Distributed Estimation of a Parametric Field Using Spatially Sparse Noisy Data.

Marwan M. Alkhweldi
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
https://researchrepository.wvu.edu/etd/11134

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Distributed Estimation of a Parametric Field Using Spatially Sparse Noisy Data

by

Marwan M. Alkhweldi

Thesis submitted to the College of Engineering and Mineral Resources at West Virginia University in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

Daryl S. Reynolds, Ph.D.
Matthew C. Valenti, Ph.D.
Natalia A. Schmid, D.Sc., Chair

Lane Department of Computer Science and Electrical Engineering

Morgantown, West Virginia 2012

Keywords: Distributed estimation, expectation-maximization algorithm, maximum likelihood estimation, distributed sensor network, sparse samples

Copyright 2012 Marwan M. Alkhweldi
Abstract

Distributed Estimation of a Parametric Field Using Spatially Sparse Noisy Data

by

Marwan M. Alkhweldi
Master of Science in Electrical Engineering
West Virginia University
Natalia A. Schmid, D.Sc., Chair

The problem of distributed estimation of a parametric physical field is stated as a maximum likelihood (ML) estimation problem. Spatially sparse sensor observations are distorted by additive white Gaussian noise. These observations are then communicated over parallel additive white Gaussian channels to the fusion center (FC) for a joint estimation. This work studies cases of both analog and digital channels.

In the case of the analog channel, each sensor transmits its observation without any prior processing. A Newton’s method is used to iteratively solve for ML estimates of unknown parameters of the field.

In the case of the digital channel, each sensor quantizes its observation to $M$ levels and transmits the quantized data to the FC. An iterative expectation-maximization (EM) algorithm to estimate the unknown parameters is formulated, and its linearized version is adopted for numerical analysis.

Numerical examples are provided for both cases of the channels. The unknown field is modeled as a Gaussian bell. Dependence of the integrated mean-square error (IMSE) between the true field and the estimated field on the number of sensors in the network and the SNR in observation and transmission channels is analyzed for both kinds of channels. In the case of the digital channel, we also evaluate the dependence of the IMSE on the number of quantization levels.

In addition, we assume that the physical field is generated by an object with the location unknown to the FC. We numerically analyze the dependence of the mean-square error (MSE) between the true object location and the estimated object location. The effect of the number of sensors, SNRs, and the number of quantization levels is evaluated. Robustness of the EM algorithm with respect to convergence of the algorithm to the true parameter values is expressed in terms of Probability of Outliers.
Acknowledgements

I would like to thank my advisor, Dr. Schmid for her support and guidance throughout my graduate study.

I also would like to thank Dr. Reynolds and Dr. Valenti for being my committee members. Thanks also go to my friends. I specially thank Mohammad Fanacci for his counsel in my research.

Finally, I want to thank my parents for their unconditional love and support.
# Contents

Acknowledgements iii  
List of Figures v  
1 Introduction 1  
2 Literature Review: Distributed Estimation 4  
3 Problem Statement 8  
3.1 Analog Method ........................................... 9  
3.2 Digital Method ........................................... 11  
4 Iterative Solution 14  
4.1 Analog Method ........................................... 14  
4.1.1 Maximum Likelihood Estimation ..................... 14  
4.1.2 Linearization ......................................... 15  
4.2 Digital Method ........................................... 16  
4.2.1 Expectation Maximization Solution .................. 16  
4.2.2 Linearization ......................................... 18  
5 Numerical Analysis 19  
5.1 Numerical Analysis of Analog Method .................. 19  
5.2 Numerical Analysis of Digital Method .................. 24  
6 Conclusion and Future Work 32  
6.1 Summary ................................................. 32  
6.2 Claimed Novelties ........................................ 33  
6.3 Future Research ......................................... 33  
References 36
List of Figures

1.1 Sensor Node Components ............................................. 2

3.1 Block-diagram of the distributed sensor network (Analog Method) ....... 10
3.2 Block-diagram of the distributed sensor network (Digital Method) .......... 12

5.1 Gaussian field parameterized by $\theta = [8, 4, 4]^T$ and spread $\sigma_x^2 = \sigma_y^2 = 4$ and the squared difference between the original and reconstructed fields. ... 20
5.2 Illustration of the convergence of the Newton’s method. ................. 21
5.3 A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area $A$. 22
5.4 Dependence of the simulated ISE on the number of sensors distributed over the area $A$. .............................................. 23
5.5 Probability of outliers $P_{\text{outliers}}(\tau) = P[SE > \tau]$ (expressed in percents) as a function of $\tau$. The plot is based on 1000 Monte Carlo simulations. ............ 23
5.6 Probability of outliers (expressed in percents) as a function of the threshold for different values of SNR in observation and transmission channels. .... 24
5.7 Illustration of the EM convergence. ................................... 25
5.8 A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 8$. ................. 26
5.9 Dependence of the simulated ISE on the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 8$. ................. 27
5.10 Probability of outliers $P_{\text{outliers}}(\tau) = P[SE > \tau]$ (expressed in percents) as a function of $\tau$. The plot is based on 1000 Monte Carlo simulations. The number of quantization levels is set to $M = 8$. ................. 27
5.11 Probability of outliers (expressed in percents) as a function of the threshold for different values of SNR in observation and transmission channels. The number of quantization levels is set to $M = 8$. ................. 28
5.12 A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 16$. ................. 29
5.13 Dependence of the simulated ISE on the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 16$. ................. 29
5.14 A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 32$. ........................................ 30

5.15 Dependence of the simulated ISE on the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 32$. ........................................ 30

5.16 A box plot of the SE between the estimated and true location of the object displayed for different number of quantization levels. The number of sensors $K = 20$. SNRs for digital case and analog case are set to equivalent values. Note the convergence of the results as $M$ increases. ................................. 31
Chapter 1

Introduction

Wireless Sensor Networks (WSN) is a collaborative environment and approach to solving various inference problems [1]. Each WSN consists of a collection of nodes spatially distributed over an area in a two-dimensional space or over a volume in a three-dimensional space. Each node is equipped with a wireless transmitter/receiver, a small CPU, and a power supply [2, 3, 4](see Fig. 1.1 for a block-diagram of a node). Given a network set up and its configuration, nodes collaborate with one another to solve an application specific inference problem.

In a network, sensors can be connected with one another, they can be connected to a set of other anchor nodes or they can be linked to a single node, called supercomputer or Fusion Center resulting in a decentralized or distributed topologies, respectively [5]. The choice of a topology depends on applications and constraints related to them.

Each sensor node (sensor) is an independent small subsystem of a WSN (see Ch. 6 in [2]). We will now provide a very brief overview of components of each sensor node and emphasize constraints that each component imposes on the node and on WSN in general.

Sensors are small and inexpensive devices that are able to take physical measurements. Sensors have the ability to sense some phenomena and convert their measurements into a voltage or resistance signals.

Processing unit (CPU) is intended to perform additional signal processing on the output of sensors. Because of a limited power allocated to each sensor node in a WSN, one of the main goals of the CPU is to intelligently compress sensory measurements and prepare them
Figure 1.1: Sensor Node Components

for transmission.

Transceiver ensures reliable transmission of processed sensory measurements. It transmits information from a node to other nodes and receives signals from them. This component consumes the most of power in a WSN. Therefore, reliable and efficient transceiver is an important component of a WSN.

WSNs are applied in a variety of fields including health care [6, 7], environmental monitoring [8, 9], structure and process monitoring [10] and multiple DoD related applications (for instance, area surveillance [11]).

Much research has been reported on each of these topics in the past two decades. In the field of distributed estimation, for example, various estimation problems have been formulated and solved.

Many works choose either to optimize a distributed sensor network with respect to energy consumption during transmission [12, 13] or impose bandwidth constraints and thus focus on designing an optimal quantization strategy for the distributed network [14], [3]. There are few that involve both constraints (see [15] as an examples). Among research groups working on the problem of distributed estimation, there are a few dealing with distributed estimation of a field (a multidimensional function, in general) [16, 17, 18]. Since in many real world applications distributed estimation of a multidimensional function may provide additional information that aids in making a high fidelity decision or in solving another inference problem, we contribute to this topic by formulating and solving the problem of a parametric field estimation from sparse noisy sensor measurements.
In this thesis, the problem of distributed estimation of a physical field from sensory data collected by a homogeneous sensor network is stated as a maximum likelihood estimation problem. The physical field is a deterministic function and has a known spatial distribution parameterized by a set of unknown parameters, such as the location of an object generating the field and the strength of the field in the region occupied by the sensors. Sensor observations are distorted by additive white Gaussian noise. Prior to transmission, each sensor quantizes its observation to $M$ levels. The quantized data are then communicated over parallel additive white Gaussian channels to the fusion center where the unknown parameters of the underlying physical field are estimated. An iterative expectation-maximization (EM) algorithm to estimate the unknown parameters is formulated and a simplified numerical solution involving additional approximations is developed. The numerical examples illustrate the developed approach for the case of the field modeled as a Gaussian bell.

The rest of the thesis is organized as follows. Chapter two presents a literature review on the topic of distributed estimation. Chapter three formulates the problem of a parametric field estimation from distributed sensor measurements. Chapter four introduces iterative solutions to the problem in chapter three. Then chapter five presents numerical results. Finally, chapter six summarizes the work and suggests few ideas for the future work.
Chapter 2

Literature Review: Distributed Estimation

A bulk of literature has been published on the topic of wireless sensor networks in the past 30 years (see, for example a review in [19], [20]). Some of these publications are concerned with particular applications [21]; other papers study the limits (scalability) of wireless networks by using asymptotic analysis [22], while the rest of publications are focused on inclusion of practical constraints in the modeling and analysis of wireless sensor networks (see for example, [23]).

Wireless sensor networks are used in many areas such as Health Care, Environmental Monitoring, Military applications etc. Dietmar et al. [24] study one of these applications related to Environment Monitoring. They use an ad hoc network of mobile sensors to estimate the temperature of the environment along a road path with the goal to obtain a temperature map. Sensors are placed on a vehicle, and the vehicle moves along the road. Sensors take measurements at uniform time intervals and produce temperature estimates. The estimates are then wirelessly transmitted to a roadside node. Sensors perform the estimation task based on their own observations and measurements by their neighboring sensors. Prior to estimating the temperature value, sensors exchange their estimates with their neighbors to achieve a consensus, which is claimed as the final estimate for this part of the road.

In many applications data acquired by wireless sensor networks need a wide bandwidth
for their transmission, but digital channels are band limited and are often shared with other applications. Practical sensor nodes are therefore equipped with a low power signal processing units (CPUs). CPUs locally process the sensor observations and output a low bandwidth signal. Since the local signal processing has to be done optimally with respect to a specific application, its design and optimization is an active research scope.

Xiao and Luo [25] proposed a distributed estimation scheme (DES) similar in its mechanisms to the best linear unbiased estimator (BLUE). DES estimates unknown values based on a combination of observations collected at a Fusion Center (FC). The DES deals with low bandwidth channels efficiently, since every sensor converts its observation to a binary form. The number of bits in these binary messages depends on the quality of sensor observations (the quality is measured in terms of $SNR_O$). More bits are assigned to observations with a high signal to noise ratio. FC assigns a weight vector to every observation. Observations represented by more bits are given a larger weight so it has stronger effect on the estimation decision.

Another practical constraint that many papers take into account is a limited amount of power or energy allocated to each sensor in a sensor network. Many of current applications of WSNs are required to serve over a long period of time (up to few years). Much research has been done on this topic in the past decade. We will mention only few works related to distributed estimation.

In Cui et al. [15], sensor observations are sent over wireless channels in analog form. The analog observations are exposed to fading. Sensors in their model are placed different distances apart from the FC, which causes variation of fading coefficients in different channels. The paper intends to minimize estimation distortions under the constraint of limited power. To save the power in the sensor network, prior to transmitting its observation each sensor evaluates the quality of its observation and makes a binary decision whether to transmit signal or not. Results related to saving power are shown in this paper. It is shown in the paper that the proposed transmission scheme results in considerable power saving compared to the traditional scheme with equal power transmissions.

Wu et al. [13], study the problem of minimizing the estimation error, mean squared error (MSE), under the constraint of a limited power. The number of bits assigned to each
sensor observation depends on the quality of the channel between that sensor and the FC (the model assumes fading in channels). Sensors located further away from the FC send fewer bits compared to the sensors that are close by.

Li and AlRegib [12] study the problem of optimizing the network performance under energy constraint using an approach similar to Wu et al. [13]. The main focus of Li and AlRegib is on establishing a tradeoff between the number of active sensors and the energy allocated to all active sensors. The number of active sensors depends on the energy need of each of them and the total available energy budget.

Wu et al. [13] and Li and AlRegib [12] study the best performance in terms of MSE under energy constraint. Li and AlRegib [12] paper derives the upper bound on the life time of WSN where the end of the network’s life is defined as the moment when it fails to keep the mean square error (MSE) above a threshold. If the MSE is larger than this threshold, then the network is not functional and it reached the last cycle. This work also proposes a methodology to increase the lifetime of the network by adopting an approach similar to the approach in Wu et al. [13], where a different number of bits is assigned to sensor observations and the number of bits is proportional to the quality of the channel between the sensor and the FC.

The work by Wang et al. [17] studies an asymptotic case, where the number of sensors is allowed to grow to infinity. Conditions for the MSE to converge to zero as the number of sensors increases are stated. This paper uses one bit distributed estimation. The area of the field is assumed to be divided into multiple cells, and the sensors are deployed uniformly among these cells. Observations at each sensor is quantized by using a random binary quantizer. A unique way of sending messages to the FC is used in this paper, where they intend to process the observation value to be one bit per snapshot. Each sensor collects a set of snapshots, which are then encoded and sent to the FC. The FC estimates the field as a constant value at each cell of sensors at every snapshot.

Patwari et al. [26] propose the Distributed Expectation-Maximization (DEM) in an ad hoc WSN to estimate the state of the environment. In this work the observations are described as a mixture of Gaussian components (means and variances) with coefficients (mixing parameters). The Gaussian components are common to all sensors, and coefficients
are varying from one sensor to another. The claimed novelty is in the way of applying the EM algorithm to estimate these coefficients where the estimation process will be performed at a different sensor during each iteration. The number of iterations that the network will take to estimate the unknown parameter is equal to the number of times sensors pass messages among themselves. Sensors will continue passing messages to each other in a cyclic path until no change in the estimated value will be observed. The messages that sensors exchange carry data about the estimated values of the preceding sensor. The next sensor uses these data to compute the Gaussian components and to update mixing parameters and then passes its message to the next sensor. The Gaussian components have the same estimated values over all nodes, and the mixing parameters are estimated locally at each node. This paper also addresses the problem of reducing the communication time between nodes with the goal to conserve the energy.

In many research publications it is assumed that positions of sensors are known to FC. However, this is not the case in practice. Patwari et al. [26], address the problem of estimating a positions of sensors (also known as localization problem). This paper suggests to group sensors in two groups. The first group includes sensors with known positions. These are reference devices. The other group involves sensors with unknown positions. They are called blindfolded devices. The sensors with unknown positions estimate their positions relatively to the reference devices by adopting either Received Signal Strength (RSS) strategy or Time of Arrival (TOA) strategy. In this paper Maximum Likelihood Estimation (MLE) is adopted to find unknown positions, and a Cramer-Rao bound (CRB) is derived. Estimation algorithm uses pair wise observations. In the case of the Received Signal Strength (RSS) strategy, the observation is the power of a received signal sent by another sensor. In the case of the Time of Arrival (TOA) strategy, the time of arrival signal sent by another sensor.
Chapter 3

Problem Statement

Consider a distributed network of homogeneous sensors monitoring the environment for the presence of a substance or an object. Assume that each substance or object is characterized by a spatially distributed physical field generated by it. As an example, a ferromagnetic object can be viewed as a single or a collection of dipoles characterized by a magnetic field that they generate. This field can be sensed by a network of magnetometers placed in the vicinity of the object. Depending on the design of the magnetometers, they may take measurements of a directional complex valued magnetic field \([27, 28]\) or of the magnitude of the field only. The field generated by a dipole decays as a function of the inverse cube of the distance to the dipole. The sensor network does not know a priori the location of the dipole as well as the type and size of the object. However, the type and size of an object can be associated with the strength of the magnetic field. Examples of other physical fields include: (1) a radioactive field that can be modeled as a stationary spatially distributed Poisson field \([29]\) with a two-dimensional intensity function decaying according to the inverse-square law or (2) a distribution of pollution or chemical fumes that, if stationary, can often be modeled as a Gaussian bell.

Consider a network of \(K\) sensors distributed over an area \(A\). The network is calibrated in the sense that the relative locations of the sensors are known. Sensors act independently of one another and take noisy measurements of a physical field \(G(x, y)\). A sample of \(G(x, y)\) at a location \((x_k, y_k)\) is denoted as \(G_k = G(x_k, y_k)\). The field \(G(x, y)\) is characterized by \(L\) unknown parameters \(\theta = [\theta_1, \ldots, \theta_L]^T\). The sensor noise, denote it by \(W_k, k = 1, \ldots, K\) is
known and modeled as Gaussian distributed with mean zero and variance $\sigma^2$. The noise of sensors is independent and identically distributed (i.i.d.). Let $R_k$, $k = 1, \ldots, K$ be the noisy samples of the field at the location of distributed sensors. Then $R_k$ is modeled as

$$R_k = G_k + W_k, \ k = 1, \ldots, K,$$

(3.1)

where sensor noise $W_k$ is a white Gaussian noise with variance $\sigma^2$, and the influence field $G_k$ is a deterministic function. Then the measurement $R_k$ is independent Gaussian distributed with mean $G_k$ and variance $\sigma^2$. Thus its probability density function can be written as

$$R_k \sim \mathcal{N}(G(x_k, y_k : \theta), \sigma^2), \ k = 1, \ldots, K.$$

(3.2)

These noisy observations are transmitted over noisy parallel channels to a processing unit called Fusion Center (FC). The method to send these observations will be chosen according to the application constraints and the channel characteristics.

In this work we consider two types of channels used for data transmission to the FC. They are analog channel and digital channel.

### 3.1 Analog Method

Analog method implies that sensor observations are transmitted as they are without any prior or post processing over analog channels. The transmission media is modeled as parallel white Gaussian noise channels. Under these assumptions the signals $Z_k$ received by the FC are modeled as

$$Z_k = R_k + N_k, \ k = 1, \ldots, K,$$

(3.3)

where the noise $N_k$ in the channel $k$ is white Gaussian with variance $\eta^2$. The random variables $R_k$ and $N_k$ are independent. Since both $R_k$ and $N_k$ are Gaussian distributed and independent of one another, $Z_k$ is also Gaussian distributed:

$$Z_k \sim \mathcal{N}(G(x_k, y_k : \theta), \sigma^2 + \eta^2), \ k = 1, \ldots, K.$$

(3.4)
The signals received at the FC are independent but not identically distributed and thus their joint probability density function is a product of the individual density functions:

\[
f_{Z_1, \ldots, Z_K}(z_1, \ldots, z_K) = \prod_{k=1}^{K} f_{Z_k}(z_k), \quad k = 1, \ldots, K.
\] (3.5)

Given the noisy measurements and the relative location of the sensors in the network, the task of the FC is to estimate the vector parameter \( \theta \). A block diagram of the distributed network used for estimation of parameters of a physical field is shown in Fig. 3.1. By taking the logarithm of both sides in eq. (3.5) we obtain:

\[
l(z) = \sum_{k=1}^{K} \log(f_{Z_k}(z_k)) = -\frac{1}{2(\sigma^2 + \eta^2)} \sum_{k=1}^{K} (z_k - G(x_k, y_k : \theta))^2 - \frac{K}{2} \log \left(2\pi(\sigma^2 + \eta^2)\right)
\] (3.6)

The function in eq.(3.6) is known as log-likelihood function. This expression will be used to find the maximum likelihood (ML) estimates of the parameters \( \theta \).
3.2 Digital Method

Due to constraints that are imposed by practical technology, each sensor may be required to quantize its measurements prior to transmitting them to the FC. Assume that a deterministic quantizer with $M$ quantization levels is involved [30]. Let $\nu_1, \nu_2, \ldots, \nu_M$ be known reproduction points of the quantizer. Denote by $q(R_k) = q_k$ the quantized version of the measurement by the $k$-th sensor. These data are modulated using a digital modulation scheme and then transmitted to the FC over noisy parallel channels. The noise in channels is due to quantization error and channel impairments, denote it by $\tilde{N}_k$, $k = 1, \ldots, K$. Denote by $m(\cdot)$ a modulation function and by $d(\cdot)$ a demodulation function. Let $Z_1, \ldots, Z_K$, be noisy observations received by the FC. Then each $Z_k$ is given by $Z_k = d(m(q_k)) + \tilde{N}_k$, $k = 1, \ldots, K$. In this work we assume that $m(\cdot)$ and $d(\cdot)$ are linear and that the demodulator recovers the quantized signal by using a soft thresholding rule. These assumptions allow $Z_k$ be approximated by its asymptotic counterpart:

$$Z_k = q_k + N_k, \ k = 1, \ldots, K, \quad (3.7)$$

where $N_k$ is a white Gaussian noise with variance $\eta^2$. The probability density function of $q_k$ can be written as:

$$f_{q_k}(\nu) = \sum_{j=1}^{M} p_{k,j} \delta(\nu - \nu_j), \quad (3.8)$$

where $p_{k,j}$ is the probability for the output of the sensor $k$ to be mapped to the $j$-th reproduction point during the encoding process:

$$p_{k,j} = \int_{\tau_j}^{\tau_{j+1}} \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(t - G_k)^2}{2\sigma^2}\right) dt, \quad (3.9)$$

where $\tau_j$ and $\tau_{j+1}$, $j = 1, \ldots, M$ are the boundaries of the $j$-th quantization region.

Since random variables $q_k$ and $N_k$ are independent, the random variable $Z_k$ has probability density function that can be obtained as a convolution of the probability density
Figure 3.2: Block-diagram of the distributed sensor network (Digital Method).

Note that the difference compared to Fig. 3.1 is that sensors quantize their observations before sending them to FC.

functions of $q_k$ and $N_k$:

\[
 f_{Z_k}(z) = f_{q_k}(z) * f_{N_k}(z) \\
 = \int_{-\infty}^{\infty} f_{N_k}(n) f_{q_k}(z-n) dn \\
 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{n^2}{2\eta^2}\right) \sum_{j=1}^{M} p_{k,j} \delta(z-n-\nu_j) dn \\
 = \sum_{j=1}^{M} p_{k,j} \frac{1}{\sqrt{2\pi\eta^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{n^2}{2\eta^2}\right) \delta(z-n-\nu_j) dn \\
 = \sum_{j=1}^{M} p_{k,j} \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{(z-\nu_j)^2}{2\eta^2}\right), \tag{3.10}
\]

where * stands for convolution.
The joint probability density function of the observed data at the FC is
\[ f_{Z_1,\ldots,Z_K}(z_1,\ldots,z_K) = \prod_{k=1}^{K} f_{Z_k}(z_k), \quad k = 1,\ldots,K. \] (3.11)

The task of the FC is to estimate the vector parameter \( \theta \). A block diagram of the distributed network used for estimation of parameters of a physical field in the case of the digital channel is shown in Fig. 3.2.

We adopt a maximum likelihood (ML) estimation approach [31], [32] to solve the problem of distributed parameter estimation. The joint likelihood function of the independent quantized noisy measurements \( Z_1, Z_2,\ldots,Z_K \) can be written as
\[
l(z) = \sum_{k=1}^{K} \log \left( \sum_{j=1}^{M} p_k(\nu_j) \frac{1}{\sqrt{2\pi\eta^2}} \exp \left( -\frac{(z_k - \nu_j)^2}{2\eta^2} \right) \right)
= \sum_{k=1}^{K} \log \left( \sum_{j=1}^{M} p_k(\nu_j) \exp \left( -\frac{(z_k - \nu_j)^2}{2\eta^2} \right) \right) - \frac{K}{2} \log (2\pi\eta^2).
\] (3.12)

The ML solution \( \hat{\theta} \) is the solution that maximizes the expression (3.12). For a numerical example in Ch. 5, the field is modeled as a Gaussian bell with three unknown parameters: the strength of the field \( \mu \) and the location \((x_c, y_c)\).
Chapter 4

Iterative Solution

Since the expressions for the log-likelihood function (3.6) and (3.12) are highly nonlinear in unknown parameters, we develop an iterative solution to the problem. In Analog method, we use the Maximum Likelihood (ML) Estimation to obtain the likelihood equations and then linearize them using Newton’s method [33], [34]. In Digital method, we first formulate a set of Expectation-Maximization (EM) iterations [35], [36] and then involve a Newton’s linearization to solve for the unknown parameters. Using EM algorithm in the digital case is due to the fact that the observations at the FC are noisy quantized versions of the samples observed by sensors.

4.1 Analog Method

4.1.1 Maximum Likelihood Estimation

Maximum Likelihood (ML) estimation is a classical method used to estimate unknown deterministic parameters [31]. The ML solution for the analog case described in Sec. 3.6 is the vector $\theta$ that maximizes the log-likelihood function in eq.(3.6).
\[ \hat{\theta} = \arg \max_{\theta \in \Theta} l(Z : \theta) \]
\[ = \arg \max_{\theta \in \Theta} \left\{ -\frac{1}{2(\sigma^2 + \eta^2)} \sum_{k=1}^{K} (Z_k - G(x_k, y_k : \theta))^2 - \frac{K}{2} \log \left( 2\pi (\sigma^2 + \eta^2) \right) \right\} \]
\[ = \arg \max_{\theta \in \Theta} \left\{ -\frac{1}{2(\sigma^2 + \eta^2)} \sum_{k=1}^{K} (Z_k - G(x_k, y_k : \theta))^2 - \text{terms not function of } \theta \right\} \tag{4.1} \]

where \( \Theta \) is a set of admissible parameters.

The necessary conditions to find the maximizer are given by:
\[ \nabla_\theta l(Z : \theta) \bigg|_{\hat{\theta}_{ML}} = 0, \tag{4.2} \]
where \( \nabla_\theta \) denotes the gradient with respect to the vector \( \theta \).

### 4.1.2 Linearization

The equations (4.2) are nonlinear in \( \hat{\theta} \). To simplify the solution, we linearize the expression in (4.2) by means of Newton’s method [33]. Newton’s method is a well known method for finding the roots of nonlinear equations. It is an iterative method and requires initialization. At each iteration of a Newton’s method, the values of the previous iteration are used to find the equation of a tangent line. Then the tangent line is applied to find an updated value of roots. Repeating this process many times will ensure convergence to a local maximum [33].

Denote by \( F(\theta) \) a vector form of the left side in (4.2):
\[ F(\theta) = \begin{bmatrix} \frac{\partial l(Z : \theta)}{\partial \theta_1} \\ \frac{\partial l(Z : \theta)}{\partial \theta_2} \\ \vdots \\ \frac{\partial l(Z : \theta)}{\partial \theta_L} \end{bmatrix}. \tag{4.3} \]

Let \( J(\theta) \) be the Jacobian of the mapping:
\[ J(\theta) = \begin{bmatrix} \frac{\partial^2 l(Z : \theta)}{\partial \theta_1^2} & \frac{\partial^2 l(Z : \theta)}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 l(Z : \theta)}{\partial \theta_1 \partial \theta_L} \\ \frac{\partial^2 l(Z : \theta)}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 l(Z : \theta)}{\partial \theta_2^2} & \cdots & \frac{\partial^2 l(Z : \theta)}{\partial \theta_2 \partial \theta_L} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(Z : \theta)}{\partial \theta_L \partial \theta_1} & \frac{\partial^2 l(Z : \theta)}{\partial \theta_L \partial \theta_2} & \cdots & \frac{\partial^2 l(Z : \theta)}{\partial \theta_L^2} \end{bmatrix}. \tag{4.4} \]
Let the index \( n \) indicates the iteration of the Newton’s solution. Then \( \theta \) solves the following linearized equation:

\[
\theta_{n+1} = \theta_n - J(\theta_n)^{-1}F(\theta_n) .
\] (4.5)

### 4.2 Digital Method

#### 4.2.1 Expectation Maximization Solution

1. Expectation Maximization (EM) is known as a convex optimization approach often used to solve a Maximum A Posteriori (MAP) and ML estimation problem.

2. An EM method relies on concepts of incomplete and complete data spaces, complete and incomplete data likelihoods and on the existence of a mapping from complete to incomplete data [37], [38].

We select the independent pairs of random variables \( (R_k, N_k) \), \( k = 1, 2, \ldots, K \) as complete data. The complete data log-likelihood, denote it by \( l_{cd}(\cdot) \), is given by

\[
l_{cd}(R, N) = \log \prod_{i=1}^{K} f_{R_i}(r_i)f_{N_i}(n_i) = -\frac{1}{2\sigma^2} \sum_{i=1}^{K} (r_i - G_i)^2 + \text{terms not function of } \theta .
\] (4.6)

The measurements \( Z_i, i = 1, \ldots, K \), form incomplete data with incomplete data likelihood function in (3.12). The mapping from complete data space to incomplete data space is given by:

\[
Z_k = q(R_k) + N_k ,
\] (4.7)

where \( q(\cdot) \) is a known quantization function.

Denote by \( \hat{\theta}^{(k)} \) an estimate of the vector \( \theta \) obtained at the \( k \)-th EM iteration. To update the estimates of the parameters we alternate the expectation and maximization steps. During the expectation step, we evaluate the conditional expectation of the complete data
log-likelihood:

\[ Q^{(k+1)} = E \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^{K} (r_i - G_i)^2 \Big| z, \hat{\theta}^{(k)} \right], \tag{4.8} \]

where the expectation is with respect to the conditional probability density function of the complete data, given the incomplete data (measurements) and the estimates of the parameters at the k-th EM iteration. During the maximization step we maximize (4.8):

\[
\frac{dQ^{(k+1)}}{d\theta_t} = E \left[ \frac{1}{\sigma^2} \sum_{i=1}^{K} (r_i - G_i) \frac{dG_i}{d\theta_t} \Big| z, \hat{\theta}^{(k)} \right] \bigg|_{\hat{\theta}^{(k+1)}} = 0, \ t = 1, \ldots, L. \tag{4.9} \]

To find the conditional expectation we note that the conditional probability density function (p.d.f.) of \( Z_i \), given \( R_i \), is Gaussian with mean \( q(R_i) \) and variance \( \eta^2 \) and the p.d.f. of \( R_i \) is Gaussian with mean \( G_i \) and variance \( \sigma^2 \). We also note that at the k-th EM iteration the conditional pdf of \( R_i \), given \( Z_i \), implicitly involves the estimates of the parameters obtained at the k-th iteration:

\[
f_{R_i|Z_i}^{(k)}(r_i|z_i) = \frac{f_{R_i}^{(k)}(z_i|r_i)f^{(k)}(r_i)}{\int_{-\infty}^{\infty} f_{R_i}^{(k)}(z_i|r)f^{(k)}(r)dr}.	ag{4.10} \]

Denote by \( G_i^{(k)} \) the estimate of the field \( G(x,y) \) at the location \((x_i, y_i)\) with the vector parameters \( \theta \) replaced by their estimates at the k-th EM iteration \( \hat{\theta}^{(k)} \). Then:

\[
E \left[ (r_i - G_i) \frac{dG_i}{d\theta_t} \Big| z, \hat{\theta}^{(k)} \right] = \int_{-\infty}^{\infty} (r_i - G_i) \frac{dG_i}{d\theta_t} f_{R_i|Z_i}^{(k)}(r_i|z_i)dr_i
\]

\[
= \sum_{j=1}^{M} \int_{\tau_j}^{\tau_{j+1}} (r_i - G_i) \frac{dG_i}{d\theta_t} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(r_i - G_i^{(k)})^2}{2\sigma^2} \right)
\]

\[
\times \frac{1}{\sqrt{2\pi\eta^2}} \exp \left( -\frac{(z_i - \nu_j)^2}{2\eta^2} \right) dr_i
\]

\[
= \sum_{j=1}^{M} \frac{1}{f_{z_i}^{(k)}(z_i)} \frac{1}{\sqrt{2\pi\eta^2}} \exp \left( -\frac{(z_i - \nu_j)^2}{2\eta^2} \right)
\]

\[
\times \int_{\tau_j}^{\tau_{j+1}} (r_i - G_i) \frac{dG_i}{d\theta_t} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(r_i - G_i^{(k)})^2}{2\sigma^2} \right) dr_i. \tag{4.11} \]

Note that the difference \((r_i - G_i)\) in the last integral can be rewritten as \((r_i - G_i^{(k)}) + \)
we obtain:

\[ E \left[ (r_i - G_i) \frac{\partial G_i}{\partial \theta_i} \right] z_i, \hat{\theta}^{(k)} \] = \sum_{j=1}^{M} \frac{1}{f_{z_i}^{(k)}(z_i)} \frac{1}{\sqrt{2\pi \eta^2}} \exp \left( -\frac{(z_i - \nu_j)^2}{2\eta^2} \right) \frac{\partial G_i}{\partial \theta_i} \]

\times \left\{ \frac{1}{\sqrt{2\pi \sigma^2}} \int_{\tau_j}^{r_{j+1}} \exp \left( -\frac{(r_i - G_i^{(k)})^2}{2\sigma^2} \right) dr_i \right\}.

Replacing the last integral with a difference of two Q-functions

\[ Q \left( \frac{\tau_j - G_i^{(k)}}{\sigma} \right) \text{ and } Q \left( \frac{\tau_{j+1} - G_i^{(k)}}{\sigma} \right) \]

we obtain:

\[ \sum_{i=1}^{K} E \left[ (r_i - G_i) \frac{\partial G_i}{\partial \theta_i} \right] z_i, \hat{\theta}^{(k)} \] = \sum_{i=1}^{K} \sum_{j=1}^{M} \frac{1}{f_{z_i}^{(k)}(z_i)} \frac{1}{\sqrt{2\pi \eta^2}} \exp \left( -\frac{(z_i - \nu_j)^2}{2\eta^2} \right) \frac{\partial G_i}{\partial \theta_i} \]

\times \left\{ \frac{\sigma^2}{\sqrt{2\pi \sigma^2}} \left[ \exp \left( -\frac{(\tau_j - G_i^{(k)})^2}{2\sigma^2} \right) \right] - \exp \left( -\frac{(\tau_{j+1} - G_i^{(k)})^2}{2\sigma^2} \right) \right\} \]

\[ + (G_i^{(k)} - G_i) \left\{ Q \left( \frac{\tau_j - G_i^{(k)}}{\sigma} \right) - Q \left( \frac{\tau_{j+1} - G_i^{(k)}}{\sigma} \right) \right\} \bigg|_{G_i = G_i^{(k+1)}} = 0, \quad (4.12) \]

where \( f_{z_i}^{(k)}(z_i) = \int f^{(k)}(z_i|r) f^{(k)}(r) dr \). The expression \( Q(\cdot) \) is used to denote the Q-function.

### 4.2.2 Linearization

The equations (4.13) are nonlinear in \( \hat{\theta}^{(k+1)} \) and have to be solved numerically for each EM iteration. To simplify the solution, we linearize the expression in (4.13) by means of Newton’s method. The procedure follows the same steps described in Sec. 4.1.2. Denote by \( F(\theta^{(k+1)}) \) the vector form of the left side in (4.13), which is a mapping from \( \theta^{(k+1)} \) to the range of \( F(\theta^{(k+1)}) \). Let \( J \left( \theta_n^{(k+1)} \right) \) be the Jacobian of the mapping. The index \( n \) indicates the iteration of the Newton’s solution. Then \( \theta^{(k)} \) solves the following linearized equation:

\[ \theta_n^{(k+1)} = \theta_n^{(k+1)} - J \left( \theta_n^{(k+1)} \right)^{-1} F \left( \theta_n^{(k+1)} \right). \quad (4.14) \]

Note the difference between eq.(4.5) and eq.(4.14). Eq.(4.5) seeks for a maximizer of the likelihood function in (4.1). Eq.(4.14) seeks for a solution of (4.13), since the equation depends on the EM iteration.
Chapter 5

Numerical Analysis

In this chapter, the performance of the distributed ML estimators in (4.1) and (4.13) is demonstrated on simulated data. A distributed network of $K$ sensors is formed by positioning sensors at random over an area $A$ of size $8 \times 8$. The location of each sensor is noted. A Gaussian field shown in Fig. 5.1(a) is sampled at the location of the $i$-th sensor, $i = 1, \ldots, K$, and a sample of randomly generated Gaussian noise with mean zero and variance $\sigma^2$ is added to each sample. In our simulations, $K$ is varied from 5 to 200 and $\sigma^2$ is selected such that the total signal-to-noise ratio (SNR) of the local observations defined as

$$ SNR_O = \frac{\int \int_A G(x, y; \theta)^2 dx dy}{A \sigma^2} $$

is 15 dB. Sensor observations are transmitted over noisy parallel channels to the Fusion Center (FC). We simulate both analog and digital channel cases.

5.1 Numerical Analysis of Analog Method

In the case of Analog channel, each sensor observation is directly transmitted to the FC over a white Gaussian noise channel with variance $\eta^2$. The variance $\eta^2$ is selected such that the total signal-to-noise ratio ($SNR_C$) in the transmission channels defined as

$$ SNR_C = \frac{\int \int_A E[R^2(x, y)] dx dy}{A \eta^2} = \frac{\int \int_A G^2(x, y) dx dy}{A \eta^2} + \frac{\sigma^2}{\eta^2} $$
is 15 dB. The ML estimation problem is solved using Newton’s method. The peak value, $x$-location and $y$-location are displayed as a function of iteration in Fig. 5.2 for two different sets of initial vector-parameters. Each plot in Fig. 5.2 is due to a single realization of the distributed network with $K=20$. We can observe that with the initial values 9 for the peak of the field, 3 for the $x$-location and 3 for the $y$-location, the algorithm takes about 42 Newton iterations to converge to the final values 7.89, 4.11, and 3.73, respectively. With the initial values 7 for the peak of the field, 5 for the $x$-location and 5 for the $y$-location, the algorithm takes about 31 Newton iterations to converge to the final values 7.88, 4.19, and 4.01, respectively. The true values of these parameters are 8, 4, and 4. The discrepancy between the vectors of estimated and true parameters are due to a low sensor density in the network, and the distortions due to sensor and channel noise.

The square distance per pixel between the original and reconstructed Gaussian fields is displayed in Fig. 5.1(b).

To further analyze the estimation performance, we evaluate the mean square error (MSE) between the estimated and true location parameters. The MSE is evaluated numerically by means of 1000 Monte Carlo simulations. Each vector of estimated parameters is substituted back in the expression for the parametric field, and an integrated square error (ISE) between the true and estimated fields is evaluated. ISE is defined as:
(a) Peak value

(b) $X$-location  

(c) $Y$-location

Figure 5.2: Illustration of the convergence of the Newton’s method.

$$ISE = \frac{\int \int_A |\hat{G}(x, y) - G(x, y)|^2 dx dy}{\int \int_A |G(x, y)|^2 dx dy}, \text{ where } \hat{G}(x, y) = G(x, y : \hat{\theta}).$$ (5.3)

The ISE statistically averaged over 1000 Monte Carlo simulations is an approximation to the integrated mean square error (IMSE).

The dependence of the SE on the number of sensors, $K$, in the distributed network is shown in Fig. 5.3. The dependence of the ISE on the number of sensors (sensor density) in the distributed network is displayed in Fig. 5.4. The number of sensors distributed over the area $A$ is varied from 5 to 200 with the step 5. Each box in Fig. 5.3 and Fig. 5.4 is
generated using 1000 Monte Carlo realizations of the network. The central mark in each box is the median. The edges of the box present the 25th and 75th percentiles. The dashed vertical lines mark the data that extend beyond the two percentiles, but not considered as outliers. The outliers are plotted individually and marked with a “+” sign. The percentage of outliers due to divergence of the Newton’s method is depicted in Fig. 5.5. For this figure outlier is defined as the estimated vector $\hat{\theta}$ with the square error $SE = \sqrt{||\theta - \hat{\theta}||^2}$ above a threshold $\tau$. The parameter $\tau$ is varied in the range between 0 and 1 to generate the plots in Fig. 5.5. Note the large percentage of outliers for small values of $K$, $K = 10, 15, 20$. These correspond to the case when one of the three parameters did not converge to its true value.

The results in Fig. 5.3, Fig. 5.4 and Fig. 5.5 indicate that the location estimation and the field reconstruction of a relatively good quality is possible with the number of sensors equal or exceeding 20.

The dependence of performance on the noise is shown in Fig. 5.6. The plots are for three different $SNRs$ in observation and transmission channels $SNR_O = SNR_C = 10$ dB, $15$ dB and $20$dB, where the number of sensors equals to 20. As expected, the percentage of outliers
Figure 5.4: Dependence of the simulated ISE on the number of sensors distributed over the area $A$.

Figure 5.5: Probability of outliers $P_{\text{outliers}}(\tau) = P[SE > \tau]$ (expressed in percents) as a function of $\tau$. The plot is based on 1000 Monte Carlo simulations.
increases as the SNR decreases.

Figure 5.6: Probability of outliers (expressed in percents) as a function of the threshold for different values of SNR in observation and transmission channels.

5.2 Numerical Analysis of Digital Method

In the case of Digital channel, each sensor observation is quantized to one of \( M \) levels using a uniform deterministic quantizer. In this section, we set the number of quantization levels to \( M = 8 \) and the quantization step to 8. \( K \) parallel white Gaussian noise channels add samples of noise with variance \( \eta^2 \) selected to set the total SNR during data transmission defined as

\[
SNR_C = \frac{\int \int_A E[q^2(R(x,y))] \, dx \, dy}{A \eta^2} \tag{5.4}
\]

to 15 dB, and the FC observes the noisy quantized samples of the field. The function \( q(G(x,y : \theta)) \) in (5.4) is a quantized version of \( G(x,y : \theta) \) and \( E[q^2(R(x,y))] \) is the expected value of the squared output of the quantizer.

First, we illustrate convergence of the EM algorithm. The value of the ML estimate as a function of iteration is shown in Fig. 5.7 for (a) the peak of the field, (b) for its \( x \)-location and (c) \( y \)-location, respectively. Each illustration is based on a single realization.
Figure 5.7: Illustration of the EM convergence.

The deviation of the estimated vector parameter from the true vector $\theta$ can be attributed to many distortions. The major of them are observation and transmission noise.
Similar to the case of the analog channel, we evaluate the mean square error (MSE) between the estimated and true location parameters and the integrated MSE between estimated and true fields:

\[
ISE = \frac{\int \int_{A} |\hat{G}(x, y) - G(x, y)|^2 dx dy}{\int \int_{A} |G(x, y)|^2 dx dy}.
\] (5.5)

The dependence of the SE on the number of sensors, \(K\), in the distributed network for the case of \(M = 8\) quantization levels is shown in Fig. 5.8. The dependence of the ISE on the number of sensors (sensor density) in the distributed network for the same value of \(M\) is displayed in Fig. 5.9. The number of sensors distributed over the area \(A\) is varied from 5 to 200 with the step 5. Each box in Fig. 5.8 and Fig. 5.9 is generated using 1000 Monte Carlo realizations of the network and EM runs. The percentage of outliers due to divergence of the EM algorithm is depicted in Fig. 5.10.

Figure 5.8: A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area \(A\). The number of quantization levels is set to \(M = 8\).

Note that in the case of Digital channel the plot of \(SE\) versus \(K\) displays a larger number of outliers compared to the case of Analog channel. Also, in the case of digital channel ISE
Figure 5.9: Dependence of the simulated ISE on the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 8$.

Figure 5.10: Probability of outliers $P_{\text{outliers}}(\tau) = P[SE > \tau]$ (expressed in percents) as a function of $\tau$. The plot is based on 1000 Monte Carlo simulations. The number of quantization levels is set to $M = 8$. 
has a large variance (especially pronounced for small values of $K$) compared to the ISE in the analog case.

Fig. 5.11 compares the percentage of outliers plotted as a function of varying threshold for three different realizations of $SNR_O$ and $SNR_C$. Note that for $M = 8$ the effect of the SNR in the observation channel is more pronounced compared to the SNR in the transmission channel. The case of high $SNR_O = 20$ dB and low $SNR_C = 10$ dB is preferred by the estimator compared to the case of a low $SNR_O = 10$ dB and a high $SNR_C = 20$ dB.

Figure 5.11: Probability of outliers (expressed in percents) as a function of the threshold for different values of SNR in observation and transmission channels. The number of quantization levels is set to $M = 8$.

A set of box plots showing dependence of the SE and the ISE on the number of sensors distributed over the area $A$ for $M = 16$ and $M = 32$ are shown in Fig. 5.12, Fig. 5.13, Fig. 5.14 and Fig. 5.15. The results are similar to those for the case of $M = 8$ with the difference that the number of outliers as a function of the threshold decays to zero faster.

Fig. 5.8, Fig. 5.12 and Fig. 5.14 show the box plot of the SE. Although they all look similar, the median of SE as well as the box size decrease as the number of quantization levels, $M$, increases. Fig. 5.16 illustrates this observation for $K=20$. 
Figure 5.12: A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 16$.

Figure 5.13: Dependence of the simulated ISE on the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 16$. 
Figure 5.14: A box plot of the SE between the estimated and true location of the object displayed as a function of the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 32$.

Figure 5.15: Dependence of the simulated ISE on the number of sensors distributed over the area $A$. The number of quantization levels is set to $M = 32$. 
Figure 5.16: A box plot of the SE between the estimated and true location of the object displayed for different number of quantization levels. The number of sensors $K = 20$. SNRs for digital case and analog case are set to equivalent values. Note the convergence of the results as $M$ increases.
Chapter 6

Conclusion and Future Work

6.1 Summary

In this thesis, we proposed and analyzed an iterative ML solution to the problem of distributed estimation of a parametric field. The accuracy in estimating these parameters of the field depends on the number of sensors (samples) that are available at the fusion center (FC), quality of these samples ($SNR$), and the signal processing applied to the samples (quantization).

The model of the network assumed (1) independent Gaussian sensor and transmission noise; (2) quantization of sensory data prior to transmission (for the case of digital channel); and (3) parametric function estimation at the FC.

In the case of analog channel, a distributed ML estimation procedure for estimating a parametric physical field is formulated. Even with small number of sensors, $K = 5$, the algorithm converges. Increasing the number of sensors results in fewer outliers and thus in increased quality of the estimated values. Also, the algorithm takes fewer iterations to converge. Varying $SNR_O$ or $SNR_C$ have the same effect on the stability of the algorithm, since the data samples observed at the FC include the sum of the observation and transmission noise.

In the case of digital channel, a distributed ML estimation procedure for estimating a parametric physical field is also formulated. An iterative linearized EM solution is presented and numerically evaluated. The stability of the EM algorithm is evaluated for three different
values of $SNR_O$ and $SNR_C$. The results show that for a small number of quantization levels (quantization error is large), $SNR_O$ dominates $SNR_C$ in terms of its effect on the performance of the estimator. Also, when the sensor network is sparse, ($K$ is small, $K = 10, 15, 20$) the EM algorithm produces a substantial number of outliers. Denser networks, $K > 20$, are more stable in terms of reliable parameter estimation. A similar analysis has been performed for $M = 16$ and $M = 32$. These cases produce better estimates of the field, and also the EM algorithm takes fewer iterations to converge.

For large $K$, increasing the number of sensors does not have a notable effect on the performance of the algorithms.

### 6.2 Claimed Novelties

This work claims the following novelties:

- An Expectation Maximization (EM) algorithm was formulated to solve for unknown parameters of a deterministic parametric field from sparse noisy distributed observations.

- The solution to the problem is numerically analyzed for the case of analog and digital channels. A number of parameters such as $SNR_O$, $SNR_C$, the number of sensors and the number of quantization levels were varied. The emphasis was placed on determining the sparseness limits of the network.

### 6.3 Future Research

In the future, this work can be extended in many different ways. A few potential extensions are summarized below.

- The distributed network in our case assumes that a FC collects all sensory data and performs estimation of the field. A natural extension would be remove the FC and perform field estimation based on local exchanges among sensors.
In this thesis, the physical field is assumed to have a field generated by a single object. Assuming that multiple objects generate a cumulative field could be an interesting problem for the future research. Also, the assumptions of knowing the shape of a field is not very practical. Dealing with nonparametric fields will be more close to reality.

Throughout this work, we assumed a homogeneous network of sensors with the same variance of the noise in each observation. This work could be extended to deal with a more general case where the network is heterogeneous.

Uniform deterministic quantizer can be replaced by a randomized quantizer. The number of levels assigned to quantizer could vary from sensor to sensor according to the quality of observations and the noise in the communication channel between sensors and the FC.

The distribution of the physical field decays as a function of distance. The assumption that we know the shape of the influence field can be used to design a more efficient quantizer.

Due to the shape of the physical field, many sensors take close measurements and the output of the quantizer has more frequent output values than others. This work could be directed to deal with bandwidth constraints by assigning shorter length code words to the frequently occurring output values, which will result in an effective transmission scheme.

In this work, we assumed that the quantized version of observations were sent using Pulse-amplitude modulation. Other types of modulation schemes could be used to send the data from sensors to the FC. Comparing the performance of different modulation schemes for data transmission could be a very interesting study.

This work could be also extended by assuming more realistic channels between sensors and the FC. In our work, we described the channel between a sensor and the FC as a white Gaussian channel. A new study could assume that distances between each sensor
and the FC are different and the channels experience different pass losses. Shadowing and the effects of multi-path can also be taken into account.
References


REFERENCES


