

2006

# Spatial signal processing in wireless sensor networks

Benhong Zhang  
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>



Part of the [Electrical and Electronics Commons](#)

## Recommended Citation

Zhang, Benhong, "Spatial signal processing in wireless sensor networks " (2006). *Retrospective Theses and Dissertations*. 3048.  
<https://lib.dr.iastate.edu/rtd/3048>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

**Spatial signal processing in wireless sensor networks**

by

Benhong Zhang

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of  
**DOCTOR OF PHILOSOPHY**

Major: Electrical Engineering

Program of Study Committee:  
Aleksandar Dogandžić, Major Professor  
Zhengdao Wang  
Yao Ma  
Sang Kim  
Huaiqing Wu

Iowa State University

Ames, Iowa

2006

UMI Number: 3243529

UMI<sup>®</sup>

---

UMI Microform 3243529

Copyright 2007 by ProQuest Information and Learning Company.  
All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## TABLE OF CONTENTS

<b>LIST OF FIGURES</b> . . . . .	vi
<b>ABBREVIATIONS AND NOTATION</b> . . . . .	ix
<b>ACKNOWLEDGEMENTS</b> . . . . .	xi
<b>ABSTRACT</b> . . . . .	xii
<b>CHAPTER 1. GENERAL INTRODUCTION</b> . . . . .	1
1.1 Introduction . . . . .	1
1.2 Literature Review . . . . .	2
1.2.1 Field parameter estimation . . . . .	3
1.2.2 Event-region detection . . . . .	4
1.2.3 Target tracking and localization . . . . .	5
1.2.4 Communication cost analysis . . . . .	6
1.3 Thesis Organization . . . . .	7
<b>CHAPTER 2. DISTRIBUTED ESTIMATION AND DETECTION FOR SENSOR NETWORKS USING HIDDEN MARKOV RANDOM FIELD MODELS</b> . . . . .	9
2.1 Introduction . . . . .	9
2.2 Hidden Markov random field framework . . . . .	11
2.2.1 HMRFs as probabilistic graphical models . . . . .	13
2.3 ICM random-field estimation . . . . .	14
2.4 Event-region detection using the HMRF framework and ICM method . . . . .	16
2.4.1 Measurement-error model . . . . .	16

2.4.2	Autologistic MRF process model . . . . .	21
2.4.3	ICM detection for Gaussian measurement-error model . . . . .	22
2.4.4	ICM detection for nonparametric measurement-error models . . . . .	24
2.5	MRF calibration . . . . .	25
2.6	Numerical examples . . . . .	26
2.6.1	Guassian measurement scenario . . . . .	28
2.6.2	Quantized Gaussian measurement scenario . . . . .	30
2.6.3	Probabilities of false alarm and miss . . . . .	31
2.6.4	MRF calibration . . . . .	33
2.7	Concluding remarks . . . . .	34
2.8	Appendix A: Empirical likelihood and CRB for estimating $\mu_k$ . . . . .	35
2.9	Appendix B: Empirical entropy and CRB for estimating $\mu_k$ . . . . .	37
2.10	Appendix C: ICM detection for the Gaussian measurement-error model . . . . .	39
2.11	Appendix D: ICM detection for nonparametric measurement-error models . . . . .	40
2.12	Appendix E: GLR tests for $\mu_k$ under nonparametric measurement-error models . . . . .	41
<b>CHAPTER 3. BAYESIAN SIGNAL ESTIMATION FOR SENSOR NETWORKS IN THE PRESENCE OF NODE LOCALIZATION ERRORS . . . . .</b>		
3.1	Introduction . . . . .	42
3.2	Measurement model and prior specifications . . . . .	43
3.2.1	Measurement-error model . . . . .	44
3.2.2	Node location error model . . . . .	45
3.2.3	Prior specifications . . . . .	46
3.3	Bayesian analysis . . . . .	47
3.4	Cramer-Rao bound for the signal parameters . . . . .	49
3.5	Numerical examples: energy-based acoustic source localization . . . . .	50
3.5.1	Node localization . . . . .	51
3.5.2	Signal and noise models for energy-based acoustic source localization . . . . .	51
3.5.3	Prior specifications . . . . .	52

3.5.4	Initialization . . . . .	52
3.5.5	Simulation examples . . . . .	53
3.6	Concluding remarks . . . . .	55
3.7	Appendix A: Step 1 and 2 of the ICM algorithm . . . . .	56
3.8	Appendix B: CRB for ICM algorithm . . . . .	58
3.9	Appendix C: Asymptotic covariance for LS algorithm . . . . .	63
<b>CHAPTER 4. EVENT-REGION ESTIMATION FOR SENSOR NETWORKS</b>		
<b>UNDER THE POISSON REGIME . . . . .</b>		<b>66</b>
4.1	Introduction . . . . .	66
4.2	Measurement and event-region models . . . . .	67
4.2.1	Measurement-error model . . . . .	67
4.2.2	Event-region and noise models and prior specifications . . . . .	69
4.3	Bayesian analysis . . . . .	70
4.3.1	Simulating the parameters from posterior pdfs . . . . .	70
4.3.2	Estimating the event-region parameters . . . . .	72
4.4	Numerical examples . . . . .	73
4.5	Conclusion remarks . . . . .	76
<b>CHAPTER 5. MEAN-FIELD ESTIMATION AND DETECTION IN COR-</b>		
<b>RELATED GAUSSIAN RANDOM FIELD . . . . .</b>		<b>77</b>
5.1	Introduction . . . . .	77
5.2	Measurement model . . . . .	78
5.3	Mean-field estimation . . . . .	80
5.3.1	Distributed implementation . . . . .	80
5.3.2	Numerical example 1: mean field estimation . . . . .	81
5.4	Mean-field detection . . . . .	82
5.4.1	Batch detector . . . . .	83
5.4.2	Sequential detector . . . . .	83
5.4.3	Numerical example 2: mean-field detection . . . . .	85

5.5	Calibration . . . . .	87
5.5.1	Numerical example 3: rainfall precipitation data . . . . .	88
<b>CHAPTER 6. GENERAL CONCLUSIONS . . . . .</b>		<b>91</b>
6.1	Concluding remarks . . . . .	91
6.2	Future work . . . . .	92
<b>BIBLIOGRAPHY . . . . .</b>		<b>93</b>

## LIST OF FIGURES

Figure 1.1	Global (right) and localized (left) estimation and detection . . . . .	2
Figure 2.1	A graphical representation of an HMRF model. . . . .	14
Figure 2.2	(Left) Noiseless field and (right) a sensor network with $K = 1000$ nodes.	27
Figure 2.3	Gaussian measurement scenario: (Left) averaged observations $\bar{y}_k, k = 1, 2, \dots, K$ as functions of the node locations and (right) one-sided $t$ -test results for $P_{FA} = 5\%$ . . . . .	27
Figure 2.4	Gaussian measurement scenario: Event-region detection results after (left) one cycle and (right) two cycles of the Gaussian ICM algorithm.	28
Figure 2.5	Gaussian measurement scenario: Event-region detection results upon convergence of the Gaussian ICM algorithm. . . . .	28
Figure 2.6	Gaussian measurement scenario: Event-region detection results for (left) the empirical likelihood and (right) empirical entropy nonparametric ICM algorithms. . . . .	29
Figure 2.7	Quantized Gaussian measurement scenario: (Left) averaged observations $\bar{y}_k, k = 1, 2, \dots, K$ as functions of the node locations and (right) event-region detection results for the Gaussian ICM algorithm. . . . .	30
Figure 2.8	Quantized Gaussian measurement scenario: Event-region detection results for (left) the empirical likelihood and (right) empirical entropy nonparametric ICM algorithms. . . . .	31
Figure 2.9	Gaussian measurement scenario: Average probabilities of (left) false alarm and (right) miss, as functions of the number of observations per sensor $N$ . . . . .	31

Figure 2.10	Quantized Gaussian measurement scenario: Average probabilities of (left) false alarm and (right) miss, as functions of $N$ . . . . .	32
Figure 2.11	Noiseless field used for calibration. . . . .	34
Figure 3.1	A 2-D sensor network with 30 nodes and a region of interest in the proximity of an acoustic source. . . . .	50
Figure 3.2	Average MSEs, CRBs, and approximate covariances for the ICM and LS source location estimates as functions of the average node-location CRB, for (Left) $\overline{\text{SNR}} = 7\text{dB}$ and (Right) $\overline{\text{SNR}} = 12\text{dB}$ . . . . .	55
Figure 3.3	Average MSEs, CRBs, and approximate covariances for the ICM and LS source location estimates as functions of $\overline{\text{SNR}}$ , for (Left) $\overline{\text{crb}}_{\text{TOA}} = 0.02\text{m}^2$ and (Right) $\overline{\text{crb}}_{\text{TOA}} = 0.2\text{m}^2$ . . . . .	56
Figure 4.1	Noiseless field (left) and a realization of the heterogeneous Poisson field formed by the alarmed nodes, for $\theta = 1$ and $\lambda = 2$ (right). . . . .	73
Figure 4.2	Estimated event region obtained using the approximate MMSE estimates of the location and shape parameters (dashed line) and exact event region (full line). . . . .	74
Figure 4.3	Average MSEs for approximate MMSE estimates of the signal strength $\theta$ (left) and event-region radius $r$ (right), as function of $\theta$ , for $K \in \{1, 4\}$ and $\lambda = 1$ . . . . .	75
Figure 4.4	Average MSEs for approximate MMSE estimates of the event-region center coordinates: $z_1$ (left) and $z_2$ (right), as function of signal strength $\theta$ , for $K \in \{1, 4\}$ and $\lambda = 1$ . . . . .	75
Figure 5.1	MSEs for the ML and sample-mean estimates of the mean field $\alpha$ as functions of the spatial-dependence level $c$ . . . . .	82

- Figure 5.2 Exact and approximate detection probabilities of the proposed and white field batch and sequential tests (left) and average number of steps in which the proposed sequential test reached the decision, as functions of the true mean field  $\alpha$ , for  $P_{FA,d} = 0.05$  and  $P_{D,d} = 0.7$ . . . . . 85
- Figure 5.3 Achieved false-alarm probabilities of the proposed and white-field batch and sequential tests (left) and exact and approximate detection probabilities of the proposed batch and sequential tests (right) as functions of the desired false-alarm probability  $P_{FA,d}$ , for  $P_{D,d} = 0.7$  and  $\alpha = 0.15$ . 87
- Figure 5.4 Annual rainfall precipitation as a function of the measurement location (left) and 95% confidence interval for  $\alpha$  as a function of the spatial-dependence level  $c$ . . . . . 89

## ABBREVIATIONS AND NOTATION

Table 1 Abbreviations

ALMP	asymptotic local most powerful
CAR	conditional autoregressive
CRB	Cramér-Rao bound
FDR	false discovery rate
FIM	fisher scoring matrix
GLR	generalized likelihood ratio
HMRF	hidden Markov random fields
ICM	iterated conditional modes
i.i.d.	independent identically distributed
LS	least square
MAC	medium access control
MAP	maximum <i>a posteriori</i>
MCMC	Markov chain Monte Carlo
ML	maximum likelihood
MMSE	minimum mean-square error
MRF	Markov random field
MSE	mean-square error
NPMLE	nonparametric maximum likelihood estimate
pdf	probability density function
pmf	probability mass function
POI	phenomena of interest
PPL	pseudo predictive likelihood
RF	radio-frequency
SNR	signal to noise ratio
TOA	time-of-arrival
WSN	wireless sensor network

Table 2 Notation

$(\cdot)^T$	transpose
$(\cdot)^H$	conjugate transpose
$ \cdot $	determinant
$I_N$	Identity matrix with dimension $N \times N$
$\mathbf{1}_N$	vector of 1s with dimension $N \times 1$
$\mathbf{0}_N$	vector of 0s with dimension $N \times 1$
$i_A(x)$	indicator function, $i_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise} \end{cases}$

## ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, my advisor Dr. Aleksandar Dogandžić for his guidance, patience and support throughout this research and the writing of this thesis. His insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. I would also like to thank my committee members for their efforts and contributions to this work: Dr. Zhengdao Wang, Dr. Yao Ma, Dr. Sang Kim and Dr. Huaiqing Wu. I would additionally like to thank all teachers and students in Communication and Signal Processing group, for their unselfish help.

## ABSTRACT

Wireless sensor networks (WSNs) are gaining attention in recent years. Considering the potential low cost of a single sensor-processor unit in the near future, it is envisioned that there will be large-scale deployments of sensor networks for various applications: environmental, medical, inventory control, energy management, structural health monitoring, etc. A WSN comprises of a large number of nodes that individually have limited energy and computational power; however, by cooperating with each other, they can jointly perform tasks that are difficult to handle by traditional centralized sensing systems.

In this dissertation, spatial and spatio-temporal signal processing methods are developed for WSNs:

- Distributed estimation and detection using hidden Markov random fields: We derive ICM algorithms for distributed estimation of the hidden random field from the noisy measurements and consider both parametric and nonparametric measurement-error models.
- Parametric signal estimation in the presence of node localization errors: We propose a Bayesian framework that accounts for the inherent uncertainties in the node locations (caused by the node localization errors) and develop an estimation method that is robust to these uncertainties.
- Event-region estimation under the Poisson regime: We propose a parametric model for the location and shape of the event region and develop a Bayesian method for event-region estimation in large-scale sensor networks.
- Sequential mean-field estimation and detection in spatially correlated Gaussian fields: We propose distributed methods for estimating and detecting the mean of a correlated

Gaussian random field observed by a sensor network.

We consider estimation and detection of both localized and global phenomena and practically important nonparametric scenarios where the distribution of the measurements is unknown.

## CHAPTER 1. GENERAL INTRODUCTION

### 1.1 Introduction

There will be large-scale deployed sensor networks for various applications: environmental, medical, inventory control, energy management, health monitoring, etc., see also in [1–3]. A WSN comprises of a large number of spatially distributed nodes that individually have finite battery lifetime and thus limited computing and communication capabilities; however, by cooperating with each other, they can jointly perform tasks that are difficult to handle by traditional centralized sensing systems [3]. Sensor networks will monitor the environment at close range with high spatial and temporal resolutions and will likely reveal previously unobserved phenomena in the physical world [4]. Furthermore, carefully designed distributed processing algorithms will ensure the robustness and energy efficiency of the sensor network.

Due to the limited communication range for single sensor, it is prohibitively costly for the sensors in the network to send all their raw measurements to the fusion center (possibly far away) for centralized processing. Instead, each sensor may only be able to process its own measurement and communicate with a small portion of other sensors in its neighborhood. Therefore, successful deployment of sensor networks calls for developing distributed signal processing methods with two goals in mind:

- efficient extraction of information from noisy measurements collected by the sensors and
- maximization of the network lifespan (subject to limited power supply at each sensor).

Many problems in sensor networks have been extensively studied, e.g. field parameter estimation [5–7], target tracking and localization [8, 9] and event-region detection [10–12], see

also references therein. In general, the estimation and detection problems in this scenario can be divided into two categories, described below (see also Figure 1.1).

- *Localized estimation and detection.* Only small number of sensors that are within the range of phenomena observe the event of interest. Here, the event of interest is often hidden due to the presence of noise. Therefore, the goal is to develop efficient distributed algorithms to retrieve the hidden process from the noisy data.
- *Global estimation and detection.* Noisy information about an event of interest is sensed by the entire sensor network. In this case, spatial and temporal correlations between the measurements are usually substantial; incorporating these correlations into the detection and estimation algorithms is crucial for avoiding erroneous conclusions.

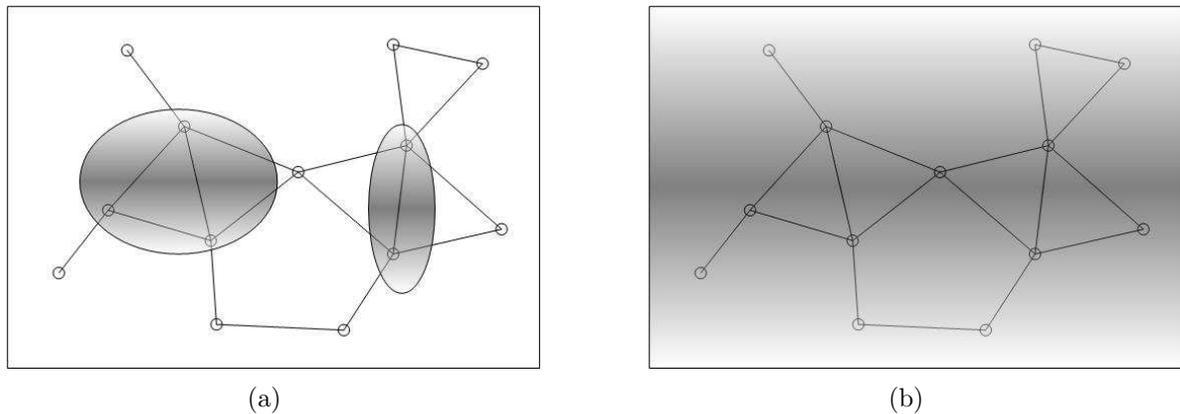


Figure 1.1 Global (right) and localized (left) estimation and detection

Both categories have their practical importance and will be discussed in this thesis.

## 1.2 Literature Review

Many research topics pertinent to WSNs have been studied in the past several years. We briefly review relevant literature in the following areas: field parameter estimation (Section 1.2.1), event-region detection (Section 1.2.2), target tracking and localization (Section 1.2.3) and communication cost analysis (Section 1.2.4).

### 1.2.1 Field parameter estimation

Sensors are typically densely deployed to measure the field of interest at its own location under this scenario. Assume that the measurement at  $n$ th sensor has the form:

$$y_n = f(\mathbf{x}_n; \boldsymbol{\theta}) + e_n \quad (1.1)$$

where  $y_n$  denotes the  $n$ th sensor's measurements,  $\mathbf{x}_n$  the sensor's location, and  $f(\mathbf{x}; \boldsymbol{\theta})$  the parameterized field strength, with parameter  $\boldsymbol{\theta}$ . The goal here is to estimate  $\boldsymbol{\theta}$  using measurements  $y_n$ s. In most literature, the field strength is assumed to be constant:  $f(\mathbf{x}_n; \boldsymbol{\theta}) = \boldsymbol{\theta}$ .

A significant amount of work has been done under the above settings. In [5], mean field estimation from coarsely quantized measurements is proposed and analyzed. It is shown that the CRB of the ML estimator under heavily quantized measurements would be surprisingly close to the ML estimator for continuous measurements, given the data quantization strategy is optimal and the number of sensors are sufficiently large. However, optimal quantization scheme can only be obtained with the actual value of field mean, which is an unknown parameter in practice. Therefore, the performance of ML estimator degrades if the quantization scheme deviates from optimal, thus reduces the value of the theoretical analysis in the paper.

In [7], correlated noise is taken into account in centralized global mean field estimation. This paper adopts a simple one-dimensional structure for spatial noise correlation and assumes that the noise is independent in time, and then shows that the relative performance loss of neglecting noise correlation would not exceed 14%. However, the correlation structure proposed in the paper is too simple for realistic problems; the authors also do not provide a solution to handle distributed estimation under spatially correlated noise.

In [6], a general algorithm using LS criterion is proposed for distributed estimation of inhomogeneous fields (corresponding to localized phenomena). Instead of assuming constant field, the authors allow the field to vary arbitrarily, with very loose smoothness constraints. In the proposed algorithm, the field is adaptively partitioned into squares with different sizes, to yield the smallest penalized function, which favors smaller sample squared error as well as fewer partitions. In addition, the proposed algorithm requires only local area communication between sensors, which makes it truly distributed.

### 1.2.2 Event-region detection

Sensor networks are often constructed to detect events within the region of interest. For example, sensors may be deployed to detect region of the environment whose concentration of a certain chemical is greater than a specified level.

Most previous work focused on detecting a global event that is measured by the whole network. Saligrama *et al.* [13] consider an  $M$ -ary detection problem using data collected by  $N$  distributed noisy sensors. In [13], MAP criterion is applied for hypothesis testing, with arbitrary but known measurement models under all  $M$  hypothesis. Testing results are achieved by reaching consensus using proper message exchange among neighboring sensors, which is defined by a connectivity graph. It is also shown that a consensus to the centralized MAP estimate can almost always be reached for any arbitrary network structure, with finite link capacity constraint.

In [12], a decentralized algorithm is proposed under the assumption that noise distributions are unknown and different across the sensors, with the existence of a fusion center which collects testing results from all sensors. Theoretical investigations on fusion architecture of global phenomena are presented in [14–16].

Recently, detecting localized phenomena where only part of the network observe the event of interest has attracted attention. In [11], Krishnamachari and Iyengar propose a distributed Bayesian algorithm for event-region detection using binary sensor measurements.

Sung *et al.* [10] also assume binary sensor measurements, with a more general setup that the POI over a region has spatially varying signal strength. In the paper, the POI over the region is modeled as a deterministic spatial signal with known shape  $s(\mathbf{x})$ , and unknown signal strength. Each sensor observes:

$$\begin{cases} y_n = \gamma(\mathbf{x}_n) + e_n = \theta s(\mathbf{x}_n) + e_n, & \mathbf{x}_n \text{ within POI} \\ y_i = e_n, & \text{otherwise} \end{cases} \quad (1.2)$$

where  $y_n$  is the sensor observation,  $\mathbf{x}_n$  is the location coordinate of sensor  $n$ ,  $\theta$  is the unknown signal amplitude and  $e_n$  denotes the i.i.d. observation noise at location  $\mathbf{x}_n$ . Then, binary

decision  $u_n$  is made at sensor  $n$  and collected by the fusion center,

$$u_n = \begin{cases} 1, & y_n > \tau_0 \\ 0, & y_n \leq \tau_0 \end{cases} \quad (1.3)$$

where  $\tau_0$  is chosen to satisfy certain false alarm rate  $\alpha_0$  for a single sensor decision. The authors then propose the detection of spatially varying signal under the Poisson regime, assuming that the initial sensor distribution is homogeneous Poisson with intensity  $\lambda_h$ . It is shown with the above decision rule, alarmed sensors (sensors with decision  $u_n = 1$ ) follow a nonhomogeneous Poisson distribution with local intensity

$$\lambda(\mathbf{x}) = \theta \lambda_h p(\mathbf{x}) \quad (1.4)$$

where  $p(\mathbf{x}_n) = \Pr\{u_n = 1\}$ . With moderate constraints on  $s(\mathbf{x})$  and  $p(\mathbf{x})$ , ALMP detector and optimal threshold for single sensor are derived in [10] under the above Poisson regime. It is claimed that the ALMP test statistic is a weighted sum of the local decisions, where the optimal weights are the shape of the spatial signal  $s(\mathbf{x})$ .

Both algorithms proposed in [10] and [11] sacrifice the measurement accuracy by utilizing binary data to reduce the communication cost between sensors and fusion center, and therefore are well suited for sensor network scenario.

In [17], Ermis and Saligrama propose a distributed implementation of an FDR approach for localized event region detection. When performing multiple comparison tests at different nodes, the FDR approach controls the the false discovery rate. This is in contrast to the traditionally used methods such as the Bonferroni procedure [15] that control the false alarm probability at each sensor. The authors present numerical simulations showing the superior performance of FDR approach in boundary detection compared with the Bonferroni procedure.

### 1.2.3 Target tracking and localization

The problem of localizing and tracking a target has been widely sighted as a canonical application of WSNs [18]. Because of its spatial coverage and multiplicity in sensing aspect and modality, a sensor network is ideally suited for tracking moving phenomena traversing the range of many sensors in a large area [19].

Zhao *et al.* [20] present information-based approaches for target localization where sensors are activated (queried) based on information content as well as constraints on resource consumption, latency, etc. Specifically, a mutual information-based sensor selection criterion has been developed and successfully tested on real data [21], which in each time slot, select a sensor whose measurements would provide the greatest amount of information about the target location.

In [22], an ML target localization estimator is proposed for binary measurements. With the assumption of spatial and temporal independent, identical Gaussian noise, it is claimed by maximizing the joint likelihood of target location, the estimator would reach CRB with small amount of communication. The estimation performance using multi-bit data, and the tradeoff between multi-bit and multi-time frame are also discussed. However, the paper starts from i.i.d. Gaussian noise assumption, which is quite a tight restriction in practice.

Rabbat and Nowak [18] provide decentralized source localization and tracking algorithm based on signal strength. In the algorithm described in [18], the parameter estimate is cycled through the network, each sensor makes small adjustment on the current estimate and passes the updated estimate to its neighbors, where the update is based on incremental subgradient optimization. The algorithm also converges with limited cycles of communications, which makes it energy efficient compare with most existing tracking algorithms.

#### 1.2.4 Communication cost analysis

Significant research efforts have focused on predicting or reducing the overall communication cost in WSNs given certain performance requirements. In [23], Krause *et al.* propose a near-optimal sensor placement strategy for maximizing information while minimizing communication cost. By assuming the field of interest to be spatial Gaussian process, an approximate algorithm is presented to minimize the overall expected number of retransmissions at the constraint of a specified estimation accuracy expected from the sensor network, quantized by the mutual information.

Giridhar *et al.* [24] present the theoretical communication cost of computing a function

$f(x_1, x_2, \dots, x_n)$  of the readings  $x_1, x_2, \dots, x_n$  taken at the  $n$  sensors under various network structures. The simplest two-node network is studied first and then expand to multiparty networks. At last, reliable computation in the network with noise is discussed.

### 1.3 Thesis Organization

Included in this thesis are two submitted journal papers as well as several conference publications describing our distributed algorithms for sensor network scenario, listed as follows:

- *Distributed signal processing using HMRF* (Chapter 2 and [25]). MRFs have been widely used to describe spatially distributed random phenomena [26, 27]. Assuming the hidden spatial process forms a random field that has simple structure with Markovian dependence (see Section 2.2 for details), we derive ICM algorithms for distributed estimation of a localized phenomenon (modeled as a *hidden random field*) from noisy measurements. The proposed ICM algorithm only requires binary data communication among neighboring sensors and does not require the existence of a fusion center.
- *Bayesian signal estimation in the presence of node localization errors* (Chapter 3 and [28, 29]). Most nodes in WSN estimate their location, see [1, 30, 31]. However, existing signal processing methods for sensor network environment have ignored the effects of node localization inaccuracies. Here, we propose a Bayesian framework that takes into account of node location errors and develop the MAP estimation. In addition, we compute the CRB for the energy-based acoustic source localization and show that the proposed algorithm nearly reaches the optimal performance predicted by the CRB.

In addition, the thesis briefs the following works, which appeared as conference publications:

- *Event-region estimation for sensor networks under Poisson regime* (Chapter 4 and [32]). Sung *et al.* [10] proposed an asymptotically optimal algorithm to estimate signal strength over a region with a large number of identical binary sensors deployed under homogeneous Poisson distribution. However, signal source location was assumed to be known in the paper. In Chapter 4 of this thesis, we introduce a MAP algorithm that can estimate the

signal strength as well as its source location; in addition, we also generalize the Poisson regime to account for the scenario where the nodes utilize multiple local thresholds to quantize the sensor measurements.

- *Mean-field estimation and detection in correlated Gaussian random field* (Chapter 5 and [33]). In large-scale sensor networks, sensors will be densely deployed on the field of interest to gain high spatial and temporal resolution. However, big density also renders the nodes' measurements highly correlated in space [7, 34–38]. Here, we propose a distributed ML method to estimate the mean of the spatially correlated field; we then propose a sequential detector to test whether the mean of the field is greater than or equal to 0, without losing generality.

The rest of the thesis is organized as follows: Chapter 2 and 3 present *Distributed estimation and detection for sensor networks using hidden Markov random field models* and *Bayesian signal estimation for sensor networks in the presence of node localization errors*, respectively. Chapter 4 shows the binary data processing under Poisson regime. The three chapters above discuss the local estimation and detection problems, shown in Figure 1.1(a). Chapter 5 presents the work on field mean estimation with correlated spatial noise, which is global decentralized detection and estimation, see Figure 1.1(b). Chapter 6 concludes the finished works and depicts the directions for future research.

## CHAPTER 2. DISTRIBUTED ESTIMATION AND DETECTION FOR SENSOR NETWORKS USING HIDDEN MARKOV RANDOM FIELD MODELS

A paper to appear in *IEEE Transactions on Signal Processing*

Aleksandar Dogandžić and Benhong Zhang

### Abstract

We develop a HMRF framework for distributed signal processing in sensor-network environments. Under this framework, spatially distributed observations collected at the sensor form a noisy realization of a underlying random field that has a simple structure with Markovian dependence. We derive ICM algorithms for distributed estimation of the hidden random field from the noisy measurements. We consider both parametric and nonparametric measurement-error models. the proposed distributed estimators are computationally simple, applicable to a wide range of sensing environments, and *localized*, implying that the nodes communicate only with their neighbors to obtain the desired results. We also develop a calibration method for estimating MRF model parameters from training data and discuss initialization of the ICM algorithms. The HMRF framework and ICM algorithms are applied to event-region detection. Numerical simulations demonstrate the performance of the proposed approach.

### 2.1 Introduction

Recent advances in integrated sensor and RF technologies, wireless communications, and signal processing allow development of sensor-network systems composed of low-cost sensor-processor elements (nodes) jointly working to collect and analyze *noisy* spatio-temporal mea-

surements. Large-scale sensor networks that can monitor an environment at close range with high spatial and temporal resolutions are expected to play an important role in various applications, including assessing "health" of machines, aerospace vehicles, and civil-engineering structures; environmental, medical, food-safety, and habitat monitoring; energy management, inventory control, home and building automation, see also [1–3, 39–41]. Each node will have limited sensing, signal processing, and communication capabilities, but by cooperating with each other they will accomplish tasks that are difficult to perform with conventional centralized sensing systems [3]. Sensor networks are expected to reveal previously unobservable phenomena in the physical world and are currently attracting considerable attention.

MRF models have been widely used to describe spatially distributed random phenomena, see e.g. [26, 42]. In this paper (see also [43]), we propose a HMRF framework for distributed signal processing in sensor-network environments. Under this framework, spatially distributed observations (collected at the sensors) form a noisy realization of a random field with Markovian dependence structure.<sup>1</sup> Previous work on distributed HMRF based signal processing for sensor networks focused on developing message passing algorithms for linear Gaussian measurement-error and MRF process models with known model parameters, see also the discussion in Section 2.2.1. In contrast, our HMRF framework allows for general measurement and random-field models with *unknown* measurement error model parameters. The unknown measurement-error model parameters vary from one node to another, thus taking into account imperfect calibration of the sensors at different nodes and permitting distributed localized processing and nonparametric measurement-error modeling. The nonparametric measurement-error models that we employ are important in practical applications where accurate parametric models are difficult to find, especially in large-scale sensor networks operating in time-varying environments [44–46].

We derive ICM algorithms for distributed estimation of a localized phenomenon (modeled as a *hidden random field*) from noisy measurements. In particular, the proposed ICM algo-

---

<sup>1</sup>Here, Markovian dependence implies that, given random-field values at all other locations, the conditional distribution of the random field at any location depends only on the field values at the neighboring locations, see also ( 2.10) in Section 2.2.

gorithms are designed to increase the *predictive likelihood* of the hidden field.<sup>2</sup> The underlying *distributed processing paradigm* ensures robustness and reliability of the proposed approach. We demonstrate our approach by applying it to event-region detection, which is an important task in wireless sensor networks [11]. We consider parametric Gaussian and nonparametric (empirical likelihood and entropy) measurement-error models and utilize an autologistic MRF process model for event-region detection.

The HMRF framework is introduced in Section 2.2 and general ICM method is presented in Section 2.3. We discuss the event-region detection problem in Section 2.4 where we first propose suitable measurement-error and random-field models (Sections 2.4.1 and 2.4.2) and then derive the corresponding ICM detection algorithms (Sections 2.4.3 and 2.4.4). Initialization of the ICM iterations is discussed in Sections 2.4.3 and 2.4.4. In Section 2.5, we develop a PPL calibration method for estimating MRF model parameters from training data and specialize it to the event-region detection problem. This method is based on maximizing the product of the full conditional predictive pdfs/pmfs of the random-field values at all the nodes. In Section 2.6, we evaluate the performance of the proposed detection algorithms via numerical simulations. Concluding remarks are given in Section 2.7.

## 2.2 Hidden Markov random field framework

Assume that each node (sensor)  $k \in \{1, 2, \dots, K\}$  in the network collects a vector of measurements

$$\mathbf{y}_k = [y_k(1), y_k(2), \dots, y_k(N)]^T \quad (2.1)$$

where  $N$  denotes the number of observations collected by each node  $k$  and "T" denotes a transpose. Define also the vector of all measurements:

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_K^T]^T \quad (2.2)$$

We assign a hidden random variable  $\beta_k$  to each node  $k$  and adopt the following hierarchical model for the collected observations:

<sup>2</sup>See [47] for the definition of predictive likelihood and examples of its use.

- $\beta_k, k = 1, 2, \dots, K$  form an MRF describing the *process model*:

$$\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_K]^T \quad (2.3)$$

- Given the MRF  $\boldsymbol{\beta}$ ,  $\mathbf{y}_k$  are conditionally independent random vectors with pdfs or pmfs  $p_{\mathbf{y}_k|\beta_k}(\mathbf{y}_k|\beta_k; \mathbf{v}_k)$  that depend only on  $\beta_k$

$$p_{\mathbf{y}|\boldsymbol{\beta}}(\mathbf{y}|\boldsymbol{\beta}; \mathbf{v}) = \prod_{k=1}^K p_{\mathbf{y}_k|\beta_k}(\mathbf{y}_k|\beta_k; \mathbf{v}_k) \quad (2.4)$$

describing the *data (measurement-error) model*.

Here,  $\mathbf{v}_k$  is the vector of *unknown* measurement-error model parameters at the  $k$ th node and

$$\mathbf{v} = [\mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_K^T]^T \quad (2.5)$$

Note that the measurement-error model parameters  $\mathbf{v}_k$  vary with the node index  $k$ , taking into account *imperfect calibration* of the sensors at different nodes. The above framework can account for both discrete and continuous measurements and random fields. The parameters  $\mathbf{v}_k$  may be used to model the entire measurement-error probability distribution  $p_{\mathbf{y}_k|\beta_k}(\mathbf{y}_k|\beta_k; \mathbf{v}_k)$  in a *nonparametric manner*, provided that the elements of  $\mathbf{y}_k$  are conditionally i.i.d.; see Section 2.4.1.

Our goal is to estimate the MRF  $\boldsymbol{\beta}$  from the observations  $\mathbf{y}_k, k = 1, 2, \dots, K$ . We define the probability distribution of  $\boldsymbol{\beta}$  via a *conditionally-specified model* suitable for distributed neighborhood-based signal processing. Before formally defining an MRF, let us introduce some terminology and notation. Throughout this paper, we assume that the *neighborhood* of a node  $k$  [denoted by  $\mathcal{N}(k)$ ] consists of all the nodes  $l \in \{1, 2, \dots, K\}$  that are within a *cutoff distance*  $d$  from that node, i.e.

$$\mathcal{N}(k) = \{l : \|\mathbf{r}_k - \mathbf{r}_l\| \leq d \text{ and } l \neq k\} \quad (2.6)$$

where

$$\|\mathbf{r}_k - \mathbf{r}_l\| = \sqrt{(\mathbf{r}_k - \mathbf{r}_l)^T (\mathbf{r}_k - \mathbf{r}_l)} \quad (2.7)$$

and  $\mathbf{r}_k$  and  $\mathbf{r}_l$  are the  $k$ th and  $l$ th node locations in Cartesian coordinates. Define the set of random-field values in this neighborhood:

$$N_\beta(k) = \beta_l, l \in \mathcal{N}(k) \quad (2.8)$$

and the conditional pdfs or pmfs of  $\beta_k$  given the neighboring MRF values:

$$p_{\beta_k|N_\beta(k)}(\beta_k|N_\beta(k)), k = 1, 2, \dots, K \quad (2.9)$$

Then, the Markov property of an MRF  $\beta$  implies that, for all  $k = 1, 2, \dots, K$ , the conditional pdfs/pmfs of  $\beta_k$  given the random-field values at all other nodes satisfy

$$p_{\beta_k|\{\beta_l, l \neq k\}}(\beta_k|\{\beta_l, l \neq k\}) = p_{\beta_k|N_\beta(k)}(\beta_k|N_\beta(k)) \quad (2.10)$$

### 2.2.1 HMRFs as probabilistic graphical models

MRF and HMRF models belong to the (broader) class of *probabilistic graphic models* (see e.g. [42, 48–52] and reference therein) and can be formulated using an undirected mathematical graph whose nodes correspond to the random variables in the field and its edges define the underlying neighborhood structure. In [48, 50], graphical-model based extended *message passing*<sup>3</sup> algorithms are developed for inference on HMRF models with linear Gaussian measurement-error and MRF process models and known model parameters, *embedded-trees* and *embedded-triangles* algorithms are developed for this scenario in [49, 50, 53]. A *belief propagation* approach is proposed in [54] for multi-hypothesis testing of global phenomena in sensor-network environments. Figure 2.1 shows a graphical representation of an HMRF model, where the filled circles depict the hidden random field (and their locations correspond to the node locations) and hollow circles the observed data. The edges in Figure 2.1 describe the (conditional) statistical dependence between the nodes in the graph, as inferred from the specifications in ( 2.6) and ( 2.4).

In the following we present a *distributed algorithm* for computing maximum predictive likelihood estimates of the random field  $\beta$ .

<sup>3</sup>See [52] for a detailed exposition on message passing algorithms for graphical models.

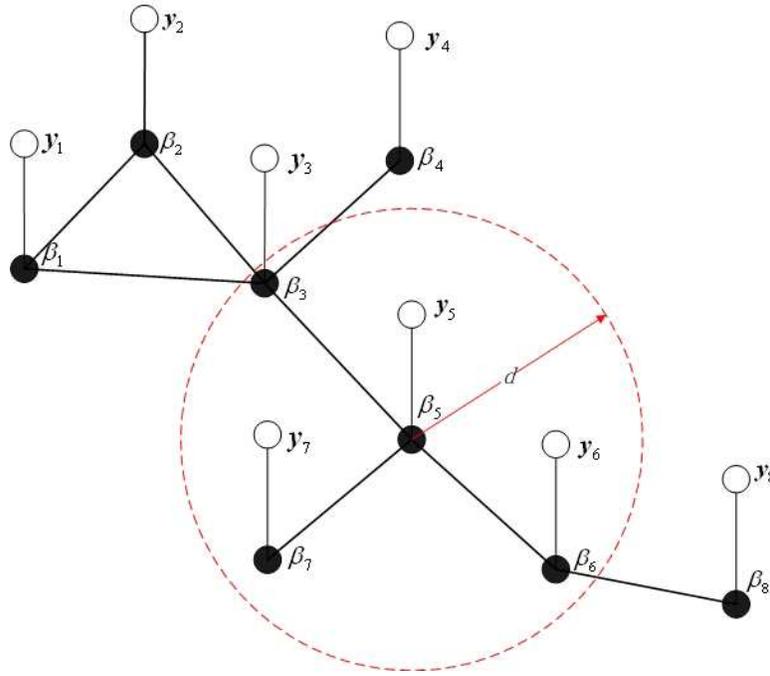


Figure 2.1 A graphical representation of an HMRF model.

### 2.3 ICM random-field estimation

We propose an ICM algorithm for estimating the MRF  $\beta$  where each node  $k \in \{1, 2, \dots, K\}$  performs the following steps:

- (ICM1) collects the current estimates of  $\beta_l$  from its neighborhood  $\mathcal{N}(k)$ ;
- (ICM2) updates its estimate of  $\beta_k$  by maximizing the conditional predictive log likelihood:

$$L_k(\beta_k | N_\beta(k)) = \max_{\mathbf{v}_k} \{ \ln p_{\mathbf{y}_k | \beta_k}(\mathbf{y}_k | \beta_k; \mathbf{v}_k) \} + \ln p_{\beta_k | N_\beta(k)}(\beta_k | N_\beta(k)) \quad (2.11)$$

with respect to  $\beta_k$ ;

- (ICM3) *broadcast* the obtained estimate of  $\beta_k$  to the nodes in the neighborhood  $\mathcal{N}(k)$ .

When applied to each node  $k$  in turn, this procedure defines a single cycle of the ICM algorithm. The cycling is performed until convergence, i.e. until the estimates of  $\beta_k$  do not change significantly for all  $k \in \{1, 2, \dots, K\}$ . The ICM approach is computationally simple and applicable to a wide range of sensing environments. It *does not* require careful treatments of loops in the

inference graphs, constructing junction trees etc. It is also *localized*, implying that the nodes communicate only with their neighbors to obtain the desired results. Localized algorithms are *robust* to node failures and the communication overhead *scales well* with increasing network size, see [2, 39]. Distributed localized algorithms and architectures also facilitate rapid data processing and information collection, and are well-suited for systems that utilize sleep modes to conserve energy [55].

Denote by  $p_{\boldsymbol{\beta}}(\boldsymbol{\beta})$  the joint pdf/pmf of  $\beta_1, \beta_2, \dots, \beta_K$ . Then, applying (ICM1)-(ICM3) at each node  $k$  increases the joint predictive log-likelihood function of  $\boldsymbol{\beta}$  (see also [47])<sup>4</sup>:

$$L(\boldsymbol{\beta}) = \max_{\mathbf{v}} \{ \ln[p_{\mathbf{y}|\boldsymbol{\beta}}(\mathbf{y}|\boldsymbol{\beta}; \mathbf{v})] \} + \ln[p_{\boldsymbol{\beta}}(\boldsymbol{\beta})] \quad (2.12)$$

in a *stepwise-ascent manner*. In particular, combining the stepwise-ascent maximization approach with the Markovian property of  $p_{\boldsymbol{\beta}}(\boldsymbol{\beta})$  leads to the distributed localized iteration (ICM1)-(ICM3). In general, this iteration converges to a *local maximum* of  $L(\boldsymbol{\beta})$ . However, if the conditional predictive log likelihoods in 2.11 are *unimodal* in  $\beta_k$  (as in the HMRFs with linear Gaussian measurement-error and MRF process models studied in [48–50, 53]), then the ICM algorithm converges to the global maximum of  $L(\boldsymbol{\beta})$ . Interestingly, its convergence to a local maximum of  $L(\boldsymbol{\beta})$  [whenn initialized with the local ML estimates of the  $\beta_k$ 's] may be preferred compared with finding the global maximum because MRFs often have undesirable large-scale properties [27].

The predictive log likelihood in ( 2.12) has a Bayesian interpretation. Here, we view  $p_{\boldsymbol{\beta}}(\boldsymbol{\beta})$  as the *prior distribution* of the hidden field  $\boldsymbol{\beta}$  and assign a flat prior distribution:  $p_{\mathbf{v}}(\mathbf{v}) \propto 1$  to the measurement-error model parameters  $\mathbf{v}$ . Then maximizing  $L(\boldsymbol{\beta})$  in ( 2.12) yields a mode of the joint posterior pdf/pmf of the unknown parameters. We emphasize that the purpose of the proposed method is to resolve ambiguous measurements. Otherwise, if the data provides strong evidence about the hidden field  $\boldsymbol{\beta}$ , the influence of the prior  $p_{\boldsymbol{\beta}}(\boldsymbol{\beta})$  disappears, which is true for Bayesian methods in general. The ICM approach to finding modes of joint posterior distributions dates back to the seminal paper by Lindley and Smith [56], see also [57, 58].

<sup>4</sup>Note that the conditional predictive log likelihood  $L_k(\beta_k|N_{\boldsymbol{\beta}}(k))$  in ( 2.11) follows from the joint predictive log-likelihood  $L(\boldsymbol{\beta})$  by substituting the identity  $p_{\boldsymbol{\beta}}(\boldsymbol{\beta}) = p_{\beta_k|N_{\boldsymbol{\beta}}(k)}(\beta_k|N_{\boldsymbol{\beta}}(k)) \cdot p_{N_{\boldsymbol{\beta}}(k)}(N_{\boldsymbol{\beta}}(k))$  into ( 2.12) and keeping the terms that depend on  $\beta_k$ .

The iteration (ICM1)-(ICM3) generalizes the ICM algorithm for more general neighborhood models and unknown measurement-error model parameters that vary from node to node. The latter extension is key for sensor-network applications where the nodes are not perfectly calibrated and data processing should be performed *locally* as much as possible. It also allows nonparametric measurement-error modeling, as discussed in Section 2.4.1.

In the following section, we demonstrate the proposed approach by applying it to event-region detection.

## 2.4 Event-region detection using the HMRF framework and ICM method

We utilize the proposed HMRF framework and ICM method to efficiently remove false alarms in event-region detection tasks. Here, our goal is to detect a region in the environment in which an event of interest has occurred. For example, if the network is capable of sensing concentration of chemical  $X$ , then it is of interest to answer the following question [11]: "In which regions in the environment is the concentration of chemical  $X$  greater than a specified level?"

We first describe measurement-error and process models suitable for event-region detection (Sections 2.4.1 and 2.4.2) and then derive the corresponding ICM algorithms for event-region detection (Sections 2.4.3 and 2.4.4).

### 2.4.1 Measurement-error model

In this section, we consider hidden fields that take two discrete values

- $\beta_k = 0$  (signal absent) and
- $\beta_k = 1$  (signal present)

and utilize a simple signal-plus-noise measurement-error model for the measurements  $y_k(1), y_k(2), \dots, y_k(N)$  collected at node  $k \in \{1, 2, \dots, K\}$ :

$$y_k(t) = \mu_k(\beta_k) + e_k(t), t = 1, 2, \dots, N \quad (2.13)$$

where

- $\mu_k(0) = 0$  (signal absent),
- $\mu_k(1)$  is the (unknown) nonzero signal, and
- $e_k(t), t = 1, 2, \dots, N$  is zero-mean i.i.d. noise.

We denote the pdf/pmf of the noise  $e_k(t)$  by  $p_{\text{noise}_k}(e_k(t))$ . Consequently, given  $\beta_k$ ,  $y_k(1), y_k(2), \dots, y_k(N)$  are conditionally i.i.d. random variables with the joint pdf/pmf

$$p_{\mathbf{y}_k|\beta_k}(\mathbf{y}_k|\beta_k) = \prod_{t=1}^N p_{\text{noise}_k}(y_k(t) - \mu_k(\beta_k)) \quad (2.14)$$

see also 2.4 for a full measurement-error model specification.

1) *Gaussian measurement-error model*: Under the Gaussian measurement-error model, we assume that the noise pdf at node  $k$  is zero-mean Gaussian:

$$p_{\text{noise}_k}(e_k(t); \sigma_k^2) = \frac{1}{\sigma_k \sqrt{2\pi}} \cdot \exp[-e_k^2(t)/(2\sigma_k^2)] \quad (2.15)$$

where  $\sigma_k^2$  is the unknown noise variance at the  $k$ th sensor. Here, the measurement-error model parameters are  $\mathbf{v}_k = \sigma_k^2$  for  $\beta_k = 0$  and  $\mathbf{v}_k = [\mu_k(1), \sigma_k^2]^T$  for  $\beta_k = 1$ .

2) *Nonparametric measurement-error models*: We now consider a scenario where the noise probability distribution  $p_{\text{noise}_k}(\cdot)$  at node  $k$  is *unknown* and utilize a class of *nonparametric* measurement-error models. This scenario is important in practical applications where accurate parametric measurement-error model are difficult to find, as is often the case in large-scale sensor networks operating in time-varying environments (see e.g. [44–46]). To simplify the notation, we omit the dependence of the mean value  $\mu_k$  on  $\beta_k$  throughout this section. Clearly, the discussion on *unknown*  $\mu_k$  corresponds to the case where  $\mu_k = \mu_k(1) \neq 0$ .

Assume that, given  $\beta_k$ ,  $y_k(1), y_k(2), \dots, y_k(N)$  are conditionally i.i.d. random variables with mean  $\mu_k = \mu_k(\beta_k)$  where each  $y_k(t)$  is assigned a multinomial probability  $p_{k,t}$ . We then construct the following nonparametric log-likelihood function of the mean  $\mu_k$  at node  $k$ :

$$l_k(\mu_k) = \sum_{t=1}^N \ln p_{k,t}(\mu_k) \quad (2.16)$$

where  $p_{k,t}(\mu_k), t = 1, 2, \dots, N$  are estimates of the probabilities  $p_{k,t}, t = 1, 2, \dots, N$  computed by minimizing the Cressie-Read power divergence<sup>5</sup> between the discrete distribution defined

<sup>5</sup>The Cressie-Read divergence is closely related to the Renyi divergence [59], see also [60, 61]

by  $p_{k,t}, t = 1, 2, \dots, N$  and the discrete uniform distribution on  $t = 1, 2, \dots, N$  (see [60–62]):

$$\min_{p_{k,t}, t=1,2,\dots,N} \frac{\sum_{t=1}^N [(Np_{k,t})^{-\kappa} - 1]}{N\kappa(1 + \kappa)} \quad (2.17)$$

subject to the constraints

$$\sum_{t=1}^N p_{k,t}[y_k(t) - \mu_k] = 0, p_{k,t} \geq 0, \sum_{t=1}^N p_{k,t} = 1 \quad (2.18)$$

Here,  $-\infty < \kappa < \infty$  is a *known* constant [define a particular choice of the discrepancy measure in ( 2.17)] and the degenerate cases  $\kappa \in \{0, -1\}$  are handled by taking limits.

In the following, we focus on the non-trivial case where<sup>6</sup>

$$y_{k,\text{MIN}} = \min_{t \in \{1,2,\dots,N\}} y_k(t) < \mu_k < \max_{t \in \{1,2,\dots,N\}} y_k(t) = y_{k,\text{MAX}} \quad (2.19)$$

and on the limiting values of  $\kappa$  (i.e.  $\kappa = 0$  and  $\kappa = -1$ ), which correspond to commonly used *least favorable* distribution families [62] and lead to the *empirical likelihood* and *empirical entropy* measurement-error models discussed below. (The concept of a least favorable family was introduced by Stein in [63].)

**Empirical likelihood:** If  $\kappa = 0$  in ( 2.17), ( 2.16) simplifies to the following *concentrated empirical log-likelihood function*<sup>7</sup> of the mean  $\mu_k$  at node  $k$ :

$$l_k(\mu_k) = \left\{ \max_{p_{k,t}, t=1,2,\dots,N} \left( \sum_{t=1}^N \ln p_{k,t} \right) \mid \sum_{t=1}^N p_{k,t}[y_k(t) - \mu_k] = 0, p_{k,t} \geq 0, \sum_{t=1}^N p_{k,t} = 1 \right\} \quad (2.20)$$

which can be viewed as a *multinomial concentrated log likelihood* [64]. In this case, the measurement-error model parameters are  $\mathbf{v}_k = [p_{k,1}, p_{k,2}, \dots, p_{k,N}]^T$  for  $\beta_k = 0$  and  $\mathbf{v}_k = [\mu_k, p_{k,1}, p_{k,2}, \dots, p_{k,N}]^T$  for  $\beta_k = 1$ , where the multinomial probabilities  $p_{k,1}, p_{k,2}, \dots, p_{k,N}$  are constrained to satisfy the conditions in ( 2.18), see also ( 2.20).

Maximizing  $l_k(\mu_k)$  with respect to  $\mu_k$  yields

$$\max_{\mu_k} [l_k(\mu_k)] = -N \ln N \quad (2.21)$$

<sup>6</sup>Note that the optimization problem in ( 2.17) does not have a solution if  $\mu_k < y_{k,\text{MIN}}$  or  $\mu_k > y_{k,\text{MAX}}$ . In such cases, we set  $l_k \mu_k = -\infty$  by convention. If  $y_{k,\text{MIN}} = y_{k,\text{MAX}} = \mu_k$ , we take  $l_k(\mu_k) = -N \ln N$  and if  $\mu_k = y_{k,\text{MIN}} < y_{k,\text{MAX}}$  or  $\mu_k = y_{k,\text{MAX}} > y_{k,\text{MIN}}$ , we set  $l_k(\mu_k) = -\infty$ .

<sup>7</sup>See also [62, 64] for the definition and properties of the empirical likelihood.

which follows by noting that, subject to  $\sum_{t=1}^N p_{k,t} = 1$ , the log-likelihood function  $\sum_{t=1}^N \ln p_{k,t}$  is maximized by choosing the discrete uniform distribution of the observations (i.e.  $p_{k,t} = 1/N, t = 1, 2, \dots, N$ ). This choice yields the NPMLE of  $\mu_k$ :

$$\bar{y}_k = \frac{1}{N} \cdot \sum_{t=1}^N y_k(t) = \operatorname{argmax}_{\mu_k} l_k(\mu_k) \quad (2.22)$$

also known as the *bootstrap estimate* of  $\mu_k$  [64]. In Section 2.8, we compute ( 2.20) by solving a *one-dimensional convex dual problem*:

$$l_k \mu_k = -N \ln N + \min_{\lambda_k} \Xi_k(\lambda_k; \mu_k) \quad (2.23)$$

where

$$\Xi_k(\lambda_k; \mu_k) = - \sum_{t=1}^N \ln \{1 + \lambda_k [y_k(t) - \mu_k]\} \quad (2.24)$$

is a convex function of  $\lambda_k$ . To ensure that the estimates of the multinomial probabilities remain in the allowed parameter space, the search for  $\lambda_k$  that minimizes ( 2.24) should be constrained to the interval (see Section 2.8):

$$\frac{1 - N^{-1}}{\mu_k - y_{k,\text{MAX}}} < \lambda_k < \frac{1 - N^{-1}}{\mu_k - y_{k,\text{MIN}}} \quad (2.25)$$

and can be efficiently performed using the damped Newton-Raphson iteration<sup>8</sup>:

$$\lambda_k^{(i+1)} = \lambda_k^{(i)} + \delta_k^{(i)} \cdot \left( \sum_{t=1}^N \frac{[y_k(t) - \mu_k]^2}{\{1 + \lambda_k^{(i)} [y_k(t) - \mu_k]\}^2} \right)^{-1} \cdot \sum_{t=1}^N \frac{y_k(t) - \mu_k}{1 + \lambda_k^{(i)} [y_k(t) - \mu_k]} \quad (2.26)$$

where the *damping factor*  $0 < \delta_k^{(i)} \leq 1$  is chosen (at every step  $i$ ) to ensure that ( 2.24) decreases and  $\lambda_k^{(i+1)}$  remains within the interval specified in ( 2.25)<sup>9</sup>. The above iteration converges to the unique solution  $\lambda_k = \lambda_k(\mu_k)$ .

In Section 2.8, we sketch a proof that the empirical=likelihood approach employs a *least favorable nonparametric distribution family* for estimating  $\mu_k$  and derive the CRB for estimating  $\mu_k$  under the empirical likelihood measurement-error model. Assuming the discrete uniform

<sup>8</sup>See e.g. [65] for an introduction to the Newton-Raphson algorithms. To simplify the notation in ( 2.26) and later in ( 2.32), we omit the dependence of  $\lambda_k^{(i)}$  and  $\delta_k^{(i)}$  on  $\mu_k$ .

<sup>9</sup>In particular, we start with  $\delta_k^{(i)} = 1$  and check if ( 2.24) decreases and  $\lambda_k^{(i+1)}$  remains within the interval ( 2.25). If these tests fail, we keep halving  $\delta_k^{(i+1)}$  until they are satisfied.

distribution of the observations, this CRB simplifies to:

$$-\left[\sum_{t=1}^N \frac{1}{N} \cdot \frac{d^2 l_k(\bar{y}_k)}{d\bar{y}_k^2}\right]^{-1} = \frac{s_k^2}{N} \quad (2.27)$$

where

$$\begin{aligned} s_k^2 &= s_{0,k}^2 - \bar{y}_k^2 \\ s_{0,k}^2 &= \frac{1}{N} \cdot \sum_{t=1}^N y_k^2(t) \end{aligned} \quad (2.28)$$

and  $\bar{y}_k$  has been defined in ( 2.22).

**Empirical entropy:** For  $\kappa = -1$ , ( 2.17) reduces to

$$\min_{p_{k,t}, t=1,2,\dots,N} \sum_{t=1}^N p_{k,t} \ln(N p_{k,t}) \quad (2.29)$$

subject to the constraints in ( 2.18). In ( 2.29), we minimize the *relative entropy*<sup>10</sup> between the multinomial distribution defined by the probabilities  $p_{k,t}, t = 1, 2, \dots, N$  and the discrete uniform distribution on  $t = 1, 2, \dots, N$ , yielding the *empirical entropy* estimates  $p_{k,t}(\mu_k), t = 1, 2, \dots, N$  of the multinomial probabilities. It can be shown that  $p_{k,t}(\mu_k)$  have the following form (see Section 2.9):

$$p_{k,t}(\mu_k) = \frac{\exp[\lambda_k(\mu_k) y_k(t)]}{\sum_{\tau=1}^N \exp[\lambda_k(\mu_k) y_k(\tau)]}, t = 1, 2, \dots, N \quad (2.30)$$

where  $\lambda_k(\mu_k)$  is obtained by minimizing

$$\zeta_k(\lambda_k; \mu_k) = \sum_{t=1}^N \exp\{\lambda_k[y_k(t) - \mu_k]\} \quad (2.31)$$

with respect to  $\lambda_k$ . Note that  $\zeta_k(\lambda_k; \mu_k)$  is a convex function of  $\lambda_k$  and can be efficiently minimized using the damped Newton-Raphson iteration:

$$\lambda_k^{(i+1)} = \lambda_k^{(i)} - \delta_k^{(i)} \cdot \left\{ \sum_{t=1}^N \exp[\lambda_k^{(i)} y_k(t)] \cdot [y_k(t) - \mu_k]^2 \right\}^{-1} \cdot \sum_{t=1}^N \exp[\lambda_k^{(i)} y_k(t)] \cdot [y_k(t) - \mu_k] \quad (2.32)$$

Here, the damping factor  $0 < \delta_k^{(i)} \leq 1$  is chosen to ensure that ( 2.31) decreases. Finally, we compute the nonparametric log-likelihood function of  $\mu_k$  by substituting ( 2.30) into ( 2.16):

$$l_k(\mu_k) = N \cdot \lambda_k(\mu_k) \bar{y}_k - N \cdot \ln\left\{ \sum_{t=1}^N \exp[\lambda_k(\mu_k) y_k(t)] \right\} \quad (2.33)$$

<sup>10</sup>Relative entropy, also known as Kullback-Leibler distance, is defined in e.g. [66].

The above empirical-entropy approach is closely related to the *nonparametric tilting* in [67, 68]. It is also known as the *empirical exponential family likelihood* [69] because it can be derived by constraining the probability distribution of  $y_k(1), y_k(2), \dots, y_k(N)$  to belong to the exponential family of distributions.

In [67], Efron presented the CRB for  $\mu_k$  under the empirical entropy measurement-error model and used it to argue that the empirical-entropy approach employs a least favorable family for estimating  $\mu_k$ . Assuming the discrete uniform distribution of the observations, the expression for this CRB simplifies to (2.27), see Section 2.9 and [67].

### 2.4.2 Autologistic MRF process model

Assume that each node  $k$  makes a binary decision about its current status, i.e. it decides between the hypothesis

$H_{0,k}$ : (signal absent,  $\beta_k = 0$ ), corresponding to  $\mu_k = 0$

versus the one-sided alternative

$H_{1,k}$ : (signal present,  $\beta_k = 1$ ), corresponding to  $\mu_k > 0$ .

This formulation is suitable for detecting event regions with elevated concentrations of chemicals, see the example at the beginning of Section 2.4. In this example, we restrict the parameter space of the mean signal  $\mu_k$  to the set of non-negative real numbers. To describe the binary MRF for event-region detection problems, we adopt the *autologistic MRF process model* specified by the conditional pmfs (see [26]):

$$p_{\beta_k|N_{\beta}(k)}(\beta_k|N_{\beta}(k)) = \frac{\exp(a_k\beta_k + \beta_k \cdot \sum_{l \in N(k)} c_{k,l}\beta_l)}{1 + \exp(a_k + \sum_{l \in N(k)} c_{k,l}\beta_l)} \quad (2.34)$$

for  $k = 1, 2, \dots, K$ , where  $a_k$  and  $c_{k,l}$  are *spatial-trend* and *spatial-dependence* MRF model parameters. Furthermore, we utilize the following simple spatial trend and dependence models:

- *constant spatial trend* (independent of  $k$ ):

$$a_k = a \quad (2.35)$$

- *homogeneous spatial dependence* with equal evidence from each neighbor:

$$c_{k,l} = \begin{cases} \eta, & \|\mathbf{r}_k - \mathbf{r}_l\| \leq d \\ 0, & \|\mathbf{r}_k - \mathbf{r}_l\| > d \end{cases} \quad (2.36)$$

where  $d$  is the cutoff distance, see also Section 2.2.

In event-region detection problems,  $\eta$  is a positive constant describing the *field strength*. This spatial-dependence model quantifies the notion that the random-field values at the nodes that are close (in terms of the spatial distance) should be more similar than the values at the nodes that are far apart. More complex spatial dependence models can be developed along the lines of [26] (for isotropic dependence) and [70] (for anisotropic dependence).

In applications where the cutoff distance  $d$  is approximately equal to the radio-transmission range of the sensor elements, the neighborhood  $\mathcal{N}(k)$  consists of those nodes with which  $k$  can communicate directly. Then, we can determine the neighborhoods without utilizing the node location information. However, the assumption that the cutoff distance coincides with the communication range may be impractical. In addition, the effective cutoff distance may vary slightly from one node to another.

In the following, we specialize the general ICM algorithm in Section 2.3 to the event-region detection problem using the measurement-error model in Section 2.4.1 and process model in (2.34).

### 2.4.3 ICM detection for Gaussian measurement-error model

We first define the indicator function

$$i_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise} \end{cases} \quad (2.37)$$

Under the Gaussian measurement-error and autologistic process models, Step (ICM2) in the

ICM algorithm simplifies to selecting  $\beta_k = 1$  if

$$\begin{aligned} L_k(1|N_\beta(k)) - L_k(0|N_\beta(k)) &= \frac{2}{N} \cdot \ln\left(\frac{s_{0,k}^2}{s_k^2}\right) \cdot i_{[0,\infty)}(\bar{y}_k) + a_k + \sum_{l \in \mathcal{N}(k)} c_{k,l} \beta_l \\ &= \frac{2}{N} \cdot \ln\left(\frac{s_{0,k}^2}{s_k^2}\right) \cdot i_{[0,\infty)}(\bar{y}_k) + a + \eta u_k \\ &\geq 0 \end{aligned} \quad (2.38)$$

and selecting  $\beta_k = 0$  otherwise; see Section 2.10 for details of the derivation. Here,

$$u_k = \sum_{l \in \mathcal{N}(k)} \beta_l \quad (2.39)$$

is the number of neighbors of  $k$  reporting the presence of signal and  $\mathcal{N}(k)$ ,  $\bar{y}_k$ ,  $s_k^2$ ,  $s_{0,k}^2$  have been defined in ( 2.6), ( 2.22), ( 2.28) and ( 2.39). The first term in ( 2.38) is the onesided  $t$ -test statistic for the mean  $\mu_k$  (based on the "local" measurements collected at node  $k$ ), whereas the second and third terms account for the spatial trend and spatial dependence effects introduced by the MRF model.

1) *Initialization:* To obtain initial decisions at each node  $k$ , we ignore the neighborhood dependence and apply the *local* one-sided  $t$  test for the mean  $\mu_k$ : select  $\beta_k = 1$  if

$$\frac{s_{0,k}^2}{s_k^2} \cdot i_{[0,\infty)}(\bar{y}_k) \geq \tau_G \quad (2.40)$$

and select  $\beta_k = 0$  otherwise. This test is also the GLR test for the hypothesis testing problem described in Section 2.4.2. Denote by  $B(0.5(N-1), 0.5)$  the *centralized beta distribution* with parameters  $0.5(N-1)$  and  $0.5$ . We select the threshold

$$\tau_G = b_{0.5(N-1), 0.5, 2P_{FA}}^{-1} \quad (2.41)$$

to guarantee a specified probability of false alarm  $P_{FA}$ . Here,  $b_{0.5(N-1), 0.5, p}$  is defined using

$$P[\beta \leq b_{0.5(N-1), 0.5, p}] = p \quad (2.42)$$

where  $\beta$  is a  $B(0.5(N-1), 0.5)$  random variable.

#### 2.4.4 ICM detection for nonparametric measurement-error models

Under the nonparametric measurement-error models in Section 2.4.1, the condition ( 2.19) implies that one-sided detection in Section 2.4.2 will be meaningful only if

$$y_{k,\text{MAX}} \geq 0 \quad (2.43)$$

with equality implying  $y_{k,\text{MIN}} = y_{k,\text{MAX}} = 0$ . For the empirical likelihood and entropy measurement-error models, Step (ICM2) simplifies to selecting  $\beta_k = 1$  if

$$\begin{aligned} & L_k(1|\beta_l \in N_\beta(k)) - L_k(0|\beta_l \in N_\beta(k)) \\ &= \max_{\mu_k > 0} [l_k(\mu_k)] = l_k(0) + \ln p_{\beta_k|N_\beta(k)}(1|N_\beta(k)) - \ln p_{\beta_k|N_\beta(k)}(0|N_\beta(k)) \\ &= [-N \ln N - l_k(0)] \cdot i_{[0,\infty)}(\bar{y}_k) + a + \eta u_k \\ &\geq 0 \end{aligned} \quad (2.44)$$

and selecting  $\beta_k = 0$  otherwise, see Section 2.11. Here, the nonparametric log likelihoods  $l_k(0)$  for the empirical likelihood and entropy models are computed using ( 2.24) and ( 2.33), respectively.

1) *initialization*: We now discuss the initialization of the ICM iteration under the empirical likelihood and entropy measurement-error models. We propose the following local GLR tests that ignore the neighborhood dependence: select  $\beta_k = 1$  if

$$\sqrt{2[-N \ln N - l_k(0)]} \cdot i_{[0,\infty)}(\bar{y}_k) \geq \tau_{\text{NP}} \quad (2.45)$$

and select  $\beta_k = 0$  otherwise. The threshold  $\tau_{\text{NP}}$  which guarantees a specified probability of false alarm  $P_{\text{FA}}$  can be approximately computed by solving (see Section 2.12):

$$\Phi(\tau_{\text{NP}}) = 1 - P_{\text{FA}} \quad (2.46)$$

where  $\Phi(\cdot)$  denotes the cumulative distribution function of the standard normal random variable. The above approximation is based on the Wilks' theorem for the empirical likelihood [62, 64] and similar results for empirical entropy [71, 72], derived under the assumption that  $N \rightarrow \infty$ , see also Section 2.12. Therefore, its accuracy improves as the number of observations per sensor increases.

In the above ICM algorithms, the nodes exchange binary messages ( $\beta_k = 0$  or  $\beta_k = 1$ ) to inform neighbors about their status; the communication cost of this exchange is low, which is important in most practical applications that require energy and bandwidth efficiency [3].

## 2.5 MRF calibration

Assume that *training data* is available containing both the observations  $\mathbf{y}_k$ ,  $k = 1, 2, \dots, K$  and the true values of the MRF  $\beta$ . we develop a calibration method for estimating the *MRF model parameters* from the training data. Denote the vector of MRF model parameters by  $\omega$ . To emphasize the dependence of the local and global predictive log-likelihood functions in ( 2.11) and ( 2.12) on  $\omega$ , we use  $L_k(\beta_k|N_\beta(k); \omega)$  and  $L(\beta; \omega)$  to denote these functions throughout this section. Similarly, we use  $p_{\beta|N_\beta(k)}(\beta|N_\beta(k); \omega)$  to denote the conditional pdfs in ( 2.9).

We denote

$$\frac{\exp[L(\beta; \omega)]}{\int \exp[L(\mathbf{b}; \omega)] d\mathbf{b}} \text{ or } \frac{\exp[L(\beta; \omega)]}{\sum_{\mathbf{b}} \exp[L(\mathbf{b}; \omega)]} \quad (2.47)$$

as the "predictive" pdfs or pmfs of  $\beta$ , see [47]. Then we may compute maximum "predictive" likelihood estimates of  $\omega$  by maximizing the expressions in ( 2.47). However, the denominators in (2.47) are usually computationally intractable. Motivated by Besag's pseudolikelihood approach in [26] and [73], we construct a computationally tractable log pseudo predictive likelihood function:

$$L_{\text{PPL}}(\omega) = \sum_{k=1}^K \ln \left\{ \frac{\exp[L_k(\beta_k|N_\beta(k); \omega)]}{\sum_{i=0}^1 \exp[L_k(i|N_\beta(k); \omega)]} \right\} \quad (2.48)$$

and estimate the MRF model parameters  $\omega$  by maximizing  $L_{\text{PPL}}(\omega)$  with respect to  $\omega$ . Here,

$$\frac{\exp[L_k(\beta_k|N_\beta(k); \omega)]}{\sum_{i=0}^1 \exp[L_k(i|N_\beta(k); \omega)]} \quad (2.49)$$

is the full conditional predictive pdf/pmf of  $\beta_k$ . The above calibration method applies to the general measurement-error and MRF models described in Section 2.2.

**Event-region detection:** we now specialize ( 2.48) to the event-region detection problem,

leading to

$$L_{\text{PPL}}(\boldsymbol{\omega}) = \text{const} + a \cdot \left( \sum_{k=1}^K \beta_k \right) + \eta \cdot \left( \sum_{k=1}^K \beta_k u_k \right) - \sum_{k=1}^K \ln \{ 1 + \exp [ L_k(1 | N_\beta(k); \boldsymbol{\omega}) - L_k(0 | N_\beta(k); \boldsymbol{\omega}) ] \} \quad (2.50)$$

where const denotes terms that do not depend on the MRF model parameters  $\boldsymbol{\omega}$ . Here, ( 2.50) follows by substituting the autologistic MRF model ( 2.34) into ( 2.48) and neglecting constant terms. Under the Gaussian and nonparametric measurement-error models in Section 2.4.1, the expression  $L_k(\beta_k = 1 | \beta_l \in N_\beta(k)) - L_k(\beta_k = 0 | \beta_l \in N_\beta(k))$  in ( 2.50) simplify to ( 2.38) and ( 2.44), respectively. To efficiently compute the last term in ( 2.50), we utilize the following approximation: for large positive  $x$ ,

$$\ln[1 + \exp(x)] \approx x \quad (2.51)$$

Interestingly, setting the data-dependent log-likelihood terms in ( 2.11) to zero and substituting the resulting expressions into ( 2.48) yields the Besag's *pseudo log-likelihood function*

$$L_{\text{PL}} = \sum_{k=1}^K \ln p_{\beta_k | N_\beta(k)}(\beta_k | N_\beta(k); \boldsymbol{\omega}) \quad (2.52)$$

for estimating the MRF model parameters, see [26, 73]. Note that ( 2.52) utilizes only the MRF  $\boldsymbol{\beta}$  and does not depend on the measurements  $\mathbf{y}_k, k = 1, 2, \dots, K$ . Maximizing the pseudo log likelihood ( 2.52) would yield reasonable estimates of the MRF parameters if the measurement-error model parameters  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K$  were *known* in the ICM estimation/detection stage. Note, however, that we assume that  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K$  are *unknown* and estimate them *locally* at each node, which is taken into account by the PPL calibration method in ( 2.48).

## 2.6 Numerical examples

To assess the performance of the proposed event-region detection methods, we consider sensor networks containing  $K = 1000$  nodes randomly (uniformly) distributed on a  $50\text{m} \times 50\text{m}$  grid with 1m spacing between the potential node locations. We assume that each sensor  $k$  collects  $N = 5$  measurements corrupted by i.i.d. zero-mean additive noise, unless specified

otherwise (see e.g. Section 2.6.3). The noiseless field containing two event regions is shown in Figure 2.2 (left) and the sensor locations (with corresponding ideal decisions) are shown in Figure 2.2 (right). Here, the filled circles correspond to the nodes in the event regions. The noisy measurements collected at the nodes outside the event region have zero means.

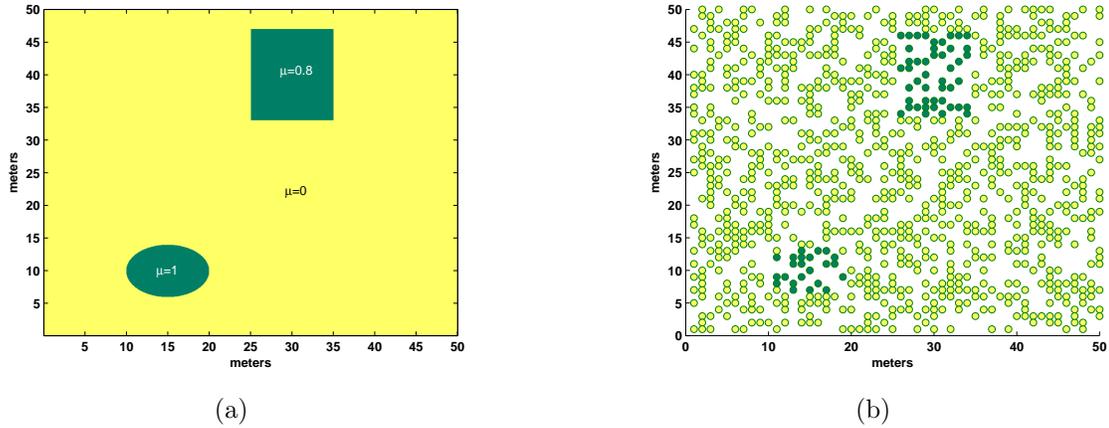


Figure 2.2 (Left) Noiseless field and (right) a sensor network with  $K = 1000$  nodes.

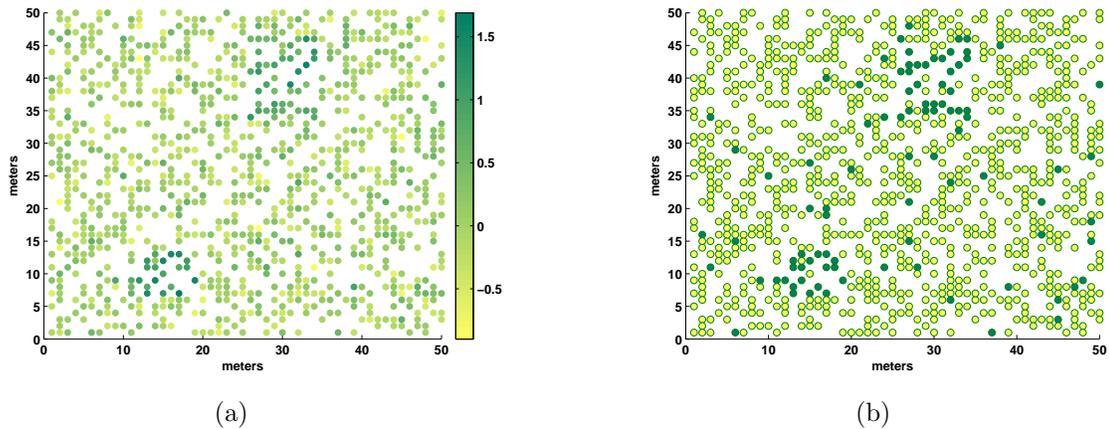


Figure 2.3 Gaussian measurement scenario: (Left) averaged observations  $\bar{y}_k, k = 1, 2, \dots, K$  as functions of the node locations and (right) one-sided  $t$ -test results for  $P_{FA} = 5\%$ .

Throughout this section, we set the cutoff distance to  $d = 3\text{m}$  and define neighborhoods according to (2.6). In all simulation examples, we estimated the MRF model parameters  $a$

(spatial trend) and  $\eta$  (field strength) using the calibration procedure in Section 2.5, where the calibration field and other details of our implementation are given in Section 2.6.4.

### 2.6.1 Gaussian measurement scenario

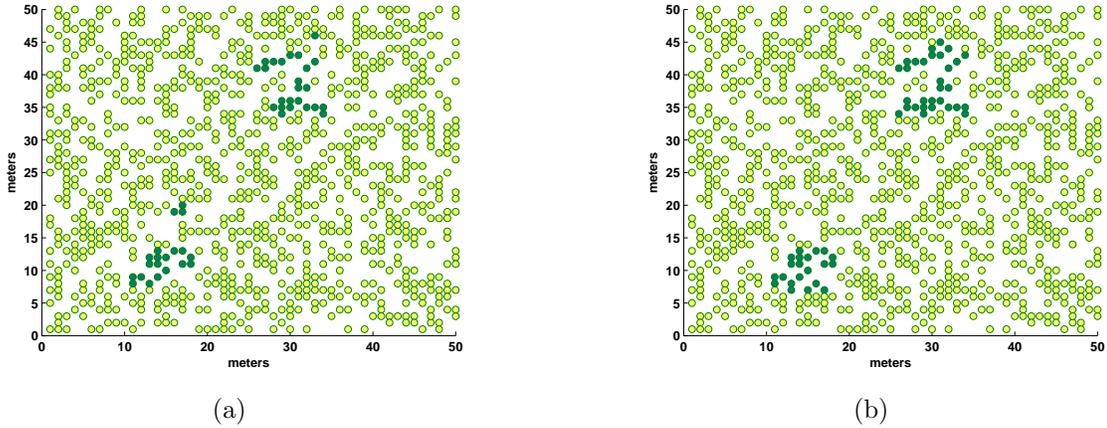


Figure 2.4 Gaussian measurement scenario: Event-region detection results after (left) one cycle and (right) two cycles of the Gaussian ICM algorithm.

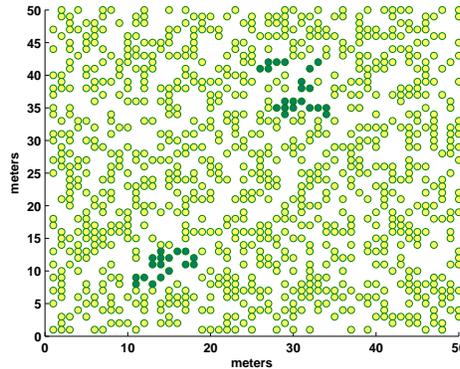


Figure 2.5 Gaussian measurement scenario: Event-region detection results upon convergence of the Gaussian ICM algorithm.

In the first set of simulations, we generated the simulated data using the Gaussian measurement-error model in Section 2.4.1 with constant noise variance  $\sigma_k^2 = 0.5$  for all  $k = 1, 2, \dots, K$ . In Figure 2.3 (left), we show the averaged observations  $\bar{y}_k$ ,  $k = 1, 2, \dots, K$  in ( 2.22) as functions

of the node locations for one realization of the noisy field. Applying the one-sided  $t$ -test in ( 2.42) yields the results in Figure 2.3 (right), where the threshold  $\tau_G$  was chosen to satisfy the false-alarm probability  $P_{FA} = 5\%$ . The filled circles correspond to the nodes declaring the presence of signal whereas hollow circles correspond to the nodes declaring the signal absence. The  $t$ -test decisions were used to initialize the Gaussian ICM detector (described in Section 2.4.3, see also Section 2.3; the decisions after one and two ICM cycles are shown in Figure 2.4. In this example, all isolated nodes reporting the presence of signal were correctly recognized as false alarms already after two ICM cycles. The Gaussian ICM algorithm converged in four cycles yielding the results in Figure 2.5.

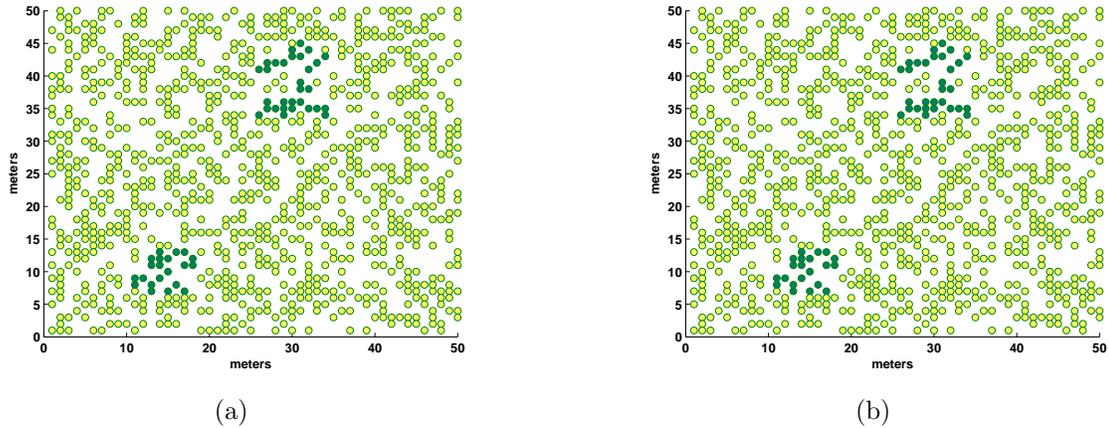


Figure 2.6 Gaussian measurement scenario: Event-region detection results for (left) the empirical likelihood and (right) empirical entropy nonparametric ICM algorithms.

Applying the nonparametric ICM detectors in Section 2.4.4 yields (upon convergence) the results in Figure 2.6. These detectors were initialized using the local GLR tests in ( 2.45) with the threshold  $\tau_{NP}$  chosen to (approximately) satisfy the false-alarm probability  $P_{FA} = 5\%$ . Both the empirical likelihood and entropy based ICM algorithms converged in four cycles and were successful in removing the false alarms.

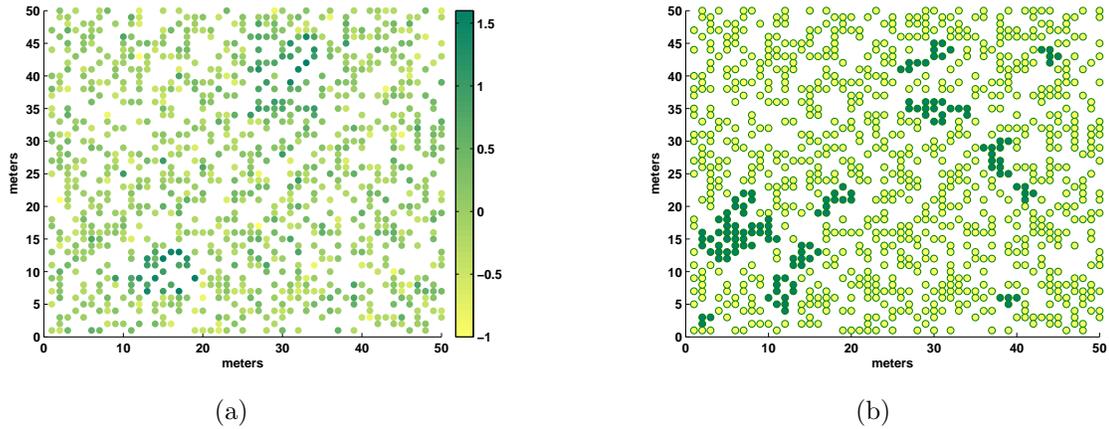


Figure 2.7 Quantized Gaussian measurement scenario: (Left) averaged observations  $\bar{y}_k, k = 1, 2, \dots, K$  as functions of the node locations and (right) event-region detection results for the Gaussian ICM algorithm.

## 2.6.2 Quantized Gaussian measurement scenario

We now study the performance of the proposed methods in the case where the Gaussian observations [generated as described in Section 2.6.1] have been coarsely quantized, leading to non-Gaussian measurements from a discrete probability distribution. Here, we quantized the measurements to the closest integer values in the interval  $[-2, 3]$ . In Figure 2.7 (left), we show the averages  $\bar{y}_k, k = 1, 2, \dots, K$  of the quantized observations [see ( 2.22)] as functions of the node locations. Applying the ICM detectors for Gaussian and nonparametric measurement-error models to the quantized measurements yields the results in Figure 2.7 (right) and Figure 2.8, respectively. The ICM algorithms were initialized using the local GLR tests in ( 2.42) and ( 2.45) with the  $\tau_G$  and  $\tau_{NP}$  chosen using ( 2.41) and ( 2.46) to satisfy the false-alarm probability  $P_{FA} = 5\%$ . The Gaussian ICM algorithm performs poorly under this scenario due to the mismatch between the quantized observations and assumed Gaussian measurement-error model, see Figure 2.7 (right). The empirical likelihood and empirical entropy based ICM methods estimated the unknown probability distributions of the measurements and successfully removed the false alarms, see Figure 2.8. Unlike the Gaussian and empirical likelihood approaches the empirical entropy method provides a connected estimated of the event region

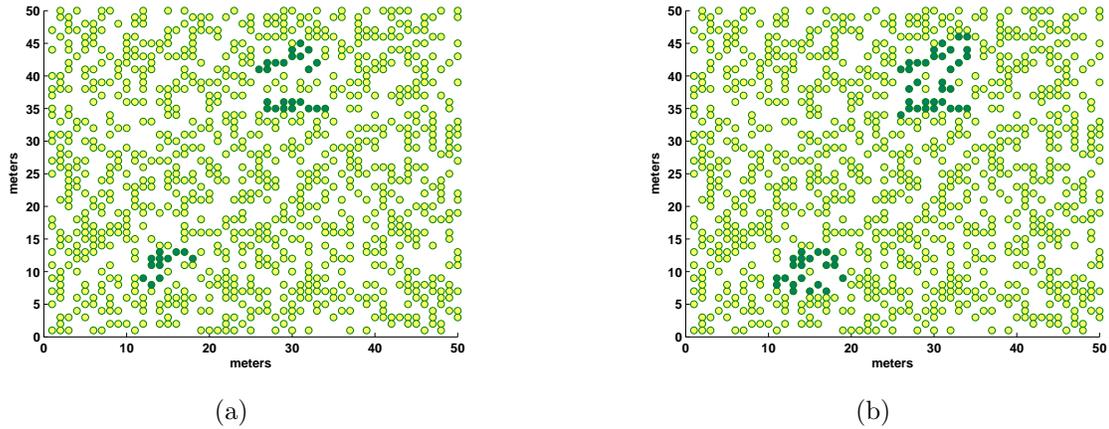


Figure 2.8 Quantized Gaussian measurement scenario: Event-region detection results for (left) the empirical likelihood and (right) empirical entropy nonparametric ICM algorithms.

in the upper right corner of the network.

### 2.6.3 Probabilities of false alarm and miss

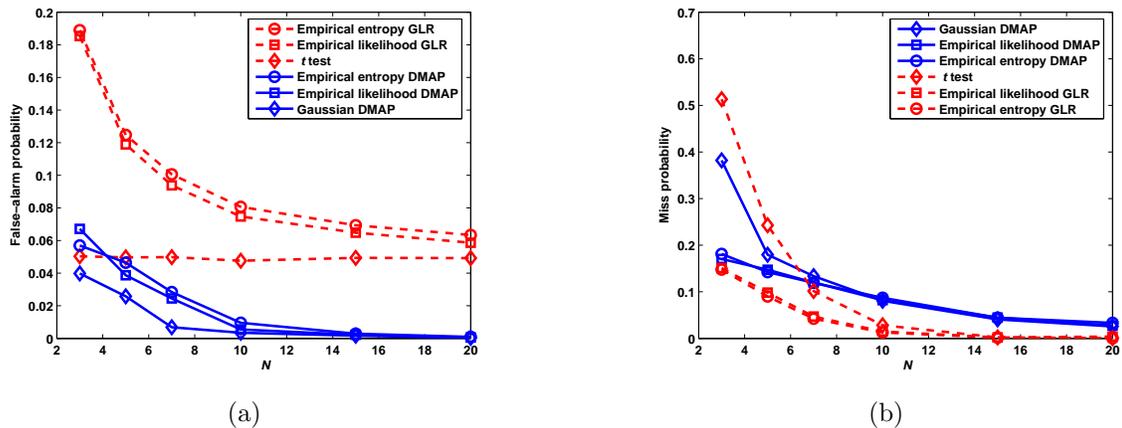


Figure 2.9 Gaussian measurement scenario: Average probabilities of (left) false alarm and (right) miss, as functions of the number of observations per sensor  $N$ .

We analyze the *average error performances* of the GLR and ICM methods under the Gaussian and quantized Gaussian measurement scenarios. Our performance metrics are the average

probabilities of *false alarm* and *miss*, calculated using 100 independent trials<sup>11</sup> where averaging has been performed over the noisy random-field realizations, random node locations and scheduling (in the ICM methods).

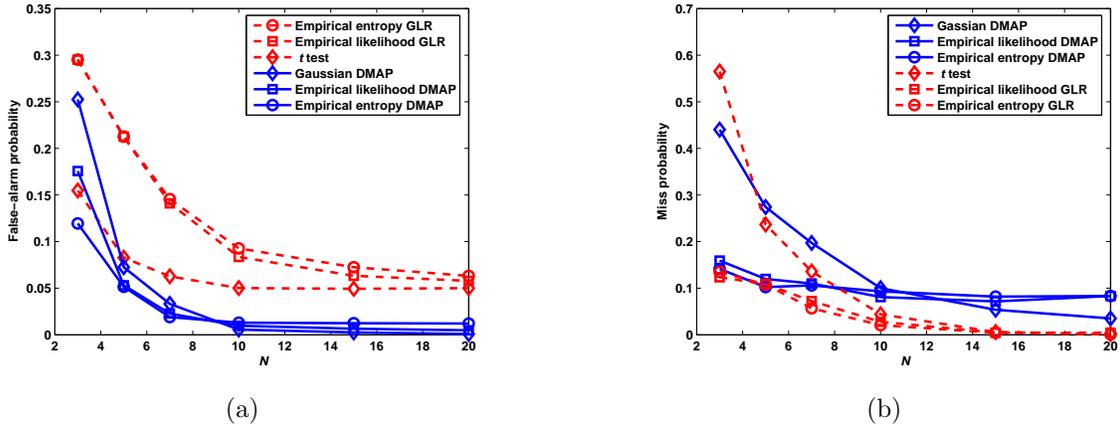


Figure 2.10 Quantized Gaussian measurement scenario: Average probabilities of (left) false alarm and (right) miss, as functions of  $N$ .

We first consider the Gaussian measurement scenario and present the average probabilities of false alarm and miss for different methods as functions of the number of observations per sensor  $N$ , see Figure 2.9. In this case,

- the average false-alarm and miss error performances of all ICM methods improve as  $N$  increases;
- the average false-alarm probability of the one-sided  $t$ -test is *constant* and equal to the specified value of 5%, verifying the validity of ( 2.41) and ( 2.42);
- the false-alarm probabilities of the local nonparametric GLR tests attain the specified level of 5% *asymptotically* (i.e. for large  $N$ , see also Section 2.4.4);
- the Gaussian ICM method achieves the smallest false-alarm probability for all  $N$  (compared with the other methods).

<sup>11</sup>Here, the two error probabilities were estimated using the ideal decisions in Figure 2.2 (right).

Consider now the quantized Gaussian measurement scenario. In Figure 2.10, we show the average probabilities of false alarm and miss for different methods as functions of  $N$ . Observe that

- as in Gaussian scenario, the average false-alarm and miss error probabilities of all ICM methods decrease with  $N$ ;
- the average false-alarm probabilities of local  $t$  and nonparametric GLR tests attain the specified level of 5% for large  $N$ ;
- for small  $N$ , the nonparametric ICM methods achieve smaller average false-alarm and miss error probabilities than the Gaussian ICM method;
- due to the averaging effect, the Gaussian ICM method performs well when  $N$  is large.

Note that the error-probability results presented in Figure 2.9 and Figure 2.10 do not show if the obtained event-region estimates were connected or not, which may be of interest in practical applications.

#### 2.6.4 MRF calibration

We utilize the calibration method in Section 2.5 to estimate the MRF model parameters  $a$  and  $\eta$ . The training data were generated by randomly placing  $K = 1000$  nodes on a  $50\text{m} \times 50\text{m}$  grid and simulating noisy realizations of a calibration field having constant mean  $\mu > 0$  within a circular event region with radius 8m, see Figure 2.11. Twenty training data sets were generated by varying the noise realizations, node locations, and values of the event-region mean  $\mu$ . We applied the calibration method proposed in Section 2.5 to fit each training data set and then averaged the obtained estimates, yielding the final calibration results. To obtain the average error probabilities in Figure 2.9 and Figure 2.10, the values of  $\mu$  in the twenty training data sets were generated by sampling from uniform  $(0.4, 1.4)$  distribution. To calibrate the ICM algorithms whose results are shown in Figure 2.4- 2.8, we sampled  $\mu$  from a wider range of values [following the uniform  $(0.4, 3.4)$  distribution]; the resulting calibration provided smaller

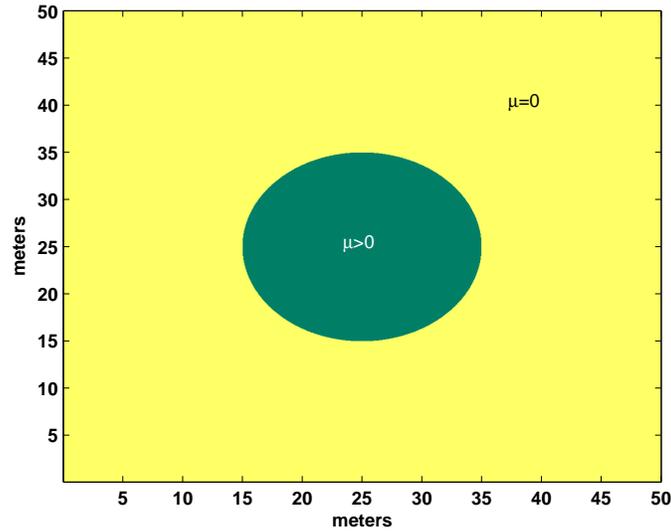


Figure 2.11 Noiseless field used for calibration.

false-alarm probabilities and larger miss probabilities [compared with the results obtained by sampling  $\mu$  from uniform  $(0.4, 1.4)$ ].

## 2.7 Concluding remarks

We presented an HMRF framework for distributed localized estimation and detection in sensor-network environments. We developed a calibration method for estimating the MRF model parameters from the training data and discussed initialization of the proposed algorithms. The proposed framework was applied to event-region detection.

Further research will include: extending the HMRF framework and ICM method to allow tracking of the field changes over time, analyzing the impact of communication errors (among the nodes) on the performance of the ICM method, comparing the ICM and message passing approaches, relaxing the conditional independence assumption in (2.4), developing data aggregation algorithms and energy-aware sensor-network design strategies for HMRFs (e.g., deciding which nodes will be in "alert" or "sleeping" modes), and studying asymptotic properties of the proposed methods as the number of measurements per node grows.

It is also of interest to relate the proposed ICM and *distributed consensus* approaches

recently proposed in [54, 55, 74]. If we select a Gaussian MRF model structure and modify the ICM iteration by replacing the measurements  $y_k$  with the estimates of the hidden field  $\beta_k$  from the previous ICM cycle, the resulting algorithm closely resembles the average-consensus scheme in [74]. Note that the consensus methods estimate *global phenomena* (e.g., the mean field) whereas the ICM methods estimate *localized features*, which is an important distinction between the two approaches.

Since the autologistic MRF model may be too simplistic for many applications, it is important to develop more general process models that will allow utilizing multiple information bits to describe the hidden field of interest. Here, it is of particular interest to derive physically based process models and corresponding ICM methods.

## 2.8 Appendix A: Empirical likelihood and CRB for estimating $\mu_k$

We derive the concentrated empirical log-likelihood expression in ( 2.24). This derivation is similar to that in [62] and [64] and is given here for completeness. We utilize the method of Lagrange multipliers to solve the constrained optimization problem in ( 2.20): Define

$$G_k = \left( \sum_{t=1}^N \ln p_{k,t} \right) + \gamma_k \cdot \left( \sum_{t=1}^N p_{k,t} - 1 \right) - N \lambda_k \cdot \sum_{t=1}^N p_{k,t} [y_k(t) - \mu_k] \quad (2.53)$$

where  $\gamma_k$  and  $\lambda_k$  are Lagrange multipliers. Forming a weighted sum of the partial derivatives of  $G_k$  with respect to  $p_{k,t}$  and setting the result to zero yields

$$0 = \sum_{t=1}^N p_{k,t} \frac{\partial G_k}{\partial p_{k,t}} = N + \gamma_k \quad (2.54)$$

where the second equality follows by using the constraints  $\sum_{t=1}^N p_{k,t} = 1$  and  $\sum_{t=1}^N p_{k,t} [y_k(t) - \mu_k] = 0$ . Therefore  $\gamma_k = -N$  implying that

$$p_{k,t} = \frac{1}{N} \cdot \frac{1}{1 + \lambda_k [y_k(t) - \mu_k]} \quad (2.55)$$

where  $\lambda_k = \lambda_k(\mu_k)$  is chosen as a solution to

$$\sum_{t=1}^N p_{k,t} [y_k(t) - \mu_k] = \frac{1}{N} \cdot \sum_{t=1}^N \frac{y_k(t) - \mu_k}{1 + \lambda_k [y_k(t) - \mu_k]} = 0 \quad (2.56)$$

Substituting ( 2.55) into the multinomial log likelihood yields

$$\sum_{t=1}^N \ln p_{k,t} = -N \ln N + \Xi_k(\lambda_k; \mu_k) \quad (2.57)$$

where  $\Xi_k(\lambda_k; \mu_k)$  was defined in ( 2.24). To satisfy ( 2.56), we need to minimize the above expression with respect to  $\lambda_k$ , yielding the complex dual formulation in ( 2.24). Assuming ( 2.19), all estimates of the multinomial probabilities need to satisfy

$$0 < p_{k,t} = \frac{1}{N} \cdot \frac{1}{1 + \lambda_k [y_k(t) - \mu_k]} < 1 \quad (2.58)$$

and ( 2.25) is obtained by using the second inequality in ( 2.58) for all  $t \in \{1, 2, \dots, N\}$ .

Finally, the first two derivative of  $\Xi_k(\lambda_k; \mu_k)$  with respect to  $\lambda_k$  are

$$\begin{aligned} \frac{\partial \Xi_k(\lambda_k; \mu_k)}{\partial \lambda_k} &= - \sum_{t=1}^N \frac{y_k(t) - \mu_k}{1 + \lambda_k [y_k(t) - \mu_k]} \\ \frac{\partial^2 \Xi_k(\lambda_k; \mu_k)}{\partial \lambda_k^2} &= - \sum_{t=1}^N \frac{[y_k(t) - \mu_k]^2}{\{1 + \lambda_k [y_k(t) - \mu_k]\}^2} \end{aligned} \quad (2.59)$$

and the Newton-Raphson iteration ( 2.26) follows.

#### **Least favorable families and CRB for $\mu_k$ under the empirical likelihood model:**

We derive the CRB for  $\mu_k$  under the empirical likelihood measurement-error model and sketch a proof that the empirical-likelihood approach employs a *least favorable nonparametric distribution family* for estimating  $\mu_k$ .

We first differentiate the empirical log likelihood in ( 2.24) with respect to  $\mu_k$ :

$$\frac{dl_k(\mu_k)}{d\mu_k} = N \lambda_k(\mu_k) \quad (2.60)$$

which follows by using ( 2.55)-( 2.56) and the constraint  $\sum_{t=1}^N p_{k,t} = 1$ . Then

$$\frac{d^2 l_k(\mu_k)}{d\mu_k^2} = N \cdot \frac{d\lambda_k(\mu_k)}{d\mu_k} \quad (2.61)$$

where  $d\lambda_k/d\mu_k$  can be computed by differentiating ( 2.56) [with  $\lambda_k$  evaluated at  $\lambda_k(\mu_k)$ ]:

$$\frac{d}{d\mu_k} \left\{ \sum_{t=1}^N \frac{y_k(t) - \mu_k}{1 + \lambda_k(\mu_k) [y_k(t) - \mu_k]} \right\} = 0 \quad (2.62)$$

leading to

$$\frac{d\lambda_k(\mu_k)}{d\mu_k} = \left\{ \sum_{t=1}^N \frac{[y_k(t) - \mu_k]^2}{\{1 + \lambda_k(\mu_k)[y_k(t) - \mu_k]\}^2} \right\}^{-1} \cdot \left\{ \lambda_k(\mu_k) \cdot \sum_{t=1}^N \frac{y_k(t) - \mu_k}{\{1 + \lambda_k(\mu_k)[y_k(t) - \mu_k]\}^2} - N \right\} \quad (2.63)$$

and consequently,

$$-\frac{d^2 l_k(\mu_k)}{d\mu_k^2} = \left\{ \sum_{t=1}^N \frac{[y_k(t) - \mu_k]^2}{\{1 + \lambda_k(\mu_k)[y_k(t) - \mu_k]\}^2} \right\}^{-1} \cdot \left\{ N - \lambda_k(\mu_k) \cdot \sum_{t=1}^N \frac{y_k(t) - \mu_k}{\{1 + \lambda_k(\mu_k)[y_k(t) - \mu_k]\}^2} \right\} \quad (2.64)$$

Then, assuming the discrete uniform distribution of the observations  $y_k(1), y_k(2), \dots, y_k(N)$ , the CRB for estimating  $\mu_k$  is given by ( 2.27), which follows from the fact that the discrete uniform distribution of the observations implies  $\mu_k = \bar{y}_k$  and  $\lambda_k(\bar{y}_k) = 0$ . Note that ( 2.27) closely resembles the well-known CRB expression for  $\mu_k$  under the *parametric* Gaussian measurement-error model in Section 2.4.1 (see e.g. [75]):

$$\text{CRB}_G = \frac{\sigma_k^2}{N} \quad (2.65)$$

In particular, ( 2.27) is a *good estimate* of. Hence the empirical likelihood approach employs a *least favorable nonparametric distribution family* for estimating  $\mu_k$ . This conclusion follows from the notion that a least favorable nonparametric family is one in which the estimation problem (i.e. estimating  $\mu_k$  in our case) is "as hard as in a parametric problem" (corresponding to the Gaussian measurement-error model in the above example), see also the discussion in [62, 63, 72, 76].

## 2.9 Appendix B: Empirical entropy and CRB for estimating $\mu_k$

We utilize Lagrange multipliers to solve the constrained optimization problem in ( 2.29) [subject to the constraints ( 2.18)]: Define

$$\begin{aligned} G_k &= \sum_{t=1}^N N p_{k,t} + \gamma_k \cdot \left( \sum_{t=1}^N p_{k,t} - 1 \right) - N \lambda_k \cdot \sum_{t=1}^N p_{k,t} [y_k(t) - \mu_k] \\ &= N \ln N + N \sum_{t=1}^N p_{k,t} \ln(p_{k,t}) + \gamma_k \cdot \left( \sum_{t=1}^N p_{k,t} - 1 \right) - N \lambda_k \sum_{t=1}^N p_{k,t} [y_k(t) - \mu_k] \end{aligned} \quad (2.66)$$

where  $\gamma_k$  and  $\lambda_k$  are Lagrange multipliers. Setting the partial derivatives of  $G_k$  with respect to  $p_{k,t}$  to zero yields

$$N + \gamma_k + N \ln(p_{k,t}) - N \lambda_k [y_k(t) - \mu_k] = 0 \quad (2.67)$$

for  $t = 1, 2, \dots, N$ . Finding  $\gamma_k$  that satisfy the constraint  $\sum_{t=1}^N p_{k,t} = 1$  leads to the following expressions for the multinomial probabilities:

$$p_{k,t} = \frac{\exp\{\lambda_k [y_k(t) - \mu_k]\}}{\sum_{\tau=1}^N \exp\{\lambda_k [y_k(\tau) - \mu_k]\}} = \frac{\exp[\lambda_k y_k(t)]}{\sum_{\tau=1}^N \exp[\lambda_k y_k(\tau)]} \quad (2.68)$$

Finally, the constraint  $\sum_{t=1}^N p_{k,t} [y_k(t) - \mu_k] = 0$  is satisfied by finding  $\lambda_k = \lambda_k(\mu_k)$  that solves

$$\sum_{t=1}^N \exp\{\lambda_k [y_k(t) - \mu_k]\} \cdot [y_k(t) - \mu_k] = 0 \quad (2.69)$$

Note that ( 2.69) is an increasing function of  $\lambda_k$  and that satisfying ( 2.69) is equivalent to minimizing  $\zeta_k(\lambda_k; \mu_k)$  in ( 2.31) with respect to  $\lambda_k$ . Finally, the first two derivations of  $\zeta_k(\lambda_k; \mu_k)$  with respect to  $\lambda_k$  are

$$\begin{aligned} \frac{\partial \zeta_k(\lambda_k; \mu_k)}{\partial \lambda_k} &= \sum_{t=1}^N \exp\{\lambda_k [y_k(t) - \mu_k]\} \cdot [y_k(t) - \mu_k] \\ \frac{\partial^2 \zeta_k(\lambda_k; \mu_k)}{\partial \lambda_k^2} &= \sum_{t=1}^N \exp\{\lambda_k [y_k(t) - \mu_k]\} \cdot [y_k(t) - \mu_k]^2 \end{aligned} \quad (2.70)$$

and the Newton-Raphson iteration ( 2.32) follows.

#### **Least favorable families and CRB for $\mu_k$ under the empirical entropy model:**

We derive the CRB for  $\mu_k$  under the empirical entropy measurement-error model and sketch a proof that the empirical-entropy approach employs a *least favorable nonparametric distribution family* for estimating  $\mu_k$ .

We first differentiate the non-parametric log likelihood ( 2.33) for the empirical entropy model with respect to  $\mu_k$ :

$$\frac{dl_k(\mu_k)}{d\mu_k} = N \cdot \frac{d\lambda_k(\mu_k)}{d\mu_k} \cdot (\bar{y}_k - \mu_k) \quad (2.71)$$

To derive ( 2.71), we have used the identity:

$$\sum_{t=1}^N \exp[\lambda_k(\mu_k) y_k(t)] \cdot [y_k(t) - \mu_k] = 0 \quad (2.72)$$

which follows from ( 2.69). We now compute  $d\lambda_k(\mu_k)/d\mu_k$  by differentiating ( 2.72):

$$\frac{d}{d\mu_k} \left\{ \sum_{t=1}^N \exp[\lambda_k(\mu_k)y_k(t)] \cdot [y_k(t) - \mu_k] \right\} = 0 \quad (2.73)$$

leading to

$$\frac{d\lambda_k(\mu_k)}{d\mu_k} = \frac{\sum_{t=1}^N \exp[\lambda_k(\mu_k)y_k(t)]}{\sum_{t=1}^N \exp[\lambda_k(\mu_k)y_k(t)] \cdot [y_k(t) - \mu_k]^2} \quad (2.74)$$

where we have used ( 2.72) to obtain ( 2.74). Finally,

$$\frac{d^2 l_k(\mu_k)}{d\mu_k^2} = N \cdot \frac{d^2 \lambda_k(\mu_k)}{d\mu_k^2} \cdot (\bar{y}_k - \mu_k) - N \cdot \frac{d\lambda_k(\mu_k)}{d\mu_k} \quad (2.75)$$

Then, assuming the discrete uniform distribution for the observations  $y_k(1), y_k(2), \dots, y_k(N)$ ,

we have  $\mu_k = \bar{y}_k$ ,  $\lambda_k(\bar{y}_k) = 0$ , and

$$-\frac{d^2 l_k(\bar{y}_k)}{d\mu_k^2} = \frac{N}{s_k^2} \quad (2.76)$$

which follows by using ( 2.74). Therefore, ( 2.27) holds, implying that estimating  $\mu_k$  is as hard as in a parametric Gaussian model and, consequently, the empirical approach employs a *least favorable nonparametric distribution family* (see also Section 2.8).

## 2.10 Appendix C: ICM detection for the Gaussian measurement-error model

Under the Gaussian measurement-error model ( 2.15) in Section 2.4.1, the conditional predictive log likelihoods in ( 2.11) simplify to

$$\begin{aligned} L_k(1|N_\beta(k)) &= \max_{\mu_k > 0, \sigma_k^2} \left\{ \sum_{t=1}^N \ln p_{\text{noise}_k}(y_k(t) - \mu_k; \sigma_k^2) \right\} + \ln p_{\beta_k|N_\beta(k)}(1|N_\beta(k)) \\ &= \begin{cases} -N/2 - (N/2) \cdot \ln(s_k^2) + \ln p_{\beta_k|N_\beta(k)}(1|N_\beta(k)), & \bar{y}_k > 0 \\ -N/2 - (N/2) \cdot \ln(s_{0,k}^2) + \ln p_{\beta_k|N_\beta(k)}(1|N_\beta(k)), & \bar{y}_k \leq 0 \end{cases} \quad (2.77) \end{aligned}$$

$$\begin{aligned} L_k(0|N_\beta(k)) &= \max_{\sigma_k^2} \left\{ \sum_{t=1}^N \ln p_{\text{noise}_k}(y_k(t); \sigma_k^2) \right\} + \ln p_{\beta_k|N_\beta(k)}(0|N_\beta(k)) \\ &= -N/2 - (N/2) \cdot \ln(s_{0,k}^2) + \ln p_{\beta_k|N_\beta(k)}(0|N_\beta(k)) \end{aligned}$$

and ( 2.38) follows.

## 2.11 Appendix D: ICM detection for nonparametric measurement-error models

We specialize Step (ICM2) of the ICM algorithm to the nonparametric measurement-error models in Section 2.4.1. Here, the conditional predictive log likelihoods in ( 2.11) simplify to

$$\begin{aligned} L_k(1|N_\beta(k)) &= \max_{\mu_k > 0} \{l_k(\mu_k)\} + \ln p_{\beta_k|N_\beta(k)}(1|N_\beta(k)) \\ L_k(0|N_\beta(k)) &= l_k(0) + \ln p_{\beta_k|N_\beta(k)}(0|N_\beta(k)) \end{aligned} \quad (2.78)$$

We now show that, for  $\kappa = 0$  and  $\kappa = 1$

$$\max_{\mu_k > 0} [l_k(\mu_k)] = \begin{cases} -N \ln N, & \bar{y}_k > 0 \\ l_k(0), & \bar{y}_k \leq 0 \end{cases} \quad (2.79)$$

**Proof of ( 2.79) for empirical likelihood:** Consider the empirical likelihood model ( $\kappa = 0$ ). Then, the result for  $\bar{y}_k > 0$  follows from ( 2.22).

We now focus on the case where  $\bar{y}_k \leq 0$ . Then, for  $\mu_k > 0$ , the expression in ( 2.56) is negative at  $\lambda_k = 0$ . Since ( 2.56) is a decreasing function of  $\lambda_k$ , the optimal  $\lambda_k$  which solves ( 2.56) for any  $\mu_k > 0$  must be negative. Then, ( 2.60) implies that, in this case,  $l_k(\mu_k)$  is a decreasing function of  $\mu_k$  and ( 2.79) follows.

**Proof of ( 2.79) for empirical entropy:** Consider now the empirical entropy model ( $\kappa = -1$ ). Then, the result for  $\bar{y}_k > 0$  follows by noting that

- $\lambda_k(\bar{y}_k) = 0$  solves ( 2.69) and
- the non-parametric log likelihood for the empirical entropy model is maximized at  $\mu_k = \bar{y}_k$ , which follows by setting  $dl_k(\mu_k)/d\mu_k$  in ( 2.71) to zero and noting that  $d\lambda_k(\mu_k)/d\mu_k$  is always positive [see ( 2.74)].

In the case where  $\bar{y}_k \leq 0$  and  $\mu_k > 0$ , the derivative  $dl_k(\mu_k)/d\mu_k$  in ( 2.71) is negative. Therefore,  $l_k(\mu_k)$  is a decreasing function of  $\mu_k$  and ( 2.79) follows.

Finally, substituting ( 2.78) and ( 2.79) into ( 2.44) yields the result.

## 2.12 Appendix E: GLR tests for $\mu_k$ under nonparametric measurement-error models

We derive the empirical likelihood and entropy GLR tests in Section 2.4.4. Under the null hypotheses  $H_{0,k}: \mu_k = 0$ , the asymptotic distribution of the GLR test statistics

$$2\max_{\mu_k > 0} \{l_k(\mu_k)\} - 2l_k(0) = [2N\ln N - 2l_k(0)] \cdot i_{[0,\infty)}(\bar{y}_k) \quad (2.80)$$

is given by, for  $l \geq 0$ ,

$$\lim_{N \rightarrow \infty} P(l_k \leq l) = \frac{1}{2}P(\chi_1^2 \leq l) + \frac{1}{2} \quad (2.81)$$

which follows by adapting the results in [62, 64] (for empirical likelihood) and [71, 72] (for empirical entropy) to the one-sided testing problem in Section 2.4. Here,  $\chi_1^2$  denotes a random variable having a central  $\chi^2$  distribution with one degree of freedom and can be obtained by squaring a standard normal random variable. The second term in ( 2.81) corresponds to the probability that  $\bar{y}_k < 0$  under  $H_{0,k}$ , which is 1/2; in this case, the GLR test statistics ( 2.80) becomes zero.

Note that ( 2.45) follows by using the square root of ( 2.80) as the test statistics, which is possible because  $-N\ln N - l_k(0)$  are non-negative. Then, ( 2.81) implies that a specified false-alarm probability  $P_{FA}$  will be achieved by comparing

$$\sqrt{2[-N\ln N - l_k(0)]} \cdot i_{[0,\infty)}(\bar{y}_k) \quad (2.82)$$

with the threshold  $\tau_{NP}$ , computing using ( 2.46).

## CHAPTER 3. BAYESIAN SIGNAL ESTIMATION FOR SENSOR NETWORKS IN THE PRESENCE OF NODE LOCALIZATION ERRORS

A paper submitted to *IEEE Transactions on Signal Processing*

Aleksandar Dogandžić and Benhong Zhang

### Abstract

Signal processing methods developed so far for sensor-network environments have ignored the effects of node localization inaccuracies. We propose a Bayesian framework that accounts for the inherent uncertainties in the node locations (caused by the node localization errors) and develop an estimation method that is robust to these uncertainties. We model the node localization errors as zero-mean Gaussian random vectors whose covariances are known up to a scaling factor. An ICM algorithm is developed to estimate the signal parameters of interest and applied to energy-based acoustic source localization. Numerical simulations demonstrate the performance of the proposed approach.

### 3.1 Introduction

Most nodes in a wireless sensor network estimate their locations [1, 30, 31]<sup>1</sup>. It is therefore important to take into account the inherent uncertainties in the node locations (caused by the localization errors) and incorporate them into the design of signal processing algorithms for sensor-network environments. The node location uncertainties have been considered in [77] in the context of coverage-oriented sensor deployment. In this paper, we propose a Bayesian

---

<sup>1</sup>Node localization is one of the canonical tasks in sensor networks [1] and is discussed in detail in an excellent review article [30], see also reference therein.

framework for taking into account the node localization errors, derive an ICM algorithm for estimating the phenomenon of interest, and apply it to energy-based acoustic source localization. Centralized and distributed Bayesian methods have been proposed in [78, 79] for node localization under the self-calibration scenario where no "anchor" nodes are present. Here, we focus on incorporating the node localization errors into the estimation of the *physical phenomenon* measured by the network.

In Section 3.2, we introduce the measurement model and prior specifications. In Section 3.3, we develop an ICM algorithm for approximate MAP estimation of the signal parameters. In Section 3.5, the proposed algorithm is applied to energy-based acoustic source localization and its performance is evaluated via numerical simulations. Concluding remarks are given in Section 3.6.

### 3.2 Measurement model and prior specifications

Assume that a *region of interest* contains  $K + L$  active nodes (sensors) at locations  $\mathbf{x}_k$ ,  $k = 1, 2, \dots, K$  and  $\mathbf{x}_{R,l}$ ,  $l = 1, 2, \dots, L$ , where each node at location  $\mathbf{x}_k$  collects a measurement  $y_k$  and each *reference node* at location  $\mathbf{x}_{R,l}$  collects a measurement  $y_{R,l}$  of the phenomenon of interest. Here, the region of interest is an area of the network *in the proximity* of the desired phenomenon, see Figure 3.1. The locations of the reference nodes are *known* exactly, which is needed for node localization.

Define  $\mathbf{y} = [y_1, y_2, \dots, y_K]^T$ ,  $\mathbf{y}_R = [y_{R,1}, y_{R,2}, \dots, y_{R,L}]^T$ ,  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K]^T$ , and  $\mathbf{x}_R = [\mathbf{x}_{R,1}, \mathbf{x}_{R,2}, \dots, \mathbf{x}_{R,L}]^T$ , where " $T$ " denotes a transpose. Furthermore, denote by

- $\phi$  the vector of unknown *signal parameters* describing the measurement phenomenon and
- $p(y_k|\phi, \mathbf{x}_k)$  the conditional pdf/pmf of the measurement  $y_k$  given  $\phi$  and sensor location  $\mathbf{x}_k$ , for  $k = 1, 2, \dots, K$ <sup>2</sup>.

---

<sup>2</sup>An analogous definition holds for  $p(y_{R,l}|\phi, \mathbf{x}_{R,l})$ .

### 3.2.1 Measurement-error model

We model the measurements  $y_k, k = 1, 2, \dots, K$  and  $y_{R,l}, l = 1, 2, \dots, L$  as conditional independent random variables given  $\phi, \mathbf{x}_k$  and  $\mathbf{x}_{R,l}$ , implying that the conditional pdf or pmf of all measurements  $\mathbf{y}$  and  $\mathbf{y}_R$  given  $\phi, \mathbf{x}$  and  $\mathbf{x}_R$  (i.e. the *likelihood function*) is<sup>3</sup>

$$p(\mathbf{y}|\phi, \mathbf{x}) \cdot p(\mathbf{y}_R|\phi, \mathbf{x}_R) = \prod_{k=1}^K p(y_k|\phi, \mathbf{x}_k) \cdot \prod_{l=1}^L p(y_{R,l}|\phi, \mathbf{x}_{R,l}) \quad (3.1)$$

**Partly linear signal in white Gaussian noise:** We now specialize the general measurement model in ( 3.1) to a practically important scenario where *partly linear signal*  $a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k)$  is corrupted by *spatially white Gaussian noise*:

$$y_k = f(\phi, \mathbf{x}_k) + e_k = a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k) + e_k \quad (3.2)$$

which also holds for the reference nodes, with  $y_k, \mathbf{x}_k$  and  $e_k$  replaced by  $y_{R,l}, \mathbf{x}_{R,l}$  and  $e_l$ . Here,  $e_k$  and  $e_l$  denote zero-mean i.i.d. Gaussian noise with known variance  $\sigma^2$  (which can be estimated in the calibration stage from noise-only data). The parametric signal  $a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k)$  is described by the  $r \times 1$  nonlinear parameter vector

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_r]^T \quad (3.3)$$

and scalar linear parameter  $a$ . Consequently, the vector of all signal parameters is

$$\phi = [\boldsymbol{\theta}, a]^T \quad (3.4)$$

Let us denote by  $N(\mathbf{z}; \boldsymbol{\mu}, \Sigma)$  the Gaussian pdf of a random vector  $\mathbf{z}$  with mean  $\boldsymbol{\mu}$  and covariance  $\Sigma$ . The model in ( 3.1) and ( 3.2) implies that, given  $\phi$  and  $\mathbf{x}$ ,  $y_k$  are conditionally independent with Gaussian pdfs

$$p(y_k|\phi, \mathbf{x}_k) = N(y_k; a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k), \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp \left\{ -\frac{[y_k - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k)]^2}{2\sigma^2} \right\} \quad (3.5)$$

and a similar expression holds for the reference nodes.

In the following, we adopt the partly linear signal and Gaussian noise models and compute the likelihood function for the measurement  $\mathbf{y}$  and  $\mathbf{y}_R$  by substituting ( 3.5) and the analogous

<sup>3</sup>If the region of interest *does not* contain reference nodes (i.e.  $L = 0$ ) then the second term in ( 3.1) is identically equal to one.

expression for  $\mathbf{y}_R$  into (3.1). However, our approach is general and applicable to non-Gaussian noise models as well.

### 3.2.2 Node location error model

We assume that the node locations  $\mathbf{x}_k, k = 1, 2, \dots, K$  are *unknown* and that their *estimates*  $\mathbf{w}_k, k = 1, 2, \dots, K$  are available and modeled as:

$$\mathbf{w}_k = \mathbf{x}_k + \mathbf{u}_k \quad (3.6)$$

Here,  $\mathbf{u}_k$  are *node location errors*, modeled as zero-mean independent Gaussian random vectors with covariances  $\sigma_u^2 \cdot U_k$ , where

- $\sigma_u^2$  is the unknown *location error variance parameter* and
- $U_k$  are known symmetric positive definite matrices.

Therefore,

$$p(\mathbf{w}_k | \mathbf{x}_k, \sigma_u^2) = N(\mathbf{w}_k; \mathbf{x}_k, \sigma_u^2 U_k) = \frac{1}{2\pi\sigma_u^2 \cdot |U_k|^{1/2}} \cdot \exp \left[ -\frac{(\mathbf{w}_k - \mathbf{x}_k)^T U_k^{-1} (\mathbf{w}_k - \mathbf{x}_k)}{2\sigma_u^2} \right] \quad (3.7)$$

where  $|\cdot|$  denotes the determinant. To facilitate node localization, we further assume that the exact locations  $\mathbf{x}_{R,l}, l = 1, 2, \dots, L$  of the reference nodes are *known*, see also Section 3.5.1.

The estimates  $\mathbf{w}_k$  are obtained using a node localization algorithm, which also provides the matrices  $U_k$ . In particular, we choose  $U_k$  as estimated CRB matrices for the node locations  $\mathbf{x}_k, k = 1, 2, \dots, K$ , see (3.26) in Section 3.5. We adopt a simplifying assumption that the node location estimates  $\mathbf{w}_k$  are *conditionally independent* given  $\mathbf{x}_k$  and  $\sigma_u^2$ :

$$p(\mathbf{w} | \mathbf{x}, \sigma_u^2) = \prod_{k=1}^K p(\mathbf{w}_k | \mathbf{x}_k, \sigma_u^2) \quad (3.8)$$

where  $\mathbf{w} = [\mathbf{w}_1^T, \mathbf{w}_2^T, \dots, \mathbf{w}_K^T]^T$ .

Finally, we assume that the node location errors are *non-differential*:

$$p(y_k | \phi, \mathbf{x}_k, \mathbf{w}_k) = p(y_k | \phi, \mathbf{x}_k), \quad k = 1, 2, \dots, K \quad (3.9)$$

or equivalently, the node location estimates  $\mathbf{w}_k$  and measurements  $y_k$  are conditionally independent given  $\mathbf{x}_k$ .

Define the vector of all unknown parameters:

$$\boldsymbol{\xi} = [\boldsymbol{\phi}^T, \mathbf{x}^T, \sigma_u^2]^T \quad (3.10)$$

and the vector of all "observations" (i.e. measurements  $\mathbf{y}$  and  $\mathbf{y}_R$  and node location estimates  $\mathbf{w}$ ):

$$\mathbf{v} = [\mathbf{y}^T, \mathbf{y}_R^T, \mathbf{w}^T]^T \quad (3.11)$$

under the signal and location-error models ( 3.1)-( 3.8).

### 3.2.3 Prior specifications

Assume that the signal and node location parameters are independent *a priori*<sup>4</sup>:

$$\pi_{\boldsymbol{\xi}} = \pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \cdot \left[ \prod_{k=1}^K \pi_{\mathbf{x}_k}(\mathbf{x}_k) \right] \cdot \pi_a(a) \cdot \pi_{\sigma_u^2}(\sigma_u^2) \quad (3.12)$$

$\mathbf{x}_k$  is the location of the  $k$ th node in Cartesian coordinates. We focus on a two-dimensional (2-D) network model with

$$\mathbf{x}_k = [x_{k,1}, x_{k,2}]^T, \quad k = 1, 2, \dots, K \quad (3.13)$$

and adopt the following simple uniform-distribution priors:

$$\begin{aligned} \pi_{\mathbf{x}_k}(\mathbf{x}_k) &= \pi_{x_{k,1}}(x_{k,1}) \cdot \pi_{x_{k,2}}(x_{k,2}) \\ \pi_{x_{k,1}}(x_{k,1}) &= \text{uniform}(x_{1,\text{MIN}}, x_{1,\text{MAX}}) \\ \pi_{x_{k,2}}(x_{k,2}) &= \text{uniform}(x_{2,\text{MIN}}, x_{2,\text{MAX}}) \\ \pi_a(a) &= \text{uniform}(a_{\text{MIN}}, a_{\text{MAX}}) \\ \pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}) &= \prod_{i=1}^r \pi_{\theta_i}(\theta_i) \\ \pi_{\theta_i}(\theta_i) &= \text{uniform}(\theta_{i,\text{MIN}}, \theta_{i,\text{MAX}}), i = 1, 2, \dots, r \\ \pi_{\sigma_u^2}(\sigma_u^2) &= \text{uniform}(1, \sigma_{u^2,\text{MAX}}) \end{aligned} \quad (3.14)$$

<sup>4</sup>Here,  $\pi_{\boldsymbol{\xi}}(\boldsymbol{\xi})$  denotes the prior pdf of  $\boldsymbol{\xi}$  and analogous notation is used for the prior pdfs of the components of  $\boldsymbol{\xi}$ .

Since  $U_k, k = 1, 2, \dots, K$  are estimated CRBs of the node locations, we have restricted  $\sigma_u^2$  to be bounded from below by unity, see ( 3.14). This choice is motivated by the CRB inequality which states that the covariances of the node location estimates are lower-bounded by their CRBs [75].

### 3.3 Bayesian analysis

We now develop a Bayesian approach for estimating the signal parameters  $\phi$  under the measurement and prior models in Section 3.2. The joint posterior distribution of all parameters  $\xi$  follows by using ( 3.1)-( 3.9):

$$\begin{aligned} p(\xi|\mathbf{v}) &\propto p(\mathbf{y}|\phi, \mathbf{x}) \cdot p(\mathbf{y}_R|\phi, \mathbf{x}_R) \cdot p(\mathbf{w}|\mathbf{x}, \sigma_u^2) \cdot \pi_\xi(\xi) \\ &= \left[ \prod_{k=1}^K p(y_k|\phi, \mathbf{x}_k) \cdot p(\mathbf{w}_k|\mathbf{x}_k, \sigma_u^2) \cdot \pi_{\mathbf{x}_k}(\mathbf{x}_k) \right] \cdot \left[ \prod_{l=1}^L p(y_{R,l}|\phi, \mathbf{x}_{R,l}) \right] \cdot \pi_\phi(\phi) \cdot \pi_{\sigma_u^2}(\sigma_u^2) \end{aligned} \quad (3.15)$$

Then, using the Gaussian likelihoods for the measurement ( 3.1) [see also ( 3.5)] and node locations ( 3.7)-( 3.8) yields

$$\begin{aligned} p(\xi|\mathbf{v}) &\propto \frac{1}{\sigma_u^{2K}} \cdot \exp \left( - \sum_{k=1}^K \left\{ \frac{[y_k - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k)]^2}{2\sigma^2} + \frac{(\mathbf{w}_k - \mathbf{x}_k)^T U_k^{-1} (\mathbf{w}_k - \mathbf{x}_k)}{2\sigma_u^2} \right\} \right) \\ &\cdot \exp \left\{ - \sum_{l=1}^L \frac{[y_{R,l} - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l})]^2}{2\sigma^2} \right\} \cdot \left[ \prod_{k=1}^K \pi_{\mathbf{x}_k}(\mathbf{x}_k) \right] \cdot \pi_a(a) \cdot \pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \cdot \pi_{\sigma_u^2}(\sigma_u^2) \end{aligned} \quad (3.16)$$

Our goal is to estimate the signal parameters  $\phi$  by maximizing the joint posterior distribution ( 3.15). We utilize the *iterative conditional modes algorithm* (see [56, 58]) to perform this maximization, yielding the posterior mode (MAP estimate) of  $\xi$ . Our ICM algorithm iterates between the following conditional maximization steps:

1. Fix  $\mathbf{x} = \hat{\mathbf{x}}$  and obtain an estimate  $\hat{\phi}$  of  $\phi$  by maximizing its full conditional posterior pdf  $p(\phi|\mathbf{x}, \mathbf{x}_R, \mathbf{y}, \mathbf{y}_R)$ :

$$\begin{aligned} \hat{\phi} &= \operatorname{argmax}_{\phi} \left\{ \left[ \prod_{k=1}^K p(y_k|\phi, \mathbf{x}_k) \right] \cdot \left[ \prod_{l=1}^L p(y_{R,l}|\phi, \mathbf{x}_{R,l}) \right] \cdot \pi_\phi(\phi) \right\} \\ &= \operatorname{argmin}_{\phi \in \Phi} \left\{ \sum_{k=1}^K [y_k - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k)]^2 + \sum_{l=1}^L [y_{R,l} - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l})]^2 \right\} \end{aligned} \quad (3.17)$$

easily performed using an  $r$ -dimensional *nested Gauss-Newton iteration*, see Section 3.7.

Here,

$$\Phi = \{\phi : a \in [a_{\text{MIN}}, a_{\text{MAX}}], \theta_i \in [\theta_{i,\text{MIN}}, \theta_{i,\text{MAX}}], i = 1, 2, \dots, r\} \quad (3.18)$$

denotes the parameter space of the signal parameters  $\phi$ .

2. Fix  $\phi = \hat{\phi}$  and  $\sigma_u^2 = \hat{\sigma}_u^2$  and estimate  $\mathbf{x}$  as

$$\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^T, \hat{\mathbf{x}}_2^T, \dots, \hat{\mathbf{x}}_K^T]^T \quad (3.19)$$

where  $\hat{\mathbf{x}}_k$  is an estimate of  $\mathbf{x}_k$  obtained by maximizing its full conditional posterior pdf  $p(\mathbf{x}_k | \phi, \sigma_u^2, y_k, \mathbf{w}_k)$ :

$$\begin{aligned} \hat{\mathbf{x}}_k &= \operatorname{argmax}_{\mathbf{x}_k} [p(y_k | \phi, \mathbf{x}_k) \cdot p(\mathbf{w}_k | \mathbf{x}_k, \sigma_u^2) \cdot \pi_{\mathbf{x}_k}(\mathbf{x}_k)] \\ &= \operatorname{argmax}_{\mathbf{x}_k \in \chi_k} \lambda_k(\mathbf{x}_k | \phi, \sigma_u^2, y_k, \mathbf{w}_k) \end{aligned} \quad (3.20)$$

easily performed using the Fisher scoring iteration derived in Section 3.8. Here,  $\chi_k = \{\mathbf{x}_k : x_{k,i} \in [x_{i,\text{MIN}}, x_{i,\text{MAX}}], i = 1, 2\}$  denotes the parameter space of  $\mathbf{x}_k$  and

$$\lambda_k(\mathbf{x}_k | \phi, \sigma_u^2, y_k, \mathbf{w}_k) = -\frac{[y_k - as(\boldsymbol{\theta}, \mathbf{x}_k)]^2}{2\sigma^2} - \frac{(\mathbf{w}_k - \mathbf{x}_k)^T U_k^{-1} (\mathbf{w}_k - \mathbf{x}_k)}{2\sigma_u^2} \quad (3.21)$$

can be viewed as a penalized log-likelihood function of  $x_k, k = 1, 2, \dots, K$ . Note that the above estimation of  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K$  *decouples* due to the simplifying assumption that the  $\mathbf{w}_k$ s are conditionally independent given  $\mathbf{x}_k$  and  $\sigma_u^2$ , see ( 3.8).

3. Fix  $\phi = \hat{\phi}$  and  $\mathbf{x} = \hat{\mathbf{x}}$  and compute an estimate of  $\sigma_u^2$ :

$$\hat{\sigma}_u^2 = \begin{cases} s_u^2, & 1 \leq s_u^2 < \sigma_{u,\text{MAX}}^2 \\ 1, & s_u^2 < 1 \\ \sigma_{u,\text{MAX}}^2, & s_u^2 > \sigma_{u,\text{MAX}}^2 \end{cases} \quad (3.22)$$

where

$$s_u^2 = \frac{\sum_{k=1}^K (\mathbf{w}_k - \mathbf{x}_k)^T U_k^{-1} (\mathbf{w}_k - \mathbf{x}_k)}{2K} \quad (3.23)$$

Here, the estimate in ( 3.22) has been obtained by maximizing the full conditional posterior pdf of  $\sigma_u^2$ , i.e.  $p(\sigma_u^2 | \mathbf{x}, \mathbf{w})$ .

The iteration 1-3 is performed until the estimates of  $\phi$ ,  $\mathbf{x}$  and  $\sigma_u^2$  do not change significantly between two consecutive cycles, indicating convergence. Upon convergence, we obtain a joint posterior mode of  $\phi$ ,  $\mathbf{x}$  and  $\sigma_u^2$ . If the algorithm converges to the global maximum of ( 3.15), then it provides the MAP estimates of these parameters<sup>5</sup>.

### 3.4 Cramer-Rao bound for the signal parameters

We derive the (non-Bayesian) CRB expression for the unknown signal parameters  $\phi$  under the measurement model in Section 3.2.1:

$$\begin{aligned} \text{CRB}_\phi &= \begin{bmatrix} \text{CRB}_\theta & \text{CRB}_{\theta,a} \\ \text{CRB}_{a,\theta} & \text{CRB}_a \end{bmatrix} \\ &= \left\{ \sum_{k=1}^K \frac{\Xi(\phi, \mathbf{x}_k)/\sigma^2}{1 + \sigma_u^2(a/\sigma)^2 \cdot \partial s(\theta, \mathbf{x}_k)/\partial \mathbf{x}_k^T \cdot U_k \cdot \partial s(\theta, \mathbf{x}_k)/\partial \mathbf{x}_k} + \sum_{l=1}^L \Xi(\phi, \mathbf{x}_{R,l})/\sigma^2 \right\}^{-1} \end{aligned} \quad (3.24)$$

where [see also ( 3.2)]

$$\begin{aligned} \Xi(\phi, \mathbf{x}_k) &= \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \\ &= \begin{bmatrix} a^2 \cdot \partial s(\theta, \mathbf{x}_k)/\partial \theta \cdot \partial s(\theta, \mathbf{x}_k)/\partial \theta^T & a \cdot s(\theta, \mathbf{x}_k) \cdot \partial s(\theta, \mathbf{x}_k)/\partial \theta \\ a \cdot s(\theta, \mathbf{x}_k) \cdot \partial s(\theta, \mathbf{x}_k)/\partial \theta^T & s^2(\theta, \mathbf{x}_k) \end{bmatrix} \end{aligned} \quad (3.25)$$

Interestingly,  $\text{CRB}_\theta$  depends on the linear parameter  $a$  and noise standard deviation  $\sigma$  only through  $(a/\sigma)^2$ , which is easily verified by applying the formula for the inverse of a partitioned matrix to ( 3.24).

**High SNR scenario:** Even when the SNR is high (i.e.  $a$  is large), ( 3.24) is *not* zero except  $L \geq 3$  (i.e. no less than 3 reference nodes are used in target localization).

Detailed CRB (and asymptotic CRB under high SNR) derivation can be found in Section 3.8

<sup>5</sup>Note that classical iterative nonlinear estimation methods (e.g. least squares, discussed in Section 3.5.5) also do not guarantee convergence to the global optimum.

### 3.5 Numerical examples: energy-based acoustic source localization

We evaluate the estimation accuracy of the ICM algorithm via numerical simulations and compare it with the existing techniques. We first describe

- the node localization scheme used to obtain the location estimates  $\mathbf{w}$  (Section 3.5.1),
- acoustic source-signal model (Section 3.5.2),
- prior specifications (Section 3.5.3)
- initialization scheme for starting the proposed iteration (Section 3.5.4)

and then present numerical simulation results in Section 3.5.5.

Throughout this section, we consider a 2-D sensor network with 30 nodes placed in a square region measuring  $15 \times 15\text{m}^2$ . Three reference nodes are located in the corners whereas the remaining nodes are randomly (uniformly) distributed within the network area, as depicted in Figure 3.1.

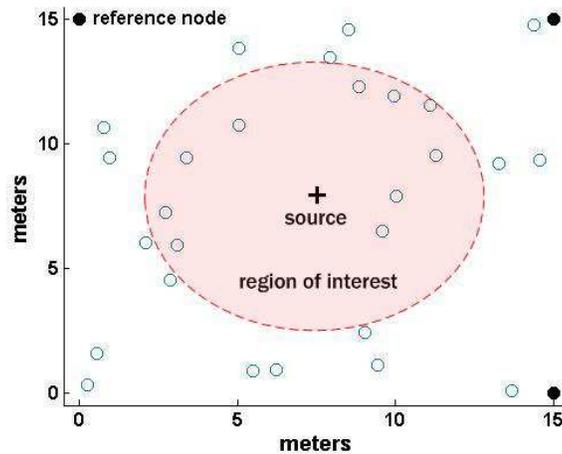


Figure 3.1 A 2-D sensor network with 30 nodes and a region of interest in the proximity of an acoustic source.

### 3.5.1 Node localization

We utilize TOA measurements between all node pairs *within the network* to localize the nodes. Note that three reference nodes are employed to facilitate the node localization, see also Figure 3.1. The TOA measurement between nodes at locations  $\mathbf{x}_k$  and  $\mathbf{x}_m$  is modeled as a Gaussian random variable with mean  $\|\mathbf{x}_k - \mathbf{x}_m\|/c$  and the constant variance  $\sigma_{\text{TOA}}^2$ , where,  $c$  denotes the speed of propagation of the TOA signal, see e.g. [31]<sup>6</sup>. To localize the nodes, we apply the ML relative location estimation algorithm in [31]. Within the region of interest, this algorithm yields the ML estimates  $\mathbf{w}_k, k = 1, 2, \dots, K$  of the node locations  $\mathbf{x}_k, k = 1, 2, \dots, K$ . The  $2 \times 2$  CRB matrices  $\text{CRB}_{\text{TOA}, \mathbf{x}_k}(\mathbf{x})$  for the node locations  $\mathbf{x}_k, k = 1, 2, \dots, K$  can be computed using [31]. We choose  $U_k$  as estimated CRB matrices of the locations  $\mathbf{x}_k$  (see also in Section 3.2.2):

$$U_k = \text{CRB}_{\text{TOA}, \mathbf{x}_k}(\mathbf{x})|_{\mathbf{x}=\mathbf{w}} = \text{CRB}_{\text{TOA}, \mathbf{x}_k}(\mathbf{w}) \quad (3.26)$$

where  $\text{CRB}_{\text{TOA}, \mathbf{x}_k}(\mathbf{w})$  are the corresponding  $2 \times 2$  blocks of the "global" CRB in [31].

An average node-location CRB

$$\text{crb}_{\text{TOA}}(\mathbf{x}) = \frac{1}{K+L} \cdot \sum_{k=1}^K \text{tr}[\text{CRB}_{\text{TOA}, \mathbf{x}_k}(\mathbf{x})] \quad (3.27)$$

is a good measure of the overall node location uncertainty within the region of interest. It scales linearly with the ranging-error variance  $c^2\sigma_{\text{TOA}}^2$ , see [31].

### 3.5.2 Signal and noise models for energy-based acoustic source localization

We consider an acoustic source localization problem where  $y_k$  is the sample mean (obtained by averaging over a time window) of the received energy measurements collected at node  $k$  [80, 81]. We adopt the partly linear signal and white Gaussian noise models in (3.2)-(3.5) with the following isotropic acoustic-energy attenuation function (see [22]):

$$s(\boldsymbol{\theta}, \mathbf{x}_k) = \frac{1}{1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha}} \quad (3.28)$$

where

<sup>6</sup>Here,  $\|\mathbf{x}_k - \mathbf{x}_m\| = \sqrt{(\mathbf{x}_k - \mathbf{x}_m)^T(\mathbf{x}_k - \mathbf{x}_m)}$  denotes the Euclidean distance between locations  $\mathbf{x}_k$  and  $\mathbf{x}_m$  in Cartesian coordinates.

- $\boldsymbol{\theta} = [\theta_1, \theta_2]^T$  is the unknown source location in Cartesian coordinates (implying that  $r = 2$ ) and
- $\alpha$  and  $\beta$  are the (known) energy attenuation function parameters.

(Similar signal and noise models were used in [18, 80–82] and have been experimentally validated in [82].) Since  $y_k, k = 1, 2, \dots, K$  correspond to energy measurements, we define the average SNR (in decibels) as follows [see also ( 3.2)]:

$$\text{SNR} = 10\log_{10} \left[ \frac{1}{K+L} \cdot \left( \sum_{k=1}^K f(\boldsymbol{\phi}, \mathbf{x}_k)/\sigma + \sum_{l=1}^L f(\boldsymbol{\phi}, \mathbf{x}_{R,l})/\sigma \right) \right] \quad (3.29)$$

To generate the simulated data, we have chosen the following source-signal energy [i.e.  $a$ , see ( 3.2)], location, and attenuation-function parameters:

$$a = 30, \quad \boldsymbol{\theta} = [7.5\text{m}, 7.5\text{m}]^T, \quad \alpha = 1, \quad \beta = 4 \quad (3.30)$$

### 3.5.3 Prior specifications

We selected the same prior pdf boundaries  $\theta_{i,\text{MIN}} = x_{i,\text{MIN}} = 0$  and  $\theta_{i,\text{MAX}} = x_{i,\text{MAX}} = 15\text{m}, i \in \{1, 2\}$  for the source and node locations, see also ( 3.14). Observe that the rectangle defined by  $\theta_{i,\text{MIN}}, \theta_{i,\text{MAX}}, i \in \{1, 2\}$  covers the entire  $15 \times 15\text{m}^2$  network area, see Figure 3.1.

Under the energy-based acoustic source localization scenario, the linear parameter  $a$  in ( 3.2) is the source-signal energy, implying that  $a_{\text{MIN}} \geq 0$ . We chose the prior pdfs in ( 3.14) with  $a_{\text{MIN}} = 0, a_{\text{MAX}} = 500$  and  $\sigma_{u,\text{MAX}} = 3$ .

Since the above prior pdfs are fairly vague, the ICM estimates of  $\boldsymbol{\phi}$  are approximately equal to their (non-Bayesian) ML estimates under the measurement model in Section 3.2.1. Hence, in this case, the signal-parameter CRB in Section 3.4 is valid benchmark for the achievable signal estimation performance.

### 3.5.4 Initialization

Following the approach in [18], we compute an initial estimate of the unknown signal parameters  $\boldsymbol{\theta}$  by *averaging* the  $\mathbf{w}_{k,S}$  and  $\mathbf{x}_{R,l,S}$  of the nodes within the region of interest whose

measurements ( $y_k$  and  $y_{R,l}$ , respectively) are higher than a specified threshold  $\tau$ :

$$\boldsymbol{\theta}^{(0)} = \frac{\sum_{k=1}^K \mathbf{w}_k \cdot i_{[\tau, \infty)}(y_k) + \sum_{l=1}^L \mathbf{x}_{R,l} \cdot i_{[\tau, \infty)}(y_{R,l})}{\sum_{k=1}^K i_{[\tau, \infty)}(y_k) + \sum_{l=1}^L i_{[\tau, \infty)}(y_{R,l})} \quad (3.31)$$

where

$$i_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise} \end{cases} \quad (3.32)$$

denotes the indicator function. In this section, we average the locations of the nodes having the  $n = 10$  largest measurements, implying the following choice of  $\tau$ :

$$\tau = y_{(10)} \quad (3.33)$$

where  $y_{(10)}$  denotes the 10th largest measurement (order statistic) observed among the nodes within the region of interest. Once  $\boldsymbol{\theta}^{(0)}$  is computed, we can obtain an initial estimate of the linear parameter  $a$  using (3.44) in Section 3.7, with  $i$  set to zero.

Initial estimates of the node locations  $\mathbf{x}_k, k = 1, 2, \dots, K$  are chosen to coincide with their estimates provided by the node localization algorithm:

$$\mathbf{x}_k^{(0)} = \mathbf{w}_k, \quad k = 1, 2, \dots, K \quad (3.34)$$

### 3.5.5 Simulation examples

We now describe the simulation scenarios and present numerical simulation results. Simulated data were generated using the acoustic measurement-error model in Section 3.2.1 and 3.5.2 with signal and noise parameters in (3.30). The node location estimates  $\mathbf{w}_k, k = 1, 2, \dots, K$  are obtained from TOA measurements using the ML relative location estimation algorithm, outlined in Section 3.5.1 and [31].

We compare our ICM algorithm in Section 3.3 with the (classical) LS signal parameter estimator. This LS estimator is derived by assuming that the node locations are exactly known and given by  $\mathbf{w}_k, k = 1, 2, \dots, K$  and  $\mathbf{x}_{R,l}, l = 1, 2, \dots, L$ . Consequently, [see also (3.17)]:

$$\begin{aligned} \hat{\boldsymbol{\phi}}_{\text{LS}} &= \operatorname{argmax}_{\boldsymbol{\phi}} [p(\mathbf{y}|\boldsymbol{\phi}, \mathbf{w}) \cdot p(\mathbf{y}_R|\boldsymbol{\phi}, \mathbf{x}_R) \cdot \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi})] \\ &= \operatorname{argmin}_{\boldsymbol{\phi} \in \Phi} \left\{ \sum_{k=1}^K [y_k - a \cdot s(\boldsymbol{\theta}, \mathbf{w}_k)]^2 + \sum_{l=1}^L [y_{R,l} - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l})]^2 \right\} \end{aligned} \quad (3.35)$$

If we utilize the initial values of  $\mathbf{x}_k^{(0)} = \mathbf{w}_k$  then Step 1 of the first ICM cycle provides exactly the above LS estimate, see ( 3.17) and ( 3.18). Using a Taylor-series expansion approach, we derive an approximate analytical expression for the covariance matrix of the LS estimator under the measurement model in Section 3.2.1 (see Section 3.9 for detailed derivation):

$$\begin{aligned} \text{Cov}(\hat{\boldsymbol{\phi}}_{\text{LS}}) \approx & \sigma^2 \cdot P(\boldsymbol{\phi}, \mathbf{x})^{-1} + \sigma_u^2 \cdot P(\boldsymbol{\phi}, \mathbf{x})^{-1} \sum_{k=1}^K \left[ \sigma^2 \cdot \frac{\partial^2 f(\boldsymbol{\phi}, \mathbf{x}_k)}{\partial \boldsymbol{\phi} \partial \mathbf{x}_k^T} U_k \frac{\partial^2 f(\boldsymbol{\phi}, \mathbf{x}_k)}{\partial \mathbf{x}_k \partial \boldsymbol{\phi}^T} \right. \\ & \left. + \frac{\partial f(\boldsymbol{\phi}, \mathbf{x}_k)}{\partial \boldsymbol{\phi}} \frac{\partial f(\boldsymbol{\phi}, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} U_k \frac{\partial f(\boldsymbol{\phi}, \mathbf{x}_k)}{\partial \mathbf{x}_k} \frac{\partial f(\boldsymbol{\phi}, \mathbf{x}_k)}{\partial \boldsymbol{\phi}^T} \right] P(\boldsymbol{\phi}, \mathbf{x}_k)^{-1} \end{aligned} \quad (3.36)$$

where

$$P(\boldsymbol{\phi}, \mathbf{x}) = \sum_{k=1}^K \Xi(\boldsymbol{\phi}, \mathbf{x}_k) + \sum_{l=1}^L \Xi(\boldsymbol{\phi}, \mathbf{x}_{R,l}) \quad (3.37)$$

where  $\Xi(\boldsymbol{\phi}, \mathbf{x}_k)$  was defined in 3.25.

In the examples presented here, we define the region of interest as a set containing  $K = 10$  nodes with the ten largest measurements within the network,  $y_k, k = 1, 2, \dots, 10$ . [Figure 3.1 depicts a typical (most likely) region of interest; observe that there are no reference nodes in this region.] Our performance metric is the *average MSE* of an estimator calculated using 2000 independent trials, where averaging is performed over independent measurement and TOA noise realization, as well as random realizations of the region of interest. Here, we present the total average MSEs for the source-location estimates:

$$\overline{\text{MSE}}(\boldsymbol{\theta}) = \overline{\text{MSE}}(\theta_1) + \overline{\text{MSE}}(\theta_2) \quad (3.38)$$

where "MSE" denotes averaging over random realizations of the region of interest. In Figure 3.2, we show the average MSEs for the ICM and LS methods and corresponding analytical accuracy measures,  $\text{tr}[\overline{\text{CRB}}_{\boldsymbol{\theta}}]$ ,  $\text{tr}[\overline{\text{CRB}}_{\boldsymbol{\theta}}|_{\text{largea}}]$ , and  $\text{tr}[\overline{\text{Cov}}(\hat{\boldsymbol{\theta}}_{\text{LS}})]$  [see Section 3.8 and 3.9 for detailed derivation] as functions of the average node-location  $\overline{\text{crb}}_{\text{TOA}}$  [computed using ( 3.27)], for the following average SNRs:  $\overline{\text{SNR}} = 7\text{dB}$  and  $\overline{\text{SNR}} = 12\text{dB}$  [computed using ( 3.29)]. The ICM method converged in 10 iteration steps. [Here, we vary  $\overline{\text{crb}}_{\text{TOA}}$  by changing the ranging-error variance  $c^2\sigma_{\text{TOA}}^2$ . If there are no reference nodes in the region of interest,  $\text{tr}[\overline{\text{CRB}}_{\boldsymbol{\theta}}|_{\text{largea}}]$  and  $\text{tr}[\overline{\text{Cov}}(\hat{\boldsymbol{\theta}}_{\text{LS}})]$  increase linearly with  $c^2\sigma_{\text{TOA}}^2$ , see also Figure 3.2.] Clearly, our ICM method

outperforms the LS estimator which ignores the node localization inaccuracies. The performance difference between the two methods increases as the node location uncertainty increases. As expected,  $\text{tr}[\overline{\text{CRB}}_{\theta}|_{\text{large } a}]$  does not depend on  $\overline{\text{SNR}}$ .

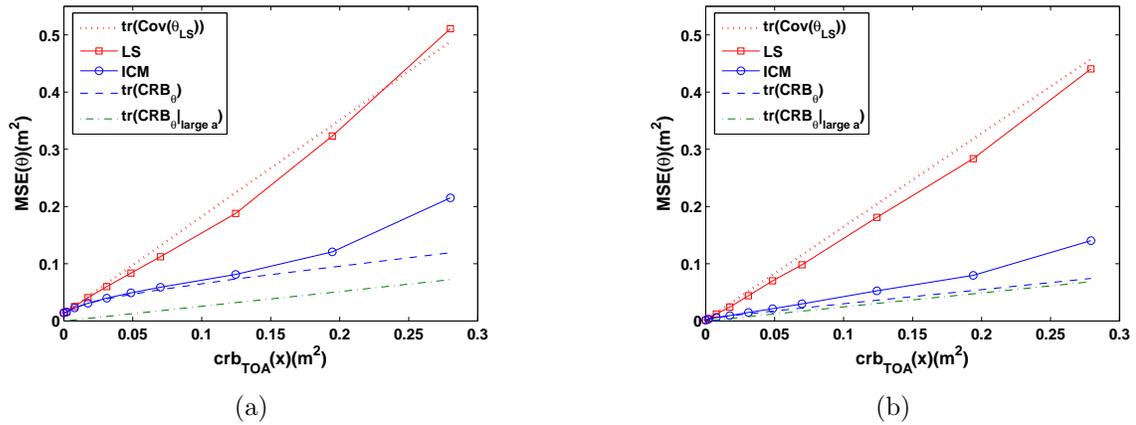


Figure 3.2 Average MSEs, CRBs, and approximate covariances for the ICM and LS source location estimates as functions of the average node-location CRB, for (Left)  $\overline{\text{SNR}} = 7\text{dB}$  and (Right)  $\overline{\text{SNR}} = 12\text{dB}$ .

In Figure 3.3, we show the average MSEs for the ICM and LS methods and corresponding analytical accuracy measures as a function of  $\overline{\text{SNR}}$ , for  $\overline{\text{crb}}_{\text{TOA}} = 0.02$  and  $\overline{\text{crb}}_{\text{TOA}} = 0.2$ . As expected, the ICM method outperforms the LS estimator. Observe that our approximate analytical expression for the covariance matrix of the LS estimator is less accurate at high SNRs.

### 3.6 Concluding remarks

We proposed a Bayesian framework for taking into account node localization inaccuracies and developed an estimation method that is robust to these inaccuracies. This framework provides significant performance improvements compared with methods that ignore the node localization errors.

Further research will include: experimentally validating the proposed approach and studying its robustness to node-location-error model assumption (i.e. choices of the  $U_k$ s and uniform

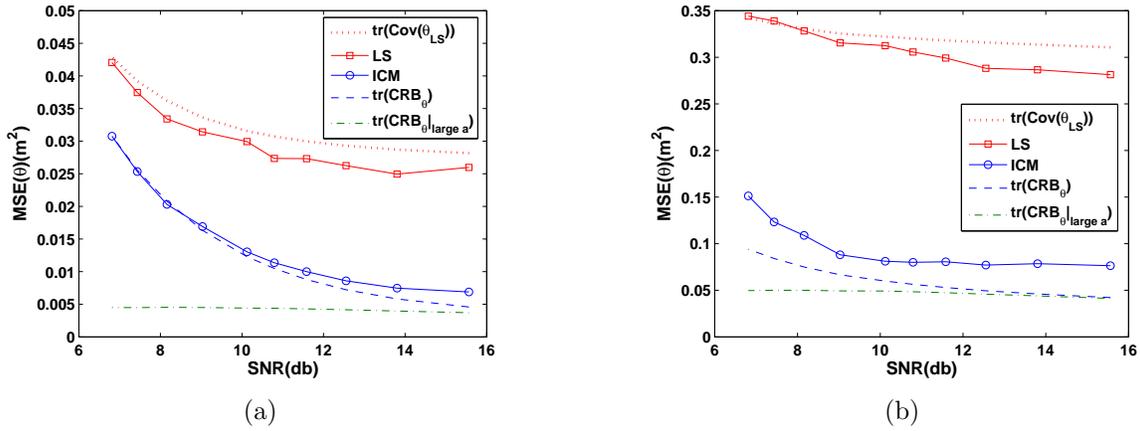


Figure 3.3 Average MSEs, CRBs, and approximate covariances for the ICM and LS source location estimates as functions of  $\overline{\text{SNR}}$ , for (Left)  $\overline{\text{crb}}_{\text{TOA}} = 0.02\text{m}^2$  and (Right)  $\overline{\text{crb}}_{\text{TOA}} = 0.2\text{m}^2$ .

degradation assumption in Section 3.2.2), and developing distributed implementations of our methods. We will also extend the proposed approach to the spatio-temporal measurement scenario where we will develop tracking algorithms (utilizing e.g. related ideas from robotics [83]) and estimate the measurement-error model (using e.g. empirical likelihood [62]).

### 3.7 Appendix A: Step 1 and 2 of the ICM algorithm

**Step 1:** We derive an  $r$ -dimensional *nested Gauss-Newton algorithm* for efficient estimation of  $\phi$  in Step 1 of the ICM algorithm, see also Section 3.3.

Let us first compute

$$\begin{aligned}
 & \sum_{k=1}^K \frac{\partial[as(\boldsymbol{\theta}, \mathbf{x}_k)]}{\partial \boldsymbol{\phi}} \cdot \frac{\partial[as(\boldsymbol{\theta}, \mathbf{x}_k)]}{\partial \boldsymbol{\phi}^T} + \sum_{l=1}^L \frac{\partial[as(\boldsymbol{\theta}, \mathbf{x}_{R,l})]}{\partial \boldsymbol{\phi}} \cdot \frac{\partial[as(\boldsymbol{\theta}, \mathbf{x}_{R,l})]}{\partial \boldsymbol{\phi}^T} \\
 & = \begin{bmatrix} a^2 \cdot A(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) & a \cdot \mathbf{b}(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) \\ a \cdot \mathbf{b}(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R)^T & c(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) \end{bmatrix} \tag{3.39}
 \end{aligned}$$

where

$$\begin{aligned}
A(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) &= \sum_{k=1}^K \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\boldsymbol{\theta}} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\boldsymbol{\theta}^T} + \sum_{l=1}^L \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\boldsymbol{\theta}} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\boldsymbol{\theta}^T} \\
\mathbf{b}(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) &= \sum_{k=1}^K \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\boldsymbol{\theta}} \cdot s(\boldsymbol{\theta}, \mathbf{x}_k) + \sum_{l=1}^L \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\boldsymbol{\theta}} \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l}) \\
c(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) &= \sum_{k=1}^K s^2(\boldsymbol{\theta}, \mathbf{x}_k) + \sum_{l=1}^L s^2(\boldsymbol{\theta}, \mathbf{x}_{R,l})
\end{aligned} \tag{3.40}$$

The upper left block of the inverse of ( 3.39) is

$$\frac{1}{a^2} \cdot Q(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) \tag{3.41}$$

where

$$Q(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) = \left[ A(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) - \frac{\mathbf{b}(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R)\mathbf{b}(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R)^T}{c(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R)} \right]^{-1} \tag{3.42}$$

Now, the nested Gauss-Newton iteration can be written as follows<sup>7</sup>:

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} + \delta_{\boldsymbol{\theta}}^{(i)} \cdot \frac{1}{a^{(i)}} \cdot Q(\boldsymbol{\theta}^{(i)}, \mathbf{x}, \mathbf{x}_R) \cdot \mathbf{q}(\boldsymbol{\theta}^{(i)}, a^{(i)}, \mathbf{x}, \mathbf{x}_R) \tag{3.43}$$

where

$$\begin{aligned}
a^{(i)} &= \begin{cases} \hat{a}(\boldsymbol{\theta}^{(i)}, \mathbf{x}, \mathbf{x}_R), & a_{\text{MIN}} \leq a(\boldsymbol{\theta}^{(i)}, \mathbf{x}, \mathbf{x}_R) < a_{\text{MAX}} \\ a_{\text{MIN}}, & a(\boldsymbol{\theta}^{(i)}, \mathbf{x}, \mathbf{x}_R) < a_{\text{MIN}} \\ a_{\text{MAX}}, & a(\boldsymbol{\theta}^{(i)}, \mathbf{x}, \mathbf{x}_R) > a_{\text{MAX}} \end{cases} \\
a(\boldsymbol{\theta}^{(i)}, \mathbf{x}, \mathbf{x}_R) &= \left\{ \sum_{k=1}^K y_k s(\boldsymbol{\theta}, \mathbf{x}_k) + \sum_{l=1}^L y_{R,l} s(\boldsymbol{\theta}, \mathbf{x}_{R,l}) \right\} / c(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_R) \\
\mathbf{q}(\boldsymbol{\theta}, a, \mathbf{x}, \mathbf{x}_R) &= \sum_{k=1}^K \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \boldsymbol{\theta}} \cdot [y_k - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k)] + \sum_{l=1}^L \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\partial \boldsymbol{\theta}} \cdot [y_k - a \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l})]
\end{aligned} \tag{3.44}$$

and the *damping factor*  $0 < \delta_{\boldsymbol{\theta}}^{(i)} \leq 1$  is chosen (at every step  $i$ ) to ensure that the conditional posterior pdf  $p(\phi|\mathbf{x}, \mathbf{x}_R, \mathbf{y}, \mathbf{y}_R)$  increases or, equivalently, that cost function in ( 3.17) *decreases* and the elements of  $\boldsymbol{\theta}$  remain within the parameter space. Upon convergence (i.e. as  $i \rightarrow \infty$ ), we obtain estimates of the signal parameters  $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(\infty)}$  and  $\hat{a} = a^{(\infty)}$  and, consequently,  $\hat{\boldsymbol{\phi}} = [(\boldsymbol{\theta}^{(\infty)})^T, a^{(\infty)}]^T$ .

<sup>7</sup>Nested Newton-type algorithms are discussed in detail in [84].

*Energy-based acoustic source localization:* For the energy-based acoustic source localization problem with isotropic acoustic-energy attenuation model ( 3.28), we have

$$\frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \boldsymbol{\theta}} = -\frac{2\alpha\beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-1)}}{(1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha})^2} \cdot (\boldsymbol{\theta} - \mathbf{x}_k) \quad (3.45)$$

and an analogous expression for the reference nodes is obtained by replacing  $\mathbf{x}_k$  by  $\mathbf{x}_{R,l}$ .

**Step 2:** We propose a Fisher scoring iteration to perform the maximizations in ( 3.20) [see also ( 3.21)]:

$$\mathbf{x}_k^{(i+1)} = \mathbf{x}_k^{(i)} + \delta_{\mathbf{x}_k}^{(i)} \cdot [\mathcal{I}_{\mathbf{x}_k}(\mathbf{x}_k^{(i)}; \boldsymbol{\phi}, \sigma_u^2)]^{-1} \frac{\partial \lambda_k(\mathbf{x}_k^{(i)} | \boldsymbol{\phi}, \sigma_u^2)}{\partial \mathbf{x}_k} \quad (3.46)$$

where

$$\begin{aligned} \frac{\partial \lambda_k(\mathbf{x}_k | \boldsymbol{\phi}, \sigma_u^2, y_k)}{\partial \mathbf{x}_k} &= a \cdot \frac{y_k - a s(\boldsymbol{\theta}, \mathbf{x}_k)}{\sigma^2} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \mathbf{x}_k} + \frac{1}{\sigma_u^2} \cdot U_k^{-1}(\mathbf{w}_k - \mathbf{x}_k) \\ \mathcal{I}_{\mathbf{x}_k}(\mathbf{x}_k; \boldsymbol{\phi}, \sigma_u^2) &= E_{y_k | \boldsymbol{\phi}, \mathbf{x}_k} \left[ -\frac{\partial^2 \lambda_k(\mathbf{x}_k | \boldsymbol{\phi}, \sigma_u^2, y_k)}{\partial \mathbf{x}_k \partial \mathbf{x}_k^T} \right] \\ &= \frac{a^2}{\sigma^2} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \mathbf{x}_k} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} + \frac{1}{\sigma_u^2} \cdot U_k^{-1} \end{aligned} \quad (3.47)$$

and the *damping factor*  $0 < \delta_{\mathbf{x}_k}^{(i)} \leq 1$  is chosen (at every step  $i$ ) to ensure that ( 3.21) increases and  $\mathbf{x}_k^{(i+1)}$  remains within the interval specified by the prior pdfs of the components of  $\mathbf{x}_k$  in ( 3.14).

*Energy-based acoustic source localization:* For the energy-based acoustic source localization problem with isotropic acoustic-energy attenuation model ( 3.28), we have

$$\frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \mathbf{x}_k} = \frac{2\alpha\beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-1)}}{(1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha})^2} \cdot (\boldsymbol{\theta} - \mathbf{x}_k) \quad (3.48)$$

### 3.8 Appendix B: CRB for ICM algorithm

Define

$$\mathbf{z} = [\boldsymbol{\phi}, \mathbf{x}]^T \quad (3.49)$$

and

$$\boldsymbol{\mu} = [f(\boldsymbol{\phi}, \mathbf{x}_1), f(\boldsymbol{\phi}, \mathbf{x}_2), \dots, f(\boldsymbol{\phi}, \mathbf{x}_K), f(\boldsymbol{\phi}, \mathbf{x}_{R,1}), f(\boldsymbol{\phi}, \mathbf{x}_{R,2}), \dots, f(\boldsymbol{\phi}, \mathbf{x}_{R,L})]^T \quad (3.50)$$



with the following simplification

$$\begin{aligned}
& FF_k - FX_k \cdot [XX_k + \sigma^2/\sigma_u^2 \cdot U_k^{-1}]^{-1} \cdot FX_k^T \\
&= \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \left\{ 1 - \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} \cdot \left[ \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} + \frac{\sigma^2}{\sigma_u^2} \cdot U_k^{-1} \right]^{-1} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k} \right\} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \\
&= \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \cdot \left[ 1 + (\sigma_u^2/\sigma^2) \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} U_k \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k} \right]^{-1}
\end{aligned} \tag{3.56}$$

Finally,

$$\begin{aligned}
\text{CRB}_\phi &= \sigma^2 \cdot \left\{ \sum_{k=1}^K \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \cdot \left[ 1 + (\sigma_u^2/\sigma^2) \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} U_k \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k} \right]^{-1} \right. \\
&\quad \left. + \sum_{l=1}^L \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \right\}^{-1}
\end{aligned} \tag{3.57}$$

Plug in the expression in (3.2), we have

$$\begin{aligned}
\text{CRB}_\phi &= \sigma^2 \cdot \left\{ \sum_{k=1}^K \frac{\begin{bmatrix} a^2 \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k)/\partial \boldsymbol{\theta} \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k)/\partial \boldsymbol{\theta}^T & a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k) \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k)/\partial \boldsymbol{\theta} \\ a \cdot s(\boldsymbol{\theta}, \mathbf{x}_k) \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k)/\partial \boldsymbol{\theta}^T & s^2(\boldsymbol{\theta}, \mathbf{x}_k) \end{bmatrix}}{1 + a^2(\sigma_u^2/\sigma^2) \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k)/\partial \mathbf{x}_k^T \cdot U_k \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k)/\partial \mathbf{x}_k} \right. \\
&\quad \left. + \sum_{l=1}^L \frac{\begin{bmatrix} a^2 \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})/\partial \boldsymbol{\theta} \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})/\partial \boldsymbol{\theta}^T & a \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l}) \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})/\partial \boldsymbol{\theta} \\ a \cdot s(\boldsymbol{\theta}, \mathbf{x}_{R,l}) \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})/\partial \boldsymbol{\theta}^T & s^2(\boldsymbol{\theta}, \mathbf{x}_{R,l}) \end{bmatrix}}{\right\}^{-1}
\end{aligned} \tag{3.58}$$

**CRB for source location at high SNR:** We now analyze the CRB expression for the source location  $\boldsymbol{\theta}$  in the scenario where the source-signal energy  $a$  is large (and, consequently,

the signal-to-noise ratio is high). Define

$$\begin{aligned}
P &= P(\boldsymbol{\theta}, \mathbf{x}) = \sum_{k=1}^K \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta} \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta}^T}{\partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k^T \cdot U_k \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k} \\
\mathbf{q} &= \mathbf{q}(\boldsymbol{\theta}, \mathbf{x}) = \sum_{k=1}^K \frac{s(\boldsymbol{\theta}, \mathbf{x}_k) \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta}}{\partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k^T \cdot U_k \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k} \\
r &= r(\boldsymbol{\theta}, \mathbf{x}) = \sum_{k=1}^K \frac{s^2(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k^T \cdot U_k \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k} \\
S_R &= S_R(\boldsymbol{\theta}) = \sum_{l=1}^L \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\partial \boldsymbol{\theta}} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\partial \boldsymbol{\theta}^T} \\
\mathbf{v}_R &= \mathbf{v}_R(\boldsymbol{\theta}) = \sum_{l=1}^L s(\boldsymbol{\theta}, \mathbf{x}_{R,l}) \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,l})}{\partial \boldsymbol{\theta}} \\
\omega_R &= \omega_R(\boldsymbol{\theta}) = \sum_{l=1}^L s^2(\boldsymbol{\theta}, \mathbf{x}_{R,l})
\end{aligned} \tag{3.59}$$

To simplify the notation, we omit the dependencies of  $P$ ,  $\mathbf{q}$ , and  $R$  on  $\boldsymbol{\theta}$  and  $\mathbf{x}$  and the dependencies of  $S_R$ ,  $\mathbf{v}_R$ , and  $\omega_R$  on  $\boldsymbol{\theta}$ . Now the CRB expressions for  $\boldsymbol{\phi}$  and  $\boldsymbol{\theta}$  at large SNR can be approximated by

$$\text{CRB}_{\boldsymbol{\phi} | \text{large } a} = \sigma^2 \cdot \begin{bmatrix} a^2 \cdot S_R + (\sigma^2 / \sigma_u^2) \cdot P & a \cdot \mathbf{v}_R + (\sigma^2 / \sigma_u^2) \cdot \mathbf{q} / a \\ a \cdot \mathbf{v}_R^T + (\sigma^2 / \sigma_u^2) \cdot \mathbf{q}^T / a & \omega_R + (\sigma^2 / \sigma_u^2) \cdot r / a^2 \end{bmatrix}^{-1} \tag{3.60}$$

and

$$\begin{aligned}
&\text{CRB}_{\boldsymbol{\theta} | \text{large } a} \\
&= \sigma^2 \cdot \left\{ a^2 \cdot S_R + (\sigma^2 / \sigma_u^2) \cdot P - \frac{[a \cdot \mathbf{v}_R + (\sigma^2 / \sigma_u^2) \cdot \mathbf{q} / a] \cdot [a \cdot \mathbf{v}_R + (\sigma^2 / \sigma_u^2) \cdot \mathbf{q} / a]^T}{\omega_R + (\sigma^2 / \sigma_u^2) \cdot r / a^2} \right\}^{-1}
\end{aligned} \tag{3.61}$$

where ( 3.61) follows by using the formula for the inverse of a partitioned matrix in e.g. [85].

Consider now the case where no reference nodes are present in the region of interest (i.e.  $L = 0$ ) and the signal amplitude  $a$  is large. Then ( 3.60) and ( 3.61) simplify to

$$\text{CRB}_{\boldsymbol{\phi} | \text{large } a} = \sigma_u^2 \cdot \begin{bmatrix} P & (1/a) \cdot \mathbf{q} \\ (1/a) \cdot \mathbf{q}^T & (1/a^2) \cdot r \end{bmatrix}^{-1} \tag{3.62}$$

$$\text{CRB}_{\boldsymbol{\theta} | \text{large } a} = \sigma_u^2 \cdot (P - \mathbf{q}\mathbf{q}^T / r)^{-1} \tag{3.63}$$

Note that ( 3.63) does not depend on  $a$  and  $\sigma^2$ . Hence, for large  $a$  and  $L = 0$ , the estimation accuracy is limited by the node location uncertainties.

Consider now the scenario where one or more reference nodes are present in the region of interest (i.e.  $L > 0$ ). If  $S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R$  is positive definite (which holds with probability one if  $L \geq 3$ ), then

$$\text{CRB}_{\boldsymbol{\theta}}|_{\text{large } a} = 0 \quad (3.64)$$

If  $1 \leq L \leq 2$ , then  $S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R$  is singular and

$$\text{CRB}_{\boldsymbol{\theta}}|_{\text{large } a} \approx \sigma^2 \cdot \left\{ a^2 \cdot (S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R) + (\sigma^2 / \sigma_u^2) \cdot \left[ P + \frac{r \cdot \mathbf{v}_R \mathbf{v}_R^T}{\omega_R^2} - \frac{\mathbf{q} \mathbf{v}_R^T}{\omega_R} - \frac{\mathbf{v}_R \mathbf{q}^T}{\omega} \right] \right\}^{-1} \quad (3.65)$$

which follows from ( 3.61) by using the following approximation:

$$\frac{a^2 \mathbf{v}_R \mathbf{v}_R^T}{\omega_R + (\sigma^2 / \sigma_u^2) \cdot r / a^2} \approx \frac{a^2 \cdot \mathbf{v}_R \mathbf{v}_R^T}{\omega_R} \cdot \left[ 1 - (\sigma^2 / \sigma_u^2) \cdot \frac{r}{a^2 \omega_R} \right] \quad (3.66)$$

For  $L = 1$ ,  $S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R = 0$ , and

$$\text{CRB}_{\boldsymbol{\theta}}|_{\text{large } a} \approx \sigma_u^2 \cdot \left[ P + \frac{r}{s^2(\boldsymbol{\theta}, \mathbf{x}_{R,1})} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,1})}{\partial \boldsymbol{\theta}} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,1})}{\partial \boldsymbol{\theta}^T} - \frac{\mathbf{q}}{s(\boldsymbol{\theta}, \mathbf{x}_{R,1})} \cdot \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,1})}{\partial \boldsymbol{\theta}^T} - \frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_{R,1})}{\partial \boldsymbol{\theta}} \cdot \frac{\mathbf{q}}{s(\boldsymbol{\theta}, \mathbf{x}_{R,1})} \right]^{-1} \quad (3.67)$$

For  $L = 2$ ,  $S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R$  is nonzero singular matrix and

$$\begin{aligned} [\text{CRB}_{\boldsymbol{\theta}}]_{1,1}|_{\text{large } a} &= \sigma^2 \cdot \frac{[S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{2,2}}{(\sigma^2 / \sigma_u^2) \cdot \Delta(\boldsymbol{\theta}, \mathbf{x})} = \sigma_u^2 \cdot \frac{[S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{2,2}}{\Delta(\boldsymbol{\theta}, \mathbf{x})} \\ [\text{CRB}_{\boldsymbol{\theta}}]_{2,2}|_{\text{large } a} &= \sigma^2 \cdot \frac{[S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{1,1}}{(\sigma^2 / \sigma_u^2) \cdot \Delta(\boldsymbol{\theta}, \mathbf{x})} = \sigma_u^2 \cdot \frac{[S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{1,1}}{\Delta(\boldsymbol{\theta}, \mathbf{x})} \end{aligned} \quad (3.68)$$

where

$$\begin{aligned} \Delta(\boldsymbol{\theta}, \mathbf{x}) &= [S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{1,1} \cdot \left[ P + \frac{r \cdot \mathbf{v}_R \mathbf{v}_R^T}{\omega_R^2} - \frac{\mathbf{q} \mathbf{v}_R^T}{\omega_R} - \frac{\mathbf{v}_R \mathbf{q}^T}{\omega_R} \right]_{2,2} \\ &+ [S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{2,2} \cdot \left[ P + \frac{r \cdot \mathbf{v}_R \mathbf{v}_R^T}{\omega_R^2} - \frac{\mathbf{q} \mathbf{v}_R^T}{\omega_R} - \frac{\mathbf{v}_R \mathbf{q}^T}{\omega_R} \right]_{1,1} \\ &- 2 \cdot [S_R - \mathbf{v}_R \mathbf{v}_R^T / \omega_R]_{1,2} \cdot \left[ P + \frac{r \cdot \mathbf{v}_R \mathbf{v}_R^T}{\omega_R^2} - \frac{\mathbf{q} \mathbf{v}_R^T}{\omega_R} - \frac{\mathbf{v}_R \mathbf{q}^T}{\omega_R} \right]_{1,2} \end{aligned} \quad (3.69)$$

Note that the expressions in ( 3.68) do not depend on  $a$  and  $\sigma^2$ .

### 3.9 Appendix C: Asymptotic covariance for LS algorithm

The LS estimator of  $\phi$  has expression:

$$\hat{\phi}_{\text{LS}} = \operatorname{argmin}_{\phi \in \Phi} \left\{ \sum_{k=1}^K [y_k - f(\phi, \mathbf{x}_k)]^2 + \sum_{l=1}^L [y_{R,l} - f(\phi, \mathbf{x}_{R,l})]^2 \right\} \quad (3.70)$$

where

$$f(\phi, \mathbf{x}_k) = as(\boldsymbol{\theta}, \mathbf{x}_k) \quad (3.71)$$

see also in ( 3.35). By taking the first derivative on ( 3.70) with respect to  $\phi$ , we have

$$\mathbf{0} = \sum_{k=1}^K [y_k - f(\hat{\phi}_{\text{LS}}, \mathbf{x}_k)] \cdot \frac{\partial f(\hat{\phi}_{\text{LS}}, \mathbf{x}_k)}{\partial \phi^T} + \sum_{l=1}^L [y_{R,l} - f(\hat{\phi}_{\text{LS}}, \mathbf{x}_{R,l})] \cdot \frac{\partial f(\hat{\phi}_{\text{LS}}, \mathbf{x}_{R,l})}{\partial \phi^T} \quad (3.72)$$

We now approximate  $f(\hat{\phi}_{\text{LS}}, \mathbf{x}_k)$ ,  $f(\hat{\phi}_{\text{LS}}, \mathbf{x}_{R,l})$ ,  $\partial f(\hat{\phi}_{\text{LS}}, \mathbf{x}_k)/\partial \phi$  and  $\partial f(\hat{\phi}_{\text{LS}}, \mathbf{x}_{R,l})/\partial \phi$  using their first order Taylor series expansion around  $\phi$  and  $\mathbf{x}_k$ , yielding

$$\begin{aligned} f(\hat{\phi}_{\text{LS}}, \mathbf{x}_k) &\approx f(\phi, \mathbf{x}_k) + \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) + \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} \cdot (\mathbf{x}_k - \mathbf{x}_k) \\ f(\hat{\phi}_{\text{LS}}, \mathbf{x}_{R,l}) &\approx f(\phi, \mathbf{x}_{R,l}) + \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) \\ \frac{\partial f(\hat{\phi}_{\text{LS}}, \mathbf{x}_k)}{\partial \phi} &\approx \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} + \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) + \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \mathbf{x}_k^T} \cdot (\mathbf{x}_k - \mathbf{x}_k) \\ \frac{\partial f(\hat{\phi}_{\text{LS}}, \mathbf{x}_{R,l})}{\partial \phi} &\approx \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi} + \frac{\partial^2 f(\phi, \mathbf{x}_{R,l})}{\partial \phi \partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) \end{aligned} \quad (3.73)$$

Therefore

$$\begin{aligned} \mathbf{0} &\approx \sum_{k=1}^K \left[ y_k - f(\phi, \mathbf{x}_k) - \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) - \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} \cdot (\mathbf{x}_k - \mathbf{x}_k) \right] \\ &\quad \cdot \left[ \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} + \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) + \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \mathbf{x}_k^T} \cdot (\mathbf{x}_k - \mathbf{x}_k) \right] \\ &+ \sum_{l=1}^L \left[ y_{R,l} - f(\phi, \mathbf{x}_{R,l}) - \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) \right] \\ &\quad \cdot \left[ \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi} + \frac{\partial^2 f(\phi, \mathbf{x}_{R,l})}{\partial \phi \partial \phi^T} \cdot (\hat{\phi}_{\text{LS}} - \phi) \right] \end{aligned} \quad (3.74)$$

We further simplify the above expression by ignoring the second-order terms:

$$\begin{aligned}
\mathbf{0} \approx & \left( \sum_{k=1}^K \left\{ [y_k - f(\phi, \mathbf{x}_k)] \cdot \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \phi^T} - \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \right\} \right. \\
& \left. - \sum_{l=1}^L \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi^T} \right) \cdot (\hat{\phi}_{\text{LS}} - \phi) \\
& + \sum_{k=1}^K \left\{ [y_k - f(\phi, \mathbf{x}_k)] \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} + [y_k - f(\phi, \mathbf{x}_k)] \cdot \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \mathbf{x}_k^T} (\mathbf{w}_k - \mathbf{x}_k) \right. \\
& \left. - \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} \cdot (\mathbf{w}_k - \mathbf{x}_k) \right\} + \sum_{l=1}^L [y_{R,l} - f(\phi, \mathbf{x}_{R,l})] \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi}
\end{aligned} \tag{3.75}$$

yielding the following approximate formula for the covariance matrix of the LS estimator

$$\begin{aligned}
\text{cov}(\hat{\phi}_{\text{LS}}) \approx & \mathcal{P}(\phi, \mathbf{x})^{-1} \cdot \left\{ \sigma^2 \cdot \mathcal{P}(\phi, \mathbf{x}) + \sigma_u^2 \cdot \sum_{k=1}^K \left[ \sigma^2 \cdot \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \mathbf{x}_k^T} U_k \frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k \partial \phi^T} \right. \right. \\
& \left. \left. + \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k^T} U_k \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \mathbf{x}_k} \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} \right] \right\} \cdot \mathcal{P}(\phi, \mathbf{x})^{-1}
\end{aligned} \tag{3.76}$$

where

$$\mathcal{P}(\phi, \mathbf{x}) = \sum_{k=1}^K \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi^T} + \sum_{l=1}^L \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi} \cdot \frac{\partial f(\phi, \mathbf{x}_{R,l})}{\partial \phi^T} \tag{3.77}$$

Recall

$$\phi = [\boldsymbol{\theta}^T, a]^T \tag{3.78}$$

implying

$$\frac{\partial f(\phi, \mathbf{x}_k)}{\partial \phi} = [a \cdot \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta}^T, s(\boldsymbol{\theta}, \mathbf{x}_k)]^T \tag{3.79}$$

Therefore, for the energy-based acoustic source localization problem with isotropic acoustic-energy attenuation model (see also in ( 3.28), we have

$$\frac{\partial s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \boldsymbol{\theta}} = - \frac{2\alpha\beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-1)}}{1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha}} \cdot (\boldsymbol{\theta} - \mathbf{x}_k) \tag{3.80}$$

and an analogous expression for the reference nodes is obtained by replacing  $\mathbf{x}_k$  by  $\mathbf{x}_{R,l}$ . Also

$$\begin{aligned}
\frac{\partial^2 s(\boldsymbol{\theta}, \mathbf{x}_k)}{\partial \boldsymbol{\theta}, \partial \boldsymbol{\theta}^T} &= \partial \left[ - \frac{2\alpha\beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-1)}}{(1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha})^2} \cdot (\boldsymbol{\theta} - \mathbf{x}_k) \right] / \partial \boldsymbol{\theta}^T \\
&= - \frac{2\alpha\beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-1)}}{(1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha})^2} \cdot I_2 - \frac{4\alpha(\alpha-1)\beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-2)}}{(1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha})^2 \cdot (\boldsymbol{\theta} - \mathbf{x}_k)(\boldsymbol{\theta}, \mathbf{x}_k)^T} \\
&\quad + \frac{8\alpha^2\beta^2 \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2(\alpha-1)}}{(1 + \beta \cdot \|\boldsymbol{\theta} - \mathbf{x}_k\|^{2\alpha})^3} \cdot (\boldsymbol{\theta} - \mathbf{x}_k)(\boldsymbol{\theta} - \mathbf{x}_k)^T
\end{aligned} \tag{3.81}$$

Therefore, for energy-based acoustic source localization problem

$$\frac{\partial^2 f(\phi, \mathbf{x}_k)}{\partial \phi \partial \mathbf{x}_k^T} = \begin{bmatrix} a \cdot \partial^2 s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta} \partial \mathbf{x}_k^T \\ \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \mathbf{x}_k^T \end{bmatrix} = - \begin{bmatrix} a \cdot \partial^2 s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T \\ \partial s(\boldsymbol{\theta}, \mathbf{x}_k) / \partial \boldsymbol{\theta}^T \end{bmatrix} \quad (3.82)$$

## CHAPTER 4. EVENT-REGION ESTIMATION FOR SENSOR NETWORKS UNDER THE POISSON REGIME

We develop a Bayesian method for event-region estimation in large-scale sensor networks under the Poisson regime. We propose a parametric model for the location and shape of the event region and assume that the unknown signal strength within this region is constant. We adopt a fusion architecture where each node in the network makes a decision locally and then conveys it to a fusion center. Both binary and quantized decisions are considered, corresponding to utilizing one or multiple thresholds (respectively) to make the local decisions. MCMC algorithms are derived for simulating from the posterior distributions of the unknown signal, location and shape parameters and for estimating these parameters. Numerical simulations demonstrate the performance of the proposed methods.

### 4.1 Introduction

We consider estimation of localized phenomena<sup>1</sup> using a large number of densely deployed sensor-processor elements (nodes). We adopt a fusion architecture where each node in the network makes a *local* decision and then conveys it to a fusion center (which can be performed efficiently using a mobile access point, as described in [86]). A MAC protocol is used to collect the local decisions, where each node has a probability  $p_{\text{MAC}}$  to transmit its decision successfully. Furthermore, we assume that the fusion center knows the locations of the nodes. The fusion center processes the collected decisions, identifies regions of interest (e.g. areas with elevated signal levels), and estimates their properties. This architecture has been studied recently in [10] where it was assumed that

---

<sup>1</sup>Here, "localized phenomena" correspond to the spatial phenomena that affect only parts of the network.

1. the nodes are randomly deployed following a homogeneous Poisson point process;
2. each node's measurement is corrupted by i.i.d. additive noise.

It was further assumed in [10] that the nodes make *binary local decisions*. Under the above assumptions, "alarming" the sensors (i.e. deciding that the signal is present) is equivalent to a location-dependent *thinning* of the original sensor distribution. Consequently, the alarmed nodes form a *heterogeneous Poisson point process*, hence the name *Poisson regime* [10]. In [10], an *asymptotic* event-region detector has been developed for the Poisson regime assuming that the location of the event region is *known* and that the shape of the spatial signal within the event region is *known* up to a scaling (signal-strength) constant.

In this Chapter, we adopt the *exact* heterogeneous Poisson process model for the alarmed nodes and develop a Bayesian method for estimating the *location* of the event region and the *strength and shape* of the underlying spatial signal. We also generalize the Poisson regime to account for the scenario where the nodes utilize multiple local thresholds to *quantize* the sensor measurements and then send the resulting quantized (multi-bit) data to the fusion center.

In Section 4.2, we introduce the measurement and event region models and prior specifications. In Section 4.3, we develop Bayesian methods for simulating and estimating the event-region parameters. In Section 4.4, we evaluate the performance of the proposed methods via numerical simulations. Concluding remarks are given in Section 4.5.

## 4.2 Measurement and event-region models

We introduce the measurement-error models for single and multiple thresholds in Section 4.2.1 and describe the prior specifications and specific event-region and noise models in Section 4.2.2.

### 4.2.1 Measurement-error model

We assume the measurement  $y_n$  obtained at node  $n$  follows the signal-plus-noise model<sup>2</sup>:

$$y_n = \theta s(\varphi, \mathbf{x}_n) + e_n \quad (4.1)$$

---

<sup>2</sup>Note that  $y_n$  may also be a sufficient statistics computed from the data collected at node  $n$ .

where the first term represents the parameterized signal of interest and  $e_n$  denotes the zero-mean additive noise. Here,  $\theta s(\boldsymbol{\varphi}, \mathbf{x})$  models the signal of interest as a function of the node location (we assume each node knows its own location)  $\mathbf{x}$ ;  $\theta$  and  $\boldsymbol{\varphi}$  are the unknown signal strength parameter and vector of signal location and shape parameters. Then, we denote the full set of parameters as:

$$\boldsymbol{\phi} = [\theta, \boldsymbol{\varphi}^T]^T \quad (4.2)$$

We first review the classical Poisson regime where the nodes make binary decisions and then extend it to the multiple threshold scenario.

**Single threshold:** If node  $n$  makes its decision by comparing  $y_n$  with a threshold  $\tau$ , with the assumption 1-2 hold, then the alarmed node locations<sup>3</sup> follow a heterogeneous Poisson process with intensity at location  $\mathbf{x}$ :

$$\lambda_{\boldsymbol{\phi}}(\mathbf{x}) = \lambda_0 \cdot p_{\text{MAC}} \cdot P_{\tau}[\theta s(\boldsymbol{\varphi}, \mathbf{x})] = \lambda \cdot P_{\tau}[\theta s(\boldsymbol{\varphi}, \mathbf{x})] \quad (4.3)$$

where  $\lambda_0$  is the (known) intensity of the initial Poisson process (describing the node deployment),  $p_{\text{MAC}}$  is the (known) constant probability of successful transmission (as in [10]), and  $P_{\tau}[\theta s(\boldsymbol{\varphi}, \mathbf{x})]$  is the probability of detecting the signal at location  $\mathbf{x}$ . Consequently,  $P_{\tau}[0]$  is the false-alarm probability of the local node decisions. The likelihood that there are  $N$  alarmed nodes within region  $A$  at locations  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  is

$$l_S(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N | \boldsymbol{\phi}) = \left\{ \prod_{n=1}^N P_{\tau}[\theta s(\boldsymbol{\varphi}, \mathbf{x}_n)] \right\} \cdot \exp \left\{ -\lambda \cdot \int_A P_{\tau}[\theta s(\boldsymbol{\varphi}, \mathbf{u})] d\mathbf{u} \right\} \quad (4.4)$$

**Multiple threshold:** Suppose now that each node  $n$  compares its measurement  $y_n$  with  $K$  thresholds:

$$\tau_1 < \tau_2 < \dots < \tau_K \quad (4.5)$$

and is *alarmed* if  $y_n$  is larger than the smallest threshold  $\tau_1$ . For notational convenience, we also define  $\tau_{K+1} = +\infty$ , implying that  $P_{\tau_{K+1}}(x) \equiv 0$  for all real arguments  $x$ . Each alarmed node  $n$  reports a number  $k(n) \in \{1, 2, \dots, K\}$  corresponding to the interval  $[\tau_{k(n)}, \tau_{k(n)+1})$  which contains the measurement  $y_n$ , i.e.

$$y_n \in [\tau_{k(n)}, \tau_{k(n)+1}) \quad (4.6)$$

<sup>3</sup>A node  $n$  is alarmed if its measurement  $y_n$  is higher than the threshold  $\tau$ .

If the Poisson-regime assumptions 1-2 and measurement-error model ( 4.1) hold, then the  $N$  alarmed nodes at locations  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  with decisions  $k(1), k(2), \dots, k(N)$  form a *multivariate heterogeneous Poisson process* with the following likelihood function:

$$l_M(k(1), k(2), \dots, k(N), \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N | \phi) \propto \left( \prod_{n=1}^N \{P_{\tau_{k(n)}}[\theta s(\boldsymbol{\varphi}, \mathbf{x}_n)] - P_{\tau_{k(n)+1}}[\theta s(\boldsymbol{\varphi}, \mathbf{x}_n)]\} \right) \cdot \exp \left\{ -\lambda \cdot \int_A P_{\tau_1}[\theta s(\boldsymbol{\varphi}, \mathbf{u})] d\mathbf{u} \right\} \quad (4.7)$$

In the following section, we describe a simple circular event-region model and Gaussian noise model that will be used in the following discussion. The proposed framework is applicable to other event-region and noise models as well.

#### 4.2.2 Event-region and noise models and prior specifications

Consider a two-dimensional (2-D) sensor network deployed to estimate an event region  $R(\boldsymbol{\varphi})$  having a circular shape with radius  $r$  and center described by Cartesian coordinates:  $\mathbf{z} = [z_1, z_2]^T$ . Hence, the vector of event-region shape and location parameters is

$$\boldsymbol{\varphi} = [r, \mathbf{z}^T]^T \quad (4.8)$$

We assume that the signal is constant (equal to  $\theta$ ) within the event region and zero outside. Equivalently,

$$s(\boldsymbol{\varphi}, \mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in R(\boldsymbol{\varphi}) \\ 0, & \mathbf{x} \notin R(\boldsymbol{\varphi}) \end{cases} = \begin{cases} 1, & \|\mathbf{x} - \mathbf{z}\| \leq r \\ 0, & \|\mathbf{x} - \mathbf{z}\| > r \end{cases} \quad (4.9)$$

where  $\|\mathbf{x} - \mathbf{z}\| = \sqrt{(\mathbf{x} - \mathbf{z})^T (\mathbf{x} - \mathbf{z})}$  denotes the Euclidean distance between  $\mathbf{x}$  and  $\mathbf{z}$  (in Cartesian coordinates).

We adopt the Gaussian noise model with known variance  $\sigma^2$ , implying

$$P_\tau(\theta) = \Phi\left(\frac{\theta - \tau}{\sigma}\right) \quad (4.10)$$

where  $\Phi(\cdot)$  denotes the cumulative distribution function of the standard normal random variable.

**Prior specifications:** We assume that  $\theta$  and  $\boldsymbol{\varphi}$  are independent *a priori*:

$$\pi_\phi(\phi) = \pi_\theta(\theta) \cdot \pi_\varphi(\boldsymbol{\varphi}) \quad (4.11)$$

and adopt the following simple uniform-distribution prior probability density functions for  $\theta$  and  $\varphi$ :

$$\begin{aligned}
\pi_{\theta}(\theta) &= \text{uniform}(\theta_{\text{MIN}}, \theta_{\text{MAX}}) \\
\pi_{\varphi}(\varphi) &= \pi_r(r) \cdot \pi_{z_1}(z_1) \cdot \pi_{z_2}(z_2) \\
\pi_r(r) &= \text{uniform}(r_{\text{MIN}}, r_{\text{MAX}}) \\
\pi_{z_1}(z_1) &= \text{uniform}(z_{1,\text{MIN}}, z_{1,\text{MAX}}) \\
\pi_{z_2}(z_2) &= \text{uniform}(z_{2,\text{MIN}}, z_{2,\text{MAX}})
\end{aligned} \tag{4.12}$$

### 4.3 Bayesian analysis

The goals of our analysis are to estimate the unknown parameters  $\phi$  under the measurement model and prior specifications in Section 4.2. In Section 4.3.1 (below), we construct a method for drawing samples from the posterior distributions of the parameters  $\phi$ . Then, we utilize these samples to compute approximate MMSE estimates of  $\phi$  (Section 4.3.2).

#### 4.3.1 Simulating the parameters from posterior pdfs

**Single threshold:** Our inference about  $\phi$  is based on its posterior pdf:

$$p_S(\phi | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \propto l_S(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N | \phi) \cdot \pi_{\phi}(\phi) \tag{4.13}$$

Under the circular even-region model in Section 4.2.2, the above pdf simplifies to:

$$p_S(\phi | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \propto q_S(\phi) \cdot \pi_{\phi}(\phi) \tag{4.14}$$

where

$$q_S(\phi) = \left( \frac{P_{\tau}[\theta]}{P_{\tau}[0]} \right)^{n(\varphi)} \cdot \exp\{-\lambda \cdot r^2 \pi \cdot (P_{\tau}[\theta] - P_{\tau}[0])\} \tag{4.15}$$

is the normalized likelihood (i.e. likelihood ratio) and  $n(\varphi)$  denotes the number of alarmed nodes in the event region  $R(\varphi)$ . We now outline our proposed scheme for simulating from  $p(\phi | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ . We utilize *shrinkage slice sampler* [87]. First, define the initial (largest) hyperrectangle:

$$\begin{aligned}
\theta_L &= \theta_{\text{MIN}}, \quad \theta_U = \theta_{\text{MAX}}, \quad r_L = r_{\text{MIN}}, \quad r_U = r_{\text{MAX}}, \\
z_{i,L} &= z_{i,\text{MIN}}, \quad z_{i,U} = z_{i,\text{MAX}} \quad i \in \{1, 2\}
\end{aligned} \tag{4.16}$$

which coincide with the parameter space of  $\phi$ . We generate  $\phi^{(t)}$  from  $p_S(\phi|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$  as follows:

1. Draw an auxiliary random variable  $u^{(t)}$  from  $\text{uniform}(0, q_S(\phi^{(t-1)}))$  pdf;
2. Draw
  - $\theta$  from  $\text{uniform}(\theta_L, \theta_U)$  pdf,
  - $r$  from  $\text{uniform}(r_L, r_U)$  pdf, and
  - $z_i$  from  $\text{uniform}(z_{i,L}, z_{i,U})$  pdfs for  $i \in \{1, 2\}$

yielding  $\phi = [\theta, r, z_1, z_2]^T$ .

3. Check if  $\phi$  is *within the slice*, i.e.

$$q_S(\phi) \geq u^{(t)} \quad (4.17)$$

If ( 4.17) holds, return

$$\phi^{(t)} = [\theta^{(t)}, r^{(t)}, z_1^{(t)}, z_2^{(t)}]^T = \phi \quad (4.18)$$

and exit the loop. If ( 4.17) does not hold, then *shrink the hyperrectangle*<sup>4</sup>:

- If  $\theta < \theta^{(t-1)}$ , set  $\theta_L = \theta$ ; else if  $\theta > \theta^{(t-1)}$ , set  $\theta_U = \theta$ .
- If  $r < r^{(t-1)}$ , set  $r_L = r$ ; else if  $r > r^{(t-1)}$ , set  $r_U = r$ .
- If  $z_1 < z_1^{(t-1)}$ , set  $z_{1,L} = z_1$ ; else if  $z_1 > z_1^{(t-1)}$ , set  $z_{1,U} = z_1$ .
- If  $z_2 < z_2^{(t-1)}$ , set  $z_{2,L} = z_2$ ; else if  $z_2 > z_2^{(t-1)}$ , set  $z_{2,U} = z_2$ .
- Go back to 2.

Cycle between Steps 1-3 until the desirable number of samples has been collected (after discarding the samples from the burn-in period). This scheme produces a *Markov chain*  $\phi^{(0)}, \phi^{(1)}, \phi^{(2)}, \dots$  with stationary distribution equal to the posterior pdf  $p_S(\phi|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ .

<sup>4</sup>Here, the hyperrectangles *shrink toward*  $\phi^{(t-1)} = [\theta^{(t-1)}, r^{(t-1)}, z_1^{(t-1)}, z_2^{(t-1)}]$ , which is clearly in the slice, see Step 1.

Since the evaluation of  $q_S(\phi)$  may cause a floating-point underflow, it is often safer to utilize its logarithm. In the  $t$ th step of the slice sampler, we then compute  $\ln q_S(\phi^{(t-1)}) - \epsilon^{(t)}$ , where  $\epsilon^{(t)}$  is exponentially distributed with mean one, and say that  $\phi$  is in the slice if

$$\ln q_S(\phi) \geq \ln q_S(\phi^{(t-1)}) - \epsilon^{(t)} \quad (4.19)$$

which is equivalent to ( 4.17).

**Multiple thresholds:** Our inference about  $\phi$  is based on its posterior pdf:

$$p_M(\phi|k(1), k(2), \dots, k(N), \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \propto l_M(k(1), k(2), \dots, k(N), \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N|\phi) \cdot \pi_\phi(\phi) \quad (4.20)$$

Under the circular event-region model in Section 4.2.2, the above pdf simplifies to:

$$p_M(\phi|k(1), k(2), \dots, k(N), \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \propto q_M(\phi) \cdot \pi_\phi(\phi) \quad (4.21)$$

where

$$q_M = \left[ \prod_{k=1}^K \left( \frac{P_{\tau_k}[\theta] - P_{\tau_{k+1}}[\theta]}{P_{\tau_k}[0] - P_{\tau_{k+1}}[0]} \right)^{n_k(\varphi)} \right] \quad (4.22)$$

and  $n_k(\varphi)$ ,  $k = 1, 2, \dots, K$  denotes the number of alarmed nodes in the event region  $R(\phi)$  reporting measurements in the interval  $[\tau_k, \tau_{k+1})$ .

To simulate from the posterior pdf in ( 4.20), we apply the above shrinkage slice sampler, with  $q_M(\phi)$  computed using ( 4.22) [which generalizes ( 4.15) used in the single-threshold scenario]. the modification in ( 4.19) applies to the multiple threshold scenario as well.

### 4.3.2 Estimating the event-region parameters

Once we have collected enough samples, we can estimate the posterior mean of  $\phi$  simply by averaging the last  $T$  draws:

$$E[\phi|\mathbf{y}] \approx \hat{\phi} = \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} \phi^{(t)} \quad (4.23)$$

where  $t_0$  defines the burn-in period. Note that  $\hat{\phi}$  is an approximate MMSE estimate of  $\phi$ .

#### 4.4 Numerical examples

We consider a 2-D network with nodes randomly (uniformly) placed in a square region measuring  $50 \times 50 \text{ m}^2$ . In Figure 4.1(a), we show the noiseless field  $\theta \cdot s(\boldsymbol{\varphi}, \mathbf{x})$  as a function of location  $\mathbf{x}$ ; in Figure 4.1(b), we show the corresponding heterogeneous Poisson field, simulated using  $K = 1$  threshold at  $\tau = 1.64$  (corresponding to the false alarm probability  $P_\tau[0] = 5\%$ ), noise variance  $\sigma^2 = 1$ , and the following event-region parameters:  $\theta = 1$ ,  $r = 5$ ,  $z_1 = 15$ ,  $z_2 = 15$ , and initial node density and successful transmission probability that yield  $\lambda = \lambda_0 \cdot p_{\text{MAC}} = 2$  nodes per  $\text{m}^2$ . We also selected the following prior parameters:  $\theta_{\text{MIN}} = -3$ ,  $\theta_{\text{MAX}} = 5$ ,  $z_{1,\text{MIN}} = 0$ ,  $z_{1,\text{MAX}} = 30$ ,  $z_{2,\text{MIN}} = 0$ ,  $z_{2,\text{MAX}} = 30$ ,  $r_{\text{MIN}} = 3$ ,  $r_{\text{MAX}} = 10$ .

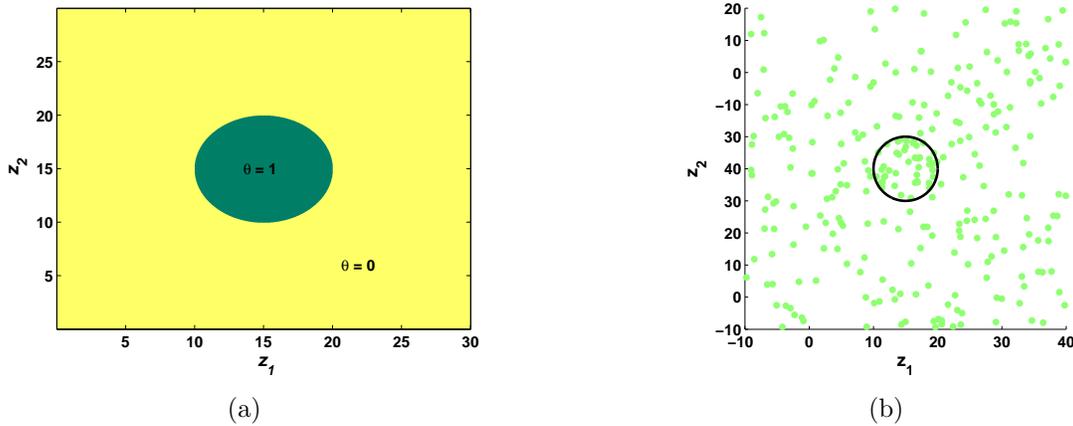


Figure 4.1 Noiseless field (left) and a realization of the heterogeneous Poisson field formed by the alarmed nodes, for  $\theta = 1$  and  $\lambda = 2$  (right).

We first analyze the field in Figure 4.1 using the shrinkage slice sampler in Section 4.3.1, where we discarded  $t_0 = 5000$  burn-in samples and used  $T = 5000$  samples for estimating the event-region parameters, see ( 4.23). In Figure 4.2, we show the contours of

- an estimated event region obtained using the approximate MMSE estimates in ( 4.23) (dashed line) and
- exact event region (full line)

which demonstrate the remarkably good performance of the proposed estimator.

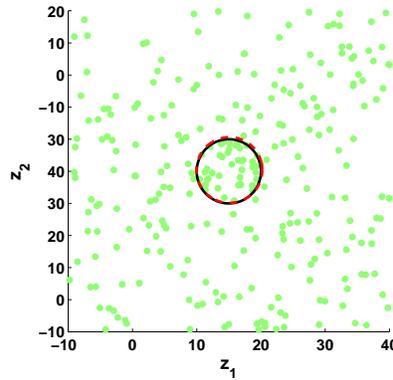


Figure 4.2 Estimated event region obtained using the approximate MMSE estimates of the location and shape parameters (dashed line) and exact event region (full line).

In the second simulation example, we study the average MSE performance of the approximate MMSE method in (4.23), calculated using 150 independent trials. In each trial, we generated independent node locations and measurement noise realizations and estimated the signal parameters using the approximate MMSE method in (4.23) with  $t_0 = 200$  and  $T = 1800$  samples. We consider the noiseless field in Figure 4.1(a) with the same event-region and noise parameters as in the previous example. However, in this example, we have chosen a more challenging deployment scenario with a smaller density:  $\lambda = 1$  nodes per  $\text{m}^2$ . We consider both single- and multiple-threshold scenario. In the single-threshold case (i.e.  $K = 1$ ), we selected the threshold  $\tau$  to guarantee the false-alarm probability of  $P_\tau[0] = 5\%$ . In the multiple-threshold case, we chose  $K = 4$  thresholds to satisfy  $P_{\tau_1}[0] = 5\%$ ,  $P_{\tau_2}[0] = 3.75\%$ ,  $P_{\tau_3}[0] = 2.5\%$ ,  $P_{\tau_4}[0] = 1.25\%$ .

Figure 4.3(a) shows the average MSEs for the approximate MMSE estimates of the signal strength  $\theta$  as a function of  $\theta$ . In the single-threshold scenario, the estimation performance deteriorates for  $\theta > 1.5$  since it is not possible to identify large signal strength using one bit only<sup>5</sup>. Figure 4.3(b), shows the average MSE for the approximate MMSE estimate of the event-region radius  $r$ . The accuracy of estimating  $r$  improves as  $\theta$  increases and is limited by

<sup>5</sup>A similar problem occurs for  $K = 4$  threshold for very large values of  $\theta$ .

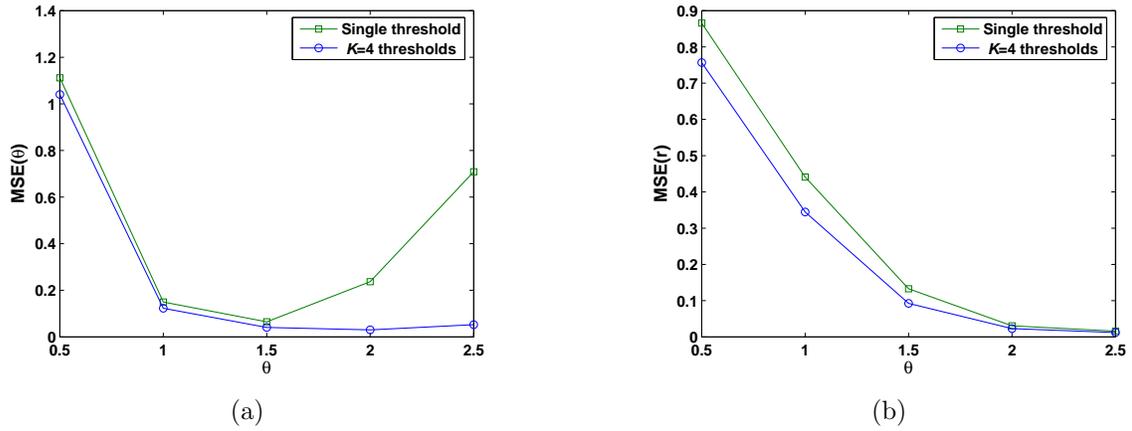


Figure 4.3 Average MSEs for approximate MMSE estimates of the signal strength  $\theta$  (left) and event-region radius  $r$  (right), as function of  $\theta$ , for  $K \in \{1, 4\}$  and  $\lambda = 1$ .

the node density  $\lambda$ .

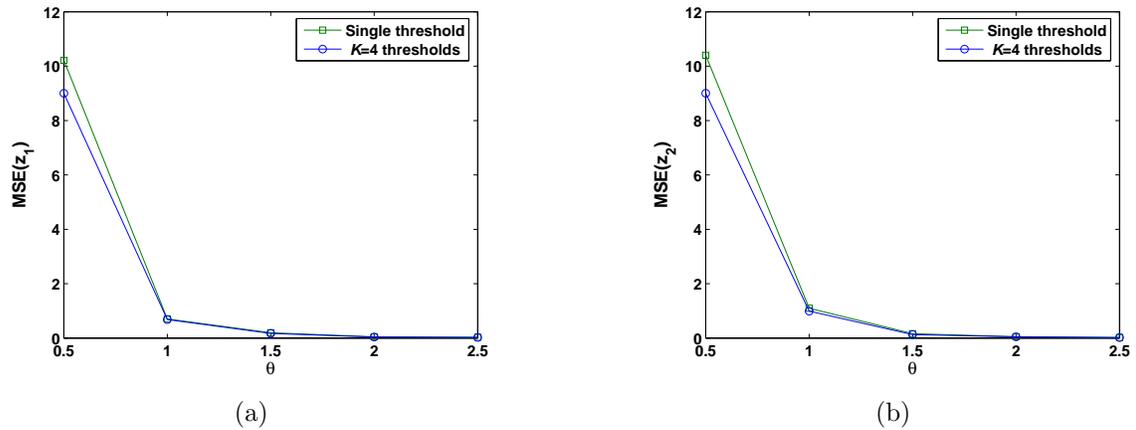


Figure 4.4 Average MSEs for approximate MMSE estimates of the event-region center coordinates:  $z_1$  (left) and  $z_2$  (right), as function of signal strength  $\theta$ , for  $K \in \{1, 4\}$  and  $\lambda = 1$ .

Figure 4.4 shows the average MSEs for the approximate MMSE estimates of the event-region center coordinates. The accuracies of estimating  $z_1$  and  $z_2$  improve as  $\theta$  increases and are limited by the node density  $\lambda$ .

#### 4.5 Conclusion remarks

We proposed Bayesian methods for event-region estimation in large-scale sensor networks under the univariate and multivariate Poisson regimes (corresponding to utilizing one or multiple thresholds to make local decisions at the nodes).

## CHAPTER 5. MEAN-FIELD ESTIMATION AND DETECTION IN CORRELATED GAUSSIAN RANDOM FIELD

We propose distributed methods for estimating and detecting the mean of a correlated Gaussian random field observed by a sensor network. The random-field correlations are assumed to follow a CAR model. First, a distributed ML estimator of the mean field is derived. We then develop batch and sequential detectors for testing the hypothesis that the mean field is greater than a specified level. We also derive exact and approximate performance measures for our methods. Numerical examples demonstrate the performance of the proposed approach.

### 5.1 Introduction

In large-scale wireless sensor networks, sensor-processor elements (nodes) will be densely deployed to monitor the environment at close range with high spatial and temporal resolutions [88]. Due to the high node density, the physical phenomena monitored by sensor networks (e.g. concentration of a chemical, temperature, or water contamination) usually yield observations that are highly correlated in space [34, 36, 37]. This spatial correlation has been utilized for (i) data aggregation, compression, routing, and querying, (ii) localization, and (iii) MAC protocol design, see [34, 36, 37] and references therein.

In this chapter, we first derive a *distributed* ML method for estimating the mean of a spatially correlated Gaussian random field in sensor-network environments where the field correlations follow a *conditional autoregressive* model [26, 89–91]. We then develop batch and sequential detectors for testing the hypothesis that the mean field is greater than a specified level. Gaussian conditional autoregressions have been used in image processing for texture analysis and many other applications, e.g. agricultural field experiments, human geography,

geographical epidemiology, and astronomy, see [90] and references therein. The CAR model can also approximate common spatial correlation models, such as exponential, squared exponential, spherical, and Matern, see [91]. We derive estimators and detectors for the general CAR model and then discuss a simple special case. Our simulation results demonstrate the importance of incorporating spatial field correlations into the design of detection algorithms in sensor-network environments.

In Section 5.2, we introduce the measurement model and in Section 5.3 we propose a distributed ML estimator of the mean field (Section 5.3.1). Mean-field detection is discussed in Section 5.4 where batch and sequential detectors are developed (Sections 5.4.1 and 5.4.2). In Section 5.5, we briefly discuss calibration of the random-field parameters under a simple connectivity-based spatial-dependence model and apply it to a rainfall precipitation data set. We evaluate the performance of the proposed methods using simulated and real-data examples (see Sections 5.3.2, 5.4.3, and 5.5.1).

## 5.2 Measurement model

Assume that the region of interest contains  $N$  nodes at locations  $\mathbf{s}_i$ ,  $i = 1, 2, \dots, N$  and denote the measurements collected at these nodes by  $z(\mathbf{s}_i)$ ,  $i = 1, 2, \dots, N$ . We model the random-field values  $z(\mathbf{s}_i)$ ,  $i = 1, 2, \dots, N$  by specifying their *full conditional* pdfs [26]:

$$p(z(\mathbf{s}_i) | \{z(\mathbf{s}_j) : j \neq i\}) = \frac{1}{(2\pi\tau_i^2)^{1/2}} \cdot \exp\left\{-\frac{1}{2\tau_i^2} \cdot [z(\mathbf{s}_i) - \mu_i]^2\right\} \quad (5.1)$$

where  $\mu_i = E[z(\mathbf{s}_i) | \{z(\mathbf{s}_j) : j \neq i\}]$  and  $\tau_i^2 = \text{var}[z(\mathbf{s}_i) | \{z(\mathbf{s}_j) : j \neq i\}]$  are the corresponding conditional mean and variances. Assuming pairwise-only dependence between the spatial locations and constant marginal mean of the field, we obtain the following linear model for the conditional means:

$$\mu_i = \alpha + \sum_{j=1}^N c_{i,j} [z(\mathbf{s}_j) - \alpha] \quad (5.2)$$

where  $\alpha$  is the (unknown) *marginal mean* of the random field and  $c_{i,j}$  are (known) *spatial dependence* parameters satisfying

$$c_{i,j}\tau_j^2 = c_{j,i}\tau_i^2, \quad c_{i,i} = 0 \quad (5.3)$$

see [26]. If a *Markovian assumption* is further imposed, then  $c_{i,j} = 0$  if the location  $\mathbf{s}_i$  and  $\mathbf{s}_j$  are not neighbors. We assume that the *neighborhood* of a node  $i$  [denote by  $\mathcal{N}(i)$ ] consists of all nodes  $j \in \{1, 2, \dots, N\}$  that are different from  $i$  and within a *cutoff distance*  $d$  from that node, i.e.

$$\mathcal{N}(i) = \{j : \|\mathbf{s}_i - \mathbf{s}_j\| \leq d \text{ and } j \neq i\} \quad (5.4)$$

where  $\|\mathbf{s}_i - \mathbf{s}_j\|$  denotes the (Euclidean) distance between the locations  $\mathbf{s}_i$  and  $\mathbf{s}_j$ .

The results in [26] imply that the random-field vector  $\mathbf{z} = [z(\mathbf{s}_1), z(\mathbf{s}_1), \dots, z(\mathbf{s}_1)]^T$  is multivariate Gaussian with mean and covariance  $E[\mathbf{z}] = \alpha \mathbf{1}_N$  and  $\text{cov}[\mathbf{z}] = (I_N - C)^{-1}D$ , where  $C$  is an  $N \times N$  matrix whose  $(i, j)$  element is  $c_{i,j}$  and  $D = \text{diag}\{\tau_1^2, \tau_2^2, \dots, \tau_N^2\}$ . The condition (5.3) ensures that  $\text{cov}[\mathbf{z}]$  is a symmetric matrix. The spatial dependence coefficients  $c_{i,j}$  need to be carefully chosen to ensure that  $\text{cov}[\mathbf{z}]$  is a positive definite matrix. To achieve appreciable correlations among the measurements at neighboring nodes, we typically need to select the  $c_{i,j}$  coefficients close to a boundary of the set of allowed values, see also [90] and Section 5.3.2.

Define the following recursive expressions:

$$\begin{aligned} q_1(n) &= \sum_{i=1}^n z(\mathbf{s}_i) \cdot \frac{1 - \sum_{j \in \mathcal{N}(i)} c_{i,j}}{\tau_i^2} = q_1(n-1) + z(\mathbf{s}_n) \cdot \frac{1 - \sum_{j \in \mathcal{N}(n)} c_{n,j}}{\tau_n^2} \\ q_0(n) &= \sum_{i=1}^n \frac{1 - \sum_{j \in \mathcal{N}(i)} c_{i,j}}{\tau_i^2} = q_0(n-1) + \frac{1 - \sum_{j \in \mathcal{N}(n)} c_{n,j}}{\tau_n^2} \end{aligned} \quad (5.5)$$

for  $n = 1, 2, \dots, N$ , where  $q_1(0) = q_0(0) = 0$ . The above recursive formulas will be used in Section 5.3 and 5.4 to implement the estimation and detection of the mean field  $\alpha$ . If the random field  $\mathbf{z}$  is *not* Markov, then the above recursions still hold, with the neighborhood  $\mathcal{N}(i)$  generally covering all the nodes in the network except the  $i$ th.

**Spatially white random field:** If the random field  $\mathbf{z}$  is *spatially white*, then  $C = 0$  and  $\tau_i^2 = \sigma^2$ , where  $\sigma^2$  is the marginal variance of the observations. Then (5.5) simplifies to

$$q_1(n) = \frac{\sum_{i=1}^n z(\mathbf{s}_i)}{\sigma^2} = q_1(n-1) + \frac{z(\mathbf{s}_n)}{\sigma^2}, \quad q_0(n) = \frac{n}{\sigma^2} \quad (5.6)$$

### 5.3 Mean-field estimation

Our goal is to estimate the marginal mean  $\alpha$  of the measured field assuming that  $D$  and  $C$  are *known*. Under the above model, the log likelihood of  $\alpha$  is

$$L(\alpha) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\text{cov}[\mathbf{z}]| - \frac{1}{2}(\mathbf{z} - \mathbf{1}_N\alpha)^T(\text{cov}[\mathbf{z}])^{-1}(\mathbf{z} - \mathbf{1}_N\alpha) \quad (5.7)$$

Then, the ML estimate of  $\alpha$  easily follows by maximizing  $L(\alpha)$ :

$$\hat{\alpha} = \frac{\mathbf{z}^T(\text{cov}[\mathbf{z}])^{-1}\mathbf{1}_N}{\mathbf{1}_N^T(\text{cov}[\mathbf{z}])^{-1}\mathbf{1}_N} = \frac{\mathbf{z}^T(I_N - C)\mathbf{1}_N}{\mathbf{1}_N^T(I_N - C)\mathbf{1}_N} = \frac{q_1(N)}{q_0(N)} \quad (5.8)$$

where  $q_1(n)$  and  $q_2(n)$  have been defined in 5.5. Note that  $\hat{\alpha}$  is an unbiased estimator of  $\alpha$ . Consequently, its mean-square error (MSE) is equal to its variance and is easily computed as

$$\text{MSE}(\hat{\alpha}) = \text{var}[\hat{\alpha}] = \frac{1}{\mathbf{1}_N^T(\text{cov}[\mathbf{z}])^{-1}\mathbf{1}_N} = \frac{1}{q_0(N)} \quad (5.9)$$

**Sample-mean estimator:** We now derive the performance of the commonly used sample-mean estimator of  $\alpha$ :

$$\bar{z} = \frac{1}{N} \cdot \mathbf{1}_N^T \mathbf{z} = \frac{1}{N} \cdot \sum_{i=1}^N z(\mathbf{s}_i) \quad (5.10)$$

which ignores the spatial dependence in the data. Here,  $\bar{z}$  is Gaussian and unbiased with mean-square error  $\text{MSE}_{\bar{z}}(\bar{z}) = \text{var}[\bar{z}] = \mathbf{1}_N^T \text{cov}[\mathbf{z}] \mathbf{1}_N / N^2$ ; recall that  $\text{cov}[\mathbf{z}] = (I_N - C)^{-1}D$ . Straightforward application of the Cauchy-Schwartz inequality yields

$$\text{MSE}(\bar{z}) \geq \text{MSE}(\hat{\alpha}) \quad (5.11)$$

implying that the MSE of the sample-mean estimator of  $\alpha$  is *always* larger than or equal to the MSE of the corresponding ML estimator.

#### 5.3.1 Distributed implementation

We now present a scheme for sequential updating quantities necessary for estimating the marginal mean of the field  $\alpha$ . Denote by  $n$  the node "visited" in the  $n$ th step of our scheme. This node *recursively* computes  $q_0(n)$  and  $q_1(n)$  using (5.5) and passes them to the next node (indexed by  $n+1$ ). After visiting all  $N$  nodes, the ML estimate of  $\alpha$  is computed using (5.8).

If the Markovian dependence assumption holds, then we can easily adapt the changes in the network topology. To update  $q_1(n)$  and  $q_0(n)$  using ( 5.5), the  $n$ th node utilizes only  $c_{n,j}$ ,  $j \in \mathcal{N}(n)$  (i.e. those coefficients that correspond to its neighbors) and its conditional variance  $\tau_n^2$ . If the *fusion center* (i.e. node  $N$ ) knows all spatial-dependence coefficients  $c_{i,j}$  and conditional variances  $\tau_i^2$ , then it can compute  $q_0(N)$  *locally*, implying that the nodes need to pass  $q_1(n)$  only. Note that  $q_1(N)$  is the *sufficient statistic* for estimating  $\alpha$ . If the Markovian dependence assumption does not hold, then adapting to changes in the network topology requires that all nodes know all spatial-dependence coefficients [to be able to update  $q_1(n)$ , see ( 5.5)].

### 5.3.2 Numerical example 1: mean field estimation

To assess the performance of the proposed methods, we consider a sensor network containing  $N = 1000$  nodes randomly (uniformly) distributed on a  $50\text{m} \times 50\text{m}$  grid with  $1\text{m}$  spacing between the potential node locations. In all simulation examples, we have selected the cutoff distance  $d = 2.3\text{m}$ . Also, the nodes are visited row by row starting from the upper left corner of the network. Here, we assume that the conditional variances  $\tau_i^2$  are constant:

$$\tau_i^2 = \tau^2, \quad i = 1, 2, \dots, N \quad (5.12)$$

implying that  $c_{i,j} = c_{j,i}$  (i.e.  $C$  is a symmetric matrix) and  $D = \tau^2 I_N$ . The above assumptions are commonly made in the literature, see e.g. [26, 89]. Let us further choose

$$C = cW \quad (5.13)$$

where  $c$  is a constant describing the level of spatial dependence among the observations and  $W$  is the *connectivity matrix* whose  $(i, j)$  element is  $w_{i,j} = \begin{cases} 1, & j \in \mathcal{N}(i) \\ 0, & j \notin \mathcal{N}(i) \end{cases}$ . Since we wish to model positive spatial dependence among the observations, we focus on the case where  $c > 0$ . Denote by  $\lambda_{\text{MAX}} = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$  the ordered eigenvalues of  $W$ . Then, to ensure that  $\text{cov}[\mathbf{z}] = \tau^2 \cdot (I - cW)^{-1}$  is positive definite,  $c$  should be such that

$$0 \leq c < \frac{1}{\lambda_{\text{MAX}}} \quad (5.14)$$

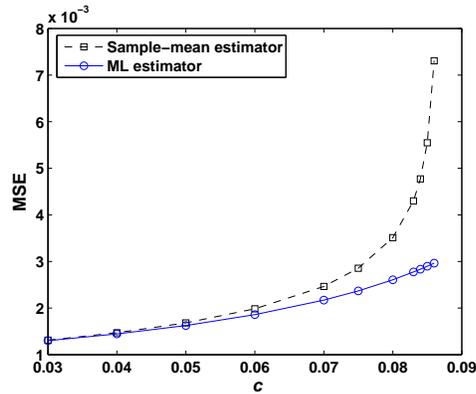


Figure 5.1 MSEs for the ML and sample-mean estimates of the mean field  $\alpha$  as functions of the spatial-dependence level  $c$ .

In this example, we have chosen the conditional variance parameter  $\tau^2 = 1$ . For  $c = 0.070, 0.083$  and  $0.086$ , the corresponding maximum correlation coefficients [over all pairs of observations] are  $0.25, 0.61$  and  $0.85$  (respectively) and  $1/\lambda_{\text{MAX}} = 0.0872$ . Observe that  $c$  needs to be very close to  $1/\lambda_{\text{MAX}}$  to achieve appreciable correlations in the data. In Figure 5.1, we compare the MSEs of the ML and sample-mean estimates of  $\alpha$  as functions of the spatial-dependence level  $c$ . The ML mean-field estimate significantly outperforms the corresponding sample-mean estimate if the maximum field correlation between neighbors is larger than  $0.61$  (corresponding to  $c$  larger than  $0.083$ ).

#### 5.4 Mean-field detection

Consider the detection problem where we wish to decide if the marginal mean  $\alpha$  is higher than a specified level. Without loss of generality, we set this level to zero. Therefore, our goal is to test the hypothesis  $H_0 : \alpha = 0$  (mean-field signal absent) versus the alternative  $H_1 : \alpha > 0$  (mean-field signal present). In the following, we develop batch and sequential detectors for the above problems (Section 5.4.1 and 5.4.2, respectively) assuming that the field covariance parameters  $\tau^2$  and  $c_{i,j}$ ,  $i, j = 1, 2, \dots, N$  are known.

### 5.4.1 Batch detector

Assuming that all  $N$  observations are available, the likelihood-ratio test for the above problem simplifies to declaring the signal presence if

$$q_1(N) > \xi \quad (5.15)$$

and signal absence otherwise. Here,  $q_1(N)$  is computed using the recursion in ( 5.5) and  $\xi$  is a threshold chosen to guarantee a specified probability of false alarm  $P_{FA,d}$ . The threshold  $\xi$  is computed by solving

$$\Phi\left(\frac{\xi}{\sqrt{q_0(N)}}\right) = 1 - P_{FA,d} \quad (5.16)$$

where  $\Phi(\cdot)$  denotes the cumulative distribution function of the standard normal random variable. For the mean-signal level  $\alpha > 0$  the probability of detection achieved by the above test is given by

$$P_D(\alpha) = \Phi\left(\frac{\alpha q_0(N) - \xi}{\sqrt{q_0(N)}}\right) = \Phi(\alpha\sqrt{q_0(N)} - \Phi^{-1}(1 - P_{FA,d})) \quad (5.17)$$

where the second equality in ( 5.17) is obtained by using ( 5.16). The results ( 5.16) and ( 5.17) follow from the fact that  $q_1(N)$  is Gaussian with mean  $\alpha q_0(N)$  and variance  $q_0(N)$ .

### 5.4.2 Sequential detector

We describe a sequential-testing approach to detecting the mean-field signal where the nodes are visited sequentially

- until, for some  $n \in \{1, 2, \dots, N\}$

$$q_1(n) > A + B \cdot q_0(n) \quad (5.18)$$

in which case we stop testing and declare the presence of the mean-field signal;

- otherwise, if

$$q_1(n) < A + B \cdot q_0(n) \quad (5.19)$$

for all  $n = 1, 2, \dots, N$ , we declare the absence of the mean-field signal.

For spatial white field, the above approach becomes similar to the Armitage's *restricted sequential procedure* in [92]. We determine the constants  $A$  and  $B$  so that the resulting tests approximately achieve false-alarm and detection performance specifications. We approximate  $x(n) = [z(\mathbf{s}_n) - B] \cdot (1 - \sum_{j \in \mathcal{N}(n)} c_{n,j}) / \tau_n^2$  as a one-dimensional diffusion process with drift  $(\alpha - B) \cdot (1 - \sum_{j \in \mathcal{N}(n)} c_{n,j})$  per unit time and growth in variance at a rate  $(1 - \sum_{j \in \mathcal{N}(n)} c_{n,j})$ , and having an absorbing barrier at  $A$ . We can then approximate the probability that  $q_1(n)$  crosses  $A + B \cdot q_0(n)$  in not more than  $N$  steps by the probability of absorption [for the diffusion process  $x(n)$ ] before time  $N$ . Unfortunately, it is not possible to find a closed-form expression for this probability. However, we can find an approximate absorption probability expression for the diffusion process with an "average" drift  $(\alpha - B) \cdot q_0(N)/N$  per unit time and "average" variance growth rate  $q_0(N)/N$ . Using the result in [93], this absorption probability is approximated by the following expression (see also in [92]):

$$P[\alpha; A, B, N] = \Phi\left(\frac{(\alpha - B)q_0(N) - A}{\sqrt{q_0(N)}}\right) + \exp[2A(\alpha - B)] \cdot \Phi\left(\frac{-(\alpha - B)q_0(N) - A}{\sqrt{q_0(N)}}\right) \quad (5.20)$$

Denote the desired probabilities of false alarm and detection by  $P_{\text{FA,d}}$  and  $P_{\text{D,d}}$  and assume that the region under test is fixed (containing a fixed number of nodes  $N$ ). Setting the likelihood ratio for testing  $H_1 : \alpha = \alpha_d$  versus  $H_0 : \alpha = 0$  to  $P_{\text{D,d}}/P_{\text{FA,d}}$  along the boundary  $q_1(n) = A + B \cdot q_0(n)$  leads to the following choices of  $A$  and  $B$ :

$$A(\alpha_d) = \frac{1}{\alpha_d} \cdot \ln\left(\frac{P_{\text{D,d}}}{P_{\text{FA,d}}}\right), \quad B(\alpha_d) = \frac{\alpha_d}{2} \quad (5.21)$$

We compute  $\alpha_d$  by solving the following equation:

$$P_{\text{D,d}} = P[\alpha_d; A(\alpha_d), B(\alpha_d), N] \quad (5.22)$$

It can be shown that  $P[\alpha_d; A(\alpha_d), B(\alpha_d), N]$  is an increasing function of  $\alpha_d$  (see [92]); hence, we can use bisection to find  $\alpha_d$  that solves (5.22).

The proposed sequential testing procedure can be summarized as:

1. First, specify the desired probabilities of false alarm and detection ( $P_{\text{FA,d}}$  and  $P_{\text{D,d}}$ ) and number of nodes in the testing region ( $N$ ).

2. Solve ( 5.22) for  $\alpha_d$ , implying that the desired detection probability  $P_{D,d}$  will be approximately achieved if the true mean-field value is  $\alpha_d$ .
3. Apply the sequential test in ( 5.18) and ( 5.19) using  $A = A(\alpha_d)$  and  $B = B(\alpha_d)$ .

If the mean signal is greater than  $\alpha_d/2$ , we approximate the average number of steps needed to detect the signal by the following expression:

$$\hat{n}_{av} = \frac{A(\alpha_d)}{[\alpha - B(\alpha_d)] \cdot q_0(N)} \cdot \ln \left( \frac{P_{FA,d}}{P_{D,d}} \right) \quad (5.23)$$

provided that  $\hat{n}_{av} < N$ .

### 5.4.3 Numerical example 2: mean-field detection

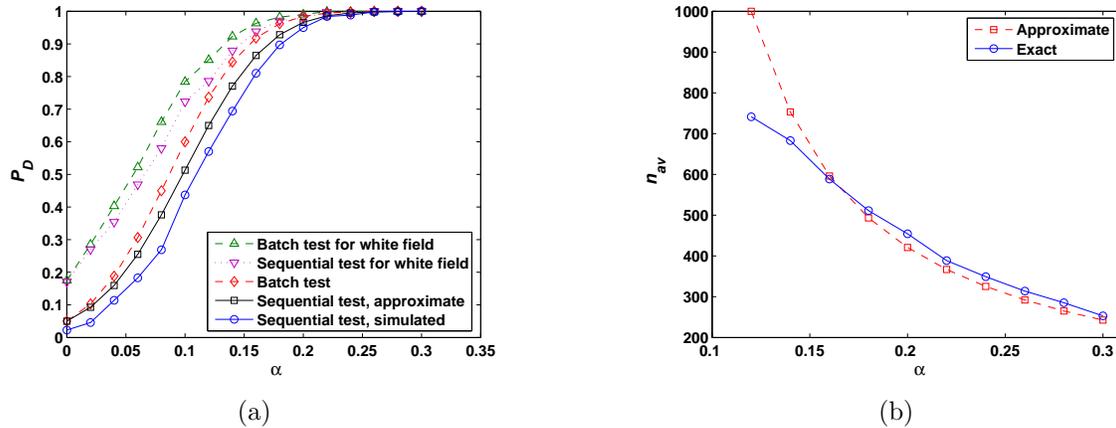


Figure 5.2 Exact and approximate detection probabilities of the proposed and white field batch and sequential tests (left) and average number of steps in which the proposed sequential test reached the decision, as functions of the true mean field  $\alpha$ , for  $P_{FA,d} = 0.05$  and  $P_{D,d} = 0.7$ .

We examine the performance of the distributed detectors in Section 5.4.1 and 5.4.2. We adopt the same measurement scenario as in Section 5.3.2 (Example 1). Our performance metrics are the probabilities of false alarm and detection and average number of steps  $n_{av}$  that the sequential test needs to reach the decision. These quantities are estimated using Monte Carlo simulations with 1000 independent trials. Throughout this example, we set the spatial-dependence and conditional-variance parameters to  $c = 0.083$  and  $\tau^2 = 1$ .

In Figure 5.2, we study the performances of the proposed sequential detector where the constants  $A = A(\alpha_d)$  and  $B = B(\alpha_d)$  have been selected according to the following false-alarm and detection probability specifications:  $P_{\text{FA,d}} = 0.05$  and  $P_{\text{D,d}} = 0.7$  [corresponding to  $\alpha_d = 0.13$ , computed by solving ( 5.22)]. We have also set the threshold  $\xi$  of the batch detector to guarantee the false-alarm probability  $P_{\text{FA,d}} = 0.05$ , see ( 5.16). In Figure 5.2(a), we present

- the detection probability of the proposed batch test [computed using ( 5.17)],
- simulated (exact) and approximate ( $P[\alpha; A(\alpha_d), B(\alpha_d), N]$ ) detection probabilities of the proposed sequential test, and
- exact detection probabilities of the batch and sequential tests that are based on the assumption that the field is spatially white, as functions of the true mean field  $\alpha$ .

Here, the white-field batch and sequential detectors have been designed and implemented by substituting ( 5.6) into the expressions in Section 5.4.1 and 5.4.2 (respectively), where the field variance  $\sigma^2$  has been chosen as the *expected sample variance* of the observations:

$$\sigma^2 = \frac{\text{tr}[(I_N - C)^{-1}D]}{N - 1} - \frac{1}{N(N - 1)} \cdot \mathbf{1}_N^T (I_N - C)^{-1} D \mathbf{1}_N \quad (5.24)$$

The proposed sequential test achieved the false-alarm probability 0.03 (which is less than the specified  $P_{\text{FA,d}}$ ) whereas the batch detector achieved exactly the specified false-alarm probability of 0.05 (as expected). The white-field batch and sequential tests failed to meet the specifications and achieved false-alarm probability of around 0.2; these tests *ignore* spatial dependence in the data and *underestimate* the field variance, which leads to a high false-alarm rate. Figure 5.2(b) presents the average number of steps  $n_{\text{av}}$  that the proposed sequential test needs to reach a decision and approximate  $\hat{n}_{\text{av}}$  in ( 5.23), as functions of the true mean field  $\alpha$ ; it demonstrates that significant savings in speed and energy consumption are possible due to the reduced average number of nodes needed to make a decision.

We now set the true mean field to  $\alpha = 0.15$  and vary the desired false-alarm probability  $P_{\text{FA,d}}$  of the batch and sequential tests, see Figure 5.3. Here,  $\alpha_d$ ,  $A(\alpha_d)$  and  $B(\alpha_d)$  for the sequential test have been chosen so that the detection probability specification  $P_{\text{D,d}}$  is constant and equal to 0.7. Figure 5.3(a) shows the simulated (exact) false-alarm probabilities of

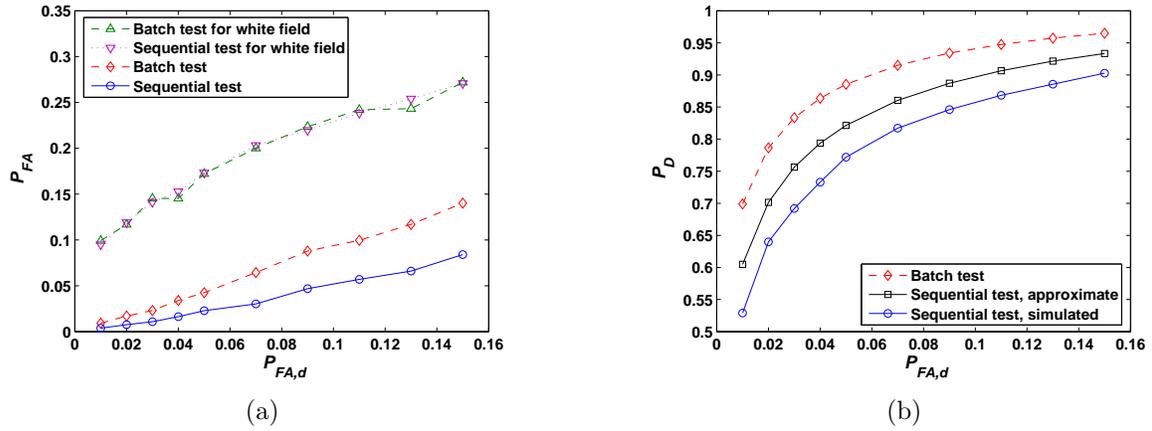


Figure 5.3 Achieved false-alarm probabilities of the proposed and white-field batch and sequential tests (left) and exact and approximate detection probabilities of the proposed batch and sequential tests (right) as functions of the desired false-alarm probability  $P_{FA,d}$ , for  $P_{D,d} = 0.7$  and  $\alpha = 0.15$ .

the proposed batch and sequential tests and exact false-alarm probabilities of the batch and sequential tests for spatially white field, as functions of  $P_{FA,d}$ . For the proposed batch test, the exact and desired false-alarm probabilities are identical, as expected. The achieved false-alarm rates of the proposed sequential test are always within the desired specifications, i.e.  $P_{FA} \leq P_{FA,d}$ . As before, the white-field batch and sequential tests did not meet the false alarm specifications. In Figure 5.3(b), we present the detection probability of the proposed batch test (which is a benchmark) and simulated (exact) and approximate ( $P[\alpha; A(\alpha_d), B(\alpha_d), N]$ ) detection probabilities of the proposed sequential test, as functions of the desired false-alarm probability  $P_{FA,d}$ . Throughout this example, the approximate detection probability curves have been computed using (5.20).

## 5.5 Calibration

We outline an ML calibration algorithm for estimating  $\alpha$ ,  $\tau^2$  and  $c$  under the connectivity-based CAR model in (5.12) and (5.13). Here, the vector of unknown calibration parameters is  $\theta = [\alpha, \tau^2, c]^T$ . For fixed spatial-dependence parameter  $c$ , there exist closed-form expressions

for the ML estimates of  $\alpha$  and  $\tau^2$  (see [26]):

$$\begin{aligned}\hat{\alpha}(c) &= \frac{\sum_{i=1}^N z(\mathbf{s}_i) \cdot (1 - c \cdot w_{i+})}{N - c \cdot \sum_{i=1}^N w_{i+}} \\ \hat{\tau}^2(c) &= \frac{1}{N} \cdot \mathbf{z}^T (I_N - cW) \mathbf{z} - \frac{1}{N} \cdot \hat{\alpha}(c) \cdot \left[ \sum_{i=1}^N z(\mathbf{s}_i) \cdot (1 - c \cdot w_{i+}) \right]\end{aligned}\quad (5.25)$$

where  $w_{i+} = \sum_{j \in \mathcal{N}(i)} w_{i,j}$  denotes the number of neighbors of node  $i$ . Substituting ( 5.25) into the log likelihood ( 5.7) and neglecting constant terms yields the *concentrated log-likelihood function*

$$L_c(c) = -\frac{N}{2} \ln[\hat{\tau}^2(c)] + \frac{1}{2} \cdot \sum_{i=1}^N \ln(1 - c\lambda_i) \quad (5.26)$$

to be maximized with respect to  $c$ . Recall that  $\lambda_i$ ,  $i = 1, 2, \dots, N$  are the eigenvalues of the connectivity matrix  $W$  and observe that the eigenvalues decomposition of  $W$  should be computed only once, prior to maximizing ( 5.26). The maximization in ( 5.26) [subject to ( 5.14)] can be efficiently performed using the damped Newton-Raphson iteration. Once the ML estimate  $\hat{c}$  is computed, the ML estimates of  $\alpha$  and  $\tau^2$  are obtained by substituting  $\hat{c}$  into ( 5.25).

### 5.5.1 Numerical example 3: rainfall precipitation data

We apply the proposed calibration algorithms to the rainfall precipitation data set [94] which has been used in recent sensor-networks literature, see e.g. [36, 37]. In particular, we analyze annual precipitation in the Pacific Northwest region of the United States, averaged over the period 1949-94. The measurements were collected at  $N = 167$  locations forming a grid with 50 km spacing, see Figure 5.4(a). We utilized the CAR model ( 5.12) and ( 5.13) and selected the cutoff distance  $d = 55$  km. In this case,  $1/\lambda_{\text{MAX}} = 0.2576$  and hence the parameter space for the spatial-dependence parameter consists of the following interval:  $0 \leq c < 0.2576 = 1/\lambda_{\text{MAX}}$ , see ( 5.14). We have obtained the following ML estimates of the random field parameters:  $\hat{\alpha} = \hat{\alpha}(\hat{c}) = 1353$  mm,  $\hat{\tau}^2 = \hat{\tau}^2(\hat{c}) = (406 \text{ mm})^2$ , and  $\hat{c} = 0.2570$ . We utilize these estimates to construct a 95% confidence interval for  $\alpha$ :

$$\alpha \in \left( \hat{\alpha} - 1.96[\text{CRB}_{\alpha,\alpha}(\hat{\tau}^2, \hat{c})]^{1/2}, \hat{\alpha} + 1.96[\text{CRB}_{\alpha,\alpha}(\hat{\tau}^2, \hat{c})]^{1/2} \right) = (1126 \text{ mm}, 1580 \text{ mm}) \quad (5.27)$$

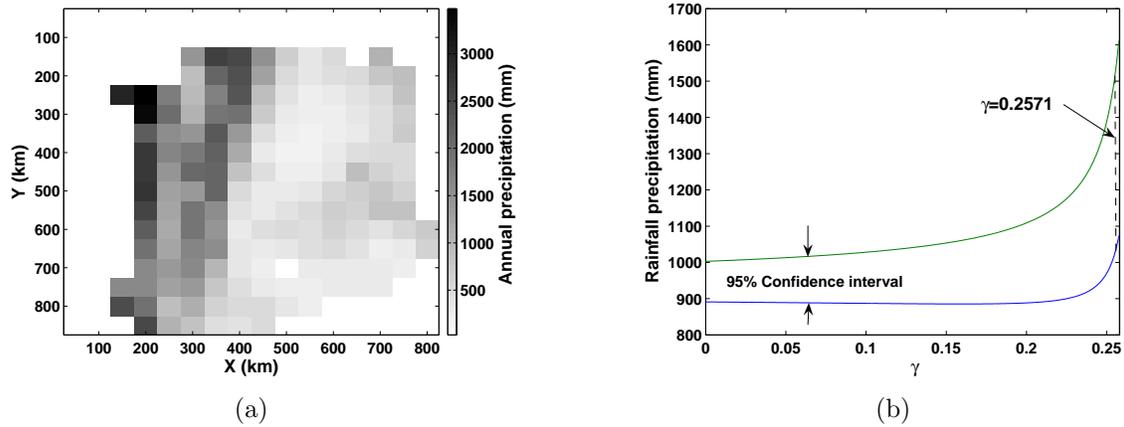


Figure 5.4 Annual rainfall precipitation as a function of the measurement location (left) and 95% confidence interval for  $\alpha$  as a function of the spatial-dependence level  $c$ .

where  $\text{CRB}_{\alpha,\alpha}(\tau^2, c)$  denotes the CRB for  $\alpha$ , which is equal to (5.8) specified to the CAR model in (5.12) and (5.13).

We now show that ignoring positive spatial dependence among the observations leads to underestimation of the field variance and, consequently, to poor performance of classical statistical inference procedures.

**Ignoring spatial dependence:** Consider now the case where we ignore the spatial dependence in the precipitation data. Then, assuming *spatially white* field, we estimate  $\alpha$  using its sample-mean estimate in (5.9):  $\bar{z} = 977$  mm. If measurements were indeed spatially white, we could construct the following 95% confidence interval for  $\alpha$ :

$$\alpha \in (\bar{z} - 1.96 \cdot s/\sqrt{N}, \bar{z} + 1.96 \cdot s/\sqrt{N}) = (860 \text{ mm}, 1094 \text{ mm}) \quad (5.28)$$

where  $s^2 = [1/(N-1)] \cdot \{\sum_{i=1}^N [z(\mathbf{s}_i)]^2\} - N \cdot \bar{z}^2/(N-1) = (769 \text{ mm})^2$  is the *sample variance* of the data. The above confidence interval is deceptively narrow, which is due to the fact that  $s^2$  is a *downward* biased estimate of  $\text{var}[z(\mathbf{s}_i)]$ . Observe that the 95% confidence interval in (5.27) and (5.28) *do not overlap*, indicating the importance of incorporating spatial dependence among the observations into the analysis of sensor-network data.

We now study the dependence of the CAR model based 95% confidence interval for  $\alpha$  on

the spatial-dependence parameter  $c$ . In Figure 5.4(b), we show the confidence interval for the CAR model with the cutoff distance  $d = 55$  km:

$$\alpha \in \left( \hat{\alpha}(c) - 1.96[\text{CRB}_{\alpha,\alpha}(\hat{\tau}^2(c), c)]^{1/2} + 1.96[\text{CRB}_{\alpha,\alpha}(\hat{\tau}^2(c), c)]^{1/2} \right) \quad (5.29)$$

as a function of  $c$ . If there are appreciable correlations among the measurements, the ML estimate  $\hat{c}$  will be very close to  $1/\lambda_{\text{MAX}}$ . Figure 5.4(b) shows that replacing  $c$  with  $1/\lambda_{\text{MAX}}$  in ( 5.29) would yield a confidence interval that is very close to that in ( 5.27). Note that confidence interval for  $c = 0$  in Figure 5.4(b) corresponds to the "classical" confidence interval for spatially white field in ( 5.28).

## CHAPTER 6. GENERAL CONCLUSIONS

### 6.1 Concluding remarks

In this thesis, the following key issues in WSNs have been studied:

- *Distributed spatial signal processing.* In WSN environment, distributed computation is preferred due to limited energy and communication ability of each single sensor. In addition, distributed processing does not require the existence of fusion center, which eliminates the risk of center failure that may disable a significant portion of the network. In Chapter 2, we presented an HMRF framework for distributed detection and estimation of localized phenomena. The algorithm has MRF assumption on the testing field and runs totally distributedly without the presence of a fusion center. The proposed algorithm uses binary data from each sensor, requires no more than 10 iterations for final detection results, and all communications are *local* (i.e., with neighbors). Calibration method for MRF parameters estimation and initialization of the algorithm were also discussed in this chapter.
- *Node localization inaccuracy.* Most nodes estimate their location in WSN scenario. Therefore it is important to take into account the node location inaccuracy and incorporate them into the signal processing algorithm design. In Chapter 3, a Bayesian framework was proposed for energy-based source localization, taking into account the node location uncertainties. Instead of using the estimated node locations as if they are known, the algorithm assumes both energy source (target) and node locations to be random variables. An ICM algorithm was then designed to achieve the MAP estimate of both target and node locations. The proposed algorithm provides significant

improvement compared with methods that ignore the node localization errors.

- *Signal processing using heavily quantized data.* In WSNs, normally sensor observations need to be quantized before sending to the processing center to save energy, therefore it is crucial to design algorithms that are optimal for quantized data other than raw measurements. Under the Poisson regime described in Chapter 4, we proposed a parametric model for location and shape of the event region and assume event signal strength to be constant. We then presented a MAP estimator for the parameters as well as the unknown event signal strength, assuming each node makes binary or quantized local decision. The algorithm is optimal in processing the quantized data in minimum mean square error sense, therefore is suitable for field parameter estimation under WSN environment.
- *Estimating global phenomena in correlated fields.* In large WSNs, sensors are usually deployed densely to monitor the environment at close range with high spatial resolution, which yields highly correlated observations in space. The existence of measurement correlation seriously degrades the detection performance if neglected, thus increases the difficulty of distributed data processing. In Chapter 5, we first derived a distributed ML method for estimating the mean of the spatially correlated Gaussian field, where the correlation follows a CAR model; then we proposed a sequential detector for testing whether the field mean is greater than a certain threshold. It was shown that the sequential detector requires fewer nodes to be involved with the increase of the actual field mean, which significantly reduces the energy consumption of the whole sensor network. Moreover, we proposed a calibration scheme for estimating field correlation parameters and apply the algorithm to the rainfall precipitation data set and construct the confidence interval for field mean, which shows the applicability of the proposed algorithm to real problems.

## 6.2 Future work

We have the following two major directions for future work:

*Extension of our algorithms to allow tracking of phenomena over time.* Most of the previous

discussion has been based on static measurements at single time point. However, monitoring the change of the interesting field over time is of great interest in WSN scenario. Temporal models (e.g. autoregressive model) could be utilized to track the phenomena of interest (such as the defect region in Chapter 2 or the movement of the energy source in Chapter 3).

*Analysis of the impact of communication errors.* In this thesis, we have ignored communication errors. In practice, however, communication errors are inevitable and may degrade the performance and increase the actual communication cost of our algorithms. Therefore further research is needed to study the effects of communication errors on our proposed algorithms and to develop proper communication protocols to reduce the communication error to a specified level.

There are several issues that require further work:

*Developing more general and realistic model for event-region detection.* In Chapter 2, the autologistic MRF model is too simple to fit real applications. More general models should be developed to take into account of multiple bits measurements, the change of field over time, etc. Proper ICM algorithms should also be designed correspondingly.

*Distributed implementation of source localization.* The MAP algorithm proposed in Chapter 3 requires full communication of raw measurements from each node to processing center, which is impractical. Therefore, it is of great importance to develop quantization schemes to reduce the communication load of each node, or to implement the algorithm in distributed way to eliminate the fusion center at all.

*Optimal threshold choice for field parameter estimation under Poisson regime.* In Chapter 4, the thresholds of each sensor are set to satisfy the equally spaced false alarm probability, for convenience. It is of interest to find optimal thresholds to minimize the MSE of field parameter estimation.

*Reversible jump MCMC method for multiple signal region estimation.* The current version of algorithm in Chapter 4 is only suitable for single signal region estimation. It is of interest to develop reversible jump MCMC method [95] to resolve multiple signal source problems, where the number of the signal source is unknown.

## BIBLIOGRAPHY

- [1] F. Zhao and L.J. Guibas. *Wireless Sensor Networks: An Information Processing Approach*. Morgan Kaufmann, 2004.
- [2] C.Y. Chong and S.P. Kumar. Sensor networks: Evolution, opportunities, and challenges. *Proceedings of the IEEE*, 91:1247–1256, August 2003.
- [3] D. Culler, D. Estrin, and M. Srivastava. Overview of sensor networks. *IEEE Transactions on Computers*, 37:41–49, August 2004.
- [4] J. Elson and D. Estrin. Sensor networks: A bridge to the physical world. In C.S. Raghavendra, K.M. Sivalingam, and T. Znati, editors, *Wireless Sensor Networks*. Kluwer, 2004.
- [5] A. Ribeiro and G.B. Giannakis. Distributed estimation in gaussian noise for bandwidth-constrained wireless sensor networks. In *Proc. 38th Asilomar Conference on Signals, Systems and Computers*, pages 1407–1411, Philadelphia, PA, November 2004.
- [6] R. Nowak, U. Mitra, and R. Willett. Estimating inhomogeneous fields using wireless sensor networks. *IEEE Journals on Selected Areas in Communications*, 22(6):999–1006, August 2004.
- [7] S.-H. Son, S.R. Kulkarni, S.C. Schwartz, and M. Roan. Communication-estimation trade-offs in wireless sensor networks. In *Proc. International Conference on Acoustics, Speech and Signal Processing*, pages 1065–1068, Philadelphia, PA, March 2005.
- [8] R. Stoleru and J.A. Stankovic. Probability grid: A location estimation scheme for wireless sensor networks. In *Proc. 1st Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communication Networks*, pages 430–438, Santa Clara, CA, October 2004.

- [9] S. Ray, D. Starobinski, A. Trachtenberg, and R. Ungrangsi. Robust location detection with sensor networks. *IEEE Journals on Selected Areas in Communications*, 22(6):1016–1025, August 2004.
- [10] Y. Sung, L. Tong, and A. Swami. Asymptotic locally optimal detector for large-scale sensor networks under the Poisson regime. *IEEE Transactions on Signal Processing*, 53(6):2005–2017, June 2005.
- [11] B. Krishnamachari and S. Iyengar. Distributed bayesian algorithms for fault-tolerant event region detection in wireless sensor networks. *IEEE Transactions on Computers*, 53(3):241–250, August 2004.
- [12] J.J. Xiao and Z.Q. Luo. Decentralized detection in a bandwidth constrained sensor network. In *Proc. Global Telecommunication Conference*, pages 123–128, November 2004.
- [13] V. Saligrama, M. Alanyali, and O. Savas. Asynchronous distributed detection in sensor networks. *IEEE Transactions on Signal Processing*, submitted for publication.
- [14] J.F. Chamberland and V.V. Veeravalli. Decentralized detection in sensor networks. *IEEE Transactions on Signal Processing*, 51(2):407–416, February 2003.
- [15] J.N. Tsitsiklis. Decentralized detection. *Advances in Statistical Signal Processing*, 2:297–344, November 1993.
- [16] P.K. Varshney. *Distributed Detection and Data Fusion*. Springer, 1997.
- [17] E.B. Ermis and V. Saligrama. Adaptive statistical sampling methods for decentralized estimation and detection of localized phenomena. In *Information Processing in Sensor Networks*, April 2005.
- [18] M.G. Rabbat and R.D. Nowak. Decentralized source localization and tracking. In *Proc. International Conference on Acoustic Speech, Signal Processing*, pages 1407–1411, March 2004.

- [19] F. Zhao, J. Shin, and J. Reich. Information-driven dynamic sensor collaboration for tracking applications. *IEEE Signal Processing Magazine*, 19(2):61–72, March 2002.
- [20] F. Zhao, J. Liu, J. Liu, L. Guibas, and J. Reich. Collaborative signal and information processing: an informative-directed approach. *Proceedings of the IEEE*, 91(8):1199–1209, August 2003.
- [21] J. Liu, J. Reich, and F. Zhao. Collaborative in-network processing for target tracking. *EURASIP Journal on Applied Signal Processing*, 23(4):378–391, April 2003.
- [22] R. Niu and P. Varshney. Target location estimation in wireless sensor networks using binary data. In *Proc. 38th Annual Conference on Information Sciences and Systems*, March 2004.
- [23] A. Krause, C. Guestrin, A. Gupta, and Jon Kleinberg. Near-optimal sensor placements: Maximizing information while minimizing communication cost. In *Proc. 5th International Conference on Information Processing in Sensor Networks*, October 2003.
- [24] A. Giridhar and P.R. Kumar. Toward a theory of in-network computation in wireless sensor networks. *IEEE Communication Magazine*, 44(4):98–107, April 2006.
- [25] A. Dogandžić and B. Zhang. Distributed maximum *a posteriori* estimation and detection for sensor networks using hidden markov random fields. *IEEE Transactions on Signal Processing*, to appear, 2006.
- [26] N.A.C. Cressie. *Statistics for Spatial Data*. Wiley, 1993.
- [27] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B*, 48:259 – 302, 1986.
- [28] A. Dogandžić and B. Zhang. Bayesian signal estimation for sensor networks in the presence of node localization errors. *IEEE Transactions on Signal Processing*, submitted for publication.

- [29] A. Dogandžić and B. Zhang. Parametric signal estimation using sensor networks in the presence of node localization errors. In *Proc. 39th Asilomar Conference on Signals, Systems and Computers*, pages 1571–1575, Pacific Grove, CA, November 2005.
- [30] N. Patwari, J. Ash, S. Kyperountas, A.O. Hero III, R.M. Moses, and N.S. Correal. Locating the nodes: Cooperative localization in wireless sensor networks. *IEEE Signal Processing Magazine*, 22:54–69, July 2005.
- [31] N. Patwari, A.O. Hero III, M. Perkins, N.S. Correal, and R.J. O’Dea. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51:2137–2148, August 2003.
- [32] A. Dogandžić and B. Zhang. Event-region estimation for sensor networks under the Poisson regime. In *Proc. 39th Asilomar Conference on Signals, Systems and Computers*, pages 951–955, Pacific Grove, CA, November 2005.
- [33] A. Dogandžić and B. Zhang. Distributed mean-field estimation and detection in correlated gaussian random fields using sensor networks. In *Proc. 43rd Annual Allerton Conference on Communication, Control and Computing*, pages 951–955, Pacific Grove, CA, November 2005.
- [34] M.C. Vuran, O.B. Akan, and I.F. Akyildiz. Spatio-temporal correlation: Theory and application for wireless sensor networks. *Computer Networks*, 45(3):245–259, 2004.
- [35] D. Ganesan, D. Estrin, and J. Heidemann. Dimensions: Why do we need a new data handling architecture for sensor networks. *Computer Communication Review*, 33:143–148, January 2003.
- [36] S. Patten, B. Krishnamachari, and R. Govindan. The impact of spatial correlation on routing with compression in wireless sensor networks. In *Proc. 3rd International Symposium on Information Processing in Sensor Networks*, pages 28–35, Berkeley, CA, April 2004.

- [37] A. Jindal and K. Psounis. Modeling spatially correlated sensor network data. In *Proc. 1st Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communication Networks*, pages 162–171, Santa Clara, CA, October 2004.
- [38] Y.A. Chau and E. Geraniotis. Distributed multisensor parameter estimation in dependent noise. *IEEE Transactions on Computers*, 40:373–384, February 1992.
- [39] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar. Next century challenges: Scalable coordination in sensor networks. In *Proc. 5th ACM Annual International Conference on Mobile Computing and Networking*, pages 263–270, Seattle, WA, August 1999.
- [40] I.F. Akyildiz, Y. Sankarasubramanian W. Su, and E. Cayirci. Wireless sensor networks: A survey. *Computer Networks*, 38(4):393–422, 2002.
- [41] B.J. Feder. Wireless sensor networks spread to new territory. *New York Times*, July 2004.
- [42] A.S. Willsky. Multiresolution markov models for signal and image processing. *Proceedings of the IEEE*, 90:1396–1458, August 2002.
- [43] A. Dogandžić and B. Zhang. Distributed signal processing for sensor networks using hidden markov random fields. In *Proc. 39th Annual Conference on Information Sciences and Systems*, pages WA7.7–WA7.12, Baltimore, MD, March 2005.
- [44] Z.Q. Luo. An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network. *IEEE Journals on Selected Areas in Communications*, 23:735–744, April 2005.
- [45] Z.Q. Luo. Universal decentralized estimation in a bandwidth constrained sensor network. *IEEE Transactions on Information Theory*, 51:2210–2219, June 2005.
- [46] J.J. Xiao, S. Cui, Z.Q. Luo, and A.J. Goldsmith. Power scheduling of universal decentralized estimation in sensor networks. *IEEE Transactions on Signal Processing*, 54(2):413–422, February 2006.

- [47] Y. Pawitan. *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, 2001.
- [48] K. Plarre and P.R. Kumar. Extended message passing algorithm for inference in loopy gaussian graphical model. *Ad Hoc Networks*, 2:153–169, April 2004.
- [49] J.W. Fisher, M.J. Wainwright, E.B. Sudderth, and A.S. Willsky. Statistical and information-theoretic methods for selforganization and fusion of multimodal, networked sensors. *International Journal of High Performance Computing Applications*, 15:337–353, 2002.
- [50] E.B. Sudderth, M.J. Wainwright, and A.S. Willsky. Embedded trees: Estimation of gaussian processes on graphs with cycles. *IEEE Transactions on Signal Processing*, 54:3136–3150, November 2004.
- [51] S.L. Lauritzen. *Graphical Models*. Oxford University Press, 1996.
- [52] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.
- [53] V. Delouille, R. Neelamani, V. Chandrasekaran, and R.G. Baraniuk. The embedded triangles algorithm for distributed estimation in sensor networks. In *Proc. 2003 IEEE Workshop Statistical Signal Processing*, pages 371–374, St. Louis, MO, September 2003.
- [54] S. Venkatesh and M. Alanyali. M-ary hypothesis testing in sensor networks. In *Proc. 38th Annual Conference on Information Sciences and Systems*, pages 696–701, Princeton, NJ, March 2004.
- [55] D.S. Scherber and H.C. Papadopoulos. Distributed computation of averages over ad hoc networks. *IEEE Journals on Selected Areas in Communications*, 23:776–787, April 2005.
- [56] D.V. Lindley and A.F.M. Smith. Bayes estimates for the linear model. *Journal of the Royal Statistical Society, Series B*, 34:1–41, 1972.

- [57] M.H. Chen, Q.M. Shao, and J.G. Ibrahim. *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag, 2000.
- [58] D.B. Rowe. *Multivariate Bayesian Statistics*. Chapman and Hall, 2002.
- [59] A. Renyi. On measures of entropy and information. In *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 547–561, 1961.
- [60] N.A.C. Cressie and T.R.C. Read. Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B*, 46:440–464, 1984.
- [61] T.R.C. Read and N.A.C. Cressie. *Goodness-of-fit Statistics for Discrete Multivariate Data*. Springer-Verlag, 1988.
- [62] A.B. Owen. *Empirical Likelihood*. Chapman and Hall, 2001.
- [63] C. Stein. Efficient nonparametric testing and estimation. In J. Neyman, editor, *Proc. 3rd Berkeley Symp. Mathematical Statistics and Probability*, volume 1. University of California Press, 1956.
- [64] P. Hall and B. La Scala. Methodology and algorithms of empirical likelihood. *International Statistical Review*, 58:109–127, August 1990.
- [65] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1992.
- [66] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [67] B. Efron. Nonparametric standard errors and confidence intervals. *Canadian Journal of Statistics*, 9(2):139–172, August 1981.
- [68] B. Efron. *The Jackknife, the Bootstrap, and Other Resampling Plans*. SIAM, 1982.
- [69] A.C. Davison and D.V. Hinkley. *Bootstrap Methods and their Application*. Cambridge Univ. Press, 1997.

- [70] P.D. Sampson and P. Guttorp. Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 87:108–119, March 1992.
- [71] S.M.S. Lee and G.A. Young. Nonparametric likelihood ratio confidence intervals. *Biometrika*, 86:107–118, March 1999.
- [72] T.J. DiCiccio and J.P. Romano. Nonparametric confidence limits by resampling methods and least favorable families. *International Statistical Review*, 58:59–76, April 1990.
- [73] J. Besag. Statistical analysis of non-lattice data. *The Statistician*, 24:179–195, 1975.
- [74] L. Xiao, S. Boyd, and S. Lall. A scheme for asynchronous distributed sensor fusion based on average consensus. In *Proc. 4th International Symposium on Information Processing in Sensor Networks*, April 2005.
- [75] S.M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.
- [76] B. Efron and R.J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall, 1993.
- [77] Y. Zou and K. Chakrabarty. Uncertainty-aware and coverage-oriented deployment for sensor networks. *Journal of Parallel and Distributed Computing*, 64:788–798, July 2004.
- [78] R. Moses and R. Patterson. Self-calibration of sensor networks. In *Proc. SPIE Unattended Ground Sensor Technologies and Applications IV*, volume 4743, April 2002.
- [79] A.T. Ihler, J.W. Fisher, R.L. Moses, and A.S. Willsky. Nonparametric belief propagation for self-localization of sensor networks. *IEEE Journals on Selected Areas in Communications*, 23:809–819, April 2005.
- [80] X.H. Sheng and Y.H. Hu. Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *IEEE Transactions on Signal Processing*, 53:44–53, January 2005.

- [81] A.O. Hero III and D. Blatt. Sensor network source localization via projection onto convex sets (pocs). In *Proc. International Conference on Acoustic Speech, Signal Processing*, March 2005.
- [82] D. Li and Y.-H. Hu. Energy-based collaborative source localization using acoustic microphone array. *EURASIP Journal on Applied Signal Processing*, 4:321–337, March 2003.
- [83] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Dsorba. A solution to the simultaneous localization and map building (slam) problem. *IEEE Transactions on Robotics and Automation*, 17:229–241, June 2001.
- [84] G.K. Smyth. Partitioned algorithms for maximum likelihood and other non-linear estimation. *Statistical Computation*, 6:201–216, September 1998.
- [85] D. A. Harville. *Matrix Algebra From a Statistician's Perspective*. Springer-Verlag, 1997.
- [86] L. Tong, Q. Zhao, and S. Adireddy. Sensor networks with mobile agents. In *IEEE Military Communications Conference*, October 2003.
- [87] R.M. Neal. Slice sampling. *The Annals of Statistics*, 31(3):705–767, 2003.
- [88] C.S. Raghavendra, K.M. Sivalingam, and T. Znati. *Wireless Sensor Networks*. Kluwer, 2004.
- [89] J. Besag. Spatial interaction and statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36:192–236, 1974.
- [90] J. Besag and C. Kooperberg. On conditional and intrinsic autoregression. *Biometrika*, 82:733–746, 1995.
- [91] H. Rue and H. Tjelmeland. Fitting gaussian markov random fields to gaussian fields. *Scandinavian Journal of Statistics*, 29:31–49, March 2002.
- [92] P. Armitage. Restricted sequential procedures. *Biometrika*, 44:9–26, 1957.

- [93] M.S. Bartlett. The large-sample theory of sequential testes. *Proceedings of the Cambridge Philosophical Society*, 42:239–244, 1946.
- [94] M. Widmann and C. Bretherton. *50 km resolution daily precipitation for the Pacific Northwest, 1949-94*. Online: [http://www.jisao.washington.edu/data\\_sets/widmann/](http://www.jisao.washington.edu/data_sets/widmann/).
- [95] P.J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, December 1995.