2016

Selected Topics in Signal Detection and Estimation in Sensor Networking

Jiangfan Zhang
Lehigh University

SELECTED TOPICS IN SIGNAL DETECTION AND ESTIMATION IN SENSOR NETWORKING

by

Jiangfan Zhang

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Electrical Engineering

Lehigh University

January 2016

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Date

Prof. Rick S. Blum
(Dissertation Director)

Accepted Date

Committee Members:

Prof. Rick S. Blum
(Committee Chair)

Prof. Tiffany Jing Li

Prof. Zhiyuan Yan

Prof. Wei-Min Huang

Acknowledgments

To me, writing these acknowledgments amounts to waving farewell to my Ph.D. studies, which is a truly joyful, invaluable and memorable experience in my life.

First of all, I would like to express my great appreciation to my advisor, Prof. Rick S. Blum. This dissertation could not be completed without his expert guidance, continuing encouragement, and constant support. I feel very lucky to pursue my Ph.D. under his supervision and guidance, and I have learned a lot from him. Prof. Blum has taught me how to tackle research problems in a critical and systematical manner, and also has taught me how to present our results in a clear and understandable way. His diligence and passion for research have strongly influenced me and inspired me to devote myself to my studies in the past years.

I am very grateful to Prof. Tiffany Jing Li, Prof. Zhiyuan Yan, and Prof. Wei-min Huang, for serving on my doctoral committee and spending their precious time on reviewing my dissertation. I would like to thank Prof. Daniel Conus, Dr. Lance Kaplan, Dr. Chuanming Wei, Basel Alnajjab, Xin Cui, and Xuanxuan Lu, for their helpful discussions and insightful advice on our joint work.

My appreciation also goes to all my fellow students in the Signal Processing and Communications Lab (PL 601) and all my friends at Lehigh, who have given me a lot of warm help and made my life easier and more colorful. My appreciation is also extended to Lehigh University and the Electrical and Computer Engineering Department, for providing me an excellent research environment.

Finally, and most importantly, I would like to specially thank my wife Xuanxuan for her selfless love, patience, and all the sacrifice that she went through. I would like to express my

deepest gratitude to my beloved parents for their unconditional love, endless support, and constant care.

I would like to dedicate this dissertation to my parents, my wife and my daughter.

Contents

List of Figures

Abstract

The work in this dissertation investigates selected topics concerning sensor networks which focus on solving signal detection and estimation problems. In the interest of complexity reduction or to facilitate efficient distributed computation using consensus, modified versions of the optimal hypothesis test are considered for a canonical multivariate Gaussian problem in the first part. As the optimal test involves all possible products of observations taken at L different times or from L different sensors, the investigations consider truncated tests which maintain only those products involving sensors or times with indices that differ by k or less. Such tests can provide significant complexity and storage reduction and facilitate efficient distributed computation using a consensus algorithm provided k is much smaller than L. The focus is on cases with a large number L of observations or sensors such that significant efficiency results with a truncation rule, k as a function of L, which increases very slowly with L. A key result provides sufficient conditions on truncation rules and sequences of hypothesis testing problems which provide no loss in deflection performance, an accepted performance measure, as L approaches infinity when compared to the optimal detector. Several popular classes of system and process models, including observations from wide-sense stationary limiting processes as $L \to \infty$ after the mean is subtracted, are employed as illustrative classes of examples to demonstrate the sufficient conditions are not overly restrictive. In these examples, we find significant truncation can be employed even when we assume the difficulty of the hypothesis testing problem scales in the least favorable manner, putting the most stringent conditions on the truncation rule. In all the cases considered, numerical results imply the fixed-false-alarm-rate detection probability of the truncated detector converges to the detection probability of the optimal detector for our asymptotically optimal truncation

in terms of deflection.

In the second part, distributed estimation of a deterministic mean-shift parameter in additive zero-mean noise is studied when using binary quantized data in the presence of man-in-the-middle attacks which falsify the data transmitted from sensors to the fusion center. Several subsets of sensors are assumed to be tampered with by adversaries using different attacks such that the compromised sensors transmit fictitious data. First, we consider the task of identifying and categorizing the attacked sensors into different groups according to distinct types of attacks. It is shown that increasing the number K of time samples at each sensor and enlarging the size N of the sensor network can both ameliorate the identification and categorization, but to different extents. As $K \to \infty$, the attacked sensors can be perfectly identified and categorized, while with finite but sufficiently large K, as $N \to \infty$, it can be shown that the fusion center can also ascertain the number of attacks and obtain an approximate categorization with a sufficiently small percentage of sensors that are misclassified. Next, in order to improve the estimation performance by utilizing the attacked observations, we consider joint estimation of the statistical description of the attacks and the parameter to be estimated after the sensors have been well categorized. When using the same quantization approach successfully employed without attacks, it can be shown that the corresponding Fisher Information Matrix (FIM) is singular. To overcome this, a time-variant quantization approach is proposed, which will provide a nonsingular FIM, provided that $K \geq 2$. Furthermore, the FIM is employed to provide necessary and sufficient conditions under which utilizing the compromised sensors in the proposed fashion will lead to better estimation performance when compared to approaches where the compromised sensors are ignored.

In the last part, estimation of an unknown deterministic vector from possible nonbina-

ry quantized sensor data is considered in the presence of spoofing attacks which alter the data presented to several sensors. Contrary to previous work, a generalized attack model is employed which manipulates the data using transformations with arbitrary functional forms determined by some attack parameters whose values are unknown to the attacked system. For the first time, necessary and sufficient conditions are provided under which the transformations provide a guaranteed attack performance in terms of Cramer-Rao Bound (CRB) regardless of the processing the estimation system employs, thus defining a highly desirable attack. Interestingly, these conditions imply that, for any such highly desirable attack when the attacked sensors can be perfectly identified by the estimation system, either the Fisher Information Matrix (FIM) for jointly estimating the desired and attack parameters is singular or the attacked system is unable to improve the CRB for the desired vector parameter through this joint estimation even though the joint FIM is nonsingular. It is shown that it is always possible to construct such a highly desirable attack by properly employing a sufficiently large dimension attack vector parameter relative to the number of quantization levels employed, which was not observed previously. For a class of spoofing attacks, a computationally efficient heuristic for the joint identification of the attacked sensors and estimation of the desired vector parameter achieves the CRB when the sensor system can perfectly identify the attacked sensors (a genie bound) for a sufficient number of observations in numerical tests.

Chapter 1

Introduction

Encouraged by the great success in applications ranging from inexpensive commercial systems to complex military and homeland defense surveillance systems, sensor systems employed for hypothesis testing and parameter estimation have seen growing interest in recent years. Sensor systems usually consist of a large number of dispersed sensors which execute multiple functions such as sensing, data processing, and communication. Several fundamental issues remain open on the topic of sensor networks focusing on signal detection and estimation problems, especially for the cases where practical concerns are taken into account.

In practical sensor systems, the communication power of each senor is limited. Hence, every sensor can only communicate with its neighbor sensors which are sufficiently close to it. For widely distributed sensor systems without a fusion center, two sensors can not directly communicate with each other when they are very far apart, and hence it is impossible to compute the optimum test statistic if single hop communications are employed. Motivated by this fact and recent advancement in consensus algorithms, we investigate the truncated multivariate Gaussian hypothesis testing problem, and show that under certain conditions,

the truncated detector can asymptotically achieve the optimum performance. It is worth mentioning that there are numerous applications of the truncated detector beyond sensor systems.

Recent technological advances in coding, digital wireless communications technology and digital electronics have lead to the dominance of digital communications using quantized data in sensor networks. Hence, a great deal of attention has focused on parameter estimation using quantized data. For this kind of system, the time samples are converted to quantized data and then transmitted to the fusion center (FC) due to the communications employed at each sensor. After collecting the quantized data from all sensors, the FC makes an estimate of the desired parameter. However, this kind of sensor system is vulnerable to malicious attackers. The last work in this dissertation focuses on attacked sensor systems attempting to perform parameter estimation by using quantized data. Two classes of malicious attacks are considered. One class of attacks are called man-in-the-middle attacks, which capture several subsets of sensors and falsify the quantized data transmitted from the attacked sensors to the FC. The other class of attacks are referred to as spoofing attacks, which modify the unquantized observations of the phenomenon presented to the attacked sensors.

In the presence of malicious attacks in sensor networks, two important issues involved in the parameter estimation are of considerable interest. One is how to identify and categorize the attacked sensors into different groups according to distinct types of attacks. The other one is how much gain we can obtain by making using of the data from the attacked sensor. For the man-in-the-middle attacks, we first study the ability of the FC to identify the attacked sensors and categorize them into different groups corresponding to distinctly different types of attacks. We only assume that the set of unattacked sensors is a larger percentage of all sensors than any set of identically attacked sensors to avoid ambiguity between a set of attacked and

a set of unattacked sensors. It can be shown that increasing the number K of time samples at each sensor and enlarging the size N of the sensor network can both improve the performance of the identification and categorization approach, but to different extents. To be specific, the FC is able to determine the number P of attacks in the sensor network and achieve the correct categorization as $K \to \infty$, while as $N \to \infty$ with finite but sufficiently large K, it can be shown that the FC can also ascertain P and obtain an approximate categorization with a very small percentage of sensors that are misclassified, so small that this misclassification impacts performance in a manner which can be tolerated. In this sense, with sufficiently many time samples at each sensor or a sufficiently large size sensor network, the FC is able to determine the number of attacks in the sensor network and categorize the sensors into different groups according to distinct types of attacks perfectly or with negligibly small misclassification. Next, we consider estimation of the desired parameter. There are two approaches: (1) ignore the data at the attacked sensors. (2) Use the data at the attacked sensors. We can easily take approach (1) without estimating any parameters describing the attacks. However, to attempt to take approach (2), and potentially do better than approach (1), we will investigate the performance of the joint estimation of the desired parameter and the unknown attack parameters. It is shown that the Fisher Information Matrix (FIM) for jointly estimating these parameters is singular when we apply exactly the same quantization approach typically used for the unattacked system. Thus, it is not possible to jointly estimate the desired and attack parameters efficiently with an estimation error that decreases with KN by employing the same quantization approach typically used for the unattacked system. In order to overcome the FIM singularity, a time-variant quantization approach has been proposed. The basic idea is that each sensor divides its observation time interval into several time slots, and in each time slot, all sensors use an identical threshold to quantize the time samples. However,

the thresholds utilized in different time slots are distinct. We can show that as long as at least two different thresholds have been employed, the FIM of the time-variant quantization approach is nonsingular. Further, this FIM has been used to provide necessary and sufficient conditions under which taking advantage of the attacked sensors in the proposed fashion will provide better estimation performance when compared to approaches where the attacked sensors are ignored. These results are obtained by also employing the FIM for the case where the attacked sensors are ignored and the comparisons were made assuming both approaches use the same set of distinct thresholds over the same different time slots to provide a fair comparison. In the numerical results, we show that for some cases, significant improvement in the estimation performance can be obtained by employing the proposed approach. The focus is on binary quantization in this chapter.

Spoofing attacks on sensor networks can occur in various engineering applications. For instance, spoofing attacks have been described for the localization problem in wireless sensor networks. Radar and sonar systems also suffer from spoofing attack threats in practice. As one example of a spoofing attack technique, the application of an electronic countermeasure (ECM), which is designed to jam or deceive the radar or sonar system, can critically degrade the detection and estimation performance of the system. One popular technique for the implementation of ECM employs digital radio frequency memory (DRFM) in radar systems to manipulate the received signal and retransmit it back to confuse the victim radar system. DRFM can mislead the estimation of the range of the target by altering the delay in transmission of pulses, and fool the system into incorrectly estimating the velocity of the target by introducing a Doppler shift in the retransmitted signal. Unlike previous work, a generalized attack model is employed which manipulates the data using transformations with arbitrary functional forms determined by some attack parameters whose values are unknown

to the attacked system. For the first time, necessary and sufficient conditions are provided under which these transformations provide a guaranteed attack performance in terms of Cramer-Rao bound (CRB) no matter what processing the estimation system employs, thus defining a highly desirable attack. These conditions imply that for any such attack when the attacked sensors can be perfectly identified by the estimation system, either the FIM for jointly estimating the desired and attack parameters is singular or that the attacked system is unable to improve the CRB for the desired vector parameter through this joint estimation even though the joint FIM is nonsingular. It is shown that it is always possible to construct such a desirable attack by properly employing a sufficiently large dimension attack vector parameter relative to the number of quantization levels employed, which was not observed previously. It is shown that when the attacked sensors can be perfectly identified, a spoofing attack can render the attacked measurements useless in terms of reducing the CRB for estimating the desired vector parameter if and only if it is such a desirable attack. For a class of such desirable attacks, a computationally efficient heuristic is developed for the joint identification of the attacked sensors and estimation of the desired vector parameter which, in numerical tests for a sufficiently large number of observations, achieves a genie bound that knows all the groups of identically attacked sensors. Possibly nonbinary quantizations are considered in this chapter.

Chapter 2

Asymptotically Optimal Truncated Multivariate Gaussian Hypothesis Testing with Application to Consensus Algorithms

2.1 Introduction

Hypothesis testing for sensor networks with observations described by a multivariate Gaussian distribution has attracted considerable attention, with applications ranging across various engineering disciplines such as spectrum sensing in cognitive radio networks [1, 2], multiple-input multiple-output radar detection [3–5], and more recently, fault and attack detection in smart grids [6–9]. Here we consider the most general formulation of the simple versus simple hypothesis test [10] for multivariate Gaussian observations which has numerous

applications beyond sensor networking. Particular example application areas include quantitative analysis of the economy [11], stochastic finite element analysis in civil and mechanical engineering [12], and medical imaging [13]. Further applications are detailed in [14–16]. Let $\mathbf{x}_L = [x_1, x_2, ..., x_L]^T$ represent an observed Gaussian random vector with real entries. Then the general hypothesis testing problem can be stated as

$$\begin{aligned} \mathcal{H}_0 : \mathbf{x}_L &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\ \mathcal{H}_1 : \mathbf{x}_L &\sim \mathcal{N}(\boldsymbol{\mu}_L, \boldsymbol{\Sigma}_L) \end{aligned} \tag{2.1}$$

where $\mathcal{N}(\mathbf{v}, \mathbf{C})$ denotes a multivariate Gaussian distribution with mean vector $\mathbf{v}$ and covariance matrix $\mathbf{C}$. A mild assumption is made throughout this chapter.

Assumption 1 *$\boldsymbol{\mu}_L$ and $\boldsymbol{\Sigma}_L$ are known, and either $\boldsymbol{\mu}_L \neq \mathbf{0}$ or $\boldsymbol{\Sigma}_L \neq \mathbf{I}$. The elements of $\boldsymbol{\mu}_L$ are finite, and the eigenvalues of $\boldsymbol{\Sigma}_L$ are bounded by $\left[\varepsilon_0, \varepsilon_0^{-1}\right]$ for some positive number $\varepsilon_0 < 1$.*

Note that, any test of $\mathcal{H}_0 : \mathbf{x}_L \sim \mathcal{N}(\boldsymbol{\mu}_{L,0}, \boldsymbol{\Sigma}_{L,0})$ versus $\mathcal{H}_1 : \mathbf{x}_L \sim \mathcal{N}(\boldsymbol{\mu}_{L,1}, \boldsymbol{\Sigma}_{L,1})$ can be reduced to the canonical test (2.1), by subtraction and whitening to define $\boldsymbol{\mu}_L = \boldsymbol{\mu}_{L,1} - \boldsymbol{\mu}_{L,0}$ and $\boldsymbol{\Sigma}_L = \boldsymbol{\Sigma}_{L,0}^{-\frac{1}{2}} \boldsymbol{\Sigma}_{L,1} \boldsymbol{\Sigma}_{L,0}^{-\frac{1}{2}}$.

The optimal test statistic to minimize error probability, risk, or one of several other criteria for the problem in (2.1) compares the log-likelihood ratio

$$T_L^{\text{opt}} = \mathbf{x}_L^T \mathbf{R}_L \mathbf{x}_L + 2\boldsymbol{\mu}_L^T \boldsymbol{\Sigma}_L^{-1} \mathbf{x}_L = \sum_{i=1}^{L} \sum_{j=1}^{L} x_i (\mathbf{R}_L)_{i,j} x_j + 2 \sum_{i=1}^{L} \zeta_i x_i \tag{2.2}$$

to a threshold [10], where $\mathbf{R}_L \triangleq \left(\mathbf{I} - \boldsymbol{\Sigma}_L^{-1}\right)$ and $\boldsymbol{\mu}_L^T \boldsymbol{\Sigma}_L^{-1} \triangleq [\zeta_1, \zeta_2, ..., \zeta_L]$. If L is large, the statistic in (2.2) is difficult to compute. Even if the components of the vector $\mathbf{x}_L$ are time

samples available at a single location [17,18], then (2.2) requires storing the entire vector $\mathbf{x}_L$ which results in unreasonable storage requirements if L is large. Further the computation of (2.2) generally requires $O(L^2)$ multiplications. On the other hand, if we were able to ignore those terms in (2.2) which involve time samples x_i and x_j that are sufficiently far apart from one another in the time sequence, thus $|i-j| > k$, then we compute (2.2) with a truncated approximation as

$$T_L^{\text{tr}} = \mathbf{x}_L^T \mathbf{B}_L^{(k)} \mathbf{x}_L + 2\boldsymbol{\mu}_L^T \boldsymbol{\Sigma}_L^{-1} \mathbf{x}_L = \sum_{i=1}^{L} \sum_{j=1}^{L} x_i \left(\mathbf{B}_L^{(k)} \right)_{i,j} x_j + 2 \sum_{i=1}^{L} \zeta_i x_i \tag{2.3}$$

where $\left(\mathbf{B}_L^{(k)} \right)_{i,j} \triangleq \begin{cases} (\mathbf{R}_L)_{i,j}, & \text{if } |i-j| \leq k \\ 0, & \text{otherwise} \end{cases}$ is the truncated matrix of $\mathbf{R}_L$, and k is the truncation length. We refer to the detector based on the test statistic shown in (2.3) as the truncated detector. To compute (2.3), we need only store a very small running window of about $2k+1$ time samples around each incoming time sample which results in considerably lower storage requirements. The number of multiplications is also reduced to grow linearly with L.

Analogous benefits can be gained in distributed sensor networking applications where the entries of $\mathbf{x}_L$ come from remotely positioned sensors and a consensus algorithm is employed. Motivated by early ground breaking work [19,20], deterministic [21,22] or randomized consensus algorithms [23,24], are known to be very efficient methods to compute a test statistic while simultaneously communicating the result to every node in the network when the number of nodes is very large, even for imperfect communication channels [25–28]. For simplicity, consider the case where the sensors are placed along a line in what is often called a linear array. Then to compute (2.2) exactly requires collecting, at a single location, observations

x_i and x_j that may be produced at sensors which are very far apart. Collecting this information implies large energy communications if single hop communications and centralized processing are employed. If multiple hop communications are employed, very large delays result and complex control is required. On the other hand, the truncated test statistic in (2.3) can be calculated in an efficient two step procedure. In the first step, each node collects the data from its k neighbors on each side so the i-th node can compute the inner sum of $\sum_{i=1}^{L}\sum_{j=1}^{L} x_i\left(\mathbf{B}_L^{(k)}\right)_{i,j}x_j$ from (2.3). In the second step, a single consensus algorithm [21–34] is used to simultaneously compute the outer sum along with the other added term in (2.3) as $\sum_{i=1}^{L}\left(x_i\sum_{j=1}^{L}\left(\mathbf{B}_L^{(k)}\right)_{i,j}x_j + 2\zeta_i x_i\right)$. The ideas extend to cases where the sensors are not located in an array, in that we would still like to have each sensor only collect observations from its closest neighbors in the first step and then run consensus in the second step. The recent flurry of activity focusing on developing the theory of consensus algorithms has produced an extremely efficient method of distributed computation of a test statistic provided the test statistic can be expressed as a linear function of local statistics which can each be computed using only local observations at each sensor. Test statistics which are quadratic forms like (2.2), which appear in some of the most basic and important signal detection problems, do not satisfy this requirement and we have not seen previous work on using consensus algorithms to compute such statistics. Our truncated test provides a method for computing the test statistic in (2.2) using consensus, motivating our study of the impact of the truncation on detection performance.

To alleviate the multiplication and storage unit requirements without consideration of consensus algorithm implementation, [17] and [18] have considered a special class of (2.1) involving signal-plus-noise hypothesis testing problems in which $\boldsymbol{\mu}_L = \mathbf{0}$ and $\boldsymbol{\Sigma}_L$ is a Toeplitz matrix. Given these assumptions, [17] and [18] investigated using a truncated test somewhat

similar to (2.3). However, they have shown that when the power of the signal is bounded, their detectors have performance loss, as measured by deflection or asymptotic relative efficiency, even as the size of the observation vector goes to infinity. In their analysis they assume the k in (2.3) is constant with L. Here we consider employing truncation in the more general hypothesis testing problem in (2.1), but unlike [17] and [18] we do not consider using a fixed truncation length k as the size of the observation vector L grows. Instead, we consider a slowly increasing function of the size of the observation vector L, denoted by $k = \varphi(L)$. We call the function $\varphi(L)$, the truncation rule of the truncated detector. Since it is infeasible to obtain a closed-form expression of the error probability of our test statistic, similar to [17] we make use of the deflection or generalized SNR [35] [36], one of most useful performance measures for quadratic detectors, to evaluate the detection performance of the truncated detector (2.3). Deflection has been extensively studied and justified for problems of the type we consider [35] [36]. In particular, in many problems of practical importance, the test statistic which optimizes the deflection criterion is exactly the celebrated likelihood ratio detector [36]. Please see [36] for a complete discussion of the properties of Deflection. Here, we are primarily interested in the asymptotic ($L \to \infty$) detection performance of the truncated detector so we focus on the asymptotic deflection ratio (ADR) of the truncated detector relative to the optimal detector (2.2). Sufficient conditions are given in this chapter for a truncation rule $\varphi(L)$ and a sequence of hypotheses tests from (2.1) which lead to no loss in asymptotic deflection ratio of the truncated detector relative to the optimal detector. Moreover, in contrast to the negative results in [17] and [18], we show that the sufficient conditions are satisfied by several important classes of system and process models [37–39]. Further, our sufficient conditions shed light on how changes in the difficulty of the hypothesis test with L will directly impact the effects of truncation. For example, if the difference

between the parameters $(\mathbf{0}, \mathbf{I})$ and $(\boldsymbol{\mu}_L, \boldsymbol{\Sigma}_L)$ becomes more considerable (in a way we define) as $L \to \infty$, generally a more severe truncation can be employed without sacrificing unity ADR when we compare to a case where the difference between the parameters $(\mathbf{0}, \mathbf{I})$ and $(\boldsymbol{\mu}_L, \boldsymbol{\Sigma}_L)$ is fixed with L. Finally, the $\varphi(L)$ satisfying our sufficient conditions are very useful for obtaining a rough idea of how the required truncation length k must increase with L in order to judge the required complexity.

Since the truncated test statistic in (2.3) eliminates some terms from the optimal statistic which may be necessary for good detection performance, its performance can seriously degrade for some scenarios in which the truncation is too severe. Extremely severe truncation can even make the detection problem singular. To avoid this, the following assumption is made throughout the chapter.

Assumption 2 *As it would not make sense to consider truncated detectors, whose truncation makes the two hypotheses indistinguishable, the truncation length k is large enough to ensure that if $\boldsymbol{\mu}_L = \mathbf{0}$ and $\mathbf{R}_L \neq \mathbf{0}$, then the matrix $\mathbf{B}_L^{(k)} \neq \mathbf{0}$.*

Some comment on the use of deflection as opposed to error probability is in order here. First, analysis using error probability would be intractable. More importantly, we employ deflection in a very constrained manner which we believe will mask any limitations of deflection. Thus we attempt to find truncation rules which render the asymptotic ($L \to \infty$) deflection of the truncated detector and the optimal detector to be equivalent. Thus in terms of deflection, the dropped terms are not important as $L \to \infty$. Employing deflection in this way is intuitively appealing, moreover in all numerical examples we tried, the sufficient conditions for the equivalence in terms of asymptotic deflection also ensure the equivalence in terms of limiting error probability.

Throughout this chapter, bold upper case letters and bold lower case letters are used to denote matrices and column vectors respectively. The symbol $\mathbf{I}$ signifies the identity matrix, while $\mathbf{0}$ and $\mathbf{1}$ stand for the all-zero and all-one column vectors respectively. The subscript of a matrix or a column vector indicates its dimension, for example $\mathbf{\Sigma}_L$ is an L-by-L matrix. The dimensions of $\mathbf{I}$, $\mathbf{0}$, and $\mathbf{1}$ are typically deducible from the context, hence are not explicitly specified. We use $\|\cdot\|$ for the ℓ^2 norm of a vector and $(\mathbf{A})_{i,j}$ for the element in the i-th row and j-th column of the matrix $\mathbf{A}$. The notation $\{\mathbf{A}_L\}$ denotes the sequence $\{\mathbf{A}_L\}_{L=1}^{\infty}$. Finally, the expectation operator is denoted as $\mathbb{E}(\cdot)$ and $\mathrm{tr}(\mathbf{A})$ is the trace of $\mathbf{A}$.

The reminder of the chapter is organized as follows. The ADR of the truncated detector relative to the optimal detector is investigated in Section 2.2. Sufficient conditions for unity ADR are developed in Section 2.3. Section 2.4 discusses some illustrative hypothesis testing problems using important classes of system and process models and provides slowly increasing truncation rules which satisfy the sufficient conditions. Section 2.5 demonstrates how to apply a two-step consensus algorithm to compute the truncated test statistic in (2.3) at each sensor. In Section 2.6, several numerical results are provided to illustrate our theoretical analysis. In Section 2.7, we extend our results to spatially and temporally correlated Gaussian hypothesis testing problems and present the corresponding sufficient conditions. Finally, Section 2.8 provides our conclusions.

2.2 Asymptotic Deflection Ratio of the Truncated Detector relative to the Optimal Detector

For a binary hypothesis testing problem like (2.1), the deflection or generalized SNR [35], [36] of a quadratic test statistic T is defined by

$$D(T) = \frac{\left[\mathbb{E}\left(T\middle|\,\mathcal{H}_1\right) - \mathbb{E}\left(T\middle|\,\mathcal{H}_0\right)\right]^2}{\mathbb{E}\left(T^2\middle|\,\mathcal{H}_0\right) - \left[\mathbb{E}\left(T\middle|\,\mathcal{H}_0\right)\right]^2}. \tag{2.4}$$

For the optimal test statistic T_L^{opt} in (2.2) for the problem in (2.1), we can obtain

$$\mathbb{E}\left(T_L^{\text{opt}}\middle|\,\mathcal{H}_1\right) = \sum_{i,j=1}^{L} (\mathbf{R}_L)_{i,j}\mathbb{E}\left(x_j x_i\middle|\,\mathcal{H}_1\right) + 2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\mathbb{E}\left(\mathbf{x}_L\middle|\,\mathcal{H}_1\right) = \text{tr}\left(\mathbf{R}_L\boldsymbol{\Sigma}_L\right) + 2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L \tag{2.5}$$

and since $\mathcal{H}_0$ is $\mathcal{H}_1$ with $\boldsymbol{\mu}_L = \mathbf{0}$ and $\boldsymbol{\Sigma}_L = \mathbf{I}$, we have

$$\mathbb{E}\left(T_L^{\text{opt}}\middle|\,\mathcal{H}_0\right) = \text{tr}\left(\mathbf{R}_L\right). \tag{2.6}$$

Since

$$\begin{aligned}\left(T_L^{\text{opt}}\right)^2 &= \left(\mathbf{x}_L^T\mathbf{R}_L\mathbf{x}_L + 2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\mathbf{x}_L\right)^2 \qquad (2.7)\\ &= \sum_{i,j,l,m=1}^{L} (\mathbf{R}_L)_{i,j}(\mathbf{R}_L)_{l,m}x_i x_j x_l x_m + 4\sum_{i,j,l=1}^{L}\zeta_l(\mathbf{R}_L)_{i,j}x_i x_j x_l + 4\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\mathbf{x}_L\mathbf{x}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\end{aligned}$$

we have

$$\mathbb{E}\left(\left(T_L^{\text{opt}}\right)^2\middle|\mathcal{H}_0\right) = \mathbb{E}\Bigg(\sum_{i,j,l,m=1}^{L}(\mathbf{R}_L)_{i,j}(\mathbf{R}_L)_{l,m}x_ix_jx_lx_m + 4\sum_{i,j,l=1}^{L}\zeta_l(\mathbf{R}_L)_{i,j}x_ix_jx_l + 4\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\mathbf{x}_L\mathbf{x}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\Big|\mathcal{H}_0\Bigg)$$

$$= \sum_{i,j,l,m=1}^{L}(\mathbf{R}_L)_{i,j}(\mathbf{R}_L)_{l,m}\left(\mathbf{I}_{i,j}\mathbf{I}_{l,m} + \mathbf{I}_{i,l}\mathbf{I}_{j,m} + \mathbf{I}_{i,m}\mathbf{I}_{j,l}\right) + 4\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\mathbf{I}\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L \tag{2.8}$$

$$= \left[\text{tr}\left(\mathbf{R}_L\right)\right]^2 + 2\text{tr}\left(\mathbf{R}_L^2\right) + 4\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-2}\boldsymbol{\mu}_L \tag{2.9}$$

where (2.8) is a consequence of Isserlis' theorem [40], and (2.9) is based on the result that $\text{tr}(\mathbf{AB}) = \sum_i\sum_j \mathbf{A}_{i,j}\mathbf{B}_{j,i}$.

As a result, the deflection of T_L^{opt} can now be given by

$$D(T_L^{\text{opt}}) = \frac{\left[\mathbb{E}\left(T_L^{\text{opt}}\middle|\mathcal{H}_1\right) - \mathbb{E}\left(T_L^{\text{opt}}\middle|\mathcal{H}_0\right)\right]^2}{\mathbb{E}\left(\left(T_L^{\text{opt}}\right)^2\middle|\mathcal{H}_0\right) - \left[\mathbb{E}\left(T_L^{\text{opt}}\middle|\mathcal{H}_0\right)\right]^2} = \frac{\left[\text{tr}\left(\mathbf{R}_L\boldsymbol{\Sigma}_L - \mathbf{R}_L\right) + 2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\right]^2}{2\text{tr}\left(\mathbf{R}_L^2\right) + 4\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-2}\boldsymbol{\mu}_L}. \tag{2.10}$$

Similarly, for the truncated test statistic T_L^{tr} with truncation matrix $\mathbf{B}_L^{(\varphi(L))}$, we can obtain

$$D(T_L^{\text{tr}}) = \frac{\left[\mathbb{E}\left(T_L^{\text{tr}}\middle|\mathcal{H}_1\right) - \mathbb{E}\left(T_L^{\text{tr}}\middle|\mathcal{H}_0\right)\right]^2}{\mathbb{E}\left(\left(T_L^{\text{tr}}\right)^2\middle|\mathcal{H}_0\right) - \left[\mathbb{E}\left(T_L^{\text{tr}}\middle|\mathcal{H}_0\right)\right]^2} = \frac{\left[\text{tr}\left(\mathbf{B}_L^{(\varphi(L))}\boldsymbol{\Sigma}_L - \mathbf{B}_L^{(\varphi(L))}\right) + 2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\right]^2}{2\text{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] + 4\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-2}\boldsymbol{\mu}_L}. \tag{2.11}$$

Now, consider the sequence of optimal test statistics $\left\{T_L^{\text{opt}}\right\}$ in (2.2) for the sequence of tests in (2.1) with $\{\mathbf{\Sigma}_L\}$, and the sequence of truncated test statistics $\left\{T_L^{\text{tr}}\right\}$ in (2.3) for the same problem using $\left\{\mathbf{B}_L^{(\varphi(L))}\right\}$. By (2.10) and (2.11), the asymptotic deflection ratio of the truncated detector relative to the optimal detector is therefore

$$\Lambda(T_\infty^{\text{tr}}, T_\infty^{\text{opt}}) \triangleq \lim_{L\to\infty} \frac{\text{tr}\left(\mathbf{R}_L^2\right) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-2}\boldsymbol{\mu}_L}{\text{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-2}\boldsymbol{\mu}_L}\left[\frac{\text{tr}\left(\mathbf{B}_L^{(\varphi(L))}\mathbf{\Sigma}_L - \mathbf{B}_L^{(\varphi(L))}\right) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L}{\text{tr}\left(\mathbf{R}_L\mathbf{\Sigma}_L - \mathbf{R}_L\right) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L}\right]^2 \tag{2.12}$$

$$= \lim_{L\to\infty}\left(1 + \frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right)\left(1 - \frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right)^2$$

where

$$\delta_1^{(\varphi(L))}(L) \triangleq \text{tr}\left(\mathbf{R}_L^2\right) - \text{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] \tag{2.13}$$

$$\delta_2^{(\varphi(L))}(L) \triangleq \text{tr}\left[\left(\mathbf{R}_L - \mathbf{B}_L^{(\varphi(L))}\right)\left(\mathbf{\Sigma}_L - \mathbf{I}\right)\right] \tag{2.14}$$

$$\psi_1^{(\varphi(L))}(L) \triangleq \text{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-2}\boldsymbol{\mu}_L \tag{2.15}$$

and

$$\psi_2(L) \triangleq \text{tr}\left(\mathbf{R}_L\mathbf{\Sigma}_L - \mathbf{R}_L\right) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L. \tag{2.16}$$

2.2.1 Upper Bounds and Lower Bounds on (2.13)-(2.16)

In order to describe sufficient conditions for $\Lambda(T_\infty^{tr}, T_\infty^{opt}) \to 1$, some upper and lower bounds on the quantities in (2.13)-(2.16) are useful. First consider lower bounds on $\psi_1^{(\varphi(L))}(L)$ and $\psi_2(L)$.

Lemma 1 *Under* Assumptions 1 and 2,

$$\psi_1^{(\varphi(L))}(L) \geq C_1 \tag{2.17}$$

and

$$\psi_2(L) \geq C_2 \tag{2.18}$$

where $C_1 > 0$ *and* $C_2 > 0$ *are constants.*

Proof: [Proof of Lemma 1] Even without *Assumption 2*, we have $\mathrm{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] \geq 0$ and $\boldsymbol{\mu}_L^T \boldsymbol{\Sigma}_L^{-2} \boldsymbol{\mu}_L \geq 0$. If $\boldsymbol{\mu}_L \neq \mathbf{0}$, then for any non-zero element of the $\boldsymbol{\mu}_L$, say $\mu_{L,j}$ which is assumed to be the j-th element of $\boldsymbol{\mu}_L$, we have

$$\psi_1^{(\varphi(L))}(L) = \mathrm{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] + 2\boldsymbol{\mu}^T \boldsymbol{\Sigma}_L^{-2} \boldsymbol{\mu} \geq 2\left(\boldsymbol{\Sigma}_L^{-2}\right)_{j,j} \mu_{L,j}^2 > 0 \tag{2.19}$$

therefore we can choose $C_1 = 2\left(\boldsymbol{\Sigma}_L^{-2}\right)_{i,i} \mu_{L,j}^2$.

Otherwise, we must have $\mathbf{B}_L^{(\varphi(L))} \neq \mathbf{0}$ according to *Assumption 2*. Then for any non-zero entry of $\mathbf{B}_L^{(\varphi(L))}$, say $\left(\mathbf{B}_L^{(\varphi(L))}\right)_{i,j}$, we can obtain

$$\psi_1^{(\varphi(L))}(L) \triangleq \mathrm{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] + 2\boldsymbol{\mu}_L^T \boldsymbol{\Sigma}_L^{-2} \boldsymbol{\mu}_L \geq \left[\left(\mathbf{B}_L^{(\varphi(L))}\right)_{i,j}\right]^2 > 0 \tag{2.20}$$

thus, $C_1 = \left[\left(\mathbf{B}_L^{(\varphi(L))}\right)_{i,j}\right]^2$.

On the other hand since $\mathbf{R}_L = \mathbf{I} - \mathbf{\Sigma}_L^{-1}$, we note that

$$\begin{aligned}\psi_2(L) &= \operatorname{tr}(\mathbf{R}_L\mathbf{\Sigma}_L - \mathbf{R}_L) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L = \operatorname{tr}\left[\mathbf{\Sigma}_L + \mathbf{\Sigma}_L^{-1} - 2\mathbf{I}\right] + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L \\ &= \sum_{i=1}^{L}\left(\lambda_i + \frac{1}{\lambda_i} - 2\right) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L\end{aligned} \tag{2.21}$$

where λ_i is the i-th eigenvalue of $\mathbf{\Sigma}_L$.

Since the minimum of $\left(\lambda_i + \frac{1}{\lambda_i} - 2\right)$ occurs at $\lambda_i = 1$, $\left(\lambda_i + \frac{1}{\lambda_i} - 2\right)$ must be non-negative. If $\boldsymbol{\mu}_L \neq \mathbf{0}$, then by the same argument for $\psi_1^{(\varphi(L))}(L)$, we know $\psi_2(L) \geq C_2$ for some positive constant C_2. Otherwise, according to *Assumption 1*, we have $\mathbf{\Sigma}_L \neq \mathbf{I}$. Thus, there exists at least one eigenvalue of $\mathbf{\Sigma}_L$ which is not equal to 1, say λ_j. Then we can obtain

$$\psi_2(L) = \sum_{i=1}^{L}\left(\lambda_i + \frac{1}{\lambda_i} - 2\right) + 2\boldsymbol{\mu}_L^T\mathbf{\Sigma}_L^{-1}\boldsymbol{\mu}_L \geq \left(\lambda_j + \frac{1}{\lambda_j} - 2\right) > 0 \tag{2.22}$$

hence $C_2 = \left(\lambda_j + \frac{1}{\lambda_j} - 2\right)$. ■

Using (2.13) and (2.14), we next describe upper bounds on the absolute values of $\delta_1^{(\varphi(L))}(L)$ and $\delta_2^{(\varphi(L))}(L)$. The upper bounds are determined by both the size of the observation vector and the truncation length.

Lemma 2 *Under* Assumption 1*, consider a truncated detector with truncation rule $\varphi(L)$ for the hypothesis testing problem (2.1) where L denotes the size of the observation vector. Upper bounds on the absolute values of $\delta_1^{(\varphi(L))}(L)$ and $\delta_2^{(\varphi(L))}(L)$ are given by*

$$\left|\delta_1^{(\varphi(L))}(L)\right| \leq L\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))} \tag{2.23}$$

and

$$\left|\delta_2^{(\varphi(L))}(L)\right| \leq \frac{1}{\varepsilon_0} L \left[\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}\right]^{\frac{1}{2}} \tag{2.24}$$

where $\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}$ *is defined by*

$$\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))} \triangleq \max_i \sum_{j:|j-i|\geq\varphi(L)+1} \left[\left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}\right]^2, \tag{2.25}$$

which is related to the part of the inverse covariance matrix that is not accounted for in the truncation rule $\varphi(L)$*.*

Proof: [Proof of Lemma 2] Inserting $\mathrm{tr}\left(\mathbf{R}_L \mathbf{B}_L^{(\varphi(L))}\right) = \mathrm{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right]$ into (2.13), we can obtain

$$\delta_1^{(\varphi(L))}(L) = \mathrm{tr}\left[\mathbf{R}_L^2 - \left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right] = \mathrm{tr}\left(\mathbf{R}_L - \mathbf{B}_L^{(\varphi(L))}\right)^2 = \sum_{i=1}^{L} \sum_{j:|j-i|\geq\varphi(L)+1} \left[\left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}\right]^2 \tag{2.26}$$

which implies

$$\left|\delta_1^{(\varphi(L))}(L)\right| = \left|\sum_{i=1}^{L} \sum_{j:|j-i|\geq\varphi(L)+1} \left[\left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}\right]^2\right| \leq L\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}. \tag{2.27}$$

Further, noting that $\mathrm{tr}\left(\mathbf{R}_L - \mathbf{B}_L^{(\varphi(L))}\right) = 0$ and applying the Cauchy-Schwarz inequality [41] to (2.14), yields

$$\begin{aligned}\left[\delta_2^{(\varphi(L))}(L)\right]^2 &= \left\{\mathrm{tr}\left[\left(\mathbf{R}_L - \mathbf{B}_L^{(\varphi(L))}\right)(\boldsymbol{\Sigma}_L - \mathbf{I})\right]\right\}^2 = \left\{\mathrm{tr}\left[\left(\mathbf{R}_L - \mathbf{B}_L^{(\varphi(L))}\right)\boldsymbol{\Sigma}_L\right]\right\}^2 \qquad (2.28)\\ &\le \mathrm{tr}\left[\left(\mathbf{R}_L - \mathbf{B}_L^{(\varphi(L))}\right)^2\right]\mathrm{tr}\left(\boldsymbol{\Sigma}_L^2\right) \le \frac{1}{\varepsilon_0^2}L^2\left(\max_i \sum_{j:|j-i|\ge\varphi(L)+1}\left[\left(\boldsymbol{\Sigma}_L^{-1}\right)_{i,j}\right]^2\right)\\ &= \frac{1}{\varepsilon_0^2}L^2\boldsymbol{\Omega}_{\boldsymbol{\Sigma}_L}^{(\varphi(L))}\end{aligned}$$

by employing (2.25) and *Assumption 1.*

Hence, the upper bound on $\left|\delta_2^{(\varphi(L))}(L)\right|$ can be expressed as

$$\left|\delta_2^{(\varphi(L))}(L)\right| \le \frac{1}{\varepsilon_0}L\left[\boldsymbol{\Omega}_{\boldsymbol{\Sigma}_L}^{(\varphi(L))}\right]^{\frac{1}{2}}. \qquad (2.29)$$

■

2.3 Sufficient Conditions for Unity ADR

Let ν denote a positive small constant, such that $0 < \nu \ll 1$. Define $\xi(L) \triangleq \sum_{i=1}^{L} \mathbb{1}(|\mu_{L,i}| \ge \nu)$ and $\eta(L) \triangleq \sum_{i=1}^{L} \mathbb{1}(|\lambda_i - 1| \ge \nu)$, where $\mathbb{1}(\cdot)$ is the indicator function and $\mu_{L,i}$ is the i-th element of $\boldsymbol{\mu}_L$. As a result, $\xi(L)$ describes the number of elements in $\boldsymbol{\mu}_L$ which are sufficiently different from zero, and $\eta(L)$ represents the number of eigenvalues of $\boldsymbol{\Sigma}_L$ which are sufficiently different from unity.

Lemma 3 *Considering the hypothesis testing problem (2.1), we can always choose a constant ν such that*

$$\xi(L) > 0, \quad or \quad \eta(L) > 0. \qquad (2.30)$$

Proof: Suppose there is no number ν which can render $\xi(L)$ or $\eta(L)$ non-zero for some L. Then we have

$$\begin{cases} \mu_{L,i} = 0 \\ \lambda_i = 1 \end{cases} \qquad \text{for } i = 1, 2, ..., L \tag{2.31}$$

Since $\mathbf{\Sigma}_L$ is a symmetric matrix, it can be diagonalized by the eigendecomposition

$$\mathbf{\Sigma}_L = \mathbf{Q}\mathbf{T}\mathbf{Q}^T \tag{2.32}$$

where $\mathbf{Q}$ is an orthogonal matrix, and $\mathbf{T}$ is a diagonal matrix with λ_i on the diagonal.

Hence, we get the contradiction that $\boldsymbol{\mu}_L = \mathbf{0}$ and $\mathbf{\Sigma}_L = \mathbf{Q}\mathbf{T}\mathbf{Q}^T = \mathbf{I}$, which implies the two hypotheses in (2.1) are indistinguishable. This completes the proof. ■

Next, we develop sufficient conditions for unity ADR.

Lemma 4 *The ADR of the truncated detector with truncation rule $\varphi(L)$ relative to the optimal detector converges to unity if and only if*

$$\begin{cases} \lim\limits_{L\to\infty} \dfrac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)} = 0 \\ \lim\limits_{L\to\infty} \dfrac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)} = 0 \end{cases} . \tag{2.33}$$

Hence, a sufficient condition for (2.33) is that upper bounds on $\left|\dfrac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right|$ *and* $\left|\dfrac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right|$ *decrease to* 0*, as the size of the observation vector L increases to infinity.*

Since *Lemma 4* is straightforward from (2.12), the proof is omitted.

Before proceeding, it is important to define a family of sequences, as it will play a significant role in our analysis. Using (2.25), let

$$\mathcal{U}(\varphi) \triangleq \left\{ \{\mathbf{\Sigma}_L\} \,\middle|\, \lim_{L\to\infty} \omega(L) \mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))} = 0 \right\} \tag{2.34}$$

describe the family of all sequences of $\{\mathbf{\Sigma}_L\}$ with the stated limit, where

$$\omega(L) \triangleq \begin{cases} \omega_1(L) = \max\left\{L, \frac{L^2}{\eta^2(L)}\right\}, \\ \qquad \text{if } \xi(L) = 0 \text{ and } \eta(L) > 0 \\ \omega_2(L) = \max\left\{\frac{L}{\xi(L)}, L^2[2\varepsilon_0(1+\nu)\xi(L) + \eta(L)]^{-2}\right\}, \\ \qquad \text{if } \xi(L) > 0 \text{ and } \eta(L) \geq 0 \end{cases} \tag{2.35}$$

and ε_0 was defined in *Assumption 1.*

We now give a Theorem providing sufficient conditions, under which the ADR converges to unity.

Theorem 1 *Given a sequence of covariance matrices* $\{\mathbf{\Sigma}_L\}$ *in (2.1) which satisfy* Assumptions 1 and 2 *and a sequence of truncated test statistics* $\left\{T_L^{\text{tr}}\right\}$ *in (2.3) with truncation rule* $\varphi_0(L)$*, sufficient conditions for* $\Lambda(T_\infty^{tr}, T_\infty^{opt}) = 1$ *are*

$$\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi_0). \tag{2.36}$$

If a given $\varphi_0(L)$ *satisfies (2.36), then it follows that any truncation rule* $\varphi(L)$*, which satisfies* $\lim_{L\to\infty} \frac{\varphi(L)}{\varphi_0(L)} \geq 1$*, will also provide* $\Lambda(T_\infty^{tr}, T_\infty^{opt}) = 1$.

Proof: [Proof of Theorem 1] Let's first deduce upper bounds on $\left|\frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right|$ and $\left|\frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right|$ respectively for the two situations enumerated in (2.35).

For the situation that $\xi(L)=0$ and $\eta(L)>0$, by (2.21), we can obtain

$$\begin{aligned}\psi_2(L)=&\sum_{i=1}^{L}\left(\lambda_i+\frac{1}{\lambda_i}-2\right)+2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\\ \geq&\sum_{i=1}^{L}\left(\lambda_i+\frac{1}{\lambda_i}-2\right)\geq\left(1+\nu+\frac{1}{1+\nu}-2\right)\eta(L)=\frac{\nu^2}{1+\nu}\eta(L).\end{aligned}\tag{2.37}$$

Consequently, (2.23), (2.24), (2.37) and *Lemma 1* yield the upper bounds on $\left|\frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right|$ and $\left|\frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right|$

$$\left|\frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right|\leq\frac{1}{C_1}L\boldsymbol{\Omega}_{\boldsymbol{\Sigma}_L}^{(\varphi(L))}\tag{2.38}$$

$$\left|\frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right|\leq\frac{\frac{1}{\varepsilon_0}L\left[\boldsymbol{\Omega}_{\boldsymbol{\Sigma}_L}^{(\varphi(L))}\right]^{\frac{1}{2}}}{\frac{\nu^2}{1+\nu}\eta(L)}=\frac{1+\nu}{\varepsilon_0\nu^2}\left[\frac{L^2}{\eta^2(L)}\boldsymbol{\Omega}_{\boldsymbol{\Sigma}_L}^{(\varphi(L))}\right]^{\frac{1}{2}}\tag{2.39}$$

where C_1 was defined in *Lemma 1*.

For the situation that $\xi(L)>0$ and $\eta(L)\geq 0$, the corresponding lower bound on $\psi_1^{(\varphi(L))}(L)$ is obtained from (2.15) as

$$\psi_1^{(\varphi(L))}(L)=\mathrm{tr}\left[\left(\mathbf{B}_L^{(\varphi(L))}\right)^2\right]+2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-2}\mu_L\geq 2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-2}\mu_L\geq 2\varepsilon_0^2\boldsymbol{\mu}_L^T\boldsymbol{\mu}_L\geq 2\varepsilon_0^2\nu^2\xi(L).\tag{2.40}$$

Similarly, we have

$$\psi_2(L)=\sum_{i=1}^{L}\left(\lambda_i+\frac{1}{\lambda_i}-2\right)+2\boldsymbol{\mu}_L^T\boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\geq\frac{\nu^2}{1+\nu}\eta(L)+2\varepsilon_0\nu^2\xi(L).\tag{2.41}$$

Hence, the corresponding upper bounds on $\left|\frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right|$ and $\left|\frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right|$ for this situation can be expressed as

$$\left|\frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right| \leq \frac{1}{2\varepsilon_0^2\nu^2}\frac{L\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}}{\xi(L)} \tag{2.42}$$

$$\left|\frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right| \leq \frac{\frac{1}{\varepsilon_0}L\left[\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}\right]^{\frac{1}{2}}}{\frac{\nu^2}{1+\nu}\eta(L)+2\varepsilon_0\nu^2\xi(L)} = \frac{1+\nu}{\varepsilon_0\nu^2}\left\{\frac{L^2}{\left[2\varepsilon_0(1+\nu)\xi(L)+\eta(L)\right]^2}\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}\right\}^{\frac{1}{2}}. \tag{2.43}$$

Taking into account the upper bounds on $\left|\frac{\delta_1^{(\varphi(L))}(L)}{\psi_1^{(\varphi(L))}(L)}\right|$ and $\left|\frac{\delta_2^{(\varphi(L))}(L)}{\psi_2(L)}\right|$ for the different situations, the conclusion that the ADR converges to unity follows from (2.36) and *Lemma 4*.

From (2.25), we can see that $\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}$ is a nonnegative decreasing function of truncation length. Thus, if $\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi_0)$ and $\lim_{L\to\infty}\frac{\varphi(L)}{\varphi_0(L)} \geq 1$, then

$$0 \leq \lim_{L\to\infty}\omega(L)\,\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))} \leq \lim_{L\to\infty}\omega(L)\,\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi_0(L))} = 0 \tag{2.44}$$

and hence, $\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi)$. This completes the proof. ■

Since $0 \leq \xi(L), \eta(L) \leq L$, then (2.35) implies that $\omega(L)$ is a non-decreasing function of L. Therefore, in order to satisfy the limit in (2.34), $\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}$ should be a decreasing function of L with a decay rate larger than the rate of increase of $\omega(L)$. Furthermore, we can see that the growth rates of $\omega_1(L)$ and $\omega_2(L)$ with L are smaller or equal to L^2, with equality if and only if $\xi(L)$ and $\eta(L)$ do not grow with L which describes the least favorable situation that the difficulty of the hypothesis testing problem does not reduce as L increases. Hence, the sufficient conditions in (2.36) involve the smallest set of solutions when $\omega(L) = L^2$ and we

call them the strongest sufficient conditions. In other words, if $\{\mathbf{\Sigma}_L\}$ satisfies the sufficient conditions in (2.36) with $\omega(L) = L^2$, then $\{\mathbf{\Sigma}_L\}$ satisfies the sufficient conditions for the other cases as well. Moreover, as (2.35) shows, $\xi(L)$ and $\eta(L)$ impact the rate of increase of $\omega(L)$ differently. For instance, if $\xi(L) = L$ and $\eta(L) = 0$, then the rate of increase of $\omega(L)$ is 0, while $\omega(L)$ is proportional to L when $\xi(L) = 0$ and $\eta(L) = L$. This seems consistent with our intuition that the shift in mean testing problems considered in (2.35) are often easier when compared to the change in covariance matrix testing problems (2.35) considers. The next section provides analysis for some specific well-accepted classes of models.

2.4 Illustrative Classes of Problems

The previous section reveals that if the sequence of covariance matrices $\{\mathbf{\Sigma}_L\}$ under hypothesis $\mathcal{H}_1$ is contained in $\mathcal{U}(\varphi)$ for some φ, we can employ a truncated detector instead of the optimal detector without performance loss when the size of the observation vector L increases to infinity. We notice that the sufficient conditions described by $\mathcal{U}(\varphi)$ in (2.34) are not expressed directly in terms of $\{\mathbf{\Sigma}_L\}$, but on $\left\{\mathbf{\Sigma}_L^{-1}\right\}$. Thus, the structure of the required $\{\mathbf{\Sigma}_L\}$ is not apparent. Nevertheless, we have found that the sufficient conditions are satisfied by several important classes of system and process models [17, 18, 37–39] with reasonable regularity conditions, provided an adequate truncation rule $\varphi(L)$ is chosen. We consider two well-accepted classes of models, which have been studied extensively in previous research projects, and elucidate that for these models, the performance of the truncated detector is asymptotically equivalent to that of the optimal untruncated detector for specific $\varphi(L)$ which increase slowly compared to L. In this section, we will just consider the strongest sufficient conditions, i.e. $\mathcal{U}(\varphi)$ with $\omega(L) = L^2$, and provide a $\varphi(L)$ which renders the

strongest sufficient conditions satisfied for the models. Following the same procedure, the corresponding results for the weaker sufficient conditions can be similarly obtained.

2.4.1 $\mathbf{\Sigma}_L$ with Banded Structure

One general class of practical models is based on the following assumption.

Assumption 3 *For a sequence of observations under hypothesis $\mathcal{H}_1$, each observation is only correlated to its neighbors with sufficiently close indices. For these sort of practical models, the covariance matrix $\mathbf{\Sigma}_L$ is a banded matrix with fixed bandwidth m. That is to say, for some $m < L$, $(\mathbf{\Sigma}_L)_{i,j} = 0$, $\forall\, |i-j| > \frac{m}{2}$.*

For spatial signals, *Assumption 3* models the situation that an observation taken at a given sensor is only correlated with observations from the sensors which are sufficiently close to it. A similar signal model has been employed for temporal signals, which deems that two observations are correlated only if the time interval between these two observations is not too long.

Generally, the inverse of a banded matrix is not banded. However the following well-known inequality for the inverse of a positive definite banded matrix from [42] will be useful in our analysis.

Lemma 5 *Let $\mathbf{A}$ be a positive definite banded matrix with band-width m, i.e. $\mathbf{A}_{i,j} = 0$ if $|i-j| > \frac{m}{2}$. Let $[a,b]$ be the smallest interval containing the spectrum of $\mathbf{A}$. Set $r = \frac{b}{a}$, $q = \frac{\sqrt{r}-1}{\sqrt{r}+1}$, $D_0 = \frac{\left(1+\sqrt{r}\right)^2}{2ar}$, and $\gamma = q^{\frac{2}{m}}$. Then we have*

$$\left|\left(\mathbf{A}^{-1}\right)_{i,j}\right| \leq D\gamma^{|i-j|} \tag{2.45}$$

where $D = \max\{a^{-1}, D_0\}$.

Using *Assumption 3* and *Lemma 5*, we give the following theorem.

Theorem 2 *As per (2.1), let* $\mathbf{\Sigma}_L$ *denote the covariance matrix of the observations under hypothesis* $\mathcal{H}_1$ *for a given* L *and consider the strongest sufficient conditions in (2.34) (*$\omega(L) = L^2$ *in (2.35)). Under* Assumptions 1, 2 and 3, $\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi_0)$ *for* $\varphi_0(L) = \left\lceil \frac{1+\kappa}{\ln \gamma^{-1}} \ln L \right\rceil$, *where* κ *is an arbitrary small positive constant and* $\gamma = \left(\frac{1-\varepsilon_0}{1+\varepsilon_0}\right)^{\frac{2}{m}}$. *Thus, the ADR of the truncated detector relative to the optimal detector converges to* 1. *Note that any other truncation rule* $\varphi(L)$ *such that* $\lim_{L\to\infty} \frac{\varphi(L)}{\varphi_0(L)} \geq 1$, *will also provide unity ADR.*

Proof: [Proof of Theorem 2] By *Lemma 5*, we have

$$\left|\left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}\right| \leq D\gamma^{|i-j|} \tag{2.46}$$

where $\gamma = \left(\frac{1-\varepsilon_0}{1+\varepsilon_0}\right)^{\frac{2}{m}} < 1$ and $D = \max\left\{\varepsilon_0^{-1}, \frac{\left(1+\varepsilon_0\right)^2}{2\varepsilon_0}\right\}$ are constants.

Employing $\omega(L) = L^2$ and $\varphi_0(L) = \left\lceil \frac{1+\kappa}{\ln \gamma^{-1}} \ln L \right\rceil$, the condition in (2.36) becomes

$$\begin{aligned} L^2 \mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi_0(L))} &= L^2 \max_i \sum_{j:|j-i|\geq\varphi_0(L)+1} \left|\left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}\right|^2 \leq L^2 \max_i \sum_{j:|j-i|\geq\varphi_0(L)+1} D^2\gamma^{2|i-j|} \\ &\leq 2L^2 \sum_{l=\varphi_0(L)+1}^{L-1} D^2\gamma^{2l} = 2L^2D^2 \frac{\left[1-\gamma^{2(L-\varphi_0(L)-1)}\right]}{1-\gamma^2}\gamma^{2(\varphi_0(L)+1)} \\ &\leq \frac{2D^2}{1-\gamma^2} L^2\gamma^{2\varphi_0(L)} \leq \frac{2D^2}{1-\gamma^2}\frac{1}{L^{2\kappa}}. \end{aligned} \tag{2.47}$$

Since $\frac{2D^2}{1-\gamma^2}$ and κ are positive constants, and $\mathbf{\Omega}_{\mathbf{\Sigma}}^{(\varphi(L))}$ is a nonnegative function, as L increases to infinity, we have

$$0 \leq \lim_{L\to\infty} L^2 \mathbf{\Omega}_{\mathbf{\Sigma}}^{(\varphi(L))} \leq \lim_{L\to\infty} \frac{2D^2}{1-\gamma^2}\frac{1}{L^{2\kappa}} = 0. \tag{2.48}$$

We therefore have proved that $\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi_0)$ for $\varphi_0(L) = \left\lceil \frac{1+\kappa}{\ln \gamma^{-1}} \ln L \right\rceil$, and hence the proof of the theorem is complete by applying *Theorem 1.* ■

Note that $\varphi_0(L) = \left\lceil \frac{1+\kappa}{\ln \gamma^{-1}} \ln L \right\rceil$ increases much slower than L, but *Theorem 2* demonstrates that its deflection performance is asymptotically equivalent to that of the optimal detector. This significant advantage of the truncated detector can provide underlying benefits in implementation in realistic problems. Furthermore, since we take the strongest sufficient conditions into account here, we can expect that the truncated detector with some truncation rule $\varphi(L)$, whose rate of increase is even slower than $\varphi_0(L)$, can also achieve unity ADR if $\xi(L)$ or $\eta(L)$ grow with L.

2.4.2 Wide-Sense Stationary Limiting Models after the Mean is Subtracted

The other general class of practical models under consideration is based on the assumption below.

Assumption 4 *Assume that as $L \to \infty$, $\mathbf{x}_L - \boldsymbol{\mu}_L$ approaches a wide-sense stationary random process with power spectral density $S(f)$ under hypothesis $\mathcal{H}_1$. Let $S^{(m)}(f)$ denote the m-th derivative of $S(f)$. We assume that $S(f)$ is bounded away from 0 and ∞, that is to say, $0 < \varepsilon \leq S(f) \leq \varepsilon^{-1}$ for some ε, and $\left\|S^{(m)}(f)\right\|_\infty \leq C$ for some $m > 1$, where $\|S(f)\|_\infty \stackrel{\Delta}{=} \sup_{f \in \left(-\frac{1}{2}, \frac{1}{2}\right)} |S(f)|$.*

Note that *Assumption 4* does not impose any restriction on $\boldsymbol{\mu}_L$. Due to this, the class of limiting processes defined by *Assumption 4* includes some that are not wide-sense stationary. The part of *Assumption 4* requiring $0 < \varepsilon \leq S(f) \leq \varepsilon^{-1}$ for some ε is similar to our previous *Assumption 1* that the eigenvalues of $\mathbf{\Sigma}_L$ are bounded by $\left[\varepsilon_0, \varepsilon_0^{-1}\right]$, since it is known from [43]

that the largest and smallest eigenvalues of $\mathbf{\Sigma}_L$ in the limit follow $\lim_{L\to\infty} \lambda_{\max} = \sup_f S(f)$ and $\lim_{L\to\infty} \lambda_{\min} = \inf_f S(f)$, where $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalues of $\mathbf{\Sigma}_L$ respectively.

Before investigating the performance of the truncated detector applied to this class of problems, we first introduce a lemma on the asymptotic behavior of the inverses of covariance matrices of wide-sense stationary processes described above. Some similar results can also be found in [44].

Lemma 6 *Define* $\tilde{r}(i-j) = \left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}$. *Under* Assumption 4, $\tilde{S}(f) = \sum_{t=-\infty}^{\infty} \tilde{r}(t)e^{-i2\pi ft} = \frac{1}{S(f)}$ *and* $\mathbf{\Sigma}_L^{-1}$ *is asymptotically a symmetric Toeplitz matrix.*

Proof: Suppose $\mathbf{\Sigma}_L^{-1}$ is a symmetric Toeplitz matrix when $L \to \infty$. Let $r(i-j)$ and $\tilde{r}(i-j)$ denote $(\mathbf{\Sigma}_L)_{i,j}$ and $\left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}$ respectively, and let $S(f) = \sum_{t=-\infty}^{\infty} r(t)e^{-i2\pi ft}$ and $\tilde{S}(f) = \sum_{t=-\infty}^{\infty} \tilde{r}(t)e^{-i2\pi ft}$ as $L \to \infty$.

Using the expression $\mathbf{\Sigma}_L\mathbf{\Sigma}_L^{-1} = \mathbf{I}$, we obtain the following equations

$$\begin{aligned} \mathbf{I}_{k,l} &= \sum_{i=-\infty}^{\infty} r(k-i)\tilde{r}(i-l) \\ &= \sum_{i=-\infty}^{\infty} r(k-l-i)\tilde{r}(i). \end{aligned} \tag{2.49}$$

Hence, we have

$$\begin{aligned} S(f)\tilde{S}(f) &= \sum_{t=-\infty}^{\infty} r(t)e^{-i2\pi ft} \sum_{k=-\infty}^{\infty} \tilde{r}(k)e^{-i2\pi fk} \\ &= \sum_{t=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} r(t)\tilde{r}(k)e^{-i2\pi f(k+t)} \\ &= \sum_{l=-\infty}^{\infty} \left[\sum_{k=-\infty}^{\infty} r(l-k)\tilde{r}(k) \right] e^{-i2\pi fl} \\ &= 1 \end{aligned} \tag{2.50}$$

Since $S(f)$ is bounded by $\left[\varepsilon, \varepsilon^{-1}\right]$, we immediately have $\tilde{S}(f) = \frac{1}{S(f)}$, and we can calculate $\tilde{r}(t) = \int_{-1/2}^{1/2} \tilde{S}(f)e^{i2\pi ft}df$ which implies $\mathbf{\Sigma}_L^{-1}$ is a symmetric Toeplitz matrix when $L \to \infty$ as assumed. By the uniqueness of the inverse of $\mathbf{\Sigma}_L$, $\mathbf{\Sigma}_L^{-1}$ is a symmetric Toeplitz matrix. ■

The following theorem addresses the asymptotic equivalence of the truncated detector to the optimal detector in terms of deflection.

Theorem 3 *Consider the strongest sufficient conditions in (2.34) ($\omega(L) = L^2$ in (2.35)). Given a sequence of covariance matrices $\{\mathbf{\Sigma}_L\}$ in (2.1) satisfying* Assumptions 1, 2 and 4, *$\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi_0)$ for $\varphi_0(L) = \left\lceil L^{\frac{2+\alpha}{2m-1}} \right\rceil$, where α is an arbitrary small positive constant. Consequently, the ADR of the truncated detector with $\varphi_0(L)$ relative to the optimal detector converges to unity. Note that any other $\varphi(L)$ such that $\lim_{L\to\infty} \frac{\varphi(L)}{\varphi_0(L)} \geq 1$, will also provide unity ADR.*

Proof: [Proof of Theorem 3] Under *Assumption 4*, $S(f)$ is bounded by $\left[\varepsilon, \varepsilon^{-1}\right]$ for some $\varepsilon > 0$ and $\left\|S^{(m)}(f)\right\|_\infty \leq C$ for some $m > 1$. Therefore by *Lemma 6*, $\tilde{S}(f)$ is also bounded by $\left[\varepsilon, \varepsilon^{-1}\right]$ and $\left\|\tilde{S}^{(m)}(f)\right\|_\infty \leq \tilde{C}$, where $\tilde{C}$ is a constant.

Let $\tilde{r}(i-j) = \left(\mathbf{\Sigma}_L^{-1}\right)_{i,j}$. Since $\tilde{r}(t) = \tilde{r}(-t)$ with $\tilde{r}(t)$ and $\tilde{S}(f)$ a Fourier transform pair, we can obtain

$$\tilde{S}^{(2l)}(f) = \tilde{S}^{(2l)}(-f) \tag{2.51}$$

$$\tilde{S}^{(2l+1)}\left(\frac{1}{2}\right) = \tilde{S}^{(2l+1)}\left(-\frac{1}{2}\right) = 0. \tag{2.52}$$

Thus, utilizing integration by parts, we can obtain the following upper bound on $|\tilde{r}(t)|$

$$\begin{aligned} |\tilde{r}(t)| &= \left| \int_{-1/2}^{1/2} \tilde{S}(f) e^{i2\pi ft} df \right| = \left| \int_{-1/2}^{1/2} \frac{\tilde{S}^{(m)}(f)}{(i2\pi t)^m} e^{i2\pi ft} df \right| \\ &\leq \left\| \tilde{S}^{(m)}(f) \right\|_\infty \frac{1}{(2\pi t)^m} \int_{-1/2}^{1/2} \left| \frac{e^{i2\pi ft}}{i^m} \right| df \leq \frac{\tilde{C}}{(2\pi t)^m}. \end{aligned} \tag{2.53}$$

As a result, for $\omega(L) = L^2$ and $\varphi_0(L) = \left\lceil L^{\frac{2+\alpha}{2m-1}} \right\rceil$, the condition in (2.36) becomes

$$\begin{aligned} 0 \leq \lim_{L\to\infty} L^2 \mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi_0(L))} &= \lim_{L\to\infty} L^2 \max_i \sum_{j:|j-i|\geq\varphi_0(L)+1} |\tilde{r}(i-j)|^2 \qquad (2.54) \\ &\leq \lim_{L\to\infty} 2L^2 \sum_{t\geq\varphi_0(L)+1}^{L-1} |\tilde{r}(t)|^2 \leq \lim_{L\to\infty} 2L^2 \sum_{t\geq\varphi_0(L)+1}^{L-1} \frac{\tilde{C}^2}{(2\pi t)^{2m}} \\ &\leq \lim_{L\to\infty} \frac{2\tilde{C}^2 L^2}{(2m-1)(2\pi)^{2m}} \left\{ \frac{1}{[\varphi_0(L)]^{2m-1}} - \frac{1}{(L-2)^{2m-1}} \right\} \qquad (2.55) \\ &\leq \lim_{L\to\infty} \frac{2\tilde{C}^2 L^2}{(2m-1)(2\pi)^{2m}} \frac{1}{[\varphi_0(L)]^{2m-1}} \\ &\leq \lim_{L\to\infty} \frac{2\tilde{C}^2}{(2m-1)(2\pi)^{2m}} \frac{1}{L^\alpha} = 0 \end{aligned}$$

where (2.55) is obtained by using an integral to bound the sum.

Consequently, it is clear that $\{\mathbf{\Sigma}_L\} \in \mathcal{U}(\varphi_0)$ for $\varphi_0(L) = \left\lceil L^{\frac{2+\alpha}{2m-1}} \right\rceil$.

Invoking *Theorem 1* again, we conclude the proof. Thus when we employ the truncated detector with $\varphi(L) \geq \varphi_0(L) = \left\lceil L^{\frac{2+\alpha}{2m-1}} \right\rceil$ for the class of problems which satisfy *Assumptions 1, 2 and 4*, the ADR of the truncated detector relative to the optimal detector converges to unity. ■

We also can see that the truncation length of the truncated detector, described by $\varphi_0(L) = \left\lceil L^{\frac{2+\alpha}{2m-1}} \right\rceil$, increases slower than the growth of L. In addition, since the strongest sufficient conditions are considered here, the minimum requirement for the truncation rule can be further reduced for the problems where $\xi(L)$ or $\eta(L)$ are increasing functions of L.

Assumption 4 is satisfied by a very large class of well-studied and well-accepted wide-sense stationary limiting models. As a particular example, we consider $\mathbf{x}_L$ generated with autoregressive moving average (ARMA) models.

Let $\{e_i\}$ denote a real sequence of independent random variables with zero mean and variance σ_S^2. An ARMA(p,q) process $\{x_i\}$ can be defined by

$$x_i = \sum_{l=1}^{p} \phi_l x_{i-l} + \sum_{t=1}^{q} \theta_t e_{i-t} + e_i. \tag{2.56}$$

Accordingly, the power spectral density of the ARMA(p,q) process can be expressed as

$$S(f) = \frac{\sigma_S^2 \left| 1 + \sum_{t=1}^{q} \theta_t \exp(-i2\pi f t) \right|^2}{\left| 1 - \sum_{l=1}^{p} \phi_l \exp(-i2\pi f l) \right|^2} \qquad \text{for} \quad |f| < \frac{1}{2}. \tag{2.57}$$

Thus, with appropriate $\{\theta_t\}$ and $\{\phi_l\}$, $S(f)$ can easily satisfy *Assumption 4*.

2.5 Implementation of Consensus based Truncated Detector

In this section, we will briefly discuss how to apply a two-step consensus algorithm to compute the truncated test statistic (2.3) with truncation rule $\varphi(L)$ at each sensor. For simplicity, we assume ideal communication channels.

2.5.1 Initialization of Local Statistics

Let $\mathbf{y}(t) = [y_1(t), y_2(t), ..., y_L(t)]^T$ denote a vector of local statistics at time t. In this step, each sensor collects the observations from $2\varphi(L)$ neighbors, and then computes the inner sum of $\sum_{i=1}^{L}\sum_{j=1}^{L} x_i\left(\mathbf{B}_L^{(\varphi(L))}\right)_{i,j} x_j$ from (2.3) as its initial local statistic. To be specific, for the i-th sensor, its initial local statistic can be written as

$$y_i(0) = \sum_{j=1}^{L} x_i\left(\mathbf{B}_L^{(\varphi(L))}\right)_{i,j} x_j + \zeta_i x_i = \sum_{j:|i-j|\leq\varphi(L)} x_i(\mathbf{R}_L)_{i,j} x_j + \zeta_i x_i \tag{2.58}$$

with ζ_i a constant from (2.2).

2.5.2 Consensus Procedure

After initialization, a standard consensus algorithm can be applied to compute the truncated test statistic (2.3). Here, we assume a synchronous time model [23], in which time is assumed to be slotted commonly across sensors. In each slot, each sensor received its neighbors' local statistics and updates its own local statistic. The updating rule for the local statistics can be expressed as

$$y_i(t+1) = \mathbf{W}_{i,i} y_i(t) + \sum_{j:|i-j|\leq\varphi(L)} \mathbf{W}_{i,j} y_j(t) \tag{2.59}$$

where $t = 1, 2, ...$ and $\mathbf{W}_{i,j}$ is the weight on the local statistic of the j-th sensor. Thus, the corresponding compact vector form is

$$\mathbf{y}(t+1) = \mathbf{W}\mathbf{y}(t) = \mathbf{W}^t\mathbf{y}(0). \tag{2.60}$$

Theorem 4 *For any doubly stochastic matrix* $\mathbf{W} \in \mathcal{W}$ *such that* $\rho\left(\mathbf{W} - \frac{1}{L}\mathbf{1}\mathbf{1}^T\right) < 1$, *then we have*

$$\lim_{t\to\infty} \mathbf{y}(t) = \frac{1}{L}T_L^{\mathrm{tr}} \cdot \mathbf{1} \tag{2.61}$$

where $\rho(\cdot)$ *denotes the spectral radius of a matrix.*

The proof is provided in [21, 22].

Theorem 4 demonstrates that every sensor's local statistic converges to a scaled version of the truncated test statistic in (2.3), and hence each sensor can make its own decision based on its own local statistic while achieving the same performance as the truncated detector implemented in centralized manner. Furthermore, as [21] indicates, if we just consider symmetric $\mathbf{W}$, we can easily find the best choice of $\mathbf{W}$ making the consensus procedure have the fastest speed of convergence by solving a convex problem.

It is worth mentioning that though we assumed ideal communication channels and a deterministic weight matrix $\mathbf{W}$ as an example here, other consensus algorithms and their corresponding performance analysis for non-ideal channels [25, 26] or random weight matrices [23] can be directly adopted in our two-step consensus algorithms to compute our truncated test statistic.

2.6 Numerical Results

To illustrate our theoretical results, here we present a few numerical examples involving the cases studied in the previous theorems.

2.6.1 Signals with Banded Covariance Matrices

We first consider a case where $\mathbf{\Sigma}_L$ is banded and the bandwidth is fixed for all L. Our particular example assumes a stationary signal with triangular correlation. For any L, $\boldsymbol{\mu}_L = \mathbf{0}$, and $\mathbf{\Sigma}_L$=$0.025\mathbf{\Sigma}_{L,s} + \mathbf{I}$, where

$$(\mathbf{\Sigma}_{L,s})_{i,j} = \begin{cases} 1 - \frac{|i-j|}{20}, & |i-j| < 20 \\ 0, & otherwise \end{cases} . \tag{2.62}$$

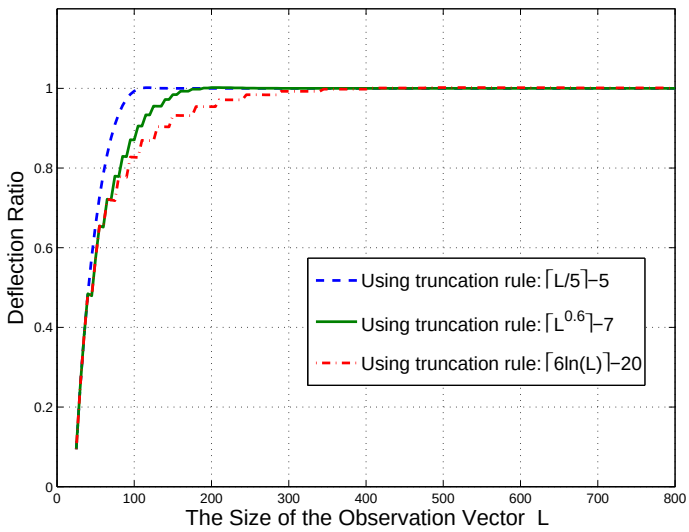

Figure 2.1: Deflection ratio of the truncated detector relative to the optimal detector for a triangularly correlated signal.

Figure 2.1 shows the deflection ratios (DR) of some truncated detectors with different $\varphi(L)$ which all satisfy our sufficient conditions for unity ADR. The deflection ratios of the

truncated detectors with truncation rules $\varphi(L) = \lceil \frac{L}{5} \rceil - 5$, $\varphi(L) = \lceil L^{0.6} \rceil - 7$, and $\varphi(L) = \lceil 6 \ln L \rceil - 20$ are plotted in dash, solid and dot-dash curves respectively. It is seen that the numerical results agree with our analytical prediction that ADR equals to 1. As expected, Figure 2.1 depicts that the larger the value of $\varphi(L)$ for a given L, the better deflection ratio performance that the corresponding truncated detector enjoys.

Figure 2.2 shows the relationship between the deflection ratio and the ROC curve[1] for various truncated detectors when $L = 1000$. It is seen that larger deflection ratio implies better performance in terms of the ROC curve in this example. Furthermore, we investigate the detection probability of some truncated detectors with different $\varphi(L)$ which all satisfy our sufficient conditions for unity ADR. Figure 2.3 illustrates that as L increases, the detection probability performance of each truncated detector converges to that of the optimal detector. Moreover, the larger the value of $\varphi(L)$ for large L, the faster the rate of convergence to the optimal detector. On the other hand, Figure 2.3 also shows that as L increases, the detection probability performance of the truncated detector with constant truncation length diverges from that of the optimal detector.

2.6.2 Autoregressive Moving Average Models

To illustrate *Theorem 3*, we use an ARMA(1,1) model. In our numerical results, $\boldsymbol{\mu}_L = \mathbf{0}$ and $\boldsymbol{\Sigma}_L = 0.01\boldsymbol{\Sigma}_L^{ARMA} + \mathbf{I}$, where $\boldsymbol{\Sigma}_L^{ARMA}$ is the covariance matrix of the ARMA(1,1) model with dimension L. The parameters in (2.56) are taken as $\phi_1 = 0.8$, $\theta_1 = 0.4$ and $\sigma_S^2 = 1$.

Figure 2.4 shows the deflection ratios of two truncated detectors with different $\varphi(L)$ which both satisfy our sufficient conditions for unity ADR. We can see that the numerical

[1]In the ROC curve, the false alarm probability $P_{\text{fa}} \triangleq \Pr(\text{Declare } \mathcal{H}_1 \,|\mathcal{H}_0 \text{ is true})$ and the detection probability is defined as $P_{\text{d}} \triangleq \Pr(\text{Declare } \mathcal{H}_1 \,|\mathcal{H}_1 \text{ is true})$.

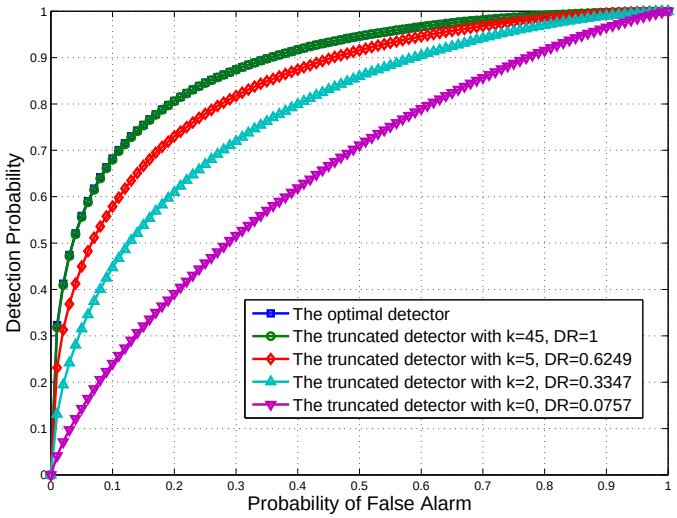

Figure 2.2: Relationship between the deflection ratio and the ROC (L=1000).

results agree with our analytical prediction that ADR equals to 1. As expected, Figure 2.4 also depicts that the truncated detector with larger $\varphi(L)$ has better deflection ratio performance for that L.

Figure 2.5 shows the relationship between the deflection ratio and the ROC curve for various truncated detectors when $L = 1000$. We can see from the figure that larger deflection ratio implies better performance in terms of the ROC curve in this example also. Figure 2.6 illustrates the detection probability performance of each truncated detector. It is seen that as L increases, the detection probability performance of the truncated detectors converges to that of the optimal detector. Moreover, the truncated detector with larger $\varphi(L)$ achieves better detection probability performance. However, the difference between the the detection probability performance of the optimal detector and that of the truncated detector with constant truncation length becomes larger and larger as L increases.

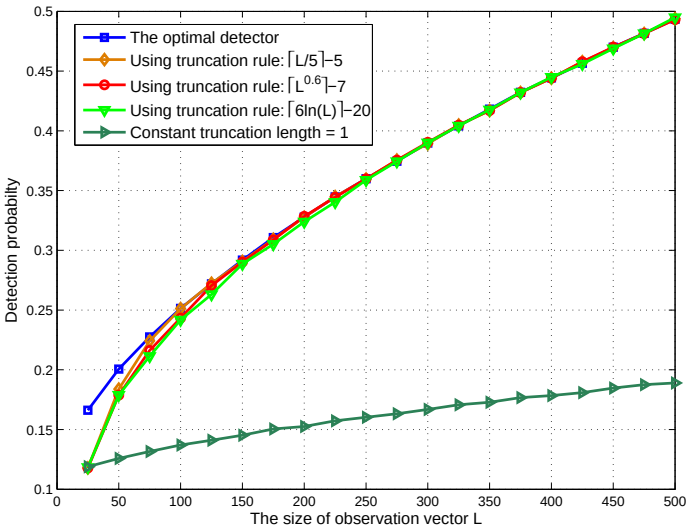

Figure 2.3: Detection probabilities of several truncated detectors ($P_{\text{fa}} = 0.1$).

2.7 Extension to Spatially and Temporally Correlated Gaussian Observations

Then the general hypothesis testing problem for spatially and temporally correlated Gaussian observations can be expressed as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{z}_{(K,L)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\ \mathcal{H}_1 &: \mathbf{z}_{(K,L)} \sim \mathcal{N}(\boldsymbol{\mu}_{(K,L)}, \boldsymbol{\Sigma}_{(K,L)}). \end{aligned} \tag{2.63}$$

The following assumption is made in this section.

Assumption 5 *$\boldsymbol{\mu}_{(K,L)}$, and $\boldsymbol{\Sigma}_{(K,L)}$ are known, and either $\boldsymbol{\mu}_{(K,L)} \neq \mathbf{0}$ or $\boldsymbol{\Sigma}_{(K,L)} \neq \mathbf{I}$. The elements of $\boldsymbol{\mu}_{(K,L)}$ are finite, and the eigenvalues of $\boldsymbol{\Sigma}_{(K,L)}$ are bounded by $\left[\varepsilon_0, \varepsilon_0^{-1}\right]$ for some positive number $\varepsilon_0 < 1$.*

Similar to (2.2), the optimal test statistic to minimize error probability, risk, or one of

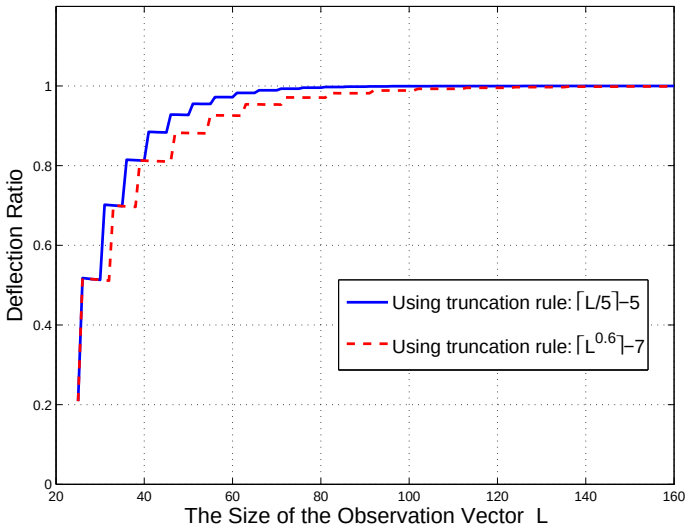

Figure 2.4: Deflection ratio of the truncated detector relative to the optimal detector for an ARMA(1,1) model with $\phi_1 = 0.8$, $\theta_1 = 0.4$ and $\sigma_S^2 = 1$.

several other criteria for the problem in (2.63) compares the log-likelihood ratio

$$T^{\text{opt}}_{(K,L)} = \mathbf{z}^T_{(K,L)}\mathbf{R}_{(K,L)}\mathbf{z}_{(K,L)} + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-1}_{(K,L)}\mathbf{z}_{(K,L)} = \sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{u}_i^T\mathbf{D}_{ij}\mathbf{u}_i + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-1}_{(K,L)}\mathbf{z}_{(K,L)} \tag{2.64}$$

to a threshold, where

$$\mathbf{R}_{(K,L)} = \mathbf{I} - \boldsymbol{\Sigma}^{-1}_{(K,L)} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} & \cdots & \mathbf{D}_{1K} \\ \mathbf{D}_{21} & \mathbf{D}_{22} & \cdots & \mathbf{D}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{D}_{K1} & \mathbf{D}_{K2} & \cdots & \mathbf{D}_{KK} \end{pmatrix} \tag{2.65}$$

and $\mathbf{D}_{ij}$ is the (i,j)-th block of $\mathbf{R}_{(K,L)}$. Since $\boldsymbol{\Sigma}_{(K,L)}$ is assumed known in *Assumption 5*, every $\mathbf{D}_{ij}$ can be calculated beforehand. As in (2.3), our truncation rule will ignore those

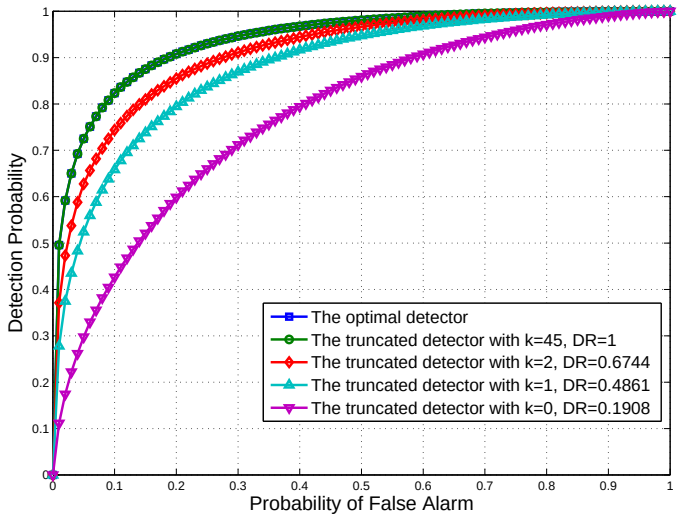

Figure 2.5: Relationship between the deflection ratio and the ROC (L=1000).

terms in (2.64) which involve $u_{i,l}$ and $u_{j,m}$ with $|l-m| > \varphi(L)$, and hence the truncated test statistic is

$$T^{\text{tr}}_{(K,L)} = \mathbf{z}^T_{(K,L)}\mathbf{B}^{(\varphi(L))}_{(K,L)}\mathbf{z}_{(K,L)} + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-1}_{(K,L)}\mathbf{z}_{(K,L)} = \sum_{i,j=1}^{K} \mathbf{u}_i^T\mathbf{E}^{(\varphi(L))}_{ij}\mathbf{u}_j + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-1}_{(K,L)}\mathbf{z}_{(K,L)} \tag{2.66}$$

where $\mathbf{B}^{(\varphi(L))}_{(K,L)}$ is the truncated matrix of $\mathbf{R}_{(K,L)}$, and the elements of the (i,j)-th block $\mathbf{E}^{(\varphi(L))}_{ij}$ of $\mathbf{B}^{(\varphi(L))}_{(K,L)}$ can be expressed as $\left(\mathbf{E}^{(\varphi(L))}_{ij}\right)_{p,q} \triangleq \begin{cases} (\mathbf{D}_{ij})_{p,q}, & \text{if } |p-q| \le \varphi(L) \\ 0, & \text{otherwise} \end{cases}$.

Consider the following generalization of *Assumption 2*.

Assumption 6 *As it would not make sense to consider truncated detectors whose truncation makes the two hypotheses indistinguishable, the truncation rule $\varphi(L)$ is required to ensure that if $\boldsymbol{\mu}_{(K,L)} = \mathbf{0}$ and $\mathbf{R}_{(K,L)} \neq \mathbf{0}$, then the matrix $\mathbf{B}^{(\varphi(L))}_{(K,L)} \neq \mathbf{0}$.*

By (2.64) and (2.66), as the number of sensors increases to infinity, the asymptotic

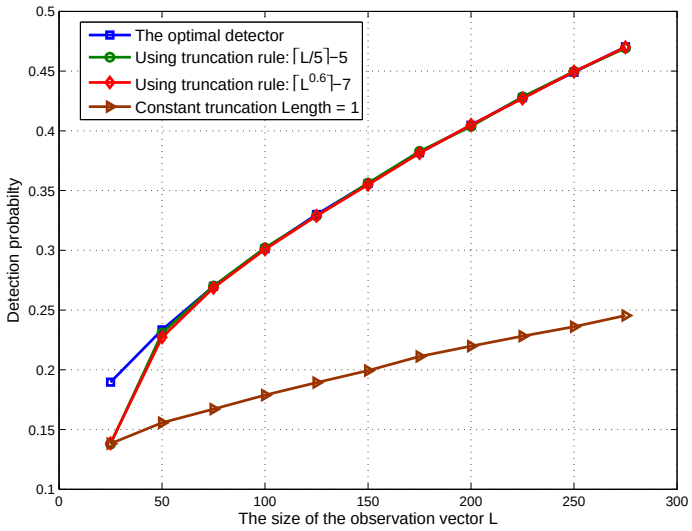

Figure 2.6: Detection probability performance of the truncated detectors ($P_{\text{fa}} = 0.1$).

deflection ratio of the truncated detector relative to the optimal detector is

$$\Lambda(T^{\text{tr}}_{(K,\infty)}, T^{\text{opt}}_{(K,\infty)}) = \lim_{L\to\infty} \frac{\text{tr}\left(\mathbf{R}^2_{(K,L)}\right) + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-2}_{(K,L)}\boldsymbol{\mu}_{(K,L)}}{\text{tr}\left[\left(\mathbf{B}^{(\varphi(L))}_{(K,L)}\right)^2\right] + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-2}_{(K,L)}\boldsymbol{\mu}_{(K,L)}} \tag{2.67}$$

$$\left[\frac{\text{tr}\left(\mathbf{B}^{(\varphi(L))}_{(K,L)}\boldsymbol{\Sigma}_{(K,L)} - \mathbf{B}^{(\varphi(L))}_{(K,L)}\right) + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-1}_{(K,L)}\boldsymbol{\mu}_{(K,L)}}{\text{tr}\left(\mathbf{R}_{(K,L)}\boldsymbol{\Sigma}_{(K,L)} - \mathbf{R}_{(K,L)}\right) + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-1}_{(K,L)}\boldsymbol{\mu}_{(K,L)}}\right]^2$$

$$= \lim_{L\to\infty}\left(1 + \frac{\tilde{\delta}^{(\varphi(L))}_1(L)}{\tilde{\psi}^{(\varphi(L))}_1(L)}\right)\left(1 - \frac{\tilde{\delta}^{(\varphi(L))}_2(L)}{\tilde{\psi}_2(L)}\right)^2$$

where

$$\tilde{\delta}^{(\varphi(L))}_1(L) \triangleq \text{tr}\left(\mathbf{R}^2_{(K,L)}\right) - \text{tr}\left[\left(\mathbf{B}^{(\varphi(L))}_{(K,L)}\right)^2\right], \tag{2.68}$$

$$\tilde{\psi}^{(\varphi(L))}_1(L) \triangleq \text{tr}\left[\left(\mathbf{B}^{(\varphi(L))}_{(K,L)}\right)^2\right] + 2\boldsymbol{\mu}^T_{(K,L)}\boldsymbol{\Sigma}^{-2}_{(K,L)}\boldsymbol{\mu}_{(K,L)}, \tag{2.69}$$

$$\tilde{\delta}^{(\varphi(L))}_2(L) \triangleq \text{tr}\left[\left(\mathbf{R}_{(K,L)} - \mathbf{B}^{(\varphi(L))}_{(K,L)}\right)\left(\boldsymbol{\Sigma}_{(K,L)} - \mathbf{I}\right)\right], \tag{2.70}$$

and

$$\tilde{\psi}_2(L) \triangleq \mathrm{tr}\left(\mathbf{R}_{(K,L)}\boldsymbol{\Sigma}_{(K,L)} - \mathbf{R}_{(K,L)}\right) + 2\boldsymbol{\mu}_{(K,L)}^T\boldsymbol{\Sigma}_{(K,L)}^{-1}\boldsymbol{\mu}_{(K,L)}. \tag{2.71}$$

Lemma 7 *Under* Assumption 5, *consider a truncated detector with truncation rule* $\varphi(L)$ *for the hypothesis testing problem (2.63). Upper bounds on the absolute values of* $\tilde{\delta}_1^{(\varphi(L))}(L)$ *and* $\delta_2^{(\varphi(\tilde{L}))}(L)$ *are given by*

$$\left|\tilde{\delta}_1^{(\varphi(L))}(L)\right| \leq K^2 L \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))} \tag{2.72}$$

and

$$\left|\tilde{\delta}_2^{(\varphi(L))}(L)\right| \leq \frac{KL}{\varepsilon_0}\left[\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))}\right]^{\frac{1}{2}} \tag{2.73}$$

where $\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))}$ *is defined by*

$$\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))} \triangleq \max_{i,j,p} \sum_{q:|p-q|\geq\varphi(L)+1} \left(\mathbf{D}_{ij}\right)_{p,q}^2. \tag{2.74}$$

Proof: [Proof of Lemma 7] Noting that $\mathrm{tr}\left(\mathbf{R}_{(K,L)}^2\right) = \sum_{i,j=1}^{K} \mathrm{tr}\left(\mathbf{D}_{ij}^2\right)$ and $\mathrm{tr}\left[\left(\mathbf{B}_{(K,L)}^{(\varphi(L))}\right)^2\right] = \sum_{i,j=1}^{K} \mathrm{tr}\left[\left(\mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right]$, we can obtain

$$\begin{aligned}\tilde{\delta}_1^{(\varphi(L))}(L) &= \mathrm{tr}\left(\mathbf{R}_{(K,L)}^2\right) - \mathrm{tr}\left[\left(\mathbf{B}_{(K,L)}^{(\varphi(L))}\right)^2\right] = \sum_{i,j=1}^{K} \mathrm{tr}\left[\mathbf{D}_{ij}^2 - \left(\mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right] \\ &= \sum_{i,j=1}^{K} \mathrm{tr}\left[\left(\mathbf{D}_{ij} - \mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right] = \sum_{i,j=1}^{K}\sum_{p=1}^{L}\sum_{q:|p-q|\geq\varphi(L)+1} \left(\mathbf{D}_{ij}\right)_{p,q}^2. \end{aligned} \tag{2.75}$$

Hence, it is easy to see that

$$\left|\tilde{\delta}_1^{(\varphi(L))}(L)\right| = \left|\sum_{i,j=1}^{K}\sum_{p=1}^{L}\sum_{q:|p-q|\geq\varphi(L)+1} \left(\mathbf{D}_{ij}\right)_{p,q}^2\right| \leq K^2 L \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))} \tag{2.76}$$

Using *Assumption 5*, (2.74) and the same bounding approach used in (2.28) yields

$$\left[\tilde{\delta}_2^{(\varphi(L))}(L)\right]^2 = \left\{\mathrm{tr}\left[\left(\mathbf{R}_{(K,L)} - \mathbf{B}_{(K,L)}^{(\varphi(L))}\right)\boldsymbol{\Sigma}_{(K,L)}\right]\right\}^2 \leq \mathrm{tr}\left[\left(\mathbf{R}_{(K,L)} - \mathbf{B}_{(K,L)}^{(\varphi(L))}\right)^2\right]\mathrm{tr}\left(\boldsymbol{\Sigma}_{(K,L)}^2\right)$$
$$\leq \frac{L}{\varepsilon_0^2}\sum_{i,j=1}^{K}\mathrm{tr}\left[\left(\mathbf{D}_{ij} - \mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right] \leq \frac{K^2L^2}{\varepsilon_0^2}\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))}, \tag{2.77}$$

which implies

$$\left|\tilde{\delta}_2^{(\varphi(L))}(L)\right| \leq \frac{KL}{\varepsilon_0}\left[\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\Sigma}_{(K,L)}}^{(\varphi(L))}\right]^{\frac{1}{2}}. \tag{2.78}$$

■

Let ν denote a positive small constant, such that $0 < \nu \ll 1$. Define

$$\tilde{\xi}(L) \triangleq \sum_{i=1}^{L}\mathbb{1}\left(\left|\mu_{(K,L),i}\right| \geq \nu\right) \text{ and } \tilde{\eta}(L) \triangleq \sum_{i=1}^{L}\mathbb{1}\left(\left|\lambda_{(K,L),i} - 1\right| \geq \nu\right), \tag{2.79}$$

where $\mu_{(K,L),i}$ and $\lambda_{(K,L),i}$ are the i-th element of $\boldsymbol{\mu}_{(K,L)}$ and the i-th largest eigenvalue of $\boldsymbol{\Sigma}_{(K,L)}$ respectively. Before proceeding, define

$$\tilde{\omega}(L) \triangleq \begin{cases} \max\left\{L, \frac{L^2}{\tilde{\eta}^2(L)}\right\}, \text{ if } \tilde{\xi}(L) = 0 \text{ and } \tilde{\eta}(L) > 0 \\ \max\left\{\frac{L}{\tilde{\xi}(L)}, L^2\left[2\varepsilon_0(1+\nu)\tilde{\xi}(L) + \tilde{\eta}(L)\right]^{-2}\right\}, \text{ if } \tilde{\xi}(L) > 0 \text{ and } \tilde{\eta}(L) \geq 0 \end{cases}. \tag{2.80}$$

We now give a Theorem providing sufficient conditions for the general hypothesis testing problem (2.63) with spatially and temporally correlated Gaussian observations, under which the ADR converges to unity.

Theorem 5 *Given a sequence of covariance matrices* $\left\{\boldsymbol{\Sigma}_{(K,L)}\right\}$ *in (2.63) which satisfy* Assumptions 5 and 6 *and a sequence of truncated test statistics* $T_{(K,\infty)}^{\mathrm{tr}}$ *in (2.66) with truncation*

rule $\varphi_0(L)$*, sufficient conditions for* $\Lambda\left(T^{\mathrm{tr}}_{(K,\infty)}, T^{\mathrm{opt}}_{(K,\infty)}\right) = 1$ *are*

$$\left\{\mathbf{\Sigma}_{(K,L)}\right\} \in \tilde{\mathcal{U}}(\varphi_0), \tag{2.81}$$

where $\tilde{\mathcal{U}}(\varphi_0) \triangleq \left\{ \left\{\mathbf{\Sigma}_{(K,L)}\right\} \middle| \lim_{L\to\infty} \tilde{\omega}\left(L\right) \tilde{\mathbf{\Omega}}^{(\varphi_0(L))}_{\mathbf{\Sigma}_{(K,L)}} = 0 \right\}$*. If a given* $\varphi_0(L)$ *satisfies (2.81), then it follows that any truncation rule* $\varphi(L)$*, which satisfies* $\lim_{L\to\infty} \frac{\varphi(L)}{\varphi_0(L)} \geq 1$*, will also provide* $\Lambda\left(T^{\mathrm{tr}}_{(K,\infty)}, T^{\mathrm{opt}}_{(K,\infty)}\right) = 1$.

Noting that K is a finite constant, the proof of *Theorem 5* is very similar to that of *Theorem 1*. Hence, the proof is omitted here.

2.7.1 Separable Space-Time Covariance Model

The separable covariance model has been widely used in statistical modeling of space-time observations, which assumes that the space-time covariance can be factored into the product of a purely spatial covariance and a purely temporal covariance. Since the separable covariance model is nicely interpretable in practical problems and can facilitate computational procedures for large space-time observations set [45–48], we will investigate the performance of the truncated detector based on this model in the following part.

Assumption 7 *The covariance between two observations* $u_{i,l}$ *and* $u_{j,m}$ *can be expressed as*

$$\mathrm{cov}\left(u_{i,l}, u_{j,m}\right) = (\mathbf{\Sigma}_K)_{i,j}(\mathbf{\Sigma}_L)_{l,m}, \tag{2.82}$$

where $\mathbf{\Sigma}_K$ *and* $\mathbf{\Sigma}_L$ *are the purely temporal covariance matrix for a given sensor and the purely spatial covariance matrix for a given time epoch respectively. In addition, we assume that the eigenvalues of positive definite matrices* $\mathbf{\Sigma}_K$*,* $\mathbf{\Sigma}_L$ *and* $\mathbf{\Sigma}_{(K,L)}$ *are bounded by* $\left[\varepsilon_0, \varepsilon_0^{-1}\right]$ *for*

some positive number $\varepsilon_0 < 1$. *Hence, the covariance matrix of* $\mathbf{z}_{(K,L)}$ *can be obtained as*[1]

$$\mathbf{\Sigma}_{(K,L)} = \mathbf{\Sigma}_K \otimes \mathbf{\Sigma}_L. \tag{2.83}$$

According to *Assumption 7*, the optimal test statistic in (2.64) can be rewritten as

$$\begin{aligned} T^{\text{opt}}_{(K,L)} &= \mathbf{z}^T_{(K,L)}\mathbf{R}_{(K,L)}\mathbf{z}_{(K,L)} + 2\boldsymbol{\mu}^T_{(K,L)}\mathbf{\Sigma}^{-1}_{(K,L)}\mathbf{z}_{(K,L)} \\ &= \sum_{i=1}^{K} \mathbf{u}_i^T \left(\mathbf{I} - \rho_{ii}\mathbf{\Sigma}_L^{-1}\right)\mathbf{u}_i - \sum_{i,j=1,i\neq j}^{K} \rho_{ij}\mathbf{u}_i^T\mathbf{\Sigma}_L^{-1}\mathbf{u}_j + 2\boldsymbol{\mu}^T_{(K,L)}\mathbf{\Sigma}^{-1}_{(K,L)}\mathbf{z}_{(K,L)}, \end{aligned} \tag{2.84}$$

where $\mathbf{R}_{(K,L)} \triangleq \mathbf{I} - \mathbf{\Sigma}^{-1}_{(K,L)} = \mathbf{I} - \mathbf{\Sigma}_K^{-1} \otimes \mathbf{\Sigma}_L^{-1}$ and $\rho_{ij} \triangleq \left(\mathbf{\Sigma}_K^{-1}\right)_{i,j}$. Furthermore, the elements of the matrix $\mathbf{E}_{ij}^{(\varphi(L))}$in (2.66) can be expressed as

$$\left(\mathbf{E}_{ij}^{(\varphi(L))}\right)_{p,q} = \begin{cases} \rho_{ij}\left(\mathbf{\Sigma}_L^{-1}\right)_{p,q}, & \text{if } i \neq j,\ |p-q| \leq \varphi(L) \\ \left(\mathbf{I} - \rho_{ii}\mathbf{\Sigma}_L^{-1}\right)_{p,q}, & \text{if } i = j,\ |p-q| \leq \varphi(L) \\ 0, & \text{otherwise} \end{cases}. \tag{2.85}$$

Lemma 8 *Under* Assumption 5, 6 and 7, *consider a truncated detector with truncation rule* $\varphi(L)$ *for the hypothesis testing problem (2.63). Upper bounds on the absolute values of* $\tilde{\delta}_1^{(\varphi(L))}(L)$ *and* $\delta_2^{(\varphi(\tilde{L}))}(L)$ *are given by*

$$\left|\tilde{\delta}_1^{(\varphi(L))}(L)\right| \leq \rho^2_{\max} K^2 L \mathbf{\Omega}^{(\varphi(L))}_{\mathbf{\Sigma}_L} \tag{2.86}$$

and

$$\left|\tilde{\delta}_2^{(\varphi(L))}(L)\right| \leq \frac{\rho_{\max} KL}{\varepsilon_0}\left[\mathbf{\Omega}^{(\varphi(L))}_{\mathbf{\Sigma}_L}\right]^{\frac{1}{2}}, \tag{2.87}$$

[1] $\otimes$ denotes kronecker product.

where $\rho_{\max} \triangleq \max_{i,j} \left|\rho_{ij}\right|$ *and* $\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}$ *is defined in (2.25). Thus, the sufficient conditions described by* $\tilde{\mathcal{U}}(\varphi_0)$ *in* Theorem 5 *can be simplified to*

$$\left\{\mathbf{\Sigma}_{(K,L)}\right\} \in \mathcal{U}^*(\varphi_0) \triangleq \left\{ \left\{\mathbf{\Sigma}_{(K,L)}\right\} \middle| \lim_{L\to\infty} \tilde{\omega}\left(L\right) \mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi_0(L))} = 0 \right\}. \tag{2.88}$$

Proof: [Proof of Lemma 7] Noting that

$$\mathrm{tr}\left(\mathbf{R}_{(K,L)}^2\right) = \mathrm{tr}\left[\left(\mathbf{I} - \mathbf{\Sigma}_K^{-1} \otimes \mathbf{\Sigma}_L^{-1}\right)^2\right] = \sum_{i=1}^{K} \mathrm{tr}\left[\left(\mathbf{I} - \rho_{ii}\mathbf{\Sigma}_L^{-1}\right)^2\right] + \sum_{i,j=1,i\neq j}^{K} \rho_{ij}^2 \mathrm{tr}\left[\left(\mathbf{\Sigma}_L^{-1}\right)^2\right] \tag{2.89}$$

and

$$\mathrm{tr}\left[\left(\mathbf{B}_{(K,L)}^{(\varphi(L))}\right)^2\right] = \sum_{i,j=1}^{K} \mathrm{tr}\left[\left(\mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right], \tag{2.90}$$

we can obtain

$$\begin{aligned}
\tilde{\delta}_1^{(\varphi(L))}\left(L\right) &= \mathrm{tr}\left(\mathbf{R}_{(K,L)}^2\right) - \mathrm{tr}\left[\left(\mathbf{B}_{(K,L)}^{(\varphi(L))}\right)^2\right] \\
&= \sum_{i=1}^{K} \mathrm{tr}\left[\left(\mathbf{I} - \rho_{ii}\mathbf{\Sigma}_L^{-1}\right)^2\right] + \sum_{i,j=1,i\neq j}^{K} \mathrm{tr}\left[\left(\rho_{ij}\mathbf{\Sigma}_L^{-1}\right)^2\right] - \sum_{i,j=1}^{K} \mathrm{tr}\left[\left(\mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right] \\
&= \sum_{i=1}^{K} \mathrm{tr}\left[\left(\mathbf{I} - \rho_{ii}\mathbf{\Sigma}_L^{-1} - \mathbf{E}_{ii}^{(\varphi(L))}\right)^2\right] + \sum_{i,j=1,i\neq j}^{K} \mathrm{tr}\left[\left(\rho_{ij}\mathbf{\Sigma}_L^{-1} - \mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right] \\
&= \sum_{i,j=1}^{K} \rho_{ij}^2 \sum_{p=1}^{L} \sum_{q:|p-q|\geq\varphi(L)+1} \left[\left(\mathbf{\Sigma}_L^{-1}\right)_{p,q}\right]^2.
\end{aligned} \tag{2.91}$$

Hence, it is easy to see that

$$\left|\tilde{\delta}_1^{(\varphi(L))}\left(L\right)\right| = \left|\sum_{i,j=1}^{K} \rho_{ij}^2 \sum_{p=1}^{L} \sum_{q:|p-q|\geq\varphi(L)+1} \left[\left(\mathbf{\Sigma}_L^{-1}\right)_{p,q}\right]^2\right| \leq \rho_{\max}^2 K^2 L \mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}, \tag{2.92}$$

where $\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}$ is defined in (2.25).

Using *Assumption 5, 7* and the same bounding approach used in (2.28) yields

$$\begin{aligned}\left[\tilde{\delta}_2^{(\varphi(L))}(L)\right]^2 &= \left\{\mathrm{tr}\left[\left(\mathbf{R}_{(K,L)} - \mathbf{B}_{(K,L)}^{(\varphi(L))}\right)\mathbf{\Sigma}_{(K,L)}\right]\right\}^2 \leq \mathrm{tr}\left[\left(\mathbf{R}_{(K,L)} - \mathbf{B}_{(K,L)}^{(\varphi(L))}\right)^2\right]\mathrm{tr}\left(\mathbf{\Sigma}_{(K,L)}^2\right) \\ &\leq \frac{L}{\varepsilon_0^2}\left\{\sum_{i=1}^{K}\mathrm{tr}\left[\left(\mathbf{I} - \rho_{ii}\mathbf{\Sigma}_L^{-1} - \mathbf{E}_{ii}^{(\varphi(L))}\right)^2\right] + \sum_{i,j=1,i\neq j}^{K}\mathrm{tr}\left[\left(\rho_{ij}\mathbf{\Sigma}_L^{-1} - \mathbf{E}_{ij}^{(\varphi(L))}\right)^2\right]\right\} \\ &\leq \frac{\rho_{\max}^2 K^2 L^2}{\varepsilon_0^2}\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}, \end{aligned} \tag{2.93}$$

which implies

$$\left|\tilde{\delta}_2^{(\varphi(L))}(L)\right| \leq \frac{\rho_{\max} KL}{\varepsilon_0}\left[\mathbf{\Omega}_{\mathbf{\Sigma}_L}^{(\varphi(L))}\right]^{\frac{1}{2}}. \tag{2.94}$$

Using *Assumption 6*, (2.92), (2.94) and the same approach used in the proof of *Theorem 1*, we can conclude the proof. ■

Theorem 6 *Given a sequence of covariance matrices* $\left\{\mathbf{\Sigma}_{(K,L)}\right\}$ *in (2.63) which satisfy* Assumption 5, 6 and 7, *if* $\{\mathbf{\Sigma}_L\}$ *satisfy* Assumption 3, *then* $\{\mathbf{\Sigma}_L\} \in \mathcal{U}^*(\varphi_0)$ *for* $\varphi_0(L) = \left\lceil \frac{1+\kappa}{\ln \gamma^{-1}} \ln L \right\rceil$, *where* κ *is an arbitrary small positive constant and* $\gamma = \left(\frac{1-\varepsilon_0}{1+\varepsilon_0}\right)^{\frac{2}{m}}$. *Similarly, if* $\{\mathbf{\Sigma}_L\}$ *satisfy* Assumption 4, $\{\mathbf{\Sigma}_L\} \in \mathcal{U}^*(\varphi_0)$ *for* $\varphi_0(L) = \left\lceil L^{\frac{2+\alpha}{2m-1}} \right\rceil$, *where* α *is an arbitrary small positive constant. As a result, for both cases, the ADR of the truncated detector with* $\varphi_0(L)$ *relative to the optimal detector converges to unity. Furthermore, any other* $\varphi(L)$ *such that* $\lim_{L\to\infty} \frac{\varphi(L)}{\varphi_0(L)} \geq 1$, *will also provide unity ADR.*

It is seen that the limit stated by $\mathcal{U}^*(\varphi_0)$ in (2.88) is the same as that described by $\mathcal{U}(\varphi_0)$ in (2.34). Thus, the proof of *Theorem 6* is very similar to *Theorem 2 and 3* and hence is omitted here.

It is worth mentioning that we adopt the separable covariance model and we let the size of the sensor network L increase to infinity while the size of temporal observations at each sensor K is fixed and finite. Hence, intuitively, the covariance matrix of total set of observations $\mathbf{\Sigma}_{(K,L)}$ should be dominated by the purely spatial covariance matrix $\mathbf{\Sigma}_L$. Thus, as the size of the sensor network L increases to infinity, it is seen that the limit in $\mathcal{U}^*(\varphi_0)$ only describes the requirement of the purely spatial covariance, which is the same as that described by $\mathcal{U}(\varphi_0)$ in (2.34). As a result, if the purely spatial covariance matrix $\mathbf{\Sigma}_L$ is contained in the two general classes of system and process models discussed in Section IV, the sufficient conditions described in *Theorem 5* can be easily satisfied, provided an adequate truncation rule $\varphi(L)$ is chosen. We can show that large classes of nonseparable covariance models also satisfy the conditions of *Theorem 5*, for example the class of all models where the covariance on the left-hand side of (2.82) can be expressed as a weighted sum of separable terms like those on the right-hand side of (2.82) provided the other assumptions in *Theorem 6* hold. However, the largest class of nonseperable covariance models satisfying the conditions of *Theorem 5* is still an open problem.

2.8 Summary

In this chapter, we study the large observation size performance of a truncated detector for a canonical multivariate Gaussian hypothesis testing problem. The benefits gained by utilizing the truncated detector instead of the optimal detector can be summarized from two viewpoints. If the components of $\mathbf{x}_L$ are time sampled observations, the truncated detector can reduce the storage and multiplications needed when compared to the optimal detector. If the components of $\mathbf{x}_L$ are obtained from distributed sensors, the truncated detector not only

reduces the communication energy requirement, it allows efficient implementation by adopting a consensus algorithm. Motivated by these benefits, we have investigated the performance of the truncated detector in terms of deflection, and derived sufficient conditions for a truncation rule and a sequence of tests which lead to no loss in ADR of the truncated detector relative to the optimal detector. The sufficient conditions provided depend on how the hypothesis testing problem scales with L. When either $\xi(L)$ or $\eta(L)$ grow with L, indicating the difficulty of the hypothesis testing problem decreases when L increases, we find a more aggressive truncation rule can be tolerated. Further, the amount of truncation which can be tolerated is different depending on which function, $\xi(L)$ or $\eta(L)$, grows with L. Moreover, we employ several well-accepted and popular classes of system and process models as examples to show that the sufficient conditions are not overly restrictive. For all the examples considered, we find truncation rules which increase slowly with L, implying significant savings, even for the least favorable case where the difficulty of the hypothesis testing problem doesn't decrease as L increases. In all the cases considered, numerical results imply that not only do the deflections of the truncated and the optimal detectors converge for large L for our asymptotically optimal truncation rules, but the probability of detections also converge for fixed false alarm probabilities.

While we have focused on asymptotic analysis in this chapter, some comments on finite L cases are in order for completeness. Finite L analysis employing (2.10) and (2.11) can be useful to evaluate deflection loss. Simplifications like $\left|\frac{\delta_2^{(k)}(L)}{\psi_2(L)}\right| \ll 1$ can be employed when justified. Even for finite L, unity deflection can be obtained in some extreme cases. The following result provides one example.

Theorem 7 *Assume the elements of $\mathbf{\Sigma}_L$ are bounded and L is finite. As $\|\boldsymbol{\mu}_L\|$ increases without bound, the deflection ratio $\Lambda(T_L^{\mathrm{tr}}, T_L^{\mathrm{opt}})$ of a truncated detector relative to the optimal*

detector converges to unity regardless of the truncation length k.

The proof is obvious. Even with bounded $\|\boldsymbol{\mu}_L\|$, we have found models satisfying *Assumption 1* that achieve unity deflection for finite k and L. However, the full class of models of this type seems to be one of a number of open problems in this area.

There are some other important problems still open. Necessary conditions for asymptotic optimality are still unknown. In this chapter, we have derived sufficient conditions which lead no loss in ADR of the truncated detector relative to the optimal detector. However, the deflection is a sub-optimal metric and the error probability is a better measure of detection performance. Our numerical results imply that the sufficient conditions for the equivalence in terms of deflection lead to an equivalence in terms of the limiting error probability in the cases we studied. A theoretic proof will be pursued in future work for the largest possible class of truncation rules and detection problems. It would be of great interest to consider tests employing constrained communications with neighbors for other hypothesis testing problems with dependent observations.

Chapter 3

Asymptotically Optimum Distributed Estimation in the Presence of Attacks

3.1 Introduction

Sensor networks employed for parameter estimation have been extremely successful in applications ranging from inexpensive commercial systems to complex military and homeland defense surveillance systems and have seen even greater interest in recent years [49]. Recent technological advances in coding, digital wireless communications technology and digital electronics have lead to the dominance of digital communications using quantized data in such systems. Hence, a great deal of attention has focused on parameter estimation using quantized data [50–56].

Under the assumption that several subsets of sensors are forced to send data corrupted

by a set of adversaries, we consider the problem of estimating a deterministic mean-shift parameter in the presence of zero-mean noise by using quantized data for a large number of observations in this chapter. Under the control of adversaries, the malicious sensors, which are called Byzantine sensors in recent literature [57–62], attempt to confuse the fusion center (FC) by sending modified quantized observations. The FC attempts to identify the different sets of malicious sensors and mitigate the impact on the estimation performance caused by the adversaries. This kind of distributed estimation problem under attacks is well motivated by the vulnerability of sensor networks in practice. For example, large scale sensor networks are typically comprised of inexpensive nodes with low computing capacity and limited battery power. Hence, highly complicated encryption algorithms cannot be implemented at each sensor which provides the adversaries an opportunity to modify the data to undermine the estimation performance of the sensor network. On the other hand, adversaries can also capture some limited set of sensors and force them to send altered data.

We assume that, without attack, all sensors make independent and identically distributed (i.i.d.) observations of the unknown deterministic parameter corrupted by zero-mean noise with known distribution. At each sensor, the time samples are converted to one-bit data and then transmitted to the FC due to the stringent energy and bandwidth limitations. However, the quantized outputs of some vulnerable subsets of sensors are hijacked by adversaries. We grant the adversaries, assumed to employ a finite number of different attacks in total at any given time, the largest power to manipulate their compromised sensors under some constraints concerning the information they have about the estimation system and the environment as well as to their access to the attacked sensors. In particular, the adversaries can only get physical access to the quantized data but they do not have access to the input of the quantizer (so called man-in-the-middle attacks) and they do not have information about what

computations the fusion system is using. Among other things they do not have information about the parameter to be estimated and the quantization thresholds. Thus, the adversaries do not really understand what the bits at each sensor actually mean since the fusion center and the sensors can agree on any interpretation they like. Thus, we assume that each adversary can modify the quantized data to bring about an arbitrary probability mass function (pmf) at the output of each sensor the adversary controls. Further, during the time window over which the estimation is performed, the statistical descriptions of the modification strategies of the adversaries are described by probability transition matrices unknown to the FC.

The communication channel between the FC and each sensor is assumed ideal, and hence the FC is able to accurately receive what was transmitted from both the unattacked and Byzantine sensors. The FC is assumed unaware of which subsets of sensors have been tampered with by adversaries. In order to avoid ambiguity between a set of attacked and a set of unattacked sensors, the set of unattacked sensors is assumed to occupy a larger percentage of the total number of sensors than any set of identically attacked sensors. The FC will attempt to identify not only the set of unattacked sensors but also each set of malicious sensors employing a distinct attack. After the identification and categorization of the sensors, an asymptotically optimum[1] estimate that involves all unknown parameters, including the parameter to be estimated and all the attack parameters, will be considered at the FC. The appropriate performance metric for the framework is the Cramer-Rao Bound (CRB) which provides a lower but asymptotically achievable bound on mean squared error (MSE). We will make use of CRB analysis to benchmark the estimation performance of unbiased estimators of these parameters.

In some parts of this chapter, we are primarily interested in the distributed estimation

[1]Maximum likelihood is one such estimate that achieves the asymptotically optimum performance.

problem for sufficiently large scale sensor networks so we pursue analytical characterization of the asymptotic estimation performance as the number of sensors grows to infinity. Attention is restricted to cases where each distinct attack occupies a nonzero percentage of all sensors in the limit to avoid consideration of attacks on sets of sensors with zero measure in the limit. Such would be the case if, for example, only a single sensor was attacked in the limit.

3.1.1 Summary of Results and Main Contributions

For the distributed estimation problem in the presence of attacks, we first study the ability of the FC to identify the attacked sensors and categorize them into different groups corresponding to distinctly different types of attacks. We only assume that the set of u-nattacked sensors is a larger percentage of all sensors than any set of identically attacked sensors to avoid ambiguity between a set of attacked and a set of unattacked sensors. It can be shown that increasing the number K of time samples at each sensor and enlarging the size N of the sensor network can both improve the performance of the identification and categorization approach, but to different extents. To be specific, the FC is able to determine the number P of attacks in the sensor network and achieve the correct categorization as $K \to \infty$, while as $N \to \infty$ with finite but sufficiently large K, it can be shown that the FC can also ascertain P and obtain an approximate categorization with a very small percentage of sensors that are misclassified, so small that this misclassification impacts performance in a manner which can be tolerated. In this sense, with sufficiently many time samples at each sensor or a sufficiently large size sensor network, the FC is able to determine the number of attacks in the sensor network and categorize the sensors into different groups according to distinct types of attacks perfectly or with negligibly small misclassification. Based on this fact, we can assume that the sensors have been well identified in the next part of the chapter.

Next, we consider estimation of the desired parameter. There are two approaches: (1) ignore the data at the attacked sensors. (2) Use the data at the attacked sensors. We can easily take approach (1) without estimating any parameters describing the attacks. However, to attempt to take approach (2), and potentially do better than approach (1), we will investigate the performance of the joint estimation of the desired parameter and the unknown attack parameters. It is shown that the Fisher Information Matrix (FIM) for jointly estimating these parameters is singular when we apply exactly the same quantization approach used for the unattacked system. Thus, it is not possible to jointly estimate the desired and attack parameters efficiently with an estimation error that decreases with KN by employing the same quantization approach used for the unattacked system.

In order to overcome the FIM singularity, a time-variant quantization approach has been proposed. The basic idea is that each sensor divides its observation time interval into several time slots, and in each time slot, all sensors use an identical threshold to quantize the time samples. However, the thresholds utilized in different time slots are distinct. We can show that as long as at least two different thresholds have been employed, the FIM of the time-variant quantization approach is nonsingular. Further, this FIM has been used to provide necessary and sufficient conditions under which taking advantage of the attacked sensors in the proposed fashion will provide better estimation performance when compared to approaches where the attacked sensors are ignored. These results are obtained by also employing the FIM for the case where the attacked sensors are ignored and the comparisons were made assuming both approaches use the same set of distinct thresholds over the same different time slots to provide a fair comparison. In the numerical results, we show that for some cases, significant improvement in the estimation performance can be obtained by employing the proposed approach.

3.1.2 Related Work

The distributed detection and estimation problem in the presence of Byzantine attacks has seen great interest in recent years, see [57–62] and references therein. The closest work we have seen to that proposed in this chapter appears in [62]. Still, there are major differences. First and foremost, the model in [62] is very different from that in our work. The model in [62] assumes that each sensor is attacked with a certain probability, which is known to the FC, and the probability that any given sensor is attacked is the same. We consider an observation time interval and assume that only some subsets of sensors can be attacked over that time due to the limited resources available. Further, the statistical description of the attack strategy in [62] is also assumed known to the FC, while we do not make this assumption and consider the joint estimation of the parameter to be estimated and the attack parameters.

An encrypted sensor network scheme is investigated in [63,64], where the stochastic encipher flips the binary sensor outputs with a certain probability to disguise the sensor outputs with the goal of confusing an enemy fusion center (EFC) while preserving the detection or estimation performance at the authorized fusion center (AFC). From the EFC perspective, the encryption process can be treated as a malicious attack, since the EFC does not know the probability a sensor output is flipped in the encryption process while the AFC knows this probability. The EFC in [63], which is the counterpart of the FC in our scheme, implements what we call the naive MLE. Hence, our proposed work is quite different. Moreover, since the focus is on a different problem, encryption, the overlap with our investigations is not significant.

The effect of attacks can also be considered as coming from sensors which transmit their original data through a bit-flipping channel to the FC [65]. Some research has investigated

distributed detection and estimation performance when the sensor network suffers from this kind of non-ideal channel, see [55, 66] and the references therein. This work assumes that the FC is aware of the set of sensors which are subject to the non-ideal channels and that the error probability of each of these binary channels is known to the FC (channel-aware). Together, these two assumptions are equivalent to assuming that the sets of attacked sensors and the attack strategies are known to the FC in our scheme. We do not make this assumption (we are channel-unaware). Instead, we investigate the identification and categorization of the attacked sensors. Further, in the previous literature on channel-aware distributed estimation, only the standard fixed-quantizer approach (fixed threshold for all time samples) is considered. However, we show that if the standard fixed-quantizer approach is used at each sensor, without information on the flipping probabilities of the binary channels, then the corresponding FIM is singular which implies that one cannot jointly estimate the desired parameter and the attack parameters with an accuracy that grows with the number of observations.

3.1.3 Notation and Organization

Throughout this chapter, bold upper case letters and bold lower case letters are used to denote matrices and column vectors respectively. The symbol $\mathbf{I}$ signifies the identity matrix, while $\mathbb{1}(\cdot)$ stands for the indicator function. Let $[\mathbf{A}]_{i,j}$ denote the element in the i-th row and j-th column of the matrix $\mathbf{A}$. $\mathbf{A} \succ 0$ and $\mathbf{A} \succeq 0$ imply that the matrix is positive definite and positive semidefinite respectively. For any set $\mathcal{S}$ of sensors, $|\mathcal{S}|$ denotes the number of sensors in the set $\mathcal{S}$, and $\mathcal{P}_{\mathcal{S}}$ represents the percentage of all sensors occupied by $\mathcal{S}$. Let $\mathcal{A}^{\mathrm{C}}$ denote the complement of the set $\mathcal{A}$. To avoid cumbersome sub-matrix and sub-vector expressions in this chapter, we introduce the following notation. Let $\mathbf{A}\left(\{i_1, i_2, ..., i_L\}, \{j_1, j_2, ..., j_M\}\right)$ denote

the sub-matrix which consists of the elements located in the $\{i_l\}_{l=1}^{L}$-th rows and $\{j_m\}_{m=1}^{M}$-th columns of matrix $\mathbf{A}$, and $\mathbf{A}\left(\{i_1, i_2, ..., i_L\}, \{:\}\right)$ represents the sub-matrix which consists of the elements located in the $\{i_l\}_{l=1}^{L}$-th rows of matrix $\mathbf{A}$. The notation $\mathbf{v}\left(i_1, i_2, ..., i_L\right)$ stands for the sub-vector which only contains the $\{i_l\}_{l=1}^{L}$-th elements of $\mathbf{v}$, and the i-th element of the vector $\mathbf{v}$ is denoted by v_i. Finally, the expectation, determinant and rank operators are denoted $\mathbb{E}\left(\cdot\right)$, $\det\left(\cdot\right)$ and $\text{rank}\left(\cdot\right)$ respectively.

The remainder of the chapter is organized as follows. The signal and adversary model is introduced in Section 3.2. The ability of the FC to identify and categorize attacked sensors is studied in Section 3.3. Section 3.4 analyzes the corresponding FIM. In Section 3.5, a time-variant quantization approach is proposed. Necessary and sufficient conditions are developed under which using the attacked sensors with this time-variant quantization approach will lead to better estimation performance. In Section 3.6, several numerical results are provided to illustrate our theoretical analysis. Finally, Section 3.7 provides our conclusions.

3.2 Signal and Adversary Models

3.2.1 Signal Model and Naive Maximum Likelihood Estimator

We consider a set of N distributed sensors, each making K observations of a deterministic scalar parameter θ corrupted by additive noise. At the j-th sensor, the observation at the k-th time instant is described by

$$x_{jk} = \theta + n_{jk},\ \forall j = 1, 2, ..., N,\ \forall k = 1, 2, ..., K, \tag{3.1}$$

where n_{jk} denotes the additive noise sample with common zero-mean probability density function (pdf) $f(n_{jk})$ and $\{n_{jk}\}$ is an independent and identically distributed sequence.

Due to the stringent energy and bandwidth limitations in realistic sensor networks, each sensor is restricted to transmit a single bit per observation x_{jk} to the fusion center (FC). In this chapter, to simplify the problem in terms of both implementation and analysis, all x_{jk} are quantized to u_{jk} by using threshold quantizers of the same design

$$u_{jk} = \mathbb{1}\left\{x_{jk} \in (\tau, \infty)\right\}. \tag{3.2}$$

We assume that the quantizer design and the threshold τ is known to the FC. The common probability mass function (pmf) at the output of the quantizer under no attack is

$$\begin{cases} \Pr\left(u_{jk} = 0 \,|\theta\right) = F\left(\tau - \theta\right) \\ \Pr\left(u_{jk} = 1 \,|\theta\right) = 1 - F\left(\tau - \theta\right) \end{cases}, \tag{3.3}$$

where $F\left(x\right) \triangleq \int_{-\infty}^{x} f\left(t\right) dt$ denotes the cumulative distribution function (cdf) corresponding to the pdf $f(x)$. We assume that $F^{-1}\left(x\right)$ is differentiable on the open interval $(0, 1)$. After collecting the binary observations $\{u_{jk}\}$ from all sensors, by employing the invariance of an ML estimate, the naive Maximum Likelihood Estimate (NMLE), the MLE formulated under the assumption of no attack, of the parameter θ can be expressed as [50, 51]

$$\hat{\theta}_{\mathrm{NML}} = \tau - F^{-1}\left(1 - \frac{1}{KN}\sum_{j=1}^{N}\sum_{k=1}^{K} u_{jk}\right), \tag{3.4}$$

which, without the presence of an adversary, can be expected to provide asymptotically unbiased and efficient estimation of θ.

3.2.2 Adversary Model

The adversaries aim at tampering with the quantized observations $\{u_{jk}\}$, hoping to cause the FC to reach an inaccurate estimate in terms of large bias and variance. Consider a set of P distinct types of malicious attacks, where each attack will sometimes modify some sensors' observations. Let $\mathcal{A}_p$ denote the set of sensors subjected to the p-th attack. Let $\tilde{u}_{jk}$ represent the after-attack quantized observation which is a modified version of u_{jk}. The statistical description of the p-th attack can be described by a probability transition matrix $\mathbf{\Psi}_p$,

$$\mathbf{\Psi}_p \triangleq \begin{bmatrix} \psi_{p,0} & 1-\psi_{p,1} \\ 1-\psi_{p,0} & \psi_{p,1} \end{bmatrix}, \tag{3.5}$$

where $\psi_{p,0} \triangleq \Pr(\tilde{u}_{jk}=0\,|u_{jk}=0)$ and $\psi_{p,1} \triangleq \Pr(\tilde{u}_{jk}=1\,|u_{jk}=1)$ are the modification probabilities. Due to the p-th attack, the after-attack pmf of the observations can be related to the before-attack pmf using

$$\begin{bmatrix} 1-\tilde{p}(\mathbf{\Psi}_p,\theta) \\ \tilde{p}(\mathbf{\Psi}_p,\theta) \end{bmatrix} \triangleq \begin{bmatrix} \Pr(\tilde{u}_{jk}=0\,|\theta) \\ \Pr(\tilde{u}_{jk}=1\,|\theta) \end{bmatrix} = \mathbf{\Psi}_p \begin{bmatrix} \Pr(u_{jk}=0\,|\theta) \\ \Pr(u_{jk}=1\,|\theta) \end{bmatrix} \tag{3.6}$$

Substituting (3.3) into (3.6), we can obtain

$$\begin{aligned} \tilde{p}(\mathbf{\Psi}_p,\theta) &= (1-\psi_{p,0})\Pr(u_{jk}=0\,|\theta) + \psi_{p,1}\Pr(u_{jk}=1\,|\theta) \\ &= (1-\psi_{p,0}-\psi_{p,1})\,F(\tau-\theta) + \psi_{p,1}. \end{aligned} \tag{3.7}$$

For the sake of expressing the after-attack pmfs of observations in a uniform form for both attacked and unattacked sensors, the set $\mathcal{A}_0$ of unattacked sensors are considered "under

attack" with probability transition matrix $\mathbf{\Psi}_0 = \mathbf{I}$, where $\mathbf{\Psi}_0$ is known to the FC.

From a practical point of view, the following assumption is made through this chapter.

Assumption 8

1. *Over the estimation time interval and for all p, the p-th attack is statistically described as in (3.6) for all the sensors in the set $\mathcal{A}_p$. The set $\mathcal{A}_p$ and $\mathbf{\Psi}_p$ are both unknown to the FC (except $\mathbf{\Psi}_0$), and for sufficiently large N the number of sensors in $\mathcal{A}_p$, $|\mathcal{A}_p|$, is a fixed percentage $\mathcal{P}_p$ of the total number N of sensors in the sensor network. Such an assumption is required so that as $N \to \infty$ the effect of an attack will not shrink to zero ($\mathcal{A}_p$ becoming a set of measure zero). Moreover, we assume that the group of unattacked sensors is the largest group and $\mathcal{P}_0 > \mathcal{P}_p + \Delta_0$ for all $p \geq 1$ where Δ_0 is a positive constant. Further the sets $\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_P$ are disjoint so that*

$$\mathcal{A}_p \cap \mathcal{A}_{p'} = \emptyset \quad \text{if} \quad p \neq p'. \tag{3.8}$$

2. Significant Attacks. *In order to give rise to sufficient impact on the statistical characterization of the outputs from attacked sensors, every attacker is required to guarantee a minimum distortion d_{impact} on $\tilde{p}\left(\mathbf{\Psi}_0, \theta\right)$ and tamper with at least Δ percent of sensors so that the following relations should be satisfied*

$$\left|\tilde{p}\left(\mathbf{\Psi}_p, \theta\right) - \tilde{p}\left(\mathbf{\Psi}_0, \theta\right)\right| \geq d_{impact}, \ \forall p = 1, 2, ..., P, \tag{3.9}$$

$$\mathcal{P}_p \geq \Delta > 0, \quad \forall p = 1, 2, ..., P. \tag{3.10}$$

3. Various Attacks. *The changes caused by two distinct types of attacks are considerably*

different, otherwise these two types of attacks can be treated as identical. To this end, we assume that

$$|\tilde{p}\left(\mathbf{\Psi}_l, \theta\right) - \tilde{p}\left(\mathbf{\Psi}_m, \theta\right)| \geq d_{diff},\ \forall l \neq m. \tag{3.11}$$

4. Non-trivial Attacks. *If the FC perceives some sensor produces a constant value of 0 or 1, then the FC can easily recognize the sensor is under attack. For this reason, in order to reduce the probability of being detected, we assume that the adversaries ensure*

$$\tilde{p}\left(\mathbf{\Psi}_p, \theta\right) \neq 0 \text{ or } 1, \quad \forall p. \tag{3.12}$$

It is worth mentioning that the adversary model assumed in (3.6) can change the after-attack pmf to have any desired valid values satisfying (4.59), (3.11), and (3.12) through proper choice of the two attack parameters $\psi_{p,0}$ and $\psi_{p,1}$. In this sense, it is a fairly general adversary model.

3.3 Identification and Categorization of Attacked Sensors

In order to mitigate the effect caused by the adversaries, we seek to identify the attacked sensors and categorize them into different groups according to the different attacks. In this section, we investigate our ability to undertake these two tasks.

Let $\mathcal{S}_T \triangleq \cup_{p=0}^{P} \mathcal{A}_p$ denote the set of all sensors in the sensor network, and let $\mathfrak{C}_0$ and $\mathfrak{C}_1(\kappa)$ define two collections of sets of the sensors, whose elements are the subsets of $\mathcal{S}_T$

$$\mathfrak{C}_0 \triangleq \{\mathcal{S} \subset \mathcal{S}_T \mid \exists\ p \text{ s.t. } \mathcal{S} \subset \mathcal{A}_p\}, \tag{3.13}$$

$$\mathfrak{C}_1(\kappa) \triangleq \big\{\mathcal{S} \subset \mathcal{S}_T \,|\, \exists\, \mathcal{S}_1,\, \mathcal{S}_2 \subset \mathcal{S} \text{ and } \mathcal{S}_1,\, \mathcal{S}_2 \in \mathfrak{C}_0,$$
$$\text{s.t.} \;\; \mathcal{P}_{\mathcal{S}_1} \geq \kappa,\; \mathcal{P}_{\mathcal{S}_2} \geq \kappa, \text{ and } \mathcal{S}_1 \cup \mathcal{S}_2 \notin \mathfrak{C}_0\big\}, \tag{3.14}$$

where $\kappa \in (0,\, \Delta)$ is a constant. $\mathfrak{C}_0$ is the collection of all homogeneous sets of sensors which are also referred to as statistically identical sets, while $\mathfrak{C}_1(\kappa)$ is the collection of all non-homogeneous sets of sensors made up from combining homogeneous parts which each occupy more than a given percentage κ of all sensors. Every subset of $\mathcal{A}_p$ for any p is an example of an element in $\mathfrak{C}_0$, and one example of an element in $\mathfrak{C}_1(\kappa)$ is $\mathcal{A}_l \cup \mathcal{A}_m$ provided $l \neq m$.

Lemma 9 *Consider a subset $\mathcal{J}$ of the sensors which includes a fixed percentage $\mathcal{P}_{\mathcal{J}} \triangleq |\mathcal{J}|/N \geq \Delta$ of all sensors in a sensor network such that $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa)$ for some κ. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ denote two disjoint subsets of $\mathcal{J}$ with $\mathcal{P}_{\mathcal{S}_1} = \mathcal{P}_{\mathcal{S}_2} = \kappa$. Let $d_{\min} \triangleq \min\{d_{\text{impact}}, d_{\text{diff}}\}$, and hence $0 < d_{\min} < 1$ due to (4.59) and (3.11). Define*

$$\lambda \triangleq \frac{d_{\min}}{2 \sup\limits_{\nu} f\,(\nu)}. \tag{3.15}$$

Now, consider the hypothesis testing problem

$$\begin{cases} \mathcal{H}_0 : \mathcal{J} \in \mathfrak{C}_0 \\ \mathcal{H}_1 : \mathcal{J} \in \mathfrak{C}_1(\kappa) \end{cases} \tag{3.16}$$

and the decision rule

$$\varpi\,(\tilde{\mathbf{u}}_{\mathcal{J}}) = \begin{cases} 0, & T\,(\tilde{\mathbf{u}}_{\mathcal{J}}, \kappa) \leq \lambda; \\ 1, & T\,(\tilde{\mathbf{u}}_{\mathcal{J}}, \kappa) > \lambda, \end{cases} \tag{3.17}$$

where

$$T\left(\tilde{\mathbf{u}}_{\mathcal{J}},\kappa\right)\triangleq\sup_{\left\{\mathcal{S}_1,\mathcal{S}_2:\mathcal{P}_{\mathcal{S}_1},\mathcal{P}_{\mathcal{S}_2}=\kappa\right\}}\left\{\left|\hat{\theta}_{\mathrm{NML}}^{\mathcal{S}_1}-\hat{\theta}_{\mathrm{NML}}^{\mathcal{S}_2}\right|\right\}\tag{3.18}$$

and $\hat{\theta}_{\mathrm{NML}}^{\mathcal{S}}$ *denotes the naive ML estimate from (3.4) based on the observations from the subset* $\mathcal{S}$*. Let* π_0 *and* π_1 *denote the prior probabilities of* $\mathcal{H}_0$ *and* $\mathcal{H}_1$ *respectively. Under* Assumption 8*, if* K *is larger than some constant* K^**, then*

$$\begin{aligned}P_{error}&\triangleq\pi_0\Pr\left(\textit{Declare }\ \mathcal{H}_1\middle|\,\mathcal{H}_0\right)+\pi_1\Pr\left(\textit{Declare }\ \mathcal{H}_0\middle|\,\mathcal{H}_1\right)\\&\leq Ce^{\kappa\gamma KN},\end{aligned}\tag{3.19}$$

where γ *is a negative constant, and* C *is a positive constant.*

Proof: Consider a subset $\mathcal{S}$ of $\mathcal{J}$ with $\mathcal{P}_{\mathcal{S}}$ percent of all sensors in the sensor network, which is only tampered with by the l-th attack, where $l\in\{0,1,...,P\}$. The naive ML estimate $\hat{\theta}_{\mathrm{NML}}^{\mathcal{S}}$ based on the observations from the sensors in $\mathcal{S}$ can be expressed as

$$\hat{\theta}_{\mathrm{NML}}^{\mathcal{S}}=\tau-F^{-1}\left(\xi_{\mathcal{S}}\right),\tag{3.20}$$

where $\xi_{\mathcal{S}}$ is defined as

$$\xi_{\mathcal{S}}\triangleq\frac{1}{K\left|\mathcal{S}\right|}\sum_{j\in\mathcal{S}}\sum_{k=1}^{K}(1-\tilde{u}_{jk}),\tag{3.21}$$

and $\tilde{u}_{jk}$, for $j\in\mathcal{S}$, follows a Bernoulli distribution with probability $\tilde{p}\left(\mathbf{\Psi}_l,\theta\right)$.

Under hypothesis $\mathcal{H}_1$, there are at least two disjoint statistically distinct groups of sensors, say $\mathcal{S}_1^*$ and $\mathcal{S}_2^*$ with $\mathcal{P}_{\mathcal{S}_1^*}=\mathcal{P}_{\mathcal{S}_2^*}=\kappa$. Without loss of generality, assume $\mathcal{S}_1^*$ and $\mathcal{S}_2^*$ are attacked by the l-th and m-th attacks respectively, and $\tilde{p}\left(\mathbf{\Psi}_l,\theta\right)>\tilde{p}\left(\mathbf{\Psi}_m,\theta\right)$, where $l\neq m$

and $l, m \in \{0, 1, ..., P\}$. Define $T_1^{(\mathcal{S}_1^*, \mathcal{S}_2^*)} \triangleq \hat{\theta}_{\text{NML}}^{\mathcal{S}_1^*} - \hat{\theta}_{\text{NML}}^{\mathcal{S}_2^*}$. By employing (3.20), we can obtain that

$$\begin{aligned}\left|T_1^{(\mathcal{S}_1^*, \mathcal{S}_2^*)}\right| &= \left|\hat{\theta}_{\text{NML}}^{\mathcal{S}_1^*} - \hat{\theta}_{\text{NML}}^{\mathcal{S}_2^*}\right| = \left|F^{-1}\left(\xi_{\mathcal{S}_2^*}\right) - F^{-1}\left(\xi_{\mathcal{S}_1^*}\right)\right| \\ &\geq \left[\inf_{\nu}\left|\frac{\partial F^{-1}(\nu)}{\partial \nu}\right|\right]\left|\xi_{\mathcal{S}_2^*} - \xi_{\mathcal{S}_1^*}\right| \\ &\geq \frac{1}{\sup\limits_{\nu} f(\nu)}\left|\xi_{\mathcal{S}_2^*} - \xi_{\mathcal{S}_1^*}\right|, \end{aligned} \tag{3.22}$$

and therefore, (3.15), (3.21), and (3.22) yield an upper bound on the error probability under hypothesis $\mathcal{H}_1$ that

$$\begin{aligned}&\Pr\left(\text{Declare } \mathcal{H}_0 \middle| \mathcal{H}_1\right) = \Pr\left(T\left(\tilde{\mathbf{u}}_{\mathcal{J}}, \kappa\right) \leq \lambda \middle| \mathcal{H}_1\right) \\ &\leq \Pr\left(\left|T_1^{(\mathcal{S}_1^*, \mathcal{S}_2^*)}\right| \leq \lambda \middle| \mathcal{H}_1\right) \\ &\leq \Pr\left(\left|\xi_{\mathcal{S}_2^*} - \xi_{\mathcal{S}_1^*}\right| \leq \lambda \sup_{\nu} f(\nu) \middle| \mathcal{H}_1\right) \\ &\leq \Pr\left(\xi_{\mathcal{S}_1^*} - \xi_{\mathcal{S}_2^*} \geq -\frac{1}{2} d_{\min} \middle| \mathcal{H}_1\right) \\ &= \Pr\left(\sum_{i=1}^{\kappa K N} X_i \geq -\frac{1}{2} d_{\min} \kappa K N \middle| \mathcal{H}_1\right) \end{aligned} \tag{3.23}$$

where $\{X_i\}$ is a sequence of i.i.d. random variables with distribution

$$p_{X_1} \triangleq \Pr\left(X_i = 1 \middle| \mathcal{H}_1\right) = \left[1 - \tilde{p}\left(\mathbf{\Psi}_l, \theta\right)\right] \tilde{p}\left(\mathbf{\Psi}_m, \theta\right), \tag{3.24}$$

$$p_{\bar{X}_1} \triangleq \Pr\left(X_i = -1 \middle| \mathcal{H}_1\right) = \left[1 - \tilde{p}\left(\mathbf{\Psi}_m, \theta\right)\right] \tilde{p}\left(\mathbf{\Psi}_l, \theta\right), \tag{3.25}$$

and

$$\begin{aligned} p_{X_0} &\triangleq \Pr\left(X_i = 0 \middle| \mathcal{H}_1\right) \\ &= \tilde{p}\left(\mathbf{\Psi}_l, \theta\right) \tilde{p}\left(\mathbf{\Psi}_m, \theta\right) + \left[1 - \tilde{p}\left(\mathbf{\Psi}_l, \theta\right)\right] \left[1 - \tilde{p}\left(\mathbf{\Psi}_m, \theta\right)\right]. \end{aligned} \tag{3.26}$$

Since $\mathbb{E}\left(X_i\right) = \tilde{p}\left(\mathbf{\Psi}_m, \theta\right) - \tilde{p}\left(\mathbf{\Psi}_l, \theta\right) \leq -d_{\min} < -\frac{1}{2} d_{\min}$, by the large deviations theory [67, 68], we know

$$\Pr\left(\sum_{i=1}^{\kappa KN} X_i \geq -\frac{1}{2}\kappa KN d_{\min} \middle| \mathcal{H}_1\right) \leq e^{\kappa \gamma_{d_{\min}}^{(l,m)} KN}, \tag{3.27}$$

where the rate function $\gamma_{d_{\min}}^{(l,m)}$ is defined as

$$\begin{aligned} &\gamma_{d_{\min}}^{(l,m)} \\ &\triangleq \lim_{\kappa KN \to \infty} \frac{1}{\kappa KN} \ln \Pr\left(\sum_{i=1}^{\kappa KN} X_i \geq -\frac{1}{2}\kappa KN d_{\min} \middle| \mathcal{H}_1\right) \\ &= \frac{1}{2} d_{\min} \eta^* + \ln \varphi_X(\eta^*) < 0, \end{aligned} \tag{3.28}$$

and

$$\varphi_X(\eta) \triangleq \mathbb{E}\left\{\exp\left(\eta X_i\right)\right\} = p_{X_0} + p_{X_1} e^{\eta} + p_{\bar{X}_1} e^{-\eta}. \tag{3.29}$$

Moreover, the quantity η^* in (3.28) is the positive solution of the equation

$$\frac{d}{d\eta} \varphi_X(\eta) = -\frac{1}{2} d_{\min} \varphi_X(\eta). \tag{3.30}$$

By employing (3.24)–(3.26) and (3.28)–(3.30), the rate function $\gamma_{d_{\min}}^{(l,m)}$ can be expressed as

$$\gamma_{d_{\min}}^{(l,m)} = \frac{1}{2} d_{\min} \ln b_X + \ln \left[p_{X_0} + p_{X_1} b_X + p_{\bar{X}_1} b_X^{-1} \right], \tag{3.31}$$

where b_X represents

$$b_X \triangleq \frac{-d_{\min} p_{X_0} + \sqrt{d_{\min}^2 p_{X_0}^2 + 4(4 - d_{\min}^2) p_{X_1} p_{\bar{X}_1}}}{2(2 + d_{\min}) p_{X_1}}. \tag{3.32}$$

Define

$$\gamma_1 \triangleq \max_{l,m:l \neq m} \gamma_{d_{\min}}^{(l,m)}, \tag{3.33}$$

then by noting (3.23) and (3.27), we can obtain

$$\Pr\left(\text{Declare } \mathcal{H}_0 | \, \mathcal{H}_1 \right) \leq e^{\kappa \gamma_1 K N}. \tag{3.34}$$

On the other hand, under hypothesis $\mathcal{H}_0$, all sensors in $\mathcal{J}$ are statistically identical to each other. Without loss of generality, we assume $\mathcal{J}$ are tampered with by the l-th attack, $l \in \{0, 1, ..., P\}$. Consider two disjoint subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathcal{J}$ with fixed percentages $\mathcal{P}_{\mathcal{S}_1} = \mathcal{P}_{\mathcal{S}_2} = \kappa$ respectively. Define

$$T_0^{(\mathcal{S}_1, \mathcal{S}_2)} \triangleq \hat{\theta}_{\text{NML}}^{\mathcal{S}_1} - \hat{\theta}_{\text{NML}}^{\mathcal{S}_2} = F^{-1}\left(\xi_{\mathcal{S}_2} \right) - F^{-1}\left(\xi_{\mathcal{S}_1} \right), \tag{3.35}$$

and therefore, the error probability under hypothesis $\mathcal{H}_0$ is bounded above as per

$$\begin{aligned}
&\Pr\left(\text{Declare } \mathcal{H}_1 \middle| \mathcal{H}_0\right) = \Pr\left(T\left(\tilde{\mathbf{u}}_{\mathcal{J}}, \kappa\right) > \lambda \middle| \mathcal{H}_0\right) \\
&= \Pr\left(\sup_{\left\{\mathcal{S}_1, \mathcal{S}_2 : \mathcal{P}_{\mathcal{S}_1}, \mathcal{P}_{\mathcal{S}_2} = \kappa\right\}} \left\{\left|T_0^{(\mathcal{S}_1, \mathcal{S}_2)}\right|\right\} > \lambda \middle| \mathcal{H}_0\right) \\
&\leq 2^{2|\mathcal{J}|} \sup_{\left\{\mathcal{S}_1, \mathcal{S}_2 : \mathcal{P}_{\mathcal{S}_1}, \mathcal{P}_{\mathcal{S}_2} = \kappa\right\}} \Pr\left(\left|T_0^{(\mathcal{S}_1, \mathcal{S}_2)}\right| > \lambda \middle| \mathcal{H}_0\right) \qquad (3.36) \\
&= 2^{2\mathcal{P}_{\mathcal{J}} N} \Pr\left(\left|T_0^{(\mathcal{S}_1, \mathcal{S}_2)}\right| > \lambda \middle| \mathcal{H}_0\right), \qquad (3.37)
\end{aligned}$$

where (3.37) is due to the fact that $\Pr\left(\left|T_0^{(\mathcal{S}_1, \mathcal{S}_2)}\right| > \lambda \middle| \mathcal{H}_0\right)$ is the same for every pair of disjoint $\mathcal{S}_1$ and $\mathcal{S}_2$ with $\mathcal{P}_{\mathcal{S}_1} = \mathcal{P}_{\mathcal{S}_2} = \kappa$, since all observations from $\mathcal{J}$ are independent and identically distributed under hypothesis $\mathcal{H}_0$.

Let E_1 and E_2 denote the events $\{\xi_{\mathcal{S}_1} \in [\varepsilon, 1-\varepsilon]\}$ and $\{\xi_{\mathcal{S}_2} \in [\varepsilon, 1-\varepsilon]\}$ respectively and $\mathsf{E} \triangleq \mathsf{E}_1 \cap \mathsf{E}_2$, where ε is a positive small constant such that

$$\begin{cases} 0 < \varepsilon < 1 - \max\limits_{p} \tilde{p}\left(\mathbf{\Psi}_p, \theta\right) \\ 0 < \varepsilon < \min\limits_{p} \tilde{p}\left(\mathbf{\Psi}_p, \theta\right) \end{cases}. \qquad (3.38)$$

Since it is assumed beforehand that $F^{-1}(x)$ is differentiable on the open interval $(0,1)$, we know $F^{-1}(x)$ is a Lipschitz continuous function with some Lipschitz constant $\mathcal{L}_F > 0$ over

the compact set $[\varepsilon, 1-\varepsilon]$. Hence, by noticing (3.21), we can obtain that

$$\begin{aligned}
&\Pr\left(\left|T_0^{(\mathcal{S}_1,\mathcal{S}_2)}\right| > \lambda \middle| \mathcal{H}_0\right) \\
&\leq \Pr\left(\left\{\left|T_0^{(\mathcal{S}_1,\mathcal{S}_2)}\right| > \lambda\right\} \cap \mathsf{E} \middle| \mathcal{H}_0\right) + \Pr\left(\mathsf{E}^{\mathrm{C}} \middle| \mathcal{H}_0\right) \\
&\leq \Pr\left(\left\{\mathcal{L}_F \left|\xi_{\mathcal{S}_2} - \xi_{\mathcal{S}_1}\right| > \lambda\right\} \cap \mathsf{E} \middle| \mathcal{H}_0\right) + 2\Pr\left(\mathsf{E}_1^{\mathrm{C}} \middle| \mathcal{H}_0\right) \\
&\leq \Pr\left(\mathcal{L}_F \left|\frac{1}{\kappa KN}\sum_{i=1}^{\kappa KN} Y_i\right| > \lambda \middle| \mathcal{H}_0\right) + 2\Pr\left(\mathsf{E}_1^{\mathrm{C}} \middle| \mathcal{H}_0\right) \\
&\leq 2\Pr\left(\sum_{i=1}^{\kappa KN} Y_i \geq \kappa KN a_Y \middle| \mathcal{H}_0\right) + 2\Pr\left(\mathsf{E}_1^{\mathrm{C}} \middle| \mathcal{H}_0\right), \qquad (3.39)
\end{aligned}$$

where

$$a_Y \triangleq \min\left\{\frac{1}{2}, \frac{\lambda}{\mathcal{L}_F}\right\} > 0, \qquad (3.40)$$

and under hypothesis $\mathcal{H}_0$, $\{Y_i\}$ is a sequence of i.i.d. random variables with distribution

$$\begin{aligned}
p_{Y_1} &\triangleq \Pr\left(Y_i = -1 \middle| \mathcal{H}_0\right) = \Pr\left(Y_i = 1 \middle| \mathcal{H}_0\right) \\
&= \tilde{p}\left(\boldsymbol{\Psi}_l, \theta\right)\left[1 - \tilde{p}\left(\boldsymbol{\Psi}_l, \theta\right)\right], \qquad (3.41)
\end{aligned}$$

and

$$\begin{aligned}
p_{Y_0} &\triangleq \Pr\left(Y_i = 0 \middle| \mathcal{H}_0\right) \\
&= \left[\tilde{p}\left(\boldsymbol{\Psi}_l, \theta\right)\right]^2 + \left[1 - \tilde{p}\left(\boldsymbol{\Psi}_l, \theta\right)\right]^2. \qquad (3.42)
\end{aligned}$$

Since $a_Y > \mathbb{E}(Y_i) = 0$, by applying a similar argument as in (3.27)–(3.32), we can obtain

$$\Pr\left(\sum_{i=1}^{\kappa KN} Y_i \geq \kappa KN a_Y \middle| \mathcal{H}_0\right) \leq e^{\kappa \gamma_{a_Y}^{(l)} KN}, \qquad (3.43)$$

where the rate function $\gamma_{a_Y}^{(l)}$ is given by

$$\gamma_{a_Y}^{(l)} = -a_Y \ln b_Y + \ln \left[p_{Y_0} + p_{Y_1} \left(b_Y + b_Y^{-1} \right) \right] < 0, \tag{3.44}$$

and b_Y represents

$$b_Y \triangleq \frac{a_Y p_{Y_0} + \sqrt{a_Y^2 p_{Y_0}^2 + 4(1-a_Y^2) p_{Y_1}^2}}{2(1-a_Y) p_{Y_1}}. \tag{3.45}$$

Similarly, by employing (3.21) and the large deviations theory, we can obtain

$$\begin{aligned}
\Pr\left(\mathsf{E}_1^{\mathrm{C}} \middle| \mathcal{H}_0 \right) &= \Pr\left(\{\xi_{\mathcal{S}_1} < \varepsilon\} \cup \{\xi_{\mathcal{S}_1} > 1-\varepsilon\} \middle| \mathcal{H}_0 \right) \\
&\leq \Pr\left(\xi_{\mathcal{S}_1} \leq \varepsilon \middle| \mathcal{H}_0 \right) + \Pr\left(\xi_{\mathcal{S}_1} \geq 1-\varepsilon \middle| \mathcal{H}_0 \right) \\
&= \Pr\left(\sum_{i=1}^{\kappa KN} \bar{Z}_i \geq \kappa KN (1-\varepsilon) \middle| \mathcal{H}_0 \right) \\
&\quad + \Pr\left(\sum_{i=1}^{\kappa KN} Z_i \geq \kappa KN (1-\varepsilon) \middle| \mathcal{H}_0 \right) \\
&\leq e^{\kappa \bar{\gamma}_\varepsilon^{(l)} KN} + e^{\kappa \gamma_\varepsilon^{(l)} KN},
\end{aligned} \tag{3.46}$$

where $\bar{Z}_i \triangleq 1 - Z_i$, and $\{Z_i\}$ is a sequence of i.i.d. random variables with Bernoulli distribution under hypothesis $\mathcal{H}_0$

$$\begin{cases} p_{Z_0} \triangleq \Pr\left(Z_i = 0 \middle| \mathcal{H}_0 \right) = \tilde{p}\left(\boldsymbol{\Psi}_l, \theta \right) \\ p_{Z_1} \triangleq \Pr\left(Z_i = 1 \middle| \mathcal{H}_0 \right) = 1 - \tilde{p}\left(\boldsymbol{\Psi}_l, \theta \right) \end{cases}. \tag{3.47}$$

The rate functions $\bar{\gamma}_\varepsilon^{(l)}$ and $\gamma_\varepsilon^{(l)}$ in (3.46) can be written as

$$\bar{\gamma}_\varepsilon^{(l)} = -(1-\varepsilon) \ln \frac{(1-\varepsilon)(1-p_{Z_0})}{\varepsilon p_{Z_0}} + \ln \frac{1-p_{Z_0}}{\varepsilon} < 0, \tag{3.48}$$

$$\gamma_{\varepsilon}^{(l)} = -(1-\varepsilon)\ln\frac{(1-\varepsilon)p_{Z_0}}{\varepsilon(1-p_{Z_0})} + \ln\frac{p_{Z_0}}{\varepsilon} < 0. \tag{3.49}$$

Taking into account (3.37), (3.39), (3.43), and (3.46), yields

$$\begin{aligned}
&\Pr\left(\text{Declare } \mathcal{H}_1 | \, \mathcal{H}_0\right) \\
&\leq 2^{1+2\mathcal{P}_{\mathcal{J}}N}\left(e^{\kappa\gamma_{a_Y}^{(l)}KN} + e^{\kappa\bar{\gamma}_{\varepsilon}^{(l)}KN} + e^{\kappa\gamma_{\varepsilon}^{(l)}KN}\right) \\
&\leq 6\exp\left[\left(\gamma^* + \frac{2\mathcal{P}_{\mathcal{J}}\ln 2}{\kappa K}\right)\kappa KN\right],
\end{aligned} \tag{3.50}$$

where

$$\gamma^* \triangleq \max\left\{\max_l \gamma_{a_Y}^{(l)}, \max_l \gamma_{\varepsilon}^{(l)}, \max_l \bar{\gamma}_{\varepsilon}^{(l)}\right\} < 0. \tag{3.51}$$

Thus, if K is large enough such that

$$K > K^* \triangleq -\frac{2\ln 2}{\kappa\gamma^*} \geq -\frac{2\mathcal{P}_{\mathcal{J}}\ln 2}{\kappa\gamma^*}, \tag{3.52}$$

then $\gamma_0 \triangleq \gamma^* + \frac{2\mathcal{P}_{\mathcal{J}}\ln 2}{\kappa K} < 0$, and hence an upper bound on the error probability under hypothesis $\mathcal{H}_0$ can be expressed as

$$\Pr\left(\text{Declare } \mathcal{H}_1 | \, \mathcal{H}_0\right) \leq 6e^{\kappa\gamma_0 KN}. \tag{3.53}$$

As a result, by (3.34) and (3.53), we conclude the proof by noting that

$$\begin{aligned}
P_{\text{error}} &= \pi_0 \Pr\left(\text{Declare } \mathcal{H}_1 | \, \mathcal{H}_0\right) + \pi_1 \Pr\left(\text{Declare } \mathcal{H}_0 | \, \mathcal{H}_1\right) \\
&\leq 6\pi_0 e^{\kappa\gamma_0 KN} + \pi_1 e^{\kappa\gamma_1 KN} \leq Ce^{\kappa\gamma KN},
\end{aligned} \tag{3.54}$$

where $\gamma \triangleq \max\{\gamma_0, \gamma_1\} < 0$ and $C \triangleq 2\max\{6\pi_0, \pi_1\}$ is a positive constant. ■

The rate functions merit some attention. Fig. 3.1 depicts the rate functions in (3.28), (3.44), (3.48) and (3.49) for different $\tilde{p}\,(\mathbf{\Psi}_l, \theta)$, where $a_Y = 1/4$, $d_{\min} = 10^{-4}$, $\tilde{p}\,(\mathbf{\Psi}_m, \theta) = 0.009$, and $\varepsilon = 10^{-3}$. It is seen that the rate functions are all smaller than 0 for every $\tilde{p}\,(\mathbf{\Psi}_l, \theta)$ and the dominant rate function can be different for different range of $\tilde{p}\,(\mathbf{\Psi}_l, \theta)$.

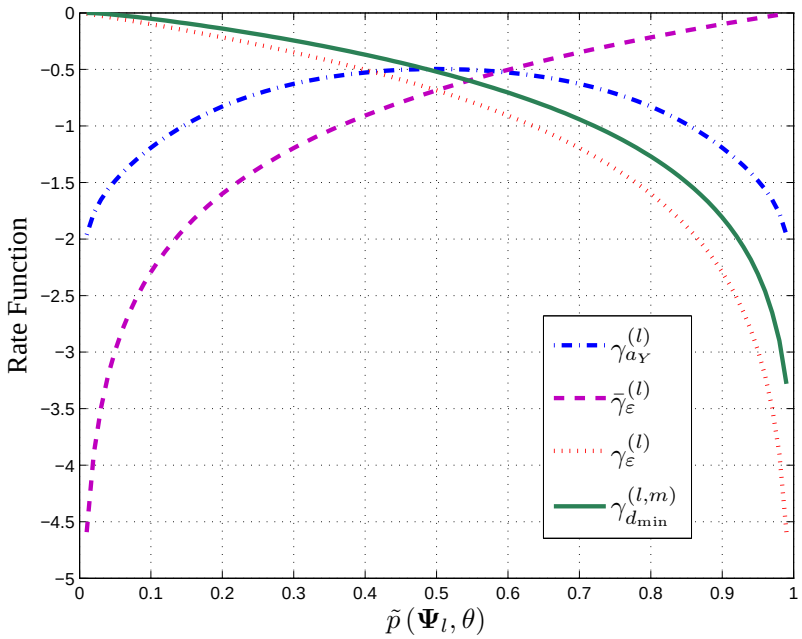

Figure 3.1: Rate functions versus $\tilde{p}\,(\mathbf{\Psi}_l, \theta)$.

Lemma 9 demonstrates that the error probability in (3.19) decreases to 0 as either $K \to \infty$ or $N \to \infty$, which implies that both enlarging the size of the sensor network and increasing the number of time observations at each sensor can improve the FC's ability to determine whether a given set $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa)$ for some κ with $\mathcal{P}_{\mathcal{J}} \geq \Delta$ is homogeneous or not. This fact motivates us to further investigate the identification and categorization performance when K or N increases.

3.3.1 The Number K of Time Samples at Each Sensor is Sufficiently Large

First, we study the identification and categorization performance we can obtain in the scenario that the number K of time samples is sufficiently large.

Theorem 8 *Take* Assumption 8 *as a given. For any N as $K \to \infty$, the FC can always identify from the observations, without further knowledge, a group of sensors which make up $\mathcal{P}_0$ percent of all sensors such that this group contains zero percent attacked sensors with probability* 1*. In this sense, one can always identify the unattacked sensors. Moreover, as $K \to \infty$, the FC is also able to identify the other P groups of sensors, which respectively make up $\{\mathcal{P}_p\}_{p=1}^{P}$ percent of all sensors, such that for $p = 1, 2, ..., P$, group p contains zero percent sensors not experiencing attack p with probability* 1*.*

Proof: By *Lemma 9*, for any given $\kappa \in (0, \Delta)$ and any set $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa)$ of sensors which includes at least Δ percent of all sensors, we know from *Lemma 9* that as $K \to \infty$, $P_{\text{error}} \to 0$. Thus, for any given κ and $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa)$,

$$\Pr\left(\varpi\left(\tilde{\mathbf{u}}_{\mathcal{J}}\right) = 1 \,|\, \mathcal{J} \in \mathfrak{C}_1(\kappa)\right) = 1, \tag{3.55}$$

and so if the decision rule in (3.17) yields $\varpi\left(\tilde{\mathbf{u}}_{\mathcal{J}}\right) = 0$, then

$$\mathcal{J} \in \left(\mathfrak{C}_1(\kappa)\right)^{\mathrm{C}}. \tag{3.56}$$

Consider

$$\kappa^* = \frac{1}{N}. \tag{3.57}$$

By the definition of $\mathfrak{C}_1(\kappa)$ in (3.14), it is easy to see that $\mathfrak{C}_1(\kappa^*) \subset \mathfrak{C}_0^{\mathrm{C}}$. On the other hand, for any $\mathcal{S} \in \mathfrak{C}_0^{\mathrm{C}}$, $\exists$ nonempty $\mathcal{S}_1$ and $\mathcal{S}_2 \subset \mathcal{S}$ such that $\mathcal{S}_1 \subset \mathcal{A}_l$ and $\mathcal{S}_2 \subset \mathcal{A}_m$ for some $l \neq m$. Since both $\mathcal{S}_1$ and $\mathcal{S}_2$ at least contain 1 sensor, $\mathcal{P}_{\mathcal{S}_1}, \mathcal{P}_{\mathcal{S}_2} \geq 1/N = \kappa^*$. Thus, $\mathcal{S} \in \mathfrak{C}_1(\kappa^*)$, and hence $\mathfrak{C}_0^{\mathrm{C}} \subset \mathfrak{C}_1(\kappa^*)$. As a result, $\mathfrak{C}_1(\kappa^*) = \mathfrak{C}_0^{\mathrm{C}}$, and therefore $\mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa^*)$ is the power set of $\mathcal{S}_T$, that is, for any set $\mathcal{J} \subset \mathcal{S}_T$ with $\mathcal{P}_{\mathcal{J}} \geq \Delta$,

$$\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa^*), \tag{3.58}$$

By the result in (3.56), if the decision rule in (3.17) yields $\varpi(\tilde{\mathbf{u}}_{\mathcal{J}}) = 0$, then

$$\mathcal{J} \in (\mathfrak{C}_1(\kappa^*))^{\mathrm{C}} = \mathfrak{C}_0. \tag{3.59}$$

Consequently, for any subset $\mathcal{J} \subset \mathcal{S}_T$ with $\mathcal{P}_{\mathcal{J}} \geq \Delta$, we can identify whether $\mathcal{J} \in \mathfrak{C}_0$ or not by checking the output of the decision rule as $K \to \infty$. In other words, we can determine whether $\mathcal{J}$ is homogeneous or not. By examining all possible such subsets in such a way, we can find the collection $\{\tilde{\mathcal{A}}_p\}_{p=0}^{\tilde{P}}$ of the largest subsets, where $\cup_{p=0}^{\tilde{P}} \tilde{\mathcal{A}}_p = \mathcal{S}_T$ and $\tilde{\mathcal{A}}_p \in \mathfrak{C}_0$ for all $p = 0, 1, ..., \tilde{P}$. Moreover, the collection of the largest subsets implies that $\tilde{\mathcal{A}}_l \cup \tilde{\mathcal{A}}_m \notin \mathfrak{C}_0$, if $l \neq m$.

According to *Assumption 8*, $\{\mathcal{A}_p\}_{p=0}^{P}$ is the collection of the largest subsets. Therefore, $\tilde{P} = P$ and $\tilde{\mathcal{A}}_p = \mathcal{A}_p$. Furthermore, the largest one in these $P + 1$ subsets is the set of unattacked sensors and its corresponding percentage equals to $\mathcal{P}_0$. The rest of subsets in $\{\mathcal{A}_p\}_{p=0}^{P}$ are the sets of attacked sensors which are taken over by different types of attacks. ■

As demonstrated by *Theorem 8*, the FC is able to perfectly identify the set of unattacked sensors and categorize the attacked sensors into different groups according to their distinct types of attacks as the number K of time samples at each sensor increases to infinity.

3.3.2 The Number N of Sensors is Sufficiently Large while the Number K of Time Samples at Each Sensor is Finite

We now give a Theorem demonstrating the ability of the FC to identify and categorize different subsets of sensors as the number N of sensors in the sensor network becomes sufficiently large.

Theorem 9 *Consider a sensor network with N sensors. Each sensor observes a finite number K of time samples which satisfies*

$$K \geq -\frac{8\ln 2}{\gamma^* \min\{\Delta\Delta_0, \Delta^2\}} + 1, \tag{3.60}$$

where γ^* are defined in (3.51). Under Assumption 8*, as $N \to \infty$, the FC can determine the number of attacks in the sensor network. Moreover, for the p-th attack $\forall p \geq 0$, the FC can identify a corresponding group of sensors $\tilde{\mathcal{A}}_p$ which satisfies*

$$0 \leq |\tilde{\mathcal{P}}_p - \mathcal{P}_p| \leq \mathcal{P}_p^* < \delta \ \text{ with probability } 1, \tag{3.61}$$

where $\tilde{\mathcal{P}}_p \triangleq |\tilde{\mathcal{A}}_p|/N$, $\mathcal{P}_p^* \triangleq |(\tilde{\mathcal{A}}_p \backslash \mathcal{A}_p) \cup (\mathcal{A}_p \backslash \tilde{\mathcal{A}}_p)|/N$, and

$$\delta \triangleq -\frac{4\ln 2}{\Delta(K-1)\gamma^*}. \tag{3.62}$$

In addition, the group $\tilde{\mathcal{A}}_0$ which approximates the set $\mathcal{A}_0$ of unattacked sensors is still the

largest group in $\{\tilde{\mathcal{A}}_p\}_{p=0}^{P}$.

Proof: See Appendix 3.8.1. ■

As *Theorem 9* demonstrates, with a finite number of time samples at each sensor, even as the number N of sensors in the sensor network grows to infinity, the FC is able to ascertain the number of attacks in the sensor network, but there is no guarantee that the FC can perfectly categorize the sensors into different groups according to distinct attack types. Essentially, *Theorem 9* provides an upper bound which quantifies the maximum percentage of sensors that are misclassified in finding every $\mathcal{A}_p$. As stipulated in (3.62), this upper bound depends on K and monotonically decreases to 0 as K increases to infinity. It is worth noting that the upper bound given in *Theorem 9* is very informative, as it reveals the relationship between the accuracy of the categorization and the requirement of the number of time samples at each sensor, which provides the FC with a tradeoff between the accuracy of the categorization and time efficiency. More specifically, based on our derived results, the FC can obtain the maximum percentage of misclassified sensors in the categorization for various K and choose K such that the maximum percentage of misclassified sensors is tolerable for the scenario of interest.

3.3.3 Discussion

It is worth mentioning that the assumption that the percentage of the unattacked sensors is larger than any percentage of similarly attacked sensors is not necessary for *Lemma 9, Theorem 8* and *9*. Once all the sensors have been categorized, the side information that the unattacked sensors constitute the largest fractions of the sensors can help the FC to identify the group of unattacked sensors.

A particular note of interest is that the performance improvements in categorization induced by increasing K or N are different. As specified in (3.19), *Lemma 9* implies that increasing K or N gives rise to the same effect on reducing the error probability of the hypothesis test. However, *Theorem 8* and *Theorem 9* indicate that under *Assumption 8*, the FC is able to determine the number P of attacks in the sensor network and achieve the correct categorization as $K \to \infty$, while as $N \to \infty$ with finite but sufficiently large K, it can be shown that the FC can also ascertain P and upper bound the maximum percentage of misclassified sensors in the categorization. In this sense, the effect of enlarging the size of the sensor network is not equivalent to that of increasing the number of time samples at each sensor in the presence of attacks. In other words, adding more spatial observations is different from adding more temporal observations in the sensor network under attack. The reason is that the additional temporal observations can provide the FC with more information than the additional spatial observations. To be specific, when the number of sensors in the sensor network increases, the set of additional observations produced by the new sensors is a mixture of differently attacked observations. Thus, the FC needs to categorize the additional sensors into different groups according to distinct types of attacks, since the FC is unaware of the attack each additional sensor undergoes. In contrast, as K increases, the information buried in the additional temporal observations not only encompasses that implied by the additional spatial observations, but also conveys that, according to *Assumption 8*, the additional temporal observations from a certain sensor are statistically identical to those original observations from the same sensor, and hence they are under the same attack. Thus, the FC doesn't need to categorize the additional temporal observations, since the additional temporal observations will be automatically categorized once the FC has categorized the original observations. For this reason, one can intuitively expect that the performance of

categorization can be better ameliorated by increasing the number of time samples at each sensor rather than increasing the number of sensors in the sensor network.

The number of hypothesis tests needed to implement the categorization approach proposed in *Theorem 8* and *Theorem 9*, which is referred to as its complexity, deserves some discussion. Admittedly, the approach we proposed in *Theorem 8* and *Theorem 9* is to check all possible subsets of the sensor network. It is observed that for a sensor network $\mathcal{S}_T$ with N sensors, the number of nonempty subsets is $(2^N - 1)$. In addition, each subset contains at most $(2^N - 1)$ nonempty sub-subsets. Thus, the complexity of the proposed approach must be smaller than 8^N. For the scenario discussed in *Theorem 8*, as K increases, the complexity remains the same and finite, and hence the proposed approach is amenable to implementation. For the scenario in *Theorem 9*, the complexity of the proposed approach can be very high when N is sufficiently large. Clearly, the task of identification and categorization, no matter what algorithm is applied, must be arduous when the number of sensors is large. The purpose of presenting the results in *Theorem 9* is twofold. On one hand, the results in *Theorem 9* are a nice complement to those results in *Theorem 8* which further the understanding of the categorization performance in the two types of asymptotic regions and highlight some differences between increasing N as opposed to K. On the other hand, we feel it is useful to demonstrate that it is possible, at least in theory, to categorize the sensors into different groups for these two cases in the specific sense previously discussed. These results can encourage further investigation on the pursuit of more efficient identification and categorization algorithms for large scale sensor networks.

3.4 Fisher Information Matrix in the Presence of Attacks

As shown in Section 3.3, when each sensor accumulates sufficiently many time samples or the size of the sensor network is sufficiently large, the FC is able to determine the number of attacks in the sensor network and categorize the sensors into different groups according to distinct types of attacks perfectly or with a tolerable small misclassification which can be ignored. Thus, in the following part, we assume that the sensors have been well identified and categorized into $\{\mathcal{A}_p\}_{p=0}^{P}$. Then we attempt to estimate θ. There are two approaches: (1) ignore the data at the attacked sensors. (2) Use the data at the attacked sensors. We can easily take approach (1) without estimating any parameters describing the attacks. However, to attempt to take approach (2), and potentially do better than approach (1), we will investigate the performance of the joint estimation of the desired parameter and the unknown attack parameters in this section.

Although Section 3.3 showed that $\{\mathcal{A}_p\}_{p=0}^{P}$ can be determined, the FC is still unaware of the attack parameters $\mathbf{\Psi}_p$ for $p = 1, 2, .., P$. Let $\mathbf{\Theta}$ denote a vector containing the parameter θ along with all the unknown parameters of the attacks

$$\mathbf{\Theta} \triangleq [\theta, \psi_{1,0}, \psi_{1,1}, ..., \psi_{P,0}, \psi_{P,1}]^T. \tag{3.63}$$

The log-likelihood function evaluated at $\tilde{\mathbf{u}} = \mathbf{r}$ is

$$\begin{aligned} L(\boldsymbol{\Theta}) &= \ln \Pr(\tilde{\mathbf{u}} = \mathbf{r} \,|\boldsymbol{\Theta}) \\ &= \ln \prod_{p=0}^{P} \prod_{j \in \mathcal{A}_p} \prod_{k=1}^{K} \prod_{r'_{jk}=0}^{1} \Pr\left(\tilde{u}_{jk} = r'_{jk} \,|\boldsymbol{\Theta}\right)^{\mathbb{1}\{r_{jk} = r'_{jk}\}} \\ &= \sum_{p=0}^{P} \sum_{j \in \mathcal{A}_p} \sum_{k=1}^{K} \sum_{r'_{jk}=0}^{1} \mathbb{1}\left\{r_{jk} = r'_{jk}\right\} \ln \Pr\left(\tilde{u}_{jk} = r'_{jk} \,|\boldsymbol{\Theta}\right). \end{aligned}$$

In order to gain insight into the impact of attacks and evaluate the estimation performance, we carry out an analysis of the FIM for $\boldsymbol{\Theta}$.

By noting that $\tilde{p}(\boldsymbol{\Psi}_p, \theta) = \Pr(\tilde{u}_{jk} = 1 \,|\theta)$, $\forall j \in \mathcal{A}_p$, the (l, m)-th element of the FIM for $\boldsymbol{\Theta}$, therefore, is given by

$$\begin{aligned} [\mathbf{J}(\boldsymbol{\Theta})]_{l,m} &= -\mathbb{E}\left\{\frac{\partial^2 L(\boldsymbol{\Theta})}{\partial \Theta_l \partial \Theta_m}\right\} \\ &= KN \sum_{p=0}^{P} \frac{\mathcal{P}_p}{\tilde{p}(\boldsymbol{\Psi}_p, \theta)\left[1 - \tilde{p}(\boldsymbol{\Psi}_p, \theta)\right]} \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p, \theta)}{\partial \Theta_l} \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p, \theta)}{\partial \Theta_m}. \end{aligned} \tag{3.64}$$

Define $\boldsymbol{\phi}_l \triangleq \frac{\partial \tilde{p}(\boldsymbol{\Psi}_l, \theta)}{\partial \boldsymbol{\Theta}}$. Then, the FIM described in (3.64) can be formulated as

$$\mathbf{J}(\boldsymbol{\Theta}) = KN \sum_{p=0}^{P} \frac{\mathcal{P}_p \boldsymbol{\phi}_p \boldsymbol{\phi}_p^T}{\tilde{p}(\boldsymbol{\Psi}_p, \theta)\left[1 - \tilde{p}(\boldsymbol{\Psi}_p, \theta)\right]} = KN \sum_{p=0}^{P} \varrho_p \boldsymbol{\Phi}_p, \tag{3.65}$$

where $\varrho_p \triangleq \frac{\mathcal{P}_p}{\tilde{p}(\boldsymbol{\Psi}_p, \theta)[1 - \tilde{p}(\boldsymbol{\Psi}_p, \theta)]}$ and $\boldsymbol{\Phi}_p \triangleq \boldsymbol{\phi}_p \boldsymbol{\phi}_p^T$. Apparently, $\operatorname{rank}(\boldsymbol{\Phi}_p) = 1$.

In the following theorem, we provide results concerning the FIM in the presence of attacks.

Theorem 10 *In the presence of attacks, the FIM* $\mathbf{J}(\boldsymbol{\Theta})$ *for* $\boldsymbol{\Theta}$ *described in (3.65) is singular.*

Proof: Since $\operatorname{rank}(\mathbf{\Phi}_p) = 1$ for all p, and noting that the dimensions of $\mathbf{J}(\mathbf{\Theta})$ are $(2P+1) \times (2P+1)$, we can obtain

$$\begin{aligned}\operatorname{rank}(\mathbf{J}(\mathbf{\Theta})) &= \operatorname{rank}\left(KN \sum_{p=0}^{P} \varrho_p \mathbf{\Phi}_p\right) \\ &\leq \sum_{p=0}^{P} \operatorname{rank}(\mathbf{\Phi}_p) = P+1 < 2P+1,\end{aligned}$$

and hence, the FIM $\mathbf{J}(\mathbf{\Theta})$ for $\mathbf{\Theta}$ is singular. ■

Theorem 10 reveals that we cannot jointly estimate the parameter θ and the attack parameters with an accuracy that grows with KN. Actually, this negative conclusion is conceivable. Reexamining the after-attack pmf in (3.7) for the sensor taken over by the p-th attack, it is observed that for any given before-attack probability $\Pr(u_{jk} = 1 \,|\theta)$ and after-attack probability $\Pr(\tilde{u}_{jk} = 1 \,|\theta)$, the pair of attack parameters $(\psi_{p,0}, \psi_{p,1})$ of the p-th attack in (3.6) is not unique. Moreover, there exists an infinite number of pairs of attack parameters $(\psi_{p,0}, \psi_{p,1})$ which can map the given before-attack probability $\Pr(u_{jk} = 1 \,|\theta)$ to the given after-attack probability $\Pr(\tilde{u}_{jk} = 1 \,|\theta)$ by using (3.6). From the perspective of the FC, even though the FC can ascertain the after-attack probability $\Pr(\tilde{u}_{jk} = 1 \,|\theta)$ as $KN \to \infty$, the FC is unable to determine the exact $(\psi_{p,0}, \psi_{p,1})$ employed by the p-th attack because of this non-uniqueness. For this reason, it is reasonable that in the presence of attackers who modify the data in the manner shown in (3.6), the corresponding FIM for $\mathbf{\Theta}$ is singular. To this end, it is of great interest to investigate approaches which lead to nonsingular FIMs to allow us to efficiently estimate $\mathbf{\Theta}$ and take advantage of attacked observations to improve the estimation performance for the parameter θ.

3.5 Time-variant Quantization Approach to Achieve Nonsingular FIM

In this section, we first develop the time-variant quantization approach (TQA) to overcome the singular FIM in the presence of attacks. We then examine the CRB performance of our approach and compare it to that of the simple estimation approach where only the set $\mathcal{A}_0$ of unattacked sensors are used to estimate the parameter θ. Furthermore, necessary and sufficient conditions are derived under which the attacked observations can be utilized in the proposed fashion to improve the CRB performance for estimating the parameter θ.

3.5.1 Time-variant Quantization Approach

In the time-variant quantization approach, the quantizer at each sensor is equipped with a set of Q distinct thresholds $\mathcal{Q} = \{\tau_1, \tau_2, ..., \tau_Q\}$. In each of Q different time slots $\{\mathcal{T}_t\}_{t=1}^{Q}$, where $\mathcal{T}_t$ contains K_t time samples and $\sum_{t=1}^{Q} K_t = K$, the quantizer employs a different threshold to quantize its time samples into one-bit observations which are sent to the FC. We assume that the length K_t of each time slot $\mathcal{T}_t$ is the same for all the sensors, and in each time slot $\mathcal{T}_t$, all sensors use an identical threshold τ_t to quantize their time samples. In this manner, the after-attack pmf of the quantized observations received at the FC in the t-th time slot can be written using, $\forall k \in \mathcal{T}_t$ and $j \in \mathcal{A}_p$,

$$\begin{aligned}
&\tilde{p}\left(\mathbf{\Psi}_p, \theta, t\right) \stackrel{\Delta}{=} \Pr\left(\tilde{u}_{jk} = 1 \,|\theta\right) \\
&= \left(1 - \psi_{p,0}\right) \Pr\left(u_{jk} = 0 \,|\theta\right) + \psi_{p,1} \Pr\left(u_{jk} = 1 \,|\theta\right) \\
&= \left(1 - \psi_{p,0} - \psi_{p,1}\right) F\left(\tau_t - \theta\right) + \psi_{p,1}, \qquad (3.66)
\end{aligned}$$

Then, the log-likelihood function of the TQA, evaluated at $\tilde{\mathbf{u}} = \mathbf{r}$, is given by

$$\begin{aligned} & L_{\mathrm{TQA}}(\boldsymbol{\Theta}) \\ & = \sum_{p=0}^{P} \sum_{j \in \mathcal{A}_p} \sum_{t=1}^{Q} \sum_{k \in \mathcal{T}_t} \sum_{r'_{jk}=0}^{1} \mathbb{1}\left\{r_{jk} = r'_{jk}\right\} \ln \Pr\left(\tilde{u}_{jk} = r'_{jk} \,|\boldsymbol{\Theta}\right). \end{aligned}$$

Applying a similar argument as in (3.64), the (l,m)-th element of the corresponding FIM for $\boldsymbol{\Theta}$ can be calculated as

$$[\mathbf{J}_{\mathrm{TQA}}(\boldsymbol{\Theta})]_{l,m} = N \sum_{p=0}^{P} \sum_{t=1}^{Q} \frac{K_t \mathcal{P}_p \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \Theta_l} \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \Theta_m}}{\tilde{p}(\boldsymbol{\Psi}_p,\theta,t)\left[1-\tilde{p}(\boldsymbol{\Psi}_p,\theta,t)\right]}. \tag{3.67}$$

Define

$$\begin{aligned} \boldsymbol{\phi}_{p,t} & \triangleq \varrho_{p,t} \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \boldsymbol{\Theta}} \\ & = \varrho_{p,t} \Bigg[\frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \theta}, \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \psi_{1,0}}, \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \psi_{1,1}}, \\ & \qquad ..., \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \psi_{P,0}}, \frac{\partial \tilde{p}(\boldsymbol{\Psi}_p,\theta,t)}{\partial \psi_{P,1}} \Bigg]^T, \end{aligned} \tag{3.68}$$

where $\varrho_{p,t} \triangleq \left(\frac{K_t \mathcal{P}_p}{\tilde{p}(\boldsymbol{\Psi}_p,\theta,t)[1-\tilde{p}(\boldsymbol{\Psi}_p,\theta,t)]}\right)^{\frac{1}{2}}$. Therefore, for all t and $p \geq 1$, we can obtain

$$\begin{aligned} \boldsymbol{\phi}_{p,t} = \varrho_{p,t} [& - \left(1 - \psi_{p,0} - \psi_{p,1}\right) f\left(\tau_t - \theta\right), 0, \\ & ..., 0, \underbrace{-F\left(\tau_t - \theta\right)}_{\text{the } (2p)\text{-th element}}, \underbrace{-F\left(\tau_t - \theta\right) + 1}_{\text{the } (2p+1)\text{-th element}}, 0, ..., 0]^T, \end{aligned} \tag{3.69}$$

while, for all t and $p = 0$,

$$\boldsymbol{\phi}_{0,t} = \varrho_{0,t} [f\left(\tau_t - \theta\right), 0, ..., 0]^T. \tag{3.70}$$

As a result, the FIM in (3.67) can be rewritten as

$$\begin{aligned}\mathbf{J}_{\mathrm{TQA}}(\boldsymbol{\Theta}) &= N\sum_{p=0}^{P}\sum_{t=1}^{Q}\boldsymbol{\phi}_{p,t}\boldsymbol{\phi}_{p,t}^{T} \\ &= N\sum_{p=0}^{P}\boldsymbol{\Xi}_{p}\boldsymbol{\Xi}_{p}^{T} = N\sum_{p=0}^{P}\boldsymbol{\Gamma}_{p},\end{aligned} \tag{3.71}$$

where

$$\boldsymbol{\Xi}_{p} \triangleq \left[\boldsymbol{\phi}_{p,1}, \boldsymbol{\phi}_{p,2}, ..., \boldsymbol{\phi}_{p,Q}\right], \tag{3.72}$$

$$\boldsymbol{\Gamma}_{p} \triangleq \boldsymbol{\Xi}_{p}\boldsymbol{\Xi}_{p}^{T} \succeq 0. \tag{3.73}$$

Theorem 11 *If $Q \geq 2$, then the FIM for $\boldsymbol{\Theta}$ in the TQA, i.e., $\mathbf{J}_{TQA}(\boldsymbol{\Theta})$ described in (3.71), is nonsingular for any value of θ.*

Proof: Since $F(\cdot)$ is a strictly monotonic function, $F(\tau_l - \theta) \neq F(\tau_m - \theta)$ provided $l \neq m$. This implies $\boldsymbol{\phi}_{p,l}$ and $\boldsymbol{\phi}_{p,m}$ in (3.69) are linearly independent for $p \geq 1$. As a result, if $Q \geq 2$, then $\forall p \geq 1$,

$$\begin{aligned}\mathrm{rank}(\boldsymbol{\Gamma}_{p}) &= \mathrm{rank}\left(\boldsymbol{\Xi}_{p}\boldsymbol{\Xi}_{p}^{T}\right) = \mathrm{rank}(\boldsymbol{\Xi}_{p}) \\ &\geq \mathrm{rank}\left(\boldsymbol{\Xi}_{p}\left(\{2p, 2p+1\}, \{:\}\right)\right) = 2.\end{aligned} \tag{3.74}$$

Noticing that

$$\boldsymbol{\Gamma}_{p}\left(\{2p, 2p+1\}, \{2p, 2p+1\}\right) = \boldsymbol{\Xi}_{p}\left(\{2p, 2p+1\}, \{:\}\right)\left[\boldsymbol{\Xi}_{p}\left(\{2p, 2p+1\}, \{:\}\right)\right]^{T} \succeq 0, \tag{3.75}$$

we can obtain

$$\begin{aligned}&\operatorname{rank}\left(\boldsymbol{\Gamma}_p\left(\{2p,2p+1\},\{2p,2p+1\}\right)\right)\\&=\operatorname{rank}\left(\boldsymbol{\Xi}_p\left(\{2p,2p+1\},\{:\}\right)\right)=2,\quad\forall p\geq 1.\end{aligned}\tag{3.76}$$

and hence

$$\boldsymbol{\Gamma}_p\left(\{2p,2p+1\},\{2p,2p+1\}\right)\succ 0,\quad\forall p\geq 1.\tag{3.77}$$

Moreover, since for all $p\geq 1$,

$$\begin{aligned}&\boldsymbol{\Gamma}_p\left(\{1,2p,2p+1\},\{1,2p,2p+1\}\right)\\&=\boldsymbol{\Xi}_p\left(\{1,2p,2p+1\},\{:\}\right)\left[\boldsymbol{\Xi}_p\left(\{1,2p,2p+1\},\{:\}\right)\right]^T\succeq 0,\end{aligned}$$

we know that for any $\mathbf{w}\neq\mathbf{0}$,

$$\mathbf{w}^T\boldsymbol{\Gamma}_p\left(\{1,2p,2p+1\},\{1,2p,2p+1\}\right)\mathbf{w}\geq 0.\tag{3.78}$$

Noting that $\boldsymbol{\Gamma}_0$ only contains one nonzero element which is $[\boldsymbol{\Gamma}_0]_{1,1}=\sum_{t=1}^{Q}\varrho_{0,t}^2 f^2\left(\tau_t-\theta\right)>0$, yields $\forall p\geq 1$ and for any $\mathbf{w}=(\omega,\boldsymbol{\nu})^T\neq\mathbf{0}$,

$$\begin{aligned}&\mathbf{w}^T\left(\left[\frac{1}{P}\boldsymbol{\Gamma}_0+\boldsymbol{\Gamma}_p\right]\left(\{1,2p,2p+1\},\{1,2p,2p+1\}\right)\right)\mathbf{w}\\&=\frac{\omega^2}{P}[\boldsymbol{\Gamma}_0]_{1,1}+\mathbf{w}^T\boldsymbol{\Gamma}_p\left(\{1,2p,2p+1\},\{1,2p,2p+1\}\right)\mathbf{w}\end{aligned}\tag{3.79}$$

If $\omega \neq 0$, then by (3.78) and (3.79), we can obtain

$$\begin{aligned}
&\mathbf{w}^T \left(\left[\frac{1}{P} \mathbf{\Gamma}_0 + \mathbf{\Gamma}_p \right] (\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}) \right) \mathbf{w} \\
&\geq \frac{\omega^2}{P} [\mathbf{\Gamma}_0]_{1,1} > 0.
\end{aligned} \tag{3.80}$$

If $\omega = 0$, then $\boldsymbol{\nu} \neq \mathbf{0}$, and hence (3.79) simplifies to

$$\begin{aligned}
&\mathbf{w}^T \left(\left[\frac{1}{P} \mathbf{\Gamma}_0 + \mathbf{\Gamma}_p \right] (\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}) \right) \mathbf{w} \\
&= \boldsymbol{\nu}^T \mathbf{\Gamma}_p (\{2p, 2p+1\}, \{2p, 2p+1\}) \boldsymbol{\nu} > 0,
\end{aligned} \tag{3.81}$$

where we have employed $\mathbf{\Gamma}_p (\{2p, 2p+1\}, \{2p, 2p+1\}) \succ 0, \ \forall p \geq 1$ in (3.77).

In consequence, we have shown that $\forall p \geq 1$,

$$\left[\frac{1}{P} \mathbf{\Gamma}_0 + \mathbf{\Gamma}_p \right] (\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}) \succ 0. \tag{3.82}$$

As a result, for any given vector $\mathbf{v} \neq \mathbf{0}$, we can obtain

$$\begin{aligned}
&\mathbf{v}^T \left(\sum_{p=0}^{P} \mathbf{\Gamma}_p \right) \mathbf{v} = \sum_{p=1}^{P} \mathbf{v}^T \left(\frac{1}{P} \mathbf{\Gamma}_0 + \mathbf{\Gamma}_p \right) \mathbf{v} \\
&= \sum_{p=1}^{P} \Bigg\{ \mathbf{v}^T (1, 2p, 2p+1) \\
&\qquad \left(\left[\frac{1}{P} \mathbf{\Gamma}_0 + \mathbf{\Gamma}_p \right] (\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}) \right) \\
&\qquad\quad \mathbf{v} (1, 2p, 2p+1) \Bigg\} > 0.
\end{aligned} \tag{3.83}$$

Thus, $\sum_{p=0}^{P} \mathbf{\Gamma}_p \succ 0$ for $Q \geq 2$, and hence $\mathbf{J}_{\mathrm{TQA}} (\mathbf{\Theta})$ described in (3.71) is nonsingular. ■

Theorem 11 implies that as long as $K \geq 2$, we can always obtain a nonsingular FIM as described in (3.71), since we can divide the time samples into at least two time slots and employ distinct thresholds to quantize the time samples in different time slots. To gain insight into the proposed TQA estimation scheme, we now revisit the relationship in (3.6) between the before-attack pmf, the after-attack pmf, and the pair of attack parameters $(\psi_{p,0}, \psi_{p,1})$. It is seen from (3.7) that for any given before-attack pmf and after-attack pmf, the attack parameters pair $(\psi_{p,0}, \psi_{p,1})$ will span a one dimensional space since if $\psi_{p,0}$ is chosen then $\psi_{p,1}$ is also determined to satisfy the equation shown in (3.7). If we change the threshold of the quantizer at each sensor, we can obtain another pair of before-attack and after-attack pmfs which must be related by the same $(\psi_{p,0}, \psi_{p,1})$ under our assumption that these parameters are fixed over the estimation interval. This second set of equations combined with the first set will yield a unique solution for $(\psi_{p,0}, \psi_{p,1})$. Consequently, it is intuitively possible for the FC to jointly estimate the parameter θ and the attack parameters simultaneously when the quantized observations are generated by at least two distinct thresholds.

3.5.2 CRB Performance Analysis of the Time-variant Quantization Approach

The main goal of the sensor network is to estimate the parameter θ. Hence, the FC can carry out a simple estimation approach (SEA) which just utilizes the set $\mathcal{A}_0$ of unattacked sensors to estimate the parameter θ rather than employing the TQA to obtain both the estimate of θ and the estimates of the attack parameters with the purpose of improving the estimation of θ. The SEA is obviously easier to implement in practice and can lower the computational complexity of the estimation by reducing the number of parameters to be estimated from $2P + 1$ to 1. However, the SEA discards the possible information on

θ buried in the attacked observations which may lead to performance loss in estimating θ when compared to the TQA. In this subsection, the CRB performance of estimating θ by employing the TQA is compared to that in the SEA. The comparisons are made assuming both approaches use the same set of Q distinct thresholds $\{\tau_1, \tau_2, ..., \tau_Q\}$ over the same Q different time slots $\mathcal{T}_t$ to provide a fair comparison. We also develop necessary and sufficient conditions under which the CRB performance of estimating θ can be improved by using the attacked observations in our proposed fashion.

For the SEA, the FC ignores the observations from the attacked sensors, and only makes use of the unattacked observations to estimate the parameter θ. By noting that

$$\begin{aligned}[\mathbf{\Gamma}_0]_{1,1} &= \sum_{t=1}^{Q} \frac{K_t \mathcal{P}_0}{\tilde{p}\left(\mathbf{\Psi}_0, \theta, t\right)\left[1-\tilde{p}\left(\mathbf{\Psi}_0, \theta, t\right)\right]}\left[\frac{\partial \tilde{p}\left(\mathbf{\Psi}_0, \theta, t\right)}{\partial \theta}\right]^2 \\ &= \mathcal{P}_0 \sum_{t=1}^{Q} \frac{K_t f^2\left(\tau_t-\theta\right)}{F\left(\tau_t-\theta\right)\left[1-F\left(\tau_t-\theta\right)\right]}, \end{aligned} \tag{3.84}$$

the Fisher Information Matrix in (3.71) degenerates to a scalar which can be expressed as

$$J_{\mathrm{SEA}}\left(\theta\right) = N[\mathbf{\Gamma}_0]_{1,1} = N\mathcal{P}_0 \sum_{t=1}^{Q} \frac{K_t f^2\left(\tau_t-\theta\right)}{F\left(\tau_t-\theta\right)\left[1-F\left(\tau_t-\theta\right)\right]}.$$

Hence, the corresponding CRB performance of SEA for θ is given by

$$\begin{aligned} CRB_{\mathrm{SEA}}\left(\theta\right) &= \frac{1}{J_{\mathrm{SEA}}\left(\theta\right)} \\ &= \frac{1}{N\mathcal{P}_0}\left\{\sum_{t=1}^{Q} \frac{K_t f^2\left(\tau_t-\theta\right)}{F\left(\tau_t-\theta\right)\left[1-F\left(\tau_t-\theta\right)\right]}\right\}^{-1}. \end{aligned} \tag{3.85}$$

On the other hand, considering the TQA which takes advantage of attacked observations to estimate the parameter θ, the CRB of estimating θ is the $(1,1)$-th element of the inverse

of the FIM described in (3.71), i.e. $CRB_{\mathrm{TQA}}(\theta) = \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\right]_{1,1}$.

Using the expression $\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\,\mathbf{J}_{\mathrm{TQA}}(\mathbf{\Theta}) = \mathbf{I}$, we have following identities,

$$\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\right]_{1,1}[\mathbf{J}_{\mathrm{TQA}}(\mathbf{\Theta})]_{1,1} + \sum_{l=2}^{2P+1}\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\right]_{1,l}[\mathbf{J}_{\mathrm{TQA}}(\mathbf{\Theta})]_{l,1} = 1, \tag{3.86}$$

and $\forall m = 2, 3, ..., 2P+1$,

$$\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\right]_{1,1}[\mathbf{J}_{\mathrm{TQA}}(\mathbf{\Theta})]_{1,m} + \sum_{l=2}^{2P+1}\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\right]_{1,l}[\mathbf{J}_{\mathrm{TQA}}(\mathbf{\Theta})]_{l,m} = 0. \tag{3.87}$$

Since there is one nonzero element in $\mathbf{\Gamma}_0$ which is the $(1,1)$ element and $\mathbf{\Gamma}_v$ only has 9 nonzero elements at the intersections of the 1st, $2v$-th, and $(2v+1)$-th rows and columns for $v = 1, 2, ..., P$, the nonzero elements of $\mathbf{\Gamma}_v$ and $\mathbf{\Gamma}_{v'}$ do not overlap except for the $(1,1)$ element provided that $v \neq v'$. As a result, taking into account $\mathbf{J}_{\mathrm{TQA}}(\mathbf{\Theta}) = N\sum_{v=0}^{P}\mathbf{\Gamma}_v$ from (3.71), (3.86) can be rewritten as

$$\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\mathbf{\Theta})\right]_{1,1}\left\{[\mathbf{\Gamma}_0]_{1,1} + \sum_{p=1}^{P} h_p(\mathbf{\Theta})\right\} = \frac{1}{N}, \tag{3.88}$$

where

$$h_p(\boldsymbol{\Theta}) \triangleq [\boldsymbol{\Gamma}_p]_{1,1} + \frac{\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p}}{\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1}}[\boldsymbol{\Gamma}_p]_{2p,1} + \frac{\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p+1}}{\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1}}[\boldsymbol{\Gamma}_p]_{2p+1,1}. \tag{3.89}$$

Similarly, (3.87) simplifies to

$$\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1}[\boldsymbol{\Gamma}_p]_{1,2p} + \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p}[\boldsymbol{\Gamma}_p]_{2p,2p} + \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p+1}[\boldsymbol{\Gamma}_p]_{2p+1,2p} = 0, \tag{3.90}$$

$$\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1}[\boldsymbol{\Gamma}_p]_{1,2p+1} + \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p}[\boldsymbol{\Gamma}_p]_{2p,2p+1} + \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p+1}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1} = 0 \tag{3.91}$$

for all $p = 1, 2, ..., P$.

Note that $\boldsymbol{\Gamma}_p$ is symmetric for all p. Then from (3.90) and (3.91), $\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p}$ and $\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p+1}$ can be determined by

$$\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p} = \frac{[\boldsymbol{\Gamma}_p]_{1,2p+1}[\boldsymbol{\Gamma}_p]_{2p,2p+1} - [\boldsymbol{\Gamma}_p]_{1,2p}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1}}{[\boldsymbol{\Gamma}_p]_{2p,2p}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1} - \left\{[\boldsymbol{\Gamma}_p]_{2p,2p+1}\right\}^2}\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1}, \tag{3.92}$$

and

$$\begin{aligned}
&\left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,2p+1} \\
&= \frac{[\boldsymbol{\Gamma}_p]_{1,2p}[\boldsymbol{\Gamma}_p]_{2p,2p+1} - [\boldsymbol{\Gamma}_p]_{1,2p+1}[\boldsymbol{\Gamma}_p]_{2p,2p}}{[\boldsymbol{\Gamma}_p]_{2p,2p}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1} - \left\{[\boldsymbol{\Gamma}_p]_{2p,2p+1}\right\}^2} \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1}.
\end{aligned} \tag{3.93}$$

Substituting (3.92) and (3.93) into (3.89), we obtain

$$\begin{aligned}
&h_p(\boldsymbol{\Theta}) \\
&= [\boldsymbol{\Gamma}_p]_{1,1} + \frac{[\boldsymbol{\Gamma}_p]_{1,2p+1}[\boldsymbol{\Gamma}_p]_{2p,2p+1} - [\boldsymbol{\Gamma}_p]_{1,2p}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1}}{[\boldsymbol{\Gamma}_p]_{2p,2p}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1} - \left\{[\boldsymbol{\Gamma}_p]_{2p,2p+1}\right\}^2} [\boldsymbol{\Gamma}_p]_{1,2p} \\
&\quad + \frac{[\boldsymbol{\Gamma}_p]_{1,2p}[\boldsymbol{\Gamma}_p]_{2p,2p+1} - [\boldsymbol{\Gamma}_p]_{1,2p+1}[\boldsymbol{\Gamma}_p]_{2p,2p}}{[\boldsymbol{\Gamma}_p]_{2p,2p}[\boldsymbol{\Gamma}_p]_{2p+1,2p+1} - \left\{[\boldsymbol{\Gamma}_p]_{2p,2p+1}\right\}^2} [\boldsymbol{\Gamma}_p]_{1,2p+1} \\
&= \frac{\det\left(\boldsymbol{\Gamma}_p\left(\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}\right)\right)}{\det\left(\boldsymbol{\Gamma}_p\left(\{2p, 2p+1\}, \{2p, 2p+1\}\right)\right)},
\end{aligned} \tag{3.94}$$

which yields

$$\begin{aligned}
&CRB_{\mathrm{TQA}}(\theta) = \left[\mathbf{J}_{\mathrm{TQA}}^{-1}(\boldsymbol{\Theta})\right]_{1,1} \\
&= \frac{1}{N}\left\{[\boldsymbol{\Gamma}_0]_{1,1} + \sum_{p=1}^{P} h_p(\boldsymbol{\Theta})\right\}^{-1} \\
&= \frac{1}{N}\Bigg\{[\boldsymbol{\Gamma}_0]_{1,1} \\
&\quad + \sum_{p=1}^{P} \frac{\det\left(\boldsymbol{\Gamma}_p\left(\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}\right)\right)}{\det\left(\boldsymbol{\Gamma}_p\left(\{2p, 2p+1\}, \{2p, 2p+1\}\right)\right)}\Bigg\}^{-1}
\end{aligned} \tag{3.95}$$

where $\boldsymbol{\Gamma}_p$ is defined in (3.73).

In the following theorem, we provide the result with regard to the necessary and sufficient

conditions under which the CRB performance of estimating θ can be improved by employing TQA.

Theorem 12 *The CRB performance for θ can be improved by utilizing the observations from the set of attacked sensors in our proposed fashion, if and only if at least one member of the set $\{\boldsymbol{\Xi}_p\}_{p=1}^{P}$ has $\operatorname{rank}(\boldsymbol{\Xi}_p) = 3$, where $\boldsymbol{\Xi}_p$ is defined in (3.72). Otherwise, there is no CRB improvement, but also no loss in performance, from utilizing the attacked observations.*

Proof: Since $\boldsymbol{\Gamma}_p(\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}) \succeq 0$ and $\boldsymbol{\Gamma}_p(\{2p, 2p+1\}, \{2p, 2p+1\}) \succ 0$, we can conclude that $\forall p = 1, 2, ..., P$,

$$\frac{\det(\boldsymbol{\Gamma}_p(\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}))}{\det(\boldsymbol{\Gamma}_p(\{2p, 2p+1\}, \{2p, 2p+1\}))} \geq 0. \tag{3.96}$$

Consequently, by noticing (3.95), we can obtain

$$CRB_{\mathrm{TQA}}(\theta) \leq \frac{1}{N}\left\{[\boldsymbol{\Gamma}_0]_{1,1}\right\}^{-1} = CRB_{\mathrm{SEA}}(\theta). \tag{3.97}$$

Moreover, the equality holds if and only if $\forall p = 1, 2, ..., P$,

$$\det(\boldsymbol{\Gamma}_p(\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\})) = 0, \tag{3.98}$$

which implies that

$$\operatorname{rank}(\boldsymbol{\Xi}_p) = 2, \quad \forall p = 1, 2, ..., P. \tag{3.99}$$

Noting that $\boldsymbol{\Xi}_p$ only contains at most 3 nonzero rows, we know $\operatorname{rank}(\boldsymbol{\Xi}_p) \leq 3$. As a result, in order to improve the CRB for θ by taking advantage of the attacked observations, we must have $\operatorname{rank}(\boldsymbol{\Xi}_p) = 3$ for some p. ■

The result in *Theorem 12* implies that a proper estimation approach will never lead to any loss in asymptotic performance from using the observations from the attacked sensors. In order to obtain rank $(\mathbf{\Xi}_p) = 3$ for some p, the number of thresholds Q cannot be less than 3. So that the number of time samples K at each sensor is required to be larger or equal to 3. Generally, if Q is large, it is easy to obtain rank $(\mathbf{\Xi}_p) = 3$ for some p. However, for some specific attacks, using the observations from the attacked sensors in the fashion of the TQA will not provide better asymptotic estimation performance. For example, if for all p, the p-th attack sets $\psi_{p,0} + \psi_{p,1} = 1$, then each $\mathbf{\Xi}_p$ only contains at most 2 nonzero rows for any thresholds. Hence rank $(\mathbf{\Xi}_p) = 2 < 3$ for all possible set of thresholds $\mathcal{Q} = \{\tau_1, \tau_2, ..., \tau_Q\}$. For this scenario, it is seen from (3.7) that the after-attack pmf is independent of the parameter θ, thus it is obvious that the attacked observations cannot improve the CRB for θ.

In order to evaluate the superiority of TQA, we are primarily interested in the relative CRB gain which is defined as the ratio of the CRB for θ when applying SEA relative to that employing TQA. From (3.85) and (3.95), the relative CRB gain can be obtained as

$$\begin{aligned} CRB_{\text{relative gain}}(\theta) &\triangleq \frac{CRB_{\text{SEA}}(\theta)}{CRB_{\text{TQA}}(\theta)} \\ &= \frac{\frac{1}{N}\left\{[\mathbf{\Gamma}_0]_{1,1}\right\}^{-1}}{\frac{1}{N}\left\{[\mathbf{\Gamma}_0]_{1,1} + \sum_{p=1}^{P} \frac{\det(\mathbf{\Gamma}_p(\{1,2p,2p+1\},\{1,2p,2p+1\}))}{\det(\mathbf{\Gamma}_p(\{2p,2p+1\},\{2p,2p+1\}))}\right\}^{-1}} \\ &= 1 + \frac{1}{[\mathbf{\Gamma}_0]_{1,1}} \sum_{p=1}^{P} \frac{\det\left(\mathbf{\Gamma}_p\left(\{1, 2p, 2p+1\}, \{1, 2p, 2p+1\}\right)\right)}{\det\left(\mathbf{\Gamma}_p\left(\{2p, 2p+1\}, \{2p, 2p+1\}\right)\right)}, \end{aligned} \tag{3.100}$$

which is not a function of N but depends on the percentage of attacked sensors, the thresholds $\mathcal{Q} = \{\tau_1, \tau_2, ..., \tau_Q\}$, the attack parameters, and the value of θ. It can be shown from (3.100) that if the attack parameters $(\psi_{p,0}, \psi_{p,1})$ are close to $(0, 0)$ or $(1, 1)$ for all p, the FC can

expect to attain significant relative CRB gain by making use of attacked observations.

3.6 Numerical Results

3.6.1 Identification and Categorization of Attacked Sensors

In this subsection, we first test the performance of the identification and categorization technique described in Section 3.3 for some example cases. Specifically, we consider a sensor network consisting of $N = 10$ sensors, which is subject to 2 attacks. The 2 attacks control 30% and 20% of sensors respectively, and modify their observations with attack parameters $(\psi_{1,0}, \psi_{1,1}) = (0.2, 0.8)$ and $(\psi_{2,0}, \psi_{2,1}) = (0.7, 0.1)$. The parameter to be estimated is $\theta = 1$, the threshold of the quantizer is $\tau = 1$, $\Delta_0 = \Delta = 20\%$, and the additive noise obeys a standard normal distribution. Fig. 3.2 depicts the Monte Carlo approximation (200 times) of the ensemble average of the percentage of all mis-categorized sensors as a function of the number K of time samples at each sensor. As expected from our analysis, the curve in Fig. 3.2 clearly shows a diminishing trend of the average percentage of mis-categorized sensors and this percentage appears to decrease towards 0 as the number K of time samples at each sensor increases. This implies that the FC can identify and categorize the attacked sensors into different groups according to distinct types of attacks to achieve any desired level of accuracy for a sufficiently large K.

Next, we present numerical results in support of our theoretical analysis which illustrate the CRB performance of our proposed TQA for estimating θ. The numerical results also corroborate the superiority of the CRB performance of the proposed TQA when compared to that of the SEA. In the following subsections, we consider a sensor network consisting of $N = 100$ sensors. The additive noise obeys a standard normal distribu-

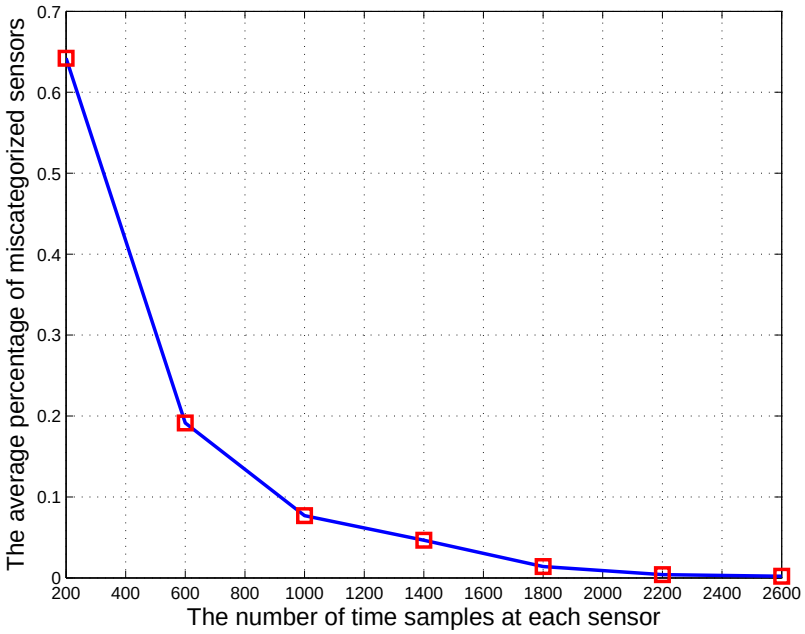

Figure 3.2: Identification and categorization of attacked sensors.

tion. The length of each time slot is fixed at $K_t = 10$, and the set of 801 thresholds is $\mathcal{Q} = \{0, -0.125, 0.125, -0.250, 0.250, ..., -5, 5\}$.

3.6.2 CRB Comparison between the TQA and the SEA

We firstly compare the CRB performance for estimating θ by using the proposed TQA to that obtained by using the SEA. Here, the parameter to be estimated is $\theta = 2$ and two different attacks ($P = 2$) are considered. The first attack tampers with 25% of the sensors with attack probabilities $\psi_{1,0} = 0.9$ and $\psi_{1,1} = 0.95$. The other attack takes over 20% of the sensors while using the attack probabilities $\psi_{2,0} = 0.15$ and $\psi_{2,1} = 0.2$. Fig. 3.3 depicts the CRB of estimating θ for the two approaches with a varying number of thresholds Q from 400 to 800. In the numerical results, for a given number of thresholds Q, each sensor observes QK_t time samples, and picks the first Q thresholds from the set of thresholds $\mathcal{Q}$ to

quantize the time samples in different time slots. It is seen that the CRB for both approaches decreases as Q, the number of thresholds, grows, since the number of time samples at each sensor increases. Moreover, it is easy to see that the relative CRB gain increases with Q. In addition, Fig. 3.3 illustrates that the TQA provides significant CRB performance gain when compared to the SEA, which implies that the set of thresholds leads to rank $(\boldsymbol{\Xi}_p) = 3$ for at least one p.

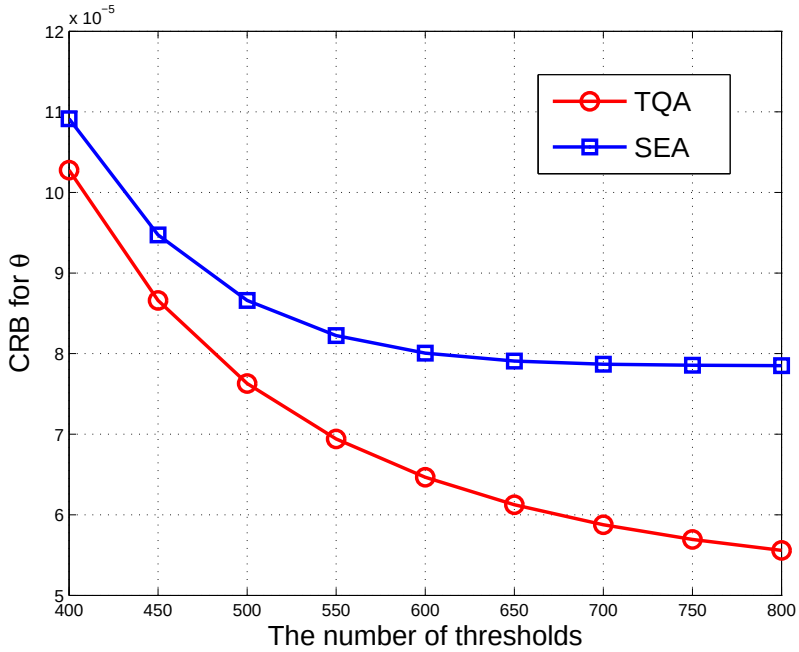

Figure 3.3: Comparison between the CRB for θ when employing either the TQA or the SEA.

3.6.3 Relative CRB Gain versus the Percentage of Attacked Sensors under One Attack

We now study the relationship between the relative CRB gain in (3.100) and the percentage of attacked sensors. To simplify the problem, we consider a scenario where $P = 1$, the parameter to be estimated is $\theta = 2$, and each sensor observes $Q = 801$ time slots of time sam-

ples which are quantized by employing the threshold set $\mathcal{Q}$. Fig. 3.4 illustrates the relative CRB gain as the percentage of the compromised sensors varies from 0% to 45% for different statistical attack matrices. The relative CRB gain for the distinct statistical attack matrices with $\psi_{1,0} = \psi_{1,1} = 0.05$, $\psi_{1,0} = \psi_{1,1} = 0.15$, $\psi_{1,0} = \psi_{1,1} = 0.25$, and $\psi_{1,0} = \psi_{1,1} = 0.35$ are marked with rectangles, circles, diamonds, and triangles respectively. It is seen that the relative CRB gain increases with the percentage of attacked sensors for all cases. As expected from the discussion after the proof of *Theorem 12*, Fig. 3.4 depicts that for a given percentage of attacked sensors, the larger the difference between $\psi_{1,0} = \psi_{1,1}$ and 0.5 (where $\psi_{1,0} + \psi_{1,1} = 1$), the larger the relative CRB gain that the corresponding TQA enjoys.

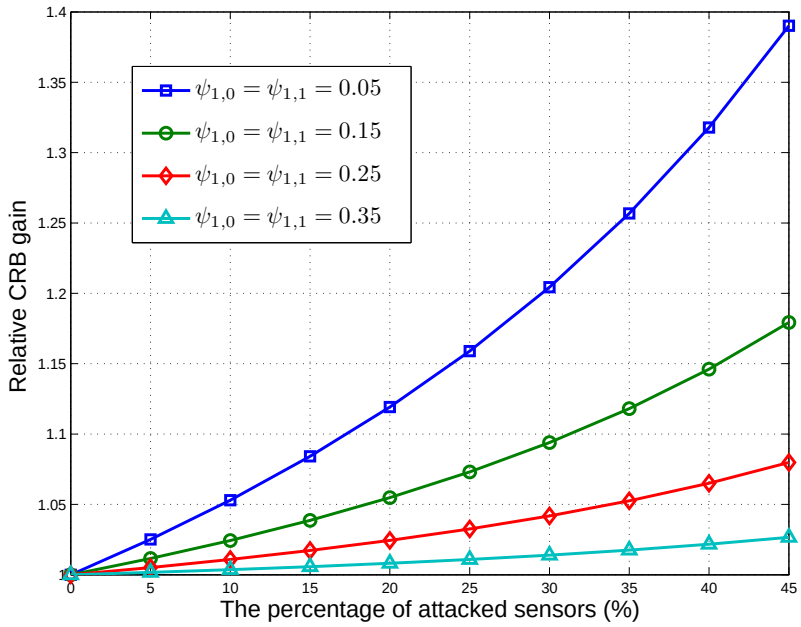

Figure 3.4: Relative CRB gain versus the percentage of attacked sensors.

3.7 Summary

In this chapter, we study the distributed estimation problem using quantized data in the presence of attacks. The sensor data modifications implemented by the adversaries are statistically characterized by a set of unknown probability transition matrices. We demonstrate that the FC is able to identify the attacked sensors and categorize these sensors into different subsets according to distinct types of attacks perfectly or with a very small percentage of misclassified sensors, as $K \to \infty$ or $N \to \infty$ respectively, provided that the set of unattacked sensors is larger than any of these subsets. In order to improve the estimation performance by utilizing the attacked sensors, a joint estimation of the statistical description of the attacks and the parameter to be estimated is considered. However, it is shown that the corresponding FIM is singular if the previously used data quantization approach is employed. Thus, it is not possible to accurately estimate the parameters using this approach with an estimate that would always become more and more accurate as we increase the number of observations. Aiming to overcome this, the TQA is proposed which divides the observation time interval at each sensor into several time slots and employs distinct thresholds to quantize the time samples in different time slots. If the number of time samples at each sensor is not less than 2, then it can be proven that the FIM for all unknown parameters in TQA is nonsingular which implies that the statistical properties of the attacks and the parameter to be estimated can be accurately estimated with a sufficiently large number of observations. We also derive necessary and sufficient condition under which the attacked observations can be taken advantage of to improve the asymptotic estimation performance. A notable fact is that for many cases, significant improvement in CRB performance for the parameter to be estimated can be attained by making use of attacked observations in our proposed fashion. However, for some

specific cases, using the attacked observations will not provide better asymptotic estimation performance. It is worth mentioning that both the theoretical analysis and numerical results illustrate that the improvement in CRB performance by utilizing attacked observations in our proposed fashion depends not only on the statistical description of the attacks and the parameter to be estimated, but also on the sets of thresholds of the quantizer, which motivates us to pursue the optimum quantizer design for distributed estimation in the presence of attacks in future work.

3.8 Appendix

3.8.1 Proof of Theorem 9

In order to satisfy the condition in *Lemma 9*, that is, $K > K^* = -2\ln 2/(\kappa\gamma^*)$ as shown in (3.52), we consider

$$\kappa \stackrel{\Delta}{=} -\frac{2\ln 2}{\gamma^*(K-1)}. \tag{3.101}$$

Hence, from (3.60), we can obtain that

$$\kappa \leq \min\left\{\Delta\Delta_0, \Delta^2\right\}/4. \tag{3.102}$$

Let $\mathfrak{C}_{\mathrm{C}}$ denote the complement of $\mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa)$, that is,

$$\begin{aligned}\mathfrak{C}_{\mathrm{C}} &\triangleq (\mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa))^{\mathrm{C}} \\ &= \{\mathcal{S} \,|\, \mathcal{S} \not\subset \mathcal{A}_p, \forall\, p\} \\ &\quad \cap \big\{\mathcal{S} \,|\, \forall\, \mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{S},\ \{\mathcal{S}_1, \mathcal{S}_2\} \subset \mathfrak{C}_0, \text{ and } \mathcal{S}_1 \cup \mathcal{S}_2 \notin \mathfrak{C}_0 \\ &\qquad \text{s.t. } \mathcal{P}_{\mathcal{S}_1} < \kappa \text{ or } \mathcal{P}_{\mathcal{S}_2} < \kappa\big\}. \end{aligned} \tag{3.103}$$

By *Lemma 9*, for any subset $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_1(\kappa)$ of sensors with percentage $\mathcal{P}_{\mathcal{J}} \geq \Delta$, we know that as $N \to \infty$, $P_{\text{error}} \to 0$, and hence

$$\Pr\left(\varpi\left(\tilde{\mathbf{u}}_{\mathcal{J}}\right) = 1 \,|\, \mathcal{J} \in \mathfrak{C}_1(\kappa)\right) = 1.$$

Therefore, for any subset $\mathcal{J} \subset \mathcal{S}_T$, if $\varpi\left(\tilde{\mathbf{u}}_{\mathcal{J}}\right) = 0$, then $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_{\mathrm{C}}$.

From the assumption in (3.10), we know that $P < 1/\Delta$ and

$$\kappa P < \Delta^2/(4\Delta) = \Delta/4 \tag{3.104}$$

by employing (3.102). Hence, if $\mathcal{J} \in \mathfrak{C}_0 \cup \mathfrak{C}_{\mathrm{C}}$ and $\mathcal{P}_{\mathcal{J}} \geq \Delta$, then by the definition of $\mathfrak{C}_0$ and $\mathfrak{C}_{\mathrm{C}}$ in (3.13) and (3.103), most of sensors in $\mathcal{J}$ must come from some unique $\mathcal{A}_l$ which constitute the core subset $\mathcal{E}_{\mathcal{J}}$ of $\mathcal{J}$, and the rest of sensors in $\mathcal{J}$ come from the other $\mathcal{A}_p$, $\forall p \neq l$ which constitute the minor part $\bar{\mathcal{E}}_{\mathcal{J}}$ of $\mathcal{J}$. To be specific, the core subset $\mathcal{E}_{\mathcal{J}}$ of $\mathcal{J}$ is defined as

$$\mathcal{E}_{\mathcal{J}} \triangleq \mathcal{J} \cap \mathcal{A}_l \tag{3.105}$$

for some unique l, which satisfies

$$\mathcal{P}_{\mathcal{E}_{\mathcal{J}}} \triangleq |\mathcal{E}_{\mathcal{J}}|/N > \mathcal{P}_{\mathcal{J}} - \kappa P > \mathcal{P}_{\mathcal{J}} - \Delta/4 \geq 3\Delta/4, \tag{3.106}$$

and the minor part $\bar{\mathcal{E}}_{\mathcal{J}}$ of $\mathcal{J}$ can be expressed as

$$\bar{\mathcal{E}}_{\mathcal{J}} \triangleq \mathcal{J} \backslash \mathcal{E}_{\mathcal{J}} = \mathop{\cup}_{p=0,\, p\neq l}^{P} \bar{\mathcal{E}}_{\mathcal{J}}^{p} \tag{3.107}$$

where $\bar{\mathcal{E}}_{\mathcal{J}}^{p} \triangleq (\mathcal{J} \backslash \mathcal{E}_{\mathcal{J}}) \cap \mathcal{A}_p = \mathcal{J} \cap \mathcal{A}_p$ for all $p \neq l$ and moreover

$$\mathcal{P}_{\bar{\mathcal{E}}_{\mathcal{J}}^{p}} \triangleq |\bar{\mathcal{E}}_{\mathcal{J}}^{p}|/N = |\mathcal{J} \cap \mathcal{A}_p|/N < \kappa. \tag{3.108}$$

The first inequality in (3.106) and the inequality in (3.108) are due to the definition of $\mathfrak{C}_0$ and $\mathfrak{C}_{\mathrm{C}}$ in (3.13) and (3.103). The second inequality in (3.106) is from (3.104). Furthermore, by (3.104), the percentage of the minor part $\bar{\mathcal{E}}_{\mathcal{J}}$ of $\mathcal{J}$ is upper bounded by

$$\begin{aligned} \mathcal{P}_{\bar{\mathcal{E}}_{\mathcal{J}}} &\triangleq \left|\bar{\mathcal{E}}_{\mathcal{J}}\right|/N = \left| \mathop{\cup}_{p=0,\, p\neq l}^{P} \bar{\mathcal{E}}_{\mathcal{J}}^{p} \right|/N \\ &= \sum_{p=1,\, p\neq l}^{P} \mathcal{P}_{\bar{\mathcal{E}}_{\mathcal{J}}^{p}} < \kappa P < \Delta/4. \end{aligned} \tag{3.109}$$

By checking all possible such subsets $\mathcal{J}$ with $\mathcal{P}_{\mathcal{J}} \geq \Delta$ and $\varpi\left(\tilde{\mathbf{u}}_{\mathcal{J}}\right) = 0$, one can determine the collection of the largest subsets $\{\tilde{\mathcal{A}}_p\}_{p=0}^{\tilde{P}} \subset \mathfrak{C}_0 \cup \mathfrak{C}_{\mathrm{C}}$ which satisfies

$$\begin{cases} \varpi\left(\tilde{\mathbf{u}}_{\tilde{\mathcal{A}}_p}\right) = 0,\ \tilde{\mathcal{P}}_p \geq \Delta \quad \forall p = 0, .., \tilde{P} \\ |\tilde{\mathcal{A}}_i \cap \tilde{\mathcal{A}}_j|/N < 2\kappa P < 2\kappa/\Delta < \Delta/2, \quad \forall i \neq j \quad , \\ \mathop{\cup}_{p=0}^{\tilde{P}} \tilde{\mathcal{A}}_p = \mathop{\cup}_{p=0}^{P} \mathcal{A}_p \end{cases} \tag{3.110}$$

where the collection of the largest subsets implies that one is unable to replace any $\tilde{\mathcal{A}}_l$ with $\tilde{\mathcal{A}}_l^*$ such that $|\tilde{\mathcal{A}}_l^*| > |\tilde{\mathcal{A}}_l|$ and the new collection of subsets still meets (3.110), and there is not any other collection of subsets $\{\tilde{\mathcal{A}}_l^*\}_{l=0}^{L}$ with $L < \tilde{P}$ which also satisfies (3.110).

Since $\tilde{\mathcal{A}}_p \in \mathfrak{C}_0 \cup \mathfrak{C}_{\mathrm{C}}$ and $\tilde{\mathcal{P}}_p \geq \Delta$ for all p as defined by (3.110), we can replace $\mathcal{J}$ with any $\tilde{\mathcal{A}}_p$ in (3.106), (3.107), (3.108), and (3.109), and the results still hold. The proof of *Theorem 9* can be completed by proving the following bullets.

- *The number of groups in* $\{\tilde{\mathcal{A}}_p\}_{p=0}^{\tilde{P}}$ *equals to the number of attacks, i.e.,* $\tilde{P} = P$.

Suppose $\tilde{P} < P$. Replacing $\mathcal{J}$ with $\tilde{\mathcal{A}}_l$ in (3.106), we know that for each $l = 0, 1, ..., \tilde{P}$, $\tilde{\mathcal{A}}_l$ has a core subset $\mathcal{E}_{\tilde{\mathcal{A}}_l} \subset \mathcal{A}_{\tilde{l}}$ for some unique $\tilde{l}$. Thus, there is at least one $\mathcal{A}_p$ which doesn't contain any core subset of $\tilde{\mathcal{A}}_l$, $\forall l = 0, 1, ..., \tilde{P}$. By employing (3.104) and replacing $\mathcal{J}$ with $\tilde{\mathcal{A}}_l$ in (3.108), we can obtain

$$\begin{aligned}\left|\bigcup_{l=0}^{\tilde{P}}(\tilde{\mathcal{A}}_l \cap \mathcal{A}_p)\right| /N &\leq \sum_{l=0}^{\tilde{P}} \left|\tilde{\mathcal{A}}_l \cap \mathcal{A}_p\right| /N = \sum_{l=0}^{\tilde{P}} \mathcal{P}_{\tilde{\mathcal{E}}^p_{\tilde{\mathcal{A}}_l}} \\ &< \kappa(\tilde{P}+1) \leq \kappa P < \Delta/4 < \mathcal{P}_p. \end{aligned} \tag{3.111}$$

Thus,

$$\mathcal{A}_p \not\subset \bigcup_{l=0}^{\tilde{P}} \tilde{\mathcal{A}}_l, \tag{3.112}$$

which contradicts the third equation in (3.110). Therefore, $\tilde{P} \geq P$.

On the other hand, noting that $\{\mathcal{A}_p\}_{p=0}^{P} \subset \mathfrak{C}_0$ are disjoint and $\mathcal{P}_p \geq \Delta$, $\forall p$ as shown in (3.10), $\{\mathcal{A}_p\}_{p=0}^{P}$ is a collection of subsets satisfying (3.110). Since $\{\tilde{\mathcal{A}}_p\}_{p=0}^{\tilde{P}}$ is the collection of the largest subsets, we know $\tilde{P} \leq P$. It follows that $\tilde{P} = P$, since we have proven $\tilde{P} \geq P$.

- *The core subsets of different groups in* $\{\tilde{\mathcal{A}}_l\}_{l=0}^{\tilde{P}}$ *are tampered with by distinct types of attacks.*

Suppose there are two core subsets controlled by the same attack type, then there exists some $\mathcal{A}_p$ which contains at least two core subsets of different groups in $\{\tilde{\mathcal{A}}_l\}_{l=0}^{\tilde{P}}$. Since $\tilde{P} = P$ and every $\tilde{\mathcal{A}}_l$ only has one core subset which comes from some unique $\mathcal{A}_p$, there is at least one $\mathcal{A}_p$ which doesn't contain any core subset of $\tilde{\mathcal{A}}_l$, $\forall l = 0, 1, ..., \tilde{P}$. By the same argument in (3.111) and (3.112), we reach a contradiction. Therefore, for different groups $\tilde{\mathcal{A}}_l$ and $\tilde{\mathcal{A}}_m$, $\forall l \neq m$, their core subsets $\mathcal{E}_{\tilde{\mathcal{A}}_l}$ and $\mathcal{E}_{\tilde{\mathcal{A}}_m}$ satisfy that if for some p, $\mathcal{E}_{\tilde{\mathcal{A}}_l} \subset \mathcal{A}_p$, then

$$\mathcal{E}_{\tilde{\mathcal{A}}_m} \not\subset \mathcal{A}_p. \tag{3.113}$$

As a result, without loss of generality, we renumber the indices of $\{\tilde{\mathcal{A}}_l\}_{l=0}^{\tilde{P}}$ to satisfy that the core subset $\mathcal{E}_{\tilde{\mathcal{A}}_l}$ of $\tilde{\mathcal{A}}_l$ is contained in $\mathcal{A}_l$ for all $l = 0, 1, ..., P$ in the following part.

- *For all* $m = 0, 1, ..., P$, $\tilde{\mathcal{P}}_0 > \tilde{\mathcal{P}}_m$ *and* $0 \leq |\tilde{\mathcal{P}}_m - \mathcal{P}_m| = \mathcal{P}_m^* < \delta$.

Since $\mathcal{E}_{\tilde{\mathcal{A}}_m} \subset \mathcal{A}_m$, by replacing $\mathcal{J}$ with $\tilde{\mathcal{A}}_m$ in (3.109), an upper bound on $|\tilde{\mathcal{A}}_m \backslash \mathcal{A}_m|/N$ is given by

$$\left|\tilde{\mathcal{A}}_m \backslash \mathcal{A}_m\right| / N = \left|\tilde{\mathcal{A}}_m \backslash \mathcal{E}_{\tilde{\mathcal{A}}_m}\right| / N = \mathcal{P}_{\bar{\mathcal{E}}_{\tilde{\mathcal{A}}_m}} < \kappa P, \tag{3.114}$$

For all $p \neq m$, by (3.113), we know that $\mathcal{E}_{\tilde{\mathcal{A}}_p} \not\subset \mathcal{A}_m$, and hence by replacing $\mathcal{J}$ with $\tilde{\mathcal{A}}_p$ in (3.108), we can obtain

$$\left|\tilde{\mathcal{A}}_p \cap \mathcal{A}_m\right| / N = \left|\bar{\mathcal{E}}^m_{\tilde{\mathcal{A}}_p}\right| / N = \mathcal{P}_{\bar{\mathcal{E}}^m_{\tilde{\mathcal{A}}_p}} < \kappa, \tag{3.115}$$

which yields

$$\left| \bigcup_{p=0,\, p\neq m}^{\tilde{P}} (\tilde{\mathcal{A}}_p \cap \mathcal{A}_m) \right| /N$$
$$\leq \sum_{p=0,\, p\neq m}^{\tilde{P}} \left| \tilde{\mathcal{A}}_p \cap \mathcal{A}_m \right| /N < \kappa P. \tag{3.116}$$

Since $\mathcal{S}_T = \bigcup_{p=0}^{P} \mathcal{A}_p = \bigcup_{p=0}^{\tilde{P}} \tilde{\mathcal{A}}_p$, we can obtain $\tilde{\mathcal{A}}_m^{\mathrm{C}} = (\bigcup_{p=0}^{\tilde{P}} \tilde{\mathcal{A}}_p) \backslash \tilde{\mathcal{A}}_m \subset \bigcup_{p=0, p\neq m}^{\tilde{P}} \tilde{\mathcal{A}}_p$, and hence, by employing (3.116),

$$\left| \mathcal{A}_m \backslash \tilde{\mathcal{A}}_m \right| /N = \left| \mathcal{A}_m \cap \tilde{\mathcal{A}}_m^{\mathrm{C}} \right| /N$$
$$\leq \left| \bigcup_{p=0,\, p\neq m}^{\tilde{P}} (\tilde{\mathcal{A}}_p \cap \mathcal{A}_m) \right| /N < \kappa P. \tag{3.117}$$

Thus, (3.114) and (3.117) yield

$$\tilde{\mathcal{P}}_0 = |\tilde{\mathcal{A}}_0|/N \geq |\mathcal{A}_0 \cap \tilde{\mathcal{A}}_0|/N$$
$$= \left(|\mathcal{A}_0| - |\mathcal{A}_0 \backslash \tilde{\mathcal{A}}_0| \right) /N > \mathcal{P}_0 - \kappa P, \text{ and} \tag{3.118}$$

$$\tilde{\mathcal{P}}_m = |\tilde{\mathcal{A}}_m|/N \leq |\mathcal{A}_m \cup \tilde{\mathcal{A}}_m|/N$$
$$= \left(|\mathcal{A}_m| + |\tilde{\mathcal{A}}_m \backslash \mathcal{A}_m| \right) /N < \mathcal{P}_m + \kappa P, \tag{3.119}$$

and hence, by employing (3.102) and noticing that $P < 1/\Delta$,

$$\tilde{\mathcal{P}}_0 - \tilde{\mathcal{P}}_m > \mathcal{P}_0 - \mathcal{P}_m - 2\kappa P > \Delta_0 - 2\kappa P$$
$$> \Delta_0 - 2\kappa/\Delta > \Delta_0/2 > 0. \tag{3.120}$$

Furthermore, we can obtain from (3.114) and (3.117) that

$$\begin{aligned}\mathcal{P}_m^* &= \left|\left(\tilde{\mathcal{A}}_m \backslash \mathcal{A}_m\right) \cup \left(\mathcal{A}_m \backslash \tilde{\mathcal{A}}_m\right)\right| / N \\ &\leq \left|\tilde{\mathcal{A}}_m \backslash \mathcal{A}_m\right| / N + \left|\mathcal{A}_m \backslash \tilde{\mathcal{A}}_m\right| / N < 2\kappa P. \end{aligned} \tag{3.121}$$

Finally, we conclude the proof by noting that

$$\begin{aligned} 0 \leq |\tilde{\mathcal{P}}_m - \mathcal{P}_m| &= \left||\tilde{\mathcal{A}}_m| - |\mathcal{A}_m|\right| / N \\ &\leq \left(|\tilde{\mathcal{A}}_m \cup \mathcal{A}_m| - |\tilde{\mathcal{A}}_m \cap \mathcal{A}_m|\right) / N = \mathcal{P}_m^* < 2\kappa P \\ &< 2\left(-\frac{2\ln 2}{(K-1)\gamma^*}\right)\frac{1}{\Delta} = \delta. \end{aligned} \tag{3.122}$$

Chapter 4

Functional Forms of Optimum Spoofing Attacks for Vector Parameter Estimation in Quantized Sensor Networks

4.1 Introduction

Recent developments in sensor technology have encouraged a large number of applications of sensor networks for parameter estimation ranging from inexpensive commercial systems to complex military and homeland defense surveillance systems [49]. Typically, large-scale sensor networks are comprised of low-cost and spatially distributed sensor nodes with limited battery power and low computing capacity, which makes the system vulnerable to cyberattacks by adversaries. This has led to great interest in studying the vulnerability

of sensor networks in various applications and from different perspectives, see [61,69–73] and the references therein. Due to the dominance of digital technology, a great deal of attention has focused on parameter estimation using quantized data [50, 51, 53, 54, 56]. The sequel considers the problem of estimating a vector parameter by using quantized data collected from a distributed sensor network under the assumption that the measurements from several subsets of sensors have been falsified by spoofing attacks, a topic that has received virtually no attention to date. To be specific, the spoofing attacks maliciously modify the temporal analog measurements of the phenomenon acquired at the subset of attacked sensors.

4.1.1 System and Adversary Models

Consider a distributed sensor network $\mathcal{S}_N$ consisting of N spatially distributed sensors, each making some measurements of a particular phenomenon. We assume that the j-th sensor acquires K_j measurements, and we denote the before-attack measurement of the j-th sensor at time instant k by x_{jk} which follows a pdf $f_j\left(x_{jk}\,|\boldsymbol{\theta}\right)$ depending on an unknown deterministic vector parameter $\boldsymbol{\theta}$ with dimension $D_{\boldsymbol{\theta}}$ that is desired to estimate for the measurements. For simplicity, we assume that the measurements $\{x_{jk}\}$ from the same sensor j but for different times ($k \neq k'$) are statistically independent and identically distributed (i.i.d.), while the measurements from different sensors are statistically independent but not necessarily identically distributed.

The adversaries alter the physical phenomenon as in Fig. 4.1, thus tampering with the measurements at a subset of sensors in the sensor network, hoping to undermine the estimation performance of the system. Let $\mathcal{V} \subset \mathcal{S}_N$ denote the set of sensors undergoing spoofing attacks while the set $\mathcal{U} \triangleq \mathcal{S}_N \backslash \mathcal{V}$ represents the set of unattacked sensors. A generalized mathematical model of spoofing attacks which maliciously modify the distribution of the analog

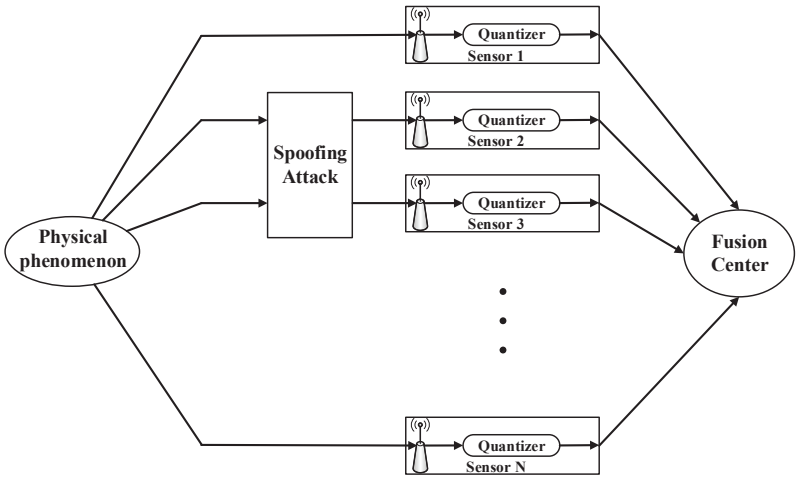

Figure 4.1: Distributed estimation system in the presence of spoofing attacks.

observations of the physical phenomenon at the attacked sensors is considered employing general probability density functions $\{f_j\}$ and $\{g_j\}$ which depend of the desired and attack vector parameters. To conform to previous work, the functional forms of the attacks, thus $\{f_j\}$ and $\{g_j\}$, are assumed known to the attacked system but the desired and attack vector parameters are not. Thus, the after-attack version $\tilde{x}_{jk}$ of x_{jk} obeys the statistical model that[1] $\{\tilde{x}_{jk}\}$ is independent over j and i.i.d. over k, moreover,

$$\tilde{x}_{jk} \sim \begin{cases} f_j\left(x_{jk}\,|\boldsymbol{\theta}\right), & \text{if } j \in \mathcal{U} \\ g_j\left(x_{jk}\,\middle|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)}\right), & \text{if } j \in \mathcal{V} \end{cases}, \tag{4.1}$$

where if $j \in \mathcal{V}$, the after-attack pdf $g_j(x_{jk}|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)})$ is parameterized by the desired vector parameter $\boldsymbol{\theta}$ and the attack vector parameter $\boldsymbol{\xi}^{(j)}$. It is worth mentioning that the notation[2] $g_j(x|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)})$ does not imply that the after-attack pdf $g_j(x|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)})$ of the measurements at the

[1]The notations $\tilde{x}_{jk}$ and $\tilde{u}_{jk}$ denote the after-attack analog measurements and the corresponding quantized measurements.

[2]Since the temporal analog measurments at the same sensor are assumed to be i.i.d, we use $f_j(x|\boldsymbol{\theta})$ and $g_j(x|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)})$ instead of $f_j(x_{jk}|\boldsymbol{\theta})$ and $g_j(x_{jk}|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)})$ in the following for the sake of notational simplicity.

j-th sensor has to depend on $\boldsymbol{\theta}$. For example, the adversaries can intercept the signal from the physical phenomenon and generate a new signal using some different pdf solely based on its attack vector parameter. A detailed example of a practical attack of the type described in (1) is provided in Section 4.2.

The set $\mathcal{V}$ of attacked sensors can be divided into disjoint subsets $\{\mathcal{A}_p\}_{p=1}^{P}$ in terms of distinct attack vector parameters $\{\boldsymbol{\xi}^{(j)}\}$ such that

$$\mathcal{V} = \bigcup_{p=1}^{P} \mathcal{A}_p, \quad \text{and} \quad \mathcal{A}_l \cap \mathcal{A}_m = \emptyset, \ \forall l \neq m, \tag{4.2}$$

where the attacked sensors in the subset $\mathcal{A}_p$ are known by the system under attack to employ an identical attack vector parameter $\boldsymbol{\tau}^{(p)}$ with dimension D_p so that $\boldsymbol{\xi}^{(j)} = \boldsymbol{\tau}^{(p)}$, $\forall j \in \mathcal{A}_p$. The identical attack vectors are possibly due to the sensors in $\mathcal{A}_p$ being attacked by the same attacker. For the sake of notational simplicity, we use $\mathcal{A}_0$ to denote the set $\mathcal{U}$ of unattacked sensors and let N_p denote the number of sensors contained in $\mathcal{A}_p$ for all $p = 0, 1, ..., P$.

Due to the communications employed, each sensor is restricted to convert analog measurements to digital data before transmitting this data to the fusion center (FC) as shown in Fig. 4.1. At the j-th sensor, each after-attack measurement $\tilde{x}_{jk}$ is quantized to $\tilde{u}_{jk}$ by using a R_j-level quantizer with quantization regions $\{I_j^{(r)}\}_{r=1}^{R_j}$, that is,

$$\tilde{u}_{jk} = \sum_{r=1}^{R_j} r \mathbb{1}\left\{\tilde{x}_{jk} \in I_j^{(r)}\right\}. \tag{4.3}$$

We adopt this general quantization model due to the fact that optimized quantization regions $\{I_j^{(r)}\}_{r=1}^{R_j}$ for different sensors can be very different, since the measurements from different sensors do not necessarily obey an identical pdf [51,74]. We assume that the quantizer design

$\{I_j^{(r)}\}_{r=1}^{R_j}$ for each sensor is predefined and known to the FC, but not the attacker.

Let $\boldsymbol{\Theta}$ denote a vector containing the unknown parameter $\boldsymbol{\theta}$ along with all the unknown attack vector parameters which parameterize the spoofing attacks in the sensor network

$$\boldsymbol{\Theta} \triangleq \left[\boldsymbol{\theta}^T, \left(\boldsymbol{\tau}^{(1)}\right)^T, ..., \left(\boldsymbol{\tau}^{(P)}\right)^T\right]^T. \tag{4.4}$$

For the sake of notational simplicity in the following parts, we use $p_j^{(r)}$ to denote the after-attack probability mass function (pmf) of the quantized measurement $\tilde{u}_{jk}$ evaluated at $\tilde{u}_{jk} = r$, that is,

$$\begin{aligned} p_j^{(r)} &\triangleq \Pr\left(\tilde{u}_{jk} = r \,|\boldsymbol{\Theta}\right) \\ &= \begin{cases} \int_{I_j^{(r)}} f_j\left(x\,|\boldsymbol{\theta}\right) dx, & \forall j \in \mathcal{A}_0 \\ \int_{I_j^{(r)}} g_j\left(x\,|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}\right) dx, & \forall j \in \mathcal{A}_p,\ \forall p \geq 1 \end{cases}. \end{aligned} \tag{4.5}$$

For simplicity, the communication channel between the FC and each sensor is assumed ideal, and hence the FC is able to accurately receive what was transmitted from both the unattacked and attacked sensors. After receiving the quantized data from all sensors, the FC attempts to make an estimate of the desired vector parameter without knowing which sensors have been tampered nor the attack parameters used by the attackers.

4.1.2 Performance Metric

It is of considerable interest to investigate the performance of spoofing attacks, and mathematically characterize the class of the most devastating spoofing attacks under the assumption that the adversaries have no information about what computations the FC is using. Hence, this chapter develops guarantees for the attacker's performance that are independent

of the computations performed at the FC. It is clear that if the FC has the information about the attack groupings for the sensors, i.e., $\{\mathcal{A}_p\}$, it can use this information to improve estimation performance over the case where this information is not employed, since the FC can always do better in estimating the desired vector parameter with extra knowledge. Therefore, for the case of spoofing attacks employing some specific $\{f_j(x|\boldsymbol{\theta})\}$ and $\{g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)})\}$, the case where the compromised sensors are well identified and categorized into P different groups according to distinct types of spoofing attacks by the FC corresponds to the case where the FC has the best chance to combat the spoofing attacks. In other words, the best possible estimation performance (smallest error) under this case provides a lower bound on the estimation performance for any other cases, which implies that the corresponding spoofing attack performance under this case provides a guaranteed attack performance in degrading the estimation performance no matter what computations the FC is using. The recent work in [73] has shown that for some classes of spoofing attacks, with a sufficient number of observations, the FC is able to perfectly identify the set of unattacked sensors and categorize the attacked sensors into different groups according to distinct types of spoofing attacks. For these reasons, we adopt the following definition of the optimal guaranteed degradation spoofing attacks in this chapter.

Definition 1 *Consider attacks employing* $\{f_j(x|\boldsymbol{\theta})\}$ *and* $\{g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)})\}$. *The optimal guaranteed degradation spoofing attack (OGDSA) maximizes the degradation of the Cramer-Rao Bound (CRB) for the vector parameter of interest at the FC when the attacked sensors are well identified and categorized according to distinct types of spoofing attacks by the FC.*

The estimation performance for a vector parameter in a distributed sensor network can be expressed using an error correlation matrix. However, in most cases, a closed form

expression for the error correlation matrix is intractable. Thus the CRB, an asymptotically achievable lower bound on the error correlation matrix, is employed in *Definition 1*. It is worth mentioning that the optimal guaranteed degradation spoofing attack defined in *Definition 1* achieves the classical definition of attack optimality (largest CRB) for the scenario where the FC has the best chance to combat the spoofing attacks. It might not be the classically optimal spoofing attack for the scenario where the FC is unable to determine which sensors are attacked, or to classify sensors into groups of distinct types of spoofing attacks. However, the OGDSAs defined in *Definition 1* can provide a guarantee that the actual degradation in the CRB must exceed some critical value no matter what computations the estimation system employs. This guarantee makes OGDSA an excellent spoofing attack from the adversaries' point of view.

4.1.3 Summary of Results and Main Contributions

Unlike previous work, a generalized attack model is employed which manipulates the data using transformations with arbitrary functional forms determined by some attack parameters whose values are unknown to the attacked system. For the first time, necessary and sufficient conditions are provided under which these transformations provide an OGDSA. These conditions imply that, for an OGDSA, either the Fisher Information Matrix (FIM) under the conditions of *Definition 1* for jointly estimating the desired and attack parameters is singular or that the attacked system is unable to improve the CRB under the conditions of *Definition 1* for the desired vector parameter through this joint estimation even though the joint FIM is nonsingular. It is shown that it is always possible to construct an OGDSA by properly employing a sufficiently large dimension attack vector parameter relative to the number of quantization levels employed, which was not observed previously. It is shown

that a spoofing attack can render the attacked measurements useless in terms of reducing the CRB under the conditions of *Definition 1* for estimating the desired vector parameter if and only if it is an OGDSA. For a class of OGDSAs, a computationally efficient heuristic is developed for the joint identification of the attacked sensors and estimation of the desired vector parameter which, in numerical tests for a sufficiently large number of observations, achieves a genie bound that knows all the groups of identically attacked sensors.

4.1.4 Related Work

In recent years, the estimation problem under different attacks has seen great interest in various engineering applications, see [9, 61, 69–73, 75–77] and the references therein. Rather than the man-in-the-middle attacks which falsify the data transmitted from the sensors to the FC [61, 72], we are primarily interested in spoofing attacks in the distributed sensor estimation system in this chapter, which maliciously modify the measurements of the physical phenomenon at a subset of sensors, see Fig. 4.1.

As previously introduced, spoofing attacks have been widely considered in wireless sensor networks, smart grids, radar systems and sonar systems [9, 69–71, 73, 75–78]. Each of these recent works takes one specific type of spoofing attack into account, and investigates the attack or the estimation performance. In this chapter, we don't focus on one specific type of spoofing attack. Instead, we consider a generalized attack model which can describe the different kinds of spoofing attacks employed in all recent work, and moreover, we make use of this generalized model to provide uniform tools to test if a spoofing attack is optimal in our defined sense. Both CRB-based analysis and a finite sample sized estimation approach are presented for a class of OGDSAs provided in this chapter.

Another difference between our work and other recent work on spoofing attacks in [9,

69–71, 75–78] is that we consider a distributed sensor estimation system which employs a quantization using a finite alphabet at each sensor which is typically the case in practice. Interestingly, we show that the quantization limits the capacity of the estimation system to combat the spoofing attacks. In partcular, it is shown that the adversaries can launch a class of quantization induced OGDSAs which are easily constructed in practice.

4.1.5 Notation

Throughout this chapter, bold upper case letters and bold lower case letters are used to denote matrices and column vectors respectively. The symbol $\mathbb{1}(\cdot)$ stands for the indicator function. Let $[\mathbf{A}]_{i,j}$ denote the element in the i-th row and j-th column of the matrix $\mathbf{A}$, and $\mathscr{R}(\mathbf{A})$ represents the range space of $\mathbf{A}$. $\mathbf{A} \succ 0$ and $\mathbf{A} \succeq 0$ imply that the matrix is positive definite and positive semidefinite respectively. To avoid cumbersome sub-matrix and sub-vector expressions in this chapter, we introduce the following notation. The notation $[\mathbf{A}]_{\mathcal{S},:}$ stands for the sub-matrix of $\mathbf{A}$ which consists of the elements with row indices in the set $\mathcal{S}$, and $[\mathbf{A}]_{1:N}$ represents the N-by-N leading principle minor of $\mathbf{A}$. The i-th element of the vector $\mathbf{v}$ is denoted by v_i, and $[\mathbf{v}]_{\mathcal{S}}$ represents the sub-vector of $\mathbf{v}$ which only contains the elements with indices in the set $\mathcal{S}$. The symbols $\nabla_{\mathbf{v}} f$ and $\nabla^2_{\mathbf{v}} f$ respectively signify the gradient and Hessian of f with respect to $\mathbf{v}$. Finally, the expectation and rank operators are denoted by $\mathbb{E}(\cdot)$ and $\mathrm{rank}(\cdot)$ respectively.

4.2 Illustrative Example of a Practical Spoofing Attack

Spoofing attacks on sensor networks can occur in various engineering applications. For instance, spoofing attacks have been described for the localization problem in wireless sensor

networks, see [69, 70] and the references therein. *Table I* in [69] provides a summary of different types of spoofing attack threats for the localization problem. Radar and sonar systems also suffer from spoofing attack threats in practice. As one example of a spoofing attack technique, the application of an electronic countermeasure (ECM), which is designed to jam or deceive the radar or sonar system, can critically degrade the detection and estimation performance of the system [79]. One popular technique for the implementation of ECM employs digital radio frequency memory (DRFM) in radar systems to manipulate the received signal and retransmit it back to confuse the victim radar system. DRFM can mislead the estimation of the range of the target by altering the delay in transmission of pulses, and fool the system into incorrectly estimating the velocity of the target by introducing a Doppler shift in the retransmitted signal [75]. An example of a spoofing attack created by nature is environmental variation in shallow water sonar systems. According to waveguide-invariant theory [80], the environmental variation, such as sound-speed or water-depth perturbations, essentially introduces an apparent shift in the position of the target of interest when the data is processed by matched field processing [78, 81]. Hence these environmental variations can be treated as spoofing attacks which falsify the physical model of the received signal in sonar systems. More recently, the data-injection attack considered in smart grids is another typical example of a spoofing attack, see [9, 71, 76, 77] and the references therein.

In order to motivate the mathematical description of spoofing attacks, we consider a spoofing attack utilizing a DRFM in a radar system as an example, which stores the received signal and strategically retransmits it back by introducing an additional delay to mislead the estimation of the range of the target. In the absence of spoofing attacks, the simplified signal model of the measurement x_j at the j-th receiver at time instant t_j, which ignores the

Doppler shift, can be expressed as

$$x_j = \sqrt{E_j} a_j s\left(t_j - \theta_j\right) + n_j, \tag{4.6}$$

where $s(\cdot)$, E_j and a_j respectively represent the transmitted signal, the transmitted energy, and the reflection coefficient. The time delay is denoted by θ_j which is the parameter to be estimated. We assume that the clutter-plus-noise n_j is a zero-mean Gaussian random variable with known variance σ_j^2 while all other quantities in (4.6) are deterministic. As a result, the probability density function (pdf) $f_j(x_j|\theta_j)$ of x_j in the absence of spoofing attacks is given by[3]

$$f_j\left(x_j\,|\theta_j\right) = \mathcal{N}\left(\sqrt{E_j} a_j s\left(t_j - \theta_j\right), \sigma_j^2\right). \tag{4.7}$$

In the presence of a spoofing attack, the after-attack measurement $\tilde{x}_j$ can be described as

$$\tilde{x}_j = \sqrt{E_j} a_j s\left(t_j - \theta_j - \xi_j\right) + n_j, \tag{4.8}$$

where ξ_j is the delay introduced by the DRFM. Therefore, we can obtain the corresponding after-attack pdf of $\tilde{x}_j$

$$\begin{aligned} g_j\left(\tilde{x}_j\,|\theta_j, \xi_j\right) &= \mathcal{N}\left(\sqrt{E_j} a_j s\left(t_j - \theta_j - \xi_j\right), \sigma_j^2\right) \\ &= f_j\left(\tilde{x}_j\,|\theta_j + \xi_j\right). \end{aligned} \tag{4.9}$$

In this example, the after-attack pdf $g_j(\tilde{x}_j|\theta_j, \xi_j)$ and the before-attack pdf $f_j(x_j|\theta_j)$ are in the same family as shown in (4.9), i.e., the family of Gaussian distributions with the same

[3] $\mathcal{N}(a, b)$ denotes a Gaussian pdf with mean a and variance b.

variance σ_j^2. While this may not always be true, the after-attack pdf is generally not only parameterized by the desired parameter θ_j but also by an unknown attack parameter ξ_j.

Motivated by this example and other popular spoofing attack examples, such as those in [9, 69–71, 73, 75–78], the essential impact of a spoofing attack, which maliciously modifies the measurements at the j-th sensor in a manner similar to (4.8), can be described as a mapping which maps the before-attack pdf $f_j(x|\boldsymbol{\theta})$ of the measurements at the j-th sensor to an after-attack pdf $g_j(x|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)})$, where $\boldsymbol{\theta}$ and $\boldsymbol{\xi}^{(j)}$ account for the desired vector parameter and the attack vector parameter at the j-th sensor which represents those deterministic unknowns which can determine the after-attack pdf.

4.3 The Optimality of Spoofing Attacks

In this section, we pursue the explicit characterization of the optimal spoofing attack defined in *Definition 1*. The adversaries can attempt to maximize CRB for $\boldsymbol{\theta}$ in the positive semidefinite sense to achieve an optimal spoofing attack as per *Definition 1*. The FIM $\mathbf{J}_{\boldsymbol{\Theta}}$ for $\boldsymbol{\Theta}$ is defined as [82]

$$[\mathbf{J}_{\boldsymbol{\Theta}}]_{l,m} \triangleq -\mathbb{E}\left\{\frac{\partial^2 L(\boldsymbol{\Theta})}{\partial\Theta_l \partial\Theta_m}\right\}, \tag{4.10}$$

where $L(\boldsymbol{\Theta})$ denotes the log-likelihood function.

When the attacked senors are well identified and categorized into different groups according to distinct types of spoofing attacks, the log-likelihood function $L(\boldsymbol{\Theta})$ in (4.10) evaluated at

$$\tilde{\mathbf{u}} \triangleq [\tilde{u}_{11}, \tilde{u}_{12}, ..., \tilde{u}_{1K_1}, \tilde{u}_{21}, ..., \tilde{u}_{NK_N}]^T = \mathbf{r}$$

can be expressed as

$$L(\boldsymbol{\Theta}) = \ln \Pr(\tilde{\mathbf{u}} = \mathbf{r} \,|\boldsymbol{\Theta}) = \sum_{p=0}^{P} \sum_{j \in \mathcal{A}_p} \sum_{k=1}^{K_j} \sum_{r=1}^{R_j} \mathbb{1}\{r_{jk} = r\} \ln p_j^{(r)} \tag{4.11}$$

by employing (4.5).

By substituting the expression of the log-likelihood function $L(\boldsymbol{\Theta})$ in (4.11) into the definition of the FIM in (4.10), it can be shown that the FIM $\mathbf{J}_{\boldsymbol{\Theta}}$ for $\boldsymbol{\Theta}$ takes the form

$$\mathbf{J}_{\boldsymbol{\Theta}} \triangleq \begin{bmatrix} \mathbf{J}_{\boldsymbol{\theta}} & \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_P \\ \mathbf{B}_1^T & \mathbf{J}_{\boldsymbol{\tau}^{(1)}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{B}_2^T & \mathbf{0} & \mathbf{J}_{\boldsymbol{\tau}^{(2)}} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{B}_P^T & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{J}_{\boldsymbol{\tau}^{(P)}} \end{bmatrix} \tag{4.12}$$

where $\mathbf{J}_{\boldsymbol{\theta}} \in \mathbb{R}^{D_{\boldsymbol{\theta}} \times D_{\boldsymbol{\theta}}}$, $\mathbf{J}_{\boldsymbol{\tau}^{(p)}} \in \mathbb{R}^{D_p \times D_p}$, and $\mathbf{B}_p \in \mathbb{R}^{D_{\boldsymbol{\theta}} \times D_p}$ for all $p = 1, 2, ..., P$. Moreover, following from (4.4) and (4.10), we can obtain that $\forall p$

$$\mathbf{J}_{\boldsymbol{\tau}^{(p)}} = \sum_{j \in \mathcal{A}_p} \sum_{r=1}^{R_j} \frac{K_j}{p_j^{(r)}} \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} \left[\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} \right]^T, \tag{4.13}$$

$$\mathbf{B}_p = \sum_{j \in \mathcal{A}_p} \sum_{r=1}^{R_j} \frac{K_j}{p_j^{(r)}} \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\theta}} \left[\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} \right]^T, \tag{4.14}$$

and

$$\mathbf{J}_{\boldsymbol{\theta}} = \sum_{p=0}^{P} \mathbf{J}_{\mathcal{A}_p}, \tag{4.15}$$

where $\mathbf{J}_{\mathcal{A}_p}$, which is contributed from the measurements observed at the sensors in $\mathcal{A}_p$, is defined as

$$\mathbf{J}_{\mathcal{A}_p} \triangleq \sum_{j\in\mathcal{A}_p}\sum_{r=1}^{R_j}\frac{K_j}{p_j^{(r)}}\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\theta}}\left[\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\theta}}\right]^T. \tag{4.16}$$

By denoting the indices of sensors in $\mathcal{A}_p$ by $\{j_i^p\}_{i=1}^{N_p}$, we define the matrices $\boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}}$ and $\boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}$ for all p as

$$\boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}} \triangleq \left[\boldsymbol{\phi}_{j_1^p 1}^{\boldsymbol{\theta}^{(p)}}, \boldsymbol{\phi}_{j_1^p 2}^{\boldsymbol{\theta}^{(p)}}, ..., \boldsymbol{\phi}_{j_1^p R_{j_1^p}}^{\boldsymbol{\theta}^{(p)}}, \boldsymbol{\phi}_{j_2^p 1}^{\boldsymbol{\theta}^{(p)}}, ..., \boldsymbol{\phi}_{j_{N_p}^p R_{j_{N_p}^p}}^{\boldsymbol{\theta}^{(p)}}\right], \tag{4.17}$$

and

$$\boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}} \triangleq \left[\boldsymbol{\phi}_{j_1^p 1}^{\boldsymbol{\tau}^{(p)}}, \boldsymbol{\phi}_{j_1^p 2}^{\boldsymbol{\tau}^{(p)}}, ..., \boldsymbol{\phi}_{j_1^p R_{j_1^p}}^{\boldsymbol{\tau}^{(p)}}, \boldsymbol{\phi}_{j_2^p 1}^{\boldsymbol{\tau}^{(p)}}, ..., \boldsymbol{\phi}_{j_{N_p}^p R_{j_{N_p}^p}}^{\boldsymbol{\tau}^{(p)}}\right], \tag{4.18}$$

where the vectors $\boldsymbol{\phi}_{jr}^{\boldsymbol{\theta}^{(p)}}$ and $\boldsymbol{\phi}_{jr}^{\boldsymbol{\tau}^{(p)}}$ in (4.17) and (4.18) are given by

$$\boldsymbol{\phi}_{jr}^{\boldsymbol{\theta}^{(p)}} \triangleq \sqrt{\frac{K_j}{p_j^{(r)}}}\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\theta}} \quad \text{and} \quad \boldsymbol{\phi}_{jr}^{\boldsymbol{\tau}^{(p)}} \triangleq \sqrt{\frac{K_j}{p_j^{(r)}}}\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}}. \tag{4.19}$$

By employing the singular value decomposition of $\boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}}$ and $\boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}$ for all p

$$\boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}=\mathbf{U}_{\boldsymbol{\tau}^{(p)}}\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}\mathbf{V}_{\boldsymbol{\tau}^{(p)}}^T \text{ and } \boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}}=\mathbf{U}_{\boldsymbol{\theta}^{(p)}}\boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(p)}}\mathbf{V}_{\boldsymbol{\theta}^{(p)}}^T, \tag{4.20}$$

the expressions of $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$, $\mathbf{B}_p$, and $\mathbf{J}_{\boldsymbol{\theta}}$ in (4.13)–(4.15) can be written in compact forms following

$$\mathbf{J}_{\boldsymbol{\tau}^{(p)}} = \boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}\boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}^T = \mathbf{U}_{\boldsymbol{\tau}^{(p)}}\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}^T\mathbf{U}_{\boldsymbol{\tau}^{(p)}}^T, \tag{4.21}$$

$$\mathbf{B}_p = \boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}}\boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}^T, \tag{4.22}$$

and

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\theta}} &= \sum_{p=0}^{P} \mathbf{J}_{\mathcal{A}_p} = \sum_{p=0}^{P} \boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}} \boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}}^T \\ &= \sum_{p=0}^{P} \mathbf{U}_{\boldsymbol{\theta}^{(p)}} \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(p)}} \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(p)}}^T \mathbf{U}_{\boldsymbol{\theta}^{(p)}}^T. \end{aligned} \tag{4.23}$$

4.3.1 Inestimable Spoofing Attacks

Next we show that just due to the sensor system employing a quantization with a limited alphabet, the adversaries can launch a class of spoofing attacks which bring about a singular FIM $\mathbf{J}_{\boldsymbol{\Theta}}$ due to the singularity of $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ for some $p \in \{1, 2, ..., P\}$. We formally define these inestimable spoofing attacks as follows.

Definition 2 (Inestimable spoofing attack) *The p-th spoofing attack is referred to as an inestimable spoofing attack (ISA) if the corresponding $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ defined in (4.13) is singular.*

From (4.13), we have the following result with regard to the singularity of $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$.

Theorem 13 *For the p-th spoofing attack, if the dimension D_p of the attack parameter $\boldsymbol{\tau}^{(p)}$ satisfies*

$$D_p > \sum_{j \in \mathcal{A}_p} R_j - |\mathcal{A}_p|, \tag{4.24}$$

then $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ is singular, and furthermore, the FIM $\mathbf{J}_{\boldsymbol{\Theta}}$ is also singular.

Proof: It is clear that

$$\sum_{r=1}^{R_j} p_j^{(r)} = 1, \tag{4.25}$$

for all j. Hence, we can obtain that

$$\sum_{r=1}^{R_j} \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} = \mathbf{0},\ \forall j, \tag{4.26}$$

which yields

$$\operatorname{rank}\left(\sum_{r=1}^{R_j} \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} \left[\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}}\right]^T\right) \leq R_j - 1,\ \forall j. \tag{4.27}$$

Thus, the rank of $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ is bounded above as per

$$\begin{aligned}
\operatorname{rank}\left(\mathbf{J}_{\tau^{(p)}}\right) &= \operatorname{rank}\left(\sum_{j\in\mathcal{A}_p}\sum_{r=1}^{R_j} \frac{K_j}{p_j^{(r)}} \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} \left[\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}}\right]^T\right) \\
&\leq \sum_{j\in\mathcal{A}_p} \operatorname{rank}\left(\sum_{r=1}^{R_j} \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}} \left[\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}}\right]^T\right) \\
&\leq \sum_{j\in\mathcal{A}_p} (R_j - 1) \\
&= \sum_{j\in\mathcal{A}_p} R_j - |\mathcal{A}_p|
\end{aligned} \tag{4.28}$$

Since $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ is a D_p-by-D_p positive semidefinite matrix, we know that $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ is singular if $D_p > \sum_{j\in\mathcal{A}_p} R_j - |\mathcal{A}_p|$. Finally, the proof concludes by noting that $\mathbf{J}_{\boldsymbol{\Theta}}$ is singular as long as $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ is singular. ■

The proof of *Theorem 13* demonstrates that the rank of $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ is upper bounded by the sum of the sizes of the alphabet sets employed at the sensors under the p-th spoofing attack minus the size of $\mathcal{A}_p$. This implies that the numbers of quantization levels employed at the compromised sensors will limit the size of the attack vector parameter the quantized estimation system can estimate with accuracy that increases with more observations. *Theorem 13* provides a sufficient condition under which inestimable spoofing attacks can be launched.

Thus, these inestimable spoofing attacks, which are quantization induced, can be easily constructed in practice, even without any information about the value of $\boldsymbol{\theta}$ and the quantization regions $\{I_j^{(r)}\}$ at each sensor. Further, if the adversaries have knowledge of the number of quantization levels of each attacked sensor, they know the minimum size of the attack vector parameter they can employ to ensure an inestimable spoofing attack. One simple example of an inestimable spoofing attack employs $D_p > \sum_{j\in\mathcal{A}_p} R_j - |\mathcal{A}_p|$ and

$$\tilde{x}_{jk} = \sum_{i=1}^{D_p} \tau_i^{(p)} \left(x_{jk}\right)^i . \tag{4.29}$$

If (4.24) is not satisfied, the inestimability is determined by the $\{I_j^{(r)}\}$ employed at the attacked sensors and the set of after-attack pdfs $\{g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)})\}$. From (4.21), it is seen that the inestimability of the p-th spoofing attack is equivalent to

$$\operatorname{rank}\left(\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}\right) < D_p. \tag{4.30}$$

In the presence of inestimable spoofing attacks, the FIM $\mathbf{J}_{\boldsymbol{\Theta}}$ for joint estimation of the desired vector parameter and the attack vector parameters is singular, which implies that the FC is unable to improve the estimation of $\boldsymbol{\theta}$ via jointly estimating $\boldsymbol{\theta}$ and the attack vector parameters in the CRB sense. If (4.30) is true for all $p = 1, 2, ..., P$, this means the best the FC can do in this sense is to estimate $\boldsymbol{\theta}$ using only unattacked data, and hence the CRB for $\boldsymbol{\theta}$ in such case can be obtained as

$$\mathrm{CRB}_{\mathrm{ISA}}\left(\boldsymbol{\theta}\right) = \mathbf{J}_{\mathcal{A}_0}^{-1} = \mathbf{U}_{\boldsymbol{\theta}^{(0)}} \left(\boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(0)}} \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(0)}}^T\right)^{-1} \mathbf{U}_{\boldsymbol{\theta}^{(0)}}^T \tag{4.31}$$

by employing (4.23).

4.3.2 Optimal Estimable Spoofing Attacks

In this section, we focus on estimable spoofing attacks which are defined as follows.

Definition 3 (Estimable spoofing attack) *The p-th spoofing attack is said to be estimable if the corresponding* $\mathbf{J}_{\boldsymbol{\tau}^{(p)}}$ *defined in (4.13) is nonsingular.*

Without loss of generality, we assume all spoofing attacks are estimable in this subsection. Otherwise, we can eliminate the observations at ISA sensors, and just consider the joint estimation of the desired vector parameter $\boldsymbol{\theta}$ and the estimable attack vector parameters.

From (4.12) and (4.15), we can obtain the CRB for $\boldsymbol{\theta}$ in the presence of estimable spoofing attacks as

$$\begin{aligned}\left[\mathbf{J}_{\boldsymbol{\Theta}}^{-1}\right]_{1:D_{\boldsymbol{\theta}}} &= \left(\mathbf{J}_{\boldsymbol{\theta}} - \sum_{p=1}^{P} \mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}^{(p)}}^{-1} \mathbf{B}_p^T\right)^{-1} \\ &= \left[\mathbf{J}_{\mathcal{A}_0} + \sum_{p=1}^{P} \left(\mathbf{J}_{\mathcal{A}_p} - \mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}^{(p)}}^{-1} \mathbf{B}_p^T\right)\right]^{-1}. \end{aligned} \tag{4.32}$$

In the following theorem, we provide an upper bound on the CRB for $\boldsymbol{\theta}$ in (4.32) in the positive semidefinite sense.

Theorem 14 *In the presence of estimable spoofing attacks, the CRB for* $\boldsymbol{\theta}$ *is bounded above as per*

$$CRB_{ESA}\left(\boldsymbol{\theta}\right) = \left[\mathbf{J}_{\boldsymbol{\Theta}}^{-1}\right]_{1:D_{\boldsymbol{\theta}}} \preceq \mathbf{J}_{\mathcal{A}_0}^{-1}. \tag{4.33}$$

Equality in (4.33) holds if and only if $\forall p = 1, 2, ..., P$,

$$\mathscr{R}\left(\mathbf{V}_{\boldsymbol{\theta}^{(p)}} \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(p)}}^T\right) \subseteq \mathscr{R}\left(\mathbf{V}_{\boldsymbol{\tau}^{(p)}} \boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}^T\right). \tag{4.34}$$

Proof: Let's first examine the term in the sum in (4.32). Noticing by (4.21), (4.22) and (4.23), we can express $\mathbf{J}_{\mathcal{A}_p} - \mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}(p)}^{-1} \mathbf{B}_p^T$ as

$$\begin{aligned} &\mathbf{J}_{\mathcal{A}_p} - \mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}(p)}^{-1} \mathbf{B}_p^T \\ =&\boldsymbol{\Phi}_{\boldsymbol{\theta}(p)} \boldsymbol{\Phi}_{\boldsymbol{\theta}(p)}^T - \boldsymbol{\Phi}_{\boldsymbol{\theta}(p)} \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)}^T \left(\boldsymbol{\Phi}_{\boldsymbol{\tau}(p)} \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)}^T \right)^{-1} \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)} \boldsymbol{\Phi}_{\boldsymbol{\theta}(p)}^T. \end{aligned} \tag{4.35}$$

Denote

$$\mathbf{D} \triangleq \left(\boldsymbol{\Phi}_{\boldsymbol{\tau}(p)} \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)}^T \right)^{-1} \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)} \boldsymbol{\Phi}_{\boldsymbol{\theta}(p)}^T, \tag{4.36}$$

then by employing (4.35), we can obtain that

$$\begin{aligned} &\mathbf{J}_{\mathcal{A}_p} - \mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}(p)}^{-1} \mathbf{B}_p^T \\ &= \left(\boldsymbol{\Phi}_{\boldsymbol{\theta}(p)}^T - \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)}^T \mathbf{D} \right)^T \left(\boldsymbol{\Phi}_{\boldsymbol{\theta}(p)}^T - \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)}^T \mathbf{D} \right) \\ &\succeq \mathbf{0}. \end{aligned} \tag{4.37}$$

What's more, the equality in (4.37) is attained if and only if

$$\boldsymbol{\Phi}_{\boldsymbol{\theta}(p)}^T - \boldsymbol{\Phi}_{\boldsymbol{\tau}(p)}^T \mathbf{D} = \mathbf{0}, \forall p \geq 1, \tag{4.38}$$

which is equivalent to $\forall p \geq 1$,

$$\mathbf{V}_{\boldsymbol{\tau}(p)} \Big[\mathbf{I} - \boldsymbol{\Lambda}_{\boldsymbol{\tau}(p)}^T \left(\boldsymbol{\Lambda}_{\boldsymbol{\tau}(p)} \boldsymbol{\Lambda}_{\boldsymbol{\tau}(p)}^T \right)^{-1} \boldsymbol{\Lambda}_{\boldsymbol{\tau}(p)} \Big] \mathbf{V}_{\boldsymbol{\tau}(p)}^T \mathbf{V}_{\boldsymbol{\theta}(p)} \boldsymbol{\Lambda}_{\boldsymbol{\theta}(p)}^T = \mathbf{0},$$

and therefore, we can obtain that

$$\mathscr{R} \left(\mathbf{V}_{\boldsymbol{\theta}(p)} \boldsymbol{\Lambda}_{\boldsymbol{\theta}(p)}^T \right) \subseteq \mathscr{R} \left(\mathbf{V}_{\boldsymbol{\tau}(p)} \boldsymbol{\Lambda}_{\boldsymbol{\tau}(p)}^T \right), \forall p \geq 1. \tag{4.39}$$

Consequently, from (4.32), (4.37), and (4.39), we can conclude that

$$\left[\mathbf{J}_{\boldsymbol{\Theta}}^{-1}\right]_{1:D_{\boldsymbol{\theta}}} \preceq \mathbf{J}_{\mathcal{A}_0}^{-1}, \tag{4.40}$$

with equality if and only if $\forall p = 1, 2, ..., P$,

$$\mathscr{R}\left(\mathbf{V}_{\boldsymbol{\theta}^{(p)}} \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(p)}}^{T}\right) \subseteq \mathscr{R}\left(\mathbf{V}_{\boldsymbol{\tau}^{(p)}} \boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}^{T}\right). \tag{4.41}$$

■

In *Theorem 14*, we provide the necessary and sufficient conditions under which the estimable spoofing attacks can deteriorate the CRB for estimating $\boldsymbol{\theta}$ to its upper bound as shown in (4.33). We formally define this class of optimal estimable spoofing attacks next.

Definition 4 (Optimal Estimable Spoofing Attack) *An estimable spoofing attack which satisfies the necessary and sufficient condition in (4.34) is called an optimal estimable spoofing attack (OESA).*

The physical meanings of the terms in (4.32) and the insight into *Theorem 14* deserve some discussion. The term $\mathbf{J}_{\mathcal{A}_0}$ represents the information on $\boldsymbol{\theta}$ embedded in the data from $\mathcal{A}_0$, while $\mathbf{J}_{\mathcal{A}_p}$ indicates the information on $\boldsymbol{\theta}$ that can be provided by the data from $\mathcal{A}_p$ if $\boldsymbol{\tau}^{(p)}$ is known to the FC. The term $\mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}^{(p)}}^{-1} \mathbf{B}_p^T$ specifies the degradation of the information on $\boldsymbol{\theta}$ from $\mathcal{A}_p$, which is induced by the uncertainty of $\boldsymbol{\tau}^{(p)}$. By considering the interpretations of these terms, the insight into *Theorem 14* is that if and only if (4.34) holds, the uncertainty of $\boldsymbol{\tau}^{(p)}$ can reduce the information on $\boldsymbol{\theta}$ conveyed by the data from $\mathcal{A}_p$ to 0 in which case the sum in the inverse does not contribute to (4.32). Moreover, *Theorem 14* points out that the degradation $\mathbf{B}_p \mathbf{J}_{\boldsymbol{\tau}^{(p)}}^{-1} \mathbf{B}_p^T$ cannot be strictly larger than $\mathbf{J}_{\mathcal{A}_p}$.

Theorem 14 also describes how to design optimal estimable spoofing attacks. The adversaries choose $\{g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)})\}$ to meet the necessary and sufficient condition in (4.34). One trivial example of OESA, which may be relatively easy to detect, is to replace the original measurements at the attacked sensors by some regenerated data obeying a distribution not parameterized by $\boldsymbol{\theta}$, which leads to $\boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}} = \mathbf{0}$ for all $p \geq 1$, and therefore, (4.34) is satisfied. In the following part, some typical OESA examples of practical interest are investigated.

Corollary 1 *If the spoofing attacks are such that for any $p \geq 1$, $\exists \lambda_p$ satisfying*

$$\boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}} = \lambda_p \boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}, \tag{4.42}$$

then the CRB $[\mathbf{J}_{\boldsymbol{\Theta}}^{-1}]_{1:D_{\boldsymbol{\theta}}}$ for $\boldsymbol{\theta}$ will be maximized in the positive semidefinite sense, more specifically

$$\left[\mathbf{J}_{\boldsymbol{\Theta}}^{-1}\right]_{1:D_{\boldsymbol{\theta}}} = \mathbf{J}_{\mathcal{A}_0}^{-1}. \tag{4.43}$$

Furthermore, the necessary and sufficient condition under which (4.42) is satisfied for any $\boldsymbol{\theta}$, $\boldsymbol{\tau}^{(p)}$ and $\{I_j^{(r)}\}$ is that $\forall j \in \mathcal{A}_p$, the after-attack pdf $g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)})$ can be expressed as

$$g_j\left(x \left| \boldsymbol{\theta}, \boldsymbol{\tau}^{(p)} \right.\right) = \tilde{g}_j\left(x \left| \lambda_p \boldsymbol{\theta} + \boldsymbol{\tau}^{(p)} \right.\right), \tag{4.44}$$

for some $\tilde{g}_j$.

Proof: Note that if for any $p \geq 1$, $\exists \lambda_p$ such that

$$\boldsymbol{\Phi}_{\boldsymbol{\theta}^{(p)}} = \lambda_p \boldsymbol{\Phi}_{\boldsymbol{\tau}^{(p)}}, \tag{4.45}$$

then $\forall p \geq 1$, $D_{\boldsymbol{\theta}} = D_p$ and

$$\mathscr{R}\left(\mathbf{V}_{\boldsymbol{\theta}^{(p)}}\boldsymbol{\Lambda}_{\boldsymbol{\theta}^{(p)}}^{T}\right) \subseteq \mathscr{R}\left(\mathbf{V}_{\boldsymbol{\tau}^{(p)}}\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}^{T}\right).$$

Thus, by *Theorem 14*, we can obtain that

$$\left[\mathbf{J}_{\boldsymbol{\Theta}}^{-1}\right]_{1:D_{\boldsymbol{\theta}}} = \mathbf{J}_{\mathcal{A}_0}^{-1}. \tag{4.46}$$

In addition, (4.45) is equivalent to

$$\frac{\partial p_j^{(r)}}{\partial \boldsymbol{\theta}} = \lambda_p \frac{\partial p_j^{(r)}}{\partial \boldsymbol{\tau}^{(p)}},\ \forall j \in \mathcal{A}_p,\ \forall r. \tag{4.47}$$

Noticing by (4.5), in order to render (4.47) be assured for any $\boldsymbol{\theta}$, $\boldsymbol{\tau}^{(p)}$ and $\{I_j^{(r)}\}$, the adversaries need to ensure that

$$\frac{\partial}{\partial \boldsymbol{\theta}} g_j\left(x \left| \boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}\right.\right) = \lambda_p \frac{\partial}{\partial \boldsymbol{\tau}^{(p)}} g_j\left(x \left| \boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}\right.\right) \tag{4.48}$$

for all $j \in \mathcal{A}_p$ and for any $\boldsymbol{\theta}$ and $\boldsymbol{\tau}^{(p)}$.

It is clear that if

$$g_j\left(x \left| \boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}\right.\right) = \tilde{g}_j\left(x \left| \lambda_p \boldsymbol{\theta} + \boldsymbol{\tau}^{(p)}\right.\right), \tag{4.49}$$

for some $\tilde{g}_j$, then (4.48) holds. On the other hand, if (4.48) is true for any $\boldsymbol{\theta}$ and $\boldsymbol{\tau}^{(p)}$, then $\forall l = 1, 2, ..., D_{\boldsymbol{\theta}}$,

$$(1, -\lambda_p)\begin{pmatrix} \frac{\partial}{\partial \theta_l} g_j\left(x \left| \{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \theta_l, \tau_l^{(p)}\right.\right) \\ \frac{\partial}{\partial \tau_l^{(p)}} g_j\left(x \left| \{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \theta_l, \tau_l^{(p)}\right.\right) \end{pmatrix} = 0$$

for any $\boldsymbol{\theta}$ and $\boldsymbol{\tau}^{(p)}$, which implies that the gradient of $g_j(x|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \theta_l, \tau_l^{(p)})$ with respect to $[\theta_l, \tau_l^{(p)}]^T$ is parallel to the vector $[\lambda_p, 1]^T$ for any θ_l and $\tau_l^{(p)}$. Therefore, for any l, if

$$(\lambda_p, 1)\begin{pmatrix} 0 \\ t \end{pmatrix} = (\lambda_p, 1)\begin{pmatrix} \theta_l \\ \tau_l^{(p)} \end{pmatrix}, \tag{4.50}$$

that is, $t = \lambda_p\theta_l + \tau_l^{(p)}$, then we can obtain that

$$\begin{aligned} & g_j\left(x\left|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, 0, t\right.\right) \\ & = g_j\left(x\left|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \theta_l, \tau_l^{(p)}\right.\right). \end{aligned} \tag{4.51}$$

As a result, for any l, by employing (4.51) and defining

$$\begin{aligned} & \tilde{g}_{j,l}\left(x\left|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, t\right.\right) \\ & \triangleq g_j\left(x\left|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, 0, t\right.\right), \end{aligned} \tag{4.52}$$

we can express $g_j(x|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \theta_l, \tau_l^{(p)})$ as

$$\begin{aligned} & g_j\left(x\left|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \theta_l, \tau_l^{(p)}\right.\right) \\ & = \tilde{g}_{j,l}\left(x\left|\{\theta_m\}_{m\neq l}, \{\tau_m^{(p)}\}_{m\neq l}, \lambda_p\theta_l + \tau_l^{(p)}\right.\right) \end{aligned} \tag{4.53}$$

for some $\tilde{g}_{j,l}$, which implies that

$$g_j\left(x\left|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}\right.\right) = \tilde{g}_j\left(x\left|\lambda_p\boldsymbol{\theta} + \boldsymbol{\tau}^{(p)}\right.\right) \tag{4.54}$$

for some $\tilde{g}_j$. ■

As demonstrated by *Corollary 1*, if the spoofing attack gives rise to an after-attack pdf $g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)})$ which is only parameterized by the sum of $\lambda_p\boldsymbol{\theta}$ and $\boldsymbol{\tau}^{(p)}$ for any λ_p, then the spoofing attack is optimal in the sense of *Definition 4*. This class of OESAs are interesting and powerful in practice, since their optimality is independent of the values of the desired vector parameter and the attack vector parameter. The DRFM example discussed in the introduction which introduces a time delay is one example of this class of OESAs (with $\lambda_p = 1$). For the scenario where the desired parameter is the mean of the observations, which is a popular signal model for sensor network estimation systems with quantized data [51, 54, 56, 72], this class of OESAs can be easily launched by just adding an offset to the measurements at each attacked sensor.

Another representative example of the class of OESAs described by (4.44) is extensively considered in smart grid systems under the name data-injection attacks, see [9, 71, 76, 77] and the references therein. At time instant k, the direct current power flow model in the absence of spoofing attacks can be expressed as

$$\mathbf{x}_k = \mathbf{H}\boldsymbol{\theta} + \mathbf{n}_k. \tag{4.55}$$

Considering the p-th data-injection attack, the after-attack measurements from the sensors in $\mathcal{A}_p$ at time instant k are given by

$$[\tilde{\mathbf{x}}_k]_{\mathcal{A}_p} = [\mathbf{x}_k]_{\mathcal{A}_p} + \mathbf{a}^{(p)} = [\mathbf{H}]_{\mathcal{A}_p,:}\boldsymbol{\theta} + \mathbf{a}^{(p)} + [\mathbf{n}_k]_{\mathcal{A}_p}, \tag{4.56}$$

where $\mathbf{a}^{(p)}$ represents the data injected by the p-th spoofing attack. If the adversaries choose $\mathbf{a}^{(p)}$ such that

$$\mathbf{a}^{(p)} = [\mathbf{H}]_{\mathcal{A}_p,:}\boldsymbol{\tau}^{(p)} \tag{4.57}$$

for some $\boldsymbol{\tau}^{(p)}$, then the after-attack measurements from the sensors in $\mathcal{A}_p$ can be equivalently written as

$$[\tilde{\mathbf{x}}_k]_{\mathcal{A}_p} = [\mathbf{H}]_{\mathcal{A}_p,:}\left(\boldsymbol{\theta} + \boldsymbol{\tau}^{(p)}\right) + [\mathbf{n}_k]_{\mathcal{A}_p}, \tag{4.58}$$

and therefore, (4.44) is satisfied by the data-injection attack. Further, by *Corollary 1*, the CRB for $\boldsymbol{\theta}$ is maximized in the positive semidefinite sense if all the attacks are of this type. Moreover, it can be shown that the stealth attack or undetectable attack in [9, 71, 76], which attracts extensive attention in recent literature on smart grids, is just such an attack with $P = 1$.

In addition to the class of OESAs described in (4.44), there are many other OESAs. For example, if the p-th spoofing attack satisfies that $\forall j \in \mathcal{A}_p$, $g_j(x|\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}) = \tilde{g}_j(x|h_j(\boldsymbol{\theta}, \boldsymbol{\tau}^{(p)}))$ for some $\tilde{g}_j$ and some symmetric function h_j of $\boldsymbol{\theta}$ and $\boldsymbol{\tau}^{(p)}$, then it can be shown that the p-th spoofing attack is an OESA provided that the values of $\boldsymbol{\tau}^{(p)}$ and $\boldsymbol{\theta}$ are equal.

4.3.3 Discussion

Under the conditions of *Definition 1*, it is clear that $\mathbf{J}_{\mathcal{A}_0}^{-1}$ is an upper bound on the CRB for $\boldsymbol{\theta}$, no matter what kind of attacks have been launched. From (4.31) and *Theorem 14*, the CRB for $\boldsymbol{\theta}$ under ISA or OESA equals to its upper bound $\mathbf{J}_{\mathcal{A}_0}^{-1}$. Therefore, according to *Definition 1*, both ISA and OESA are OGDSAs. Furthermore, note that $\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}$ is a $D_p \times (\sum_{j\in\mathcal{A}_p} R_j)$ matrix, and hence, $\text{rank}(\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}) \leq D_p$. Thus, any OGDSA is either an ISA when $\text{rank}(\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}) < D_p$, or an OESA when $\text{rank}(\boldsymbol{\Lambda}_{\boldsymbol{\tau}^{(p)}}) = D_p$.

A particular note of interest is that the results in *Section 4.3.1* and *4.3.2* can be used to judge whether the attacked measurements are useful or not in terms of reducing CRB under the conditions of *Definition 1*. In particular, it is seen from (4.31) and *Theorem 14* that the CRB for $\boldsymbol{\theta}$ in the presence of ISA or OESA is the same as the CRB for $\boldsymbol{\theta}$ when only unattacked data is used. Thus, we obtain the following corollary.

Corollary 2 *Under the conditions of* Definition 1*, the necessary and sufficient condition under which the attacked measurements are useless in terms of reducing CRB is that the spoofing attacks belong to either ISA or OESA which are defined in* Definition 2 *and* 4 *respectively.*

However, the fundamental mechanisms of ISA and OESA for making the attacked measurements useless in terms of reducing CRB are very different. To be specific, ISA renders the task of estimating the attacks beyond the capabilities of the quantized estimation system by causing the FIM for jointly estimating the desired and attack parameters to be singular, thus preventing the FC from potentially improving the CRB of the estimate of $\boldsymbol{\theta}$ by jointly estimating $\boldsymbol{\theta}$ and the attacks. In contrast, even though the FC is able to estimate the attacks, paying a big price in computational complexity for jointly estimating $\boldsymbol{\theta}$ and the attacks, the FC is not able to obtain any improvement in the CRB performance for $\boldsymbol{\theta}$ under OESA.

It is worth mentioning that (4.31) and *Theorem 14* demonstrate that under the conditions of *Definition 1*, the CRB for $\boldsymbol{\theta}$ reaches its upper bound in the presence of ISA or OESA. In practice, however, the FC may not be able to well identify the set of unattacked sensors and categorize the attacked sensors into different groups according to distinct types of spoofing attacks. Thus, the actual estimation performance under ISA and OESA can be expected to be inferior to the analytical results in this section. To illustrate this, consider the example of the data-injection attack in smart grids as described in (4.56). It can be shown that if

the adversaries only employ one spoofing attack, then it is possible for the adversaries to dramatically impact the estimate of the desired vector parameter to produce an arbitrarily large bias [9, 76].

4.4 Joint Identification and Estimation under Optimal Estimable Spoofing Attack

In this section, we focus on a class of OESAs in which for any p, $\forall j \in \mathcal{A}_p$, the FIM for $\boldsymbol{\tau}^{(p)}$ based on the data from the j-th sensor is nonsingular. Further, we assume that $\mathbf{J}_{\mathcal{A}_0}$ defined in (4.16) is nonsingular in the presence of spoofing attacks. This could occur, for example, if only a small subset of sensors can be attacked in a distributed sensor setting or if a subset of sensors can be well protected in advance to give rise to a nonsingular $\mathbf{J}_{\mathcal{A}_0}$.

Before proceeding, the following assumptions are made from a practical viewpoint.

Assumption 9 *As the sensors are assumed to be spread over a wide area and typically adversaries have limited resources, we assume that no more than half of sensors are attacked.*

Assumption 10 (Significant Attack) *In order to give rise to sufficient impact on the statistical characterization of the measurements at each attacked sensors, every attacker is required to guarantee a minimum norm of the attack vector parameter, that is,*

$$\|\boldsymbol{\tau}^{(p)}\|_2 > d_{\boldsymbol{\tau}}, \ \forall p. \tag{4.59}$$

We do not conisder modifications smaller than (4.59) as attacks and assume they have little impact on performance.

The results in *Section 4.3* demonstrate that under OESA, the CRB for $\boldsymbol{\theta}$ which employs the data from both attacked and unattacked sensors is equal to the CRB for $\boldsymbol{\theta}$ which only makes use of unattacked data. Since CRB is the adopted performance metric, we only need to identify the set of unattacked sensors, and the categorization of the attacked sensors according to distinct types of spoofing attacks is no longer necessary for estimating $\boldsymbol{\theta}$ in the presence of OESA. To this end, we use $\{\boldsymbol{\xi}^{(j)}\}_{j=1}^{N}$ instead of $\{\boldsymbol{\tau}^{(p)}\}_{p=1}^{P}$ to denote the attack vector parameters employed by the adversaries in the following part. To be specific, $\boldsymbol{\xi}^{(j)}$ denotes the attack vector parameter employed at the j-th sensor. For the sake of notational simplicity, we introduce the following notations

$$q_{jr} \triangleq \int_{I_j^{(r)}} f_j\left(x\,|\boldsymbol{\theta}\right) dx \text{ and } \tilde{q}_{jr} \triangleq \int_{I_j^{(r)}} g_j\left(x\,\middle|\boldsymbol{\theta}, \boldsymbol{\xi}^{(j)}\right) dx \tag{4.60}$$

where $\tilde{q}_{jr}$ and q_{jr} represent the r-th value of the after-attack pmf at the j-th sensor when it is attacked and unattacked respectively.

Let $\boldsymbol{\Omega}$ denote a vector containing the desired vector parameter $\boldsymbol{\theta}$, the set of unknown attack vector parameters $\{\boldsymbol{\xi}^{(j)}\}$ as well as a set of unknown binary state variables $\{\eta_j\}$ that

$$\boldsymbol{\Omega} \triangleq \left[\boldsymbol{\Xi}^T, \boldsymbol{\eta}^T\right]^T, \tag{4.61}$$

where

$$\boldsymbol{\Xi} \triangleq \left[\boldsymbol{\theta}^T, \left(\boldsymbol{\xi}^{(1)}\right)^T, \left(\boldsymbol{\xi}^{(2)}\right)^T, ..., \left(\boldsymbol{\xi}^{(N)}\right)^T\right]^T \tag{4.62}$$

and

$$\boldsymbol{\eta} \triangleq \left[\eta_1, \eta_2, ..., \eta_N\right]^T. \tag{4.63}$$

The j-th element of $\boldsymbol{\eta}$ is zero, i.e., $\eta_j = 0$, if the j-th sensor is unattacked, while $\eta_j = 1$ implies the j-th sensor is attacked. The log-likelihood function evaluated at $\tilde{\mathbf{u}} = \mathbf{r}$ is

$$\begin{aligned} L(\boldsymbol{\Omega}) &\triangleq \ln \Pr(\tilde{\mathbf{u}} = \mathbf{r} \,|\boldsymbol{\Omega}) \\ &= \sum_{j=1}^{N} \sum_{k=1}^{K_j} \left[\eta_j \ln \tilde{q}_{jr_{jk}} + (1-\eta_j) \ln q_{jr_{jk}}\right]. \end{aligned} \tag{4.64}$$

Based on this setting, the FC can jointly identify the state of each sensor and estimate the desired vector parameter $\boldsymbol{\theta}$ by solving the following constrained optimization problem

$$\hat{\boldsymbol{\Omega}} = \arg\max_{\boldsymbol{\Omega}} \sum_{j=1}^{N} \sum_{k=1}^{K_j} \left[\eta_j \ln \tilde{q}_{jr_{jk}} + (1-\eta_j) \ln q_{jr_{jk}}\right] \tag{4.65a}$$

$$\text{s. t.} \quad \eta_j \in \{0,1\}, \ \forall j, \tag{4.65b}$$

$$\sum_{j=1}^{N} \eta_j < \frac{N}{2}, \tag{4.65c}$$

$$\|\boldsymbol{\xi}^{(j)}\|_2 > d_{\boldsymbol{\tau}}, \ \forall \eta_j = 1, \tag{4.65d}$$

where the constraints in (4.65c) and (4.65d) are due to *Assumption 9* and *Assumption 10*.

The integer constraint in (4.65b) makes the optimization problem difficult to solve. For small N, it may be solved exactly simply by exhaustively searching through all possible combinations of $\{\eta_j\}$, while for large N, this is not feasible in practice, since the number of all possible combination of $\{\eta_j\}$ is on the order of 2^N. To this end, it is of considerable practical interest to develop an efficient algorithm to solve the optimization problem in (4.65). In this section, we propose a heuristic for solving (4.65).

4.4.1 Random Relaxation with the EM Algorithm

According to the constraint in (4.65b), η_j is an unknown deterministic binary variable, and hence, (4.65b) is equivalent to

$$\pi_j \stackrel{\Delta}{=} \Pr\left(\eta_j = 1\right) \in \{0, 1\} \text{ and } \Pr\left(\eta_j = 0\right) = 1 - \pi_j, \forall j. \tag{4.66}$$

Further, by dropping the constraint (4.65c) as well as (4.65d), and then relaxing the deterministic $\{\eta_j\}$ to be random, that is, allowing $\pi_j = \Pr\left(\eta_j = 1\right) \in [0, 1]$ for all $j = 1, 2, ..., N$, the problem in (4.65) reduces to

$$\hat{\mathbf{\Omega}}_\pi = \arg\max_{\mathbf{\Omega}_\pi} \sum_{j=1}^{N} \sum_{k=1}^{K_j} \ln\left[\pi_j \tilde{q}_{jr_{jk}} + \left(1 - \pi_j\right) q_{jr_{jk}}\right] \tag{4.67a}$$

$$\text{s. t. } \pi_j \in [0, 1], \ \forall j = 1, 2, ..., N, \tag{4.67b}$$

where $\mathbf{\Omega}_\pi \stackrel{\Delta}{=} [\mathbf{\Xi}^T, \boldsymbol{\pi}^T]^T$ and $\boldsymbol{\pi} \stackrel{\Delta}{=} [\pi_1, \pi_2 ..., \pi_N]^T$.

The physical interpretation behind (4.67) is that via random relaxation of the deterministic binary vector state variable $\boldsymbol{\eta}$, the set $\mathcal{A}_0$ of unattacked sensors is no longer deterministic, and moreover, each sensor in the sensor network is attacked with a certain probability π_j at every time instant.

By introducing a latent vector variable

$$\mathbf{z} = [z_{11}, z_{12}, ..., z_{1K_1}, z_{21}, ..., z_{NK_N}]^T, \tag{4.68}$$

where $z_{jk} = 1$ indicates that the k-th measurement at the j-th sensor was attacked, and $z_{jk} = 0$ implies that the k-th measurement at the j-th sensor was unattacked, we can employ

the Expectation-Maximization (EM) algorithm [**?**, 83], which is an iterative method that alternates between performing an expectation (E) step and a maximization (M) step, to solve the relaxed problem in (4.67).

E-step

The *E-step* computes the expected log-likelihood function $Q(\mathbf{\Omega}_\pi|\mathbf{\Omega}'_\pi)$, with respect to $\mathbf{z}$ given the quantized data $\tilde{\mathbf{u}} = \mathbf{r}$ and the current estimate of the vector parameter $\hat{\mathbf{\Omega}}'_\pi = [(\hat{\mathbf{\Xi}}')^T, (\hat{\boldsymbol{\pi}}')^T]^T$, as following

$$Q\left(\mathbf{\Omega}_\pi \left| \hat{\mathbf{\Omega}}'_\pi \right.\right) \triangleq \mathbb{E}_{\mathbf{z}|\hat{\mathbf{\Omega}}'_\pi, \tilde{\mathbf{u}}=\mathbf{r}} \left\{ L\left(\mathbf{\Omega}_\pi\right) \right\}, \tag{4.69}$$

where the log-likelihood function $L\left(\mathbf{\Omega}_\pi\right)$ is given by

$$\begin{aligned} L\left(\mathbf{\Omega}_\pi\right) &= \ln \Pr\left(\mathbf{z}, \tilde{\mathbf{u}} = \mathbf{r} \middle| \mathbf{\Omega}_\pi\right) \\ &= \ln \Pr\left(\tilde{\mathbf{u}} = \mathbf{r} \middle| \mathbf{\Omega}_\pi, \mathbf{z}\right) + \ln \Pr\left(\mathbf{z} \middle| \mathbf{\Omega}_\pi\right) \\ &= \sum_{j=1}^{N} \sum_{k=1}^{K_j} \Big\{ \mathbb{1}_{\{z_{jk}=1\}} \left(\ln \tilde{q}_{jr_{jk}} + \ln \pi_j \right) \\ &\qquad + \mathbb{1}_{\{z_{jk}=0\}} \left[\ln q_{jr_{jk}} + \ln\left(1 - \pi_j\right) \right] \Big\}. \end{aligned} \tag{4.70}$$

Define

$$\upsilon_{jk}^{(1)} \triangleq \mathbb{E}_{\mathbf{z}|\hat{\mathbf{\Omega}}'_\pi, \tilde{\mathbf{u}}=\mathbf{r}} \left\{ \mathbb{1}_{\{z_{jk}=1\}} \right\} = \frac{\hat{\pi}'_j \tilde{q}_{jr_{jk}}}{\hat{\pi}'_j \tilde{q}_{jr_{jk}} + \left(1 - \hat{\pi}'_j\right) q_{jr_{jk}}} \tag{4.71}$$

and

$$\upsilon_{jk}^{(0)} \triangleq \mathbb{E}_{\mathbf{z}|\hat{\mathbf{\Omega}}'_\pi, \tilde{\mathbf{u}}=\mathbf{r}} \left\{ \mathbb{1}_{\{z_{jk}=0\}} \right\} = 1 - \upsilon_{jk}^{(1)}, \tag{4.72}$$

then by employing (4.69) and (4.70), we can obtain the expected log-likelihood function

$$Q\left(\boldsymbol{\Omega}_\pi \middle| \hat{\boldsymbol{\Omega}}'_\pi\right) = \sum_{j=1}^{N}\sum_{k=1}^{K_j}\left\{\upsilon_{jk}^{(1)}\left(\ln \tilde{q}_{jr_{jk}} + \ln \pi_j\right) + \upsilon_{jk}^{(0)}\left[\ln q_{jr_{jk}} + \ln\left(1-\pi_j\right)\right]\right\}. \tag{4.73}$$

M-step

The *M-step* seeks to find a new estimate of the vector parameter $\hat{\boldsymbol{\Omega}}_\pi$ to update the current estimate of the vector parameter $\hat{\boldsymbol{\Omega}}'_\pi$ by maximizing the expected log-likelihood function $Q(\boldsymbol{\Omega}_\pi|\hat{\boldsymbol{\Omega}}'_\pi)$, that is,

$$\hat{\boldsymbol{\Omega}}_\pi = \left[\hat{\boldsymbol{\Xi}}^T, \hat{\boldsymbol{\pi}}^T\right]^T = \arg\max Q\left(\boldsymbol{\Omega}_\pi \middle| \hat{\boldsymbol{\Omega}}'_\pi\right). \tag{4.74}$$

Updated estimate of $\boldsymbol{\pi}$ According to (4.74), the updated estimate $\hat{\pi}_j$ should satisfy

$$\frac{\partial Q\left(\boldsymbol{\Omega}_\pi \middle| \hat{\boldsymbol{\Omega}}'_\pi\right)}{\partial \pi_j} = \frac{1}{\pi_j}\sum_{k=1}^{K_j}\upsilon_{jk}^{(1)} - \frac{1}{1-\pi_j}\sum_{k=1}^{K_j}\upsilon_{jk}^{(0)} = 0, \tag{4.75}$$

which yields, by employing (4.72),

$$\hat{\pi}_j = \frac{1}{K_j}\sum_{k=1}^{K_j}\upsilon_{jk}^{(1)}. \tag{4.76}$$

Updated estimate of $\boldsymbol{\Xi}$ Similarly, the updated estimate $\hat{\boldsymbol{\Xi}}$ is the solution of the following equation

$$\nabla_{\boldsymbol{\Xi}} Q\left(\boldsymbol{\Omega}_\pi \middle| \hat{\boldsymbol{\Omega}}'_\pi\right) = \mathbf{0}. \tag{4.77}$$

Generally, a closed-form solution for the above equation may not exist. To solve (4.77) in such cases, Newton's method can be employed with an initial point $\hat{\boldsymbol{\Xi}}^{(0)} = \hat{\boldsymbol{\Xi}}'$. At the $(i+1)$-th

iteration of Newton's Method, the updated point $\hat{\boldsymbol{\Xi}}^{(t+1)}$ can be expressed as

$$\begin{aligned}&\hat{\boldsymbol{\Xi}}^{(t+1)}\\&= \hat{\boldsymbol{\Xi}}^{(t)} - \kappa_t \left[\nabla^2_{\boldsymbol{\Xi}} Q\left(\boldsymbol{\Omega}_\pi^{(t)} \middle| \hat{\boldsymbol{\Omega}}'_\pi\right)\right]^{-1} \nabla_{\boldsymbol{\Xi}} Q\left(\boldsymbol{\Omega}_\pi^{(t)} \middle| \hat{\boldsymbol{\Omega}}'_\pi\right)\end{aligned} \tag{4.78}$$

where $\boldsymbol{\Omega}_\pi^{(t)} = [(\hat{\boldsymbol{\Xi}}^{(t)})^T, (\hat{\boldsymbol{\pi}}')^T]^T$, and $\kappa_t \in (0,1)$ is the t-th step size computed by using a backtracking line search [84].

For completeness, the explicit expressions for the gradient and Hessian of the expected log-likelihood function with respect to $\boldsymbol{\Xi}$ are provided. The gradient $\nabla_{\boldsymbol{\Xi}} Q(\boldsymbol{\Omega}_\pi^{(t)} | \hat{\boldsymbol{\Omega}}'_\pi)$ consists of the quantities $\frac{\partial}{\partial \theta_l} Q(\boldsymbol{\Omega}_\pi^{(t)} | \hat{\boldsymbol{\Omega}}'_\pi)$ and $\frac{\partial}{\partial \xi_l^{(j)}} Q(\boldsymbol{\Omega}_\pi^{(t)} | \hat{\boldsymbol{\Omega}}'_\pi)$ for different j and l, which can be computed by

$$\begin{aligned}&\frac{\partial}{\partial \theta_l} Q\left(\boldsymbol{\Omega}_\pi^{(t)} \middle| \hat{\boldsymbol{\Omega}}'_\pi\right)\\&= \sum_{j=1}^{N} \sum_{k=1}^{K_j} \left\{ \upsilon_{jk}^{(1)} \frac{1}{\tilde{q}_{jr_{jk}}} \frac{\partial}{\partial \theta_l} \tilde{q}_{jr_{jk}} + \upsilon_{jk}^{(0)} \frac{1}{q_{jr_{jk}}} \frac{\partial}{\partial \theta_l} q_{jr_{jk}} \right\}\end{aligned} \tag{4.79}$$

and

$$\frac{\partial}{\partial \xi_l^{(j)}} Q\left(\boldsymbol{\Omega}_\pi^{(t)} \middle| \hat{\boldsymbol{\Omega}}'_\pi\right) = \sum_{k=1}^{K_j} \upsilon_{jk}^{(1)} \frac{1}{\tilde{q}_{jr_{jk}}} \frac{\partial}{\partial \xi_l^{(j)}} \tilde{q}_{jr_{jk}}. \tag{4.80}$$

The elements of the Hessian $\nabla^2_{\boldsymbol{\Xi}} Q(\boldsymbol{\Omega}_\pi^{(t)}|\hat{\boldsymbol{\Omega}}'_\pi)$ can be calculated by the following expressions

$$\begin{aligned}
&\frac{\partial^2}{\partial\theta_l\partial\theta_m} Q\left(\boldsymbol{\Omega}_\pi^{(t)}\,\middle|\,\hat{\boldsymbol{\Omega}}'_\pi\right) \\
&= \sum_{j=1}^{N}\sum_{k=1}^{K_j}\left\{ \upsilon_{jk}^{(1)}\left(\frac{1}{\tilde{q}_{jr_{jk}}}\frac{\partial^2\tilde{q}_{jr_{jk}}}{\partial\theta_l\partial\theta_m} - \frac{1}{\tilde{q}^2_{jr_{jk}}}\frac{\partial\tilde{q}_{jr_{jk}}}{\partial\theta_l}\frac{\partial\tilde{q}_{jr_{jk}}}{\partial\theta_m}\right)\right. \\
&\qquad \left. +\upsilon_{jk}^{(0)}\left(\frac{1}{q_{jr_{jk}}}\frac{\partial^2 q_{jr_{jk}}}{\partial\theta_l\partial\theta_m} - \frac{1}{q^2_{jr_{jk}}}\frac{\partial q_{jr_{jk}}}{\partial\theta_l}\frac{\partial q_{jr_{jk}}}{\partial\theta_m}\right)\right\},
\end{aligned} \tag{4.81}$$

$$\begin{aligned}
&\frac{\partial^2}{\partial\theta_l\partial\xi_m^{(j)}} Q\left(\boldsymbol{\Omega}_\pi^{(t)}\,\middle|\,\hat{\boldsymbol{\Omega}}'_\pi\right) \\
&= \sum_{k=1}^{K_j}\upsilon_{jk}^{(1)}\left(\frac{1}{\tilde{q}_{jr_{jk}}}\frac{\partial^2\tilde{q}_{jr_{jk}}}{\partial\theta_l\partial\xi_m^{(j)}} - \frac{1}{\tilde{q}^2_{jr_{jk}}}\frac{\partial\tilde{q}_{jr_{jk}}}{\partial\theta_l}\frac{\partial\tilde{q}_{jr_{jk}}}{\partial\xi_m^{(j)}}\right),
\end{aligned} \tag{4.82}$$

$$\begin{aligned}
&\frac{\partial^2}{\partial\xi_l^{(j)}\partial\xi_m^{(j)}} Q\left(\boldsymbol{\Omega}_\pi^{(t)}\,\middle|\,\hat{\boldsymbol{\Omega}}'_\pi\right) \\
&= \sum_{k=1}^{K_j}\upsilon_{jk}^{(1)}\left(\frac{1}{\tilde{q}_{jr_{jk}}}\frac{\partial^2\tilde{q}_{jr_{jk}}}{\partial\xi_l^{(j)}\partial\xi_m^{(j)}} - \frac{1}{\tilde{q}^2_{jr_{jk}}}\frac{\partial\tilde{q}_{jr_{jk}}}{\partial\xi_l^{(j)}}\frac{\partial\tilde{q}_{jr_{jk}}}{\partial\xi_m^{(j)}}\right),
\end{aligned} \tag{4.83}$$

and

$$\frac{\partial^2}{\partial\xi_l^{(i)}\partial\xi_m^{(j)}} Q\left(\boldsymbol{\Omega}_\pi^{(t)}\,\middle|\,\hat{\boldsymbol{\Omega}}'_\pi\right) = 0, \text{ if } i \neq j. \tag{4.84}$$

The quantities in (4.79)–(4.84) are all evaluated at $\boldsymbol{\Omega}_\pi^{(t)}$. Repeating the calculation of (4.78) until $\{\hat{\boldsymbol{\Xi}}^{(t)}\}$ converges, the limit point $\hat{\boldsymbol{\Xi}}$ of $\{\hat{\boldsymbol{\Xi}}^{(t)}\}$ is the solution for (4.77), and also the updated estimate of $\boldsymbol{\Xi}$.

The convergence of the EM algorithm is guaranteed and the detailed analysis can be found in [83, 85], that is to say, by iteratively alternating between *E-step* and *M-step*, the

solution for (4.67) can be obtained. It is worth mentioning that since we do not require a very accurate solution for the relaxed optimization problem in (4.67), once the difference between the updated and current estimates is sufficiently small, we can terminate the iterations in the EM algorithm and utilize the current estimate of $\mathbf{\Omega}_\pi$ in the following rounding step.

4.4.2 Constrained Variable Threshold Rounding and Barrier Method

By utilizing the EM algorithm as illustrated in *Section 4.4.1*, we can obtain the solution $\hat{\mathbf{\Omega}}_\pi$ for the relaxed optimization problem in (4.67). The element $\hat{\pi}_j$ of $\hat{\mathbf{\Omega}}_\pi$ specifies the probability of the j-th sensor being attacked over time. However, according to (4.65c) and (4.66), we know that before relaxation, $\hat{\pi}_j \in \{0, 1\}$ and $\mathbf{1}^T \hat{\boldsymbol{\pi}} < N/2$. To this end, we consider the task of rounding $\hat{\boldsymbol{\pi}}$ to a valid binary vector. To accomplish this task, we propose a constrained variable threshold rounding (CVTR) approach which is based on the heuristic developed by Zymnis *et al.* [86]. The basic idea of the CVTR is that we first round $\hat{\boldsymbol{\pi}}$ to generate a set of most likely probability vectors $\{\tilde{\boldsymbol{\pi}}^{(l)}\}$ with binary elements which satisfy the constraints in (4.65c). Then, under constraint (4.65d), the joint maximum likelihood estimate of the desired vector parameter and attack vector parameters are pursued over the generated set of valid probability binary vectors $\{\tilde{\boldsymbol{\pi}}^{(l)}\}$.

We first generate the set of the most likely valid binary probability vectors $\{\tilde{\boldsymbol{\pi}}^{(l)}\}$ by employing the CVTR which can be described as

$$\left\{\tilde{\boldsymbol{\pi}}^{(l)}\right\} \triangleq \Bigg\{ \operatorname{sgn}\left(\hat{\boldsymbol{\pi}} - \lambda \mathbf{1}\right) : 0 \leq \lambda \leq 1, \\ \left\|\operatorname{sgn}\left(\hat{\boldsymbol{\pi}} - \lambda \mathbf{1}\right)\right\|_1 < \frac{N}{2} \Bigg\}. \tag{4.85}$$

Since the j-th element $\tilde{\pi}_j^{(l)}$ of $\tilde{\boldsymbol{\pi}}^{(l)}$ denotes the probability the j-th sensor is attacked, each

probability vector $\tilde{\boldsymbol{\pi}}^{(l)}$ with binary values corresponds to a deterministic state variable vector $\tilde{\boldsymbol{\eta}}^{(l)}$ as following

$$\tilde{\boldsymbol{\eta}}^{(l)} = \tilde{\boldsymbol{\pi}}^{(l)},\ \forall l. \tag{4.86}$$

We refer to $\{\tilde{\boldsymbol{\eta}}^{(l)}\}$ as the set of the most likely state variable vectors, and we only consider the combinations of $\{\eta_j\}$ in this set. Further, it is seen from (4.85) that as λ increases from 0 to 1, this approach only generates up to $\lfloor N/2 \rfloor$ distinct valid binary probability vectors. Thus, it is feasible to exhaustively evaluate the maximum likelihood function, which is maximized with respect to $\boldsymbol{\Xi}$, for each given $\tilde{\boldsymbol{\eta}}^{(l)}$. As a result, the optimization problem in (4.65) can be reduced to

$$\hat{\boldsymbol{\Omega}}_R = \left[\hat{\boldsymbol{\Xi}}_R^T, \hat{\boldsymbol{\eta}}_R^T\right]^T = \arg \max_{\boldsymbol{\eta} \in \left\{\tilde{\boldsymbol{\eta}}^{(l)}\right\}} \max_{\boldsymbol{\Xi}} L\left(\boldsymbol{\Omega}\right) \tag{4.87a}$$

$$\text{s. t.} \quad \|\boldsymbol{\xi}^{(j)}\|_2 > d_{\boldsymbol{\tau}},\ \forall \eta_j = 1, \tag{4.87b}$$

As (4.87) demonstrates, we need to solve the inner maximization for each candidate state variable vector $\tilde{\boldsymbol{\eta}}^{(l)}$, and then keep the solution which gives rise to the maximal objective function in (4.87). Noticing that the constraint in (4.87b) only has effects on the inner maximization, the inner constrained maximization for each $\tilde{\boldsymbol{\eta}}^{(l)}$ in (4.87) can be converted to an unconstrained problem by employing a logarithmic barrier function as

$$\max_{\boldsymbol{\Xi}} \left\{ \sum_{j=1}^{N} \sum_{k=1}^{K_j} \left[\tilde{\eta}_j^{(l)} \ln \tilde{q}_{jr_{jk}} + \left(1 - \tilde{\eta}_j^{(l)}\right) \ln q_{jr_{jk}} \right] + \mu \sum_{j=1}^{N} \tilde{\eta}_j^{(l)} \ln \left(\|\boldsymbol{\xi}^{(j)}\|_2 - d_{\boldsymbol{\tau}} \right) \right\}, \tag{4.88}$$

where the positive barrier parameter μ determines the accuracy with which (4.88) approx-

imates the inner constrained maximization in (4.87). Since the objective function in (4.88) is differentiable, the unconstrained problem in (4.88) can be similarly solved by Newton's Method as in *Section 4.4.1* for any given μ.

Let $\hat{\boldsymbol{\Xi}}_{\mu}^{(l)}$ denote the solution of (4.88) for any given $\tilde{\boldsymbol{\eta}}^{(l)}$ and μ, and let $L_{*}^{(l)}$ represent the optimal objective value of the inner constrained maximization in (4.87a) for any given $\tilde{\boldsymbol{\eta}}^{(l)}$. It can be shown that as $\mu \to 0$, any limit point $\hat{\boldsymbol{\Xi}}_{*}^{(l)}$ of the sequence $\{\hat{\boldsymbol{\Xi}}_{\mu}^{(l)}\}_{\mu}$ is a solution of the inner constrained maximization in (4.87) [87]. Thus, we can obtain an accurate solution of the inner constrained maximization in (4.87) by iteratively solving (4.88) for a sequence $\{\mu_m\}$ of positive barrier parameters, which decrease monotonically to zero, such that the solution $\hat{\boldsymbol{\Xi}}_{\mu_m}^{(l)}$ for μ_m is chosen as the starting point for the next iteration with barrier parameter μ_{m+1}. By defining $l^* \stackrel{\Delta}{=} \max_l \; L_{*}^{(l)}$, the solution of the constrained optimization problem in (4.87) can be obtained as

$$\hat{\boldsymbol{\Omega}}_R = \left[\hat{\boldsymbol{\Xi}}_R, \hat{\boldsymbol{\eta}}_R\right]^T = \left[\left(\hat{\boldsymbol{\Xi}}_{*}^{(l^*)}\right)^T, \left(\tilde{\boldsymbol{\eta}}^{(l^*)}\right)^T\right]^T. \tag{4.89}$$

4.4.3 Discussion

The random relaxation and constrained variable threshold rounding approach proposed for solving the joint identification and estimation problem in (4.65) is a heuristic approach. Further improvement in the identification and estimation can sometimes be obtained by performing a local optimization by searching around $\hat{\boldsymbol{\eta}}_R$ [84]. To be specific, we cycle through $j = 1, 2, ..., N$, and at j-th step, we flip the j-th element of $\hat{\boldsymbol{\eta}}_R$. If this change can result in an increase in the optimal value in (4.87a) for some $\boldsymbol{\Xi}$, then we accept this change, otherwise we move on to the next index. We continue checking each element of the state variable vector until we have rejected any new change. After this local optimization, the estimate of state variable vector is at least 1-OPT, since no change in one element of the estimate can increase

the likelihood function.

It is well known that the condition number of the Hessian matrix of the logarithmic barrier function in (4.88) might become increasingly larger as the barrier parameter decreases to 0. In order to overcome the ill-conditioning issue in practical computation, the numerically stable approximation of the Newton direction can be utilized in Newton's method for solving (4.88) with small barrier parameter, see [87] and the references therein. It is worth mentioning that to preserve the generality, we don't make additional assumptions to ensure the convexity of the objective functions in the section. Hence, the EM algorithm and Newton's method involved in our approach might converge to a locally optimal point if the starting point is not close to the globally optimal point. To avoid this possibility, multiple starting points can be employed and we choose the one that yields the maximal objective function at convergence [82].

It is seen from (4.87)–(4.89) that the proposed approach employs joint estimation of the desired vector parameter and the attack vector parameters in the identification process, and moreover, as shown in (4.89), the final estimate of the desired vector parameter is directly obtained from the joint estimation once l^* is determined. Thus, for some scenarios where the spoofing attacks are not OGDSAs, one can expect that the proposed approach is able to outperform the estimation approach which just utilizes the unattacked data to estimate the desired vector parameter, since the attacked data is employed in the proposed approach.

4.5 Numerical Results

In this section, we investigate the performance of the approaches proposed in *Section 4.4* for some example cases. Specifically, we consider a sensor network consisting of $N = 10$

sensors. Each sensor makes K measurements of the physical phenomenon, and employs an identical 4-bit quantizer with a set of thresholds $\{0, \pm 1, \pm 2, ..., \pm 7, \pm\infty\}$ to convert analog measurements to quantized data before transmitting them to the FC. We take the following signal model of the before-attack measurements, which has been studied the most in sensor network area,

$$x_{jk} = \theta + n_{jk}, \ \forall k \text{ and } \forall j, \tag{4.90}$$

where θ is a deterministic unknown parameter, and $\{n_{jk}\}$ is an i.i.d. zero-mean Gaussian noise sequence with distribution $\mathcal{N}(0, \sigma^2)$. Further, we assume that the first 3 sensors in the sensor network are under data-injection spoofing attacks. The after-attack measurements are described as

$$\tilde{x}_{jk} = \theta + a_{jk} + n_{jk}, \ \forall k \text{ and } \forall j = 1, 2, 3, \tag{4.91}$$

where a_{jk} is the unknown attack injected at the j-th sensor at time k.

4.5.1 Scalar Parameter Case with Deterministic $\{a_{jk}\}$

In this subsection, we assume that the injected attacks $a_{1k} = -2$, $a_{2k} = -1$, and $a_{3k} = 1$ are deterministic unknown for all k, and the variance of the noise $\sigma^2 = 5$ is known to the FC. The desired scalar parameter is $\theta = 1$, and the constraint on the attack parameter defined in (4.59) is $d_{\boldsymbol{\tau}} = 0.8$. We first test the performance of the approach of the random relaxation (RR) with the EM and CVTR in identifying the attacked and unattacked sensors. Fig. 4.2 illustrates the Monte Carlo approximation (1000 times) of the ensemble average of the percentage of all mis-classified sensors as a function of the number K of measurements at each sensor. As Fig. 4.2 shows, the average percentage of mis-classified sensors decreases towards 0 as K increases.

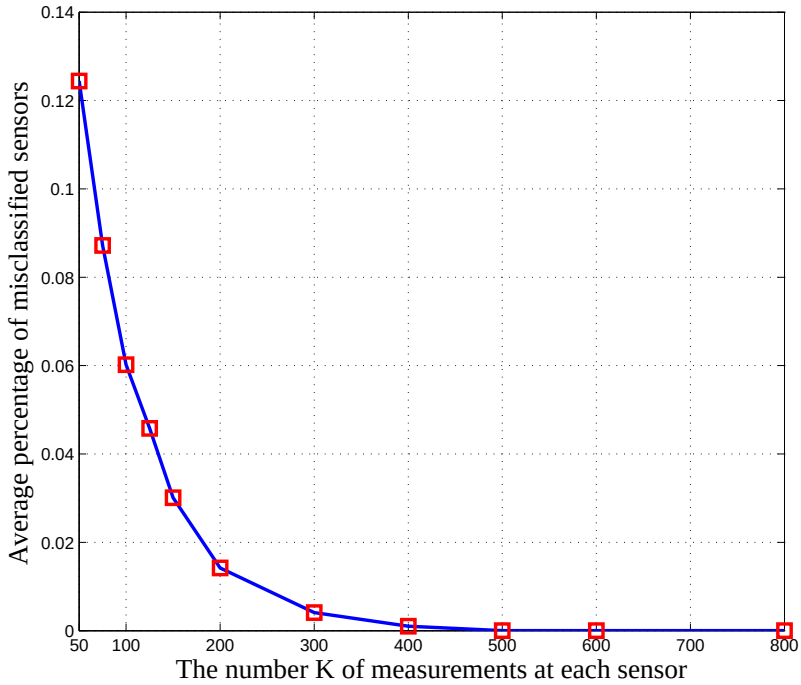

Figure 4.2: Performance of identifying the attacked and unattacked sensors for scalar parameter.

Next, we examine the estimation performance of the proposed approaches in *Section 4.4*, that is, the approach of the RR with the EM, and the approach of the RR with the EM and CVTR. Fig. 4.3 depicts the mean squared error (MSE) performance of the two approaches for estimating θ. For comparison, the genie CRB for θ which assumes the FC is aware of the true states of sensors and only utilizes the unattacked data to estimate θ is also provided in Fig. 4.3. It is seen that as K increases, the MSE performance of the approach with CVTR for estimating θ converges to the genie CRB for θ from above, which would be the case if the proposed estimator for the desired parameter is asymptotically efficient for this case. The results in Fig. 4.3 also corroborates the previous theoretical results in *Section 4.3* that under OESA, jointly estimating the desired parameter and the attack parameter does not improve the estimation performance for the desired parameter in the CRB sense when

compared to the case where only unattacked data is employed to estimate θ. In addition, the MSE performance of the approach with CVTR is shown to be better than the approach which only employs the RR with the EM algorithm, which implies that the proposed constrained variable threshold rounding can further improve the estimation performance for the deisred parameter.

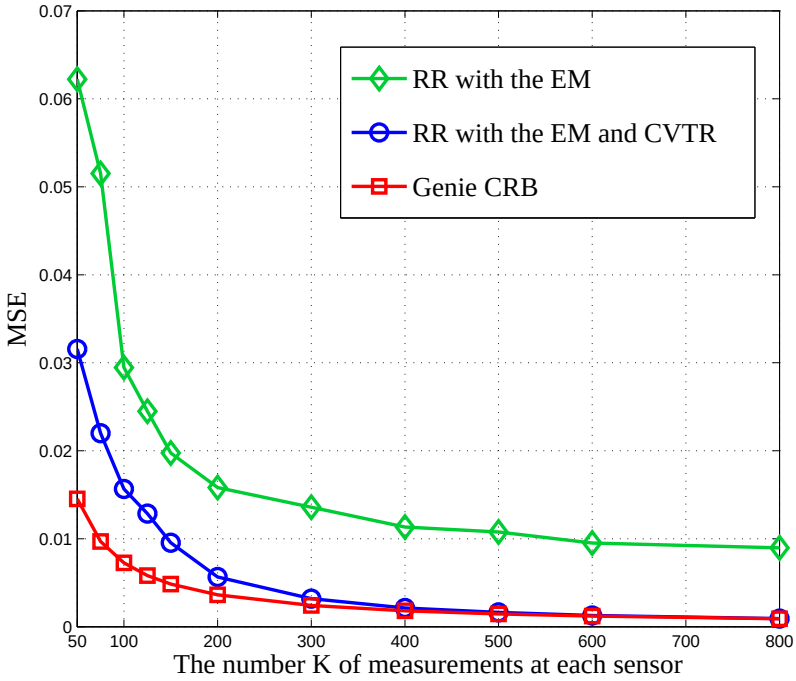

Figure 4.3: Estimation performance of the proposed approaches for scalar parameter.

4.5.2 Vector Parameter Case with Random $\{a_{jk}\}$

In this subsection, we extend the results in *Section 4.5.1* to the vector parameter case. We consider the scenario that the parameter $\theta = 1$ and the variance of the noise $\sigma^2 = 3$ are both the parameters of interest. Moreover, the unknown injected attacks $\{a_{jk}\}$ are independent random variables, where a_{jk} obeys the Gaussian distribution $\mathcal{N}(\alpha_j, \beta_j)$ for all k. The desired

vector parameter $\boldsymbol{\theta} \triangleq [\theta, \sigma^2]^T$ and the attack vector parameters $\{\boldsymbol{\xi}^{(j)} \triangleq [\alpha_j, \beta_j]^T\}_{j=1,2,3}$ are $\boldsymbol{\theta} = [1,3]^T$, $\boldsymbol{\xi}^{(1)} = [-2,1]^T$, $\boldsymbol{\xi}^{(2)} = [-1,2]^T$, and $\boldsymbol{\xi}^{(3)} = [1,2]^T$. In our simulations, the constraint on the attack vector parameter defined in (4.59) is $d_{\boldsymbol{\tau}} = 2$. We first study the performance of the approach of the RR with the EM and CVTR in identifying the attacked and unattacked sensors for the vector parameter case. Fig. 4.4 depicts the Monte Carlo approximation (1000 times) of the ensemble average of the percentage of all mis-classified sensors versus K. It is seen from Fig. 4.4 that the average percentage of mis-classified sensors reduces towards 0 as K increases. In Fig. 4.5, we plot the MSE performance of our proposed approaches for $\boldsymbol{\theta}$. The genie CRB performance for $\boldsymbol{\theta}$ which assumes the FC is aware of the true states of sensors and only utilizes the unattacked data to estimate $\boldsymbol{\theta}$ is also plotted for comparison. As Fig. 4.5 shows, we obtain similar results to those for the scalar parameter case in *Section 4.5.1*. To be specific, the MSE performance of the RR with the EM and CVTR is very close to the genie CRB and outperforms the approach which only employs the RR with the EM algorithm.

4.6 Summary

In this chapter, we study the distributed estimation of a deterministic vector parameter by using quantized data in the presence of spoofing attacks. A generalized attack model is employed which manipulates the data using transformations with arbitrary functional forms determined by some attack parameters whose values are unknown to the attacked system. Novel necessary and sufficient conditions are provided under which these transformations provide an OGDSA. It is shown that an OGDSA implies that either the FIM under the conditions of *Definition 1* for jointly estimating the desired and attack parameters is singular

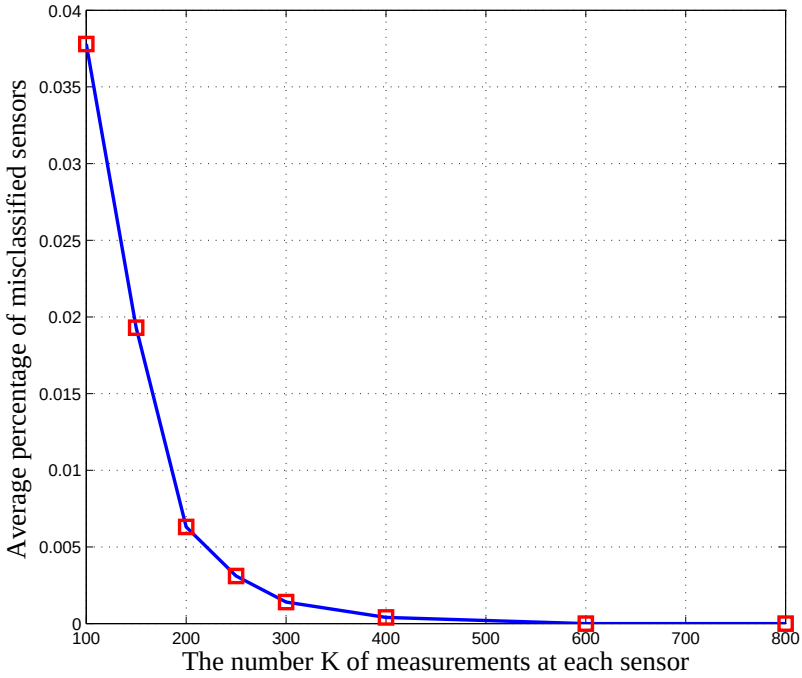

Figure 4.4: Performance of identifying the attacked and unattacked sensors for vector parameter.

or that the attacked system is unable to improve the CRB under the conditions of *Definition 1* for the desired vector parameter through this joint estimation even though the joint FIM is nonsingular. It is demonstrated that it is always possible to construct an OGDSA by properly employing a sufficiently large dimension attack vector parameter relative to the number of quantization levels employed, which was not observed previously. In addition, we demonstrate that under the conditions of *Definition 1*, a spoofing attack can corrupt the original measurements to make them useless in terms of reducing the CRB for estimating the desired vector parameter if and only if it is an OGDSA. For a class of OGDSAs, a computationally efficient heuristic which employs the Expectation-Maximization algorithm and the constrained variable threshold rounding is proposed for the joint identification of attacked sensors and estimation of the desired vector parameter. The proposed heuristic

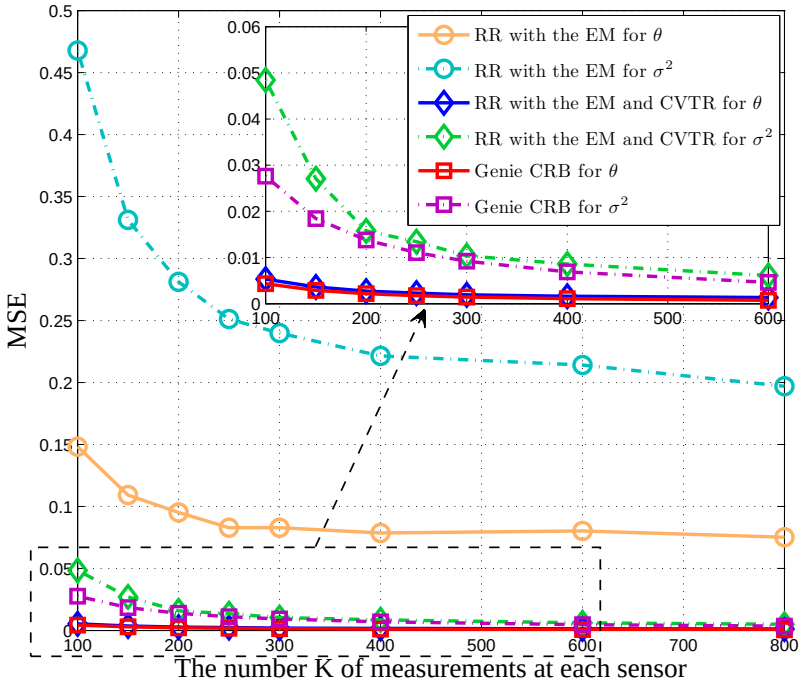

Figure 4.5: Estimation performance of the proposed approaches for vector parameter.

approach is guaranteed to provide a locally optimal solution, but will find globally optimal solutions when they exist when suitable conditions are satisfied. Numerical results show cases where the proposed approach can correctly identify the attacked sensors while providing an estimate whose mean squared error converges to the genie bound based on knowledge of the set of attacked sensors, provided a sufficient number of measurements are available.

Chapter 5

Conclusions

This dissertation presents our research on several selected issues concerning sensor networks which focus on signal detection and estimation problems.

In Chapter 2, the large observation size performance of a truncated detector for a canonical multivariate Gaussian hypothesis testing problem is studied. If the observations consist of data taken at different times, the truncated detector can reduce the storage and multiplications needed when compared to the optimal detector. If the observations are obtained from distributed sensors, the truncated detector not only reduces the communication energy requirement for computing the test statistic, it allows efficient implementation by adopting a consensus algorithm. Motivated by these benefits of utilizing the truncated detector, the performance of the truncated detector in terms of deflection is investigated. Sufficient conditions for a truncation rule and a sequence of tests which lead to no loss in asymptotic deflection ratio of the truncated detector relative to the optimal detector are derived. Several well-accepted and popular classes of system and process models are employed as examples to show that the sufficient conditions are not overly restrictive. For all the examples considered,

we find truncation rules which increase slowly with the number of the observations, implying significant savings. In all the cases considered, numerical results imply that not only do the deflections of the truncated and the optimal detectors converge to the same values for large number of observations for our asymptotically optimal truncation rules, but the probability of detections also converge to the same values for fixed false alarm probabilities.

In Chapter 3, the distributed estimation problem using binary quantized data in the presence of man-in-the-middle attacks is studied. In this work, the sensor data modifications implemented by the adversaries are statistically characterized by a set of unknown probability transition matrices. We demonstrate that the fusion center is able to identify the attacked sensors and categorize these attacked sensors into different subsets according to distinct types of attacks perfectly or with a very small percentage of misclassified sensors, as the number of temporal observations at each sensor grows to infinity or the number of sensors increases to infinity respectively, provided that the set of unattacked sensors is larger than any set of identically attacked sensors. In order to improve the estimation performance by utilizing the attacked sensors, a joint estimation of the statistical description of the attacks and the parameter to be estimated is considered. However, it is shown that the corresponding Fisher information matrix (FIM) is singular if a standard data quantization approach is employed. Thus, it is not possible to accurately estimate the parameters using this approach with an estimate that would always become more and more accurate as we increase the number of observations. Aiming to overcome this, the time-variant quantization approach is proposed which divides the observation time interval at each sensor into several time slots and employs distinct thresholds to quantize the time samples in different time slots. If the number of time samples at each sensor is not less than 2, then it can be shown that the FIM for all unknown parameters in time-variant quantization approach is nonsingular which implies that the statis-

tical properties of the attacks and the parameter to be estimated can be accurately estimated with a sufficiently large number of observations. A necessary and sufficient condition under which the attacked observations can be taken advantage of to improve the asymptotic estimation performance is derived. A notable fact is that for many cases, significant improvement in Cramer-Rao Bound (CRB) performance for the parameter to be estimated can be attained by making use of attacked observations in our proposed fashion. However, for some specific cases, using the attacked observations will not provide better asymptotic estimation performance. It is worth mentioning that both the theoretical analysis and numerical results illustrate that the improvement in CRB performance by utilizing attacked observations in our proposed fashion depends not only on the statistical description of the attacks and the parameter to be estimated, but also on the sets of thresholds of the quantizer, which motivates us to pursue the optimum quantizer design for distributed estimation in the presence of man-in-the-middle attacks in future work.

In Chapter 4, we investigate the distributed estimation of a deterministic vector parameter by using possibly nonbinary quantized data in the presence of spoofing attacks. A generalized attack model is employed which manipulates the data using transformations with arbitrary functional forms determined by some attack parameters whose values are unknown to the attacked system. Novel necessary and sufficient conditions are provided under which these transformations provide a guaranteed attack performance in terms of CRB regardless of the processing the estimation system employs, thus defining a highly desirable attack. It is shown that this highly desirable attack implies that for any such attack when the fusion center can perfectly identify the attacked sensors, either the FIM for jointly estimating the desired and attack parameters is singular or the attacked system is unable to improve the CRB for the desired vector parameter through this joint estimation even though the joint

FIM is nonsingular. It is demonstrated that it is always possible to construct such a desirable attack by properly employing a sufficiently large dimension attack vector parameter relative to the number of quantization levels employed, which was not observed previously. In addition, we demonstrate that when the fusion center can perfectly identify the attacked sensors, a spoofing attack can corrupt the original measurements to make them useless in terms of reducing the CRB for estimating the desired vector parameter if and only if it is such a desirable attack. For a class of such desirable attacks, a computationally efficient heuristic which employs the Expectation-Maximization algorithm and the constrained variable threshold rounding is proposed for the joint identification of attacked sensors and estimation of the desired vector parameter. For the cases considered, numerical results illustrate that the proposed approach can asymptotically correctly identify the attacked sensors while providing an estimate whose mean squared error converges to the genie bound based on knowledge of the set of attacked sensors, provided a sufficient number of measurements are available.

Bibliography

[1] R. Chen, J. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE.* IEEE, 2008, pp. 1876–1884.

[2] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-p signals in cognitive radio networks with uncalibrated multiple antennas," *Signal Processing, IEEE Transactions on*, vol. 59, no. 8, pp. 3764–3774, 2011.

[3] I. Bekkerman and J. Tabrikian, "Target detection and localization using mimo radars and sonars," *Signal Processing, IEEE Transactions on*, vol. 54, no. 10, pp. 3873–3883, 2006.

[4] E. Fishler, A. Haimovich, R. Blum, L. Cimini Jr, D. Chizhik, and R. Valenzuela, "Spatial diversity in radars-models and detection performance," *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 823–838, 2006.

[5] A. Haimovich, R. Blum, and L. Cimini, "MIMO radar with widely separated antennas," *Signal Processing Magazine, IEEE*, vol. 25, no. 1, pp. 116–129, 2008.

[6] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on.* IEEE, 2010, pp. 220–225.

[7] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *Smart Grid, IEEE Transactions on*, vol. 2, no. 2, pp. 342–351, 2011.

[8] Y. Huang, H. Li, K. A. Campbell, and Z. Han, "Defending false data injection attack on smart grid network using adaptive cusum test," in *Information Sciences and Systems (CISS), 2011 45th Annual Conference on.* IEEE, 2011, pp. 1–6.

[9] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *Smart Grid, IEEE Transactions on*, vol. 2, no. 4, pp. 645–658, 2011.

[10] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag New York, Inc., 1994.

[11] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative risk management: concepts, techniques, and tools.* Princeton university press, 2005.

[12] F. Yamazaki, A. Member, M. Shinozuka, and G. Dasgupta, "Neumann expansion for stochastic finite element analysis," *Journal of Engineering Mechanics*, vol. 114, no. 8, pp. 1335–1354, 1988.

[13] W. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of mri data," *Medical Imaging, IEEE Transactions on*, vol. 15, no. 4, pp. 429–442, 1996.

[14] H. Rue and L. Held, *Gaussian Markov random fields: theory and applications*. Chapman & Hall, 2005, vol. 104.

[15] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis*. Prentice hall Upper Saddle River, NJ, vol. 5.

[16] W. Härdle and L. Simar, *Applied multivariate statistical analysis*. Springer Verlag, 2007.

[17] H. Poor and C. Chang, "A reduced-complexity quadratic structure for the detection of stochastic signals," *The Journal of the Acoustical Society of America*, vol. 78, no. 5, pp. 1652–1657, 1985.

[18] Y. Sung, L. Tong, and H. Poor, "Optimal and suboptimal detection of gaussian signals in noise: asymptotic relative efficiency," in *Proceedings of SPIE*, vol. 5910, 2005, p. 591002.

[19] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Massachusetts Institute of Technology, 1984.

[20] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *Automatic Control, IEEE Transactions on*, vol. 31, no. 9, pp. 803–812, 1986.

[21] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[22] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*. IEEE, 2005, pp. 63–70.

[23] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2508–2530, 2006.

[24] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 4, pp. 634–649, 2008.

[25] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *Signal Processing, IEEE Transactions on*, vol. 57, no. 1, pp. 355–369, 2009.

[26] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc wsns with noisy links-part I: Distributed estimation of deterministic signals," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 350–364, 2008.

[27] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, and A. Ribeiro, "Consensus in ad hoc wsns with noisy links-part II: Distributed estimation and smoothing of random signals," *Signal Processing, IEEE Transactions on*, vol. 56, no. 4, pp. 1650–1666, 2008.

[28] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1383–1400, 2010.

[29] A. Dimakis, S. Kar, J. Moura, M. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[30] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1520–1533, 2004.

[31] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[32] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *Signal Processing, IEEE Transactions on*, vol. 57, no. 7, pp. 2748–2761, 2009.

[33] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 1917–1932, 2011.

[34] P. Braca, S. Marano, V. Matta, and P. Willett, "Asymptotic optimality of running consensus in testing binary hypotheses," *Signal Processing, IEEE Transactions on*, vol. 58, no. 2, pp. 814–825, 2010.

[35] C. Baker, "Optimum quadratic detection of a random vector in gaussian noise," *Communication Technology, IEEE Transactions on*, vol. 14, no. 6, pp. 802–805, 1966.

[36] B. Picinbono, "On deflection as a performance criterion in detection," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 31, no. 3, pp. 1072–1081, 1995.

[37] J. Chamberland and V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?" *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5099–5106, 2006.

[38] Y. Sung, L. Tong, and H. Poor, "Neyman-pearson detection of gauss-markov signals in noise: closed-form error exponent and properties," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1354–1365, 2006.

[39] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *Signal Processing, IEEE Transactions on*, vol. 55, no. 3, pp. 822–833, 2007.

[40] J. Michalowicz, J. Nichols, F. Bucholtz, and C. Olson, "An isserlis' theorem for mixed gaussian variables: Application to the auto-bispectral density," *Journal of Statistical Physics*, vol. 136, no. 1, pp. 89–102, 2009.

[41] J. Magnus and H. Neudecker, "Matrix differential calculus with applications in statistics and econometrics," 1988.

[42] S. Demko, W. Moss, and P. Smith, "Decay rates for inverses of band matrices," *Mathematics of Computation*, vol. 43, no. 168, pp. 491–499, 1984.

[43] U. Grenander and G. Szegő, *Toeplitz forms and their applications*. Univ of California Press, 1958.

[44] P. Bickel and E. Levina, "Regularized estimation of large covariance matrices," *The Annals of Statistics*, vol. 36, no. 1, pp. 199–227, 2008.

[45] C. W. Therrien and K. Fukunaga, "Properties of separable covariance matrices and their associated gaussian random processes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 5, pp. 652–656, 1984.

[46] J. A. Ritcey and A. Chindapol, "A kronecker product improvement to pca for space time adaptive processing," in *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, vol. 1. IEEE, 2000, pp. 651–655.

[47] M. G. Genton, "Separable approximations of space-time covariance matrices," *Environmetrics*, vol. 18, no. 7, pp. 681–695, 2007.

[48] M. Sherman, *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*. Wiley, 2011.

[49] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications magazine, IEEE*, vol. 40, no. 8, pp. 102–114, 2002.

[50] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *Information Theory, IEEE Transactions on*, vol. 47, no. 3, pp. 978–1002, 2001.

[51] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case," *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 1131–1143, 2006.

[52] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *Information Theory, IEEE Transactions on*, vol. 51, no. 6, pp. 2210–2219, 2005.

[53] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 23, no. 4, pp. 27–41, 2006.

[54] R. Niu and P. K. Varshney, "Target location estimation in sensor networks with quantized data," *Signal Processing, IEEE Transactions on*, vol. 54, no. 12, pp. 4519–4528, 2006.

[55] O. Ozdemir, R. Niu, and P. K. Varshney, "Channel aware target localization with quantized data in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 57, no. 3, pp. 1190–1202, 2009.

[56] J. Fang and H. Li, "Hyperplane-based vector quantization for distributed estimation in wireless sensor networks," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5682–5699, 2009.

[57] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of Byzantine attacks," *Signal Processing, IEEE Transactions on*, vol. 57, no. 1, pp. 16–29, 2009.

[58] A. Vempaty, K. Agrawal, H. Chen, and P. Varshney, "Adaptive learning of Byzantines' behavior in cooperative spectrum sensing," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. IEEE, 2011, pp. 1310–1315.

[59] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney, "Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks," *Signal Processing, IEEE Transactions on*, vol. 59, no. 2, pp. 774–786, 2011.

[60] X. He, H. Dai, and P. Ning, "A Byzantine attack defender in cognitive radio networks: The conditional frequency check," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 5, pp. 2512–2523, 2013.

[61] A. Vempaty, L. Tong, and P. Varshney, "Distributed inference with Byzantine data: State-of-the-art review on data falsification attacks," *Signal Processing Magazine, IEEE*, vol. 30, no. 5, pp. 65–75, 2013.

[62] A. Vempaty, O. Ozdemir, K. Agrawal, H. Chen, and P. K. Varshney, "Localization in wireless sensor networks: Byzantines and mitigation techniques," *IEEE Transactions on Signal Processing*, vol. 61, pp. 1495–1508, 2013.

[63] T. C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *Information Forensics and Security, IEEE Transactions on*, vol. 3, no. 2, pp. 273–289, 2008.

[64] R. Soosahabi and M. Naraghi-Pour, "Scalable phy-layer security for distributed detection in wireless sensor networks," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 4, pp. 1118–1126, 2012.

[65] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[66] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 23, no. 4, pp. 16–26, 2006.

[67] F. den Hollander, *Large deviations (Fields Institute Monographs)*. Providence, RI: American Mathematical Soc., 2000.

[68] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd ed. New York: Springer-Verlag, 2009.

[69] Z. Li, W. Trappe, Y. Zhang, and B. Nath, "Robust statistical methods for securing wireless localization in sensor networks," in *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, April 2005, pp. 91–98.

[70] J. H. Lee and R. Buehrer, "Characterization and detection of location spoofing attacks," *Communications and Networks, Journal of*, vol. 14, no. 4, pp. 396–409, Aug 2012.

[71] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *Signal Processing Magazine, IEEE*, vol. 29, no. 5, pp. 106–115, 2012.

[72] J. Zhang, R. S. Blum, X. Lu, and D. Conus, "Asymptotically optimum distributed estimation in the presence of attacks," *Signal Processing, IEEE Transactions on*, vol. 63, no. 5, pp. 1086–1101, March 2015.

[73] B. Alnajjab, J. Zhang, and R. S. Blum, "Attacks on sensor network estimation systems with quantization: Performance and optimum processing," *accepted to IEEE Transactions on Signal Processing*.

[74] P. Venkitasubramaniam, L. Tong, and A. Swami, "Quantization for maximin ARE in distributed estimation," *Signal Processing, IEEE Transactions on*, vol. 55, no. 7, pp. 3596–3605, July 2007.

[75] S. Roome, "Digital radio frequency memory," *Electronics Communication Engineering Journal*, vol. 2, no. 4, pp. 147–153, Aug 1990.

[76] D. Liu, P. Ning, A. Liu, C. Wang, and W. K. Du, "Attack-resistant location estimation in wireless sensor networks," *ACM Transactions on Information and System Security (TISSEC)*, vol. 11, no. 4, p. 22, 2008.

[77] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *Smart Grid, IEEE Transactions on*, vol. 2, no. 2, pp. 326–333, 2011.

[78] S. Kim, W. Kuperman, W. Hodgkiss, H. Song, G. Edelmann, and T. Akal, "Robust time reversal focusing in the ocean," *The Journal of the Acoustical Society of America*, vol. 114, p. 145, 2003.

[79] M. I. Skolnik, *Introduction to Radar Systems*, 2nd ed. New York: McGraw Hill Book Co., 1980.

[80] G. Grachev, "Theory of acoustic field invariants in layered waveguides," *Acoustical physics*, vol. 39, no. 1, pp. 33–35, 1993.

[81] G. D'spain, J. Murray, W. Hodgkiss, N. Booth, and P. Schey, "Mirages in shallow water matched field processing," *The Journal of the Acoustical Society of America*, vol. 105, no. 6, pp. 3245–3265, 1999.

[82] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation theory*. Upper Saddle River, NJ: Prentice Hall, 1993.

[83] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. pp. 1–38, 1977.

[84] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[85] C. F. J. Wu, "On the convergence properties of the EM algorithm," *The Annals of Statistics*, vol. 11, no. 1, pp. pp. 95–103, 1983.

[86] A. Zymnis, S. Boyd, and D. Gorinevsky, "Relaxed maximum a posteriori fault identification," *Signal Process.*, vol. 89, no. 6, pp. 989–999, Jun. 2009.

[87] S. Nash, R. Polyak, and A. Sofer, "A numerical comparison of barrier and modified barrier methods for large-scale bound-constrained optimization," in *Large Scale Optimization*, W. Hager, D. Hearn, and P. Pardalos, Eds. Springer US, 1994, pp. 319–338.

Vita

Jiangfan Zhang received the B.Eng. degree in communication engineering from Huazhong University of Science and Technology, Wuhan, China, in 2008, and the M.Eng. degree in information and communication engineering from Zhejiang University, Hangzhou, China, 2011. Since 2011, he has been working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, Lehigh University, Bethlehem, PA. His research interests include signal processing for sensor networking, smart grid, communications, radar, and sonar processing. Mr. Zhang is a recipient of the Dean's Doctoral Student Assistantship, Gotshall Fellowship, and a P. C. Rossin Doctoral Fellow at Lehigh University.