate a number of non-anchor sensors selected to improve the accuracy of the source location estimate the most. Distributed compression of measurement data prior to transmission is also employed at the non-anchor sensors to further reduce the energy consumption. It is shown that this iterative scheme reduces the communication requirements by selecting only the most informative sensors and compressing their data prior to transmission to the fusion center. An iterative non-linear least-square received signal strength (RSS)-based location estimation technique is also proposed in<sup>10</sup> for joint estimation of unknown location coordinates and distance-power gradient, which is a parameter of radio propagation path-loss model. Salman et al.<sup>11</sup> have proposed a low-complexity version of this approach. Although the illustrative case study for the numerical simulation results in this paper considers estimating the location of the center of a generic Gaussian function, among its other parameters, its scope is not limited to estimating only the location. In other words, this paper considers a more general framework than the ones considered in the aforementioned works and is not based on the signal propagation model in the observation environment.

The rest of this paper is organized as follows: Section 2 describes the model of the distributed parallel fusion WSN that will be considered in our analysis and defines the precise problem that we are considering in this paper. In Section 3, the ML estimate of a vector of unknown parameters associated with a deterministic two-dimensional function is derived for the case of analog local processing scheme. Section 4 considers the same problem for the case of digital signal processing method. As it is shown in this section, the ML estimate in this case is infeasible to be found efficiently in a distributed manner. Therefore, a linearized version of the expectation maximization (EM) algorithm will be proposed in Section 5 to numerically find the ML estimate of the vector of unknown parameters in the case of digital local processing scheme. Section 6 presents the numerical results of our simulations to study the performance of the proposed distributed estimation framework for a special two-dimensional Gaussian-shaped function, whose associated parameters to be estimated are its height and center. The effects of different parameters of the WSN on the performance of the proposed system will be studied in this section. Finally, in Section 7 we conclude our discussions and summarize the main achievements of this work.

**Notations.** Throughout this paper, the following notations are used. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and standard lower-case letters denote scalars. The superscripts  $(\cdot)^T$  and  $(\cdot)^{-1}$  denote the matrix transpose and matrix inverse operators, respectively. The operator  $\text{diag}(\cdot)$  manipulates the diagonal elements of a matrix or puts a column vector into a diagonal matrix.  $|\cdot|$  is the absolute value of a complex number or the determinant of a square matrix.

## 2. SYSTEM MODEL AND PROBLEM STATEMENT

Consider a set of  $K$  spatially-distributed sensors forming a WSN as shown in Fig. 1. Suppose that the sensors are located within the domain of a two-dimensional function  $g(x, y)$  that is completely known except for a set of unknown deterministic parameters denoted by  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_p]^T$ . The ultimate goal of the underlying WSN is to reliably estimate the vector of unknown parameters  $\boldsymbol{\theta}$  using distributed noisy samples of function  $g(x, y)$  provided by local sensors to a fusion center through parallel impaired flat-fading channels.

Assume that the  $i$ th sensor,  $i = 1, 2, \dots, K$ , observes a noisy version of a sample of function  $g(x, y)$  at its location as

$$r_i = g(x_i, y_i) + w_i, \quad (1)$$

where  $r_i$  is the local sensor's noisy observation,  $(x_i, y_i)$  is the location of the  $i$ th sensor in the network,  $g_i \stackrel{\text{def}}{=} g(x_i, y_i)$  is the sample of function  $g(x, y)$  at location  $(x_i, y_i)$ , and  $w_i$  is zero-mean spatially-uncorrelated additive white Gaussian noise with variance  $\sigma_{O_i}^2$ , i.e.  $w_i \sim \mathcal{N}(0, \sigma_{O_i}^2)$ . Sensors can be placed on a uniform lattice or distributed at random over the observation area covered by the WSN. Throughout the discussions in this paper, we assume that the location of distributed sensors is known at the fusion center.

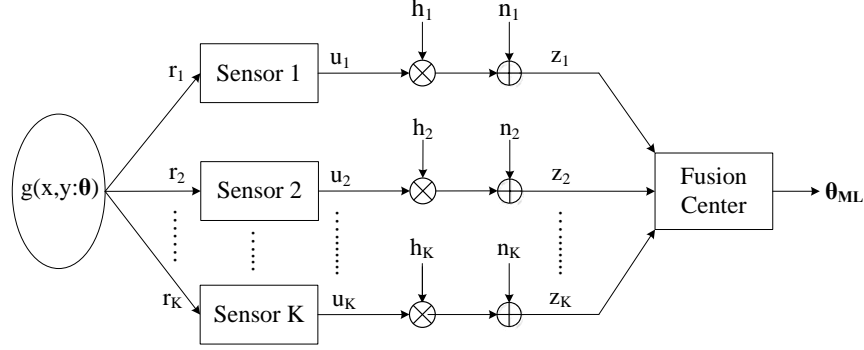


Figure 1: System model of a typical WSN for distributed estimation of a deterministic parameter vector  $\theta$ .

Note that the observation model introduced in Equation (1) covers a number of variations in different applications. For example, it is a generalization of the well-studied linear observation model in which  $r_i = \mathbf{a}_i^T \theta + w_i$ , where  $\mathbf{a}_i = [a_{i_1}, a_{i_2}, \dots, a_{i_p}]^T$  is the local observation gain vector for sensor  $i$  (considered for example by Ribeiro and Giannakis<sup>5</sup>). Furthermore, it is also a generalized version of the observation model considered by Niu and Varshney<sup>7</sup> and Ozdemir et al.<sup>8</sup>

Each sensor  $i$  processes its local observation based on its local processing rule  $\gamma_i(\cdot)$  and sends its result, denoted by  $u_i = \gamma_i(r_i)$ , to the fusion center through an impaired communication channel. In this paper, two local processing schemes, namely *analog* and *digital*, are considered. In the analog local processing scheme, each sensor acts as a pure relay and transmits an amplified version of its raw analog local observation to the fusion center. In the digital local processing method, each sensor quantizes its local observation and sends its quantized sample to the fusion center using a digital modulation format. We will formulate detailed analysis of the local analog processing scheme in Section 3 and that of the local digital processing scheme in Sections 4 and 5.

Suppose that locally-processed sample is transmitted to the fusion center over parallel independent flat fading channels. In other words, assume that the received signal from sensor  $i$  at the fusion center is

$$z_i = h_i u_i + n_i, \quad i = 1, 2, \dots, K, \quad (2)$$

where  $h_i$  is the fading gain of the channel between sensor  $i$  and the fusion center and  $n_i$  is zero-mean spatially-uncorrelated additive white Gaussian noise with variance  $\sigma_{C_i}^2$ , i.e.  $n_i \sim \mathcal{N}(0, \sigma_{C_i}^2)$ . It is assumed that channel fading gains are uncorrelated and *completely known* at the fusion center.

Upon receiving the vector of locally-processed observations from distributed sensors  $\mathbf{z} = [z_1, z_2, \dots, z_K]^T$ , the fusion center combines them to estimate the vector of unknown deterministic parameters  $\theta$ . In the following three sections, we will formulate the ML estimation of  $\theta$  at the fusion center for the two cases of analog and digital local processing schemes and the EM algorithm to numerically find the ML estimate of  $\theta$  at the fusion center for the digital local processing scheme.

### 3. ML ESTIMATION OF $\theta$ : ANALOG LOCAL PROCESSING CASE

Suppose that each sensor acts as a pure relay and transmits an amplified version of its local raw analog observation to the fusion center. In other words,

$$u_i = \alpha_i r_i, \quad i = 1, 2, \dots, K, \quad (3)$$

where  $\alpha_i$  is the local amplification gain at sensor  $i$ , assumed to be known at the fusion center. Note that in this paper, we do not try to optimize the local amplification gains as it is considered by Banavar et al.<sup>12</sup> for the problem of distributed scalar random parameter estimation. The received signal from sensor  $i$  at the fusion center can then be represented as

$$\begin{aligned} z_i &= h_i \alpha_i r_i + n_i \\ &= h_i \alpha_i g_i + h_i \alpha_i w_i + n_i, \quad i = 1, 2, \dots, K. \end{aligned} \quad (4)$$

Assuming complete knowledge of the local observation amplification gains and channel fading gains, the fusion center can make a linear transformation on the received signal from each local sensor to find

$$\begin{aligned} z'_i &= (h_i \alpha_i)^{-1} z_i \\ &= g_i + w_i + (h_i \alpha_i)^{-1} n_i \\ &= g_i + v_i \quad i = 1, 2, \dots, K, \end{aligned} \quad (5)$$

where  $v_i \stackrel{\text{def}}{=} w_i + \frac{1}{h_i \alpha_i} n_i$  is a zero-mean white Gaussian random variable with variance  $\sigma_i^2 \stackrel{\text{def}}{=} \sigma_{O_i}^2 + \frac{\sigma_{C_i}^2}{|h_i \alpha_i|^2}$ . The estimation process at the fusion center can then be performed based on vector  $\mathbf{z}' = [z'_1, z'_2, \dots, z'_K]^T$ .

The probability density function of the linearly-processed received vector of local observations from distributed sensors at the fusion center is given by

$$f_{\mathbf{z}'}(\mathbf{z}' : \boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^K |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{z}' - \mathbf{g})^T \boldsymbol{\Sigma}^{-1} (\mathbf{z}' - \mathbf{g})\right), \quad (6)$$

where  $\boldsymbol{\Sigma} \stackrel{\text{def}}{=} \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2)$  is the diagonal matrix of cumulative noise variances and  $\mathbf{g} = [g_1, g_2, \dots, g_K]^T$  is the vector of samples of function  $g(x, y)$  at sensor locations. The joint log-likelihood function of the vector of unknown parameters  $\boldsymbol{\theta}$  can then be found as

$$\begin{aligned} l(\boldsymbol{\theta}) &= \ln f_{\mathbf{z}'}(\mathbf{z}' : \boldsymbol{\theta}) \\ &\equiv -\left[\ln(|\boldsymbol{\Sigma}|) + (\mathbf{z}' - \mathbf{g})^T \boldsymbol{\Sigma}^{-1} (\mathbf{z}' - \mathbf{g})\right] \\ &\stackrel{(a)}{\equiv} -(\mathbf{z}' - \mathbf{g})^T \boldsymbol{\Sigma}^{-1} (\mathbf{z}' - \mathbf{g}) \\ &= -\sum_{i=1}^K \frac{1}{\sigma_i^2} (z'_i - g_i)^2, \end{aligned} \quad (7)$$

where (a) is based on the fact that  $|\boldsymbol{\Sigma}|$  is independent of  $\boldsymbol{\theta}$ .

The maximum-likelihood estimate of the vector of unknown parameters  $\boldsymbol{\theta}$  is the maximizer of its log-likelihood function given in Equation (7). In other words,

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{\text{ML}} &= \arg \max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}) \\ &= \arg \min_{\boldsymbol{\theta}} (\mathbf{z}' - \mathbf{g})^T \boldsymbol{\Sigma}^{-1} (\mathbf{z}' - \mathbf{g}) \\ &= \arg \min_{\boldsymbol{\theta}} \sum_{i=1}^K \frac{1}{\sigma_i^2} (z'_i - g_i)^2. \end{aligned} \quad (8)$$

Based on Equation (8), the ML estimate of the vector of unknown parameters  $\hat{\boldsymbol{\theta}}_{\text{ML}}$  can be found as the solution of the following system of equations:

$$\nabla_{\boldsymbol{\theta}} l(\boldsymbol{\theta}) \big|_{\hat{\boldsymbol{\theta}}_{\text{ML}}} = \mathbf{0}, \quad (9)$$

where  $\nabla_{\boldsymbol{\theta}} \stackrel{\text{def}}{=} \left[ \frac{\partial}{\partial \theta_1}, \frac{\partial}{\partial \theta_2}, \dots, \frac{\partial}{\partial \theta_p} \right]^T$  is the gradient operator with respect to  $\boldsymbol{\theta}$ . This system of equations can be simplified in the vector form by substituting  $l(\boldsymbol{\theta})$  from Equation (7) as

$$\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} \boldsymbol{\Sigma}^{-1} (\mathbf{z}' - \mathbf{g}) = \mathbf{0}, \quad (10)$$

where  $\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}$  is a  $p \times K$  matrix of partial derivatives of the components of vector  $\mathbf{g}$  with respect to the components of  $\boldsymbol{\theta}$  defined as

$$\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial g_1}{\partial \theta_1} & \frac{\partial g_2}{\partial \theta_1} & \dots & \frac{\partial g_K}{\partial \theta_1} \\ \frac{\partial g_1}{\partial \theta_2} & \frac{\partial g_2}{\partial \theta_2} & \dots & \frac{\partial g_K}{\partial \theta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial \theta_p} & \frac{\partial g_2}{\partial \theta_p} & \dots & \frac{\partial g_K}{\partial \theta_p} \end{bmatrix}. \quad (11)$$

The system of equations derived in Equation (10) can be rewritten in the scalar format as

$$\sum_{i=1}^K \left[ \frac{1}{\sigma_i^2} \left( \frac{\partial g_i}{\partial \theta_j} \right) (z'_i - g_i) \right] = 0, \quad j = 1, 2, \dots, p. \quad (12)$$

This system of equations is highly non-linear in unknown parameters for most practical applications and does not have a closed-form solution. In Section 6, this system will be solved using numerical methods to estimate a vector of unknown parameters associated with a specific two-dimensional Gaussian-shaped function of interest  $g(x, y)$ .

#### 4. ML ESTIMATION OF $\boldsymbol{\theta}$ : DIGITAL LOCAL PROCESSING CASE

Suppose that sensor  $i$  quantizes its local observation  $r_i$  to  $b_i = \log_2 M_i$  bits and sends the index of its quantized multi-bit sample to the fusion center, where  $M_i$  is the number of quantization levels for sensor  $i$ . Let  $\mathcal{L}_i = \{\beta_i(0), \beta_i(1), \dots, \beta_i(M_i)\}$  be the set of quantization thresholds for sensor  $i$ , where  $\beta_i(\ell)$  is the  $\ell$ th quantization threshold of the  $i$ th sensor,  $\beta_i(0) = -\infty$ , and  $\beta_i(M_i) = \infty$ ,  $i = 1, 2, \dots, K$ . The local processing rule of sensor  $i$  is then defined as a function  $\gamma_i : \mathbb{R} \mapsto \{0, 1, \dots, M_i - 1\}$ , whose values are determined as

$$u_i = \ell \iff \beta_i(\ell) \leq r_i < \beta_i(\ell + 1), \quad \ell = 0, 1, \dots, M_i - 1, \text{ and } i = 1, 2, \dots, K. \quad (13)$$

Based on the local observation model introduced in Equation (1), the probability density function of each sensor's quantized output sample can be found as

$$f_{U_i}(u_i : \boldsymbol{\theta}) = \sum_{\ell=0}^{M_i-1} \Delta Q_i(\ell) \delta[u_i - \ell], \quad u_i = 0, 1, \dots, M_i - 1, \text{ and } i = 1, 2, \dots, K, \quad (14)$$

where  $\delta[\cdot]$  denotes the discrete Dirac delta function,  $\Delta Q_i(\ell)$  is defined as

$$\Delta Q_i(\ell) \stackrel{\text{def}}{=} Q\left(\frac{\beta_i(\ell) - g_i}{\sigma_{O_i}}\right) - Q\left(\frac{\beta_i(\ell + 1) - g_i}{\sigma_{O_i}}\right), \quad \ell = 0, 1, \dots, M_i - 1, \text{ and } i = 1, 2, \dots, K, \quad (15)$$

and  $Q(\cdot)$  is the complementary distribution function of the standard Gaussian random variable defined as

$$Q(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt. \quad (16)$$

Like in the case of analog local processing, assuming complete knowledge of the channel fading gains, the fusion center can make a linear transformation on the received signal from each local sensor to find

$$\begin{aligned} z'_i &= h_i^{-1} z_i \\ &= u_i + \frac{1}{h_i} n_i \\ &= u_i + v_i \end{aligned} \quad i = 1, 2, \dots, K, \quad (17)$$

where  $v_i \stackrel{\text{def}}{=} \frac{1}{h_i} n_i$  is a zero-mean Gaussian random variable with variance  $\sigma_i^2 \stackrel{\text{def}}{=} \frac{\sigma_{C_i}^2}{|h_i|^2}$ . The estimation process at the fusion center can then be performed based on vector  $\mathbf{z}' = [z'_1, z'_2, \dots, z'_K]^T$ .

The probability density function of the linearly-processed received vector of local observations from distributed sensors at the fusion center is given by

$$\begin{aligned}
f_{\mathbf{z}'}(\mathbf{z}' : \boldsymbol{\theta}) &\stackrel{(a)}{=} \prod_{i=1}^K f_{Z'_i}(z'_i : \boldsymbol{\theta}) \\
&\stackrel{(b)}{=} \prod_{i=1}^K \sum_{u_i=0}^{M_i-1} f_{Z'_i|U_i}(z'_i|u_i) f_{U_i}(u_i : \boldsymbol{\theta}) \\
&= \prod_{i=1}^K \sum_{u_i=0}^{M_i-1} \left[ \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left\{ -\frac{(z'_i - u_i)^2}{2\sigma_i^2} \right\} \right] \left[ \sum_{\ell=0}^{M_i-1} \Delta Q_i(\ell) \delta[u_i - \ell] \right] \\
&= \prod_{i=1}^K \sum_{\ell=0}^{M_i-1} \sum_{u_i=0}^{M_i-1} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[ -\frac{(z'_i - u_i)^2}{2\sigma_i^2} \right] \Delta Q_i(\ell) \delta[u_i - \ell] \\
&\stackrel{(c)}{=} \prod_{i=1}^K \frac{1}{\sqrt{2\pi\sigma_i^2}} \sum_{\ell=0}^{M_i-1} \Delta Q_i(\ell) \exp \left[ -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right], \tag{18}
\end{aligned}$$

where (a) is based on the fact that the received local observations from distributed sensors are spatially uncorrelated, the summation in (b) is based on the theorem of total probability and over all  $M_i$  possible realizations of each  $u_i$ , and (c) is based on the fact that the inner summation over  $u_i$  can be simplified using the properties of discrete Dirac delta function as

$$\sum_{u_i=0}^{M_i-1} \exp \left[ -\frac{(z'_i - u_i)^2}{2\sigma_i^2} \right] \Delta Q_i(\ell) \delta[u_i - \ell] = \exp \left[ -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right] \Delta Q_i(\ell). \tag{19}$$

Therefore, the joint log-likelihood function of the vector of unknown parameters  $\boldsymbol{\theta}$  can be found as

$$\begin{aligned}
l(\boldsymbol{\theta}) &= \ln f_{\mathbf{z}'}(\mathbf{z}' : \boldsymbol{\theta}) \\
&\equiv \sum_{i=1}^K \ln \left( \sum_{\ell=0}^{M_i-1} \Delta Q_i(\ell) \exp \left[ -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right] \right). \tag{20}
\end{aligned}$$

The maximum-likelihood estimate of the vector of unknown parameters  $\boldsymbol{\theta}$  is the maximizer of its log-likelihood function given in Equation (20). In other words,

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \arg \max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}). \tag{21}$$

Therefore,  $\hat{\boldsymbol{\theta}}_{\text{ML}}$  can be found as the solution of the following system of equations:

$$\nabla_{\boldsymbol{\theta}} l(\boldsymbol{\theta}) \big|_{\hat{\boldsymbol{\theta}}_{\text{ML}}} = \mathbf{0}, \tag{22}$$

which can be simplified by the substitution of  $l(\boldsymbol{\theta})$  from Equation (20) as

$$\sum_{i=1}^K \frac{\sum_{\ell=0}^{M_i-1} A_{i,j}(\ell) \exp \left[ -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right]}{\sum_{\ell=0}^{M_i-1} \Delta Q_i(\ell) \exp \left[ -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right]} = 0, \quad j = 1, 2, \dots, p, \tag{23}$$

where  $A_{i,j}(\ell)$  is the partial derivative of  $\Delta Q_i(\ell)$  with respect to  $\theta_j$  defined as

$$\begin{aligned} A_{i,j}(\ell) &= \frac{\partial}{\partial \theta_j} \{\Delta Q_i(\ell)\} \\ &= \frac{1}{\sqrt{2\pi\sigma_{O_i}^2}} \left( \frac{\partial g_i}{\partial \theta_j} \right) \exp\left(-\frac{g_i^2}{2\sigma_{O_i}^2}\right) B_i(\ell), \end{aligned} \quad (24)$$

where  $B_i(\ell)$  is defined as

$$B_i(\ell) = \exp\left(\frac{2g_i\beta_i(\ell) - \beta_i^2(\ell)}{2\sigma_{O_i}^2}\right) - \exp\left(\frac{2g_i\beta_i(\ell+1) - \beta_i^2(\ell+1)}{2\sigma_{O_i}^2}\right). \quad (25)$$

This system of equations given in Equation (23) is highly non-linear in unknown parameters for most practical applications and does not have a closed-form solution. In practice, an efficient numerical approach has to be devised to find the ML estimate formulated in this section. In the next section, the expectation-maximization algorithm<sup>13</sup> will be developed as an efficient iterative approach to numerically find the ML estimate of a vector of unknown parameters for the case of digital local processing scheme.

## 5. LINEARIZED EM SOLUTION: DIGITAL LOCAL PROCESSING CASE

Based on the expectation-maximization algorithm introduced by Dempster et al.<sup>13</sup>, we consider linearly-processed received vector of local observations from distributed sensors at the fusion center  $\mathbf{z}'$  defined in Equation (17) as *incomplete data set*, and the pair of the vector of local observations  $\mathbf{r} = [r_1, r_2, \dots, r_K]^T$  defined in Equation (1) and the vector of linearly-transformed *channel* additive white Gaussian noise variables  $\mathbf{v} = [v_1, v_2, \dots, v_K]^T$ , i.e.  $\{\mathbf{r}, \mathbf{v}\}$ , as *complete data set*. The mapping  $\mathbf{z}' = \mathbf{u} + \mathbf{v}$  relates the incomplete and complete data spaces, where  $\mathbf{u} = [u_1, u_2, \dots, u_K]^T$  is the vector of locally-quantized observations of distributed sensors based on  $u_i = \gamma_i(r_i)$  as the known local quantization rule of sensor  $i$  defined in Equation (13).

The joint probability density function of the complete data set parametrized by the vector of unknown parameters  $\boldsymbol{\theta}$  is found as

$$\begin{aligned} f_{\text{CD}}(\mathbf{r}, \mathbf{v} : \boldsymbol{\theta}) &\stackrel{(a)}{=} f_{\mathbf{r}}(\mathbf{r} : \boldsymbol{\theta}) f_{\mathbf{v}}(\mathbf{v}) \\ &= \left[ \frac{1}{\sqrt{(2\pi)^K |\boldsymbol{\Sigma}_O|}} \exp\left(-\frac{1}{2}(\mathbf{r} - \mathbf{g})^T \boldsymbol{\Sigma}_O^{-1}(\mathbf{r} - \mathbf{g})\right) \right] \left[ \frac{1}{\sqrt{(2\pi)^K |\boldsymbol{\Sigma}_C|}} \exp\left(-\frac{1}{2}\mathbf{v}^T \boldsymbol{\Sigma}_C^{-1}\mathbf{v}\right) \right] \end{aligned} \quad (26)$$

where (a) is based on the fact that  $\mathbf{r}$  and  $\mathbf{v}$  are independent Gaussian random vectors,  $\boldsymbol{\Sigma}_O \stackrel{\text{def}}{=} \text{diag}(\sigma_{O_1}^2, \sigma_{O_2}^2, \dots, \sigma_{O_K}^2)$  and  $\boldsymbol{\Sigma}_C \stackrel{\text{def}}{=} \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2)$  are the diagonal matrices of observation noise variances and transformed channel noise variances  $\sigma_i^2 \stackrel{\text{def}}{=} \frac{\sigma_{C_i}^2}{|h_i|^2}$ , respectively. The joint log-likelihood function of the complete data set is defined as

$$\begin{aligned} l_{\text{CD}}(\mathbf{r}, \mathbf{v} : \boldsymbol{\theta}) &\stackrel{\text{def}}{=} \ln f_{\text{CD}}(\mathbf{r}, \mathbf{v} : \boldsymbol{\theta}) \\ &\equiv -(\mathbf{r} - \mathbf{g})^T \boldsymbol{\Sigma}_O^{-1}(\mathbf{r} - \mathbf{g}) \\ &= -\sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} (r_i - g_i)^2, \end{aligned} \quad (27)$$

where the terms independent of the vector of unknown parameters  $\boldsymbol{\theta}$  are omitted.

Let  $\hat{\boldsymbol{\theta}}^{(k)}$  be the estimate of the unknown vector of parameters at the  $k$ th iteration of the expectation-maximization algorithm. To further refine and update the estimates of unknown parameters, we alternate the *expectation* and *maximization* steps defined as follows:



**Expectation Step (E-Step):** During the expectation step, the conditional expectation of the joint log-likelihood function of the complete data set, given the incomplete data set and  $\hat{\boldsymbol{\theta}}^{(k)}$ , is found as

$$\begin{aligned} F\left(\boldsymbol{\theta} \mid \hat{\boldsymbol{\theta}}^{(k)}\right) &\stackrel{\text{def}}{=} \mathbb{E}\left[l_{\text{CD}}(\mathbf{r}, \mathbf{v} : \boldsymbol{\theta}) \mid \mathbf{z}', \hat{\boldsymbol{\theta}}^{(k)}\right] \\ &= \mathbb{E}\left[-\sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} (r_i - g_i)^2 \mid \mathbf{z}', \hat{\boldsymbol{\theta}}^{(k)}\right], \end{aligned} \quad (28)$$

where  $\mathbb{E}[\cdot]$  denotes the expectation operation with respect to the conditional pdf of the complete data set, given the incomplete data set and the estimate of the vector of parameters at the  $k$ th iteration as

$$f_{\text{CD}|\text{ID}}\left(\mathbf{r}, \mathbf{v} : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{z}'\right) = f_{\mathbf{r}|\mathbf{z}'}\left(\mathbf{r} : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{z}'\right) f_{\mathbf{v}|\mathbf{z}'}\left(\mathbf{v} : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{z}'\right), \quad (29)$$

where  $\mathbf{r}$  and  $\mathbf{v}$  are two independent Gaussian random vectors,  $f_{\mathbf{v}|\mathbf{z}'}\left(\mathbf{v} : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{z}'\right)$  is independent of the argument inside the expectation operation, and based on the Bayes' rule,

$$f_{\mathbf{r}|\mathbf{z}'}\left(\mathbf{r} : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{z}'\right) = \frac{f_{\mathbf{z}'|\mathbf{r}}\left(\mathbf{z}' : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{r}\right) f_{\mathbf{r}}\left(\mathbf{r} : \hat{\boldsymbol{\theta}}^{(k)}\right)}{f_{\mathbf{z}'}\left(\mathbf{z}' : \hat{\boldsymbol{\theta}}^{(k)}\right)}. \quad (30)$$

Based on the theorem of total probability,

$$f_{\mathbf{z}'}\left(\mathbf{z}' : \hat{\boldsymbol{\theta}}^{(k)}\right) = \underbrace{\int \cdots \int}_{K} f_{\mathbf{z}'|\mathbf{r}}\left(\mathbf{z}' : \hat{\boldsymbol{\theta}}^{(k)} \mid \mathbf{r}\right) f_{\mathbf{r}}\left(\mathbf{r} : \hat{\boldsymbol{\theta}}^{(k)}\right) d\mathbf{r}. \quad (31)$$

Note that  $f_{\mathbf{z}'}\left(\mathbf{z}' : \hat{\boldsymbol{\theta}}^{(k)}\right)$  is independent of the argument inside the expectation operation in Equation (28).

**Maximization Step (M-Step):** During the maximization step, the next estimate of the vector of unknown parameters is found as the maximizer of the result of the expectation step. In other words,

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \arg \max_{\boldsymbol{\theta}} F\left(\boldsymbol{\theta} \mid \hat{\boldsymbol{\theta}}^{(k)}\right), \quad (32)$$

which can be rewritten as

$$\left. \frac{\partial}{\partial \theta_j} \left\{ F\left(\boldsymbol{\theta} \mid \hat{\boldsymbol{\theta}}^{(k)}\right) \right\} \right|_{\hat{\boldsymbol{\theta}}^{(k+1)}} = 0, \quad j = 1, 2, \dots, p. \quad (33)$$

This system of equations can be specified more precisely by substitution of  $F\left(\boldsymbol{\theta} \mid \hat{\boldsymbol{\theta}}^{(k)}\right)$  from Equation (28) as

$$\left. \mathbb{E} \left[ \sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \mid \mathbf{z}', \hat{\boldsymbol{\theta}}^{(k)} \right] \right|_{\hat{\boldsymbol{\theta}}^{(k+1)}} = 0, \quad j = 1, 2, \dots, p. \quad (34)$$

The above conditional expectation can be found using the probability density functions defined in Equa-

tions (29)–(31) as

$$\begin{aligned}
\mathbb{E} \left[ \sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \middle| \mathbf{z}', \widehat{\boldsymbol{\theta}}^{(k)} \right] &= \sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} \mathbb{E} \left[ (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \middle| z'_i, \widehat{\boldsymbol{\theta}}^{(k)} \right] \\
&= \sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} \iint \left[ (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \right] f_{\text{CD|ID}} \left( r_i, v_i : \widehat{\boldsymbol{\theta}}^{(k)} \middle| z'_i \right) dr_i dv_i \\
&\stackrel{(a)}{=} \sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} \int \left[ (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \right] f_{R_i|Z'_i} \left( r_i : \widehat{\boldsymbol{\theta}}^{(k)} \middle| z'_i \right) dr_i \\
&= \sum_{i=1}^K \frac{1}{\sigma_{O_i}^2} \left[ \frac{1}{f_{Z'_i} \left( z'_i : \widehat{\boldsymbol{\theta}}^{(k)} \right)} \right] T_{i,j}^{(k)}(z'_i), \tag{35}
\end{aligned}$$

where (a) is based on Equation (29) and the independence of  $v_i$  from the argument of the expectation operation and  $T_{i,j}^{(k)}(z'_i)$  is defined as

$$T_{i,j}^{(k)}(z'_i) = \int_{-\infty}^{\infty} \left[ (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \right] f_{Z'_i|R_i} \left( z'_i : \widehat{\boldsymbol{\theta}}^{(k)} \middle| r_i \right) f_{R_i} \left( r_i : \widehat{\boldsymbol{\theta}}^{(k)} \right) dr_i. \tag{36}$$

Note that based on Equation (17), the conditional pdf of  $Z'_i$ , given  $R_i$ , is Gaussian with mean  $u_i$  and variance  $\sigma_i^2 \stackrel{\text{def}}{=} \frac{\sigma_{C_i}^2}{|h_i|^2}$ , i.e.  $Z'_i|R_i \sim \mathcal{N}(u_i, \sigma_i^2)$ . Moreover, based on Equation (1), the pdf of  $R_i$  is Gaussian with mean  $g_i^{(k)}$  and variance  $\sigma_{O_i}^2$ , i.e.  $R_i \sim \mathcal{N}(g_i^{(k)}, \sigma_{O_i}^2)$ , where  $g_i^{(k)} = g(x_i, y_i : \widehat{\boldsymbol{\theta}}^{(k)})$  is the estimate of the underlying function  $g(x, y)$  at location  $(x_i, y_i)$  with the vector of unknown parameters  $\boldsymbol{\theta}$  replaced by its estimate at the  $k$ th iteration of the EM algorithm. Using these two probability density functions,  $T_{i,j}^{(k)}(z'_i)$  can be found as

$$\begin{aligned}
T_{i,j}^{(k)}(z'_i) &= \int_{-\infty}^{\infty} \left[ (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \right] \left[ \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{(z'_i - u_i)^2}{2\sigma_i^2} \right) \right] \left[ \frac{1}{\sqrt{2\pi\sigma_{O_i}^2}} \exp \left\{ -\frac{(r_i - g_i^{(k)})^2}{2\sigma_{O_i}^2} \right\} \right] dr_i \\
&= \frac{1}{2\pi\sqrt{\sigma_i^2\sigma_{O_i}^2}} \sum_{\ell=0}^{M_i-1} \int_{\beta_i(\ell)}^{\beta_i(\ell+1)} \left[ (r_i - g_i) \frac{\partial g_i}{\partial \theta_j} \right] \exp \left( -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right) \exp \left\{ -\frac{(r_i - g_i^{(k)})^2}{2\sigma_{O_i}^2} \right\} dr_i \\
&= \frac{1}{\sqrt{2\pi\sigma_i^2}} \sum_{\ell=0}^{M_i-1} \exp \left( -\frac{(z'_i - \ell)^2}{2\sigma_i^2} \right) \left( \frac{\partial g_i}{\partial \theta_j} \right) \Gamma_i^{(k)}(\ell), \tag{37}
\end{aligned}$$

where  $\Gamma_i^{(k)}(\ell)$  is defined as

$$\begin{aligned}
\Gamma_i^{(k)}(\ell) &= \frac{1}{\sqrt{2\pi\sigma_{O_i}^2}} \int_{\beta_i(\ell)}^{\beta_i(\ell+1)} (r_i - g_i) \exp \left[ -\frac{(r_i - g_i^{(k)})^2}{2\sigma_{O_i}^2} \right] dr_i \\
&= \sqrt{\frac{\sigma_{O_i}^2}{2\pi}} \exp \left[ -\frac{(g_i^{(k)})^2}{2\sigma_{O_i}^2} \right] \Lambda_i^{(k)}(\ell) + (g_i^{(k)} - g_i) \Delta Q_i^{(k)}(\ell), \tag{38}
\end{aligned}$$

where  $\Delta Q_i^{(k)}(\ell) = \Delta Q_i(\ell) \Big|_{g_i=g_i^{(k)}}$  and  $\Lambda_i^{(k)}(\ell)$  is defined as

$$\Lambda_i^{(k)}(\ell) = \exp \left( \frac{2g_i^{(k)}\beta_i(\ell) - \beta_i^2(\ell)}{2\sigma_{O_i}^2} \right) - \exp \left( \frac{2g_i^{(k)}\beta_i(\ell+1) - \beta_i^2(\ell+1)}{2\sigma_{O_i}^2} \right). \tag{39}$$

Combing Equations (35) and (37)–(39) and replacing them in Equation (34) will result in a non-linear system of equations in terms of  $\widehat{\boldsymbol{\theta}}^{(k+1)}$ . In other words, Equation (34) can be rewritten as follows to give a new update of the vector of unknown parameters at the  $(k+1)$ th step of the EM algorithm using its values derived in the  $k$ th step:

$$\sum_{i=1}^K \frac{1}{\sigma_{O_i}^2 \sqrt{2\pi\sigma_i^2}} \left[ \frac{1}{f_{Z_i'}(z_i' : \widehat{\boldsymbol{\theta}}^{(k)})} \right] \left( \frac{\partial g_i^{(k+1)}}{\partial \theta_j} M_i^{(k)}(z_i') - g_i^{(k+1)} \frac{\partial g_i^{(k+1)}}{\partial \theta_j} N_i^{(k)}(z_i') \right) = 0, \quad j = 1, 2, \dots, p, \quad (40)$$

where  $M_i^{(k)}(z_i')$  and  $N_i^{(k)}(z_i')$  are defined as

$$M_i^{(k)}(z_i') = \sum_{\ell=0}^{M_i-1} \exp\left(-\frac{(z_i' - \ell)^2}{2\sigma_i^2}\right) \left( \sqrt{\frac{\sigma_{O_i}^2}{2\pi}} \exp\left[-\frac{(g_i^{(k)})^2}{2\sigma_{O_i}^2}\right] \Lambda_i^{(k)}(\ell) + g_i^{(k)} \Delta Q_i^{(k)}(\ell) \right) \quad (41a)$$

$$N_i^{(k)}(z_i') = \sum_{\ell=0}^{M_i-1} \Delta Q_i^{(k)}(\ell) \exp\left(-\frac{(z_i' - \ell)^2}{2\sigma_i^2}\right). \quad (41b)$$

It can easily be seen that the system of equations given in Equation (40) is highly non-linear with respect to the components of the vector of unknown parameters to be estimated  $\widehat{\boldsymbol{\theta}}^{(k+1)}$ . Furthermore, the solution set for this system of equations is non-convex. Therefore, a numerical solution should be found for this system of equations at each iteration of the EM algorithm. In Section 6, this approach will be utilized to numerically find the ML estimate of a vector of unknown parameters associated with a specific two-dimensional Gaussian-shaped function of interest  $g(x, y)$ . Newton's method is applied to linearize the system of equations given in (40) as briefly described in the following.

Let  $\mathbf{f}(\boldsymbol{\theta}^{(k+1)}) = \mathbf{0}$  be the vector form of the system of equations given in (40), where each component of the vector  $\mathbf{f}(\cdot)$  is a non-linear function of the components of the vector of unknown parameters  $\boldsymbol{\theta}^{(k+1)}$ . At each step  $n$  of the Newton's linearization method, the following system of *linear* equations is solved to find a new update for  $\boldsymbol{\theta}_{n+1}^{(k+1)}$  based on the previous estimate of  $\boldsymbol{\theta}_n^{(k+1)}$ :

$$\mathbf{J}(\boldsymbol{\theta}_n^{(k+1)}) \left[ \boldsymbol{\theta}_{n+1}^{(k+1)} - \boldsymbol{\theta}_n^{(k+1)} \right] = -\mathbf{f}(\boldsymbol{\theta}_n^{(k+1)}), \quad n = 0, 1, 2, \dots, \quad (42)$$

where  $\mathbf{J}(\boldsymbol{\theta}_n^{(k+1)})$  is the  $p \times p$  Jacobian matrix of vector function  $\mathbf{f}(\cdot)$  evaluated at  $\boldsymbol{\theta}_n^{(k+1)}$ . The element at the  $r$ th row and  $c$ th column of the Jacobian matrix  $\mathbf{J}(\cdot)$  evaluated at  $\boldsymbol{\theta}_n^{(k+1)}$  is defined as

$$J_{r,c}(\boldsymbol{\theta}_n^{(k+1)}) = \left. \frac{\partial f_r(\boldsymbol{\theta})}{\partial \theta_c} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_n^{(k+1)}}, \quad r, c = 1, 2, \dots, p, \quad (43)$$

where  $f_r(\cdot)$  is the  $r$ th component of the vector function  $\mathbf{f}(\cdot)$ .

## 6. CASE STUDY AND NUMERICAL ANALYSIS

In this section, we will numerically evaluate the performance of the ML estimation framework developed in previous sections for both analog and digital local processing schemes for a specific function of interest  $g(x, y)$ .

## 6.1 Simulation Setup, Parameter Specification, and Performance Measure Definition

Suppose that a two-dimensional Gaussian-shaped function defined as

$$g(x, y) = h \exp \left[ -\frac{1}{2} \left( \frac{(x - x_c)^2}{\sigma_x^2} + \frac{(y - y_c)^2}{\sigma_y^2} \right) \right] \quad (44)$$

is being sampled by the WSN under study, where  $h$  is the maximum intensity or height of the function,  $(x_c, y_c)$  is the location of its center, and  $\sigma_x^2$  and  $\sigma_y^2$  are *known* spread of the function in the  $x$  and  $y$  directions, respectively. In our simulations, we fix these values to be  $\sigma_x^2 = 4$  and  $\sigma_y^2 = 1$ . Let  $\boldsymbol{\theta} \stackrel{\text{def}}{=} [h, x_c, y_c]^T$  be the vector of unknown deterministic parameters associated with function  $g(x, y)$ . The goal is to estimate these  $p = 3$  deterministic parameters using distributed (possibly sparse) noisy samples of the function acquired by a WSN and transmitted to a fusion center through impaired parallel flat-fading channels.

Suppose that  $K$  sensors are *randomly* distributed in the observation environment, which is assumed to at least cover the area  $A \stackrel{\text{def}}{=} [x_c - 3\sigma_x, x_c + 3\sigma_x] \times [y_c - 3\sigma_y, y_c + 3\sigma_y]$ , where  $X \times Y$  denotes the Cartesian product between sets  $X$  and  $Y$ . This choice of the observation environment guarantees that all of the sensors are within the domain of the underlying function  $g(x, y)$ , and that almost all of the domain of this function is covered by distributed sensors. In our simulations, we assume that the observations made by distributed sensors are *homogeneous*. In other words, it is assumed that the variances of additive white Gaussian observation noises for all sensors are the same, i.e.

$$\sigma_O^2 \stackrel{\text{def}}{=} \sigma_{O_1}^2 = \sigma_{O_2}^2 = \dots = \sigma_{O_K}^2. \quad (45)$$

In our simulations, the observation signal-to-noise ratio (SNR) is defined as

$$\text{SNR}_O \stackrel{\text{def}}{=} \frac{1}{2\sigma_O^2}. \quad (46)$$

Note that this definition of  $\text{SNR}_O$  is the SNR at a reference distance from the center of the function  $(x_c, y_c)$  at which the strength of the function samples is unit.

As it was mentioned in Section 2, there are two main classes of local processing schemes, namely analog and digital. Both of these schemes are considered in our simulations. For the analog local processing scheme, it is assumed that the observation amplification gain is absorbed in the channel fading gain and normalized to one as described in the following paragraph. For the digital local processing scheme, it is assumed that the quantization rule at all of the sensors is the same. Therefore, all sensors quantize their local analog observations to  $b \stackrel{\text{def}}{=}} b_1 = b_2 = \dots = b_K$  bits and use the same number of quantization levels  $M \stackrel{\text{def}}{=} M_1 = M_2 = \dots = M_K$ . Each sensor uses a deterministic uniform scalar quantizer, whose set of quantization thresholds is known and is the same for all sensors. In other words,  $\beta(\ell) \stackrel{\text{def}}{=} \beta_i(\ell)$ ,  $\forall i$ , and  $\ell = 0, 1, \dots, M$ . In our simulations, the set of quantization thresholds for all sensors is chosen as  $\beta(\ell) = \frac{\ell h}{M}$ ,  $\ell = 1, 2, \dots, M - 1$ .

The parallel independent channels between distributed sensors and the fusion center are assumed to be flat Rayleigh fading channels. The observation and channel SNRs and the channel fading gains are normalized to ensure that  $\mathbb{E} [ |h_i|^2 ] = 1$  (or more generally,  $\mathbb{E} [ |\alpha_i h_i|^2 ] = 1$ , when applicable),  $i = 1, 2, \dots, K$ . When the sensors are located close to each other and the fusion center is far away from them, the distance between the sensors and the fusion center is approximately the same for all sensors, and this assumption is valid. In our simulations, it is assumed that the parallel channels between distributed sensors and the fusion center have homogeneous noises. In other words, it is assumed that the variances of additive white Gaussian channel noises in the channels between all sensors and the fusion center are the same, i.e.

$$\sigma_C^2 \stackrel{\text{def}}{=} \sigma_{C_1}^2 = \sigma_{C_2}^2 = \dots = \sigma_{C_K}^2. \quad (47)$$

In our simulations, the channel signal-to-noise ratio (SNR) is defined as

$$\text{SNR}_C \stackrel{\text{def}}{=} \frac{1}{2\sigma_C^2}. \quad (48)$$

Again, this definition of  $\text{SNR}_C$  is the SNR at a reference distance from each sensor at which the strength of the faded transmitted signal is unit.

For the analog local processing scheme, the ML estimate of the vector of unknown parameters  $\boldsymbol{\theta}$  is found based on the system of non-linear equations given in Equation (12). This system of equations is linearized and solved iteratively using Newton's method as briefly described at the end of Section 5. For the digital local processing scheme, the linearized EM algorithm given in Equations (40)–(41) is used as an iterative, efficient approach to numerically find the ML estimate of  $\boldsymbol{\theta}$ . To solve this system of non-linear equations, Newton's linearization method is applied, as briefly described at the end of Section 5, based on Equation (42).

We have chosen the estimation *integrated mean-squared error* (IMSE) defined as

$$\text{IMSE} \stackrel{\text{def}}{=} \iint_A \text{MSE}(x, y) dx dy \quad (49)$$

as the performance measure to evaluate the ML estimation techniques developed in this paper, where  $A$  is the two-dimensional observation environment and  $\text{MSE}(x, y)$  is the location-dependent estimation mean-squared error at location  $(x, y) \in A$  defined as

$$\text{MSE}(x, y) \stackrel{\text{def}}{=} \mathbb{E}_{\mathbf{h}, \mathbf{w}, \mathbf{n}} \left[ (g(x, y) - \hat{g}(x, y))^2 \right], \quad (50)$$

where  $\mathbb{E}[\cdot]$  denotes the expectation with respect to observation noise, channel noise, and channel fading gain realizations,  $g(x, y)$  is the sample of the underlying function at location  $(x, y)$ , and  $\hat{g}(x, y)$  is its estimated value based on the estimate of the vector of unknown parameters. In our simulations, the result of  $\text{MSE}(x, y)$  is also averaged with respect to the random location of distributed sensors in the observation environment.

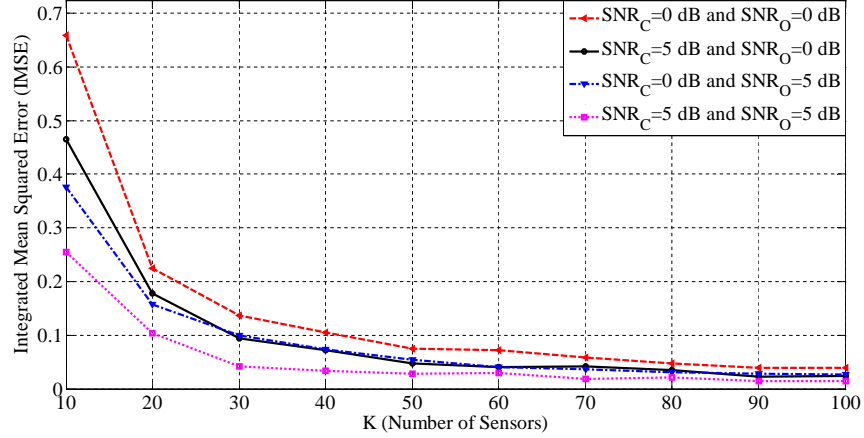
In our simulations, a Monte Carlo method has been used to calculate the IMSE averaged based on Equation (49). In this Monte Carlo simulation, 100 random placements of  $K$  local sensors in the observation environment are generated. For each random sensor placement,  $10^4$  realizations of observation noise, channel noise, and channel fading gain have been generated based on the statistical models described in Section 2. The IMSE values shown in the figures depicted in the rest of this section are averaged over all of these realizations. Therefore, our simulation results average the effects of random sensor placement in the observation environment, observation noise, channel noise, and channel fading gains on the performance of the proposed distributed parameter estimation system.

## 6.2 Effects of $K$ , $\text{SNR}_O$ , and $\text{SNR}_C$ on Distributed Parameter Estimation Performance

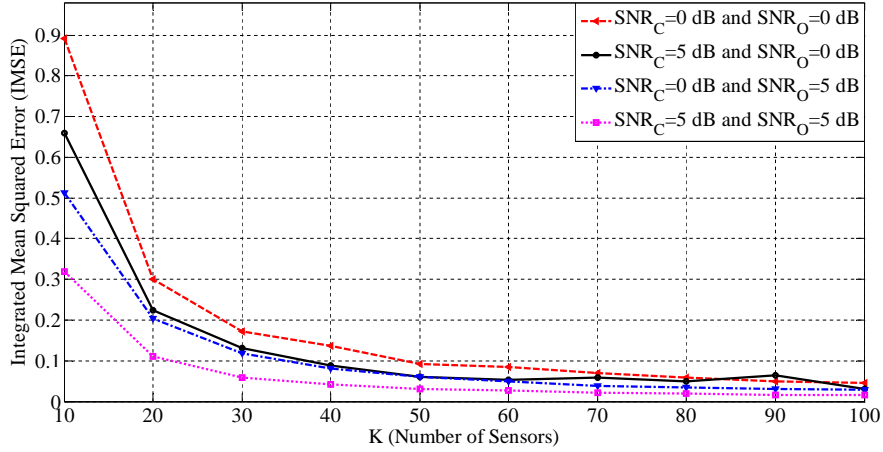
The performance of the proposed distributed parameter estimation framework is a function of, in part, the number of distributed sensors in the observation environment, observation SNR, and channel SNR.

In Fig. 2, the integrated mean-squared error at the fusion center versus the number of distributed sensors in the observation environment ( $K$ ) is shown for different values of observation SNR ( $\text{SNR}_O$ ) and channel SNR ( $\text{SNR}_C$ ). Figure 2a shows this performance measure for applying ML estimation technique at the fusion center, when the analog local processing scheme is used. Figure 2b shows the IMSE as the estimation performance measure for applying the linearized EM estimation algorithm at the fusion center, when the digital local processing scheme is used with  $M = 8$  quantization levels. As it can be seen in this figure, as the number of local sensors in the observation environment increases, the IMSE decreases monotonically. This conclusion is valid for both analog and digital local processing schemes. Furthermore, it can be observed from Fig. 2 that the performance improvement due to the increase in the density of the local sensors in the observation environment is more considerable when there are small number of sensors. As the number of sensors increases, the percentage of this performance improvement decreases, and the distributed parameter estimation system achieves an acceptable performance in terms of the IMSE at a moderate number of local sensors in the observation environment.

Figure 3 shows the integrated mean-squared error at the fusion center versus the observation SNR ( $\text{SNR}_O$ ) for different values of the number of distributed sensors in the observation environment ( $K$ ) and channel SNR ( $\text{SNR}_C$ ). Figure 3a shows the IMSE as the performance measure for applying ML estimation technique at the fusion center, when the analog local processing scheme is used. Figure 3b shows the IMSE as the estimation



(a) ML estimation at the fusion center based on *analog* local processing.

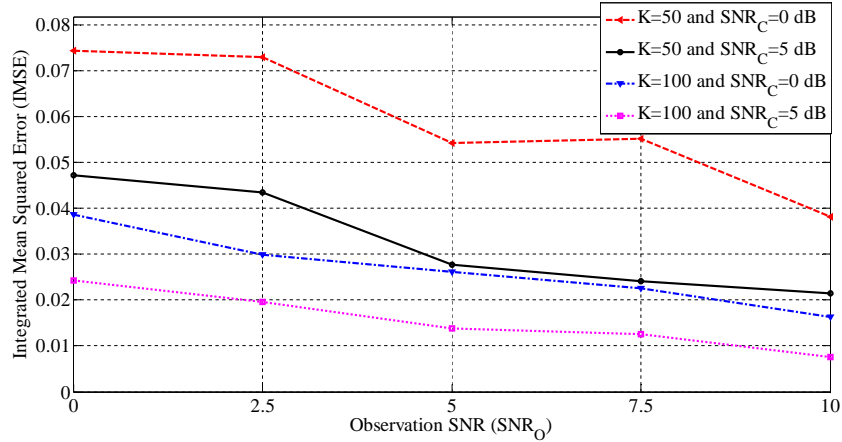


(b) Linearized EM estimation at the fusion center based on *digital* local processing with  $M = 8$  quantization levels.

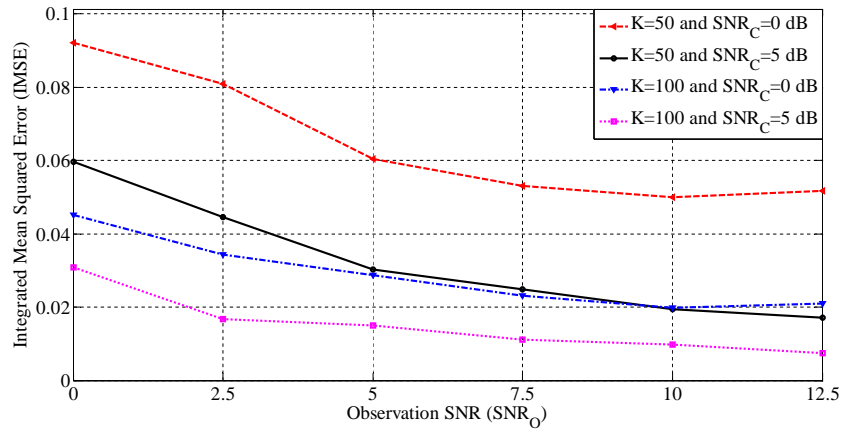
Figure 2: Integrated mean-squared error (IMSE) versus the number of distributed sensors in the observation environment ( $K$ ) for different values of observation SNR ( $\text{SNR}_O$ ) and channel SNR ( $\text{SNR}_C$ ).

performance measure for applying the linearized EM estimation algorithm at the fusion center, when the digital local processing scheme is used with  $M = 8$  quantization levels. Similar to detailed discussions provided in analyzing Fig. 2, as the observation SNR increases, the IMSE decreases monotonically for both cases of analog and digital local processing schemes. Furthermore, it can be observed from Fig. 3 that the performance improvement due to the increase in the observation SNR is more considerable at low SNRs. As the observation SNR increases, the percentage of this performance improvement decreases, and the distributed parameter estimation system achieves an acceptable performance in terms of the IMSE at a moderate observation SNR. Similar results and discussions could have been provided for analyzing the effects of channel SNR on the performance of the distributed parameter estimation framework, which is omitted because of the limited space.

One of the most important points to be noticed in Figs. 2 and 3 is that for the same values of  $K$ ,  $\text{SNR}_O$ , and  $\text{SNR}_C$ , the ML estimation at the fusion center based on the analog local processing scheme outperforms the linearized EM estimation at the fusion center based on the digital local processing schemes with  $M = 8$  quantization levels. This observation is expected for two main reasons. First, the linearized EM algorithm is an efficient iterative algorithm for *numeric* calculation of the ML estimate and therefore, it shows a degraded performance compared to the exact ML estimate. Second, by performing digital signal processing and quantizing



(a) ML estimation at the fusion center based on *analog* local processing.



(b) Linearized EM estimation at the fusion center based on *digital* local processing with  $M = 8$  quantization levels.

Figure 3: Integrated mean-squared error (IMSE) versus observation SNR ( $\text{SNR}_O$ ) for different values of the number of distributed sensors in the observation environment ( $K$ ) and channel SNR ( $\text{SNR}_C$ ).

their noisy observation, local sensors are introducing some form of quantization noise in the processed samples to become available at the fusion center. Therefore, the results of estimation at the fusion center based on the received analog samples is more reliable than those based on the quantized versions of local samples.

### 6.3 Effects of $M$ on Distributed Linearized EM Parameter Estimation Performance

Besides  $K$ ,  $\text{SNR}_O$ , and  $\text{SNR}_C$ , one of the major parameters that affects the performance of the linearized EM estimation at the fusion center, when the digital local processing scheme is used, is the number of quantization levels at local sensors, i.e.  $M$ . Figure 4 shows the integrated mean-squared error at the fusion center versus the number of quantization levels at local sensors ( $M$ ) for different values of the number of distributed sensors in the observation environment ( $K$ ), observation SNR ( $\text{SNR}_O$ ), and channel SNR ( $\text{SNR}_C$ ), when the linearized EM estimation algorithm is applied at the fusion center and the digital local processing scheme is used at local sensors. As it can be seen in this figure, as the number of quantization levels at local sensors increases, the IMSE decreases monotonically. Again, it can be observed from Fig. 4 that the performance improvement due to the increase in the number of quantization levels at local sensors is more considerable for small values of  $M$ . As the number of quantization levels at local sensors increases, the percentage of this performance improvement decreases, and the distributed parameter estimation system achieves an acceptable performance in terms of

the IMSE at a reasonably low number of quantization levels. It is worth mentioning that Fig. 4 shows that even  $M = 4$  quantization levels at local sensors can achieve a very good performance in terms of the IMSE. In other words, even if the local sensors quantize their noisy observations just to two bits, the system shows an acceptable performance. This conclusion emphasizes on the *energy efficiency* of the proposed parameter estimation framework that could lead to a higher lift-time of distributed sensors in the network. In other words, local sensors do not need to waste a lot of energy to send very high-resolution observations quantized to a large number of quantization levels to achieve an acceptable estimation error performance at the fusion center in terms of the IMSE.

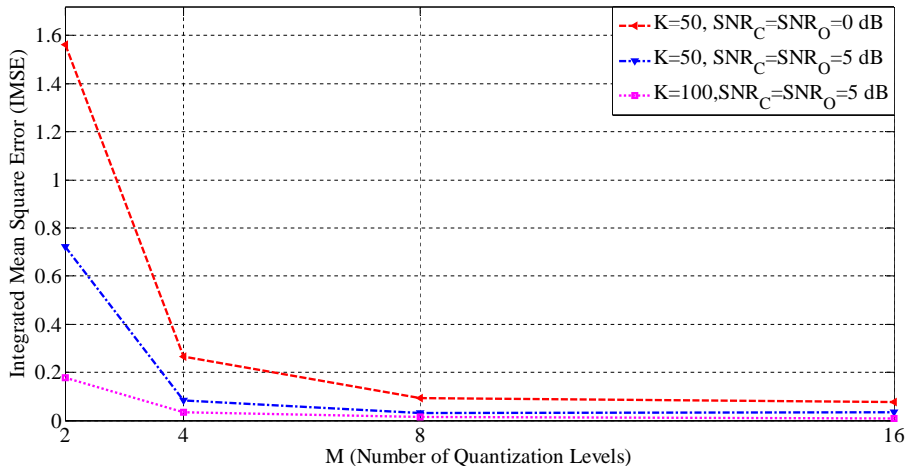


Figure 4: Integrated mean-squared error (IMSE) versus the number of quantization levels at local sensors for linearized EM estimation at the fusion center based on *digital* local processing for different values of the number of distributed sensors in the observation environment ( $K$ ), observation SNR ( $\text{SNR}_O$ ), and channel SNR ( $\text{SNR}_C$ ).

## 7. CONCLUSIONS

In this paper, the problem of distributed estimation of a vector of unknown deterministic parameters associated with a two-dimensional function was considered in the context of WSNs. Each sensor observes a sample of the underlying function at its location corrupted by an additive white Gaussian observation noise, whose samples are spatially uncorrelated across sensors. After local processing, each sensor transmits its locally-processed sample to the fusion center of WSN through parallel flat-fading channels. Two local processing schemes were considered, namely analog and digital. In the analog local processing scheme, each sensor transmits an amplified version of its local analog noisy observation to the fusion center, acting like a relay in a wireless network. In the digital local processing scheme, each sensor quantizes its noisy observation using a deterministic uniform scalar quantizer before transmitting its digitally-modulated version to the fusion center. The ML estimate of the vector of unknown parameters at the fusion center was derived for both analog and digital local processing schemes. Since the ML estimate for the case of digital local processing scheme was too complicated to efficiently be implemented, an efficient iterative EM algorithm was proposed to numerically find the ML estimate in this case. Numerical simulation results were provided to evaluate the performance of the proposed distributed parameter estimation framework in a typical WSN application scenario. As shown in the results of these simulation, the proposed distributed estimation framework can achieve a very good performance in terms of the integrated mean-squared error for reasonable values of the parameters of the system including the number of distributed local sensors in the observation environment, observation SNR, channel SNR, and the number of quantization levels for digital local processing scheme. In particular, numerical performance analysis showed that even with a low number of quantization levels at distributed sensors, i.e. high energy-efficiency, the estimation framework could provide a very good performance in terms of the integrated mean-squared error.



## REFERENCES

- [1] Xiao, J.-J., Cui, S., Luo, Z.-Q., and Goldsmith, A. J., “Linear coherent decentralized estimation,” *IEEE Transactions on Signal Processing* **56**, 757–770 (February 2008).
- [2] Gastpar, M., “Uncoded transmission is exactly optimal for a simple gaussian sensor network,” *IEEE Transactions on Information Theory* **54**, 5247–5251 (November 2008).
- [3] Ishwar, P., Puri, R., Ramchandran, K., and Pradhan, S., “On rate-constrained distributed estimation in unreliable sensor networks,” *IEEE Journal on Selected Areas in Communications* **23**, 765–775 (April 2005).
- [4] Nowak, R., Mitra, U., and Willett, R., “Estimating inhomogeneous fields using wireless sensor networks,” *IEEE Journal on Selected Areas in Communications* **22**, 999–1006 (August 2004).
- [5] Ribeiro, A. and Giannakis, G. B., “Bandwidth-constrained distributed estimation for wireless sensor networks—part I: Gaussian case,” *IEEE Transactions on Signal Processing* **54**, 1131–1143 (March 2006).
- [6] Ribeiro, A. and Giannakis, G. B., “Bandwidth-constrained distributed estimation for wireless sensor networks—part II: unknown probability density function,” *IEEE Transactions on Signal Processing* **54**, 2784–2796 (July 2006).
- [7] Niu, R. and Varshney, P. K., “Target location estimation in sensor networks with quantized data,” *IEEE Transactions on Signal Processing* **54**, 4519–4528 (December 2006).
- [8] Ozdemir, O., Niu, R., and Varshney, P. K., “Channel aware target localization with quantized data in wireless sensor networks,” *IEEE Transactions on Signal Processing* **57**, 1190–1202 (March 2009).
- [9] Maşazade, E., Niu, R., Varshney, P. K., and Keskinöz, M., “Energy aware iterative source localization for wireless sensor networks,” *IEEE Transactions on Signal Processing* **58**, 4824–4835 (September 2010).
- [10] Li, X., “RSS-based location estimation with unknown path-loss model,” *IEEE Transactions on Wireless Communications* **5**, 3626–3633 (December 2006).
- [11] Salman, N., Ghogho, M., and Kemp, A. H., “On the joint estimation of the RSS-based location and path-loss exponent,” *IEEE Wireless Communications Letters* **1**, 34–37 (February 2012).
- [12] Banavar, M. K., Tepedelenlioglu, C., and Spanias, A., “Estimation over fading channels with limited feedback using distributed sensing,” *IEEE Transactions on Signal Processing* **58**, 414–425 (January 2010).
- [13] Dempster, A. P., Laird, N. M., and Rubin, D. B., “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society, Series B (Methodological)* **39**, 1–38 (December 1977).