ARRAY SIGNAL PROCESSING IN THE KNOWN WAVEFORM AND STEERING VECTOR CASE

By

YI JIANG

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2003

I dedicate this work to my parents.

ACKNOWLEDGMENTS

This thesis began as a project for the course in spectral analysis taught by my advisor (Dr. Jian Li) in Spring 2002. The original problem was to derive a maximum likelihood estimate of the amplitude of signal with known waveform and steering vector. The derivation itself is easy. In fact I finished that project with ease. Then Dr. Li suggested that I make a comparative study of the asymptotic properties of the ML and Capon estimates. At that time, I had noticed interesting differences in the performance of the two methods. But Dr. Li's encouragement pushed me to find the underlying mathematical structures. I learned much from this experience. Dr. Li also devoted much time to helping me, an ungifted English writer, improve my writing. I am deeply grateful.

I was fortunate to have the opportunity to work with Dr. Petre Stoica over the past year. Chapter 4 of this thesis was made possible by his great instruction. I also express my gratitude to Dr. John M. Shea and Dr. M. Nechyba for serving on my committee.

I express my thanks to all of the members of the Spectral Analysis Laboratory for providing a pleasant work environment.

I dedicate this work to my parents with gratitude for everything.

TABLE OF CONTENTS

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

ARRAY SIGNAL PROCESSING IN THE KNOWN WAVEFORM
AND STEERING VECTOR CASE

By

Yi Jiang

August 2003

Chair: Jian Li
Major Department: Electrical and Computer Engineering

The amplitude estimation of a signal which is known only up to an unknown scaling factor, with interference and noise present, is of interest in several applications, including using the emerging Quadrupole Resonance (QR) technology for detecting explosives. In such applications a sensor array is often deployed for interference suppression. This thesis considers the complex amplitude estimation of a known waveform signal whose array response is also known *a priori*. Two approaches (Capon and maximum likelihood (ML) methods) were considered for signal amplitude estimation in the presence of temporally white but spatially colored interference and noise. We derived closed-form expressions for the expected values and mean-squared errors (MSEs) of the two estimators. A comparative study showed that the ML estimate is unbiased while the Capon estimate is biased downwards for finite data sample lengths. We found that both methods are asymptotically statistically efficient when the number of data samples is large, but not when the signal-to-noise ratio (SNR) is high. We also considered a more general scenario where interference and noise are both spatially and temporally correlated. We modeled the interference and noise vector as a multichannel autoregressive (AR) random process. An alternating least squares (ALS) method for parameter estimation is presented. We show

that in most cases the ALS method is superior to the model-mismatched ML (M^3L) method which ignores the temporal correlation of the interference and noise.

CHAPTER 1
INTRODUCTION

1.1 Thesis Motivation and Objectives

Quadrupole Resonance (QR) technology is an emerging technology for detecting explosives [1, 2]. In that QR application, the ^{14}N in TNT, when stimulated by a sequence of pulses, gives a characteristic response specific to the TNT. We refer to this response (which consists of a sequence of echoes) as the QR signal. Each echo of the QR signal is a back-to-back exponentially damped sinusoid separated by an interval T_e from the adjacent echoes; and with a damping rate determined by a parameter T_2^* (Figure 1–1). The echo amplitude also damps exponentially as e^{-t/T_2}. The QR signal frequency, T_2, and T_2^* are all known quite accurately *a priori*; and the precise echo timing T_e is also available in practice. Hence the QR signal is known *a priori*, to within a multiplicative constant.

The main challenge of using QR technology for landmine detection is that the QR signal is in the same frequency band as the AM/FM radio signals; and the QR signal frequency cannot be changed. To suppress radio interferences, an antenna array can be deployed with one of the sensors (the main antenna) receiving the QR signal besides the interference and noise and the remaining sensors (reference antennas) receiving the interference and noise only. Hence one of the elements of the array steering vector for the QR signal is equal to one; and the remaining elements are zero. Thus in this application, both temporal and spatial information are available *a priori*. One of the main tasks of signal processing in QR technology for landmine detection is to estimate the amplitude of QR signal with high accuracy. It is a mandatory step for signal detection. Our study aimed to improve estimate accuracy by fully exploiting both the temporal information and spatial information of the signal.

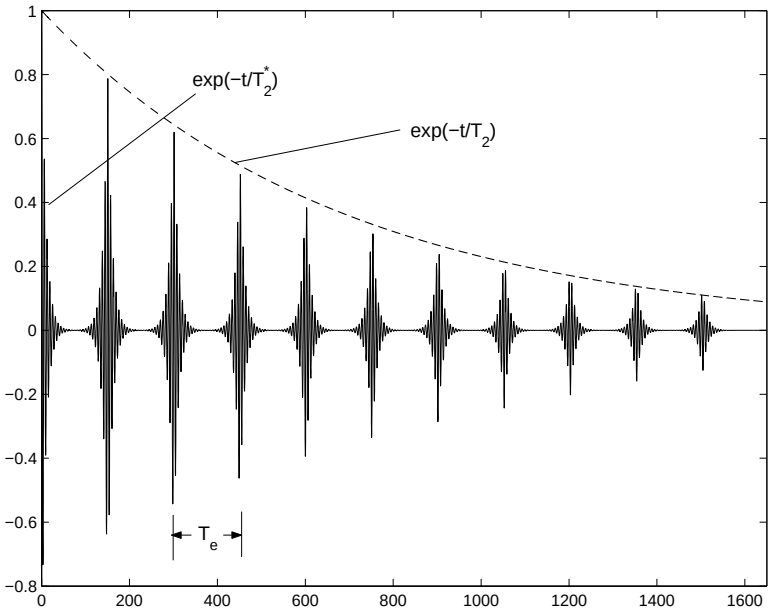

Figure 1–1: The QR response of ^{14}N in TNT

1.2 Related Previous Work

In practical applications, signal parameters oftern must be estimated in the presence of interference and noise by using array processing [3, 4]. It is well known that temporal information about the signal can be used to effectively suppress interference and noise; thus significantly improving the estimation accuracy. For example, Li and Compton [5] and Li *et al* [6] studied the estimation of directions of arrival (DOAs) of signals with known waveforms; and showed that significant improvements in accuracy, interference suppression capability, and spatial resolution can be obtained. Through Cramer-Rao bound (CRB) analysis, Zeira and Friedlander [7] studied the DOA estimation of a parametric signal; and showed that exploiting temporal information about the signal can improve the DOA estimation (except when no interference is present or when the interferences are coherent with the signal). Other studies addressed the estimation of parameters of known waveform signals [8, 9, 10, 11, 12]. But all of those studies concentrated on estimating signal parameters by exploiting temporal information only.

1.3 Contributions

We considered the two most popular approaches, the Capon and the maximum likelihood (ML) methods, that use both the temporal and spatial information to estimate

the signal amplitude in the presence of temporally white but spatially colored interference and noise. Based on the results of Reed *et al* [13] and Kelly [14], we derived closed-form expressions for the expected values and mean-squared errors (MSEs) of the two estimators. A comparative study showed that the ML estimate is unbiased while the Capon estimate is biased downwards for finite data sample lengths. We also show that both methods approach the corresponding Cramer-Rao bound (CRB). That is, they are asymptotically statistically efficient, when the number of data samples is large. However, they are not asymptotically statistically efficient when the signal-to-noise ratio (SNR) is high. We also considered a more general scenario where the interference and noise vectors are both spatially and temporally correlated. We modeled the interference and noise vector as a multichannel autoregressive (AR) random process. We proposed an alternating least squares (ALS) method to tackle the amplitude-estimation problem in this situation. We found that in most cases the ALS method is superior to the model-mismatched ML (M^3L) method that ignores the temporal correlation of the interference and noise.

1.4 Summary of Chapters

The remainder of the paper is organized as follows. Chapter 2 formulates the problem of interest. The ML and Capon estimators are also given in that section. Chapter 3 gives the closed-form expressions of the expected values and MSEs of the ML and Capon estimators; and compares their statistical properties. In Chapter 4, we propose an alternating least squares (ALS) method to deal with the more general scenario of both spatially and temporally correlated interference and noise. Our theoretical findings are verified via numerical examples in Chapter 5. Finally, Chapter 6 gives the conclusions of this thesis.

CHAPTER 2
PRELIMINARY WORK

The Capon and ML estimators are two widely used methods in array processing [4]. In this chapter, we first set up the data model and then give the Capon and ML estimators.

2.1 Data Model

We consider the problem of estimating the complex amplitude of a known waveform signal in the presence of interference and noise:

$$\mathbf{x}_l = \mathbf{a}\beta s_l + \mathbf{e}_l, \tag{2.1}$$

where $\mathbf{x}_l \in \mathcal{C}^{M\times 1}$, $l = 1, 2, \cdots, L$, denotes the lth array output vector (with M being the number of sensors and L being the number of snapshots), the array steering vector $\mathbf{a} \in \mathcal{C}^{M\times 1}$ of the signal of interest is known and β is the unknown complex amplitude of the signal whose temporal waveform $\{s_l\}_{l=1}^{L}$ is known. First we model the interference and noise term $\mathbf{e}_l \in \mathcal{C}^{M\times 1}$ as a zero-mean temporally white but spatially colored circularly symmetric complex Gaussian random process with an unknown and arbitrary, but fixed, spatial covariance matrix $\mathbf{Q}$. We define the SNR by lumping the interference and noise together in a "generalized noise" term:

$$\text{SNR} = \frac{M|\beta|^2 P_s}{\text{tr}(\mathbf{Q})}, \tag{2.2}$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix and

$$P_s = \frac{1}{L}\sum_{l=1}^{L} |s_l|^2 \tag{2.3}$$

is the average power of the known waveform. Chapter 4 extends this thinking to a case in which the interference and noise are both temporally and spatially correlated.

2.2 Capon Estimate

The Capon method, also known as the minimum variance distortionless response (MVDR) beamformer, was proposed by Capon in 1969 [15] and has been widely used in radar, sonar ([16, 17] and the references therein), speech signal processing [18] and wireless communications [19, 20, 21, 22]. The basic idea of the Capon method is to form a filter that lets the signal with steering vector $\mathbf{a}$ pass with unitary gain and minimizes the overall array output power:

$$\hat{\mathbf{w}}_{Capon} = \arg\min_{\mathbf{w}} \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} \quad \text{subject to} \quad \mathbf{w}^H \mathbf{a} = 1, \tag{2.4}$$

with $(\cdot)^H$ denoting the conjugate transpose,

$$\hat{\mathbf{R}} = \frac{1}{L} \sum_{l=1}^{L} \mathbf{x}_l \mathbf{x}_l^H. \tag{2.5}$$

By using Lagranger multiplier, it is easy to show [4]

$$\hat{\mathbf{w}}_{Capon} = \frac{\hat{\mathbf{R}}^{-1}\mathbf{a}}{\mathbf{a}^H \hat{\mathbf{R}}^{-1} \mathbf{a}}. \tag{2.6}$$

Applying Capon filter in the spatial domain and matched filter in the temporal domain, we get the Capon estimate

$$\hat{\beta}_{Capon} = \hat{\mathbf{w}}_{Capon}^H \bar{\mathbf{x}} / P_s \tag{2.7}$$

where

$$\bar{\mathbf{x}} = \frac{1}{L} \sum_{l=1}^{L} \mathbf{x}_l s_l^*. \tag{2.8}$$

and P_s is defined in (2.3). Combing (2.6) and (2.7) yields

$$\hat{\beta}_{Capon} = \frac{\mathbf{a}^H \hat{\mathbf{R}}^{-1} \bar{\mathbf{x}}}{P_s \mathbf{a}^H \hat{\mathbf{R}}^{-1} \mathbf{a}}. \tag{2.9}$$

2.3 Maximum Likelihood Estimate

The ML method estimates the signal amplitude by maximizing the likelihood function of the random vectors $\{\mathbf{x}_l\}_{l=1}^{L}$. We show in Appendix A that

$$\hat{\beta}_{ML} = \frac{\mathbf{a}^H \mathbf{T}^{-1} \bar{\mathbf{x}}}{P_s \mathbf{a}^H \mathbf{T}^{-1} \mathbf{a}}, \tag{2.10}$$

where

$$\mathbf{T} = \hat{\mathbf{R}} - \frac{\bar{\mathbf{x}}\bar{\mathbf{x}}^H}{P_s}. \tag{2.11}$$

Note that the only difference between the Capon and ML estimators is that the matrix $\hat{\mathbf{R}}$ in (2.9) is replaced by $\mathbf{T}$ in (2.10). We show in Chapter 3 that this seemingly minor difference in fact leads to significant and interesting performance differences between the two estimators.

CHAPTER 3
PERFORMANCE ANALYSIS OF MAXIMUM LIKELIHOOD AND CAPON

3.1 Introduction

In Chapter 2, we have set up the data model and proposed two methods, i.e., ML and Capon, for signal amplitude estimation. In this Chapter, we will compare the statistical performance of the two estimators thoroughly. In Section 3.2, we study the statistical properties, including the expected value and mean-squared error of the ML estimate. In Section 3.3, the statistical properties of the Capon estimate are investigated. Based on the results in Section 3.3, we modify the Capon estimate to be unbiased and study the mean-squared error of the unbiased Capon estimate in Section 3.4. Finally, we summarize our findings in Section 3.5.

3.2 Performance Analysis of ML Estimator

We present below a statistical performance analysis of the ML estimator. We prove that the ML estimator is unbiased. By deriving the MSE of the ML estimate and comparing it with the corresponding CRB, we show that the ML estimator is asymptotically statistically efficient when the number of snapshots is large, but also that this is not the case for high SNR.

3.2.1 Bias Analysis

The matrix $\mathbf{T}$ defined in (2.11) and the vector $\bar{\mathbf{x}}$ in (2.8) are both functions of the vectors $\{\mathbf{x}_l\}_{l=1}^{L}$, which might suggest that they are correlated with each other. However, the lemma below somewhat surprisingly shows that they are in fact statistically independent of each other.

Lemma 1. *Under the assumption made on the data model in (2.1), the vector* $\bar{\mathbf{x}}$ *and the matrix* $\mathbf{T}$ *are statistically independent of each other.*

Proof. See Appendix C. □

Using this lemma and the conditional expectation rule, we can easily show that the ML estimator is unbiased, i.e.,

$$\mathcal{E}\left[\hat{\beta}_{ML}\right] = \mathcal{E}_{\mathbf{T}}\left[\mathcal{E}_{\bar{\mathbf{x}}|\mathbf{T}}\left[\hat{\beta}_{ML}\right]\right] = \mathcal{E}_{\mathbf{T}}\left[\beta\right] = \beta, \tag{3.1}$$

where $\mathcal{E}_{\bar{\mathbf{x}}|\mathbf{T}}[\cdot]$ denotes calculating the expected value with respect to $\bar{\mathbf{x}}$ for a fixed $\mathbf{T}$.

Lemma 1 and its proof also help in the performance analyses that follow.

3.2.2 Mean-Squared Error Analysis

Before calculating the MSE of the ML estimate, we first introduce the best possible performance bound for any unbiased estimator of β, i.e., the CRB. Appendix B shows that the CRB has the following compact form:

$$\mathrm{CRB}(\beta) = \frac{1}{LP_s\mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}}. \tag{3.2}$$

Note that

$$\begin{aligned}
\hat{\beta}_{ML} &= \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}} \\
&= \frac{\mathbf{a}^H\mathbf{T}^{-1}\frac{1}{L}\sum_{l=1}^{L}\left(\mathbf{a}\beta s_l s_l^* + \mathbf{e}_l s_l^*\right)}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}} \\
&= \beta + \frac{\mathbf{a}^H\mathbf{T}^{-1}\frac{1}{L}\sum_{l=1}^{L}\mathbf{e}_l s_l^*}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}}.
\end{aligned} \tag{3.3}$$

That is, the error of the ML estimate of β is

$$\hat{\beta}_{ML} - \beta = \frac{\mathbf{a}^H\mathbf{T}^{-1}\frac{1}{L}\sum_{l=1}^{L}\mathbf{e}_l s_l^*}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}}. \tag{3.4}$$

Hence the MSE of the ML estimate is

$$\mathrm{MSE}(\hat{\beta}_{ML}) = \mathcal{E}\left[\frac{\mathbf{a}^H\mathbf{T}^{-1}\left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{e}_l s_l^*\right)\left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{e}_l s_l^*\right)^H\mathbf{T}^{-1}\mathbf{a}}{\left(\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}\right)^2 P_s^2}\right] \tag{3.5}$$

$$= \mathcal{E}\left[\frac{\mathbf{a}^H\mathbf{T}^{-1}\mathbf{Q}\mathbf{T}^{-1}\mathbf{a}}{\left(\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}\right)^2 LP_s}\right]. \tag{3.6}$$

To obtain (3.6) from (3.5), we have used the facts that $\mathbf{T}$ and $\sum_{l=1}^{L} \mathbf{e}_l s_l^*$ are statistically independent of each other (Lemma 1); and that $\{\mathbf{e}_l\}_{l=1}^{L}$ are independently and identically distributed.

Let

$$\zeta = \frac{\mathbf{a}^H \mathbf{T}^{-1} \mathbf{Q} \mathbf{T}^{-1} \mathbf{a} \mathbf{a}^H \mathbf{Q}^{-1} \mathbf{a}}{\left(\mathbf{a}^H \mathbf{T}^{-1} \mathbf{a}\right)^2}. \tag{3.7}$$

According to (3.6) and (3.2),

$$\mathcal{E}\left[\zeta\right] = \frac{\mathrm{MSE}(\hat{\beta}_{ML})}{\mathrm{CRB}(\beta)}. \tag{3.8}$$

As shown in the proof of Lemma 1

$$L\mathbf{T} = \sum_{l=2}^{L} \mathbf{z}_l \mathbf{z}_l^H, \tag{3.9}$$

where $\mathbf{z}_l \sim N(\mathbf{0}, \mathbf{Q})$ $l = 2, \ldots, L$, are independently and identically distributed. Hence $\mathbf{T}$ has a complex Wishart distribution [23][13][14], i.e., $\mathbf{T} \sim CW(L-1, M; \mathbf{Q}/L)$. The paper [13] has provided the probability density function (PDF) of $\rho \triangleq \frac{1}{\zeta}$:

$$f(\rho) = \frac{(L-1)!}{(M-2)!(L-M)!}(1-\rho)^{M-2}\rho^{L-M} \quad \text{for} \quad 0 \le \rho \le 1. \tag{3.10}$$

Thus

$$\begin{aligned} \mathcal{E}[\zeta] = \mathcal{E}[1/\rho] &= \frac{(L-1)!}{(M-2)!(L-M)!} \int_0^1 \rho^{L-M-1}(1-\rho)^{M-2} d\rho & (3.11) \\ &= \frac{L-1}{L-M}. & (3.12) \end{aligned}$$

Therefore, it follows from (3.8) and (3.12) that

$$\mathrm{MSE}(\hat{\beta}_{ML}) = \frac{L-1}{L-M}\mathrm{CRB}(\beta). \tag{3.13}$$

The above equation shows that (1) the ML estimate is asymptotically statistically efficient for large snapshot lengths L, which was expected, and (2) the ML estimate is not asymptotically statistically efficient for high SNR values, when L is fixed. In fact we can infer from (3.13) that the MSE (in dB)-versus-SNR (in dB) line is parallel to the

CRB-versus-SNR line and thus no "threshold effect" exits at low SNR. This theoretical result is verified via a numerical example in Section 5.

3.3 Performance Analysis of Capon Estimator

We now establish the theoretical properties of the Capon estimator via an analysis that parallels the one in Section 3.2. Our analysis relies on the results in [24, 13, 14].

3.3.1 Bias Analysis

We have proved that the ML estimate is unbiased. We investigate the bias of the Capon estimate by studying the relationship between the two estimators.

Lemma 2. *Under the assumptions made on the data model in (2.1), we have*

$$\hat{\beta}_{Capon} = \lambda \hat{\beta}_{ML}, \quad \lambda = \frac{1}{1+u} \leq 1, \tag{3.14}$$

where

$$u = \frac{1}{P_s} \left(\bar{\mathbf{x}}^H \mathbf{T}^{-1} \bar{\mathbf{x}} - \frac{\mathbf{a}^H \mathbf{T}^{-1} \bar{\mathbf{x}} \bar{\mathbf{x}}^H \mathbf{T}^{-1} \mathbf{a}}{\mathbf{a}^H \mathbf{T}^{-1} \mathbf{a}} \right). \tag{3.15}$$

and the equality holds iff $\bar{\mathbf{x}} = \alpha \mathbf{a}$ *for some constant* α.

Proof. See Appendix D. □

Lemma 3. *Under the assumption made on the data model in (2.1), any polynomial function of the* λ *defined in (3.14) is uncorrelated with the ML estimate, i.e.,*

$$\mathcal{E}\left[f(\lambda)(\hat{\beta}_{ML} - \beta) \right] = 0,$$

where $f(\lambda)$ *is an arbitrary polynomial.*

Proof. See Appendix E. □

Based on Lemmas 2 and 3, we have

$$\begin{aligned} \mathcal{E}[\hat{\beta}_{Capon}] &= \mathcal{E}[\lambda \hat{\beta}_{ML}] \\ &= \mathcal{E}[\lambda]\beta. \end{aligned} \tag{3.16}$$

We derive below the PDF of λ. Recall that $\bar{\mathbf{e}} \sim N(\mathbf{0}, \mathbf{Q}/L)$ (cf (E.2)) and $\mathbf{T} \sim CW(L-1, M; \mathbf{Q}/L)$. Following the techniques used in [13, 14], we define

$$\boldsymbol{\xi} = L^{1/2}\mathbf{Q}^{-1/2}\bar{\mathbf{e}}, \tag{3.17}$$

and

$$\mathbf{C} = L\mathbf{Q}^{-1/2}\mathbf{T}\mathbf{Q}^{-1/2}. \tag{3.18}$$

Then $\boldsymbol{\xi} \sim N(\mathbf{0}, \mathbf{I})$ and $\mathbf{C} \sim CW(L-1, M; \mathbf{I})$ [23].

Inserting (3.17) and (3.18) into (E.3) yields

$$u = \boldsymbol{\xi}^H \left(\mathbf{C}^{-1} - \frac{\mathbf{C}^{-1}\mathbf{d}\mathbf{d}^H\mathbf{C}^{-1}}{\mathbf{d}^H\mathbf{C}^{-1}\mathbf{d}} \right) \boldsymbol{\xi}, \tag{3.19}$$

where

$$\mathbf{d} = L^{1/2}\mathbf{Q}^{-1/2}\mathbf{a}. \tag{3.20}$$

We next introduce an $M \times M$ unitary matrix $\tilde{\mathbf{U}}$ such that

$$\mathbf{t} = \tilde{\mathbf{U}}\mathbf{d} = [\|\mathbf{d}\|, \mathbf{0}_{1\times M-1}]^T. \tag{3.21}$$

Note that $\tilde{\mathbf{U}}\boldsymbol{\xi}$ and $\tilde{\mathbf{U}}\mathbf{C}\tilde{\mathbf{U}}^H$ have the same PDFs as $\boldsymbol{\xi}$ and $\mathbf{C}$, respectively. Then (3.19) can be further simplified as

$$u = \boldsymbol{\xi}^H\mathbf{C}^{-1}\boldsymbol{\xi} - \boldsymbol{\xi}^H\frac{\mathbf{C}^{-1}\mathbf{t}\mathbf{t}^H\mathbf{C}^{-1}}{\mathbf{t}^H\mathbf{C}^{-1}\mathbf{t}}\boldsymbol{\xi}. \tag{3.22}$$

We partition $\mathbf{C}$, $\mathbf{C}^{-1}$ and $\boldsymbol{\xi}$ as

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \tag{3.23}$$

$$\mathbf{C}^{-1} = \begin{pmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{pmatrix} \tag{3.24}$$

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{3.25}$$

where C_{11}, C^{11} and ξ_1 are scalars, $C_{21} = C_{12}^H, C^{21} = [C^{12}]^H$ and ξ_2 are $(M-1)\times 1$ vectors, and both C_{22} and C^{22} are $(M-1)\times(M-1)$ matrices. Some useful relationships are as follows [13, 14]:

$$C^{11} = (C_{11} - C_{12}C_{22}^{-1}C_{21})^{-1} \tag{3.26}$$

$$C^{21} = -C_{22}^{-1}C_{21}C^{11} \tag{3.27}$$

$$\begin{aligned} C^{22} &= \left(C_{22} - C_{11}^{-1}C_{21}C_{12}\right)^{-1} && (3.28)\\ &= C_{22}^{-1} + \frac{C_{22}^{-1}C_{21}C_{12}C_{22}^{-1}}{C_{11} - C_{12}C_{22}^{-1}C_{21}}. && (3.29) \end{aligned}$$

To obtain (3.29) from (3.28), we have used the matrix inversion lemma.

Based on the partitions defined in (3.23), (3.24), and (3.25), the equation (3.22) can be written as

$$\begin{aligned} u &= \xi_1^* C^{11}\xi_1 + \xi_1^* C^{12}\xi_2 + \xi_2^H C^{21}\xi_1 + \xi_2^H C^{22}\xi_2 - \frac{|\xi_1^* C^{11} + \xi_{21}^H C^{21}|^2}{C^{11}} \\ &= \xi_1^* C^{11}\xi + \xi_1^* C^{12}\xi_2 + \xi_2^H C^{21}\xi_1 + \xi_2^H C^{22}\xi_2 - |\xi_1^* - \xi_2^H C_{22}^{-1}C_{21}|^2 C^{11} \\ &= \xi_2^H C^{22}\xi_2 - \xi_2^H C_{22}^{-1}C_{21}C^{11}C_{12}C_{22}^{-1}\xi_2 && (3.30)\\ &= \xi_2^H C^{22}\xi_2 - \frac{\xi_2^H C_{22}^{-1}C_{21}C_{12}C_{22}^{-1}\xi_2}{C_{11} - C_{12}C_{22}^{-1}C_{21}} && (3.31)\\ &= \xi_2^H C_{22}^{-1}\xi_2 && (3.32) \end{aligned}$$

Hence we have

$$\lambda = \frac{1}{1 + \xi_2^H C_{22}^{-1}\xi_2} \tag{3.33}$$

where $\xi_2 \sim N(\mathbf{0}, \mathbf{I})$ and $C_{22} \sim CW(L-1, M-1; \mathbf{I})$ and they are statistically independent of each other. It can been shown ([14]) that the PDF of λ is

$$f(\lambda) = \frac{(L-1)!}{(L-M)!(M-2)!}(1-\lambda)^{M-2}\lambda^{L-M} \quad \text{for} \quad 0 \le \lambda \le 1 \tag{3.34}$$

and the first and second moments of this distribution are

$$\begin{aligned} \mathcal{E}[\lambda] &= \frac{L-M+1}{L} && (3.35)\\ \mathcal{E}[\lambda^2] &= \frac{(L-M+1)(L-M+2)}{L(L+1)}. && (3.36) \end{aligned}$$

Thus it follows from (3.35) and (3.16) that

$$\mathcal{E}[\hat{\beta}_{Capon}] = \frac{L-M+1}{L}\beta. \tag{3.37}$$

The above equation shows that the Capon estimate of β is biased downwards for a finite data sample number L while it is asymptotically unbiased for large L. We can also see that the biasedness of the Capon estimate is not related to SNR. We note that the result in (3.37) is not completely new. The papers [25] and [26] have given results similar to (3.37) when studying the convergence rate of the Capon beamformer, which was used to estimate the signal power and waveform. However, our assumptions are different from those in [25, 26], where the signals were assumed to be independently and identically distributed complex Gaussian random vectors.

3.3.2 Mean-Squared Error Analysis

Based on Lemma 2, we can obtain the MSE($\hat{\beta}_{Capon}$) as:

$$\begin{aligned}
\text{MSE}(\hat{\beta}_{Capon}) &= \mathcal{E}\left[|\beta - \lambda\hat{\beta}_{ML}|^2\right] && (3.38)\\
&= \mathcal{E}\left[|(1-\lambda)\beta - \lambda(\hat{\beta}_{ML}-\beta)|^2\right] && (3.39)\\
&= \mathcal{E}\left[\lambda^2|\hat{\beta}_{ML}-\beta|^2\right] + \mathcal{E}\left[(1-\lambda)^2\right]|\beta|^2, && (3.40)
\end{aligned}$$

where to obtain (3.40) from (3.39), we have used Lemma 3 and (3.1).

Combing (3.4) and (E.2) gives

$$\hat{\beta}_{ML} - \beta = \frac{\mathbf{a}^H\mathbf{T}^{-1}\sqrt{P_s}\bar{\mathbf{e}}}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}}. \tag{3.41}$$

Using the notations defined in (3.17)-(3.29), we can rewrite (3.41) as

$$\begin{aligned}
\hat{\beta}_{ML} - \beta &= \frac{\mathbf{t}^H\mathbf{C}^{-1}\boldsymbol{\xi}}{\sqrt{P_s}\mathbf{t}^H\mathbf{C}^{-1}\mathbf{t}} && (3.42)\\
&= \frac{C^{11}\xi_1 + C^{12}\xi_2}{\sqrt{LP_s\mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}}C^{11}} && (3.43)\\
&= \frac{\xi_1 - C_{12}C_{22}^{-1}\xi_2}{\sqrt{LP_s\mathbf{a}\mathbf{Q}^{-1}\mathbf{a}}}. && (3.44)
\end{aligned}$$

Thus

$$\begin{aligned}
\mathcal{E}\left[\lambda^2|\hat{\beta}_{ML}-\beta|^2\right] &= \mathcal{E}\left[\left|\frac{\xi_1 - C_{12}C_{22}^{-1}\xi_2}{1+\xi_2^H C_{22}^{-1}\xi_2}\right|^2\right]\mathrm{CRB}(\beta) && (3.45)\\
&= \mathcal{E}\left[\frac{|\xi_1|^2+\xi_2^H C_{22}^{-1}C_{21}C_{12}C_{22}^{-1}\xi_2}{(1+\xi_2^H C_{22}^{-1}\xi_2)^2}\right]\mathrm{CRB}(\beta) && (3.46)\\
&= \mathcal{E}\left[\frac{1+\xi_2^H C_{22}^{-1}\xi_2}{(1+\xi_2^H C_{22}^{-1}\xi_2)^2}\right]\mathrm{CRB}(\beta) && (3.47)\\
&= \mathcal{E}\left[\lambda\right]\mathrm{CRB}(\beta) && (3.48)\\
&= \frac{L-M+1}{L}\mathrm{CRB}(\beta). && (3.49)
\end{aligned}$$

We got (3.47) from (3.46) using the fact that

$$\mathcal{E}_{C_{21}|C_{22}}\left[C_{21}C_{12}\right] = C_{22}. \tag{3.50}$$

It follows from (3.35) and (3.36) that the second term of (3.40) is

$$\begin{aligned}
\mathcal{E}\left[(1-\lambda)^2\right]|\beta|^2 &= \left[1-\frac{2(L-M+1)}{L}+\frac{(L-M+1)(L-M+2)}{L(L+1)}\right]|\beta|^2 && (3.51)\\
&= \frac{M^2-M}{L^2+L}|\beta|^2. && (3.52)
\end{aligned}$$

Thus

$$\mathrm{MSE}(\hat{\beta}_{Capon}) = \frac{L-M+1}{L}\mathrm{CRB}(\beta)+\frac{M^2-M}{L^2+L}|\beta|^2. \tag{3.53}$$

We can conclude from (3.53) that the Capon estimate is asymptotically statistically efficient for a large data sample number L. In the finite L case, the MSE of the Capon estimate may be smaller than the CRB for any unbiased estimator if the first term in (3.53) dominates the second one. However, the MSE of the Capon estimate has an error floor equal to $\frac{M^2-M}{L^2+L}|\beta|^2$ at high SNR.

3.4 Performance Analysis of Unbiased Capon Estimator

We infer from equation (3.37) that the error floor of the MSE of the Capon estimate $\frac{M^2-M}{L^2+L}|\beta|^2$ is at least partly caused by the biasedness. Since the biasedness of Capon is determined only by the dimensionality of the data set, we can get a modified Capon

estimate [1]

$$\hat{\beta}^u_{Capon} = \frac{L}{L-M+1}\hat{\beta}_{Capon}. \tag{3.54}$$

whose expected value

$$\mathcal{E}[\hat{\beta}^u_{Capon}] = \beta. \tag{3.55}$$

Now we derive the MSE of the unbiased Capon estimate.

$$\begin{aligned} \mathrm{MSE}(\hat{\beta}^u_{Capon}) &= \mathcal{E}\left[\left|\hat{\beta}^u_{Capon} - \beta\right|^2\right] \\ &= \left(\frac{L}{L-M+1}\right)^2 \mathcal{E}\left[\left|\hat{\beta}_{Capon} - \frac{L-M+1}{L}\beta\right|^2\right]. \end{aligned} \tag{3.56}$$

On the other hand,

$$\begin{aligned} \mathrm{MSE}(\hat{\beta}_{Capon}) &= \mathcal{E}\left[\left|\hat{\beta}_{Capon} - \frac{L-M+1}{L}\beta - \frac{M-1}{L}\beta\right|^2\right] \\ &= \mathcal{E}\left[\left|\hat{\beta}_{Capon} - \frac{L-M+1}{L}\beta\right|^2\right] + \left|\frac{M-1}{L}\beta\right|^2. \end{aligned} \tag{3.57}$$

Thus

$$\mathcal{E}\left[\left|\hat{\beta}_{Capon} - \frac{L-M+1}{L}\beta\right|^2\right]$$

$$= \mathrm{MSE}(\hat{\beta}_{Capon}) - \left|\frac{M-1}{L}\beta\right|^2 \tag{3.58}$$

$$= \frac{L-M+1}{L}\mathrm{CRB}(\beta) + \frac{M^2-M}{L^2+L}|\beta|^2 - \left|\frac{M-1}{L}\beta\right|^2 \tag{3.59}$$

$$= \frac{L-M+1}{L}\left[\mathrm{CRB}(\beta) + \frac{M-1}{L(L+1)}|\beta|^2\right]. \tag{3.60}$$

Inserting (3.60) into (3.56) yields

$$\mathrm{MSE}(\hat{\beta}^u_{Capon}) = \frac{L}{L-M+1}\mathrm{CRB}(\beta) + \frac{M-1}{(L-M+1)(L+1)}|\beta|^2. \tag{3.61}$$

(3.61) shows that error floor of the MSE of the unbiased Capon estimate still exists though it is smaller than that of the biased one.

[1] We use subscript u to denote unbiased estimate

It is interesting to compare the MSEs of the ML estimate and the unbiased Capon estimate. Rewrite (3.13):

$$\mathrm{MSE}(\hat{\beta}_{ML}) = \frac{L-1}{L-M}\mathrm{CRB}(\beta). \tag{3.62}$$

Unbiased Capon estimate can yield a smaller MSE than that of ML.

$$\mathrm{MSE}(\hat{\beta}^u_{Capon}) < \mathrm{MSE}(\hat{\beta}_{ML}) \tag{3.63}$$

$$\Leftrightarrow \quad \frac{L}{L-M+1}\mathrm{CRB}(\beta) + \frac{M-1}{(L-M+1)(L+1)}|\beta|^2 < \frac{L-1}{L-M}\mathrm{CRB}(\beta) \tag{3.64}$$

$$\Leftrightarrow \quad \mathrm{CRB} > \frac{L-M}{L+1}|\beta|^2. \tag{3.65}$$

However, for this case

$$\mathrm{MSE}(\hat{\beta}^u_{Capon}) > \frac{L-1}{L+1}|\beta|^2, \tag{3.66}$$

i.e., both methods fail in this situation. Hence, ML is always a preferable choice over Capon.

3.5 Summary of the Capon and ML Statistical Properties

- The ML estimate is unbiased.
- The ML estimate is asymptotically statistically efficient for large number of snapshots.
- The ML estimate is *not* asymptotically statistically efficient for high SNR.
- The Capon estimate is biased downwards.
- The Capon estimate is asymptotically unbiased and statistically efficient for large number of snapshots.
- The Capon estimate is *neither* asymptotically unbiased *nor* asymptotically statistically efficient for high SNR.

The above properties are summarized in Table 3–1.

Table 3–1: M sensors, L snapshots, and signal amplitude β

	ML	Capon	Unbiased Capon
Mean	β	$\frac{L-M+1}{L}\beta$	β
MSE	$\frac{L-1}{L-M}\mathrm{CRB}$	$\frac{L-M+1}{L}\mathrm{CRB} + \frac{M^2-M}{L^2+L}\lvert\beta\rvert^2$	$\frac{L}{L-M+1}\mathrm{CRB} + \frac{M-1}{(L-M+1)(L+1)}\lvert\beta\rvert^2$

CHAPTER 4
EXTENSION TO MULTICHANNEL AR INTERFERENCE AND NOISE

The previous study assumed that the interference and noise term in (2.1) is spatially colored but temporally white. However the interference and noise can be temporally correlated [3, 27]. In this chapter, we model the interference and noise vector as a multichannel autoregressive (AR) random process and propose an alternating least squares (ALS) method based on the cyclic optimization approach [28]. For discussion purposes, we refer to the ML method in Section II that ignores the temporal correlation of the interference and noise as the model-mismatched ML (M^3L) method.

4.1 Data Model

Consider the data model:

$$\mathbf{x}_l = \mathbf{a}\beta s_l + \mathbf{v}_l, \quad l = 1, 2, \ldots, L, \tag{4.1}$$

which is the same as the one in (2.1) except that the interference and noise term now satisfies the next AR equation [29]

$$\mathbf{A}(z^{-1})\mathbf{v}_l = \mathbf{e}_l, \tag{4.2}$$

where z^{-1} is the unit delay operator,

$$\mathbf{A}(z^{-1}) = \mathbf{I} + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} + \ldots + \mathbf{A}_p z^{-p}, \tag{4.3}$$

and

$$\mathcal{E}[\mathbf{e}_l \mathbf{e}_n^H] = \mathbf{Q}\delta_{ln}, \tag{4.4}$$

where δ_{ln} denotes the Kronecker delta function:

$$\delta_{ln} = \begin{cases} 1 & l = n \\ 0 & l \neq n \end{cases}. \tag{4.5}$$

Note that if the interference component in $\mathbf{v}_l$ is a multichannel AR process while the noise component in $\mathbf{v}_l$ is white temporally, then the interference and noise term will be a multichannel autoregressive and moving average (ARMA) random process, which can still be approximated by a multichannel AR process. The SNR for the data model in (4.1) is defined as

$$\text{SNR} = \frac{MP_s|\beta|^2}{\text{tr}\,(\mathbf{R_v})}, \tag{4.6}$$

where $\mathbf{R_v}$ is the covariance matrix of $\{\mathbf{v}_l\}$.

4.2 The ALS Algorithm

Conditioned on the first p data vectors $\{\mathbf{x}_l\}_{l=1}^{p}$, the log-likelihood function is proportional to

$$C_1 = -(L-p)\ln|\mathbf{Q}| - \text{tr}\left[\mathbf{Q}^{-1}\sum_{l=p+1}^{L}\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]^H\right]. \tag{4.7}$$

Maximizing the above function with respect to $\mathbf{Q}$ gives [30]

$$\hat{\mathbf{Q}} = \frac{1}{L-p}\sum_{l=p+1}^{L}\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]^H. \tag{4.8}$$

After substituting (4.8) into (4.7), we need to minimize

$$C_2 = \left|\sum_{l=p+1}^{L}\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]^H\right| \tag{4.9}$$

with respect to both β and $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p]$. Hence the optimization problem becomes more complicated than the one in Section 2. Here we propose an alternating least squared (ALS) approach to solve this problem.

To begin with, we obtain an initial estimate $\hat{\beta}^{(0)}$ of β by using the M^{3}L method (cf (2.10)). For a given estimate $\hat{\beta}$, let $\mathbf{z}_l = \mathbf{x}_l - \mathbf{a}\hat{\beta}s_l$. From (4.9), we get

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}}\left|\sum_{l=p+1}^{L}\left[\mathbf{A}(z^{-1})\mathbf{z}_l\right]\left[\mathbf{A}(z^{-1})\mathbf{z}_l\right]^H\right| \tag{4.10}$$

where

$$\mathbf{A}(z^{-1})\mathbf{z}_l = \mathbf{z}_l + [\mathbf{A}_1 \dots \mathbf{A}_p] \begin{bmatrix} \mathbf{z}_{l-1} \\ \vdots \\ \mathbf{z}_{l-p} \end{bmatrix} \triangleq \mathbf{z}_l + \mathbf{A}\boldsymbol{\phi}_l.$$

Note that

$$\begin{aligned}
& \sum_{l=p+1}^{L} \left[\mathbf{A}(z^{-1})\mathbf{z}_l\right] \left[\mathbf{A}(z^{-1})\mathbf{z}_l\right]^H \\
= & \sum_{l=p+1}^{L} \left[\mathbf{z}_l + \mathbf{A}\boldsymbol{\phi}_l\right] \left[\mathbf{z}_l + \mathbf{A}\boldsymbol{\phi}_l\right]^H \\
= & \sum_{l=p+1}^{L} \mathbf{z}_l\mathbf{z}_l^H + \sum_{l=p+1}^{L} \mathbf{z}_l\boldsymbol{\phi}_l^H\mathbf{A}^H + \mathbf{A}\sum_{l=p+1}^{L} \boldsymbol{\phi}_l\mathbf{z}_l^H + \mathbf{A}\left(\sum_{l=p+1}^{L} \boldsymbol{\phi}_l\boldsymbol{\phi}_l^H\right)\mathbf{A}^H \\
\triangleq & \hat{\mathbf{R}}_{\mathbf{z}} + \hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}\mathbf{A}^H + \mathbf{A}\hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}^H + \mathbf{A}\hat{\mathbf{R}}_{\boldsymbol{\phi}}\mathbf{A}^H \\
= & \left(\mathbf{A} + \hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}\hat{\mathbf{R}}_{\boldsymbol{\phi}}^{-1}\right)\hat{\mathbf{R}}_{\boldsymbol{\phi}}\left(\mathbf{A} + \hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}\hat{\mathbf{R}}_{\boldsymbol{\phi}}^{-1}\right)^H + \hat{\mathbf{R}}_{\mathbf{z}} - \hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}\hat{\mathbf{R}}_{\boldsymbol{\phi}}^{-1}\hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}^H. \qquad (4.11)
\end{aligned}$$

Hence the solution to (4.10) is

$$\hat{\mathbf{A}} = -\hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}\hat{\mathbf{R}}_{\boldsymbol{\phi}}^{-1}, \qquad (4.12)$$

which is recognized as the multichannel Prony estimate of $\mathbf{A}$ [29]. We assume that the order p of the multichannel random process AR(p) is known. If p is unknown, it can be estimated, for instance, by using the Generalized Akaike Information Criterion (GAIC) [30].

For a given $\hat{\mathbf{A}}$, we obtain an improved estimate of β as follows:

$$\hat{\beta} = \arg\min_{\beta} \left| \sum_{l=p+1}^{L} \left[\hat{\mathbf{A}}(z^{-1})\left(\mathbf{x}_l - \mathbf{a}\beta s_l\right)\right] \left[\hat{\mathbf{A}}(z^{-1})\left(\mathbf{x}_l - \mathbf{a}\beta s_l\right)\right]^H \right|. \qquad (4.13)$$

First, we consider the case of a known damped (or undamped) sinusoidal signal, i.e., $s_l = e^{(-\alpha_s + j\omega_s)l}$ with known frequency ω_s and damping factor α_s.

Let

$$\mathbf{y}_l = \hat{\mathbf{A}}(z^{-1})\mathbf{x}_l, \quad l = p+1, \dots, L, \quad \mathbf{b} = \hat{\mathbf{A}}(z^{-1})|_{z=e^{-\alpha_s + j\omega_s}}\mathbf{a}.$$

Note that the length of the new data sequence $\mathbf{y}_l,\ l = p+1, \ldots, L$, is $L-p$ instead of L. The solution to the above problem is given by the ML estimator proposed in Section 2:

$$\hat{\beta} = \frac{\mathbf{b}^H \mathbf{T}_y^{-1} \bar{\mathbf{y}}}{P_s \mathbf{b}^H \mathbf{T}_y^{-1} \mathbf{b}} \tag{4.14}$$

where

$$\mathbf{T}_y = \frac{1}{L-p} \sum_{l=p+1}^{L} \mathbf{y}_l \mathbf{y}_l^H - \frac{\bar{\mathbf{y}} \bar{\mathbf{y}}^H}{P_s}, \qquad \bar{\mathbf{y}} = \frac{1}{L-p} \sum_{l=p+1}^{L} \mathbf{y}_l s_l^* \tag{4.15}$$

and $P_s = \frac{1}{L-p} \sum_{l=1}^{L-p} |s_l|^2$ is the average power of the known waveform $\{s_l\}_{l=p+1}^{L}$. The ALS approach maximizes the likelihood function cyclically. We set $\hat{\mathbf{A}}^{(0)} = \mathbf{0}$ and obtain $\hat{\beta}^{(0)} = \hat{\beta}_{\mathrm{M^3L}}$. We then iterate the next two steps until the solution converges, i.e., until the two consecutive estimates $\hat{\beta}^{(i)}$ and $\hat{\beta}^{(i+1)}$ are sufficiently close:

$$\hat{\mathbf{A}}^{(i+1)} = \arg\max_{\mathbf{A}} f(\mathbf{x}|\mathbf{A}; \hat{\beta}^{(i)}, \{\mathbf{x}_l\}_{l=1}^{p}), \tag{4.16}$$

which is given by (4.12) with $\hat{\beta}$ replaced by $\hat{\beta}^{(i)}$, and

$$\hat{\beta}^{(i+1)} = \arg\max_{\beta} f(\mathbf{x}|\beta; \hat{\mathbf{A}}^{(i+1)}, \{\mathbf{x}_l\}_{l=1}^{p}), \tag{4.17}$$

which is given by (4.14) with $\hat{\mathbf{A}}$ replaced by $\hat{\mathbf{A}}^{(i+1)}$.

Obviously the likelihood function never decreases in any iteration. In the simulations reported in the next section we found that ALS converges in 2 or 3 iterations. Hence the ALS estimator is computationally quite efficient.

Next, we consider the case of an arbitrary known waveform signal. Let

$$\mathbf{y}_l = \hat{\mathbf{A}}(z^{-1})\mathbf{x}_l, \quad \mathbf{a}_l = \hat{\mathbf{A}}(z^{-1})\mathbf{a}s_l, \quad l = p+1, \ldots, L$$

$$\mathbf{Y}_{M\times(L-p)} = [\mathbf{y}_{p+1}\mathbf{y}_{p+2}\ldots\mathbf{y}_L] \quad \text{and} \quad \mathbf{G}_{M\times(L-p)} = [\mathbf{a}_{p+1}\mathbf{a}_{p+2}\ldots\mathbf{a}_L].$$

Also let $\mathbf{P}$ be an orthogonal projection matrix defined as

$$\mathbf{P} = \mathbf{G}^H (\mathbf{G}^H)^{\dagger}, \tag{4.18}$$

where $(\mathbf{G}^H)^{\dagger}$ is the Moore-Penrose pseudo-inverse of $\mathbf{G}^H$ [31], and let $\mathbf{P}^{\perp} = \mathbf{I} - \mathbf{P}$.

Then (4.13) can be written concisely in a matrix form:

$$\begin{aligned}
\hat{\beta} &= \arg\min_{\beta} \left|(\mathbf{Y}-\beta\mathbf{G})(\mathbf{Y}-\beta\mathbf{G})^H\right| \\
&= \arg\min_{\beta} \left|(\mathbf{Y}-\beta\mathbf{G})(\mathbf{P}+\mathbf{P}^{\perp})(\mathbf{Y}-\beta\mathbf{G})^H\right| && (4.19)\\
&= \arg\min_{\beta} \left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H+\mathbf{YP}^{\perp}\mathbf{Y}^H\right| && (4.20)\\
&\triangleq \arg\min_{\beta} \left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H+\mathbf{T}_y\right| && (4.21)\\
&= \arg\min_{\beta} \left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H\mathbf{T}_y^{-1}+\mathbf{I}\right|\left|\mathbf{T}_y\right|, && (4.22)
\end{aligned}$$

Note that minimizing the cost function in (4.22) requires a two-dimensional search (since β is complex-valued). To avoid the search, we use Lemma 4 to obtain an approximate estimate of β.

Lemma 4. *For a large data sample number L, minimizing*

$$F_1 = \left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H\mathbf{T}_y^{-1}+\mathbf{I}\right|, \tag{4.23}$$

is asymptotically equivalent to minimizing

$$F_2 = \operatorname{tr}\left[(\mathbf{YP}-\beta\mathbf{G})^H\mathbf{T}_y^{-1}(\mathbf{YP}-\beta\mathbf{G})\right]. \tag{4.24}$$

Proof. See Appendix F. □

It follows from (4.24) that

$$F_2 = |\beta|^2\operatorname{tr}(\mathbf{G}^H\mathbf{T}_y^{-1}\mathbf{G}) - \operatorname{tr}(\mathbf{G}^H\mathbf{T}_y^{-1}\mathbf{Y})\beta^* - \beta\operatorname{tr}(\mathbf{Y}^H\mathbf{T}_y^{-1}\mathbf{G}) + \operatorname{tr}(\mathbf{PY}^H\mathbf{T}_y^{-1}\mathbf{Y}). \tag{4.25}$$

Minimizing F_2 with respect to β yields

$$\hat{\beta} \approx \frac{\operatorname{tr}(\mathbf{G}^H\mathbf{T}_y^{-1}\mathbf{Y})}{\operatorname{tr}(\mathbf{G}^H\mathbf{T}_y^{-1}\mathbf{G})}, \tag{4.26}$$

where we remind the reader that $\mathbf{T}_y = \mathbf{YP}^{\perp}\mathbf{Y}^H$. Because (4.26) is only an approximate solution to (4.22) in this more general case, ALS is not theoretically guaranteed to yield a more accurate solution than the M^{3}L method. However, in our numerical examples, ALS outperforms M^{3}L in most cases even for modest data sample lengths. To avoid any

"divergence problem" in this case in which ALS is no longer an iterative maximizer, we simply pre-impose the number of iterations to be 3.

CHAPTER 5
NUMERICAL AND EXPERIMENTAL EXAMPLES

We provide both simulated-data and real-life data examples to demonstrate the performance the ML and Capon estimates. In all the simulated-data examples, we consider the case where the steering vector is given by $\mathbf{a} = [1\ 0\ 0\ 0]^T$ and $\beta = 1$. This corresponds to the case where the first of the $M = 4$ sensors receives the signal besides the interference and noise while the other three sensors receive the interference and noise only. In all but the last example, we assume that $s_l = 1$, $l = 1, 2, \cdots, L$, for simplicity. We obtain the empirical MSEs of the estimates by using 500 Monte-Carlo trials.

5.1 Spatially Colored but Temporally White Interference and Noise

We first consider simulated-data examples. We assume that the interference and noise term is a spatially colored but temporally white Gaussian random vector with the spatial covariance matrix $\mathbf{Q}$ given by

$$[\mathbf{Q}]_{ij} = \rho \cdot (0.9)^{|i-j|}, \tag{5.1}$$

where $\rho = \frac{1}{\mathrm{SNR}}$.

Figure 5–1 shows the MSEs of ML estimates and the biased and the unbiased Capon (cf. (3.54)) estimates obtained from both theoretical predictions (based on (3.13) (3.53)) and (3.61) and Monte Carlo trials besides the corresponding CRB as a function of L when the SNR is 10 dB. As expected, both the ML and the two Capon estimates approach the corresponding CRB as L increases since both methods are asymptotically statistically efficient for large L. It is also shown that the convergence rate of the ML estimate is much faster than the two Capon estimates and the unbiased Capon estimate is superior to the biased Capon. Figure 5–2 gives the MSEs of the Capon and ML estimates besides the corresponding CRB as a function of SNR when $L = 10$. The unbiased Capon estimate has a lower error floor than the biased one. Note also that the biased Capon estimate

can have lower MSE than the unbiased ML estimate at low SNR (yet this happens at MSE values that are too large to be of practical value). As predicted by our theoretical analyses, for a fixed data length L, MSE($\hat{\beta}_{ML}$) is parallel to the CRB and no "threshold effect" occurs. Furthermore, the ML estimate is not asymptotically statistically efficient for high SNR.

Next we present a real-life data example based on experimentally measured QR data. The main antenna of a QR landmine detector receives a QR signal which consists of 40 echoes besides AM/FM interferences. We apply a fast Fourier transform (FFT) to each echo and only pick the value corresponding to the echo frequency ω_e. In this way, we compress the QR signal into a signal with known waveform $s_l = e^{-lT_e/T_2},\ l = 1, \ldots, 40$. Next the data received at the 3 reference antennas is segmented into 40 blocks which occupy the same period of time as the 40 echoes. We apply FFT to each block and pick 3 values corresponding to ω_e and the two adjacent frequency bins. Doing so, we get a virtual array with one main antenna and 9 reference antennas, i.e.,$\mathbf{a} = [1\ \mathbf{0}_{1\times 9}]^T$. Although the aforementioned pre-processing method might seem somewhat *ad hoc*, it worked well in our experiments. Figure 5–3 shows the ML and th biased Capon estimates of the QR signal amplitude in 30 experimental trials. Since we do not know the true value of the signal amplitude, we cannot compare the MSEs of the two estimators. Nevertheless we let

$$\epsilon(\hat{\beta}) = \frac{\sigma[\hat{\beta}]}{\mathcal{E}[\hat{\beta}]} \tag{5.2}$$

where $\sigma[\hat{\beta}]$ and $\mathcal{E}[\hat{\beta}]$ are the empirical standard deviation and the mean of $\hat{\beta}$ (over the 30 trials). A small $\epsilon[\hat{\beta}]$ is desirable for signal detection. Based on the 30 trials, we get $\epsilon(\hat{\beta}_{ML}) = 0.1746$ which is smaller than $\epsilon(\hat{\beta}_{Capon}) = 0.2039$. We don't need to differentiate the biased and unbiased Capon estimate for the quantity ϵ since it is invariant to scaling.

5.2 Both Spatially and Temporally Colored Interference and Noise

We now consider the case of spatially and temporally correlated interference and noise. We generate a multichannel AR(2) random process with the method in [27]. The

autocorrelation matrices are given by

$$[\mathbf{R}_{\mathbf{v}}(l)]_{mn} = \mathcal{E}[\mathbf{v}_t \mathbf{v}_{t-l}^H] = \rho \rho_s^{|m-n|} \exp\{-\rho_t l^2 + j(m-n+l)\omega\}, \tag{5.3}$$

and

$$\mathbf{R}_{\mathbf{v}}(l) = \mathbf{R}_{\mathbf{v}}{}^H(-l), \quad l = 0, 1, \ldots, p, \tag{5.4}$$

where $\rho = \frac{1}{\mathrm{SNR}}$, ρ_s controls the spatial correlation, ρ_t partly decides the temporal correlation, and ω defines the spectral peak location of the colored interference and noise in each channel. The data sample number is $L = 50$. When we use the true autoregressive matrix $\mathbf{A}$ in the ALS instead of the estimated one, we refer to the method as the known-AR ML (KML) approach. We include KML for comparison purposes only. Note that unlike the temporally white interference and noise case discussed previously, the performance of the M^{3}L and ALS estimators depends on the temporal frequency characteristics of the known signal. The simulations are performed for both a constant signal and a known BPSK signal.

First, we consider the relationship between the cost function C_2 defined in (4.9) and β. Because $\mathbf{A}(z^{-1})$ can be concentrated out by using its estimate given in (4.12), C_2 is a function of β only. Consider the constant signal case. Figure 5–4 shows the mesh plot of $-C_2$ versus the real and imaginary part of $\hat{\beta}$. We can see that only one local maximum exits around the true value of $\beta = 1$.

For the constant signal case, our simulations show that the spatial correlation coefficient ρ_s is not closely related to the gap between the performance of M^{3}L and ALS. However, the temporal correlation coefficient ρ_t and the position of the spectral peak ω have an impact on the relative performance of the two methods (Figures 5–5 and 5–6). We summarize our observations as follows:

A: Both ALS and M^{3}L work better for large ω and/or small ρ_t,

B: ALS is slightly worse than M^{3}L for small ω and/or large ρ_t,

C: ALS is significantly better than M^{3}L for large ω and/or small ρ_t.

To explain these observations, we examine the signal besides the interference and noise term in the temporal frequency domain. The signal is a constant and hence has power at zero frequency only. The power of the interference and noise is concentrated around ω especially for small ρ_t. For large ω, the signal is separated from the interference and noise in the temporal frequency domain, which benefits both methods. Similarly, smaller ρ_t means higher correlation in the temporal domain or more peaky spectra in the temporal frequency domain. Hence both estimators perform better for this case when ω is away from zero. This explains Observation A. Next, we note that a large ρ_t means low correlation in the temporal domain and hence the interference and noise vector is approximately temporally white. For small ω, the signal and the interference and noise terms are not well separated in the temporal frequency domain. KML behaves approximately as M^3L in either case. Since ALS is inferior to KML, ALS is also slightly worse than M^3L. This explains Observation B. Observation C was expected since ALS estimates the temporal correlation of the interference and noise and can suppress the interference and noise more efficiently in this case in which the temporal correlation is significant.

Finally, we consider the known BPSK signal case. Because a BPSK signal is wideband in the temporal frequency domain, the impact of the interference spectral peak location ω on the performance of the two methods is not as significant as in the constant signal case, which was verified in our simulations. However, the temporal correlation parameter ρ_t still controls the relative performance of the two methods as shown in Figure 5–7. We also see from Figure 5–7 that the ALS method significantly outperforms M^3L (by over 10 dB in SNR) even for modestly temporally correlated interference and noise ($\rho_t = 0.1$) although it performs similarly to M^3L when the temporal correlation of the interference and noise is weak ($\rho_t = 2$). Our simulations also suggest that a known wideband signal makes suppressing temporally correlated interference and noise easier than a narrowband one, in the sense that better estimates of β can be obtained in the wideband case.

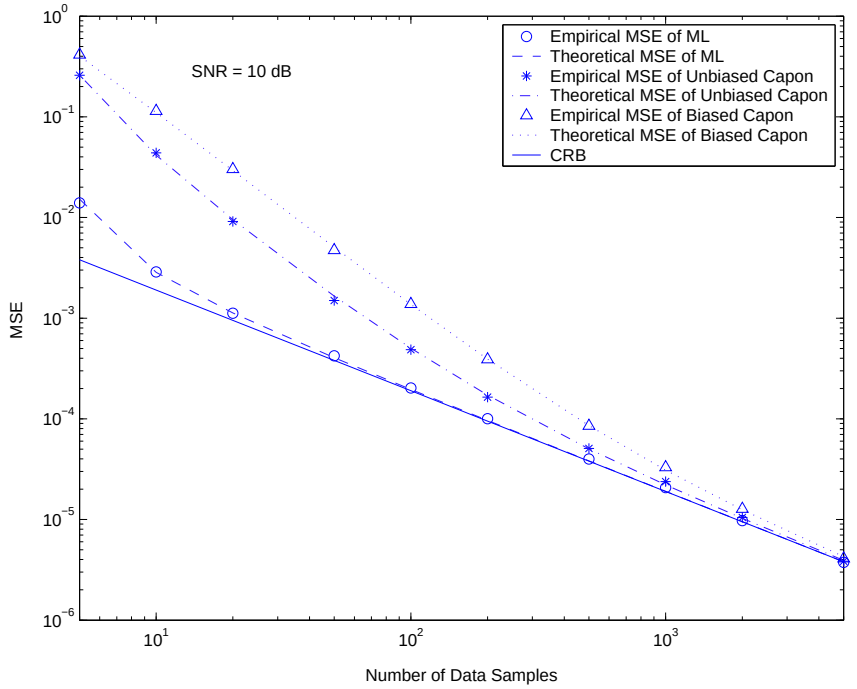

Figure 5–1: MSEs of $\hat{\beta}_{ML}$, $\hat{\beta}_{Capon}$ and $\hat{\beta}^{u}_{Capon}$ and the corresponding CRB vs. L when SNR= 10 dB

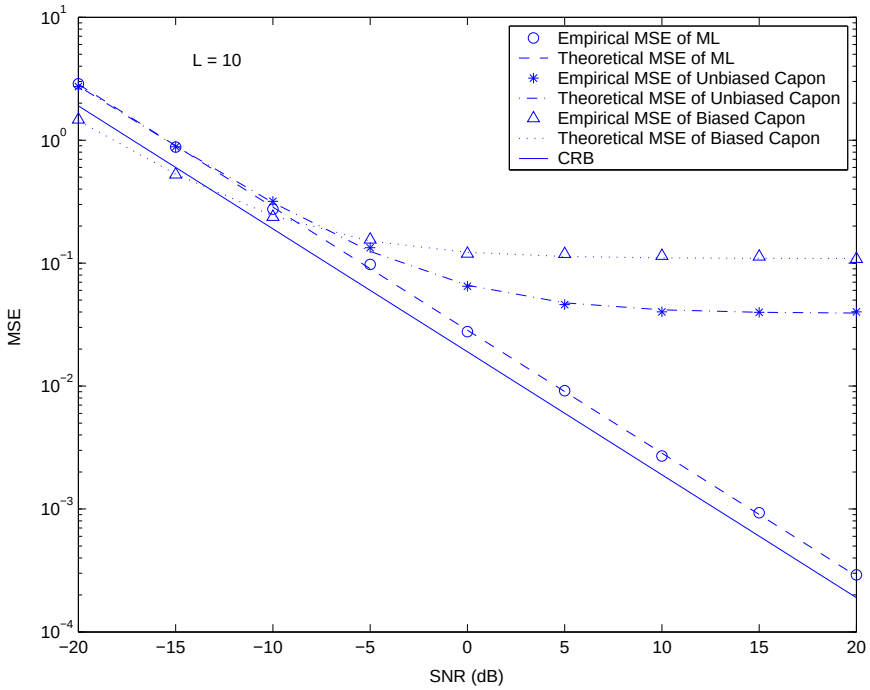

Figure 5–2: MSEs of $\hat{\beta}_{ML}$, $\hat{\beta}_{Capon}$ and $\hat{\beta}^{u}_{Capon}$ and the corresponding CRB vs. SNR when $L = 10$

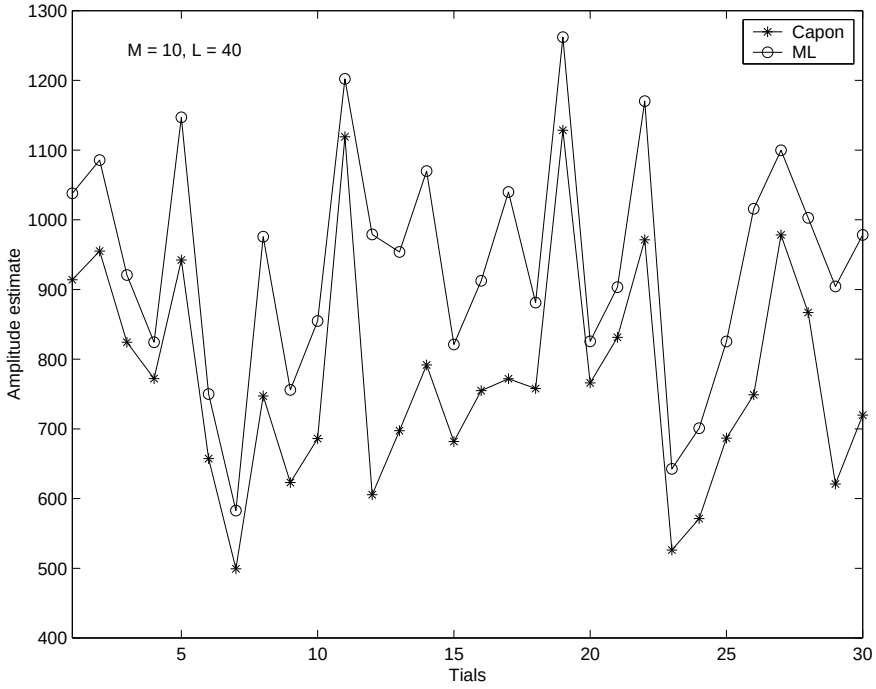

Figure 5–3: Amplitude estimates of a QR signal obtained via ML and Capon using experimentally measured data

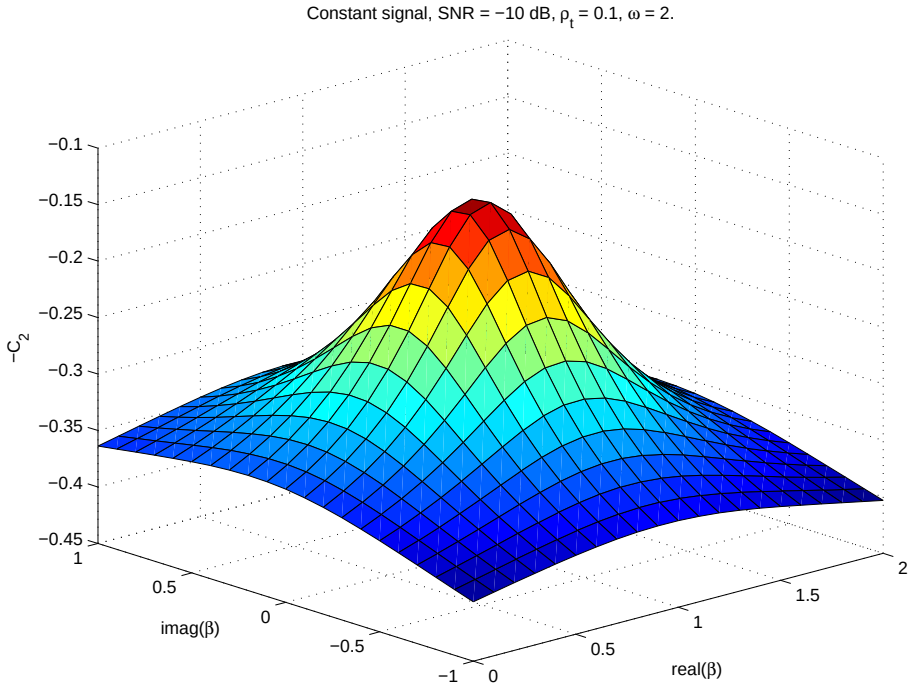

Figure 5–4: Mesh plot of $-C_2$ (with C_2 defined in (4.9)) for a constant signal vs. the real and imaginary part of β when SNR$= -10$ dB, $\rho_t = 0.1$, $\rho_s = 0.6$, $\omega = 2$, and the true value of β is 1

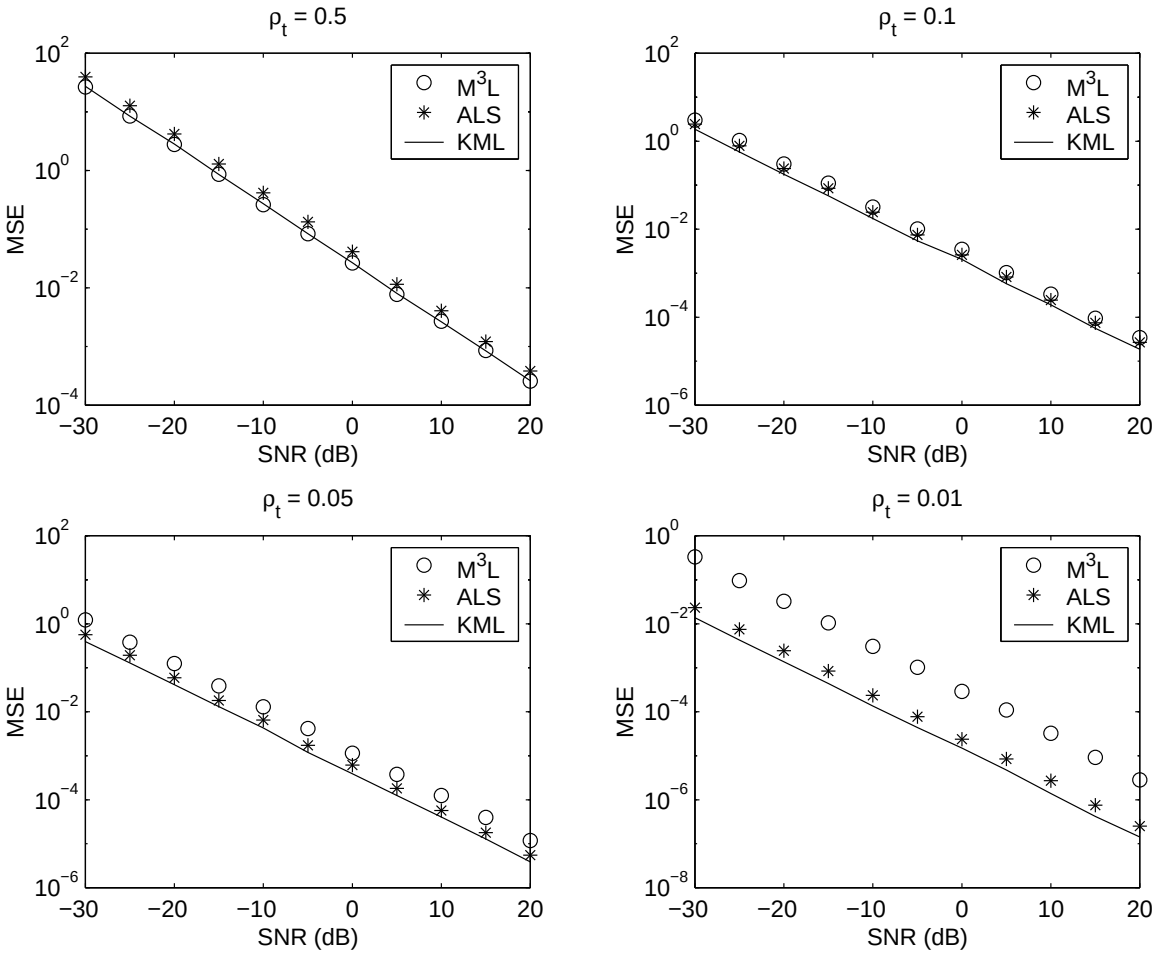

Figure 5–5: MSEs of ALS, M^{3}L and KML estimates for a constant signal vs. SNR when $\omega = 1$ and $\rho_s = 0.6$

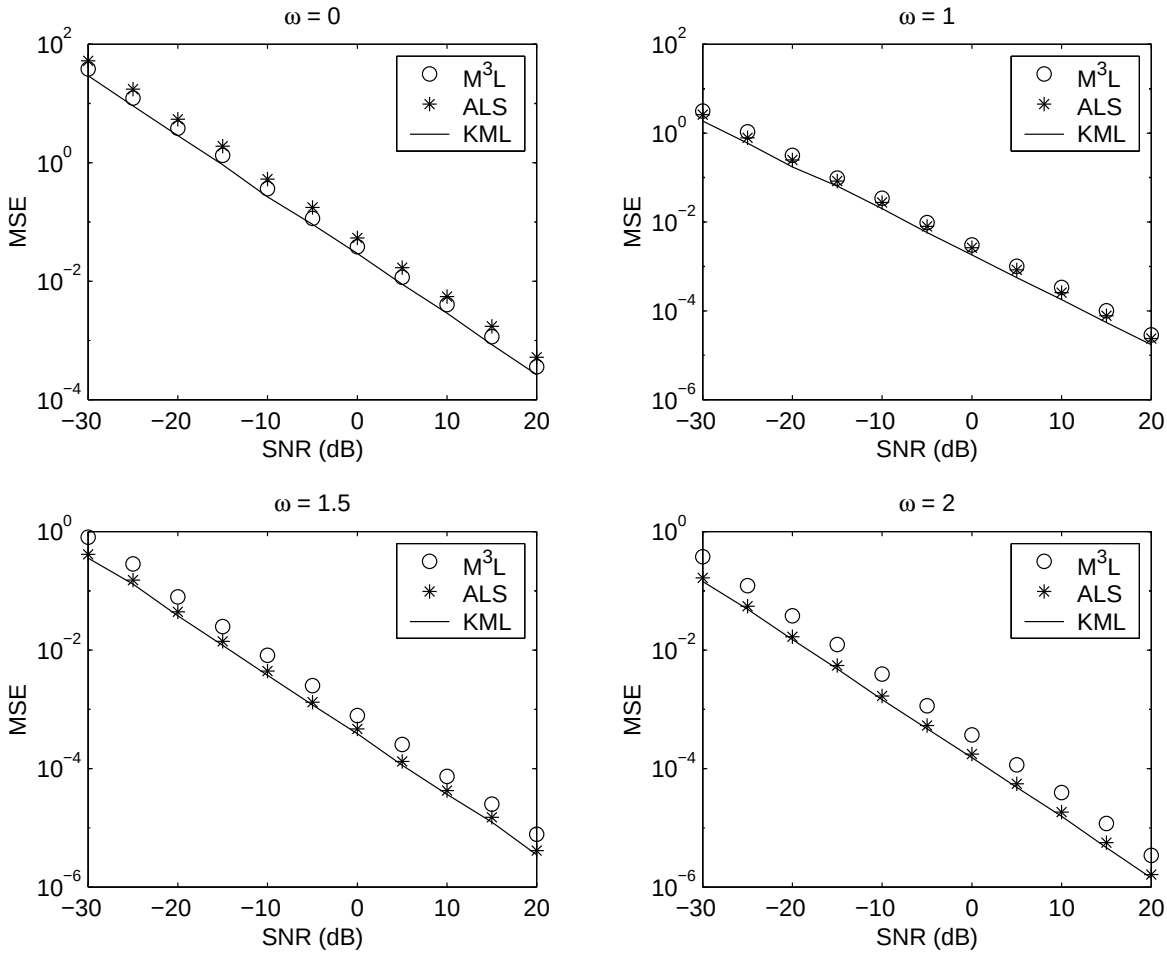

Figure 5–6: MSEs of ALS, M^{3}L and KML estimates for a constant signal vs. SNR when $\rho_t = 0.1$ and $\rho_s = 0.6$

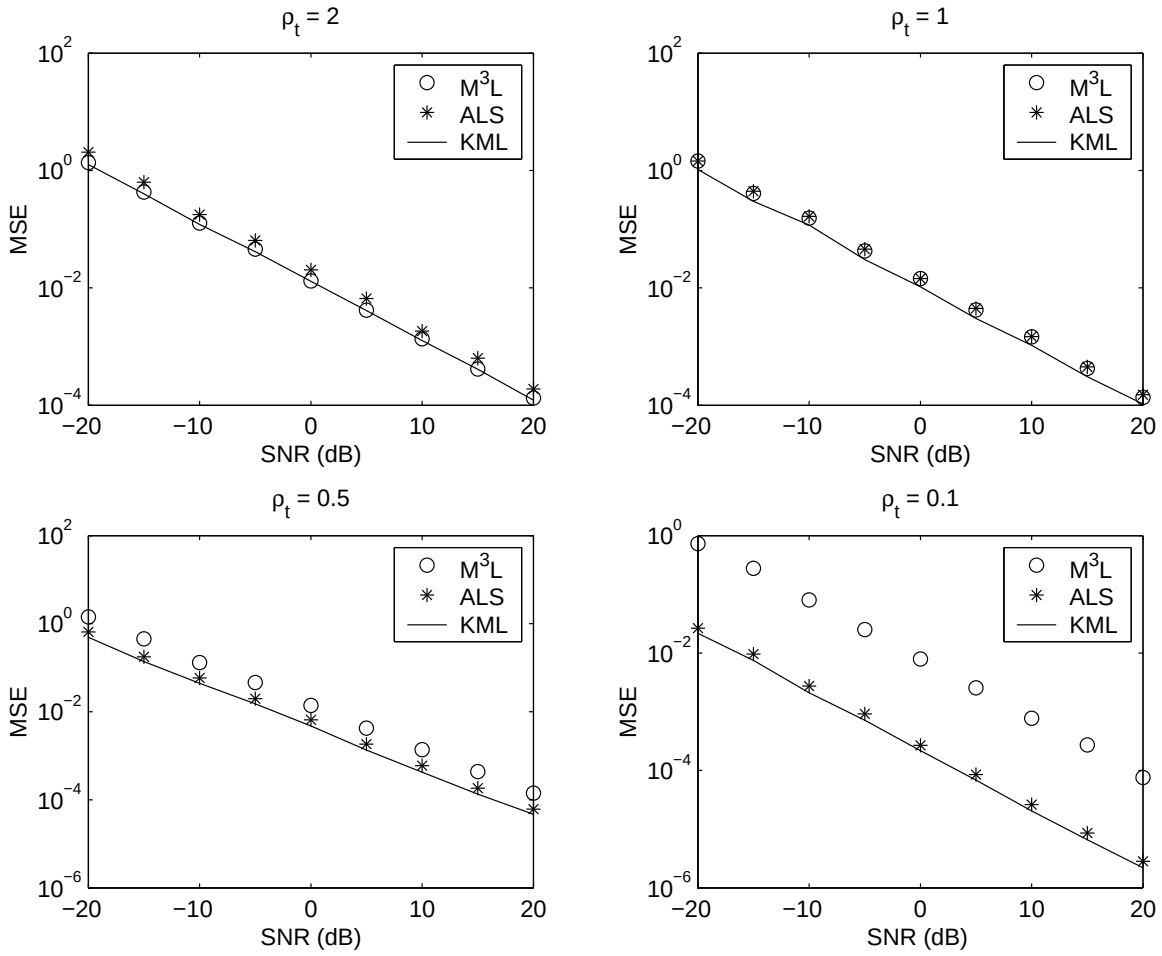

Figure 5–7: MSEs of ALS, M^{3}L and KML estimates for a known BPSK signal vs. SNR when $\omega = 0$ and $\rho_s = 0.6$

CHAPTER 6
CONCLUSIONS

We have investigated the problem of amplitude estimation for a signal with known waveform and steering vector in the presence of interference and noise. We first assumed that the interference and noise vector was spatially colored but temporally white. The ML and Capon methods and the closed-form expressions of the expected values and MSEs of the two estimators have been derived. We have shown that the ML estimate is unbiased and asymptotically statistically efficient for large data sample sets although it is not asymptotically statistically efficient for high SNR. We have also shown that the Capon method is biased downwards but it is asymptotically unbiased and efficient for large data sample lengths. The bias of the Capon estimate dominates its variance for high SNR, which results in an mean-squared error floor in high SNR area. At low SNR, however, Capon can even outperform ML and the CRB for any unbiased estimator. We then considered a more general scenario where the interference and noise vector was both spatially and temporally colored. We have proposed an ALS method based on the idea of cyclic optimization. We have shown that in most cases ALS outperforms the M^3L estimator which ignores the temporal correlation of the interference and noise.

APPENDIX A
DERIVATION OF THE ML ESTIMATOR

The normalized log-likelihood function of $\{\mathbf{x}_l\}_{l=1}^L$ is proportional to

$$C = -\ln|\mathbf{Q}| - \mathrm{tr}\left[\mathbf{Q}^{-1}\frac{1}{L}\sum_{l=1}^{L}(\mathbf{x}_l - \mathbf{a}\beta s_l)(\mathbf{x}_l - \mathbf{a}\beta s_l)^H\right] \tag{A.1}$$

where $|\cdot|$ denotes the determinant of a matrix. Maximizing the above cost function with respect to $\mathbf{Q}$ gives [3]

$$\hat{\mathbf{Q}}_{ML} = \frac{1}{L}\sum_{l=1}^{L}(\mathbf{x}_l - \mathbf{a}\beta s_l)(\mathbf{x}_l - \mathbf{a}\beta s_l)^H. \tag{A.2}$$

Hence the ML estimate of β is obtained as follows:

$$\begin{aligned}
\hat{\beta}_{ML} &= \arg\min_{\beta}\left|\frac{1}{L}\sum_{l=1}^{L}(\mathbf{x}_l - \mathbf{a}\beta s_l)(\mathbf{x}_l - \mathbf{a}\beta s_l)^H\right| && \text{(A.3)}\\
&= \arg\min_{\beta}\left|\frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l\mathbf{x}_l^H - \left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l s_l^*\right)(\mathbf{a}\beta)^H\right.\\
&\qquad\left. -(\mathbf{a}\beta)\left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l s_l^*\right)^H + \mathbf{a}\beta\beta^H\mathbf{a}^H\frac{1}{L}\sum_{l=1}^{L}|s_l|^2\right|\\
&\triangleq \arg\min_{\beta}\left|\hat{\mathbf{R}} - \bar{\mathbf{x}}(\mathbf{a}\beta)^H - (\mathbf{a}\beta)\bar{\mathbf{x}}^H + (\mathbf{a}\beta)P_s(\mathbf{a}\beta)^H\right|\\
&= \arg\min_{\beta}\left|\hat{\mathbf{R}} - \frac{\bar{\mathbf{x}}\bar{\mathbf{x}}^H}{P_s} + \left(\mathbf{a}\beta\sqrt{P_s} - \frac{\bar{\mathbf{x}}}{\sqrt{P_s}}\right)\left(\mathbf{a}\beta\sqrt{P_s} - \frac{\bar{\mathbf{x}}}{\sqrt{P_s}}\right)^H\right| && \text{(A.4)}\\
&\triangleq \arg\min_{\beta}|\mathbf{T}|\left|\mathbf{I} + \mathbf{T}^{-1}\left(\mathbf{a}\beta\sqrt{P_s} - \frac{\bar{\mathbf{x}}}{\sqrt{P_s}}\right)\left(\mathbf{a}\beta\sqrt{P_s} - \frac{\bar{\mathbf{x}}}{\sqrt{P_s}}\right)^H\right| && \text{(A.5)}\\
&= \arg\min_{\beta}\left(\mathbf{a}\beta\sqrt{P_s} - \frac{\bar{\mathbf{x}}}{\sqrt{P_s}}\right)^H\mathbf{T}^{-1}\left(\mathbf{a}\beta\sqrt{P_s} - \frac{\bar{\mathbf{x}}}{\sqrt{P_s}}\right) && \text{(A.6)}\\
&= \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}} && \text{(A.7)}
\end{aligned}$$

where from (A.5) to (A.6), we have used the fact that $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$ [4].

APPENDIX B
CRAMER-RAO BOUND

Let $\boldsymbol{\eta}$ be a vector containing all of the real-valued unknowns in the data model in (2.1). Let $\boldsymbol{\mu}_l = \mathbf{a}\beta s_l$. Then the Fisher's Information Matrix (FIM) for $\boldsymbol{\eta}$ is [4]

$$\mathrm{FIM}_{ij}(\boldsymbol{\eta}) = L \,\mathrm{tr}\left(\mathbf{Q}^{-1}\frac{\partial \mathbf{Q}}{\partial \eta_i}\mathbf{Q}^{-1}\frac{\partial \mathbf{Q}}{\partial \eta_j}\right) + 2\mathrm{Re}\sum_{l=1}^{L}\left[\left(\frac{\partial \boldsymbol{\mu}_l}{\partial \eta_i}\right)^H \mathbf{Q}^{-1}\left(\frac{\partial \boldsymbol{\mu}_l}{\partial \eta_j}\right)\right], \tag{B.1}$$

where η_i denotes the ith element of $\boldsymbol{\eta}$. Because $\boldsymbol{\mu}_l$ and $\mathbf{Q}$ depend on different elements of $\boldsymbol{\eta}$, $\mathrm{FIM}(\boldsymbol{\eta})$ will be block diagonal with respect to $[\ \mathrm{Re}\{\beta\}\ \ \mathrm{Im}\{\beta\}\ ]^T$ and the elements of $\mathbf{Q}$, where Re and Im denote the real and imaginary parts of a complex variable, respectively. Hence the first term of (B.1) does not affect the CRB of β and

$$\begin{aligned}
\mathrm{FIM}([\ \mathrm{Re}\{\beta\}\ \ \mathrm{Im}\{\beta\}\ ]^T) &= 2\mathrm{Re}\begin{bmatrix} \mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2 & j\mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2 \\ -j\mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2 & \mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2 \end{bmatrix} \\
&= 2\begin{bmatrix} \mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2 & 0 \\ 0 & \mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2 \end{bmatrix}.
\end{aligned} \tag{B.2}$$

Hence

$$\begin{aligned}
\mathrm{CRB}_\beta &= \mathrm{CRB}_{\mathrm{Re}\{\beta\}} + \mathrm{CRB}_{\mathrm{Im}\{\beta\}} && \text{(B.3)} \\
&= \frac{1}{\mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}\sum_{l=1}^{L}|s_l|^2} && \text{(B.4)} \\
&= \frac{1}{\mathbf{a}^H\mathbf{Q}^{-1}\mathbf{a}LP_s}. && \text{(B.5)}
\end{aligned}$$

APPENDIX C
PROOF OF LEMMA 1

In the following, a subscript is used to indicate the dimension of a matrix for the sake of clarity and is dropped whenever convenient. Consider an $M \times L$ matrix

$$\mathbf{X}_{M\times L} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_L] = \mathbf{a}\beta\mathbf{s}^T + \mathbf{E}_{M\times L}, \tag{C.1}$$

where $\mathbf{s} = [s_1, \ldots, s_L]^T$ with $(\cdot)^T$ denoting the transpose and

$$\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_L]\,. \tag{C.2}$$

We can construct a unitary matrix

$$\mathbf{U}_{L\times L} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_L]$$

whose first column $\mathbf{u}_1$ is chosen as $\mathbf{u}_1 = \frac{1}{\sqrt{LP_s}}\,[s_1, s_2, \ldots, s_L]^H$.

Let

$$\mathbf{Y}_{M\times L} = \mathbf{X}_{M\times L}\mathbf{U}_{L\times L} \tag{C.3}$$

and

$$\mathbf{Z}_{M\times L} = \mathbf{E}_{M\times L}\mathbf{U}_{L\times L}. \tag{C.4}$$

Inserting (C.1) into (C.3) gives

$$\mathbf{Y} = \left[\sqrt{LP_s}\beta\mathbf{a}, \mathbf{0}_{M\times(L-1)}\right] + \mathbf{Z}. \tag{C.5}$$

From (2.8), the first column of $\mathbf{Y}$

$$\mathbf{y}_1 = \sqrt{LP_s}\beta\mathbf{a} + \mathbf{z}_1 = \sqrt{\frac{L}{P_s}}\bar{\mathbf{x}}, \tag{C.6}$$

and

$$\mathbf{y}_l = \mathbf{z}_l,\ \ l = 2, \ldots, L, \tag{C.7}$$

where $\mathbf{y}_l$ (or $\mathbf{z}_l$) denotes the lth column of $\mathbf{Y}$ (or $\mathbf{Z}$). Since $\mathbf{U}\mathbf{U}^H = \mathbf{I}$,

$$\mathbf{Y}\mathbf{Y}^H = \mathbf{X}\mathbf{U}\mathbf{U}^H\mathbf{X}^H = \mathbf{X}\mathbf{X}^H. \tag{C.8}$$

It follows from (2.5)(2.11) and(C.8) that

$$\begin{aligned} L\mathbf{T} &= \mathbf{X}\mathbf{X}^H - \frac{L}{P_s}\bar{\mathbf{x}}\bar{\mathbf{x}}^H \\ &= \mathbf{Y}\mathbf{Y}^H - \mathbf{y}_1\mathbf{y}_1^H \\ &= \sum_{l=2}^{L} \mathbf{y}_l\mathbf{y}_l^H \\ &= \sum_{l=2}^{L} \mathbf{z}_l\mathbf{z}_l^H. \end{aligned} \tag{C.9}$$

Hence $\mathbf{T}$ is a function of $\mathbf{z}_l, l = 2, 3, \ldots, L$, while $\bar{\mathbf{x}} = P_s\beta\mathbf{a} + \sqrt{\frac{P_s}{L}}\mathbf{z}_1$.

We now show that $\{\mathbf{z}_l\}_{l=1}^{L}$ are statistically independent of each other. The cross correlation matrix of any two columns of $\mathbf{Z}$ has the form

$$\mathbf{B} = \mathcal{E}[\mathbf{z}_i\mathbf{z}_j^H] = \mathcal{E}[\mathbf{E}\mathbf{u}_i\mathbf{u}_j^H\mathbf{E}^H], \tag{C.10}$$

which is an $M \times M$ matrix with the $(k,l)^{\text{th}}$ element

$$\begin{aligned} [\mathbf{B}]_{kl} &= \mathcal{E}[\tilde{\mathbf{e}}_k^H\mathbf{u}_i\mathbf{u}_j^H\tilde{\mathbf{e}}_l] \\ &= \text{tr}\{[\mathbf{u}_i\mathbf{u}_j^H]\mathcal{E}[\tilde{\mathbf{e}}_l\tilde{\mathbf{e}}_k^H]\} \end{aligned} \tag{C.11}$$

where $\text{tr}(\cdot)$ denotes the trace and $\tilde{\mathbf{e}}_k^H$ denotes the kth row of $\mathbf{E}$. To obtain (C.11), we have used $\text{tr}(\mathbf{A}\mathbf{B}) = \text{tr}(\mathbf{B}\mathbf{A})$. Next, note that

$$\begin{aligned} \left[\mathcal{E}[\tilde{\mathbf{e}}_l\tilde{\mathbf{e}}_k^H]\right]_{ij} &= \mathcal{E}[e_{li}^* e_{kj}] \\ &= q_{kl}\delta_{ij}, \end{aligned} \tag{C.12}$$

where q_{kl} is the $(k,l)^{th}$ element of the covariance matrix $\mathbf{Q}$ and δ_{ij} is the Kronecker delta function defined in (4.5). Substituting (C.12) into (C.11) yields

$$[\mathbf{B}]_{kl} = \mathrm{tr}\{[\mathbf{u}_i \mathbf{u}_j^H] q_{kl} \mathbf{I}_L\} \tag{C.13}$$

$$= q_{kl}\delta_{ij}, \quad \text{for} \quad k,l = 1,2,\ldots,M. \tag{C.14}$$

Therefore,

$$\mathcal{E}[\mathbf{z}_i \mathbf{z}_j^H] = \mathbf{Q}\delta_{ij}. \tag{C.15}$$

Furthermore, based on the circularly symmetric property of $\{\mathbf{e}_l\}_{l=1}^{L}$, it is easy to show that $\mathcal{E}[\mathbf{z}_i \mathbf{z}_j^T] = \mathbf{0}$ for $1 \leq i,j \leq L$. Since $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L$ are Gaussian random vectors, they are statistically independent of each other. Since $\mathbf{T}$ is a function of $\mathbf{z}_l, l = 2,\ldots,L$, and $\bar{\mathbf{x}}$ is a function of $\mathbf{z}_1$, we conclude that $\mathbf{T}$ and $\bar{\mathbf{x}}$ are statistically independent of each other.

APPENDIX D
PROOF OF LEMMA 2

We know from (2.11) that

$$\hat{\mathbf{R}} = \mathbf{T} + \frac{\bar{\mathbf{x}}\bar{\mathbf{x}}^H}{P_s}. \tag{D.1}$$

Using the matrix inversion lemma on (D.1) gives

$$\hat{\mathbf{R}}^{-1} = \mathbf{T}^{-1} - \frac{\mathbf{T}^{-1}\bar{\mathbf{x}}\bar{\mathbf{x}}^H\mathbf{T}^{-1}}{P_s + \bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}. \tag{D.2}$$

Substituting (D.2) into (2.9), we have

$$\begin{aligned}
\hat{\beta}_{Capon} &= \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}} - \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}{P_s+\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}}{P_s(\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a} - \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}\bar{\mathbf{x}}^H\mathbf{T}^{-1}\mathbf{a}}{P_s+\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}}})} \\
&= \hat{\beta}_{ML}\frac{1 - \frac{\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}{P_s+\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}}{1 - \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}\bar{\mathbf{x}}^H\mathbf{T}^{-1}\mathbf{a}}{\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}(P_s+\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}})}} \\
&= \hat{\beta}_{ML}\frac{1}{1 + \left(\bar{\mathbf{x}}^H\mathbf{T}^{-1}\bar{\mathbf{x}} - \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}\bar{\mathbf{x}}^H\mathbf{T}^{-1}\mathbf{a}}{\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}}\right)/P_s} \qquad \text{(D.3)} \\
&= \hat{\beta}_{ML}\frac{1}{1+u}, \qquad \text{(D.4)}
\end{aligned}$$

where u is defined in (3.15) and by the Cauchy-Schwartz inequality, $u \geq 0$. Hence the lemma is proved.

APPENDIX E
PROOF OF LEMMA 3

We rewrite the $\bar{\mathbf{x}}$ in (2.8) as

$$\bar{\mathbf{x}} = P_s \mathbf{a}\beta + \sqrt{P_s}\bar{\mathbf{e}}, \tag{E.1}$$

where

$$\bar{\mathbf{e}} = \frac{1}{L\sqrt{P_s}} \sum_{l=1}^{L} \mathbf{e}_l s_l^*, \tag{E.2}$$

and $\bar{\mathbf{e}} \sim N(0, \mathbf{Q}/L)$. Then (3.15) can be reduced to

$$u = \bar{\mathbf{e}}^H \left(\mathbf{T}^{-1} - \frac{\mathbf{T}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{T}^{-1}}{\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}} \right) \bar{\mathbf{e}}. \tag{E.3}$$

From (3.4) and (E.2), we get

$$\hat{\beta}_{ML} - \beta = \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{e}}}{\sqrt{P_s}\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}}. \tag{E.4}$$

According to the conditional expectation rule,

$$\mathcal{E}\left[\lambda^n\left(\hat{\beta}_{ML} - \beta\right)\right] = \mathcal{E}_{\mathbf{T}}\left\{\mathcal{E}_{\bar{\mathbf{e}}|\mathbf{T}}\left[\lambda^n\left(\hat{\beta}_{ML} - \beta\right)\right]\right\}, \tag{E.5}$$

Since $\bar{\mathbf{e}}$ and $\mathbf{T}$ are statistically independent of each other, the conditional probability density function $f_{\bar{\mathbf{e}}|\mathbf{T}}(\bar{\mathbf{e}}|\mathbf{T}) = f_{\bar{\mathbf{e}}}(\bar{\mathbf{e}})$ is an even function of $\bar{\mathbf{e}}$. It follows from (E.3), (E.4) and (3.14) that

$$\mathcal{E}_{\bar{\mathbf{e}}|\mathbf{T}}\left[\lambda^n\left(\hat{\beta}_{ML} - \beta\right)\right] = 0, \tag{E.6}$$

since $\lambda^n\left(\hat{\beta}_{ML} - \beta\right)$ is an odd function of $\bar{\mathbf{e}}$. From (E.5), we obtain

$$\mathcal{E}\left[\lambda^n\left(\hat{\beta}_{ML} - \beta\right)\right] = 0, \quad \text{for} \quad n = 0, 1, 2, \cdots. \tag{E.7}$$

Hence the lemma is proved.

APPENDIX F
PROOF OF LEMMA 4

Let $\tilde{\mathbf{E}} = \mathbf{Y} - \beta\mathbf{G}$. Then we have $(\mathbf{YP} - \beta\mathbf{G})(\mathbf{YP} - \beta\mathbf{G})^H = \tilde{\mathbf{E}}\mathbf{P}\tilde{\mathbf{E}}^H$. Let $\mathbf{P} = \mathbf{V}\mathbf{\Gamma}\mathbf{V}^H$ denote the singular value decomposition of $\mathbf{P}$. Since $\mathbf{P}$ is an orthogonal projection matrix, the first r ($r = \text{rank}(\mathbf{P}) \leq M$) diagonal elements of $\mathbf{\Gamma}$ are ones and rest are zeros. Hence we get $\mathbf{P} = \bar{\mathbf{V}}_{L\times r}\bar{\mathbf{V}}^H$ where $\bar{\mathbf{V}} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r]$ consists of the first r columns of the unitary matrix $\mathbf{V}$. Let $\tilde{\mathbf{e}}_i^H$ denote the ith row of $\tilde{\mathbf{E}}$. Then

$$\left[\tilde{\mathbf{E}}\mathbf{P}\tilde{\mathbf{E}}^H\right]_{ii} = \sum_{j=1}^{r} |\tilde{\mathbf{e}}_i^H \mathbf{v}_j|^2. \tag{F.1}$$

Because $\mathcal{E}\{\tilde{\mathbf{e}}_i\tilde{\mathbf{e}}_i^H\} \to q_{ii}\mathbf{I}_L$ as $L \to \infty$, where q_{ii} denotes the ith diagonal element of the $\mathbf{Q}$ defined in (4.4), we have $\mathcal{E}\{|\tilde{\mathbf{e}}_i^H\mathbf{v}_j|^2\} \to q_{ii}||\mathbf{v}_j||^2 = q_{ii}$ as $L \to \infty$. It follows from (F.1) that the diagonal elements of $\tilde{\mathbf{E}}\mathbf{P}\tilde{\mathbf{E}}^H$ are bounded. Since $\tilde{\mathbf{E}}\mathbf{P}\tilde{\mathbf{E}}^H$ is a positive semi-definite matrix, we have $\tilde{\mathbf{E}}\mathbf{P}\tilde{\mathbf{E}}^H = O(1)$. However, $\mathbf{T}_y = \tilde{\mathbf{E}}\mathbf{P}^{\perp}\tilde{\mathbf{E}}^H = \tilde{\mathbf{E}}\tilde{\mathbf{E}}^H - \tilde{\mathbf{E}}\mathbf{P}\tilde{\mathbf{E}}^H = O(L)$. Hence $\mathbf{C} = (\mathbf{YP} - \beta\mathbf{G})(\mathbf{YP} - \beta\mathbf{G})^H\mathbf{T}_y^{-1} = O(\frac{1}{L})$. Let $\{\lambda_m\}_{m=1}^M$ denote the eigenvalues of $\mathbf{C}$. Then

$$F_1 = \left|(\mathbf{YP} - \beta\mathbf{G})(\mathbf{YP} - \beta\mathbf{G})^H\mathbf{T}_y^{-1} + \mathbf{I}\right| \tag{F.2}$$

$$= \Pi_{m=1}^{M}(1 + \lambda_m). \tag{F.3}$$

Since $\lambda_m \ll 1$ for large L, (F.3) can be approximated as

$$F_1 \approx 1 + \sum_{m=1}^{M} \lambda_m. \tag{F.4}$$

(Note that if $\lambda_m = 0$ for $m = 2, 3, \cdots, M$, which is true for the case of a known damped or undamped sinusoidal signal, (F.3) and (F.4) are exactly equal.) Thus minimizing F_1 is asymptotically (for large L) equivalent to minimizing

$$F_2 = \text{tr}\left[(\mathbf{YP} - \beta\mathbf{G})^H\mathbf{T}_y^{-1}(\mathbf{YP} - \beta\mathbf{G})\right]. \tag{F.5}$$

REFERENCES

[1] M. Rowe and J. Smith, "Mine detection by nuclear quadrupole resonance," *EUREL International Conference*, no. 431, pp. 62–66, 1996.

[2] S. Tantum, L. Collines, and L. Carin, "Signal processing for NQR discrimination of buried landmines," *Proceedings of SPIE*, vol. 3710, 1999.

[3] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proceedings of the IEEE*, vol. 86, pp. 421–441, February 1998.

[4] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[5] J. Li and R. T. Compton, Jr., "Maximum likelihood angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, vol. ASSP-41, pp. 2850–2862, September 1993.

[6] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2154–2163, September 1995.

[7] A. Zeira and B. Friedlander, "Direction of arrival estimation using parametric signal models," *IEEE Transactions on Signal Processing*, vol. 44, pp. 339–350, February 1996.

[8] M. Cedervall and R. L. Moses, "Efficient maximum likelihood DOA estimation for signals with known waveforms in the presence of multipath," *IEEE Transactions on Signal Processing*, vol. 45, pp. 808–811, March 1997.

[9] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal," *IEEE Transactions on Signal Processing*, pp. 2477–2484, October 1997.

[10] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Transactions on Signal Processing*, vol. 46, pp. 449–462, February 1998.

[11] A. Jakobsson, A. Swindlehurst, and P. Stoica, "Subspace-based estimation of time delays and doppler shifts," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2472–2483, September 1998.

[12] U. Spagnolini, "An iterative quadratic method for high resolution delay estimation with known waveform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 1134–1137, March 2000.

[13] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 10, pp. 853–863, November 1974.

[14] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 22, pp. 115–127, March 1986.

[15] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408–1418, August 1969.

[16] B. D. V. Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, pp. 4–24, April 1988.

[17] H. L. Van Trees, *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory.* New York, NY: John Wiley and Sons, Inc., 2002.

[18] M. Murthi and B. Rao, "All-pole modeling of speech based on the minimum variance distortionless response spectrum," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 221 –239, May 2000.

[19] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Transactions on Information Theory*, vol. 41, pp. 944–960, July 1995.

[20] J. Schodorf and D. Williams, "A constrained optimization approach to multiuser detection," *IEEE Transactions on Signal Processing*, pp. 258–262, January 1997.

[21] M. Tsatsanis and Z. Xu, "Performance analysis of minimum variance cdma receiver," *IEEE Transactions on Signal Processing*, vol. 46, pp. 3014–3022, November 1998.

[22] H. Li, X. Lu, and G. Giannakis, "Capon multiuser receiver for CDMA systems with spatial-time coding," *IEEE Transactions on Signal Processing*, vol. 50, May 2002.

[23] N. R. Goodman, "Statistical analysis based on a certain multi-variate complex Gaussian distribution," *Ann. Math. Stat.*, vol. 34, pp. 152–177, March, 1963.

[24] J. Capon and N. R. Goodman, "Probability distribution for estimators of the frequency-wavenumber spectrum,," *Proceedings of the IEEE*, vol. 58, pp. 1785–1786, October 1970.

[25] B. Van Veen, "Adaptive convergence of linearly constrained beamformers based on the sample covariance matrix," *IEEE Transactions on Signal Processing*, vol. 39, pp. 1470–1473, June 1991.

[26] J. L. Krolik and D. N. Swingler, "On the mean-square error performance of adaptive minimum variance beamformers based on the sample covariance matrix," *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 445–448, February 1994.

[27] J. H. Michels, P. Varshney, and D. D. Weiner, "Synthesis of correlated multichannel random processes," *IEEE Transactions on Signal Processing*, vol. 42, pp. 367–375, February 1994.

[28] W. I. Zangwill, *Nonlinear Programming: A Unified Approach.* Prentice-Hall, Inc., Englewood Cliffs, NJ 07632, 1969.

[29] S. L. Marple, Jr., *Digital Spectral Analysis with Applications.* Prentice Hall NJ: Upper Saddle River, 1987.

[30] T. Söderström and P. Stoica, *System Identification.* London, U.K.: Prentice-Hall International, 1989.

[31] S. R. Searle, *Matrix Algebra Useful for Statistics.* New York, NY: John Wiley and Sons, Inc., 1982.

BIOGRAPHICAL SKETCH

Yi Jiang was born in Yixing, Jiangsu Province, People's Republic of China, on November 8, 1978. He received his Bachelor of Science degree from the University of Science and Technology of China (USTC) in 2001.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Jian Li, Chair
Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Michael C. Nechyba
Assistant Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

John M. Shea
Assistant Professor of Electrical and
Computer Engineering

This thesis was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Master of Science.

August 2003

Pramod P. Khargonekar
Dean, College of Engineering

Winfred M. Phillips
Dean, Graduate School

TITLE OF THE DISSERTATION

Yi Jiang
(352) 392-5241
Department of Electrical and Computer Engineering
Chair: Jian Li
Degree: Master of Science
Graduation Date: August 2003

Signal detection and estimation is a mandatory task in many military and civil applications, including using the emerging Quadrupole Resonance (QR) technology for explosive detection. It is found that the TNT molecules in any explosive can emit a characteristic response (we refer to it as QR signal) if stimulated by a specific stimulus. The main challege is that the QR signal falls in the AM/FM frequency bands. Hence we need to suppress the strong interferences from the nearby AM/FM radio stations before implementing signal estimation and detection. In this paper, we use array processing technique to suppress the interference. We come up two methods, ML and Capon, and study thoroughly the performance of the two methods. It has been shown that the new method based on the results of this thesis can signficantly improve the explosive detection performance over the method used in the current QR system.