NASA-CR-197485

NCC3-319

FINAL
IN-32-CR
OCIT
33901
P-119

SYNTHETIC APERTURE RADAR TARGET DETECTION, FEATURE EXTRACTION, AND IMAGE FORMATION TECHNIQUES

Jian Li
Department of Electrical Engineering
405 CSE, Bldg. 42
University of Florida, Gainesville, FL 32611.
Phone: (904) 392-2642; Fax: (904) 392-0044; E-mail: li@saturn.ee.ufl.edu

Submitted to:
Ohio Aerospace Institute
As the Final Report for the Contract from
WL/AARA, Wright Laboratory, Wright Patterson Air Force Base
September 1993 - September 1994

(NASA-CR-197485) SYNTHETIC APERTURE RADAR TARGET DETECTION, FEATURE EXTRACTION, AND IMAGE FORMATION TECHNIQUES Final Report, Sep. 1993 - Sep. 1994 (Florida Univ.) 119 p

N95-17651

Unclas

G3/32 0033901

TABLE OF CONTENTS

Abstract

This final report presents new algorithms for target detection, feature extraction, and image formation with the synthetic aperture radar (SAR) technology. For target detection, we consider target detection with SAR and coherent subtraction. We also study how the image false alarm rates are related to the target template false alarm rates when target templates are used for target detection. For feature extraction from SAR images, we present a computationally efficient eigenstructure-based 2D-MODE algorithm for two-dimensional frequency estimation. For SAR image formation, we present a robust parametric data model for estimating high resolution range signatures of radar targets and for forming high resolution SAR images.

1. Introduction

This final report presents new algorithms for target detection, feature extraction, and image formation with the synthetic aperture radar (SAR) technology.

In Chapter 2, we consider target detection with SAR and coherent subtraction. We shall show with some limited experimental data that the coherent subtraction technique may be used to suppress outliers and obtain approximate Gaussian distributions for clutter and noise. We shall also derive generalized likelihood ratio (GLR) detection algorithms that may be used with SAR images that have Gaussian distributions. We shall analytically compare the performance of a) a single pixel detector, b) a detector using complete knowledge of the target signature information and known orientation information, c) a detector using incomplete knowledge of the target signature information and known orientation information, d) a detector using unknown target signature information and known orientation information, and e) a detector using unknown target signature information and unknown orientation information.

In Chapter 3, we study how the image false alarm rates are related to the target template false alarm rates when target templates are used for target detection. In particular, we shall show a simple way of determining the probability of false alarm of a target template when a low constant false alarm rate is desired for an image and when the image size is much larger than the size of the target template.

In Chapter 4, we present a computationally efficient eigenstructure-based 2D-MODE algorithm for two-dimensional frequency estimation or feature extraction from SAR images. We derive the theoretical performance of the 2D-MODE estimator and show that it is asymptotically statistically efficient under either the assumption that the number of temporal snapshots is large or the signal-to-noise ratio is high. Numerical examples showing the performance of this algorithm and comparing it with the computationally efficient subspace rotation algorithms are also given. We show that the statistical performance of the 2D-MODE algorithm is better than that of the subspace rotation methods, whereas the amount of computations required by the former is usually no more than a few times of that needed by the latter.

In Chapter 5, we present a robust parametric data model for estimating high resolution range signatures of radar targets and for forming high resolution SAR images. This paper also presents an estimation algorithm for the data model. The algorithm is referred to as the APES (Amplitude and Phase Estimation of a Sinusoid in unknown colored noise) algorithm. We shall describe how the APES algorithm can be used to estimate range signatures and to form SAR images. We shall show, with both numerical and experimental examples, that our modeling and estimation approach yields better resolution and lower sidelobes than the conventional nonparametric FFT (fast Fourier transform) method. We shall also show that our approach is more robust than modeling the radar data as a certain number of complex sinusoids in noise and estimating the frequencies, amplitudes, and phases of the sinusoids with one of the best sinusoidal parameter estimation methods.

Each of the afore-mentioned chapters is self-contained with its own introductions, formulations of the problems of interests, approaches, conclusions, and references.

The results we present in this report are obtained with the partial support from WL/AARA, Wright Laboratory, Wright Patterson Air Force Base as a subcontract from the Ohio Aerospace Institute. Our work is also supported in part by the 1993 and 1994 AFOSR Summer Faculty Research Programs, by the National Science Foundation, and by the Göran Gustafsson Foundation.

Those who contributed to this report include Dr. Jian Li, Mr. Syed M. Rahman, Dr. Petre Stoica, Mr. Edmund G. Zelnio, and Mr. Dunmin Zheng.

2. Target Detection with Synthetic Aperture Radar and Coherent Subtraction

2.1 Introduction

Synthetic aperture radar (SAR) technology may be used to detect radar targets of interest. High resolution SAR technology is especially useful for detecting small radar targets embedded in strong ground clutter such as in foliage. In this paper, we shall consider target detection algorithms that may be used with high resolution SAR.

Target detection from SAR or optical images has been considered by many authors. For example, Novak, Burl, and Irving [1] considered target detection with a polarimetric SAR. The three output images of the polarimetric SAR are first processed by a polarimetric whitening filter, which is derived by assuming a K-distribution for clutter and noise. The output image of the filter is next used with a two-parameter detector for target detection. The target detection in [1] is performed one pixel at a time even though the target may occupy more than one pixel; i.e., even though the target size may be larger than the resolution of the SAR image.

Reed and Yu [2] considered generalized likelihood ratio target detection from a sequence of optical images, which are first preprocessed by removing local means so that the clutter and noise will approximately have the Gaussian distribution. In [2], each target in an image is described by a completely known template or signature with an unknown gain, which is a scalar. The algorithm, however, may not work well with SAR images. For example, for a target in foliage, the SAR target signature may change due to the interaction between target and surrounding clutter.

Stotts [3] considered detecting several dim targets in an image simultaneously. The image is also first preprocessed by removing local means so that the clutter and noise will approximately have the Gaussian distribution. Each dim target is described by a known template or signature with an unknown gain, which is a scalar. Stotts has shown that simultaneous detection of multiple targets may perform better than separate detection of individual target. In this paper, we shall extend this idea of the simultaneous detection of multiple targets to the detection of a target with multiple pixels in a SAR image.

We shall derive generalized likelihood ratio (GLR) detection algorithms that may be used with multiple SAR images that are obtained with coherent subtraction or have Gaussian distributions. We shall show with some limited and well calibrated experimental data that we may eliminate Gaussian outliers of the clutter and noise through coherent subtraction between two complex SAR images of the same area of interest. One of the two SAR images is assumed to be target free and the other is to be sought for the presence of target. In each image obtained with coherent subtraction, the target of interest is modeled with a target template, which is large enough to cover the entire target. The size of the target template and the number of complex unknowns in it are determined by the knowledge of the target orientation information and the amount of target signature information known to the detector. Using this unifying framework, we shall analytically compare the performance of a) a single pixel detector, b) a detector using complete knowledge of the target signature information and known orientation information, c) a detector using incomplete knowledge of the target signature information and known orientation information, d) a detector using unknown target signature information and known orientation information, and e) a detector using unknown target signature information and unknown orientation information. We shall derive the probability of detection and the probability of false alarm of each detector. To achieve a constant false alarm rate (CFAR), each detector threshold is simply a function of the dimensional parameters of the detection problem and the desired probability of false alarm.

In Section II, we discuss the effects of Gaussian assumption on target detection. In

Section III, we formulate the target detection problem. In Section IV and Appendices A, B, and C, we assume that the statistics of the clutter and noise are known and derive a GLR detector under the assumption and discuss its performance. In Section V and Appendices D and E, we assume that the statistics of the clutter and noise are unknown and present a more practical GLR detector. In Section VI, we apply the more practical detector to the experimental data. Finally, Section VII contains our conclusions.

2.2 Gaussian Assumption

A key problem in radar target detection is the description of the statistical properties of radar clutter and noise. In general, radar clutter and noise do not satisfy the conditions of a Gaussian distribution. Many statistical models, such as the well-known log-normal, Weibull, and K-distributions, have been proposed to describe the clutter and noise statistics. Although these distributions may provide better statistical models for the clutter and noise than the Gaussian distribution, the detectors that are derived based on these models may be very complicated and may involve an expensive multidimensional search over the parameter space. It is also difficult to analyze the performance of these detectors. It is difficult, if not impossible, to derive the probability of detection and probability of false alarm for such a detector. It is thus difficult to analytically determine the correct threshold for a prescribed probability of false alarm.

Alternatively, the SAR images may be preprocessed so that the Gaussian assumption for clutter and noise is approximately valid. For example, Hunt and Cannon [4] and Reed and Yu [2] considered removing local means as such a preprocessing technique. The detectors obtained from Gaussian assumptions may avoid the multidimensional search over the parameter space. It is also quite tractable to analyze the performance of such detectors and many Gaussian assumption based results exist in the literature. Yet if the distribution of the preprocessed images is still non-Gaussian, then the detectors obtained under the Gaussian assumption will suffer performance degradation. The selection of a detector, therefore, must balance these tradeoffs.

In this paper, we consider target detection by assuming the radar clutter and noise in a SAR image have a Gaussian distribution. Such a SAR image may be obtained by taking the difference between two SAR images of the same area, with one assumed to be target free and the other is to be sought for the presence of target. We show below that the coherent subtraction method can suppress Gaussian outliers when the SAR images are obtained by moving the radar antenna along a fixed rail. These rail SAR images can be used in many applications including monitoring the environmental changes of a cetain area of interest. For an airborne SAR, however, we do not know how the coherent subtraction method may perform. The presence of errors such as the flight path errors and the existence of atmospheric turbulence can degrade the performance of the coherent subtraction.

The limited and well calibrated experimental data we have are obtained by ERIM (Environmental Research Institute of Michigan) with a portable rail SAR that has foliage penetration capabilities [5]. More specifically, a horizontal 36 feet long aluminum truss was used to support a rail and an antenna carriage. The data was collected by moving an antenna along a fixed rail to obtain the synthetic aperture. The rail SAR is an FM-CW radar system based on an HP-8501 network analyzer [5]. Our results are, therefore, tentative and the effectiveness of coherent subtraction needs to be further studied with more extensive experimental data obtained under more operational conditions.

The data we shall use were obtained when both the transmitter and receiver of the SAR are horizontally linearly polarized. Figures 2.1(a) and (b) show the 3-dimensional (3-D) plots of the magnitudes of two complex SAR images obtained with two identical synthetic apertures. The frequency band used for the image is between

400 and 1300 MHz and the depression angle is 30°. We have downsampled the original images presented in [5] by a factor of two in range and by a factor of six in cross range since the original images are oversampled. The range and cross range resolutions in the SAR images are 0.34 and 3 meters, respectively. Figure 2.1(a) shows the 3-D plot with foliage only. The peaks in the figure correspond to the radar returns from tree trunks. Figure 2.1(b) shows the 3-D plot of a target in foliage. The target is a pickup truck rotated 24° counterclockwise from end-on. Figure 2.1(c) shows the 3-D plot of the magnitude of the coherent subtraction between the two complex SAR images. Figure 2.2 is similar to Figure 2.1 except that the truck is broadside. Thus the target return in Figure 2.2 is much stronger than in Figure 2.1.

We note that coherent subtraction can effectively suppress the large clutter returns due to tree trunks, which result in false alarms and causes CFAR detectors to fail. The large returns left in Figures 2.1(c) and 2.2(c) are due to the target and its surroundings. For example, the darkened peak to the left of the target in Figure 2.2(c) occurs in Figure 2.2(a) but not in Figure 2.2(b). This result occurs because the main response of a tree comes when the radar energy bounces off the tree onto the ground and returns to the radar or vice versa. The presence of the target interrupts this path and results in no tree return in the image in Figure 2.2(b). Subtracting the image in Figure 2.2(a), which contains the tree return, from the image in Figure 2.2(b), which does not contain the tree return, results in a tree return not canceled by the subtraction process thus yielding the darkened peak. The presence of the darkened peak in Figure 2.2(c) is additional information showing the presence of a target because its presence is due to the interaction between the target and clutter. This information may be especially useful when the target return is weak.

Figure 2.3 shows the magnitude of the correlation coefficient of the clutter and noise (i.e., the target-free) pixels in Figure 2.1(c) as a function of the spatial distance (as measured by the number of pixels) between two pixels. An unbiased autocorrelation estimator [6] was used to estimate the correlation coefficient. We note that the clutter and noise pixels are approximately independent of each other.

Consider next the test of normality of the clutter and noise pixels before and after coherent subtraction. Let $\mathbf{x}_i$, $i = 1, 2, \cdots, Q$, denote a 2×1 vector containing the real and imaginary parts of the ith pixel used for the test. Let

$$\beta_i = \frac{1}{Q-1}\left(\mathbf{x}_i - \hat{\hat{\mathbf{x}}}\right)^T \hat{\Sigma}^{-1}\left(\mathbf{x}_i - \hat{\hat{\mathbf{x}}}\right), \tag{2.1}$$

where $(\cdot)^T$ denotes the transpose,

$$\hat{\hat{\mathbf{x}}} = \frac{1}{Q}\sum_{i=1}^{Q} \mathbf{x}_i, \tag{2.2}$$

and

$$\hat{\Sigma} = \frac{1}{Q}\sum_{i=1}^{Q}\left(\mathbf{x}_i - \hat{\hat{\mathbf{x}}}\right)\left(\mathbf{x}_i - \hat{\hat{\mathbf{x}}}\right)^T. \tag{2.3}$$

For true Gaussian random vectors, the variance of β_i is approximately equal to $4/(Q-1)^2$. Thus the test of normality of $\mathbf{x}_i$ may be performed by comparing the sample variance of β_i with $4/(Q-1)^2$ [7]; i.e., by comparing how close ϵ is to 1, where ϵ denotes the ratio between the sample variance of β_i and $4/(Q-1)^2$. For the clutter and noise pixels in Figure 2.1, the ϵ is equal to 7.97 before coherent subtraction. This large ϵ is caused by the large tree trunk returns. After coherent subtraction, however, the ϵ is equal 1.10 and is much closer to 1. The histogram of the clutter and noise pixels after coherent subtraction can also be shown to match a zero-mean complex Gaussian probability density function well.

Finally, since broadcast stations often occupy low frequency bands, SAR images of the same area may be formed with two or more separate frequency bands instead of one wide frequency band. The separate bands may be chosen to avoid the jamming from the radio stations. The clutter and noise pixels in different images obtained with different frequency bands can also be shown to be approximately independent of each other.

The problem formulation below will take these results into account.

2.3 Formulation of the Target Detection Problem

Consider J high resolution SAR images obtained with coherent subtraction in which a target may be present. The J images may be obtained from polarimetric SAR, different frequency bands, and/or different aspect angles. For each image, the target may be modeled with a template consisting of N_j pixels, $j = 1, 2, \cdots, J$. The shape of the templates may be *arbitrary* and the templates may consist of areas that are not connected. Among the N_j pixels, K_j ($K_j \leq N_j$) pixels are assumed to be deterministic and *arbitrary* unknown complex scalars that correspond to the bright returns of target scatterers. The remaining $N_j - K_j$ pixels are assumed to contain clutter and noise only; i.e., they correspond to the areas of the target that does not generate radar returns.

The locations of the K_j pixels may be assumed known, not completely known, or unknown. If the locations of the unknown scalars are not known exactly, the dimension K_j (and N_j correspondingly since $N_j \geq K_j$) may be increased to include all possible locations of the scalars. If the locations are unknown and we only know the approximate size of a target, we may choose a template large enough to cover the target and assume that all pixels in the template are unknown, i.e, $K_j = N_j$. Yet increasing K_j decreases the detection performance due to the increased number of unknowns in the target template. This result will be shown in the following sections.

Let $\mathbf{z}$ denote an $N \times 1$ vector consisting of the pixels of such templates in the presence of clutter and noise, where $N = N_1 + N_2 + \cdots + N_J$. Under hypothesis H_1, the target presence hypothesis, the $\mathbf{z}$ may be written

$$\mathbf{z} = \mathbf{S}\mathbf{b} + \mathbf{n}. \tag{2.4}$$

The $\mathbf{b}$ is the $K \times 1$ vector consisting of the K deterministic unknown complex scalars in the target templates, where $K = K_1 + K_2 + \cdots + K_J$. The $\mathbf{b}$ may be written

$$\mathbf{b} = \left[\begin{array}{cccc} \mathbf{b}_1^T & \mathbf{b}_2^T & \cdots & \mathbf{b}_J^T \end{array} \right]^T, \tag{2.5}$$

where $\mathbf{b}_j$, $j = 1, 2, \cdots, J$, is the $K_j \times 1$ vector consisting of the K_j deterministic unknown complex scalars in the target template of the jth image, or the jth template.

The $\mathbf{S}$ is a full-rank $N \times K$ matrix describing the locations of the unknown scalars. Only one element in each row and each column of $\mathbf{S}$ is one and the remaining of the elements are zero. Thus we have

$$\mathbf{S}^H\mathbf{S} = \mathbf{I}_K, \tag{2.6}$$

where $(\cdot)^H$ denotes the complex conjugate transpose and $\mathbf{I}_K$ denotes the identity matrix of dimension K. The $\mathbf{S}$ may be written

$$\mathbf{S} = \left[\begin{array}{cccc} \mathbf{S}_1 & & & \mathbf{0} \\ & \mathbf{S}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{S}_J \end{array} \right], \tag{2.7}$$

where $\mathbf{S}_j$, $j = 1, 2, \cdots, J$, is a full-rank $N_j \times K_j$ matrix describing the locations of the unknown scalars of the jth target template and

$$\mathbf{S}_j^H\mathbf{S}_j = \mathbf{I}_{K_j}, \tag{2.8}$$

The $\mathbf{n}$ denotes the $N \times 1$ clutter and noise random vector and may be written

$$\mathbf{n} = \left[\begin{array}{cccc} \mathbf{n}_1^T & \mathbf{n}_2^T & \cdots & \mathbf{n}_J^T \end{array} \right]^T, \tag{2.9}$$

where $\mathbf{n}_j$ are $N_j \times 1$ clutter and noise vectors of the jth image and are assumed zero-mean complex Gaussian with covariance matrix $\sigma_j^2\mathbf{I}_{N_j}$ and are statistically independent of each other. Under hypothesis H_0, the target absence hypothesis, the $\mathbf{z}$

may be written

$$\mathbf{z} = \mathbf{n}. \tag{2.10}$$

The problem of interest is to develop a CFAR detector for the data model (2.4). We shall consider the generalized likelihood ratio (GLR) target detection algorithms for the purpose. We shall consider both a detector, where the clutter and noise variance is assumed known, and a more practical detector, where the clutter and noise variance is unknown. The effects of the dimensional parameters such as N, K, and J on the performance of the detectors will also be considered through performance analysis of the detectors.

We remark that the above detection problem is a generalized version of the approach of considering one pixel at a time and the approach of using the complete knowledge of the target signature. When we consider detection by using one pixel at a time, we have $N_j = K_j = 1$, $j = 1, 2, \cdots, J$. When the complete knowledge of the target signature is known except for a complex gain, the $\mathbf{S}_j$ becomes the signature vector $\tilde{\mathbf{s}}_j$ whose elements are arbitrary except that its Euclidean norm is constrained to be 1. Also, the $\mathbf{b}_j$ becomes the unknown complex scalar gain $\tilde{b}_j$ and $K_j = 1$, $j = 1, 2, \cdots, J$. The elements of $\tilde{\mathbf{s}}_j$ describe the locations and relative return strengths and phases of the scatterers of a target.

We next remark on the knowledge of $\mathbf{S}_j$ and $\tilde{\mathbf{s}}_j$, which may be obtained through experiments or simulations with software packages similar to XPATCH [8]. The target signature vector $\tilde{\mathbf{s}}_j$ may be easily altered by the surrounding environment of a target, but the location matrix $\mathbf{S}_j$ is more robust to the target surroundings. For a given target, a set of signature vectors or location matrices may be obtained for different orientation angles of the target relative to the radar. These vectors or matrices may be used with GLR detectors to form a filter bank.

For example, consider detecting the target shown in Figures 2.1 and 2.2. Since $J = 1$, for example, we drop the subscript j for simplicity. For this example, we may use four different $\mathbf{S}$'s to form a filter bank: (i) when the target is oriented broadside (or +180° due to symmetry), (ii) when the target is oriented end-on, (iii) when the target is oriented head-on, and (iv) for the rest of the target orientations. We note from Figure 2.1 that for Case (iv), the target return consists of only one bright pixel and hence we should let $\mathbf{S} = 1$, which results in a single pixel detector.

When a filter bank is used for target detection, the probability of false alarm may be increased as compared with using a single filter; i.e., a single $\bar{\mathbf{s}}_j$ or $\mathbf{S}_j$, for target detection. When a set of $\mathbf{S}_j$'s is used, for example, the increase will depend on how similar the $\mathbf{S}_j$'s are. The more non-overlapping elements in $\mathbf{S}_j$'s, the larger the increase of the probability of false alarm. Yet the increase is no more than the number of filters times the probability of false alarm of using a single filter. It appears that the exact probability of false alarm of using a filter bank for target detection cannot be determined analytically in general and must be obtained with Monte-Carlo simulations. In other words, although a constant false alarm rate can be achieved by using a filter bank, we may not know exactly what the achieved false alarm rate is.

The design of a filter bank is more flexible using $\mathbf{S}_j$ than using $\bar{\mathbf{s}}_j$. We may increase the template size N_j and the number of unknown template pixels K_j to reduce the number of filters in the filter bank. Increasing the number unknown template pixels K_j has the effect of decreasing the sensitivity of the target detector to target signature information. Increasing the template size N_j has the effect of decreasing the sensitivity of the target detector to target aspect. By selecting appropriate K_j and N_j for each image based on the amount of the *a priori* knowledge about the target, we may use a single filter instead of a bank of filters for target detection. The next paragraph illustrates these issues with several examples using different combinations of K_j and N_j.

Assume that we have $J = 1$ image and we drop the subscript j for simplicity. Consider, for example, the four scenarios shown in Figure 2.4. In Figure 2.4(a), we assume we have the complete knowledge of the target signature except for the unknown gain. Also, the target orientation is assumed known. For this case, we have $K = 1$ unknown in the target template and we should use $\tilde{s}\tilde{b}$ instead of $\mathbf{Sb}$ to model the target return. In Figure 2.4(b), we assume that we have the incomplete knowledge of the target signature; i.e., we assume that we know the locations of the unknown scalars that represent the bright returns from the target scatterers. The number of the unknown scalars in the target template is assumed to be $K = 10$ for this case. We also assume that the target orientation is known and the number of pixels in the target template is, say, $N = 100$. For this case, then, $\mathbf{S}$ is a 100×10 matrix and $\mathbf{b}$ is a 10×1 vector. In Figure 2.4(c), we assume that the target orientation and the target approximate size are known. For this case, all pixels in the target template are assumed unknown and we assume that there are $K = 100$ unknowns in the target template. For this case, then, $\mathbf{S}$ is a 100×100 identity matrix and $\mathbf{b}$ is a 100×1 vector. In Figure 2.4(d), we assume that only the approximate size of the target is known. Since the target orientation is unknown, we choose a target template that is large enough to cover all possible target orientations and all pixels in the template are assumed unknown. We assume that there are $K = 1000$ unknowns in the target template. For this case, then, $\mathbf{S}$ is a 1000×1000 identity matrix and $\mathbf{b}$ is a 1000×1 vector. In the following sections, we shall show how the target detection performance is affected by the use of these target templates.

2.4 A CFAR Detector Based on Known Clutter and Noise Variance and Its Performance

We shall present below a generalized likelihood ratio (GLR) detector under the assumption that the clutter and noise variance in each image; i.e., σ_j^2, is known, and also present its performance. This detector is referred to as *Detector A* below and its performance is easy to analyze.

It is shown in Appendix A that Detector A has the form

$$\sum_{j=1}^{J} \frac{\mathbf{z}_j^H \mathbf{S}_j \mathbf{S}_j^H \mathbf{z}_j}{\sigma_j^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{2.11}$$

The threshold parameter γ is determined according to a given probability of false alarm and thus the detector is a CFAR detector.

It is shown in Appendix B that the probability of false alarm of Detector A is

$$P_F = \sum_{k=0}^{K-1} \frac{\gamma^{K-1-k}}{(K-1-k)!} \exp(-\gamma). \tag{2.12}$$

We note that the P_F depends only on K, the sum of the numbers of unknown parameters in the target templates, and the threshold parameter γ. The P_F is independent of N_j, the sizes of the templates. For a given probability of false alarm, the γ in the detector (2.11) is obtained with (2.12). It is also shown in Appendix B that the probability of detection of Detector A is

$$P_D = \exp(-\delta - \gamma) \sum_{i=0}^{\infty} \frac{\delta^i}{i!} \sum_{k=0}^{i+K-1} \frac{\gamma^{i+K-1-k}}{(i+K-1-k)!}, \tag{2.13}$$

where

$$\delta = \sum_{j=1}^{J} \frac{\mathbf{b}_j^H \mathbf{b}_j}{\sigma_j^2}. \tag{2.14}$$

The δ is the sum of the signal-to-clutter-and-noise ratios (SCNRs) of the templates. Note that δ is the sum of the signal-to-clutter-and-noise ratios of the non-zero pixels in all templates. We also note that the P_D is also independent of the template sizes N_j. The P_D depends on K, SCNR δ, and the threshold γ.

In most of the following examples, we consider $P_F = 10^{-10}$ for the target templates. This is because we intend to achieve a false alarm rate of 10^{-4} per km^2, where km denotes the kilometer. (ARPA's goal is to achieve a false alarm rate of 10^{-3} per km^2.) For a SAR image with range and cross range resolutions of 0.34 and 3 meters, respectively, the probability of false alarm of a single pixel detector is approximately

10^{-10}. When the target template size is much smaller than the image covering the one km^2 area, the template false alarm rate should also be approximately 10^{-10} [9].

To compare the performance of using different target templates with the Detector A, let us consider the effect of K on the performance of Detector A. Figure 2.5 shows the probability of detection as a function of SCNR for different K when $P_F = 10^{-10}$. When we have one image; i.e., $J = 1$, for example, the four performance curves shown in Figure 2.5 could correspond to the four scenarios shown in Figure 2.4. It is shown that for the given P_F and a fixed SCNR δ, the P_D of the detector in (2.11) decreases as K increases. We note from Figure 2.5 that the best performance occurs when we use the complete knowledge of the target signature except for the unknown gain and when the target orientation is known. The worst performance occurs for the case of unknown target signature and unknown target orientation. This result occurs due to the non-coherent integration loss [10, p.71] since as may be seen from (2.43) in Appendix B, the energies of the template pixels are summed up non-coherently.

Note also that for SCNR = 15 dB in Figure 2.5, the average SCNR *per pixel* is 15 dB for $K = 1$ and is 5 dB for $K = 10$. Thus Figure 2.5 also shows the effect of smearing a given amount target energy among more pixels and hence lowering the SCNR per pixel.

Figure 2.6 shows the extra SCNR needed to achieve $P_D = 0.5$ for different probabilities of false alarm P_F. We note that we have similar curves for different probabilities of false alarm P_F. The extra SCNR needed decreases slightly as P_F decreases. We also note that the extra SCNR needed increases slowly as K increases.

Although it is the best to use the complete knowledge of the target signature for target detection, the target signature may not be completely known for SAR images and even when known, the signature may change due to the interaction between the target and clutter and other factors. The change of signature may result in severe detector performance degradation. Let $\tilde{\mathrm{s}}_j$ and $\hat{\mathrm{s}}_j$ be the assumed and true target signature vectors, respectively, where both $\tilde{\mathbf{s}}_j$ and $\hat{\mathbf{s}}_j$ have Euclidean norm 1. It is shown in Appendix C that the probability of false alarm for this case is the same as (2.12) with $K = J$. The probability of detection for this case has the form of (2.13) with $K = J$ and

$$\delta = \sum_{j=1}^{J} \frac{|\tilde{b}_j|^2}{\sigma_j^2} \rho_j, \tag{2.15}$$

where ρ_j $(0 \leq \rho_j \leq 1)$ are the SCNR loss factors as a result of the signature mismatch; i.e.,

$$\rho_j = |\hat{\mathbf{s}}_j^H \tilde{\mathbf{s}}_j|^2. \tag{2.16}$$

Figure 2.7 shows the SCNR loss as a function of ρ_1 when $J = 1$.

We now make the comparison between using the incomplete knowledge of the target signature information and the approach of using the complete knowledge of the target signature information except for the unknown gain. The comparison may be made most easily with $J = 1$. For $J = 1$, we drop the subscript of ρ and K. Consider the example where $J = 1$ and we use an incomplete target signature described by $K = 20$ unknown parameters. The extra SCNR needed to achieve the same probability of detection as when using the complete target signature information ($K = 1$) is about 3 dB, as shown in Figure 2.6. If we know that the mismatch between the assumed and true target signatures is small (possibly through experiments), or more specifically, if the mismatch will not result in $\rho < 0.5$, then it is better to use the complete knowledge of the target signature information in the detector. Otherwise, it is better to use the incomplete signature information and assume $K = 20$ unknowns in the target template for this example.

The comparison between considering one target template at a time and one pixel at a time is also clear. As shown in Appendix C, for the same target SCNR, the performance of assuming the complete knowledge of the target signature is the same as $K_j = 1$, $j = 1, 2, \cdots, J$, in our problem formulation. For $J = 1$ and a template with $K = 20$ unknown parameters, for example, the extra SCNR needed to achieve the same probability of detection as for $K = 1$ is about 3 dB, as shown in Figure 2.6. Then for the case where the target return is a single bright pixel, the net loss of using the target template is about 3 dB, as compared with using one pixel at a time. If the template SCNR δ is at least 3 dB larger than the SCNR of the highest pixel in the template, then, using the target template model is better than the single pixel detector. For the best possible case where all $K = 20$ pixels in the target template have equal magnitude, the template SCNR is about 13 dB more than the individual pixel SCNR. Thus the net gain of using the target template for target detection is about 10 dB because of the 3 dB loss due to the increased number of unknowns.

We now compare the performance of using a bank of target templates at a time and using a single pixel at a time for target detection. Note that the maximum probability of false alarm of using a filter bank for target detection is the number of target templates or filters times the probability of false alarm of using a single target template or filter. Thus we may set the probability of false alarm of each target template to be the desired probability false alarm divided by the number of filters. The probability of false alarm of the single pixel detector, however, is the same as the the desired probability false alarm. Thus we encounter an additional SCNR loss when using a bank of target templates for target detection due to the decreased probability of false alarm for each target template. Yet when the number of filters in a filter bank is small, this SCNR loss is negligible. When the number of filters is 4, for example, this SCNR loss is much less than 1 dB. (See Figure 2.10 for example.)

Finally, let us consider the advantages of using multiple SAR images for target detection. Consider the example where $J = 2$ and $K_1 = K_2 = 20$. For the best case where the target template SCNRs for both images are assumed the same, the net gain of using both images for target detection is about 2 dB because of the 1 dB loss due to doubling the number of unknowns, as shown in Figure 2.6.

2.5 A CFAR Detector Based on Unknown Clutter and Noise Variance and Its Performance

In the previous section, we have studied the performance of a CFAR detector that assumes that the clutter and noise variance in each image is known. In practice, however, the clutter and noise variances are unknown. We present below a detector for this practical situation. This CFAR detector is referred to as *Detector B.* Our approach is similar to the one developed by Kelly [11] for target detection with a phased array airborne surveillance radar.

The Detector B we shall present utilizes both primary and secondary data of a SAR image for target detection. The data vector $\mathbf{z}_j$, from which the target presence is sought, is referred to as the primary data. For the jth SAR image, the secondary data vectors are denoted $\mathbf{z}_j(1)$, $\mathbf{z}_j(2)$, $\cdots$, $\mathbf{z}_j(L_j)$. These vectors are assumed to be target free; i.e., they represent the target free background of the jth SAR image. They are assumed to have the same statistics as $\mathbf{z}_j$, the subvector of the primary data vector $\mathbf{z}$, under hypothesis H_0 and are statistically independent of each other and $\mathbf{z}$. The secondary data are useful for estimating the clutter and noise variance in Detector B.

It is shown in Appendix D that Detector B has the form

$$\prod_{j=1}^{J}\left[\frac{\mathbf{z}_j^H\mathbf{z}_j+\sum_{l=1}^{L_j}\mathbf{z}_j^H(l)\mathbf{z}_j(l)}{\mathbf{z}_j^H\left(\mathbf{I}_{N_j}-\mathbf{S}_j\mathbf{S}_j^H\right)\mathbf{z}_j+\sum_{l=1}^{L_j}\mathbf{z}_j^H(l)\mathbf{z}_j(l)}\right]^{(L_j+1)N_j} \underset{H_0}{\overset{H_1}{\gtrless}} \xi. \tag{2.17}$$

The threshold parameter ξ is determined according to a given probability of false alarm and thus Detector B is a CFAR detector.

In general, we simply set the number of the secondary data vectors to be equal to each other; i.e., we let

$$L_1 = L_2 = \cdots = L_J \triangleq L_0. \tag{2.18}$$

Taking the natural logarithm of the (L_0+1)st root of (2.17) yields

$$\eta = \sum_{j=1}^{J} \ln \left[\frac{\mathbf{z}_j^H \mathbf{z}_j + \sum_{l=1}^{L_0} \mathbf{z}_j^H(l)\mathbf{z}_j(l)}{\mathbf{z}_j^H \left(\mathbf{I}_{N_j} - \mathbf{S}_j\mathbf{S}_j^H\right)\mathbf{z}_j + \sum_{l=1}^{L_0} \mathbf{z}_j^H(l)\mathbf{z}_j(l)} \right] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{2.19}$$

where $\gamma = (\ln \xi)/[(L_0+1)N_j]$.

It is shown in Appendix E that except for $J = 1$, there are no closed form expressions for the probability of detection P_D and probability of false alarm P_F. It is shown that under hypothesis H_0, η has the pdf

$$f_\eta(\eta|H_0) = f_{\eta_1}(\eta|H_0) * f_{\eta_2}(\eta|H_0) * \cdots * f_{\eta_J}(\eta|H_0), \tag{2.20}$$

where $*$ denotes the convolution and for $j = 1, 2, \cdots, J$,

$$f_{\eta_j}(\eta|H_0) = \frac{[\exp(\eta)-1]^{K_j-1}\,(M_j+K_j-1)!}{(K_j-1)!(M_j-1)!\exp[(M_j+K_j-1)\eta]}, \tag{2.21}$$

with

$$M_j = N_j(L_0+1) - K_j. \tag{2.22}$$

Under hypothesis H_1, η has the pdf

$$f_\eta(\eta|H_1) = f_{\eta_1}(\eta|H_1) * f_{\eta_2}(\eta|H_1) * \cdots * f_{\eta_J}(\eta|H_1), \tag{2.23}$$

where for $j = 1, 2, \cdots, J$,

$$f_{\eta_j}(\eta|H_1) = \exp(-\delta_j) \sum_{i=0}^{\infty} \frac{\delta_j^i [\exp(\eta)-1]^{K_j+i-1}(M_j+K_j+i-1)!}{i!(M_j-1)!(K_j+i-1)!\exp[(M_j+K_j+i-1)\eta]}, \tag{2.24}$$

where δ_j is the SCNR of the jth template.

$$\delta_j = \frac{\mathbf{b}_j^H \mathbf{b}_j}{\sigma_j^2}. \tag{2.25}$$

The P_F and P_D may be obtained numerically from $f_\eta(\eta|H_0)$ and $f_\eta(\eta|H_1)$, respectively; i.e.,

$$P_F = \int_\gamma^\infty f_\eta(\eta|H_0)d\eta, \tag{2.26}$$

and

$$P_D = \int_{\gamma}^{\infty} f_\eta(\eta|H_1)d\eta. \tag{2.27}$$

The $f_\eta(\eta|H_0)$ given in (2.20) and $f_\eta(\eta|H_1)$ given in (2.23) may be obtained more efficiently by first calculating the characteristic functions of $f_{\eta_j}(\eta|H_0)$ and $f_{\eta_j}(\eta|H_1)$ with FFT (Fast Fourier Transform). For a given probability of false alarm; i.e., to achieve CFAR, the γ in the detector (2.19) is obtained with (2.26).

We remark that the above analysis also holds when L_0 is not an integer but a rational number such that N_jL_0, $j = 1, 2, \cdots, J$, are integers since the number of target free pixels in each image does not have to be the multiples of the size of the corresponding target template. The detector (2.17) or (2.19) may be changed slightly to accommodate the fact that L_0 is not an integer.

Consider next $J = 1$, where we have the closed form expressions for P_D and P_F. For this case, we drop the subscripts of $\mathbf{z}$, $\mathbf{z}(l)$, δ, M, K, and N. Taking the $(L+1)$st root of (2.17) yields the detector for $J = 1$:

$$\frac{\mathbf{z}^H\mathbf{z} + \sum_{l=1}^{L} \mathbf{z}^H(l)\mathbf{z}(l)}{\mathbf{z}^H\left(\mathbf{I}_N - \mathbf{S}\mathbf{S}^H\right)\mathbf{z} + \sum_{l=1}^{L} \mathbf{z}^H(l)\mathbf{z}(l)} \overset{H_1}{\underset{H_0}{\gtrless}} \zeta. \tag{2.28}$$

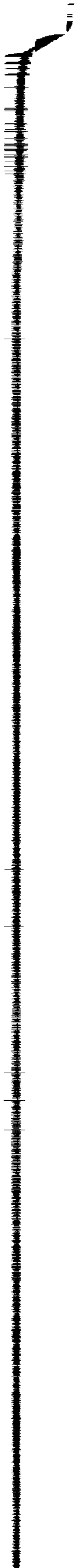
It is shown in Appendix E that the probability of false alarm of the above detector is

$$P_F = \sum_{k=0}^{K-1} \frac{(\zeta-1)^{K-k-1}}{\zeta^{M+K-k-1}} \begin{pmatrix} M+K-k-2 \\ K-k-1 \end{pmatrix}. \tag{2.29}$$

We note that for $J = 1$, the P_F of Detector B depends on the dimensional parameters M and K and the threshold parameter ζ. For a given probability of false alarm, the ζ in the detector (2.17) is obtained with (2.29).

It is also shown in Appendix E that the probability of detection of the above detector when $J = 1$ is

$$P_D = 1 - \exp(-\delta/\zeta) \sum_{k=0}^{M-1} \sum_{i=0}^{M-k-1} \begin{pmatrix} M+K-k-2 \\ K+i-1 \end{pmatrix} \frac{(\delta/\zeta)^i}{i!} \frac{(\zeta-1)^{K+i}}{\zeta^{M+K-k-1}}. \tag{2.30}$$



We note that for $J = 1$, P_D depends on the dimensional parameters M and K, the SCNR δ, and the threshold ζ.

We now examine how the different dimensional parameters affect the performance of the detector in (2.17). These effects of the parameters may be most easily explained with the case of $J = 1$. The effects are then generalized to the case of $J > 1$. Let $J = 1$ and consider $M = NL + N - K$, the total number of target free pixels in the primary and secondary data vectors. The larger the number of target free pixels M, the better the estimate of the clutter and noise variance, and hence the closer the performance of Detector B to that of Detector A. This result may be observed from Figure 2.8, which shows the probability of detection as a function of SCNR for different M when $J = 1$, $K = 2$ and $P_F = 10^{-10}$.

Consider next the effect of K, the number of unknown parameters in the target template, on the performance of Detector B. We first explain that the larger the K, the more number of target free pixels M is needed by Detector B to achieve similar performance as Detector A. We shall consider $J = 1$ and drop the subscript j for convenience. As shown in Appendix B, we may rewrite Detector A in (2.11) as

$$\eta' = \frac{\bar{\bar{\mathbf{z}}}_A^H \bar{\bar{\mathbf{z}}}_A}{K} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma'. \tag{2.31}$$

Under hypothesis H_0, $\bar{\bar{\mathbf{z}}}_A$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_K$. Under hypothesis H_1, $\bar{\bar{\mathbf{z}}}_A$ has the complex Gaussian distribution with mean $\mathbf{b}/\sigma$ and covariance matrix $\mathbf{I}_K$. Thus under hypothesis H_0, the mean and variance of η' are 1 and $1/K$, respectively. Under hypothesis H_1, the mean and variance of η' are $1 + \delta/K$ and $1/K + \delta/K$, respectively, where δ is the SCNR given in (2.14). As shown in Appendix E, we may rewrite Detector B in (2.17) as

$$\eta_1'' = \frac{\bar{\mathbf{z}}_A^H \bar{\mathbf{z}}_A / K}{\left[\bar{\mathbf{z}}_B^H \bar{\mathbf{z}}_B + \sum_{l=1}^{L} \bar{\mathbf{z}}^H(l) \bar{\mathbf{z}}(l)\right] / M} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta''. \tag{2.32}$$

The statistical properties of the *numerator* of η_1'' is the same as the properties of η' in

(2.31). Since under both hypotheses, the $\bar{\mathbf{z}}(l)$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_N$, the mean and variance of the *denominator* of η_1' are 1 and $1/M$, respectively. We note that for large K and small M, the performance of (2.32) will be affected by the variance $1/M$ of the denominator of η_1''. Thus for large K, M must also be large in order for Detector B in (2.32) to achieve similar performance as Detector A in (2.31) that assumes that the clutter and noise variance is known. We found through numerical examples that for $1 \leq K \leq 1000$, the performance differences between Detectors A and B are similar for different K when M is proportional to $K^{2/3}$. Figure 2.9 shows the probability of detection as a function of SCNR for different K when $J = 1$, $M = \lfloor 48K^{2/3} \rfloor$ and $P_F = 10^{-10}$, where $\lfloor \times \rfloor$ denotes the integer part of $\times$. The figure also shows the performance of Detector A for comparison. Note that Detectors A and B have similar performances.

Figure 2.10 shows the probability of detection as a function of probability of false alarm; i.e., the receiver operating characteristic of Detector B, for different K and SCNR when $J = 1$, $M = 118$, and $P_F = 10^{-10}$. We note that the 5 dB change in SCNR or the change in the number of unknowns K from 1 to 10 has a significant effect on the P_F for a given P_D. This significant change is due to the exponentially decreasing tail associated with the assumed Gaussian model for the clutter distribution.

The discussions for the case of a single image ($J = 1$) may be easily generalized to the case of multiple images ($J > 1$). Consider $J = 2$, for example. To achieve similar performance as Detector A, the number of target free pixels in each image must be large for Detector B. Figure 2.11 shows the probability of detection as a function of SCNR for different $M_1 = M_2$ when $J = 2$, $K_1 = K_2 = 2$ and $P_F = 10^{-10}$. We note that the larger the M_1 and M_2, the closer the performance of Detector B to that of Detector A, which is expected.

2.6 Target Detection with Experimental Data

We consider first the performance of Detector B when used with the experimental data shown in Figure 2.1. Figure 2.12 shows the detection results before and after coherent subtraction when $M = 48$, $N = K = 1$, and $P_F = 10^{-10}$. Figure 2.12 shows the generalized likelihood ratio (GLR) obtained with the left side of (2.28) when it is above the detection threshold and zero when it is below the threshold. We note that before coherent subtraction, although the presence of the target is detected, the large tree trunk returns also result in a false alarm. Thus in the presence of large tree trunk returns, Detector B is no longer a CFAR detector and its probability of false alarm is also determined by the number of large tree trunk returns. With coherent subtraction, however, the false alarm due to the large tree trunk return is eliminated and Detector B is truly CFAR.

Consider next the experimental data shown in Figure 2.2. We note that the target occupies more than one pixel. Let us assume that the target orientation is known and use an incomplete target signature described by $K = 20$ unknowns. (Note that the incomplete target signature may be determined by simulation softwares such as XPATCH [8] or by experiments in a laboratory. They should not be determined from images from which the target presence is sought. When the target SCNR is small, determining target templates from such images is impossible.) Then the template SCNR for the target is approximately 35 dB for the data shown in Figure 2.2(c). The largest pixel SCNR for the target is about 25 dB. Thus compared with using each single pixel for target detection, using the template with $K = 20$ unknowns results in a net gain of about 7 dB because of the 3 dB loss due to the increased number of unknowns.

Finally, we could add noise to the data shown in Figure 2.1 to simulate a weak target in clutter and noise. Assume that with the added noise, the template SCNR for the target is approximately 20 dB and the largest pixel SCNR for the target is about 10 dB. Then from Figure 2.9, we note that the probability of detection of using the target template is 1 while the probability of detection of using the single pixel detector is approximately 0.

2.7 Conclusions

We have considered target detection with synthetic aperture radar. We have derived generalized likelihood ratio (GLR) detection algorithms that may be used with SAR images that are obtained with coherent subtraction or have Gaussian distributions. Through performance analysis, we have analytically compared the performance of a) a single pixel detector, b) a detector using a complete knowledge of the target signature information and known orientation information, c) a detector using an incomplete knowledge of the target signature information and known orientation information, d) a detector using unknown target signature information and known orientation information, and e) a detector using unknown target signature information and unknown orientation information.

Acknowledgments

The first author gratefully acknowledges Maj. R. Williams for sponsoring her for the AFOSR Summer Faculty Research Program. The authors sincerely thank S. Wei for her expert help with the ERIM data. The authors are also grateful for the helpful discussions with R. Schindel, M. Minardi, and J. Leonard. The detailed and constructive comments from the reviewers and the editor, especially those from the AES Editor for Radar, have improved the quality of the paper and are deeply appreciated.

Appendix A: Derivation of Detector A

Under hypothesis H_1, the probability density function (pdf) of the complex Gaussian random vector $\mathbf{z}$ may be written

$$f_1(\mathbf{z}) = \prod_{j=1}^{J} \frac{1}{\pi^{N_j} \sigma_j^{2N_j}} \exp\left[-\frac{(\mathbf{z}_j - \mathbf{S}_j b_j)^H (\mathbf{z}_j - \mathbf{S}_j b_j)}{\sigma_j^2}\right]. \tag{2.33}$$

Under hypothesis H_0, the pdf of the complex Gaussian random vector $\mathbf{z}$ may be written

$$f_0(\mathbf{z}) = \prod_{j=1}^{J} \frac{1}{\pi^{N_j} \sigma^{2N_j}} \exp\left[-\frac{\mathbf{z}_j^H \mathbf{z}_j}{\sigma_j^2}\right]. \tag{2.34}$$

The generalized likelihood ratio has the form $\frac{\max_{\mathbf{b}} f_1}{f_0}$. Maximizing f_1 with respect to $\mathbf{b}_j$ yields

$$\hat{\mathbf{b}}_j = \mathbf{S}_j^H \mathbf{z}_j, \tag{2.35}$$

and

$$\max_{\mathbf{b}} f_1(\mathbf{z}) = \prod_{j=1}^{J} \frac{1}{\pi^{N_j} \sigma^{2N_j}} \exp\left[-\frac{\mathbf{z}_j^H (\mathbf{I}_{N_j} - \mathbf{S}_j \mathbf{S}_j^H) \mathbf{z}_j}{\sigma_j^2}\right], \tag{2.36}$$

where we have used (2.8). Then the generalized likelihood ratio test becomes

$$\prod_{j=1}^{J} \exp\left[\frac{\mathbf{z}_j^H \mathbf{S}_j \mathbf{S}_j^H \mathbf{z}_j}{\sigma_j^2}\right] \underset{H_0}{\overset{H_1}{\gtrless}} \xi. \tag{2.37}$$

Taking the natural logarithm of both sides of the above test yields the optimal detector

$$\sum_{j=1}^{J} \frac{\mathbf{z}_j^H \mathbf{S}_j \mathbf{S}_j^H \mathbf{z}_j}{\sigma_j^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{2.38}$$

where $\gamma = \ln \xi$.

Appendix B: Performance of Detector A

Consider the GLR test in (2.11). Let $\bar{\mathbf{z}}_j = \mathbf{z}_j / \sigma_j$. Then under hypothesis H_0, $\bar{\mathbf{z}}_j$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_{N_j}$.

Under hypothesis H_1, $\bar{\mathbf{z}}_j$ has the complex Gaussian distribution with mean $\mathbf{S}_j\mathbf{b}_j/\sigma_j$ and covariance matrix $\mathbf{I}_{N_j}$. The GLR test in (2.11) may now be written

$$\sum_{j=1}^{J} \bar{\mathbf{z}}_j^H \mathbf{S}_j \mathbf{S}_j^H \bar{\mathbf{z}}_j \mathop{\gtrless}_{H_0}^{H_1} \gamma. \tag{2.39}$$

Since $\mathbf{S}_j^H\mathbf{S}_j = \mathbf{I}_{N_j}$, there exists a unitary matrix $\mathbf{U}_j$ such that

$$\mathbf{U}_j\mathbf{S}_j = \begin{bmatrix} \mathbf{I}_{K_j} \\ \mathbf{0} \end{bmatrix}. \tag{2.40}$$

Let $\bar{\bar{\mathbf{z}}}_j = \mathbf{U}_j\bar{\mathbf{z}}_j$. Then under hypothesis H_0, $\bar{\bar{\mathbf{z}}}_j$ still has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_{N_j}$. Under hypothesis H_1, $\bar{\bar{\mathbf{z}}}_j$ has the complex Gaussian distribution with mean $\mathbf{U}_j\mathbf{S}_j\mathbf{b}_j/\sigma_j$ and covariance matrix $\mathbf{I}_{N_j}$. The GLR test in (2.39) may now be written

$$\sum_{j=1}^{J} \bar{\bar{\mathbf{z}}}_j^H \mathbf{U}_j \mathbf{S}_j \mathbf{S}_j^H \mathbf{U}_j^H \bar{\bar{\mathbf{z}}}_j \mathop{\gtrless}_{H_0}^{H_1} \gamma. \tag{2.41}$$

Let

$$\bar{\bar{\mathbf{z}}}_j = \begin{bmatrix} \bar{\bar{\mathbf{z}}}_{jA} \\ \bar{\bar{\mathbf{z}}}_{jB} \end{bmatrix}, \tag{2.42}$$

where $\bar{\bar{\mathbf{z}}}_{jA}$ and $\bar{\bar{\mathbf{z}}}_{jB}$ are $K_j \times 1$ and $(N_j - K_j) \times 1$ vectors, respectively. By using (2.40) and (2.42), (2.41) may be written

$$\eta = \sum_{j=1}^{J} \bar{\bar{\mathbf{z}}}_{jA}^H \bar{\bar{\mathbf{z}}}_{jA} \mathop{\gtrless}_{H_0}^{H_1} \gamma. \tag{2.43}$$

Under hypothesis H_0, $\bar{\bar{\mathbf{z}}}_{jA}$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_{K_j}$. Under hypothesis H_1, $\bar{\bar{\mathbf{z}}}_{jA}$ has the complex Gaussian distribution with mean $\mathbf{b}_j/\sigma_j$ and covariance matrix $\mathbf{I}_{K_j}$. Thus under hypothesis H_0, 2η has the central χ^2 distribution with $2K$ degrees of freedom since $K = K_1 + K_2 + \cdots + K_J$; i.e.,

$$f_\eta(\eta|H_0) = \frac{\eta^{K-1}\exp(-\eta)}{(K-1)!}. \tag{2.44}$$

Then the probability of false alarm is

$$\begin{aligned} P_F &= \int_\gamma^\infty f_\eta(\eta|H_0)d\eta \\ &= \sum_{k=0}^{K-1} \frac{\gamma^{K-1-k}}{(K-1-k)!}\exp(-\gamma). \end{aligned} \quad (2.45)$$

Under hypothesis H_1, 2η has the noncentral χ^2 distribution with $2K$ degrees of freedom and noncentrality parameter $\delta = \sum_{j=1}^{J} \mathbf{b}_j^H \mathbf{b}_j / \sigma_j^2$; i.e.,

$$f_\eta(\eta|H_1) = \exp(-\delta - \eta) \sum_{i=0}^{\infty} \frac{\delta^i}{i!} \frac{\eta^{i+K-1}}{(i+K-1)!}. \quad (2.46)$$

Then the probability of detection is

$$\begin{aligned} P_D &= \int_\gamma^\infty f_\eta(\eta|H_1)d\eta \\ &= \exp(-\delta - \gamma) \sum_{i=0}^{\infty} \frac{\delta^i}{i!} \sum_{k=0}^{i+K-1} \frac{\gamma^{i+K-1-k}}{(i+K-1-k)!}. \end{aligned} \quad (2.47)$$

Appendix C: Effect of Target Signature Mismatch

When the target signatures $\tilde{\mathbf{s}}_j$ are used in the GLR test in (2.11), (2.11) becomes

$$\sum_{j=1}^{J} \frac{\mathbf{z}_j^H \tilde{\mathbf{s}}_j \tilde{\mathbf{s}}_j^H \mathbf{z}_j}{\sigma_j^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \quad (2.48)$$

where $\tilde{\mathbf{s}}_j^H \tilde{\mathbf{s}}_j = 1$. Under hypothesis H_0, $\mathbf{z}_j$ has the complex Gaussian distribution with zero-mean and covariance matrix $\sigma^2 \mathbf{I}_{N_j}$. Under hypothesis H_1, $\mathbf{z}_j$ has the complex Gaussian distribution with mean $\hat{\mathbf{s}}_j \tilde{b}_j$ and covariance matrix $\sigma_j^2 \mathbf{I}_{N_j}$.

Through similar transformations as in Appendix A, the test (2.48) may be written

$$\eta = \sum_{j=1}^{J} |\bar{\bar{z}}_{jA}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \quad (2.49)$$

Under hypothesis H_0, $\bar{\bar{z}}_{jA}$ has the complex Gaussian distribution with zero-mean and variance 1. Under hypothesis H_1, $\bar{\bar{z}}_{jA}$ has the complex Gaussian distribution with mean $\tilde{\mathbf{s}}_j^H \hat{\mathbf{s}}_j \tilde{b}_j / \sigma_j$ and variance 1. Thus under hypothesis H_0, 2η has the central χ^2

distribution with 2 degrees of freedom. Under hypothesis H_1, 2η has the noncentral χ^2 distribution with 2 degrees of freedom and noncentrality parameter

$$\delta = \sum_{j=1}^{J} \frac{|\tilde{b}_j|^2}{\sigma_j^2} |\hat{\mathbf{s}}_j^H \tilde{\mathbf{s}}_j|^2. \tag{2.50}$$

We note that in the absence of target signature mismatch; i.e., when $\hat{\mathbf{s}}_j = \tilde{\mathbf{s}}_j$, the δ in (2.50) becomes

$$\delta = \sum_{j=1}^{J} \frac{|\tilde{b}_j|^2}{\sigma_j^2}, \tag{2.51}$$

which is the SCNR for the case where the target signature is known completely.

Appendix D: Derivation of Detector B

The derivations below are both a specialization (due to the uncorrelated clutter and noise pixels) and a generalization (due to multiple images) of [11]. Under hypothesis H_1, the probability density function (pdf) of the complex Gaussian random vectors $\mathbf{z}_j$, $\mathbf{z}_j(1)$, $\cdots$, $\mathbf{z}_j(L_j)$, $j = 1, 2, \cdots, J$, may be written

$$f_1\left(\mathbf{z}_j, \mathbf{z}_j(1), \cdots, \mathbf{z}_j(L_j), j = 1, 2, \cdots, J\right) = \prod_{j=1}^{J} \left[\frac{1}{\pi^{N_j} \sigma_j^{2N_j}} \exp\left(-\frac{T_{1j}}{\sigma_j^2} \right) \right]^{L_j+1}, \tag{2.52}$$

where

$$T_{1j} = \frac{1}{L_j + 1} \left[(\mathbf{z}_j - \mathbf{S}_j b_j)^H (\mathbf{z}_j - \mathbf{S}_j b_j) + \sum_{l=1}^{L_j} \mathbf{z}_j^H(l) \mathbf{z}_j(l) \right]. \tag{2.53}$$

Under hypothesis H_0, the pdf of the complex Gaussian random vectors $\mathbf{z}_j$, $\mathbf{z}_j(1)$, $\cdots$, $\mathbf{z}_j(L_j)$, $j = 1, 2, \cdots, J$, may be written

$$f_0\left(\mathbf{z}_j, \mathbf{z}_j(1), \cdots, \mathbf{z}_j(L_j), j = 1, 2, \cdots, J\right) = \prod_{j=1}^{J} \left[\frac{1}{\pi^{N_j} \sigma_j^{2N_j}} \exp\left(-\frac{T_{0j}}{\sigma_j^2} \right) \right]^{L_j+1}, \tag{2.54}$$

where

$$T_{0j} = \frac{1}{L_j + 1} \left[\mathbf{z}_j^H \mathbf{z}_j + \sum_{l=1}^{L_j} \mathbf{z}_j^H(l) \mathbf{z}_j(l) \right]. \tag{2.55}$$

The generalized likelihood ratio has the form $\frac{\max_{\sigma_1^2,\cdots,\sigma_J^2,\mathbf{b}} f_1}{\max_{\sigma_1^2,\cdots,\sigma_J^2} f_0}$. Maximizing f_0 with respect to σ_j^2 yields

$$\hat{\sigma}_j^2 = \frac{1}{N_j} T_{0j}. \tag{2.56}$$

Thus

$$\max_{\sigma_1^2,\cdots,\sigma_J^2} f_0 = \prod_{j=1}^{J} \left(\frac{N_j}{e\pi T_{0j}} \right)^{(L_j+1)N_j}. \tag{2.57}$$

Maximizing f_1 with respect to σ_j^2 yields

$$\hat{\sigma}_j^2 = \frac{1}{N_j} T_{1j}. \tag{2.58}$$

Thus

$$\max_{\sigma_1^2,\cdots,\sigma_J^2} f_1 = \left(\frac{N_j}{e\pi T_{1j}} \right)^{(L_j+1)N_j}. \tag{2.59}$$

Minimizing T_{1j} with respect to $\mathbf{b}_j$ gives

$$\hat{\mathbf{b}}_j = \mathbf{S}_j^H \mathbf{z}_j, \tag{2.60}$$

and

$$\min_{\mathbf{b}_j} T_{1j} = \frac{1}{L_j+1} \left[\mathbf{z}_j^H \left(\mathbf{I}_{N_j} - \mathbf{S}_j \mathbf{S}_j^H \right) \mathbf{z}_j + \sum_{l=1}^{L_j} \mathbf{z}_j^H(l) \mathbf{z}_j(l) \right]. \tag{2.61}$$

Then the generalized likelihood ratio test becomes

$$\prod_{j=1}^{J} \left[\frac{\mathbf{z}_j^H \mathbf{z}_j + \sum_{l=1}^{L_j} \mathbf{z}_j^H(l) \mathbf{z}_j(l)}{\mathbf{z}_j^H \left(\mathbf{I}_{N_j} - \mathbf{S}_j \mathbf{S}_j^H \right) \mathbf{z}_j + \sum_{l=1}^{L_j} \mathbf{z}_j^H(l) \mathbf{z}_j(l)} \right]^{(L_j+1)N_j} \underset{H_0}{\overset{H_1}{\gtrless}} \xi. \tag{2.62}$$

Appendix E: Performance of Detector B

The derivations below are again both a specialization (due to the uncorrelated clutter and noise pixels) and a generalization (due to multiple images) of [11].

$\underline{J = 1:}$

Consider first $J = 1$ and the GLR test in (2.28). Let

$$\eta_1' = \eta' - 1. \tag{2.63}$$

Note that η_1' is a monotonically increasing function of η'. Then (2.17) may be rewritten

$$\eta_1' = \frac{\mathbf{z}^H \mathbf{S}\mathbf{S}^H \mathbf{z}}{\mathbf{z}^H \left(\mathbf{I}_N - \mathbf{S}\mathbf{S}^H\right) \mathbf{z} + \sum_{l=1}^{L} \mathbf{z}^H(l)\mathbf{z}(l)} \overset{H_1}{\underset{H_0}{\gtrless}} \zeta - 1. \tag{2.64}$$

Let

$$\bar{\mathbf{z}} = \frac{\mathbf{U}\mathbf{z}}{\sigma}, \tag{2.65}$$

and

$$\bar{\mathbf{z}}(l) = \frac{\mathbf{U}\mathbf{z}(l)}{\sigma}, \quad l = 1, 2, \cdots, L, \tag{2.66}$$

where $\mathbf{U}$ is a unitary matrix that satisfies (2.40). Then under hypothesis H_0, $\bar{\mathbf{z}}$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_N$. Under hypothesis H_1, $\bar{\mathbf{z}}$ has the complex Gaussian distribution with mean $\mathbf{U}\mathbf{S}\mathbf{b}/\sigma$ and covariance matrix $\mathbf{I}_N$. Under both hypotheses, the $\bar{\mathbf{z}}(l)$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_N$. By using (2.65) and (2.66), (2.64) becomes

$$\eta_1' = \frac{\bar{\mathbf{z}}_A^H \bar{\mathbf{z}}_A}{\bar{\mathbf{z}}_B^H \bar{\mathbf{z}}_B + \sum_{l=1}^{L} \bar{\mathbf{z}}^H(l)\bar{\mathbf{z}}(l)} \overset{H_1}{\underset{H_0}{\gtrless}} \zeta - 1, \tag{2.67}$$

where

$$\bar{\mathbf{z}} = \begin{bmatrix} \bar{\mathbf{z}}_A \\ \bar{\mathbf{z}}_B \end{bmatrix}, \tag{2.68}$$

with $\bar{\mathbf{z}}_A$ and $\bar{\mathbf{z}}_B$ denoting $K \times 1$ and $(N-K) \times 1$ vectors, respectively. Under hypothesis H_0, $\bar{\mathbf{z}}_A$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_K$. Under hypothesis H_1, $\bar{\mathbf{z}}_A$ has the complex Gaussian distribution with mean $\mathbf{b}/\sigma$ and covariance matrix $\mathbf{I}_K$. Under both hypotheses, $\bar{\mathbf{z}}_B$ has the complex Gaussian distribution with zero-mean and covariance matrix $\mathbf{I}_{N-K}$. Let

$$t = \bar{\mathbf{z}}_A^H \bar{\mathbf{z}}_A, \tag{2.69}$$

and

$$\tau = \bar{\mathbf{z}}_B^H \bar{\mathbf{z}}_B + \sum_{l=1}^{L} \bar{\mathbf{z}}^H(l)\bar{\mathbf{z}}(l). \tag{2.70}$$

Then (2.67) becomes

$$t \underset{H_0}{\overset{H_1}{\gtrless}} (\zeta - 1)\tau. \tag{2.71}$$

Under hypothesis H_0, $2t$ has the central χ^2 distribution with $2K$ degrees of freedom. Under hypothesis H_1, $2t$ has the noncentral χ^2 distribution with $2K$ degrees of freedom and noncentrality parameter $\delta = \mathbf{b}^H\mathbf{b}/\sigma^2$. Under both hypotheses, 2τ has the central χ^2 distribution with $2(NL + N - K) = 2M$ degrees of freedom. The t and τ are independent of each other. Thus under hypothesis H_0, $\eta_1' M/K = Mt/(K\tau)$ has the central F-distribution with $2K$ and $2M$ degrees of freedom. Under hypothesis H_1, $\eta_1' M/K = Mt/(K\tau)$ has the noncentral F-distribution with $2K$ and $2M$ degrees of freedom and noncentrality parameter δ.

The probability of false alarm may now be calculated as

$$\begin{aligned} P_F &= P(t > (\zeta - 1)\tau | H_0) \\ &= \int_0^\infty \left[\int_{(\zeta-1)\tau}^\infty f_t(t|H_0)dt\right] f_\tau(\tau)d\tau \\ &= \sum_{k=0}^{K-1} \frac{(\zeta - 1)^{K-k-1}}{\zeta^{M+K-k-1}} \frac{(M+K-k-2)!}{(K-k-1)!(M-1)!}, \end{aligned} \tag{2.72}$$

where $f_t(t|H_i)$ and $f_{\tau|H_i}(\tau)$ denote the conditional probability density functions of t and τ, respectively, under hypothesis H_i, $i = 1, 2$.

The probability of detection may be calculated as

$$\begin{aligned} P_D &= 1 - \int_0^\infty \left[\int_{t/(\zeta-1)}^\infty f_\tau(\tau)d\tau\right] f_t(t|H_1)dt \\ &= 1 - \exp(-\delta) \sum_{k=0}^{M-1} \frac{(\zeta - 1)^{-(M-k-1)}}{(M-k-1)!} \\ &\qquad \sum_{i=0}^{\infty} \frac{\delta^i}{i!} \frac{(M+K+i-k-2)!}{(i+K-1)!} \left(\frac{\zeta - 1}{\zeta}\right)^{M+K+i-k-1}. \end{aligned} \tag{2.73}$$

Using the finite sum expression derived by Kelly in [12], we have

$$P_D = 1 - \exp(-\delta/\zeta) \sum_{k=0}^{M-1} \sum_{i=0}^{M-k-1} \begin{pmatrix} M+K-k-2 \\ K+i-1 \end{pmatrix} \frac{(\delta/\zeta)^i}{i!} \frac{(\zeta - 1)^{K+i}}{\zeta^{M+K-k-1}}. \tag{2.74}$$

$\underline{J > 1:}$

Consider below $J > 1$ and the GLR test in (2.19). Let (2.19) be rewritten as

$$\eta = \sum_{j=1}^{J} \ln(q_j + 1) \underset{H_0}{\overset{H_1}{\gtrless}} \ln \xi, \tag{2.75}$$

where

$$q_j = \frac{\mathbf{z}_j^H \mathbf{S}_j \mathbf{S}_j^H \mathbf{z}_j}{\mathbf{z}_j^H \left(\mathbf{I}_{N_j} - \mathbf{S}_j \mathbf{S}_j^H\right) \mathbf{z}_j + \sum_{l=1}^{L_0} \mathbf{z}_j^H(l) \mathbf{z}_j(l)}. \tag{2.76}$$

Note that the q_j is similar to η_1' in (2.64). Thus under hypothesis H_0, $q_j M_j / K_j$ has the central F-distribution with $2K_j$ and $2M_j$ degrees of freedom; i.e.,

$$f_{q_j}(q_j | H_0) = \frac{q_j^{K-1}(M_j + K_j - 1)!}{(K_j - 1)!(M_j - 1)!(1 + q_j)^{M_j + K_j}}. \tag{2.77}$$

Under hypothesis H_1, $q_j M_j / K_j$ has the noncentral F-distribution with $2K_j$ and $2M_j$ degrees of freedom and noncentrality parameter δ_j; i.e.,

$$f_{q_j}(q_j | H_1) = \exp(-\delta_j) \sum_{i=0}^{\infty} \frac{\delta_j^i q_j^{K_j + i - 1}(M_j + K_j + i - 1)!}{i!(M_j - 1)!(K_j + i - 1)!(1 + q_j)^{M_j + K_j + i}}. \tag{2.78}$$

Note also that q_1, q_2, $\cdots$, q_J are statistically independent of each other. Let

$$\eta_j = \ln(q_j + 1), \quad j = 1, 2, \cdots, J. \tag{2.79}$$

It is easy to show that

$$f_{\eta_j}(\eta_j | H_0) = \frac{[\exp(\eta_j) - 1]^{K_j - 1} (M_j + K_j - 1)!}{(K_j - 1)!(M_j - 1)! \exp[(M_j + K_j - 1)\eta_j]}, \tag{2.80}$$

and

$$f_{\eta_j}(\eta_j | H_1) = \exp(-\delta_j) \sum_{i=0}^{\infty} \frac{\delta_j^i [\exp(\eta_j) - 1]^{K_j + i - 1}(M_j + K_j + i - 1)!}{i!(M_j - 1)!(K_j + i - 1)! \exp[(M_j + K_j + i - 1)\eta_j]}. \tag{2.81}$$

The η_1, η_2, $\cdots$, η_J are also statistically independent of each other. Thus

$$f_\eta(\eta | H_0) = f_{\eta_1}(\eta | H_0) * f_{\eta_2}(\eta | H_0) * \cdots * f_{\eta_J}(\eta | H_0), \tag{2.82}$$

and

$$f_\eta(\eta|H_1) = f_{\eta_1}(\eta|H_1) * f_{\eta_2}(\eta|H_1) * \cdots * f_{\eta_J}(\eta|H_1). \tag{2.83}$$

References

[1] L. M. Novak, M. C. Burl, and W. W. Irving, "Optimal polarimetric processing for enhanced target detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, pp. 234–244, January 1993.

[2] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Signal Processing*, vol. 38, pp. 1760–1770, October 1990.

[3] L. B. Stotts, *Moving-Target Detection Techniques for Optical-Image Sequences.* Ph.D. thesis, University of California, San Diego, California, 1988.

[4] B. R. Hunt and T. M. Cannon, "Nonstationary assumptions for gaussian models of images," *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 876–881, December 1976.

[5] D. R. Sheen, S. C. Wei, T. B. Lewis, and S. R. DeGraaf, "Ultrawide bandwidth polarimetric SAR imagery of foliage-obscured objects," *SPIE Proceedings on OE/LASE,* Vol. 1875, Los Angeles, CA, January 1993.

[6] S. L. Marple, Jr., *Digital Spectral Analysis with Applications.* Englewood Cliffs, NJ 07632: Prentice-Hall, Inc., 1987.

[7] K. Fukunaga, *Introduction to Statistical Pattern Recognition.* New York, NY: Academic Press, Inc., 1990.

[8] D. J. Andersh, M. Hazlett, S. W. Lee, D. D. Reeves, D. P. Sullivan, and Y. Chu, "XPATCH:, a high-frequency electromagnetic scattering prediction code and environment for complex three-dimensional objects," *IEEE Antennas and Propagation Magazine*, vol. 36, pp. 65–69, February 1994.

[9] J. Li and S. M. Rahman, "On image and template false alarm rates when using target templates for target detection," submitted to *IEEE Signal Processing Letters.*

[10] D. K. Barton, *Modern Radar System Analysis.* Norwood, MA: Artech House, Inc., 1988.

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

[12] E. J. Kelly, "Finite-sum expressions for signal detection probabilities," Technical Report 566, Lincoln Laboratory, M.I.T., May 1981.

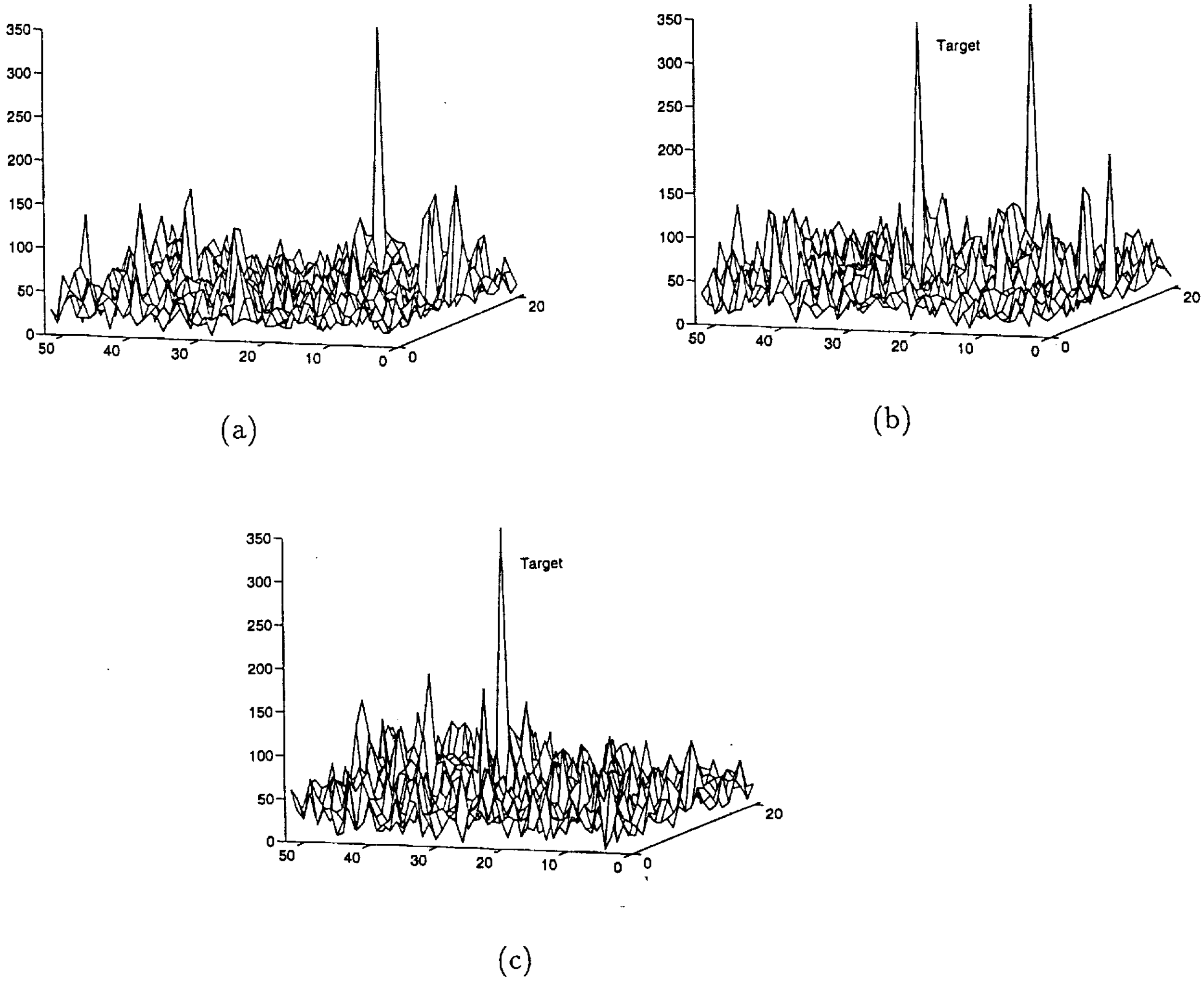
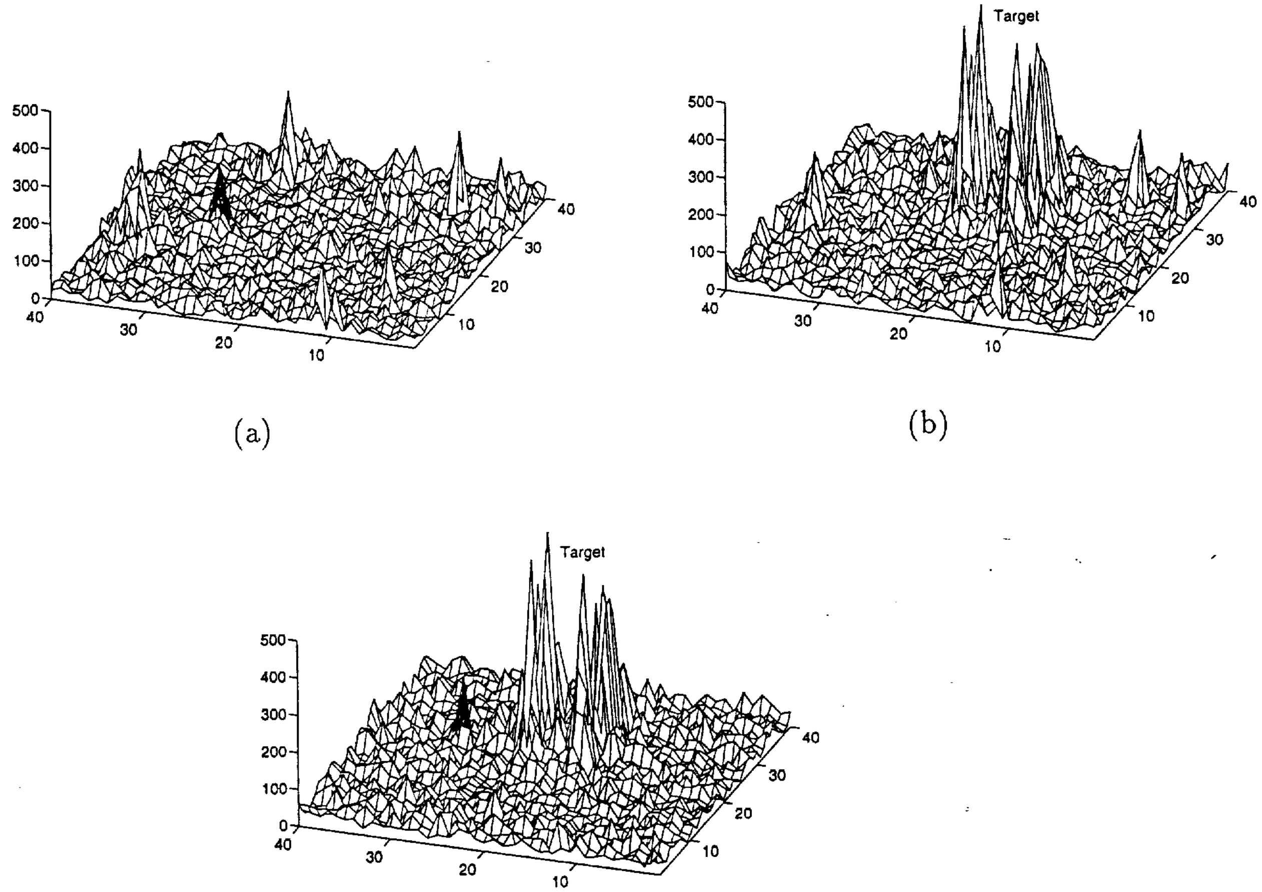


Figure 2.1: 3-dimensional plots of the magnitudes of complex SAR images when the target is oriented 24° from end-on and frequency band between 400 and 1300 MHz. (a) Foliage only. (b) Target in foliage. (c) After coherent subtraction.



(a)

(b)

(c)

Figure 2.2: 3-dimensional plots of the magnitudes of complex SAR images when the target is oriented broadside and frequency band between 400 and 1300 MHz. (a) Foliage only. (b) Target in foliage. (c) After coherent subtraction.

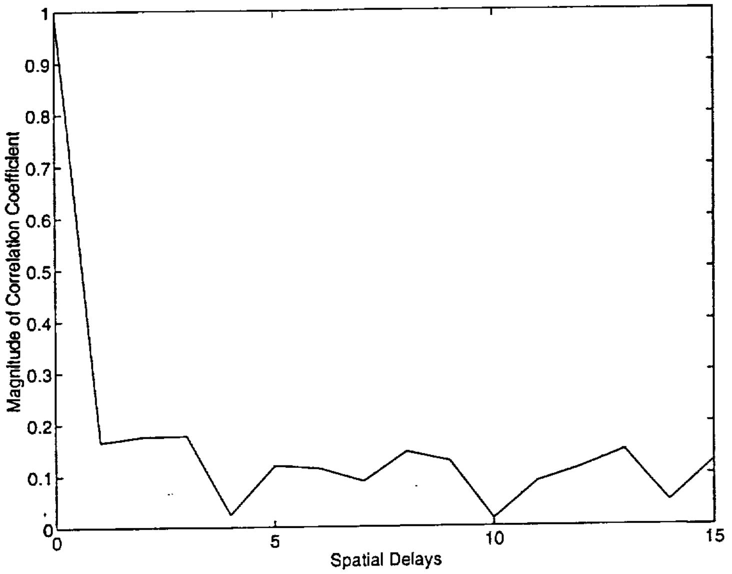


Figure 2.3: Magnitude of clutter and noise correlation coefficient as a function of spatial delays measured by the number of pixels.

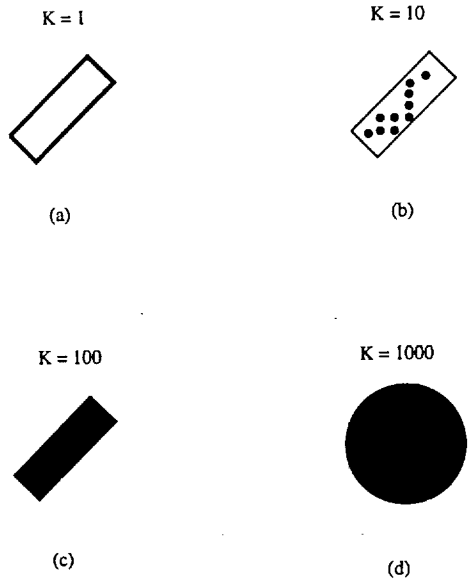


Figure 2.4: Target template examples for (a) complete knowledge of target signature information except for the unknown complex gain and known orientation information, (b) incomplete knowledge of target signature information and known orientation information, (c) unknown target signature information and known orientation information, and (d) unknown target signature information and unknown orientation information.

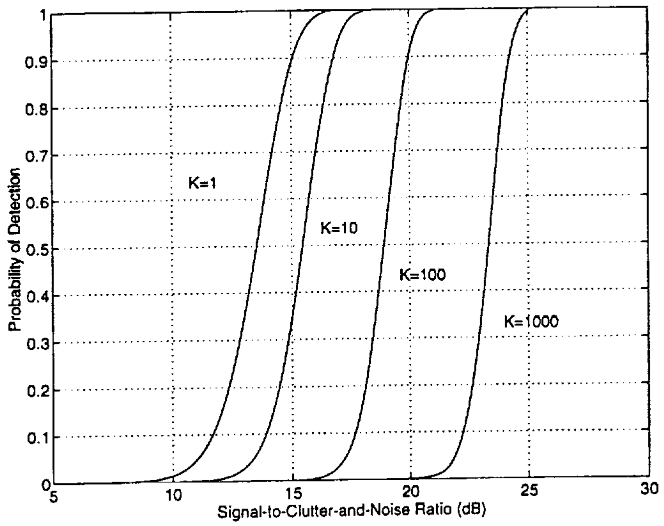


Figure 2.5: Probability of detection of Detector A vs. signal-to-clutter-and-noise ratio for different number of unknown parameters K with $P_F = 10^{-10}$.

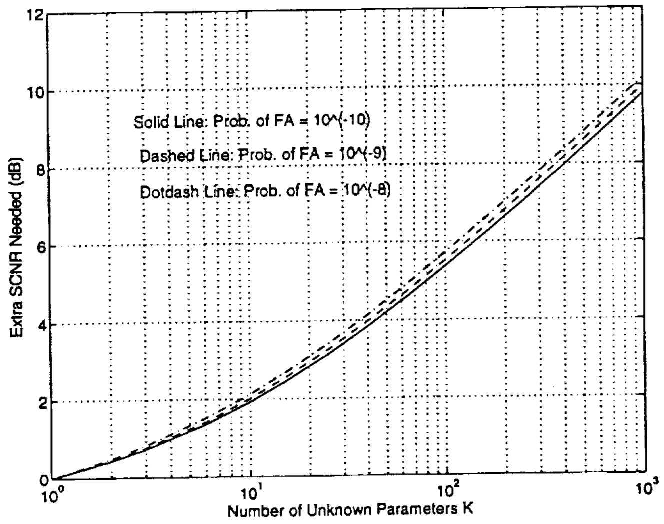


Figure 2.6: As compared with $K = 1$, extra SCNR needed for Detector A to achieve $P_D = 0.5$ vs. K for different probabilities of false alarm P_F.

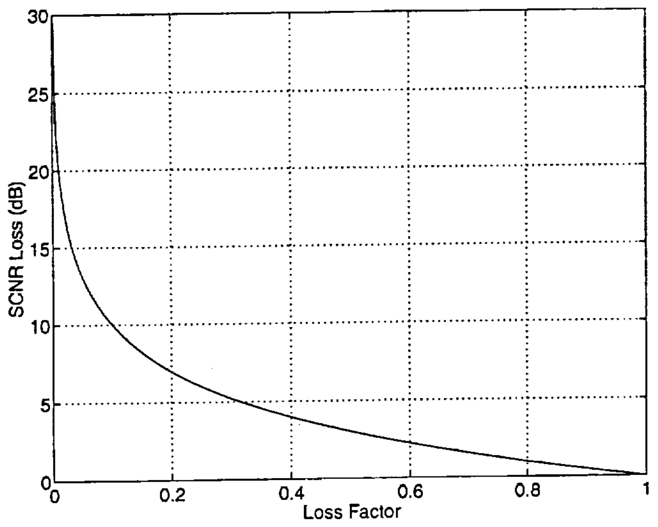


Figure 2.7: Signal-to-clutter-and-noise ratio (SCNR) loss for Detector A as a result of target signature mismatch vs. SCNR loss factor when $J = 1$.

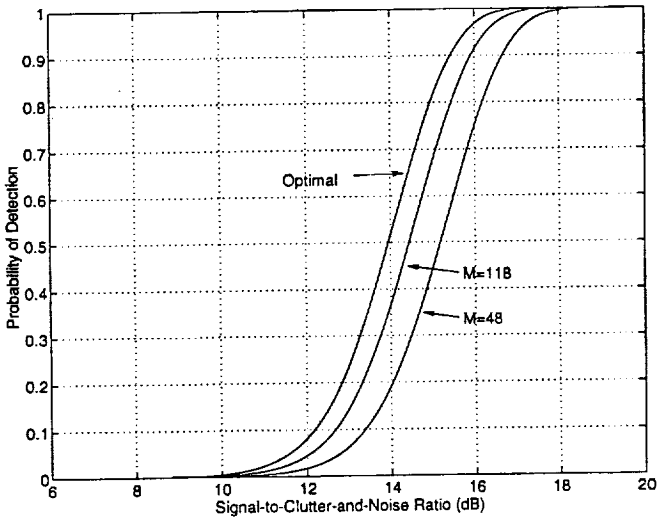


Figure 2.8: Probability of detection of Detector B vs. signal-to-clutter-and-noise ratio for different M when $J = 1$, $K = 2$, and $P_F = 10^{-10}$.

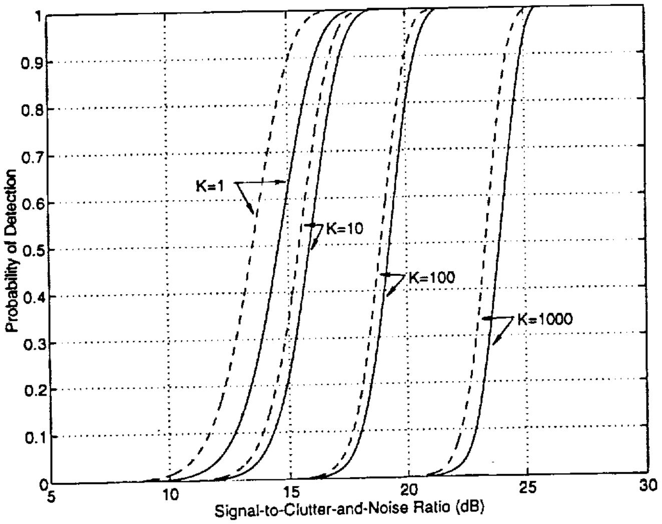


Figure 2.9: Probability of detection vs. signal-to-clutter-and-noise ratio for different K when $J = 1$, $M = \lfloor 48K^{2/3} \rfloor$, and $P_F = 10^{-10}$. The dashed and solid lines are for Detectors A and B, respectively.

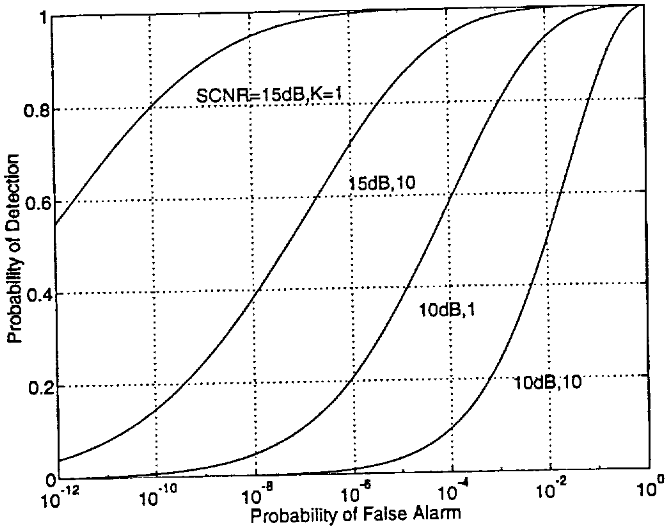


Figure 2.10: Receiver Operating Characteristic of Detector B for different K and different signal-to-clutter-and-noise ratio when $J = 1$, $M = 118$, and $P_F = 10^{-10}$.

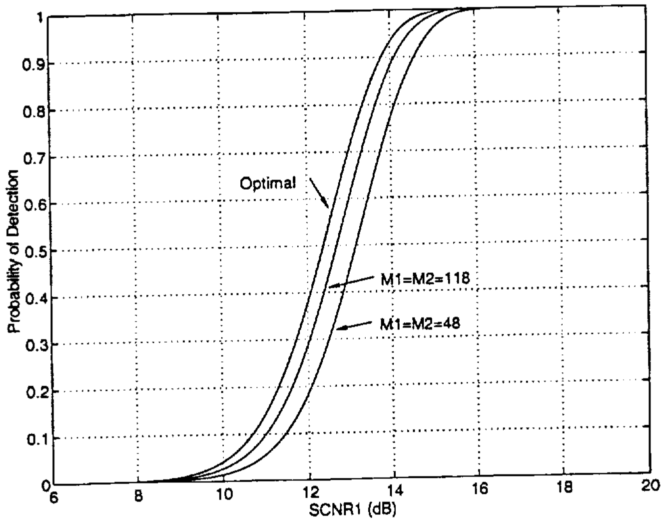
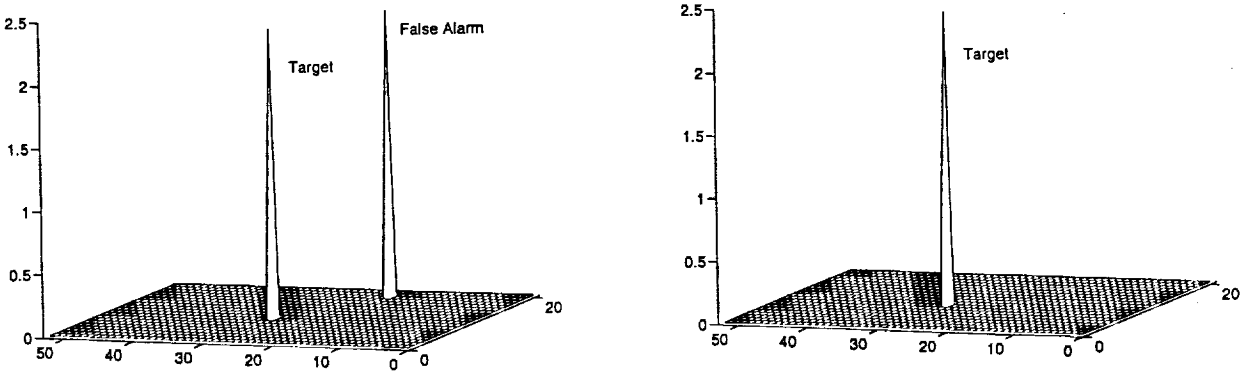


Figure 2.11: Probability of detection of Detector B vs. signal-to-clutter-and-noise ratio for different $M_1 = M_2$ when $J = 2$, $L_1 = L_2 = M_1/2 = M_2/2$, $K_1 = K_2 = 2$ and $P_F = 10^{-10}$.



(a) (b)

Figure 2.12: Target detection (end-on) with $J = N = K = 1$, $M = 48$, and $P_F = 10^{-10}$. (a) Before coherent subtraction. (b) After coherent subtraction.

3. On Image and Template False Alarm Rates When Using Target Templates for Target Detection

3.1 Introduction

Due to the availability of high resolution synthetic aperture radar (SAR) and optical images, ground targets such as trucks and tanks are often described by target templates with each template consisting of one or more pixels in an image. As a result, target templates are used for target detection [1, 2]. It has been shown in [1] that when a target is described by more than one bright pixel, using an appropriate target template for target detection may give better performance than using one pixel at a time.

For a given probability of false alarm of a target template, generalized likelihood ratio detectors may be derived to achieve the constant false alarm rate (CFAR) [1, 2]. Yet achieving CFAR for each target template may not be what a radar designer is asked to achieve. For example, a radar designer may be asked to achieve a constant false alarm rate per kilometer squared. For this case, the size or dimension of an image that covers an area of one kilometer squared is determined by the image resolution. Thus a CFAR detector should be derived to achieve CFAR for the *image* rather than for the *target template.*

However, there does not appear to exist an analytical expression that shows how the false alarm rate of a target template is related to the false alarm rate of an image except for a couple of special cases. One of the special cases is when the size of the image is equal to the size of the target template. For this case, the two false alarm rates are the same. The other special case is when the target template is a single pixel and the pixel false alarm rate is small. For this case, the image false alarm rate is approximately equal to the template or pixel false alarm rate multiplied by the number of pixels in the image.

In this letter, we shall show with computer simulations how the false alarm rate of a target template is related to the false alarm rate of an image when we wish to achieve an image CFAR, which is much less than 1. We shall show that for the case where the radar clutter and noise pixels are independently and identically distributed and the image size is much larger than the target template size, the image false alarm rate is also approximately equal to the template false alarm rate multiplied by the number of pixels in the image. This case is often encountered in practice. For the example where each image covers an area of one kilometer squared, the image size is easily much larger than the target template size for most ground targets such as trucks and tanks.

3.2 Simulations Results

The simulation results below are obtained for the following target detection problem [1]. Let $\mathbf{z}$ denote an $N \times 1$ vector consisting of the pixels of a target template in the presence of clutter and noise. Under hypothesis H_1, the target presence hypothesis, the $\mathbf{z}$ may be written

$$\mathbf{z} = \mathbf{S}\mathbf{b} + \mathbf{n}.$$

The $\mathbf{b}$ is the $K \times 1$ vector consisting of K deterministic unknown complex scalars in the target templates. The $\mathbf{S}$ is a full-rank $N \times K$ matrix describing the locations of the unknown scalars. Only one element in each row and each column of $\mathbf{S}$ is one and the remaining elements are zero. The $\mathbf{n}$ denotes the $N \times 1$ clutter and noise random vector and is assumed zero-mean complex Gaussian with covariance matrix $\sigma^2\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. Under hypothesis H_0, the target absence hypothesis, the $\mathbf{z}$ may be written

$$\mathbf{z} = \mathbf{n}.$$

It has been shown in [1] that the generalized likelihood ratio detector for the above target detection problem has the form

$$\frac{\mathbf{z}^H \mathbf{S}\mathbf{S}^H \mathbf{z}}{\sigma^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma.$$

The threshold parameter γ is determined according to a given probability of false alarm of the target template P_F(Target Template), where

$$P_F(\text{Target Template}) = \sum_{k=0}^{K-1} \frac{\gamma^{K-1-k}}{(K-1-k)!} \exp(-\gamma).$$

To detect the presence of a target from a SAR or optical image, the target template, to which 3.2 is applied, is shifted both horizontally and vertically and one pixel at a time across the entire image. Note that as the target template moves across the entire image, the areas covered by the template may overlap. When more than one detections occur over a group of overlapping areas of the image, all detections are combined together as one detection since targets do not overlap.

In the simulation examples below, the images from which the presence of a target is sought are assumed to be $L \times L$ square images. The target templates are also assumed to be square templates with $N = K$. Both the image and the target template false alarm rates are assumed to be much less than 1. Also, the template false alarm rate is assumed to be so small that the probability of two or more false alarms occuring over non-overlapping areas of an image is negligible as compared to the probability of one false alarm per image. Thus in the simulation examples below, the image false alarm rates are computed as the number of images that contain false alarms divided by the total number of independent images used in the Monte-Carlo simulations.

Figure 3.1 shows the image false alarm rate as a function of $\log_{10}(L)$ when the target is described by a 2×2 template ($N = K = 4$). The threashold γ is determined from 3.2 when the target template false alarm rate[1] satisfies P_F(Target Template) =

---

[1] Note that if the target template were a single pixel, the image false alarm rate would be approximately 0.1 for all image sizes when P_F(Target Template) $= 0.1/L^2$.

$0.1/L^2$. The results shown in Figure 3.1 are obtained with 10,000 Monte-Carlo simulations. Note from Figure 3.1 that the image false alarm rate increases almost linearly with $\log_{10}(L)$ and approaches 0.1 when $L \gg 2$, i.e., when the image size is much larger than the template size.

Figure 3.2 is similar to Figure 3.1 except that P_F(Target Template) is ten times smaller, i.e., P_F(Target Template) $= 0.01/L^2$. The results shown in Figure 3.2 are obtained with 100,000 Monte-Carlo simulations. Note that a similar relation exists between the image false alarm rate and the target template false alarm rate.

Figure 3.3 is also similar to Figure 3.1 except that the target is described by a 3×3 template ($N = K = 9$). Note again that a similar relation exists between the image false alarm rate and the target template false alarm rate.

The implication we obtain from the above three simple simulation examples is that when the image size is much larger than the target template size and a constant image false alarm rate P_F(Image) (P_F(Image) $\ll 1$) is desired, the target template false alarm rate P_F(Target Template) may simply be chosen as P_F(Image) divided by the number of pixels in the image. The so obtained image false alarm rate will be approximately the same as or slightly smaller than the desired image false alarm rate.

Acknowledgments

The first author gratefully acknowledges the helpful discussions with E. G. Zelnio.

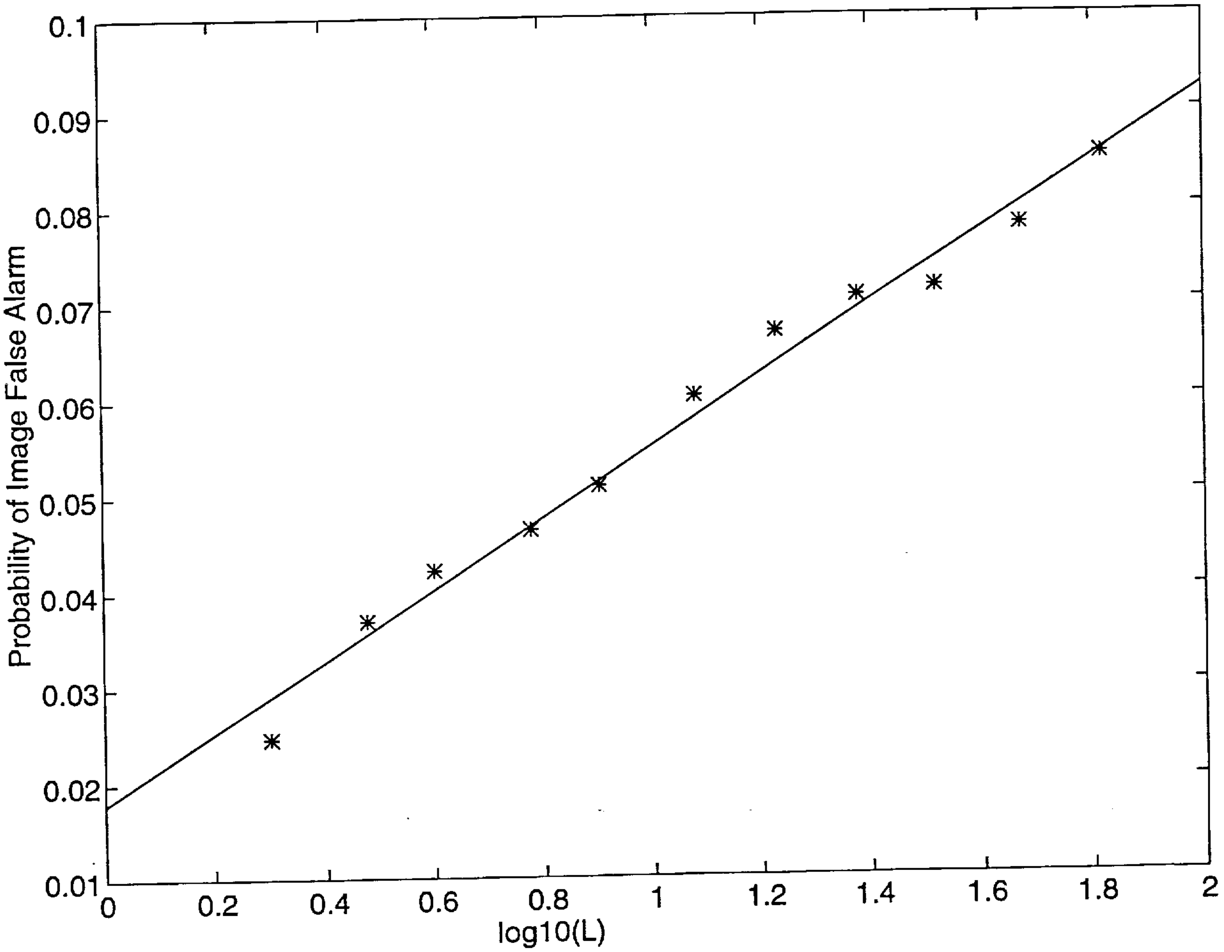


Figure 3.1: Image false alarm rate as a function of $\log_{10}(L)$ when the target template false alarm rate is $0.1/L^2$. The image size is $L \times L$ and the target template size is 2×2. The "*" symbols denote the Monte-Carlo simulation results. The solid line denotes the straight line fitting, which has the form $0.0372 \log_{10}(L) + 0.0179$, of the symbols.

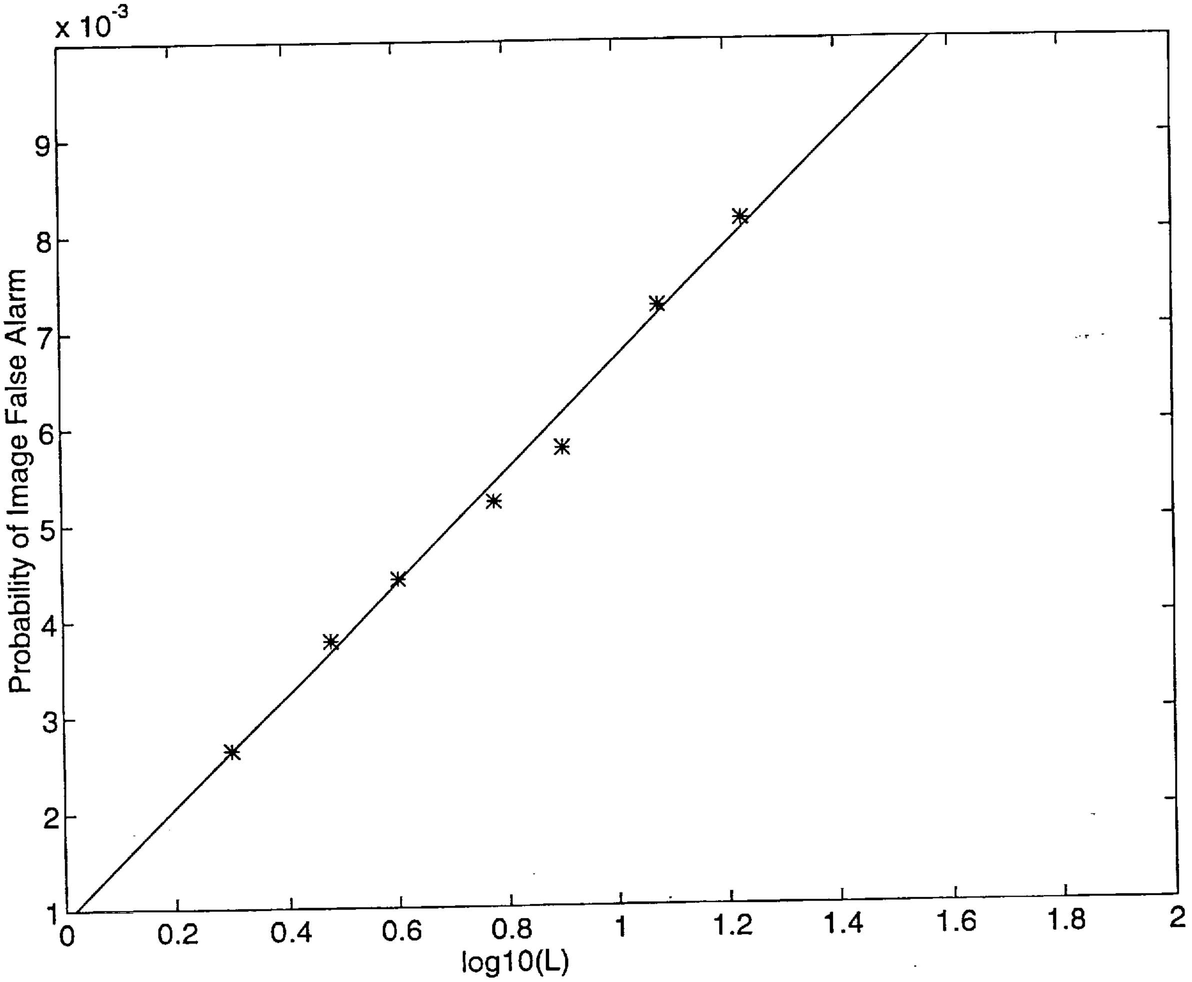


Figure 3.2: Image false alarm rate as a function of $\log_{10}(L)$ when the target template false alarm rate is $0.01/L^2$. The image size is $L \times L$ and the target template size is 2×2. The "*" symbols denote the Monte-Carlo simulation results. The solid line denotes the straight line fitting, which has the form $0.0058\log_{10}(L) + 0.0009$, of the symbols.

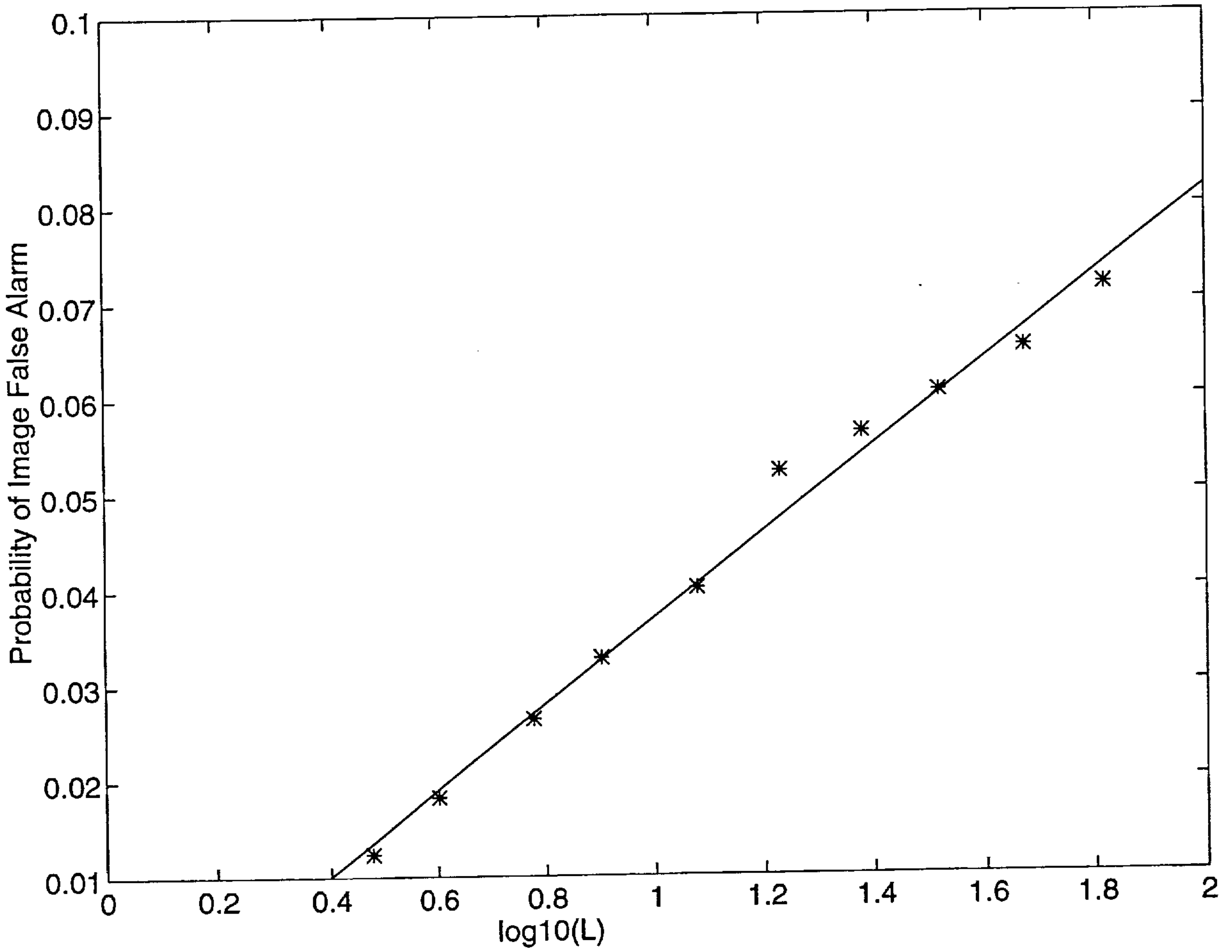


Figure 3.3: Image false alarm rate as a function of $\log_{10}(L)$ when the target template false alarm rate is $0.1/L^2$. The image size is $L \times L$ and the target template size is 3×3. The "*" symbols denote the Monte-Carlo simulation results. The solid line denotes the straight line fitting, which has the form $0.0448 \log_{10}(L) - 0.0078$, of the symbols.

References

[1] J. Li and E. G. Zelnio, "Target detection with synthetic aperture radar and coherent subtraction," *SPIE Proceedings on Optical Engineering in Aerospace Sensing,* Orlando , FL, April 1994 (also submitted to IEEE Transactions on Aerospace and Electronic Systems).

[2] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Signal Processing*, vol. 38, pp. 1760–1770, October 1990.

4. An Efficient Algorithm for Two-Dimensional Frequency Estimation

4.1 Introduction

The two-dimensional (2-D) frequency/angle estimation algorithms, which have been studied recently, include the minimum variance method [1], the linear prediction methods [2, 3, 4, 5], the subspace rotation methods [6, 7, 8, 9, 10, 11, 12, 13, 14], the nonlinear least squares fitting method [15], and the maximum likelihood (ML) methods [16, 17]. These algorithms consider one of the following two data models. The first data model assumes that the frequencies of the 2-D complex sinusoids occur at arbitrary points of the 2-D frequency domain and the data model has the form

$$x_{m,\overline{m}}(t_n) = \sum_{\check{k}=1}^{\check{K}} \alpha_{\check{k}} e^{j(\omega_{\check{k}} m + \mu_{\check{k}} \overline{m})}, \tag{4.1}$$

where $\{\alpha_{\check{k}}\}$ denote the complex amplitudes and $\{\omega_{\check{k}}\}$ and $\{\mu_{\check{k}}\}$ denote the 2-D frequencies. For this data model, the number of unknowns is $4\check{K}$. The second data model assumes that the frequencies of the 2-D complex sinusoids occur at the intersections of some unknown grid lines in the 2-D frequency domain and the data model has the form

$$x_{m,\overline{m}}(t_n) = \sum_{k=1}^{K} \sum_{\overline{k}=1}^{\overline{K}} \alpha_{k,\overline{k}}(t_n) e^{j(\omega_k m + \mu_{\overline{k}} \overline{m})}. \tag{4.2}$$

For this data model, the number of unknowns is $K + \overline{K} + 2K\overline{K}$. Depending on the true distribution of the frequencies of the 2-D complex sinusoids, one model may result in fewer unknowns than the other, which, according to the parsimony principle [18], makes one model yield more accurate frequency estimates than the other. Consider an example where the frequency locations of the complex sinusoids are shown in Figure 4.1. The numbers of unknowns in the first data model for the cases in Figures 4.1(a) - (g) are 12, 16, 20, 24, 28, 32, and 36, respectively. The numbers of unknowns in the second data model for these cases are the same and are 24.

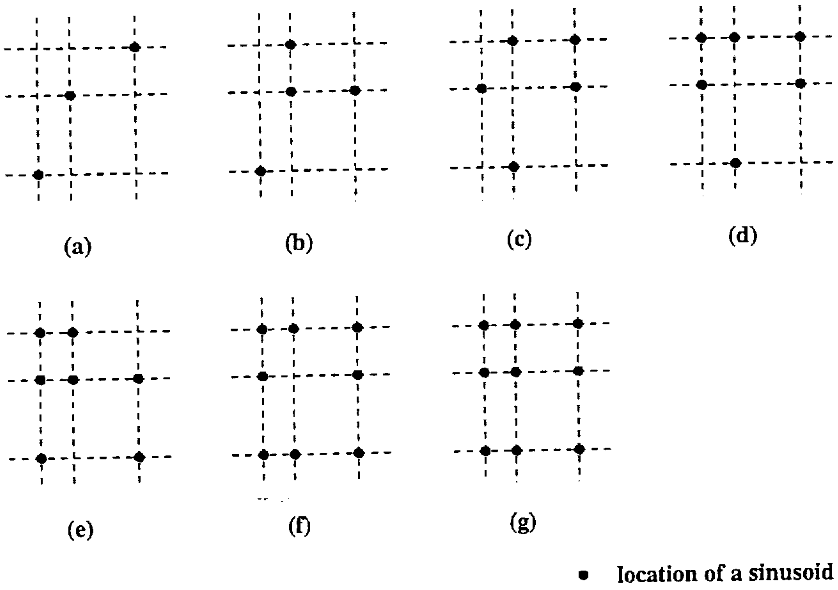


Figure 4.1: Locations of sinusoids in the two-dimensional frequency domain.

This paper considers the second data model, which is also considered in [3, 5, 13, 17]. This data model may be used when a vertical 2-D rectangular uniform linear array is used to estimate the incident angle of a signal arriving from a low angle relative to a smooth reflecting surface [19]. For this case, the signals arriving at the array consist of both the original incident signal and the signal reflected from the smooth surface. The same data model may also be used in synthetic aperture radar (SAR) or inverse SAR (ISAR) imaging [20, 21] to estimate the locations of the scattering centers of an object being imaged [5, 13]. These two cases will be further addressed in the paper.

Among the previously studied methods, the subspace rotation methods, which include the state-space [6, 7, 9], ESPRIT [8, 10], and matrix pencil methods [11, 13, 14], are known to be computationally efficient and have high resolution. Yet their estimation performances are usually not as good as that of the ML methods. The ML methods are asymptotically statistically efficient [17], but they require a computationally intensive multidimensional search over the parameter space.

This paper presents a computationally efficient eigenstructure-based 2D-MODE algorithm for estimating the frequencies of 2-D complex sinusoids. Unlike the subspace rotation methods, the algorithm is statistically efficient under either the assumption that the number of temporal snapshots is large or the signal-to-noise ratio (SNR) is high. The statistical performance of the 2D-MODE algorithm is better than that of the computationally efficient subspace rotation methods, whereas the amount of computations required by the 2D-MODE algorithm is usually no more than a few times of that needed by the subspace rotation methods.

The remaining of this paper is organized as follows. In Section 2, we describe the problem of interest. In Section 3, we present the computationally efficient 2D-MODE estimator. In Section 4, we describe the properties and the applications of the 2D-MODE estimator. In Section 5, we establish the asymptotic statistical performance of the estimator under either the assumption that the number of temporal snapshots is large or the SNR is high. In Section 6, we provide several numerical examples comparing the performance of the 2D-MODE algorithm with that of the subspace rotation algorithms. To make the paper self-contained, the subspace rotation algorithms we use are briefly described in the Appendix. Finally, Section 7 contains our conclusions.

4.2 Problem Formulation

Consider the following model of 2-D complex sinusoids in additive noise:

$$y_{m,\overline{m}}(t_n) = \sum_{k=1}^{K} \sum_{\overline{k}=1}^{\overline{K}} \alpha_{k,\overline{k}}(t_n) e^{j(\omega_k m + \mu_{\overline{k}} \overline{m})} + e_{m,\overline{m}}(t_n), \tag{4.3}$$

where $m = 1, 2, \cdots, M$, $\overline{m} = 1, 2, \cdots, \overline{M}$, and $n = 1, 2, \cdots, N$. We refer to M ($M > K$) and $\overline{M}$ ($\overline{M} > \overline{K}$) as the numbers of *spatial measurements*, and to N as the number of *temporal snapshots*. The additive noise $e_{m,\overline{m}}(t_n)$ is assumed to be a complex Gaussian random process with zero-mean and

$$E\{e_{m,\overline{m}}(t_{n_1})e^*_{m,\overline{m}}(t_{n_2})\} = \sigma^2\delta_{n_1,n_2}, \tag{4.4}$$

where $(\cdot)^*$ denotes the complex conjugate and δ_{n_1,n_2} denotes the Kronecker delta. The $e_{m,\overline{m}}(t_n)$, $m = 1, 2, \cdots, M$, $\overline{m} = 1, 2, \cdots, \overline{M}$, are assumed to be independent of each other and the complex sinusoids. The complex amplitudes $\alpha_{k,\overline{k}}(t_n)$, $k = 1, 2, \cdots, K$, $\overline{k} = 1, 2, \cdots, \overline{K}$, may be modeled as zero-mean complex Gaussian random processes that may or may not be correlated with each other. This model is referred to as the *stochastic* (or unconditional) signal model [22, 23]. For the stochastic signal model, the temporal signal snapshots $\alpha_{k,\overline{k}}(t_n)$ at different sampling times t_n are assumed to be independent of each other. The complex amplitudes $\alpha_{k,\overline{k}}(t_n)$ may also be modeled as deterministic unknowns. This model is referred to as the *deterministic* (or conditional) signal model [22, 23].

Let

$$\mathbf{y}(t_n) = [\ y_{1,1}(t_n) \ \cdots \ y_{1,\overline{M}}(t_n) \ \cdots\cdots \ y_{M,1}(t_n) \ \cdots \ y_{M,\overline{M}}(t_n)\ ]^T, \tag{4.5}$$

$$\mathbf{e}(t_n) = [\ e_{1,1}(t_n) \ \cdots \ e_{1,\overline{M}}(t_n) \ \cdots\cdots \ e_{M,1}(t_n) \ \cdots \ e_{M,\overline{M}}(t_n)\ ]^T, \tag{4.6}$$

$$\mathbf{s}(t_n) = [\ \alpha_{1,1}(t_n) \ \cdots \ \alpha_{1,\overline{K}}(t_n) \ \cdots\cdots \ \alpha_{K,1}(t_n) \ \cdots \ \alpha_{K,\overline{K}}(t_n)\ ]^T, \tag{4.7}$$

$$\mathbf{A} = [\ \mathbf{a}(\omega_1) \ \cdots \ \mathbf{a}(\omega_K)\ ], \quad \mathbf{a}(\omega_k) = [\ e^{j\omega_k} \ \cdots \ e^{jM\omega_k}\ ]^T, \quad k = 1, 2, \cdots, K, \tag{4.8}$$

and

$$\mathbf{B} = [\ \mathbf{b}(\mu_1) \ \cdots \ \mathbf{b}(\mu_{\overline{K}})\ ], \quad \mathbf{b}(\mu_{\overline{k}}) = [\ e^{j\mu_{\overline{k}}} \ \cdots \ e^{j\overline{M}\mu_{\overline{k}}}\ ]^T, \quad \overline{k} = 1, 2, \cdots, \overline{K}, \tag{4.9}$$

where $(\cdot)^T$ denotes the transpose. Then ((4.3)) may be rewritten as

$$\mathbf{y}(t_n) = (\mathbf{A} \otimes \mathbf{B})\mathbf{s}(t_n) + \mathbf{e}(t_n), \tag{4.10}$$

where $\otimes$ denotes the Kronecker product.

The problem of interest herein is to estimate $\omega_1, \omega_2, \cdots, \omega_K$ and $\mu_1, \mu_2, \cdots, \mu_{\overline{K}}$ from $\mathbf{y}(t_n)$, $n = 1, 2, \cdots, N$. Once the frequency estimates are obtained, the complex amplitudes of the 2-D complex sinusoids can also be estimated with a simple least-squares method [3].

4.3 The Computationally Efficient 2D-MODE Estimator

The exact ML estimator of the 2-D frequencies requires a multidimensional search over the parameter space [17]. Since the search over the parameter space is computationally prohibitive, computationally more efficient methods are attractive. We describe below a computationally efficient eigenstructure-based technique, which is an extension of a 1-D algorithm known as MODE [24] or, in a related form, WSF [25]. This estimator is herein referred to as the *2D-MODE*.

Let

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}(t_n)\mathbf{y}^H(t_n), \tag{4.11}$$

where $(\cdot)^H$ denotes the complex conjugate transpose and $\hat{\mathbf{R}}$ is the estimate of the spatial covariance matrix $\mathbf{R}$,

$$\mathbf{R} = E\{\mathbf{y}(t_n)\mathbf{y}^H(t_n)\}. \tag{4.12}$$

It has been shown in [24, 25, 23, 26] that an asymptotically (for large N or high SNR) statistically efficient estimator of the 2-D frequencies $\boldsymbol{\omega} = [\omega_1, \omega_2, \cdots, \omega_K]^T$ and $\boldsymbol{\mu} = [\mu_1, \mu_2, \cdots, \mu_{\overline{K}}]^T$ can be obtained by minimizing the following function:

$$f(\boldsymbol{\omega}, \boldsymbol{\mu}) = \mathrm{tr}\left[\mathbf{P}^{\perp}_{\mathbf{A}\otimes\mathbf{B}}(\boldsymbol{\omega}, \boldsymbol{\mu})\hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H\right], \tag{4.13}$$

where, for some matrix $\mathbf{X}$, the symbol $\mathbf{P}^{\perp}_{\mathbf{X}}$ stands for the orthogonal projector onto the null space of $\mathbf{X}$, and where the columns in $\hat{\mathbf{E}}_s$ are the signal subspace eigenvectors of $\hat{\mathbf{R}}$ that correspond to the $\tilde{K}$ largest eigenvalues of $\hat{\mathbf{R}}$, with $\tilde{K}$ defined as

$$\tilde{K} = \min[N, \text{rank}(\mathbf{S})]. \tag{4.14}$$

Here $\mathbf{S}$ is the signal covariance matrix,

$$\mathbf{S} = E\{\mathrm{s}(t_n)\mathrm{s}^H(t_n)\}. \tag{4.15}$$

We assume that $\tilde{K}$ is known. (If $\tilde{K}$ is unknown, it can be estimated from the data as described, for example, in [27].) Note that if no components of the signal vector $\mathrm{s}(t)$ are fully correlated to one another and if N is large enough, then $\tilde{K} = K\overline{K}$. Further, in ((4.13)) $\hat{\mathbf{\Lambda}}$ is a diagonal matrix with diagonal elements $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_{\tilde{K}}$, which are the $\tilde{K}$ largest eigenvalues of $\hat{\mathbf{R}}$, and

$$\hat{\mathbf{\Lambda}}_s = \hat{\mathbf{\Lambda}} - \hat{\sigma}^2\mathbf{I}, \tag{4.16}$$

with $\mathbf{I}$ denoting the identity matrix and

$$\hat{\sigma}^2 = \frac{1}{M\overline{M} - \tilde{K}} \sum_{i=\tilde{K}+1}^{M\overline{M}} \hat{\lambda}_i = \frac{1}{M\overline{M} - \tilde{K}} \left[\text{tr}\left(\hat{\mathbf{R}}\right) - \sum_{i=1}^{\tilde{K}} \hat{\lambda}_i \right]. \tag{4.17}$$

It is worth noting that the evaluation of $\hat{\mathbf{E}}_s$, $\hat{\mathbf{\Lambda}}$, and $\hat{\mathbf{\Lambda}}_s$ introduced above requires only the computation of the $\tilde{K}$ principal eigenpairs of $\hat{\mathbf{R}}$. Since usually $\tilde{K} \ll M\overline{M}$, the involved computational burden is of the order $O(M^2\overline{M}^2)$ and hence much reduced compared with what would be required for a full eigendecomposition. The orthogonal projector $\mathbf{P}^{\perp}_{\mathbf{A}\otimes\mathbf{B}}$ may be written as (see [28] for the properties of the Kronecker product)

$$\begin{aligned}
\mathbf{P}^{\perp}_{\mathbf{A}\otimes\mathbf{B}} &= \mathbf{I} - (\mathbf{A}\otimes\mathbf{B})\left[(\mathbf{A}^H\otimes\mathbf{B}^H)(\mathbf{A}\otimes\mathbf{B})\right]^{-1}(\mathbf{A}^H\otimes\mathbf{B}^H) \\
&= \mathbf{I} - \left[\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\right]\otimes\left[\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\right] \\
&= \mathbf{I} - (\mathbf{I} - \mathbf{P}^{\perp}_{\mathbf{A}})\otimes(\mathbf{I} - \mathbf{P}^{\perp}_{\mathbf{B}}) \\
&= \mathbf{I}\otimes\mathbf{P}^{\perp}_{\mathbf{B}} + \mathbf{P}^{\perp}_{\mathbf{A}}\otimes\mathbf{I} - \mathbf{P}^{\perp}_{\mathbf{A}}\otimes\mathbf{P}^{\perp}_{\mathbf{B}}.
\end{aligned} \tag{4.18}$$

Thus ((4.13)) may be rewritten as

$$f = \operatorname{tr}\left[\left(\mathbf{I} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right) \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right] + \operatorname{tr}\left[\left(\mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I}\right) \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right] - \operatorname{tr}\left[\left(\mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right) \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right]. \tag{4.19}$$

It is shown in Appendix A that the third term above is a higher-order term for both the case of large N and the case of high SNR. Thus minimizing ((4.19)) is asymptotically equivalent to minimizing

$$f = \operatorname{tr}\left[\left(\mathbf{I} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right) \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right] + \operatorname{tr}\left[\left(\mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I}\right) \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right]. \tag{4.20}$$

We show below how to avoid the search over the parameter space needed to minimize ((4.20)). The projector $\mathbf{P}_{\mathbf{A}}^{\perp}$ above may be reparameterized in terms of the coefficients $\mathbf{c} = [\ c_0\ \ c_1\ \ \cdots\ \ c_K\ ]^T$ of a polynomial defined as

$$\sum_{k=0}^{K} c_k z^{K-k} = c_0 \prod_{k=1}^{K} \left(z - e^{j\omega_k}\right); \quad c_0 \neq 0. \tag{4.21}$$

Let $\mathbf{C}^H$ be the following $(M-K) \times M$ matrix

$$\mathbf{C}^H = \begin{bmatrix} c_K & \cdots & c_1 & c_0 & 0 \\ & \ddots & & \ddots & \ddots \\ 0 & c_K & \cdots & c_1 & c_0 \end{bmatrix}. \tag{4.22}$$

Then $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{P}_{\mathbf{C}}$, where $\mathbf{P}_{\mathbf{C}}$ denotes the orthogonal projector onto the range space of $\mathbf{C}$ (see [24, 29]). By defining $\mathbf{d}$ and $\mathbf{D}$ for $\mathbf{B}$ similarly as $\mathbf{c}$ and $\mathbf{C}$ for $\mathbf{A}$, we obtain $\mathbf{P}_{\mathbf{B}}^{\perp} = \mathbf{P}_{\mathbf{D}}$. Thus by using the reparameterization above, ((4.20)) may be rewritten as

$$f = \operatorname{tr}\left\{\left[\mathbf{I} \otimes \left(\mathbf{D}(\mathbf{D}^H\mathbf{D})^{-1}\mathbf{D}^H\right)\right] \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right\} + \operatorname{tr}\left\{\left[\left(\mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H\right) \otimes \mathbf{I}\right] \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s^2 \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{E}}_s^H\right\}. \tag{4.23}$$

Since $\hat{\mathbf{E}}_s$ is a consistent estimate of the true signal eigenvector matrix $\mathbf{E}_s$ and the columns of $\mathbf{E}_s$ belong to the range space of $\mathbf{A} \otimes \mathbf{B}$, it can be shown [24] that minimizing ((4.23)) is asymptotically equivalent to minimizing

$$f = \operatorname{tr}\left\{\left[\mathbf{I} \otimes \left(\mathbf{D}\mathbf{W}_2^{-1}\mathbf{D}^H\right)\right] \hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H\right\} + \operatorname{tr}\left\{\left[\left(\mathbf{C}\mathbf{W}_1^{-1}\mathbf{C}^H\right) \otimes \mathbf{I}\right] \hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H\right\}, \tag{4.24}$$

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are consistent estimates of $\mathbf{C}^H\mathbf{C}$ and $\mathbf{D}^H\mathbf{D}$, respectively. The 2D-MODE algorithm minimizes ((4.24)), in the following two steps. In the first step, let $\mathbf{W}_1 = \mathbf{I}$ and $\mathbf{W}_2 = \mathbf{I}$ in ((4.24)) and minimize ((4.24)) to obtain consistent estimates $\hat{\mathbf{c}}$ and $\hat{\mathbf{d}}$ of $\mathbf{c}$ and $\mathbf{d}$ [29]. In the second step, let $\mathbf{W}_1 = \hat{\mathbf{C}}^H\hat{\mathbf{C}}$ and $\mathbf{W}_2 = \hat{\mathbf{D}}^H\hat{\mathbf{D}}$, where $\hat{\mathbf{C}}$ and $\hat{\mathbf{D}}$ are formed from the $\hat{\mathbf{c}}$ and $\hat{\mathbf{d}}$ obtained in the first step, and minimize ((4.24)) again. The $\hat{\mathbf{c}}$ and $\hat{\mathbf{d}}$ obtained in the second step are then used to compute the 2-D frequency estimates by rooting the corresponding polynomials of the form of ((4.21)).

Let

$$\hat{\mathbf{Z}} = \hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H = \begin{bmatrix} \hat{\mathbf{Z}}_{11} & \cdots & \hat{\mathbf{Z}}_{1M} \\ \vdots & \cdots & \vdots \\ \hat{\mathbf{Z}}_{M1} & \cdots & \hat{\mathbf{Z}}_{MM} \end{bmatrix}, \tag{4.25}$$

where $\hat{\mathbf{Z}}_{ij}$ are $\overline{M} \times \overline{M}$ matrices. Then

$$\hat{\mathbf{c}} = \min_{\mathbf{c}} \operatorname{tr}\left\{\left(\mathbf{C}\mathbf{W}_1^{-1}\mathbf{C}^H\right)\begin{bmatrix} \operatorname{tr}(\hat{\mathbf{Z}}_{11}) & \cdots & \operatorname{tr}(\hat{\mathbf{Z}}_{1M}) \\ \vdots & \cdots & \vdots \\ \operatorname{tr}(\hat{\mathbf{Z}}_{M1}) & \cdots & \operatorname{tr}(\hat{\mathbf{Z}}_{MM}) \end{bmatrix}\right\} \triangleq \min_{\mathbf{c}} \operatorname{tr}\left\{\left(\mathbf{C}\mathbf{W}_1^{-1}\mathbf{C}^H\right)\hat{\mathbf{Z}}_{\omega}\right\}, \tag{4.26}$$

and

$$\hat{\mathbf{d}} = \min_{\mathbf{d}} \operatorname{tr}\left\{\left(\mathbf{D}\mathbf{W}_2^{-1}\mathbf{D}^H\right)\left(\sum_{m=1}^{M} \hat{\mathbf{Z}}_{mm}\right)\right\} \triangleq \min_{\mathbf{d}} \operatorname{tr}\left\{\left(\mathbf{D}\mathbf{W}_2^{-1}\mathbf{D}^H\right)\hat{\mathbf{Z}}_{\mu}\right\}. \tag{4.27}$$

As shown in Appendix B, as either N or SNR increases, $\hat{\mathbf{Z}}_{\mu}$ approaches a matrix $\mathbf{Z}_{\mu}$ whose rank is no more than $\overline{K}$. However, the rank of $\hat{\mathbf{Z}}_{\mu}$ may be more than $\overline{K}$. To improve the accuracy of the estimates of μ, we may replace $\hat{\mathbf{Z}}_{\mu}$ with its nearest rank-$\overline{K}$ approximation in the Frobenius norm metric [30]. Similarly, $\hat{\mathbf{Z}}_{\omega}$ may be replaced with its nearest rank-K approximation in the Frobenius norm metric since the rank of the limiting matrix $\mathbf{Z}_\omega$ is no more than K (as also shown in Appendix B). For detailed implementation procedures of minimizing the quadratic functions ((4.26)) and ((4.27)), see [29, 31].

The polynomial coefficients in ((4.26)) and ((4.27)) can be constrained by making use of the so-called conjugate symmetry conditions [29, 31]

$$c_k = c^*_{K-k}, \quad k = 0, 1, \cdots, K, \tag{4.28}$$

and

$$d_{\overline{k}} = d^*_{\overline{K}-\overline{k}}, \quad \overline{k} = 0, 1, \cdots, \overline{K}, \tag{4.29}$$

These constraints, considered in detail in the cited works, leads to the most parsimonious parameterization of the estimation problem under discussion, and hence can yield enhanced estimation performance [18]. In the simulation examples given in Section 6, we present results obtained by imposing the conjugate symmetry constraints on the polynomial coefficients in ((4.26)) and ((4.27)).

4.4 Applications and Properties of the 2D-MODE Estimator

We consider below the use and the properties of the 2D-MODE estimator in the case where the number of temporal snapshots N is large and the case where the SNR is high. Since in practical applications, particularly those involving sensor arrays, it is less likely to have large M or $\overline{M}$, we shall not consider such a case herein.

4.4.1 The Case of Large N

The case where the number of temporal snapshots N is "large" may occur when estimating 2-D incident angles with a 2-D $M \times \overline{M}$ rectangular uniform linear array. For this application, the $\{\omega_k\}$ and $\{\mu_{\overline{k}}\}$ are the phase factors from which we can calculate the 2-D incident angles (i.e., the azimuth and elevation angles). Equation ((4.3)) models certain special cases where the incident angles are related to each other. For example, such a special case occurs when a vertical 2-D rectangular uniform linear array is used to estimate the incident angles of a signal arriving from a low angle relative to a smooth reflecting surface such as the calm sea [19]. For this case, the signals arriving at the array consist of both the original incident signal and the signal reflected from the surface, i.e., the specular path, as shown in Figure 4.2. The incident angle and its reflected angle are related to each other and can be shown to fit into the data model in ((4.3)). For this application, N denotes the number of independent snapshots taken at the output of the array. Furthermore,

$$\omega_k = \frac{2\pi\delta_1}{\lambda_0} \sin\theta_k \sin\phi_k, \tag{4.30}$$

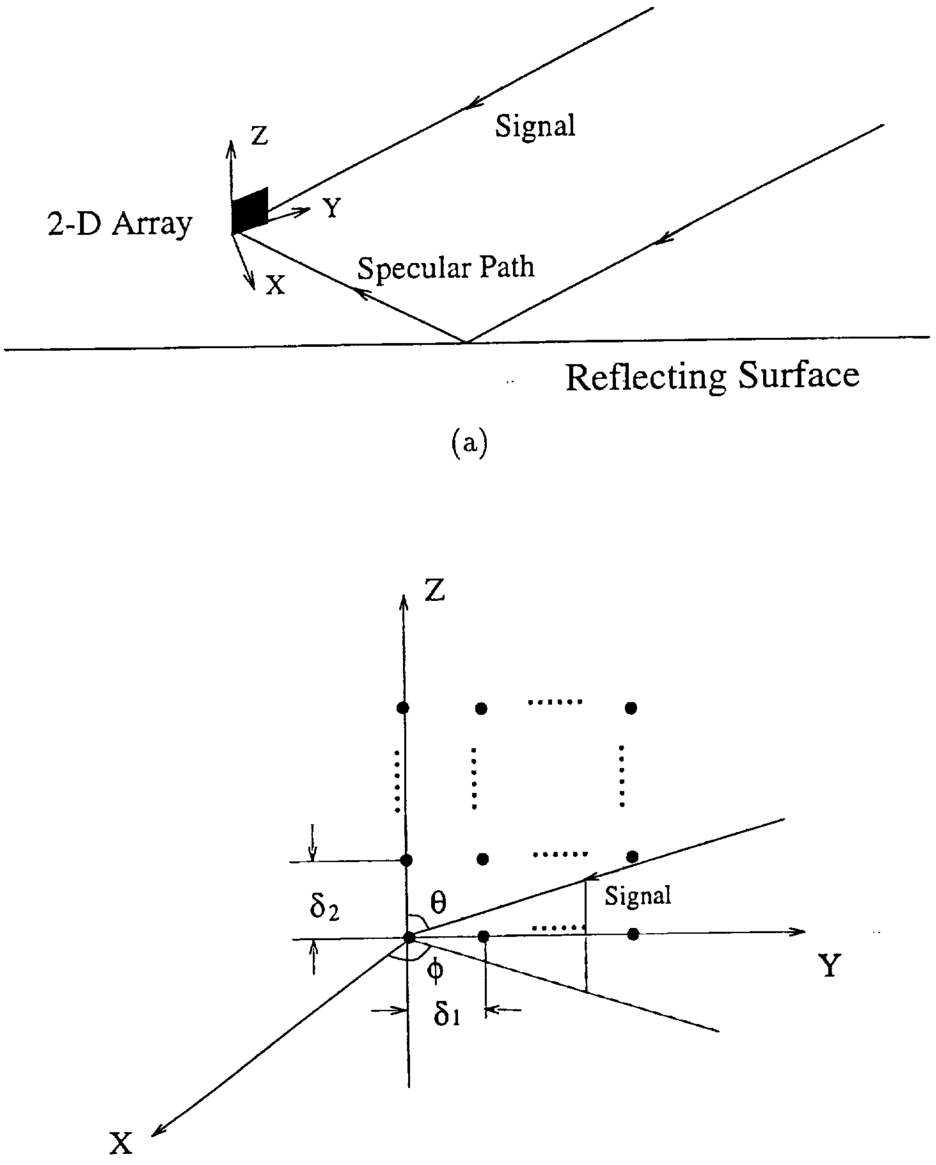
and

$$\mu_{\overline{k}} = \frac{2\pi\delta_2}{\lambda_0} \cos\theta_{\overline{k}}, \tag{4.31}$$

where λ_0 denotes the wavelength of the incident signals and θ_k and ϕ_k denote the elevation and azimuth angles, respectively. Since $\phi_1 = \phi_2$ and $\theta_2 = 180° - \theta_1$, we have $K = 1$, $\overline{K} = 2$, and $\mu_1 = -\mu_2$ for this example.

We can also use the knowledge that $\mu_1 = -\mu_2$ to obtain improved estimates when minimizing ((4.27)). With $\mu_1 = -\mu_2$, we can constrain the polynomial coefficients d_i in ((4.27)) so that $d_0 = d_2 = 1$ and d_1 is real. Note that these conditions are more strict than the conjugate symmetry condition in ((4.29)). Under these conditions, the minimization of ((4.27)) may be carried out as follows. Let $\tilde{\mathbf{Z}}_{\boldsymbol{\mu}}\tilde{\mathbf{Z}}_{\boldsymbol{\mu}}^H$ denote the rank-2 approximation of $\hat{\mathbf{Z}}_{\boldsymbol{\mu}}$, where $\tilde{\mathbf{Z}}_{\boldsymbol{\mu}}$ is an $\overline{M} \times 2$ matrix. Let $\tilde{\mathbf{z}}_1^{\boldsymbol{\mu}}$ and $\tilde{\mathbf{z}}_2^{\boldsymbol{\mu}}$ be the first and second columns of $\tilde{\mathbf{Z}}_{\boldsymbol{\mu}}$, respectively. Then for $\overline{k} = 1, 2$,

$$\mathbf{D}^H \tilde{\mathbf{z}}_{\overline{k}}^{\boldsymbol{\mu}} = \begin{bmatrix} \tilde{z}_{3,\overline{k}}^{\boldsymbol{\mu}} + \tilde{z}_{1,\overline{k}}^{\boldsymbol{\mu}} & \tilde{z}_{2,\overline{k}}^{\boldsymbol{\mu}} \\ \tilde{z}_{4,\overline{k}}^{\boldsymbol{\mu}} + \tilde{z}_{2,\overline{k}}^{\boldsymbol{\mu}} & \tilde{z}_{3,\overline{k}}^{\boldsymbol{\mu}} \\ \vdots & \vdots \\ \tilde{z}_{\overline{M},\overline{k}}^{\boldsymbol{\mu}} + \tilde{z}_{\overline{M}-2,\overline{k}}^{\boldsymbol{\mu}} & \tilde{z}_{\overline{M}-1,\overline{k}}^{\boldsymbol{\mu}} \end{bmatrix} \begin{bmatrix} 1 \\ d_1 \end{bmatrix} \triangleq \hat{\mathbf{H}}_{\overline{k}}^{\boldsymbol{\mu}} \begin{bmatrix} d_0 \\ d_1 \end{bmatrix}, \tag{4.32}$$



(a)

(b)

Figure 4.2: Direction-of-arrival estimation with a 2-D array.

where $\tilde{z}^{\mu}_{i,\bar{k}}$, $i = 1, 2, \cdots, \overline{M}$, denotes the ith element of $\tilde{\mathbf{z}}^{\mu}_{\bar{k}}$. Let

$$\hat{\mathbf{G}}^{\mu} = \begin{bmatrix} \mathrm{Re}\{\mathbf{W}_2^{1/2}\hat{\mathbf{H}}_1^{\mu}\} \\ \mathrm{Re}\{\mathbf{W}_2^{1/2}\hat{\mathbf{H}}_2^{\mu}\} \\ \mathrm{Im}\{\mathbf{W}_2^{1/2}\hat{\mathbf{H}}_1^{\mu}\} \\ \mathrm{Im}\{\mathbf{W}_2^{1/2}\hat{\mathbf{H}}_2^{\mu}\} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{g}}_1^{\mu} & \hat{\mathbf{g}}_2^{\mu} \end{bmatrix}, \tag{4.33}$$

where $\hat{\mathbf{g}}_j^{\mu}$, $j = 1, 2$, denotes the jth column of $\hat{\mathbf{G}}^{\mu}$. Then minimizing ((4.27)) becomes equivalent to minimizing $\left\| \hat{\mathbf{G}}^{\mu} \begin{bmatrix} 1 \\ d_1 \end{bmatrix} \right\|^2$, where $\|\cdot\|$ denotes the Euclidean norm, and the solution is

$$\hat{d}_1 = -\left[(\hat{\mathbf{g}}_2^{\mu})^H \hat{\mathbf{g}}_2^{\mu}\right]^{-1} \left(\hat{\mathbf{g}}_2^{\mu}\right)^H \hat{\mathbf{g}}_1^{\mu}. \tag{4.34}$$

In the case where the number of temporal snapshots N is large, similar to the results obtained in [24, 23] for the 1-D angle/frequency estimation case, the computationally efficient 2D-MODE estimator can be shown to be statistically equivalent to the stochastic maximum likelihood (ML) estimator (i.e., the ML estimator obtained under the stochastic signal model). Since the stochastic ML estimator asymptotically achieves the stochastic Cramer-Rao bound (CRB), it follows that the 2D-MODE estimator is an asymptotically (for large N) statistically efficient estimator.

We remark that the assumption that the temporal signal snapshots $\alpha_{k,\bar{k}}(t_n)$ are independent of each other is not a necessary condition for the application of the 2D-MODE algorithm. In particular, the asymptotic accuracy of the 2D-MODE algorithm is the *same* for both temporally correlated and temporally uncorrelated signal snapshots (see Remark 3 on p. 1787 of [22]). For the temporally correlated signals, however, more accurate estimators may be devised *if* the temporal correlation function is known or *if* it is known to have some parameterized functional form.

4.4.2 The Case of High SNR

The case where the SNR is "high" and N is small may occur in synthetic aperture radar (SAR) or inverse SAR (ISAR) imaging [20, 21, 13, 5]. In this application, the radar usually transmits linear frequency modulated (chirp) pulses. Upon receiving each pulse returned by an object being imaged, the radar mixes the pulse with a reference chirp signal and low-pass filter the mixed signal. As a result, the scattering centers of the object at different ranges correspond to different frequencies of the output of this operation. Since either the radar or the object is moving or rotating, the pulses received at different angles between the radar and the object are used to form a synthetic aperture. The scattering centers of the object at the same range but different cross-ranges correspond to different (Doppler) frequencies over the synthetic aperture. Thus for this application, the $\{\omega_k\}$ and $\{\mu_{\bar{k}}\}$ describe the locations, i.e., the ranges and cross-ranges, of the scattering centers of the object being imaged.

Similar to the results obtained in [26] for the 1-D angle/frequency estimation case, the computationally efficient 2D-MODE estimator can be shown to be asymptotically (for high SNR) statistically equivalent to the deterministic ML estimator. It has been shown in [26] that for high SNR, the deterministic ML estimator achieves the deterministic CRB. Thus the 2D-MODE estimator is an asymptotically (for high SNR) statistically efficient estimator in this case also.

We remark that in the case of high SNR, the N may take any value greater than or equal to 1. The case of $N = 1$ has a particular relevance for the SAR/ISAR application mentioned previously [20, 21]. In such a case, we have $\hat{\mathbf{E}}_s = \mathbf{y}(t_1)/\|\mathbf{y}(t_1)\|$ and $\hat{\mathbf{\Lambda}}_s$ and $\hat{\mathbf{\Lambda}}$ are scalars. Thus for $N = 1$, minimizing ((4.13)) is equivalent to minimizing

$$f = \mathbf{y}^H(t_1)\mathbf{P}^{\perp}_{\mathbf{A}\otimes\mathbf{B}}\mathbf{y}(t_1). \tag{4.35}$$

We note that this function is exactly the one to be minimized by the deterministic ML estimator given in [17]. Since the 2D-MODE algorithm determines an asymptotic realization of the minimizer of ((4.13)), it readily follows that the 2D-MODE and the deterministic ML estimators are asymptotically (for high SNR) equivalent. (We note that showing this equivalence in the case of $N > 1$ is a more complicated operation [26].)

4.4.3 Parameter Identifiability Conditions

We note that our previous results are valid if M and $\overline{M}$ are constrained by the *parameter identifiability* requirements [23, 32, 33]. Let $\boldsymbol{\eta}$ denote a real-valued vector containing all unknowns of the data model in ((4.3)). Then our problem is parameter identifiable if

$$\mathbf{R} = \mathbf{R}_0 \iff \boldsymbol{\eta} = \boldsymbol{\eta}_0. \tag{4.36}$$

As shown in Appendix C, the inequalities

$$M > 2K - \frac{\tilde{K}}{\overline{\overline{M}}}, \tag{4.37}$$

and

$$\overline{M} > 2\overline{K} - \frac{\tilde{K}}{K}, \tag{4.38}$$

are sufficient conditions for parameter identifiability.

By rearranging the elements of $\mathbf{y}(t_n)$, $\mathrm{s}(t_n)$, and $\mathbf{e}(t_n)$, we can rewrite ((4.10)) as $\mathbf{y}(t_n) = (\mathbf{B} \otimes \mathbf{A})\mathrm{s}(t_n) + \mathbf{e}(t_n)$. We can then show similarly that

$$\overline{M} > 2\overline{K} - \frac{\tilde{K}}{M}, \tag{4.39}$$

and

$$M > 2K - \frac{\tilde{K}}{\overline{\overline{K}}}, \tag{4.40}$$

are also sufficient conditions for parameter identifiability.

4.5 Statistical Performance Analysis

In this section, we establish the asymptotic statistical performance of the 2D-MODE estimator for both the case of large N and the case of high SNR.

4.5.1 The Case of Large N

As argued in Appendix D, the asymptotic (for large N) statistical distribution of $\hat{\boldsymbol{\omega}}$ is Gaussian with mean $\boldsymbol{\omega}$ and covariance matrix equal to the corresponding stochastic Cramer-Rao bound (CRB), $\mathbf{CRB}^{s}_{\boldsymbol{\omega}}$. The ijth element of $(\mathbf{CRB}^{s}_{\boldsymbol{\omega}})^{-1}$ is given by

$$\left[(\mathbf{CRB}^{s}_{\boldsymbol{\omega}})^{-1}\right]_{ij} = \frac{2N}{\sigma^2}\mathrm{Re}\left(\mathrm{tr}\left\{\left[\left(\mathbf{A}_j^H\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{A}_i\right)\otimes\left(\mathbf{B}^H\mathbf{B}\right)\right]\mathbf{S}\left(\mathbf{A}^H\otimes\mathbf{B}^H\right)\mathbf{R}^{-1}\left(\mathbf{A}\otimes\mathbf{B}\right)\mathbf{S}\right\}\right), \tag{4.41}$$

where $\mathbf{A}_i = \partial\mathbf{A}/\partial\omega_i$. Similarly, the asymptotic (for large N) statistical distribution of $\hat{\boldsymbol{\mu}}$ is shown to be Gaussian with mean $\boldsymbol{\mu}$ and covariance matrix equal to the corresponding stochastic CRB, $\mathbf{CRB}^{s}_{\boldsymbol{\mu}}$. The ijth element of $(\mathbf{CRB}^{s}_{\boldsymbol{\mu}})^{-1}$ is (see Appendix D)

$$\left[(\mathbf{CRB}^{s}_{\boldsymbol{\mu}})^{-1}\right]_{ij} = \frac{2N}{\sigma^2}\mathrm{Re}\left(\mathrm{tr}\left\{\left[\left(\mathbf{A}^H\mathbf{A}\right)\otimes\left(\mathbf{B}_j^H\mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{B}_i\right)\right]\mathbf{S}\left(\mathbf{A}^H\otimes\mathbf{B}^H\right)\mathbf{R}^{-1}\left(\mathbf{A}\otimes\mathbf{B}\right)\mathbf{S}\right\}\right), \tag{4.42}$$

where $\mathbf{B}_i = \partial\mathbf{B}/\partial\mu_i$.

4.5.2 The Case of High SNR

It is shown in Appendix D that the asymptotic (for high SNR) statistical distribution of $\hat{\boldsymbol{\omega}}$ is Gaussian with mean $\boldsymbol{\omega}$ and covariance matrix equal to the deterministic CRB given by

$$\left[\left(\mathbf{CRB}^{d}_{\boldsymbol{\omega}}\right)^{-1}\right]_{ij} = \frac{2N}{\sigma^2}\mathrm{Re}\left\{\mathrm{tr}\left(\left[\left(\mathbf{A}_{k_2}^H\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{A}_{k_1}\right)\otimes\left(\mathbf{B}^H\mathbf{B}\right)\right]\hat{\mathbf{S}}\right)\right\}, \tag{4.43}$$

where

$$\hat{\mathbf{S}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{s}(t_n)\mathbf{s}^H(t_n). \tag{4.44}$$

It is also shown in Appendix D that the asymptotic (for high SNR) statistical distribution of $\hat{\boldsymbol{\mu}}$ is Gaussian with mean $\boldsymbol{\mu}$ and covariance matrix equal to the corresponding deterministic CRB given by

$$\left[\left(\mathbf{CRB}^{d}_{\boldsymbol{\mu}}\right)^{-1}\right]_{ij} = \frac{2N}{\sigma^2}\mathrm{Re}\left(\mathrm{tr}\left\{\left[\left(\mathbf{A}^H\mathbf{A}\right) \otimes \left(\mathbf{B}_j^H\mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{B}_i\right)\right]\hat{\mathbf{S}}\right\}\right). \tag{4.45}$$

4.6 Numerical Examples

In this section, we illustrate the performance achievable by the 2D-MODE algorithm and compare it with the performance of the two-dimensional subspace rotation methods (2D-SRMs). (The 2D-SRMs, both with and without spatial smoothing, are briefly described in Appendix E.) The empirical performance of the algorithms is obtained from 100 independent trials and it is compared with the theoretical statistical performance given by the CRBs.

In the first three examples, we consider the case of large N. In particular, we consider the scenario shown in Figure 4.2 where a vertical rectangular uniform linear array with $M = 8$ and $\overline{M} = 10$ is used to estimate the angles of arrival of an incident signal arriving from (ϕ, θ) and of its reflected signal arriving from $(\phi, 180° - \theta)$. The signal-to-noise ratio (SNR) of the direct signal used in the examples is assumed to be -7 dB. The SNR of the reflected signal is assumed to be 3dB less than that of the direct signal. The spacings δ_1 and δ_2 between two adjacent sensors in the array are assumed to be a half wavelength. The CRBs of the estimates of ϕ and θ are readily obtained from the CRBs of the estimates of ω_k and $\mu_{\overline{k}}$ given in Section 5 and the Equations ((4.30)) and ((4.31)) relating ϕ and θ to ω_k and $\mu_{\overline{k}}$.

Figure 4.3 shows the root-mean-squared errors (RMSEs) of the angle estimates as a function of the elevation angle θ when $\phi = 45°$ and $N = 60$. The incident and reflected signals are assumed to be uncorrelated with each other, which may happen when the rectangular array is well above the reflecting surface. For this example, the 2D-SRM is used without spatial smoothing. We note that the performance of the 2D-MODE is always better than that of the 2D-SRM. Figure 4.3 also shows the asymptotic (large N) statistical performance, corresponding to the CR-bound (CRB). We note that the 2D-MODE estimator performance is very close to its asymptotic statistical performance, which is also the best achievable performance in the class of (asymptotically) unbiased estimators.

Figure 4.4 shows the performance of the 2-D estimators as a function of the correlation coefficient between the incident and reflected signals when $\phi = 45^\circ$, $\theta = 85^\circ$, and $N = 500$. For this example, the 2D-SRMs are used both with and without spatial smoothing. We note that the 2D-MODE algorithm has the most significant advantage over the 2D-SRM without spatial smoothing when the signals are highly correlated or coherent. When the incident signals are completely correlated with each other or coherent, the 2D-SRM without spatial smoothing fails. For this case, the 2D-SRM must be used with spatial smoothing. For the 2D-SRM with spatial smoothing, we set both dimensions of the subarrays to $L = \overline{L} = 5$. Figure 4.4 shows that the 2D-SRM with spatial smoothing is also outperformed by the 2D-MODE. In our implementations, the amount of computations needed by the 2D-MODE is about 7.1 times of that needed by the 2D-SRM with or without spatial smoothing.

Figure 4.5 shows the RMSEs of the angle estimates as a function of N when $\phi = 45^\circ$, $\theta = 85^\circ$, and the incident and reflected signals are 99% correlated. For the 2D-SRM with spatial smoothing, both dimensions of the subarrays are set to $L = \overline{L} = 5$. We note again that the 2D-MODE algorithm performs better than the 2D-SRM both with and without spatial smoothing. We also note that the larger the N, the closer the 2D-MODE estimator performance to its CRB. Note from Figures 4.3 and 4.5 that for the 2D-MODE estimator to achieve its CRB, a larger N is needed for highly correlated signals than for uncorrelated signals. For $N = 10,000$, the amount of computations needed by the 2D-MODE is about 4.0 times of that needed by the 2D-SRM with or without spatial smoothing.

In the next three examples, we consider the performance of the 2D-MODE frequency estimator for the case of $N = 1$, which is the radar imaging case. We shall consider the case where there are two complex sinusoids in each of the two dimensions, i.e., $K = \overline{K} = 2$, and all 2-D complex sinusoids have the same SNR. We also assume that $M = 8$ and $\overline{M} = 10$. The performance of the 2D-MODE estimator is compared with that of the 2D-SRM with spatial smoothing (with $L = \overline{L} = 5$). Note that as for the case of coherent incident signals and large N, the 2D-SRM must be used with spatial smoothing for this $N = 1$ case, no matter what the SNR is.

Figure 4.6 shows the RMSEs of the frequency estimates as a function of SNR when $(\omega_1, \omega_2) = (2\pi \times 0.28, 2\pi \times 0.33)$ and $(\mu_1, \mu_2) = (2\pi \times 0.30, 2\pi \times 0.35)$. We note that the performance of the 2D-MODE algorithm is always better than that of the 2D-SRM with spatial smoothing, especially when the SNR is around 0 dB. Figure 4.6 also shows the asymptotic (high SNR) statistical performance, corresponding to the CR-bound (CRB). We note that the 2D-MODE estimator performance is very close to its asymptotic statistical performance, which is also the best achievable performance in the class of (asymptotically) unbiased estimators, when the SNR is greater than or equal to 5 dB.

Figure 4.7 shows the RMSEs of the frequency estimates as a function of the frequency separation $\Delta\omega/(2\pi)$ when $(\omega_1, \omega_2) = (2\pi \times 0.28, 2\pi \times 0.28 + \Delta\omega)$, $(\mu_1, \mu_2) = (2\pi \times 0.30, 2\pi \times 0.35)$, and SNR = 10 dB. Figure 4.8 shows the frequency estimates as a function of frequency separation $\Delta\mu/(2\pi)$ when $(\omega_1, \omega_2) = (2\pi \times 0.28, 2\pi \times 0.33)$, $(\mu_1, \mu_2) = (2\pi \times 0.30, 2\pi \times 0.30 + \Delta\mu)$, and SNR = 10 dB. We note that the performance of the 2D-MODE algorithm again is always better than that of the 2D-SRM with spatial smoothing, especially when the frequency separation, either $\Delta\omega/(2\pi)$

or $\Delta\mu/(2\pi)$, is small. We also note that as the frequency separation increases, the 2D-MODE estimator performance becomes closer to its asymptotic statistical performance.

Because the eigendecomposition of $\hat{\mathbf{R}}$ is no longer needed for $N = 1$, the amount of computations needed by the 2D-MODE for $N = 1$ is reduced to 0.35% of the amount required by the example in Figure 4.4. The amount of computations needed by the 2D-SRM with spatial smoothing for $N = 1$ is also reduced to 0.80% of the amount required by the example in Figure 4.4. For $N = 1$, $L = \overline{L} = 5$, $M = 8$, and $\overline{M} = 10$, the amount of computations needed by the 2D-MODE is about 3.1 times of that needed by the 2D-SRM with spatial smoothing.

4.7 Conclusions

We have presented a computationally efficient eigenstructure-based 2D-MODE algorithm for two-dimensional frequency estimation. We have shown that this estimator is asymptotically statistically efficient under the assumption that either the number of temporal snapshots is large or the signal-to-noise ratio is high. Numerical examples showing the comparative performances of this algorithm and of the computationally efficient subspace rotation algorithms have also been given in this paper. We have shown that the performance of the 2D-MODE algorithm is better than that of the subspace rotation methods, whereas the amount of computations required by the 2D-MODE algorithm is usually no more than a few times of that needed by the subspace rotation methods.

Appendix A – The Simplification of the 2D-MODE Cost Function

We show below that the third term of ((4.19)) is a higher-order term and may be neglected for large N or high SNR.

The Case of Large N

Consider first the case of $N \gg 1$. Let f_1, f_2, and f_3 denote the first, second, and third terms, respectively, of the right side of ((4.19)). Let $(\cdot)'$ denote the gradient of $(\cdot)$ with respect to $\boldsymbol{\mu}$. Since $\hat{\boldsymbol{\mu}}$ minimizes ((4.19)), we have for $N \gg 1$.

$$\hat{\boldsymbol{\mu}} - \boldsymbol{\mu} = -(f'')^{-1}f' = -(f_1'' + f_3'')^{-1}(f_1' + f_3'). \tag{4.46}$$

The ith element of f_1' is

$$[f_1]'_i = \mathrm{tr}\left\{\left(\mathbf{I}\otimes\left[\mathbf{P}_{\mathbf{B}}^{\perp}\right]'_i\right)\hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H\right\}, \tag{4.47}$$

where [25]

$$\left[\mathbf{P}_{\mathbf{B}}^{\perp}\right]'_i = -\mathbf{B}^{\dagger H}[\mathbf{B}^H]'_i\mathbf{P}_{\mathbf{B}}^{\perp} - \left(\mathbf{B}^{\dagger H}[\mathbf{B}^H]'_i\mathbf{P}_{\mathbf{B}}^{\perp}\right)^H, \tag{4.48}$$

with

$$\mathbf{B}^{\dagger} = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}. \tag{4.49}$$

Thus $[f_1]'_i$ may be rewritten as

$$[f_1]'_i = -2\ \mathrm{Re}\left(\mathrm{tr}\left\{\left[\mathbf{I}\otimes\left(\mathbf{B}^{\dagger H}[\mathbf{B}^H]'_i\mathbf{P}_{\mathbf{B}}^{\perp}\right)\right]\hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H\right\}\right). \tag{4.50}$$

Since the columns of $\mathbf{E}_s$ span the same signal subspace as the columns of $\mathbf{A}\otimes\mathbf{B}$ and $\mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{B} = \mathbf{0}$, we have $\left(\mathbf{I}\otimes\mathbf{P}_{\mathbf{B}}^{\perp}\right)\mathbf{E}_s = \mathbf{0}$. Thus

$$[f_1]'_i = -2\ \mathrm{Re}\left(\mathrm{tr}\left\{\left[\mathbf{I}\otimes\left(\mathbf{B}^{\dagger H}[\mathbf{B}^H]'_i\mathbf{P}_{\mathbf{B}}^{\perp}\right)\right]\left(\hat{\mathbf{E}}_s - \mathbf{E}_s\right)\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{E}}_s^H\right\}\right). \tag{4.51}$$

Since $\left(\hat{\mathbf{E}}_s - \mathbf{E}_s\right) = O(1/\sqrt{N})$ [34, 35], we have $[f_1]'_i = O(1/\sqrt{N})$. The ith element of f_3' may be written as

$$[f_3]'_i = -2\ \mathrm{Re}\left(\mathrm{tr}\left\{\left[\mathbf{P}_{\mathbf{A}}^{\perp}\otimes\left(\mathbf{B}^{\dagger H}[\mathbf{B}^H]'_i\mathbf{P}_{\mathbf{B}}^{\perp}\right)\right]\left(\hat{\mathbf{E}}_s - \mathbf{E}_s\right)\hat{\mathbf{\Lambda}}_s^2\hat{\mathbf{\Lambda}}^{-1}\left(\hat{\mathbf{E}}_s - \mathbf{E}_s\right)^H\right\}\right), \tag{4.52}$$

where we have used the fact that $\left(\mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I}\right) \mathbf{E}_s = \mathbf{0}$. Thus $[f_3]'_i = O(1/N)$ is negligible for $N \gg 1$ as compared with $[f_1]'_i$. It can be shown similarly that $[f_3]''_{ij}$ is also negligible for $N \gg 1$ as compared with $[f_1]''_{ij}$.

Thus for $N \gg 1$,

$$\hat{\boldsymbol{\mu}} - \boldsymbol{\mu} = -(f''_1)^{-1} f'_1, \tag{4.53}$$

which means that the third term of ((4.19)) is a higher-order term when estimating $\boldsymbol{\mu}$ and may be neglected asymptotically (for $N \gg 1$). We can show similarly that this result is also true when estimating $\boldsymbol{\omega}$.

The Case of High SNR

Consider next the case of high SNR, i.e., $\sigma \ll 1$, where σ is the standard deviation of the additive noise. For $\sigma \ll 1$, Equation ((4.46)) also holds. Since $\left(\mathbf{I} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right) \hat{\mathbf{E}}_s = \left(\mathbf{I} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right) \left(\hat{\mathbf{E}}_s - \mathbf{E}_s\right) = O(\sigma)$ and $\left(\mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I}\right) \hat{\mathbf{E}}_s = \left(\mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I}\right) \left(\hat{\mathbf{E}}_s - \mathbf{E}_s\right) = O(\sigma)$, we obtain $[f_1]'_i = O(\sigma)$ from Equation ((4.51)) and $[f_3]'_i = O(\sigma^2)$ from Equation ((4.52)). Thus $[f_3]'_i$ is negligible for $\sigma \ll 1$ as compared with $[f_1]'_i$. It can be shown similarly that $[f_3]''_{ij}$ is also negligible for $\sigma \ll 1$ as compared with $[f_1]''_{ij}$. Then the third term of ((4.19)) is a higher-order term when estimating $\boldsymbol{\mu}$ and may be neglected asymptotically (for $\sigma \ll 1$). It can be shown similarly that this result is also true when estimating $\boldsymbol{\omega}$.

Appendix B – The Ranks of $\mathbf{Z}_{\boldsymbol{\omega}}$ and $\mathbf{Z}_{\boldsymbol{\mu}}$.

To determine the ranks of $\mathbf{Z}_{\boldsymbol{\omega}}$ and $\mathbf{Z}_{\boldsymbol{\mu}}$, let us consider

$$\mathbf{Z} = \mathbf{E}_s \mathbf{\Lambda}_s^2 \mathbf{\Lambda}^{-1} \mathbf{E}_s^H = \begin{bmatrix} \mathbf{Z}_{11} & \cdots & \mathbf{Z}_{1M} \\ \vdots & \cdots & \vdots \\ \mathbf{Z}_{M1} & \cdots & \mathbf{Z}_{MM} \end{bmatrix}, \tag{4.54}$$

where $\mathbf{Z}_{ij}$ are $\overline{M} \times \overline{M}$ submatrices of $\mathbf{Z}$. Note first that $\mathbf{E}_s$ may be written as

$$\mathbf{E}_s = (\mathbf{A} \otimes \mathbf{B})\mathbf{Q}, \tag{4.55}$$

where $\mathbf{Q}$ is some $(K\overline{K}) \times \tilde{K}$ matrix of full rank $\tilde{K} = \min[N, \mathrm{rank}(\mathbf{S})]$ with probability 1. Thus $\mathbf{Z}$ may be rewritten as

$$\mathbf{Z} = (\mathbf{A} \otimes \mathbf{B})\mathbf{Q}\mathbf{\Lambda}_s^2\mathbf{\Lambda}^{-1}\mathbf{Q}^H(\mathbf{A}^H \otimes \mathbf{B}^H). \tag{4.56}$$

Consider first the rank of $\mathbf{Z}_{\boldsymbol{\mu}}$. Note that

$$(\mathbf{A} \otimes \mathbf{B}) = \begin{bmatrix} \mathbf{B}a_{11} & \cdots & \mathbf{B}a_{1K} \\ \vdots & \cdots & \vdots \\ \mathbf{B}a_{M1} & \cdots & \mathbf{B}a_{MK} \end{bmatrix}, \tag{4.57}$$

where a_{ij} denotes the ijth element of $\mathbf{A}$. Thus $\mathbf{Z}_{ij}$ has the form $\mathbf{B}(\cdots)\mathbf{B}^H$ and

$$\mathbf{Z}_{\boldsymbol{\mu}} = \sum_{m=1}^{M} \mathbf{Z}_{mm} = \mathbf{B}(\cdots)\mathbf{B}^H. \tag{4.58}$$

Then the rank of $\mathbf{Z}_{\boldsymbol{\mu}}$ is no more than the rank of $\mathbf{B}$.

Consider next the rank of $\mathbf{Z}_{\boldsymbol{\omega}}$. Using $(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{A} \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{B})$, we get

$$\mathbf{Z} = (\mathbf{A} \otimes \mathbf{I})\mathbf{\Gamma}(\mathbf{A}^H \otimes \mathbf{I}), \tag{4.59}$$

where

$$\mathbf{\Gamma} = (\mathbf{I} \otimes \mathbf{B})\mathbf{Q}\mathbf{\Lambda}_s^2\mathbf{\Lambda}^{-1}\mathbf{Q}^H(\mathbf{I} \otimes \mathbf{B}^H) = \begin{bmatrix} \mathbf{\Gamma}_{11} & \cdots & \mathbf{\Gamma}_{1K} \\ \vdots & \cdots & \vdots \\ \mathbf{\Gamma}_{K1} & \cdots & \mathbf{\Gamma}_{KK} \end{bmatrix}, \tag{4.60}$$

where $\mathbf{\Gamma}_{ij}$ are $\overline{M} \times \overline{M}$ submatrices of $\mathbf{\Gamma}$. Thus

$$\mathrm{tr}(\mathbf{Z}_{ij}) = \mathrm{tr}\left\{\left(\mathbf{a}_i^{(r)} \otimes \mathbf{I}\right)\mathbf{\Gamma}\left[(\mathbf{a}_j^{(r)})^H \otimes \mathbf{I}\right]\right\} = \mathbf{a}_i^{(r)} \begin{bmatrix} \mathrm{tr}(\mathbf{\Gamma}_{11}) & \cdots & \mathrm{tr}(\mathbf{\Gamma}_{1K}) \\ \vdots & \cdots & \vdots \\ \mathrm{tr}(\mathbf{\Gamma}_{K1}) & \cdots & \mathrm{tr}(\mathbf{\Gamma}_{KK}) \end{bmatrix} (\mathbf{a}_j^{(r)})^H, \tag{4.61}$$

where $\mathbf{a}_i^{(r)}$ denotes the ith row vector of $\mathbf{A}$. Hence $\mathbf{Z}_\omega$ has the form $\mathbf{Z}_\omega = \mathbf{A}(\cdots)\mathbf{A}^H$. Then the rank of $\mathbf{Z}_\omega$ is no more than the rank of $\mathbf{A}$.

Appendix C – A Sufficient Condition for Parameter Identifiability

We prove below the sufficiency of the parameter identifiability requirements ((4.37)) and ((4.38)). Our proof follows closely the one for 1-D angle/frequency estimation in [32, 33].

As in [32, 33], to prove the sufficiency of ((4.37)) and ((4.38)), we need to show that ((4.37)) and ((4.38)) are sufficient conditions for $\mathbf{A} = \mathbf{A}_0$, $\mathbf{B} = \mathbf{B}_0$, and $\mathbf{Q} = \mathbf{Q}_0$ to be the unique solution of

$$(\mathbf{A} \otimes \mathbf{B})\mathbf{Q} = (\mathbf{A}_0 \otimes \mathbf{B}_0)\mathbf{Q}_0, \tag{4.62}$$

where $\mathbf{Q}$ is some $(K\overline{K}) \times \tilde{K}$ matrix of full rank $\tilde{K}$. The Equation ((4.62)) may be written as

$$(\mathbf{A} \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{B})\mathbf{Q} = (\mathbf{A}_0 \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{B}_0)\mathbf{Q}_0. \tag{4.63}$$

Let

$$(\mathbf{I} \otimes \mathbf{B})\mathbf{Q} = \mathbf{\Omega}, \quad \text{and} \quad (\mathbf{I} \otimes \mathbf{B}_0)\mathbf{Q}_0 = \mathbf{\Omega}_0. \tag{4.64}$$

The $\mathbf{\Omega}$ is of dimension $(K\overline{M}) \times \tilde{K}$ and is of full rank $\tilde{K}$ since $\overline{M} > \overline{K}$ and $\mathbf{B}$ is a Vandermonde matrix. Then ((4.63)) may be written as

$$(\mathbf{A} \otimes \mathbf{I})\mathbf{\Omega} = (\mathbf{A}_0 \otimes \mathbf{I})\mathbf{\Omega}_0. \tag{4.65}$$

Consider first a solution $\mathbf{A}$ and $\mathbf{\Omega}$ of ((4.65)), which is such that the columns i_1, i_2, $\cdots$, i_c in $\mathbf{A}$ and $\mathbf{A}_0$ coincide, where $0 \le c < K$. Since $c < K$, the solution $\mathbf{A} = \mathbf{A}_0$ and $\mathbf{\Omega} = \mathbf{\Omega}_0$ is excluded. Let $\tilde{\mathbf{A}}_0$ be the submatrix of $\mathbf{A}_0$ without the columns i_1, i_2, $\cdots$, i_c. Then $(\tilde{\mathbf{A}}_0 \otimes \mathbf{I})$ is of dimension $(M\overline{M}) \times (K\overline{M} - c\overline{M})$. Let $\tilde{\mathbf{\Omega}}_0$ be the submatrix of $\mathbf{\Omega}_0$ without the rows that correspond to the columns of $(\mathbf{A} \otimes \mathbf{I})$ which are missing in $(\tilde{\mathbf{A}}_0 \otimes \mathbf{I})$. Let $\tilde{\mathbf{\Omega}}$ be the same as $\mathbf{\Omega}$ except that the rows of $\mathbf{\Omega}_0$ that are missing in $\tilde{\mathbf{\Omega}}_0$ are subtracted from the corresponding rows of $\mathbf{\Omega}$. Thus ((4.65)) can be written as

$$\begin{bmatrix} (\mathbf{A} \otimes \mathbf{I}) & (\tilde{\mathbf{A}}_0 \otimes \mathbf{I}) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Omega}} \\ -\tilde{\mathbf{\Omega}}_0 \end{bmatrix} = \mathbf{0}. \tag{4.66}$$

Let

$$\gamma_1 = \dim\left\{\mathcal{N}\left(\begin{bmatrix} (\mathbf{A} \otimes \mathbf{I}) & (\tilde{\mathbf{A}}_0 \otimes \mathbf{I}) \end{bmatrix}\right)\right\}, \tag{4.67}$$

where $\dim\{\mathcal{N}(\cdot)\}$ denotes the dimension of the null space of a matrix. Since $\mathbf{A}$ is a Vandermonde matrix, the vectors in $\mathbf{A}$ for any M different values of ω_k are linearly independent. Thus

$$\gamma_1 = 2K\overline{M} - c\overline{M} - \min\left[M\overline{M}, \left(2K\overline{M} - c\overline{M}\right)\right]. \tag{4.68}$$

Let

$$\gamma_2 = \dim\left\{\mathcal{R}\left(\begin{bmatrix} \tilde{\mathbf{\Omega}} \\ -\tilde{\mathbf{\Omega}}_0 \end{bmatrix}\right)\right\}, \tag{4.69}$$

where $\dim\{\mathcal{R}(\cdot)\}$ denotes the dimension of the range space of a matrix. Our goal is to prove that ((4.66)) has no solution under ((4.37)), or equivalently

$$\gamma_1 < \gamma_2. \tag{4.70}$$

If $M\overline{M} \geq \left(2K\overline{M} - c\overline{M}\right)$, then $\gamma_1 = 0$ and ((4.70)) follows since $\gamma_2 \geq 1$. If $M\overline{M} < \left(2K\overline{M} - c\overline{M}\right)$, then $\gamma_1 = 2K\overline{M} - c\overline{M} - M\overline{M}$ and from ((4.64)), we have $\gamma_2 \geq \tilde{K} - c\overline{M}$. From ((4.37)), we have $2K\overline{M} - c\overline{M} - M\overline{M} < \tilde{K} - c\overline{M}$. Thus under ((4.37)), the unique solution to ((4.65)) is $\mathbf{A} = \mathbf{A}_0$ and $\mathbf{\Omega} = \mathbf{\Omega}_0$.

Consider next the solution of

$$(\mathbf{I} \otimes \mathbf{B})\mathbf{Q} = (\mathbf{I} \otimes \mathbf{B}_0)\mathbf{Q}_0. \tag{4.71}$$

We can show similarly that under ((4.38)), the unique solution to ((4.71)) is $\mathbf{B} = \mathbf{B}_0$ and $\mathbf{Q} = \mathbf{Q}_0$.

Thus the proof is finished.

Appendix D – Derivation of the Cramer-Rao Bounds

We present the derivations of the Cramer-Rao Bounds (CRBs) for both stochastic and deterministic signal models.

Stochastic CRB

It follows from [22, 23, 24] that the asymptotic (for large N) statistical distribution of the parameter estimates $[\ \hat{\boldsymbol{\omega}}^T \quad \hat{\boldsymbol{\mu}}^T\ ]^T$ obtained with the 2D-MODE algorithm is Gaussian with mean $[\ \boldsymbol{\omega}^T \quad \boldsymbol{\mu}^T\ ]^T$ and covariance matrix equal to the stochastic CRB, $\mathbf{CRB}^s$. Let

$$\mathcal{A} = \mathbf{A} \otimes \mathbf{B}. \tag{4.72}$$

Using Equation (4.63) in [23], we obtain the ijth element of $(\mathbf{CRB}^s)^{-1}$ as

$$\left[(\mathbf{CRB}^s)^{-1}\right]_{ij} = \frac{2N}{\sigma^2}\mathrm{Re}\left[\mathrm{tr}\left(\mathcal{A}_j^H \mathbf{P}_{\mathcal{A}}^{\perp} \mathcal{A}_i \mathbf{S} \mathcal{A}^H \mathbf{R}^{-1} \mathcal{A} \mathbf{S}\right)\right], \tag{4.73}$$

with $\mathcal{A}_i$ denoting the derivative of $\mathcal{A}$ with respect to the ith element of $[\ \boldsymbol{\omega}^T \quad \boldsymbol{\mu}^T\ ]^T$.

We first show that the CRBs for $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{\mu}}$ are decoupled. Let $\mathcal{A}_k$ denote the derivative of $\mathcal{A}$ with respect to the kth element of $\boldsymbol{\omega}$. Then

$$\mathcal{A}_k = \mathbf{A}_k \otimes \mathbf{B}, \tag{4.74}$$

where $\mathbf{A}_k = \partial \mathbf{A} / \partial \omega_k$. Let $\mathcal{A}_{\bar{k}}$ denote the derivative of $\mathcal{A}$ with respect to the $\bar{k}$th element of $\boldsymbol{\mu}$. Then

$$\mathcal{A}_{\bar{k}} = \mathbf{A} \otimes \mathbf{B}_{\bar{k}}, \tag{4.75}$$

where $\mathbf{B}_{\bar{k}} = \partial \mathbf{B} / \partial \mu_{\bar{k}}$. Using ((4.18)), we obtain

$$\begin{aligned}
\mathcal{A}_k^H \mathbf{P}_{\mathcal{A}}^{\perp} \mathcal{A}_{\bar{k}} &= \left(\mathbf{A}_k^H \otimes \mathbf{B}^H\right)\left(\mathbf{I} \otimes \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I} - \mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right)\left(\mathbf{A} \otimes \mathbf{B}_{\bar{k}}\right) \\
&= \left(\mathbf{A}_k^H \mathbf{A}\right) \otimes \left(\mathbf{B}^H \mathbf{P}_{\mathbf{B}}^{\perp} \mathbf{B}_{\bar{k}}\right) + \left(\mathbf{A}_k^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{A}\right) \otimes \left(\mathbf{B}^H \mathbf{B}_{\bar{k}}\right) - \left(\mathbf{A}_k^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{A}\right) \otimes \left(\mathbf{B}^H \mathbf{P}_{\mathbf{B}}^{\perp} \mathbf{B}_{\bar{k}}\right)
\end{aligned}$$

$$= \mathbf{0}, \tag{4.76}$$

which shows that the $\mathbf{CRB}^s$ is block-diagonal.

Next note that

$$\begin{aligned} \mathcal{A}_{k_2}^H \mathbf{P}_{\mathcal{A}}^{\perp} \mathcal{A}_{k_1} &= \left(\mathbf{A}_{k_2}^H \otimes \mathbf{B}^H\right) \left(\mathbf{I} \otimes \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{I} - \mathbf{P}_{\mathbf{A}}^{\perp} \otimes \mathbf{P}_{\mathbf{B}}^{\perp}\right) \left(\mathbf{A}_{k_1} \otimes \mathbf{B}\right) \\ &= \left(\mathbf{A}_{k_2}^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{A}_{k_1}\right) \otimes \left(\mathbf{B}^H \mathbf{B}\right). \end{aligned} \tag{4.77}$$

Hence the (k_1, k_2)th element of $(\mathbf{CRB}_{\boldsymbol{\omega}}^s)^{-1}$ is

$$\left[(\mathbf{CRB}_{\boldsymbol{\omega}}^s)^{-1}\right]_{k_1 k_2} = \frac{2N}{\sigma^2} \mathrm{Re}\left\{\mathrm{tr}\left(\left[\left(\mathbf{A}_{k_2}^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{A}_{k_1}\right) \otimes \left(\mathbf{B}^H \mathbf{B}\right)\right] \mathbf{S} \mathcal{A}^H \mathbf{R}^{-1} \mathcal{A} \mathbf{S}\right)\right\}. \tag{4.78}$$

Thus using ((4.72)), we complete the proof of ((4.41)). The proof of ((4.42)) is similar.

Deterministic CRB

It follows from [26] that for high SNR, the statistical distribution of the estimate $[\ \hat{\boldsymbol{\omega}}^T \quad \hat{\boldsymbol{\mu}}^T\ ]^T$ obtained with the 2D-MODE algorithm is Gaussian with mean $[\ \boldsymbol{\omega}^T \quad \boldsymbol{\mu}^T\ ]^T$ and covariance matrix equal to the deterministic CRB, $\mathbf{CRB}^d$. Using Equation (4.68) in [23], we obtain the ijth element of $(\mathbf{CRB}^d)^{-1}$ as

$$\left[(\mathbf{CRB}^d)^{-1}\right]_{ij} = \frac{2}{\sigma^2} \left\{\sum_{n=1}^{N} \mathrm{Re}\left[\mathbf{s}^H(t_n) \mathcal{A}_j^H \mathbf{P}_{\mathcal{A}}^{\perp} \mathcal{A}_i \mathbf{s}(t_n)\right]\right\}. \tag{4.79}$$

Using ((4.76)), we can show the $\mathbf{CRB}^d$ is also block-diagonal. Thus the CRBs for $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{\mu}}$ are decoupled for the deterministic signal model as well.

Using ((4.77)), we get the (k_1, k_2)th element of $(\mathbf{CRB}_{\boldsymbol{\omega}}^d)^{-1}$ as

$$\begin{aligned} \left[(\mathbf{CRB}_{\boldsymbol{\omega}}^d)^{-1}\right]_{k_1 k_2} &= \frac{2}{\sigma^2} \left\{\sum_{n=1}^{N} \mathrm{Re}\left[\mathbf{s}^H(t_n) \left(\left(\mathbf{A}_{k_2}^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{A}_{k_1}\right) \otimes \left(\mathbf{B}^H \mathbf{B}\right)\right) \mathbf{s}(t_n)\right]\right\} \\ &= \frac{2N}{\sigma^2} \mathrm{Re}\left\{\mathrm{tr}\left(\left[\left(\mathbf{A}_{k_2}^H \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{A}_{k_1}\right) \otimes \left(\mathbf{B}^H \mathbf{B}\right)\right] \hat{\mathbf{S}}\right)\right\}. \end{aligned} \tag{4.80}$$

Thus ((4.43)) is proven. The proof of ((4.45)) is similar.

Appendix E – The 2D-SRM Algorithms

We briefly describe below the two-dimensional subspace rotation methods (2D-SRMs) we used in the numerical examples. The methods are similar to the matrix pencil method detailed in [13], which was derived for $N = 1$.

The 2D-SRM without Spatial Smoothing

For uncorrelated or slightly correlated signals and large N, the following method, which we refer to as the *2D-SRM without Spatial Smoothing*, may be used to estimate the 2-D frequencies.

Let $\mathbf{Y}(t_n)$ be an $M \times \overline{M}$ matrix whose $(m, \overline{m})$th element is $y_{m,\overline{m}}(t_n)$. To estimate $\boldsymbol{\omega}$, let

$$\tilde{\mathbf{Y}}^{\boldsymbol{\omega}} = \left[\begin{array}{cc} \mathbf{Y}^{\boldsymbol{\omega}} & \mathbf{J}\left(\mathbf{Y}^{\boldsymbol{\omega}}\right)^* \end{array} \right], \tag{4.81}$$

where $\mathbf{J}$ denotes the exchange matrix (with ones on the antidiagonal and zeros elsewhere) and

$$\mathbf{Y}^{\boldsymbol{\omega}} = \left[\begin{array}{cccc} \mathbf{Y}(t_1) & \mathbf{Y}(t_2) & \cdots & \mathbf{Y}(t_N) \end{array} \right]. \tag{4.82}$$

Let the columns of the $M \times K$ matrix $\mathbf{U}_s$ be the left singular vectors of $\tilde{\mathbf{Y}}^{\boldsymbol{\omega}}$ that correspond to the largest singular values of $\tilde{\mathbf{Y}}^{\boldsymbol{\omega}}$. Let $\mathbf{U}_{s1}$ and $\mathbf{U}_{s2}$ denote the $(M - 1) \times K$ submatrices of $\mathbf{U}_s$ consisting of the first and last $(M-1)$ rows of $\mathbf{U}_s$. Then the eigenvalues of $(\mathbf{U}_{s1}^H \mathbf{U}_{s1})^{-1} \mathbf{U}_{s1}^H \mathbf{U}_{s2}$ are the estimates of $e^{j\omega_1}$, $e^{j\omega_2}$, $\cdots$, $e^{j\omega_K}$. From these eigenvalues, we obtain the estimate of $\boldsymbol{\omega}$.

The steps of estimating $\boldsymbol{\mu}$ are similar to those of estimating $\boldsymbol{\omega}$. The main difference is that the $\mathbf{Y}(t_n)$, $n = 1, 2, \cdots, N$, in ((4.82)) above are replaced by $\mathbf{Y}^T(t_n)$. The parameters M and K are also changed to $\overline{M}$ and $\overline{K}$, respectively.

The 2D-SRM with Spatial Smoothing

For highly correlated signals or small N, the following method, which we refer to as the *2D-SRM with Spatial Smoothing*, may be used to estimate the 2-D frequencies.

To estimate $\boldsymbol{\omega}$, let

$$\tilde{\mathbf{Y}}_{ss}^{\boldsymbol{\omega}} = \left[\begin{array}{cc} \mathbf{Y}_{ss}^{\boldsymbol{\omega}} & \mathbf{J}\left(\mathbf{Y}_{ss}^{\boldsymbol{\omega}}\right)^{*} \end{array} \right], \tag{4.83}$$

where

$$\mathbf{Y}_{ss}^{\boldsymbol{\omega}} = \left[\begin{array}{cccc} \mathbf{Y}_{ss}(t_1) & \mathbf{Y}_{ss}(t_2) & \cdots & \mathbf{Y}_{ss}(t_N) \end{array} \right], \tag{4.84}$$

with

$$\mathbf{Y}_{ss}(t_n) = \left[\begin{array}{cccc} \mathbf{Y}_{ss1}(t_n) & \mathbf{Y}_{ss2}(t_n) & \cdots & \mathbf{Y}_{ss(M-L+1)}(t_n) \end{array} \right], \tag{4.85}$$

and $\mathbf{Y}_{ssi}(t_n)$, $i = 1, 2, \cdots, M - L + 1$, denoting the submatrix of $\mathbf{Y}(t_n)$ that consists of the ith to $(i + L - 1)$th rows of $\mathbf{Y}(t_n)$. The remaining steps of the 2D-SRM with Spatial Smoothing are similar to those of the 2D-SRM without Spatial Smoothing.

The steps of estimating $\boldsymbol{\mu}$ are again similar to those of estimating $\boldsymbol{\omega}$. The main difference is that the $\mathbf{Y}_{ssi}(t_n)$, $i = 1, 2, \cdots, M - L + 1$, in ((4.85)) above are obtained as submatrices of $\mathbf{Y}^T(t_n)$. The parameters M, K, and L are also changed to $\overline{M}$, $\overline{K}$, and $\overline{L}$, respectively.

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

[2] L. B. Jackson and H. C. Chien, "Frequency and bearing estimation by two-dimensional linear prediction," *Proceedings of ICASSP 79*, pp. 665–669, April, 1979.

[3] M. M. Barbieri and P. Barone, "A two-dimensional Prony's method for spectral estimation," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2747–2756, November 1992.

[4] J. J. Sacchini, W. M. Steedly, and R. L. Moses, "Two-dimensional Prony modeling and parameter estimation," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3127–3137, November 1993.

[5] I. J. Gupta, "High-resolution radar imaging using 2-D linear prediction," *IEEE Transactions on Antennas and Propagation*, vol. 42, pp. 31–37, January 1994.

[6] S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, "State-space and singular-value decomposition-based approximation methods for harmonic retrieval problem," *Journal of the Optical Society of America*, vol. 73, pp. 1799–1811, December 1983.

[7] D. V. Bhaskar Rao and S. Y. Kung, "A state space approach for the 2-d harmonic retrieval problem," *Proceedings of ICASSP 84*, pp. 4.10.1–4.10.4, 1984.

[8] M. D. Zoltowski and D. Stavrinides, "Sensor array signal processing via a procrustes rotations based eigenanalysis of the ESPRIT data pencil," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 832–861, June 1989.

[9] F. Li and R. J. Vaccaro, "On frequency-wavenumber estimation by state-space realization," *IEEE Transactions on Circuits and Systems*, vol. 38, pp. 800–803, July 1991.

[10] J. Li and R. T. Compton, Jr., "Two-dimensional angle and polarization estimation using the ESPRIT algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 40, pp. 550–555, May 1992.

[11] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2267–2280, September 1992.

[12] A. L. Swindlehurst and T. Kailath, "Algorithms for azimuth/elevation direction finding using regular array geometries," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, pp. 145–156, January 1993.

[13] Y. Hua, "High resolution imaging of continuously moving object using stepped frequency radar," *Signal Processing*, vol. 35, pp. 33–40, January 1994.

[14] Y. Hua, "A pencil-MUSIC algorithm for finding two-dimentional angles and polarizations using crossed dipoles," *IEEE Transactions on Antennas and Propagation*, vol. 41, pp. 370–376, March 1993.

[15] S. R. DeGraaf, "Parametric estimation of complex 2-d sinusoids," *IEEE Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, pp. 391–396, August, 1988.

[16] M. P. Clark and L. L. Scharf, "Two-dimensional modal analysis based on maximum likelihood," *IEEE Transactions on Signal Processing*, vol. 42, pp. 1443–1452, June 1994.

[17] C. R. Rao, L. C. Zhao, and B. Zhou, "Maximum likelihood estimation of two-dimensional superimposed exponential signals," to appear in *IEEE Transactions on Signal Processing.*

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

[19] M. D. Zoltowski and T. Lee, "Maximum likelihood based sensor array signal processing in the beamspace domain for low angle radar tracking," *IEEE Transactions on Signal Processing*, vol. 39, pp. 656–671, March 1991.

[20] D. C. Munson, Jr., J. D. O'Brien, and W. K. Jenkins, "A tomographic formulation of spotlight-mode synthetic aperture radar," *Proceedings of IEEE*, vol. 71, pp. 917–925, August 1983.

[21] D. A. Ausherman, A. Kozma, J. L. Walker, H. M. Jones, and E. C. Poggio, "Developments in radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, pp. 363–400, July 1984.

[22] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, pp. 1783–1795, October 1990.

[23] B. Ottersten. M. Viberg, P. Stoica, and A. Nehorai, "Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing." In *Radar Array Processing,* S. Haykin *et. al.* (eds.), Chapter 4. New York, NY: Springer-Verlag Inc., 1993.

[24] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, pp. 1132–1143, July 1990.

[25] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-39, pp. 1110–1121, May 1991.

[26] M. Viberg, "Sensitivity of parametric direction finding to colored noise fields and undermodeling," *Signal Processing*, vol. 34, pp. 207–222, 1993.

[27] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-33, pp. 387–392, April 1985.

[28] A. Graham, *Kronecker Products and Matrix Calculus with Applications.* Chichester, UK: Ellis Horwood Litd., 1981.

[29] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *IEE Proceedings, Pt. F*, vol. 137, pp. 19–26, February 1990.

[30] G. W. Stewart, *Introduction to Matrix Computations.* New York, NY: Academic Press, Inc., 1973.

[31] J. Li and P. Stoica, "Efficient parameter estimation of partially polarized electromagnetic waves," to appear in *IEEE Transactions on Signal Processing.*

[32] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 996–1000, July 1989.

[33] A. Nehorai, D. Starer, and P. Stoica, "Direction-of-arrival estimation in applications with multipath and few snapshots," *Circuits, Systems, and Signal Processing*, vol. 10, pp. 327–342, 1991.

[34] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, pp. 331–341, April 1986.

[35] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 720–741, May 1989.

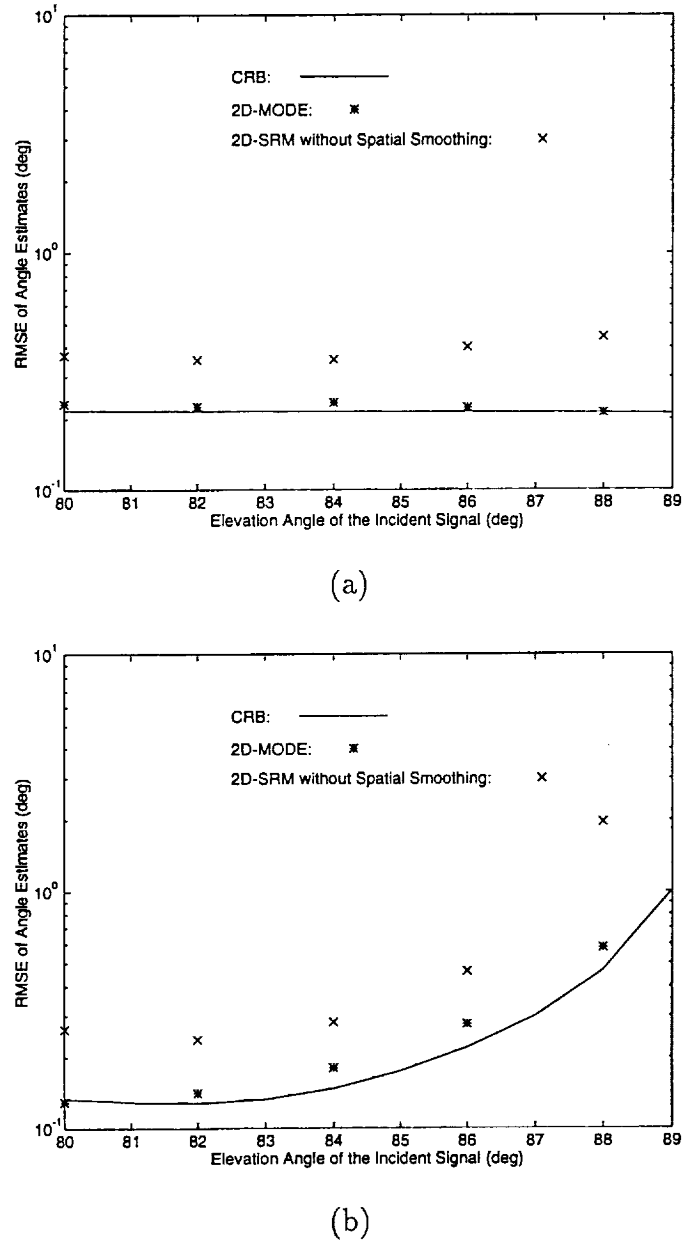
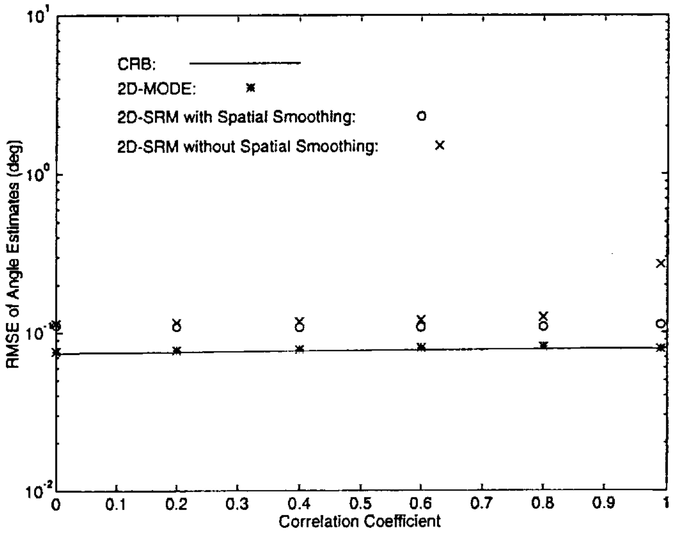
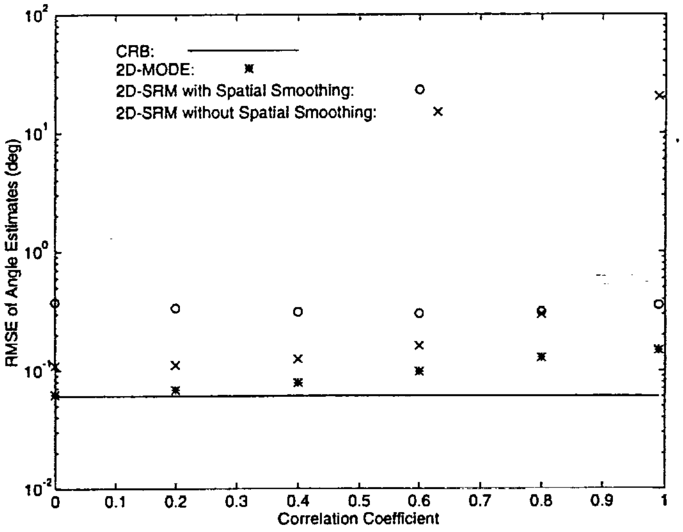


Figure 4.3: Root-mean-squared errors of angle estimates as a function of the elevation angle θ when the direct and reflected signals arrive from $(45°, \theta)$ and $(45°, 180° - \theta)$, respectively. The SNRs for the direct and incident signals are -7 dB and -10 dB, respectively. The signals are assumed to be uncorrelated with each other. Further, $M = 8$, $\overline{M} = 10$, and $N = 60$. (a) For the estimates of the azimuth angle ϕ. (b) For the estimates of the elevation angle θ.

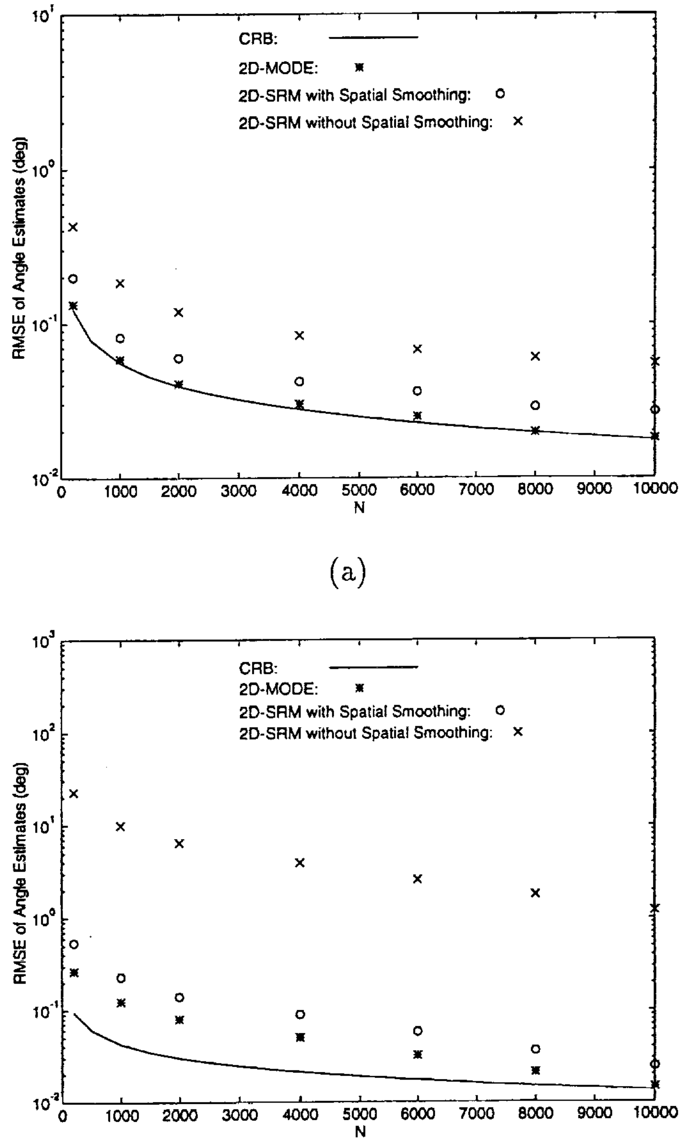


(a)



(b)

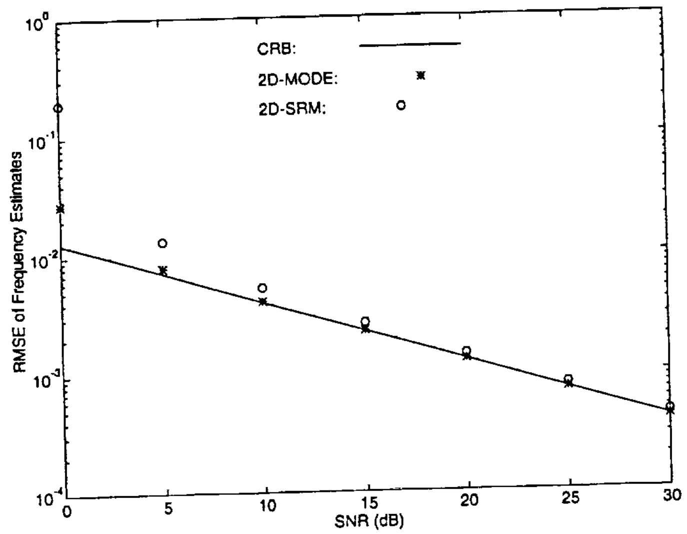
Figure 4.4: Root-mean-squared errors of angle estimates as a function of the correlation coefficient between the direct and reflected signals when the direct and reflected signals arrive from (45°, 85°) and (45°, 95°), respectively. The SNRs for the direct and incident signals are −7 dB and −10 dB, respectively. Further, $M = 8$, $\overline{M} = 10$, $L = \overline{L} = 5$ (for 2D-SRM with spatial smoothing), and $N = 500$. (a) For the estimates of the azimuth angle ϕ. (b) For the estimates of the elevation angle θ.



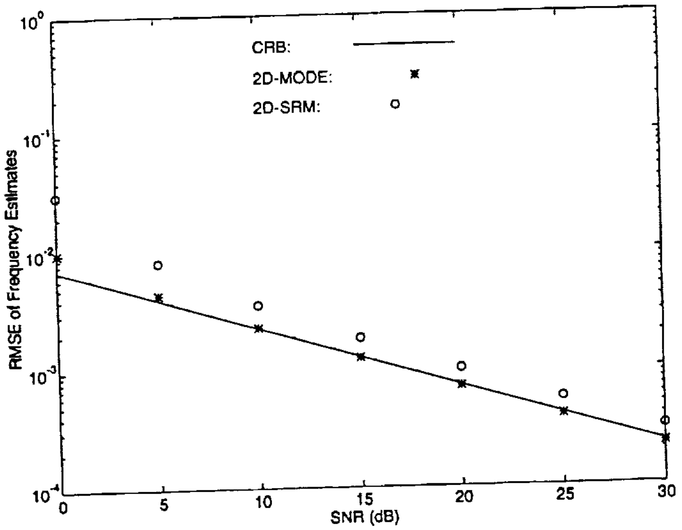
(a)

(b)

Figure 4.5: Root-mean-squared errors of angle estimates as a function of the number of temporal snapshots N when the direct and reflected signals arrive from $(45°, 85°)$ and $(45°, 95°)$, respectively. The SNRs for the direct and incident signals are -7 dB and -10 dB, respectively. The correlation coefficient between the direct and reflected signals is 0.99. Further, $M = 8$, $\overline{M} = 10$, and $L = \overline{L} = 5$ (for 2D-SRM with spatial smoothing). (a) For the estimates of the azimuth angle ϕ. (b) For the estimates of the elevation angle θ.

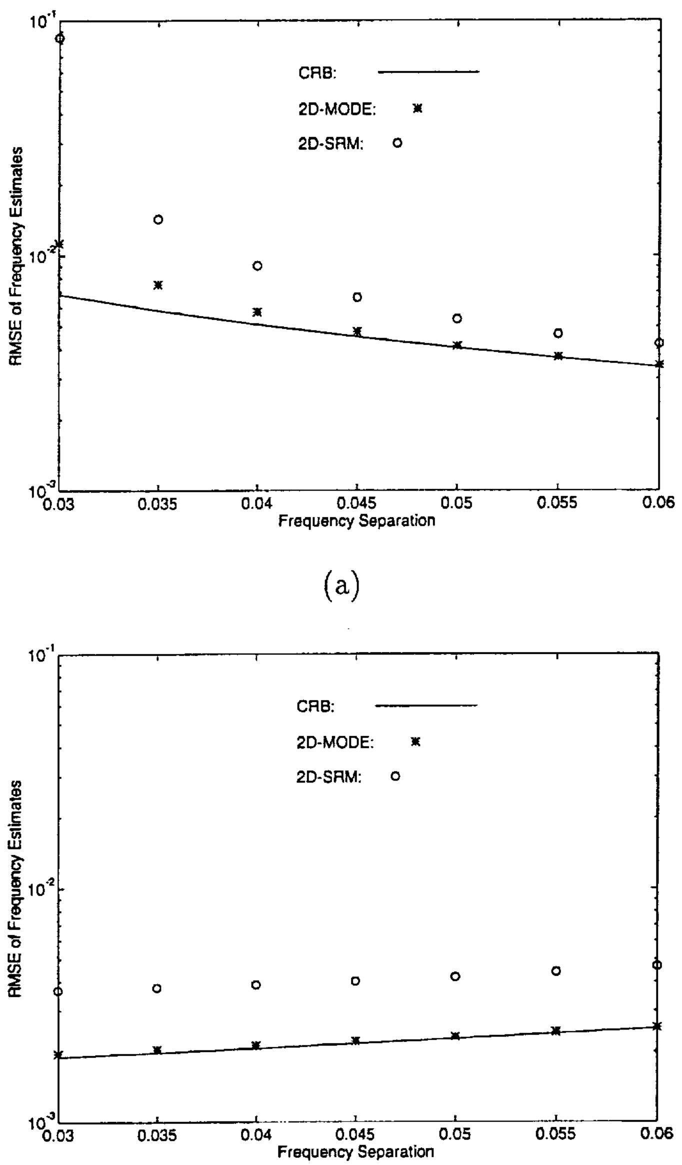


(a)



(b)

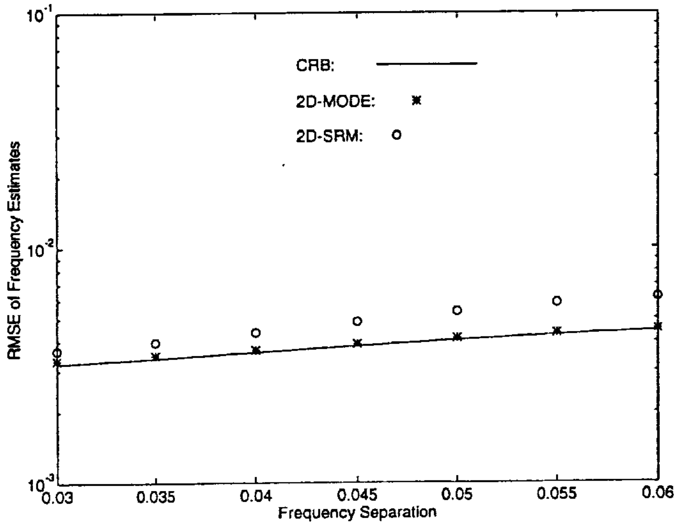
Figure 4.6: Root-mean-squared errors of frequency estimates as a function of SNR when $(\omega_1, \omega_2) = (2\pi \times 0.28, 2\pi \times 0.33)$, $(\mu_1, \mu_2) = (2\pi \times 0.30, 2\pi \times 0.35)$, $M = 8$, $\overline{M} = 10$, $L = \overline{L} = 5$ (for 2D-SRM with spatial smoothing), and $N = 1$. (a) For the estimates of $\omega_1/(2\pi)$. (b) For the estimates of $\mu_1/(2\pi)$.



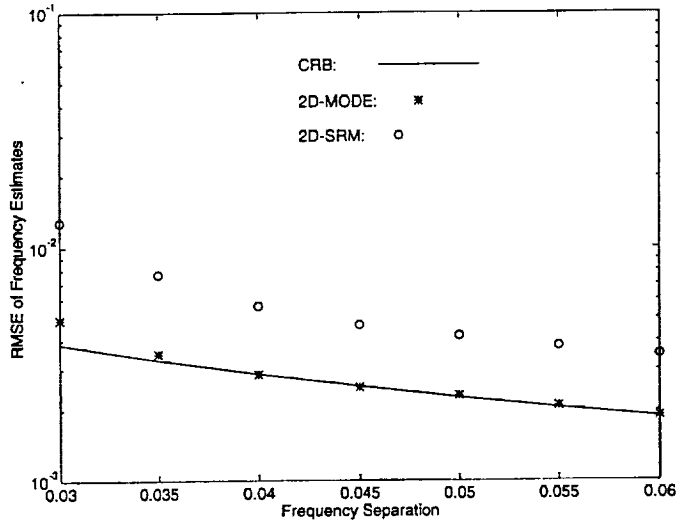
(a)

(b)

Figure 4.7: Root-mean-squared errors of frequency estimates as a function of frequency separation $\Delta\omega/2\pi$ when $(\omega_1, \omega_2) = (2\pi \times 0.28, 2\pi \times 0.28 + \Delta\omega)$, $(\mu_1, \mu_2) = (2\pi \times 0.30, 2\pi \times 0.35)$, SNR $= 10$ dB, $M = 8$, $\overline{M} = 10$, $L = \overline{L} = 5$ (for 2D-SRM with spatial smoothing), and $N = 1$. (a) For the estimates of $\omega_1/(2\pi)$. (b) For the estimates of $\mu_1/(2\pi)$.



(a)



(b)

Figure 4.8: Root-mean-squared errors of frequency estimates as a function of frequency separation $\Delta\mu/2\pi$ when $(\omega_1, \omega_2) = (2\pi \times 0.28, 2\pi \times 0.33)$, $(\mu_1, \mu_2) = (2\pi \times 0.30, 2\pi \times 0.30 + \Delta\mu)$, SNR = 10 dB, $M = 8$, $\overline{M} = 10$, $L = \overline{L} = 5$ (for 2D-SRM with spatial smoothing), and $N = 1$. (a) For the estimates of $\omega_1/(2\pi)$. (b) For the estimates of $\mu_1/(2\pi)$.

5. High Resolution Range Signature Estimation and Synthetic Aperture Radar Imaging

5.1 Introduction

This paper proposes a robust parametric data model for estimating high resolution range signatures of radar targets and for forming high resolution synthetic aperture radar (SAR) images. For the range signature estimation, estimating the radar cross section (RCS) of the scattering center of a radar target at a certain range is modeled as estimating the amplitude and phase of a complex sinusoid with known frequency in unknown colored Gaussian noise. For the SAR imaging application, estimating the complex intensity of a pixel in an SAR image is modeled as estimating the amplitude and phase of a two-dimensional complex sinusoid with known frequency in unknown colored Gaussian noise. This new modeling approach models the thermal noise and the interferences from other scattering centers of a radar target or other pixels in a SAR image as unknown colored noise.

This paper also presents a parameter estimation algorithm for the data model. The algorithm is referred to as the APES (Amplitude and Phase Estimation of a Sinusoid in unknown colored noise) algorithm. We shall describe how the APES algorithm can be used to estimate the range signatures and to form SAR images. The APES algorithm avoids the search over the parameter space and requires only simple matrix multiplications and matrix inverses.

We will show that our modeling and estimation approach yields better resolution and lower sidelobes than the conventional nonparametric FFT (fast Fourier transform) method. We will also show that our approach is more robust than modeling the radar data as a certain number of complex sinusoids in noise and estimating the frequencies, amplitudes, and phases of the sinusoids with one of the best sinusoidal parameter estimation methods. We will present both numerical and experimental examples comparing the performance of our approach with the two afore-mentioned approaches.

In Section 2, we consider high resolution range signature estimation. We shall formulate the problem of interest, present the APES algorithm, and demonstrate the performance of our approach with both simulated and experimental data. (In Appendix A, we derive a computationally and asymptotically statistically efficient large sample maximum likelihood estimator for estimating the complex gain of a signal with known waveform and known steering vector in unknown colored Gaussian noise, which the APES estimator mimics.) In Section 3, we consider high resolution SAR imaging with the APES method. We shall formulate the problem of interest, describe how to use the APES method for SAR imaging, and demonstrate the performance of our approach with XPATCH (a computational electromagnetics software) data. Finally, Section 4 contains our conclusions.

5.2 High Resolution Range Signature Estimation

In this section, we propose a parametric data model and present an estimation algorithm for estimating high resolution range signatures of a radar target. High resolution range signatures are useful for many applications including the non-cooperative target identification (NCTI).

5.2.1 Problem Formulation

The range resolution of a radar is determined by the radar bandwidth. To achieve high resolution in range, the radar must transmit wideband pulses, which are often linear frequency modulated (chirp) pulses. Upon receiving each pulse returned by a radar target, the radar demodulates the pulse by mixing the pulse with a reference chirp signal and low-pass filter the mixed signal. As a result, the scattering centers of a radar target at different ranges correspond to different frequencies of the output of this demodulation operation.

Let $\mathbf{z}$ denote an $M \times 1$ vector containing the samples of the output of this operation. We assume that the radar bandwidth is moderate so that the radar cross sections (RCSs) of the scattering centers of the radar target do not change with frequency. Then $\mathbf{z}$ may be written as

$$\mathbf{z} = \sum_{k=1}^{K} \gamma_k \mathbf{f}_k + \mathbf{n}_0, \tag{5.1}$$

where γ_k, $k = 1, 2, \cdots, K$, denotes the RCS of the kth scattering center of the radar target, $\mathbf{n}_0$ denotes the additive noise vector, and

$$\mathbf{f}_k = \left[\begin{array}{cccc} 1 & e^{-jt_k} & \cdots & e^{-j(M-1)t_k} \end{array} \right]^T, \tag{5.2}$$

with $(\cdot)^T$ denoting the transpose and t_k denoting the time delay proportional to the range of the kth scattering center.

A simple nonparametric method of estimating γ_k and t_k is to use FFT (Fast Fourier Transform), which is both computationally efficient and robust to model errors. However, FFT is known for its high sidelobes and poor resolution. Many different types of windows may be applied to $\mathbf{z}$ to reduce the sidelobes. Yet using windows with FFT further reduces the already poor resolution of FFT.

When the number of scattering centers K of a radar target is small, one may use many existing frequency estimation methods, such as the computationally and asymptotically (for high signal-to-noise ratio) statistically efficient MODE method [1, 2, 3], to estimate the t_k. Once t_k are estimated, the straightforward least-squares fitting method may be used to estimate the corresponding γ_k.

However, there are two cases where the above parametric modeling approach may result in poor performance. First, for a complicated radar target such as an airplane, the number of scattering centers K may be very large and may even be larger than the number of samples M in $\mathbf{z}$. For this case, the frequency estimation methods cannot be used since to uniquely determine the time delays t_k, even in the absence of the additive noise $\mathbf{n}_0$, the K must be less than $(M-1)/2$ [4, 5]. Second, when two or more scattering centers are very closely spaced, the frequency estimation methods may not be able to resolve all of the scattering centers. When this result occurs, the least-squares fitting method yields poor RCS estimates. As a result, the estimated range signature, i.e., γ_k versus t_k, may be distorted.

We consider below a more robust parametric data model. To estimate the RCS $\gamma(t)$ of a scattering center whose range corresponds to a time delay t, we model the received data $\mathbf{z}$ as

$$\mathbf{z} = \gamma(t)\mathbf{f}(t) + \mathbf{n}(t), \tag{5.3}$$

where $\mathbf{f}(t)$ is defined in ((5.2)) with t_k replaced by t and $\mathbf{n}(t)$ denotes the additive noise. The $\mathbf{n}(t)$ is assumed to be circularly symmetric complex Gaussian random vector with zero-mean and unknown covariance matrix $\mathbf{Q}(t)$. The unknown $\mathbf{Q}(t)$ models both the thermal noise and the interferences from other scattering centers. The $\mathbf{Q}(t)$ is assumed to be a positive definite and Toeplitz matrix.

The problem of interest herein is to determine the range signature $\gamma(t)$ from $\mathbf{z}$. We remark that to compute a discrete range signature with M_s samples, we set $t = 2\pi(m-1)/M_s$ for the mth sample. We shall show below that the M_s RCS estimates in the signature can be computed in parallel.

5.2.2 The APES Algorithm

We present below an algorithm that may be used to estimate the unknown $\gamma(t)$ in model ((5.3)). We note that the amplitude and phase of $\gamma(t)$ may be considered as the amplitude and phase of the complex sinusoid with frequency t described by $\mathbf{f}(t)$. Thus we refer to our estimator below as the *Amplitude and Phase Estimation of a Sinusoid in unknown colored noise*, or simply the *APES*, method.

The APES algorithm is obtained by mimicking the computationally and asymptotically statistically efficient large sample maximum likelihood estimator obtained in Appendix A for estimating the complex gain of a signal with known waveform and known steering vector in unknown colored noise. To obtain the APES algorithm, we first divide the measurement vector $\mathbf{z}$ into $\tilde{N}$ overlapping subvectors $\tilde{\mathbf{z}}(n)$ of dimension $\tilde{M} \times 1$, where $\tilde{N} = M - \tilde{M} + 1$. The nth subvector $\tilde{\mathbf{z}}(n)$ contains the nth to the $(n + \tilde{M} - 1)$th element of $\mathbf{z}$. Let $\tilde{\mathbf{n}}(t,n)$ be formed from $\mathbf{n}(t)$ in the same way $\tilde{\mathbf{z}}(n)$ are formed from $\mathbf{z}$. The $\tilde{\mathbf{n}}(t,n)$ are assumed to be circularly symmetric complex Gaussian random vectors with zero-mean and the same unknown covariance matrix $\mathbf{Q}_1(t)$, which is a submatrix of $\mathbf{Q}(t)$. Then $\tilde{\mathbf{z}}(n)$, $n = 1, 2, \cdots, \tilde{N}$, may be written as

$$\tilde{\mathbf{z}}(n) = \gamma(t)\mathbf{a}_0(t)y(t,n) + \tilde{\mathbf{n}}(t,n), \tag{5.4}$$

where $\mathbf{a}_0(t)$ is referred to as the steering vector and has the form

$$\mathbf{a}_0(t) = \left[\begin{array}{cccc} 1 & e^{jt} & \cdots & e^{j(\tilde{M}-1)t} \end{array} \right]^T, \tag{5.5}$$

and

$$y(t,n) = e^{j(n-1)t}, \quad n = 1, 2, \cdots, \tilde{N}. \tag{5.6}$$

Let $\mathbf{J}$ be the $\tilde{M} \times \tilde{M}$ exchange matrix (with ones on the antidiagonal and zeros elsewhere). Let

$$\bar{\mathbf{z}}(n) = \mathbf{J}\tilde{\mathbf{z}}^*(\tilde{N} - n + 1), \tag{5.7}$$

where $(\cdot)^*$ denotes the complex conjugate. Let

$$\bar{\mathbf{n}}(t,n) = \mathbf{J}\tilde{\mathbf{n}}^*(t, \tilde{N} - n + 1), \tag{5.8}$$

where $\bar{\mathbf{n}}(t,n)$ are assumed to have the same statistics as $\tilde{\mathbf{n}}(t,n)$. Then $\bar{\mathbf{z}}(n)$ may be written as

$$\bar{\mathbf{z}}(n) = \gamma^*(t)\beta(t)\mathbf{a}_0(t)y(t,n) + \bar{\mathbf{n}}(t,n), \tag{5.9}$$

where

$$\beta(t) = e^{-j(N-1)t}. \tag{5.10}$$

Let

$$\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t) = \frac{1}{\tilde{N}} \sum_{n=1}^{\tilde{N}} \tilde{\mathbf{z}}(n) y^H(t,n), \tag{5.11}$$

where $(\cdot)^H$ denotes the complex conjugate transpose for matrices and complex conjugate for scalars. Let $\hat{r}_{yy}(t)$, $\hat{\mathbf{R}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$, $\hat{\mathbf{R}}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}$, and $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t)$ be defined similarly as $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)$. Note from ((5.6)) that $\hat{r}_{yy}(t) = 1$. Note also that as may be seen from ((5.11)), the mth element of $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)$ is the discrete-time Fourier transform of the mth sequence $\{\tilde{z}_m(1), \cdots, \tilde{z}_m(\tilde{N})\}$ divided by $\tilde{N}$, where $\tilde{z}_m(n)$ denotes the mth element of $\tilde{\mathbf{z}}(n)$. Thus to compute a discrete range signature with M_s samples, where M_s is a power of 2, the $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)$, and similarly the $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t)$, may be computed with the FFT method and zero-padding. The amount of computations required for each sequence is $O[M_s \log_2(M_s)]$. Since the 2nd through the $\tilde{N}$th elements of the mth sequence are the 1st through the $(\tilde{N}-1)$th elements of the $(m+1)$th sequence, the FFT of the $(m+1)$th sequence may be updated from the FFT of the mth sequence with $O(M_s)$ operations.

Let $\tilde{\mathbf{b}}(t) = \gamma(t)\mathbf{a}_0(t)$. Applying the results in Appendix A to $\tilde{\mathbf{z}}(n)$, we obtain the estimate of $\tilde{\mathbf{b}}(t)$ as

$$\hat{\tilde{\mathbf{b}}}(t) = \hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t). \tag{5.12}$$

Similarly, let $\bar{\mathbf{b}}(t) = \gamma^*(t)\beta(t)\mathbf{a}_0(t)$. Applying the results in Appendix A to $\bar{\mathbf{z}}(n)$, we have

$$\hat{\bar{\mathbf{b}}}(t) = \hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t). \tag{5.13}$$

By applying the results in Appendix A to both $\tilde{\mathbf{z}}(n)$ and $\bar{\mathbf{z}}(n)$, the covariance matrix $\mathbf{Q}_1(t)$ may be estimated as

$$\begin{aligned}
\hat{\mathbf{Q}}_1(t) &= \frac{1}{2}\left[\hat{\mathbf{R}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} - \hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}^H(t) + \hat{\mathbf{R}}_{\bar{\mathbf{z}}\bar{\mathbf{z}}} - \hat{\mathbf{R}}_{\bar{\mathbf{z}}y}(t)\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}^H(t)\right] \\
&= \frac{1}{2}\left\{\hat{\mathbf{R}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \hat{\mathbf{R}}_{\bar{\mathbf{z}}\bar{\mathbf{z}}} - \left[\begin{array}{cc} \hat{\tilde{\mathbf{b}}}(t) & \hat{\bar{\mathbf{b}}}(t) \end{array}\right]\left[\begin{array}{cc} \hat{\tilde{\mathbf{b}}}(t) & \hat{\bar{\mathbf{b}}}(t) \end{array}\right]^H\right\}
\end{aligned}$$

$$\triangleq \frac{1}{2}\left[\hat{\mathbf{R}} + \hat{\mathbf{B}}(t)\hat{\mathbf{B}}^H(t)\right]. \tag{5.14}$$

Then the estimate of $\gamma(t)$ may be computed as (see Appendix A)

$$\hat{\gamma}(t) = \frac{1}{2}\left[\frac{\mathbf{a}_0^H(t)\hat{\mathbf{Q}}_1^{-1}(t)\hat{\tilde{\mathbf{b}}}(t) + \beta(t)\hat{\bar{\mathbf{b}}}^H(t)\hat{\mathbf{Q}}_1^{-1}(t)\mathbf{a}_0(t)}{\mathbf{a}_0^H(t)\hat{\mathbf{Q}}_1^{-1}(t)\mathbf{a}_0(t)}\right]. \tag{5.15}$$

Note that since only $\hat{\mathbf{B}}(t)$ is a function of the range related time delay t, the $\hat{\mathbf{Q}}_1^{-1}(t)$ can be computed more efficiently with

$$\hat{\mathbf{Q}}_1^{-1}(t) = 2\left\{\hat{\mathbf{R}}^{-1} - \hat{\mathbf{R}}^{-1}\hat{\mathbf{B}}(t)\left[\hat{\mathbf{B}}^H(t)\hat{\mathbf{R}}^{-1}\hat{\mathbf{B}}(t) - \mathbf{I}\right]^{-1}\hat{\mathbf{B}}^H(t)\hat{\mathbf{R}}^{-1}\right\}. \tag{5.16}$$

We note that computing $\hat{\mathbf{R}}^{-1}$ requires $O(\tilde{M}^3)$ computations. Once $\hat{\mathbf{R}}^{-1}$ is obtained, computing $\hat{\mathbf{Q}}_1^{-1}(t)$ for each t requires $O(\tilde{M}^2)$ computations. Computing $\hat{\gamma}(t)$ for each t also requires $O(\tilde{M}^2)$ computations.

Then the APES estimator for estimating $\gamma(t)$ may be summarized as follows:

Step 1: Compute $\hat{\tilde{\mathbf{b}}}(t)$, $\hat{\bar{\mathbf{b}}}(t)$, and $\hat{\mathbf{Q}}_1(t)$ with ((5.12)), ((5.13)), and ((5.16)), respectively.

Step 2: Determine $\hat{\gamma}(t)$ with ((5.15)).

We remark that the vectors $\tilde{\mathbf{n}}(t,n)$ and $\bar{\mathbf{n}}(t,n)$ are not independent of each other since they are formed as the overlapping subvectors of $\mathbf{n}(t)$. Yet since these vectors are not completely correlated with each other, it can be shown that as $\tilde{N}$ goes to infinity, $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)$, $\hat{\mathbf{R}}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}$, $\hat{\mathbf{R}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$, and $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t)$ are all consistent estimates. As a result, the $\hat{\gamma}(t)$, for all possible t, is also a consistent estimate of $\gamma(t)$.

We also remark that the APES algorithm only requires simple matrix multiplications and matrix inverses. Moreover, many computations of the APES algorithm can be done in parallel. It is thus possible to achieve real-time range signature estimation by implementing the APES algorithm with parallel processors and/or specially designed hardwares.

We note that for the special case of $\tilde{M} = 1$, the $\hat{\gamma}(t)$ in ((5.15)) can be shown to be the Fourier transform of $\mathbf{z}$. Thus the Fourier transform method is a special case of the APES method.

Finally, the parameter $\tilde{M}$ has the following effects on the APES performance. Note first that the larger the $\tilde{M}$, the larger the dimension of $\mathbf{Q}_1(t)$, and thus the better the modeling of the interferences. We shall show in the following subsection that the larger the $\tilde{M}$, the better the resolution of the APES method. On the other hand, the larger the $\tilde{M}$, the smaller the $\tilde{N} = M - \tilde{M} + 1$. Thus increasing $\tilde{M}$ increases the variance of $\hat{\mathbf{Q}}_1(t)$ since $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)$, $\hat{\mathbf{R}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$, $\hat{\mathbf{R}}_{\bar{\mathbf{z}}\mathbf{z}}$, and $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t)$ are poorer estimates for larger $\tilde{M}$. We shall show in the following subsection that for very large $\tilde{M}$, the variance of $\hat{\gamma}(t)$ may increase. Also, increasing $\tilde{M}$ increases the amount of computations needed by the APES method.

5.2.3 Numerical and Experimental Results

We present below both numerical and experimental examples showing the performance of the APES algorithm. In the simulated numerical examples, the root-mean-squared errors (RMSEs) of the APES estimator are obtained with 100 independent Monte-Carlo trials and are compared with the corresponding Cramer-Rao bounds (CRBs) derived in Appendix A.

We first use simple numerical examples to illustrate the performance of the APES estimator. Consider first an example where the true range signature is shown in Figure 5.1(a). We assume that there are 66 scattering centers in the signature. The number of data samples in the received data vector $\mathbf{z}$ in ((5.3)) is assumed to be $M = 128$. The additive thermal noise in $\mathbf{z}$ is assumed to be a zero-mean white Gaussian random process with variance 1. Figure 5.1(b) shows the estimated range signature (amplitude only) with the FFT method, which is also equivalent to the APES method with $\tilde{M} = 1$. Note that due to the large sidelobes of the FFT method, the fourth scattering center is not discernible in Figure 5.1(b). Figure 5.1(c) shows the estimated range signature with the APES method with $\tilde{M} = 48$. Comparing Figures 5.1(b) and (c), we note that using APES method with $\tilde{M} = 48$ yields a much better range signature estimate, which has much reduced sidelobes and better resolution for the large scattering centers, than using the FFT method. For the 62 very small and very closely spaced scattering centers between 0.97 and 2.47, the APES method with $\tilde{M} = 48$ tends to suppress them. Figure 5.1(d) shows the estimated range signature obtained when using the MODE with the least-squares fitting (MODE-LSF) method by assuming that there are 48 complex sinusoids plus white noise in the received data vector $\mathbf{z}$. We note that since the data model is incorrect for the MODE-LSF method and some of the estimated scattering centers are very close to each other, using the MODE-LSF method may yield very large false peaks. For this case, they occur near $t = 2.5$. (Note that this type of large false peaks do not occur in every Monte-Carlo simulation. They occur in about 20% of the Monte-Carlo simulations.) The presence of these large false peaks makes the MODE-LSF method less preferable even than the FFT method. Figure 5.1(e) shows the estimated range signature obtained by using FFT with Kaiser window and shape parameter 4. Comparing Figures 5.1(b), (c), and (e), we note that using FFT with Kaiser window can reduce the FFT sidelobes, but the already poor resolution of the FFT method is made poorer due to the windowing.

We now consider the effects of $\tilde{M}$ on the performance of the APES estimator. Figure 5.2 shows the RMSEs of $\hat{\gamma}(t = 0.2332)$ (the first scatterer in Figure 5.1(a)) obtained with the APES estimator as a function of $\tilde{M}$. We note from Figure 5.2 that the APES estimator with a proper $\tilde{M}$ can give much more accurate estimates of $\gamma(t = 0.2332)$ than the FFT method, which is equivalent to the APES method with $\tilde{M} = 1$. With a proper $\tilde{M}$, the performance of the APES estimator can be close the corresponding CRB, which is also the best unbiased performance that can be achieved by an estimator. Note also that for very large $\tilde{M}$, the RMSE of $\hat{\gamma}(t = 0.2332)$

may increase due to the poor estimate of $\bar{\mathbf{Q}}_1(t)$. To achieve the best estimate of $\gamma(t = 0.2332)$, the $\tilde{M}$ should be within $20 \leq \tilde{M} \leq 64$.

Figures 5.3(a) and (b) show the range signature estimates obtained with the APES method when $\tilde{M} = 28$ and $\tilde{M} = 64$, respectively. We note from Figure 5.2(a) that using the APES method with any $\tilde{M}$ within $20 \leq \tilde{M} \leq 64$ yields similar RMSEs for $\hat{\gamma}(t = 0.2332)$. Yet Figure 5.3 shows that the larger the $\tilde{M}$ within $20 \leq \tilde{M} \leq 64$, the better the resolution of the estimated range signature. The larger the $\tilde{M}$, however, the larger the amount of computations needed by the APES method.

Consider next an example where the true range signature is shown in Figure 5.4(a). We assume that there are 75 scattering centers in the signature. The number of data samples in the received data vector $\mathbf{z}$ in ((5.3)) is assumed to be $M = 128$. The additive thermal noise in $\mathbf{z}$ is assumed to be a zero-mean white Gaussian random process with variance 10. Figure 5.4(b) shows the estimated range signature with the FFT method. Note that due to the large sidelobes of the FFT method, the two small scattering centers that are to left and right of the first group of the large scattering centers are not discernible in Figure 5.4(b). Figure 5.4(c) shows the estimated range signature with the APES method with $\tilde{M} = 48$. We note that the two small scattering centers that are to left and right of the first group of the large scattering centers are resolved in Figure 5.4(c) since the sidelobes are very small for the APES method with $\tilde{M} = 48$. Comparing Figures 5.4(b) and (c), we note that neither the FFT nor the APES method with $\tilde{M} = 48$ can resolve the three groups of very closely spaced scattering centers. Both methods tend to combine the RCSs of very closely spaced scattering centers together to yield a large peak in the range signature since these scattering centers have the same phase (see Figure 5.4(a)). Figure 5.4(d) shows the estimated range signature with the MODE-LSF method by assuming that there are 48 complex sinusoids plus white noise in the received data vector $\mathbf{z}$. We note that since the data model is incorrect for the MODE-LSF method and some of the estimated scatterer locations are very close to each other, using the MODE-LSF method may yield large false peaks. For this example, they occur near $t = 1.7$. Figure 5.4(e) shows the estimated range signature obtained by using FFT with Kaiser window and shape parameter 4. Comparing Figures 5.4(b) and (e), we note that using FFT with Kaiser window further widens the large peaks in 5.4(b).

Finally, we apply below the APES method to the experimental data measured by a ground-to-air radar. (All other information about the radar and the radar target is not available for public release.) The measured data consists of $M = 128$ samples and is also degraded with the zero-mean white Gaussian noise with variance 0.1. Figure 5.5(a) shows the target range signature of an aircraft obtained with the FFT method. Figure 5.5(b) shows the target range signature of the aircraft obtained by using the FFT method with the Kaiser window and shape parameter 4. Figure 5.5(c) shows the target range signature obtained with the APES method with $\tilde{M} = 48$. Comparing Figures 5.5(a) and (c), we note that using the APES method with $\tilde{M} = 48$ yields much lower sidelobes than using the FFT method. Comparing Figures 5.5(b) and (c), we note that using the APES method with $\tilde{M} = 48$ yields better resolution and sharper peaks than using the windowed FFT. Figure 5.5(d) shows the target range signature obtained when using the MODE-LSF method and assuming that there are 48 complex sinusoids plus white noise in the measured data. We note that using the MODE-LSF method can yield poor target range signatures due to inaccurate data models.

5.3 Using the APES Algorithm for SAR Imaging

In this section, we extend the parametric model we proposed in the previous section to synthetic aperture radar (SAR) imaging. We also describe how the APES algorithm presented in the previous section may be used for SAR imaging.

5.3.1 Problem Formulation

In synthetic aperture radar (SAR) or inverse SAR (ISAR) imaging [6, 7], the radar usually transmits linear frequency modulated (chirp) pulses. Upon receiving each pulse returned by an object being imaged, the radar demodulates the pulse by mixing the pulse with a reference chirp signal and low-pass filter the mixed signal. As a result, the scattering centers of the object at different ranges correspond to different frequencies of the output of this demodulation operation. Since either the radar or the object is moving or rotating, the pulses received at different angles between the radar and the object are used to form a synthetic aperture. After Polar-to-Cartesian interpolation, the scattering centers of the object at the same range but different cross-ranges correspond to different (Doppler) frequencies over the synthetic aperture.

Let **Z** denote an $M_1 \times M_2$ matrix containing the samples of the demodulated and Polar-to-Cartesian interpolated data, from which a SAR image is computed. A simple nonparametric method of computing the SAR image is to use the two-dimensional (2-D) FFT, which is again both computationally efficient and robust to model errors. However, FFT is known for its high sidelobes and poor resolution. Many different types of windows may be applied to **Z** to reduce the sidelobes. Yet using windows with FFT further reduces the already poor resolution of FFT. Thus to use 2-D windowed FFT to obtain high resolution SAR images, a large radar bandwidth and a large synthetic aperture are needed. One of the disadvantages of this requirement is that it makes it difficult to use synthetic aperture radar technology for wide-area surveillance.

To achieve high resolution with limited radar bandwidth and synthetic aperture, many parametric and nonparametric spectral estimation methods have been used for SAR imaging [8, 9, 10, 11, 12, 13, 14]. For example, the **Z** may be written as the sum of 2-D complex sinusoids in additive noise. The 2-D frequencies of the sinusoids may be estimated with computationally and asymptotically (for high signal-to-noise ratio) statistically efficient 2D-MODE method [14]. The complex amplitudes of the sinusoids may then be estimated with the straightforward least-squares fitting [9]. However, similar to the case of the high resolution range estimation, this parametric modeling approach may result in poor performance when the number of sinusoids is too large to uniquely determine the parameters of the sinusoids [14] or when two or more sinusoids are so closely spaced that 2-D frequency estimation methods cannot resolve all of them.

We consider below a more robust parametric data model, which is similar to the one proposed in the previous section. Let $\gamma(t_1, t_2)$ denote the complex intensity of the (i, j)th pixel of an $M_{s1} \times M_{s2}$ SAR image we intend to form, where $t_1 = 2\pi(i-1)/M_{s1}$ and $t_2 = 2\pi(j-1)/M_{s2}$. To estimate $\gamma(t_1, t_2)$, we model the received data $\mathbf{Z}$ as

$$\mathbf{Z} = \gamma(t_1, t_2)\mathbf{f}(t_1) \otimes \mathbf{f}^T(t_2) + \mathbf{N}(t_1, t_2), \tag{5.17}$$

where $\otimes$ denotes the Kronecker matrix product, $\mathbf{f}(t_i)$, $i = 1, 2$, is defined in ((5.2)) with t_k replaced by t_i and M replaced by M_i, and $\mathbf{N}(t_1, t_2)$ denotes the additive noise. The vec$[\mathbf{N}(t_1, t_2)]$, where vec$(\cdot)$ denotes stacking all columns of a matrix into a single column vector, is assumed to be a circularly symmetric complex Gaussian random vector with zero-mean and unknown covariance matrix $\mathbf{Q}(t_1, t_2)$. The unknown $\mathbf{Q}(t_1, t_2)$ models both the thermal noise and the interferences from other pixels of the SAR image. The $\mathbf{Q}(t_1, t_2)$ is assumed to be a positive definite and block-Toeplitz matrix whose blocks are also Toeplitz matrices.

The problem of interest herein is to determine the range signature $\gamma(t_1, t_2)$ from $\mathbf{Z}$. We remark that the complex intensities $\gamma(t_1, t_2)$ of all pixels in a SAR image can be computed in parallel.

5.3.2 SAR Imaging with the APES Algorithm

The APES algorithm may be applied to the rows (columns) and then to the columns (rows) of the data matrix $\mathbf{Z}$ to form the SAR image. We referred to this approach as the *one-dimensional APES* or the *1-D APES* algorithm. We also present below a *two-dimensional APES* or the *2-D APES* algorithm. We shall show in the next subsection that the 2-D APES algorithm may be computationally more intensive but may yield better SAR images than the 1-D APES algorithm.

To obtain the 2-D APES algorithm, we first divide the data matrix $\mathbf{Z}$ into $\tilde{N}_1\tilde{N}_2$ overlapping submatrices $\tilde{\mathbf{Z}}(n)$ of dimension $\tilde{M}_1 \times \tilde{M}_2$ (in raster sequence), where $\tilde{N}_1 = M_1 - \tilde{M}_1 + 1$ and $\tilde{N}_2 = M_2 - \tilde{M}_2 + 1$. Let $\tilde{\mathbf{N}}(t_1, t_2, n)$ be formed from $\mathbf{N}(t_1, t_2)$ in the same way $\tilde{\mathbf{Z}}(n)$ are formed from $\mathbf{Z}$. Let

$$\tilde{\mathbf{n}}(t_1, t_2, n) = \text{vec}\left[\tilde{\mathbf{N}}(t_1, t_2, n)\right]. \tag{5.18}$$

The $\tilde{\mathbf{n}}(t_1, t_2, n)$ are assumed to be circularly symmetric complex Gaussian random vectors with zero-mean and the same unknown covariance matrix $\mathbf{Q}_1(t_1, t_2)$, which has the same form as $\mathbf{Q}(t_1, t_2)$. Let

$$\tilde{\mathbf{z}}(n) = \text{vec}\left[\tilde{\mathbf{Z}}(n)\right]. \tag{5.19}$$

Then $\tilde{\mathbf{z}}(n)$, $n = 1, 2, \cdots, \tilde{N}_1\tilde{N}_2$, may be written as

$$\tilde{\mathbf{z}}(n) = \gamma(t_1, t_2)\mathbf{a}_0(t_1, t_2)y(t_1, t_2, n) + \tilde{\mathbf{n}}(t_1, t_2, n), \tag{5.20}$$

where $\mathbf{a}_0(t_1, t_2)$ is referred to as the *steering vector* and has the form

$$\mathbf{a}_0(t_1, t_2) = \mathbf{a}_1(t_1) \otimes \mathbf{a}_2(t_2), \tag{5.21}$$

where

$$\mathbf{a}_i(t_i) = \begin{bmatrix} 1 & e^{jt_i} & \dots & e^{j(\tilde{M}_i - 1)t_i} \end{bmatrix}^T, \quad i = 1, 2, \tag{5.22}$$

and $y(t_1, t_2, n)$ is the nth element of the vector $\mathbf{y}_0(t_1, t_2)$,

$$\mathbf{y}_0(t_1, t_2) = \mathbf{y}_1(t_1) \otimes \mathbf{y}_2(t_2), \tag{5.23}$$

with

$$\mathbf{y}_i(t_i) = \begin{bmatrix} 1 & e^{jt_i} & \dots & e^{j(\tilde{N}_i - 1)t_i} \end{bmatrix}^T, \quad i = 1, 2. \tag{5.24}$$

Let $\mathbf{J}$ be the $\tilde{M}_1 \tilde{M}_2 \times \tilde{M}_1 \tilde{M}_2$ exchange matrix. Let

$$\bar{\mathbf{z}}(n) = \mathbf{J} \tilde{\mathbf{z}}^*(\tilde{N}_1 \tilde{N}_2 - n + 1), \tag{5.25}$$

Let

$$\bar{\mathbf{n}}(t_1, t_2, n) = \mathbf{J} \tilde{\mathbf{n}}^*(t_1, t_2, \tilde{N}_1 \tilde{N}_2 - n + 1), \tag{5.26}$$

where $\bar{\mathbf{n}}(t_1, t_2, n)$ are assumed to have the same statistics as $\tilde{\mathbf{n}}(t_1, t_2, n)$. Then $\bar{\mathbf{z}}(n)$ may be written as

$$\bar{\mathbf{z}}(n) = \gamma^*(t_1, t_2) \beta(t_1, t_2) \mathbf{a}_0(t_1, t_2) y(t_1, t_2, n) + \bar{\mathbf{n}}(t_1, t_2, n), \tag{5.27}$$

where

$$\beta(t_1, t_2) = e^{-j[(N_1 - 1)t_1 + (N_2 - 1)t_2]}. \tag{5.28}$$

With these new notation definitions, the remaining steps of computing $\hat{\gamma}(t_1, t_2)$ in the 2-D APES algorithm are similar to those of computing $\hat{\gamma}(t)$ in the APES algorithm presented in the previous section. Let

$$\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t_1, t_2) = \frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{n=1}^{\tilde{N}_1 \tilde{N}_2} \tilde{\mathbf{z}}(n) y(t_1, t_2), \tag{5.29}$$

and $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t_1, t_2)$ be defined similarly. Then similar to using FFT to compute $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t)$ and $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t)$ in the previous section, the 2-D FFT method can be used to compute $\hat{\mathbf{r}}_{\bar{\mathbf{z}}y}(t_1, t_2)$ and $\hat{\mathbf{r}}_{\tilde{\mathbf{z}}y}(t_1, t_2)$ more efficiently.

We remark that for the special case of $\tilde{M}_1 = \tilde{M}_2 = 1$, the $\hat{\gamma}(t_1, t_2)$ obtained with either the 1-D or the 2-D APES algorithm can be shown to be the 2-D Fourier transform of **Z**. Thus the Fourier transform method is again a special case of the APES methods.

We also remark that both the 1-D and 2-D APES algorithms only require simple matrix multiplications and matrix inverses. Moreover, many computations of the APES algorithms can be done in parallel. It is thus possible to achieve real-time SAR image formation by implementing the APES algorithm with parallel processors and/or specially designed hardwares.

5.3.3 XPATCH Examples

We present below simulated examples showing the performance of the 1-D and 2-D APES algorithms for SAR imaging. The data we use are generated by XPATCH, which is a computational electromagnetics software [15].

Figure 5.6(a) shows the SAR image (the logarithm of the amplitude) of a firetruck obtained with the 2-D FFT method when $M_1 = M_2 = 128$. The firetruck is simulated with the XPATCH software [15] and is added with the zero-mean white Gaussian noise with variance 0.1. Figure 5.6(b) shows the SAR image obtained by using the 2-D FFT method with the circularly symmetric Kaiser window and shape parameter 6. Figure 5.6(c) shows the SAR image obtained by using the 1-D APES algorithm with $\tilde{M} = 48$. Comparing Figures 5.6(a), (b), and (c), we note that using the 1-D APES algorithm with $\tilde{M} = 48$ yields lower sidelobes and better resolution than the 2-D FFT methods. Figure 5.6(d) shows the SAR image obtained by using the 2D-MODE algorithm with the least-squares fitting (2D-MODE-LSF) and assuming that there are 121 two-dimensional complex sinusoids in additive white noise. We note that the 2D-MODE-LSF algorithm performs poorly and many features that are important for target identification, such as the target pixels to the upper right of the brightest pixels in Figure 5.6(b), are left out in Figure 5.6(d) by the 2D-MODE-LSF algorithm.

Figure 5.7 is similar to Figure 5.6 except that only 1/2 of the bandwidth and 1/2 of the synthetic aperture are used for SAR imaging. Further, Figure 5.7(c) is obtained with the 1-D APES algorithm with $\tilde{M} = 24$. Figure 5.8 is also similar to Figure 5.6 except that only 1/3 of the bandwidth and 1/3 of the synthetic aperture are used for SAR imaging. Further, Figure 5.8(c) is obtained with the 1-D APES algorithm with $\tilde{M} = 16$. Figure 5.8(d) is obtained with the 2-D APES algorithm with $\tilde{M}_1 = \tilde{M}_2 = 16$. Figure 5.8(e) is obtained similarly as Figure 5.7(d). We again note that using the APES algorithms yield lower sidelobes and better resolution than the 2-D FFT methods. The APES algorithm is also more robust than the 2D-MODE-LSF method. Finally, we note from Figures 5.8(c) and (d) that the 2-D APES algorithm gives better performance than the 1-D APES algorithm at the cost of much more computations.

5.4 Conclusions

We have presented a robust parametric data model for estimating high resolution range signatures of radar targets and for forming high resolution synthetic aperture radar (SAR) images. We have described how the APES (Amplitude and Phase Estimation of a Sinusoid in unknown colored noise) algorithm can be used to estimate the range signatures and to form SAR images. With both numerical and experimental examples, we have shown that our modeling and estimation approach yields better resolution and lower sidelobes than the conventional nonparametric FFT (fast Fourier transform) method. We have also shown that our approach is more robust than modeling the radar data as a certain number of complex sinusoids in noise and estimating the frequencies, amplitudes, and phases of the sinusoids with one of the best sinusoidal parameter estimation methods.

Acknowledgments

The author grateful acknowledges the many helpful discussions with Mr. E. G. Zelnio. The author is grateful to Mr. D. Zheng for generating the SAR images with the 2D-MODE-LSF estimator.

Appendix A – Large Sample Maximum Likelihood Estimator and Cramer-Rao Bound

We consider below the estimation of the complex gain of a signal with known waveform and known steering vector in unknown colored Gaussian noise. Let $\mathbf{x}(n)$, $n = 1, 2, \cdots, N$, denote some received data vectors of dimensions $M \times 1$. Let $y(n)$, $n = 1, 2, \cdots, N$, denote a known waveform. Assume that the received data vectors $\mathbf{x}(n)$ can be written as

$$\mathbf{x}(n) = \gamma \mathbf{a}_0 y(n) + \mathbf{n}(n), \tag{5.30}$$

where $\mathbf{n}$ is the additive noise vector of dimensions $M \times 1$ and $\mathbf{a}_0$ is referred to as the *steering vector* and is known. The noise vectors $\mathbf{n}(n)$ are assumed to be circularly symmetric complex Gaussian random vectors with zero-mean and arbitrary covariance matrix $\mathbf{Q}$ and are temporally white, i.e.,

$$\mathrm{E}[\mathbf{n}(i)\mathbf{n}^H(j)] = \mathbf{Q}\delta_{i,j}, \tag{5.31}$$

where $(\cdot)^H$ denotes the complex conjugate transpose and $\delta_{i,j}$ is the Kronecker delta.

We consider below a large sample ($N \gg 1$) maximum likelihood (ML) estimator for estimating γ from $\mathbf{x}(n)$, $n = 1, 2, \cdots, N$. It is easy to show that an exact ML estimator requires a multidimensional search over the parameter space and is computationally burdensome. We shall present below a large sample ML estimator that is both computationally and asymptotically (for large N) statistically efficient. The approach we use to derive the large sample ML estimator is similar to the one in [16, 17], which is devised for a different problem.

The log-likelihood function of the received vectors $\mathbf{x}(n)$, $n = 1, 2, \cdots, N$, is proportional to (within an additive constant):

$$-\ln|\mathbf{Q}| - \operatorname{tr}\left\{\mathbf{Q}^{-1}\frac{1}{N}\sum_{n=1}^{N}[\mathbf{x}(n) - \mathbf{b}y(n)][\mathbf{x}(n) - \mathbf{b}y(n)]^H\right\},$$

where $|\cdot|$ denotes the determinant of a matrix and

$$\mathbf{b} = \gamma\mathbf{a}_0. \tag{5.32}$$

Consider first the estimate of $\mathbf{Q}$ and the unstructured estimate of $\mathbf{b}$. It is easy to show that

$$\hat{\mathbf{Q}} = \frac{1}{N}\sum_{n=1}^{N}[\mathbf{x}(n) - \hat{\mathbf{b}}y(n)][\mathbf{x}(n) - \hat{\mathbf{b}}y(n)]^H, \tag{5.33}$$

and $\hat{\mathbf{b}}$ may be obtained by minimizing the following cost function

$$F = \left|\frac{1}{N}\sum_{n=1}^{N}[\mathbf{x}(n) - \mathbf{b}y(n)][\mathbf{x}(n) - \mathbf{b}y(n)]^H\right|. \tag{5.34}$$

Let

$$\hat{\mathbf{r}}_{y\mathbf{x}} = \frac{1}{N}\sum_{n=1}^{N} y(n)\mathbf{x}^H(n), \tag{5.35}$$

and

$$\hat{r}_{yy} = \frac{1}{N}\sum_{n=1}^{N} y(n)y^H(n). \tag{5.36}$$

Let $\hat{\mathbf{R}}_{\mathbf{xx}}$ be defined similarly as $\hat{r}_{yy}$. Then let [18]

$$\begin{aligned}
\mathbf{G} &= \frac{1}{N}\sum_{n=1}^{N}[\mathbf{x}(n) - \mathbf{b}y(n)][\mathbf{x}(n) - \mathbf{b}y(n)]^H & (5.37)\\
&= \hat{\mathbf{R}}_{\mathbf{xx}} - \mathbf{b}\hat{\mathbf{r}}_{y\mathbf{x}} - \hat{\mathbf{r}}_{y\mathbf{x}}^H\mathbf{b}^H + \mathbf{b}\hat{r}_{yy}\mathbf{b}^H & (5.38)\\
&= \left[\mathbf{b} - \hat{\mathbf{r}}_{y\mathbf{x}}^H\hat{r}_{yy}^{-1}\right]\hat{r}_{yy}\left[\mathbf{b} - \hat{\mathbf{r}}_{y\mathbf{x}}^H\hat{r}_{yy}^{-1}\right]^H + \hat{\mathbf{R}}_{\mathbf{xx}} - \hat{\mathbf{r}}_{y\mathbf{x}}^H\hat{r}_{yy}^{-1}\hat{\mathbf{r}}_{y\mathbf{x}}. & (5.39)
\end{aligned}$$

Since $\hat{r}_{yy}$ is a positive number and the second and third terms in ((5.39)) do not depend on $\mathbf{b}$, it follows that

$$\mathbf{G} \geq \mathbf{G} \mid_{\mathbf{b}=\hat{\mathbf{b}}}, \tag{5.40}$$

where

$$\hat{\mathbf{b}} = \hat{\mathbf{r}}_{y\mathbf{x}}^{H} \hat{r}_{yy}^{-1}. \tag{5.41}$$

Since the whole sample covariance matrix $\mathbf{G}$ is minimized, the unstructured estimate $\hat{\mathbf{b}}$ of $\mathbf{b}$ in ((5.41)) will minimize any nondecreasing function of $\mathbf{G}$ including the determinant of $\mathbf{G}$, which is F in ((5.34)). It is easy to see that $\hat{\mathbf{b}}$ is a consistent estimate of $\mathbf{b}$.

By using ((5.41)) with ((5.33)), the $\hat{\mathbf{Q}}$ may be rewritten as

$$\hat{\mathbf{Q}} = \hat{\mathbf{R}}_{\mathbf{xx}} - \hat{\mathbf{r}}_{y\mathbf{x}}^{H} \hat{r}_{yy}^{-1} \hat{\mathbf{r}}_{y\mathbf{x}}. \tag{5.42}$$

It is easy to see that $\hat{\mathbf{Q}}$ is a consistent estimate of $\mathbf{Q}$.

Let us now consider the structure of $\mathbf{b}$. The cost function in ((5.34)) may be rewritten as

$$\begin{aligned} F &= \left|\hat{\mathbf{R}}_{\mathbf{xx}} - \mathbf{b}\hat{\mathbf{r}}_{y\mathbf{x}} - \hat{\mathbf{r}}_{y\mathbf{x}}^{H}\mathbf{b}^{H} + \mathbf{b}\hat{r}_{yy}\mathbf{b}^{H}\right| && (5.43) \\ &= \left|\hat{\mathbf{R}}_{\mathbf{xx}} - \hat{\mathbf{b}}\hat{r}_{yy}\hat{\mathbf{b}}^{H} + (\mathbf{b}-\hat{\mathbf{b}})\hat{r}_{yy}(\mathbf{b}-\hat{\mathbf{b}})^{H}\right| && (5.44) \\ &= \left|\hat{\mathbf{Q}}\right| \left|\mathbf{I} + \hat{\mathbf{Q}}^{-1}(\mathbf{b}-\hat{\mathbf{b}})\hat{r}_{yy}(\mathbf{b}-\hat{\mathbf{b}})^{H}\right|. && (5.45) \end{aligned}$$

The ML estimate of γ may be obtained by minimizing F in ((5.45)) or equivalently $\ln F$. In [16, 17], we have shown that minimizing $\ln F$, with F defined in ((5.45)), is asymptotically (for large N) equivalent to minimizing

$$F_1 = \operatorname{tr}\left[\hat{r}_{yy}(\mathbf{b}-\hat{\mathbf{b}})^{H}\hat{\mathbf{Q}}^{-1}(\mathbf{b}-\hat{\mathbf{b}})\right]. \tag{5.46}$$

Using $\mathbf{b} = \gamma\mathbf{a}_0$ and minimizing F_1 in ((5.46)) with respect to γ yields

$$\hat{\gamma} = \frac{\mathbf{a}_0^{H}\hat{\mathbf{Q}}^{-1}\hat{\mathbf{b}}}{\mathbf{a}_0^{H}\hat{\mathbf{Q}}^{-1}\mathbf{a}_0}. \tag{5.47}$$

The large sample ML estimator for estimating γ may be summarized as follows:

Step 1: Compute $\hat{\mathbf{b}}$ and $\hat{\mathbf{Q}}$ with ((5.41)) and ((5.42)), respectively.

Step 2: Determine $\hat{\gamma}$ with ((5.47)).

We remark that since $\hat{\gamma}$ is a consistent and large sample realization of the ML estimate, it follows that $\hat{\gamma}$ is asymptotically (for large N) statistically efficient according to the general properties of ML estimators [19]. Using the results in [16, 17], we can show that the asymptotic (for large N) distribution of $\hat{\gamma}$ is complex Gaussian with mean γ and variance

$$\mathrm{Var}[\hat{\gamma}] = \frac{1}{\sum_{n=1}^{N} \mathbf{y}^H(n)\mathbf{a}_0^H \mathbf{Q}^{-1}\mathbf{a}_0 \mathbf{y}(n)}, \tag{5.48}$$

which is also the CRB.

[1] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, pp. 1132–1143, July 1990.

[2] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *IEE Proceedings, Pt. F*, vol. 137, pp. 19–26, February 1990.

[3] M. Viberg, "Sensitivity of parametric direction finding to colored noise fields and undermodeling," *Signal Processing*, vol. 34, pp. 207–222, 1993.

[4] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 996–1000, July 1989.

[5] A. Nehorai, D. Starer, and P. Stoica, "Direction-of-arrival estimation in applications with multipath and few snapshots," *Circuits, Systems, and Signal Processing*, vol. 10, pp. 327–342, 1991.

[6] D. C. Munson, Jr., J. D. O'Brien, and W. K. Jenkins, "A tomographic formulation of spotlight-mode synthetic aperture radar," *Proceedings of IEEE*, vol. 71, pp. 917–925, August 1983.

[7] D. A. Ausherman, A. Kozma, J. L. Walker, H. M. Jones, and E. C. Poggio, "Developments in radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, pp. 363–400, July 1984.

[8] S. R. DeGraaf, "Parametric estimation of complex 2-d sinusoids," *IEEE Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, pp. 391–396, August, 1988.

[9] Y. Hua, "High resolution imaging of continuously moving object using stepped frequency radar," *Signal Processing*, vol. 35, pp. 33–40, January 1994.

[10] I. J. Gupta, "High-resolution radar imaging using 2-D linear prediction," *IEEE Transactions on Antennas and Propagation*, vol. 42, pp. 31–37, January 1994.

[11] S. R. DeGraaf, "SAR imaging via modern 2-d spectral estimation methods," *SPIE Proceedings on Optical Engineering in Aerospace Sensing,* Orlando , FL, April 1994.

[12] G. R. Benitz, "Adaptive high-definition imaging," *SPIE Proceedings on Optical Engineering in Aerospace Sensing,* Orlando , FL, April 1994.

[13] S. R. DeGraaf, "Sidelobe reduction via adaptive FIR filtering in SAR imagery," *IEEE Transactions on Image Processing*, vol. 3, pp. 292–301, May 1994.

[14] J. Li, P. Stoica, and D. Zheng, "An efficient algorithm for two-dimensional frequency estimation," submitted to *IEEE Transactions on Signal Processing.*

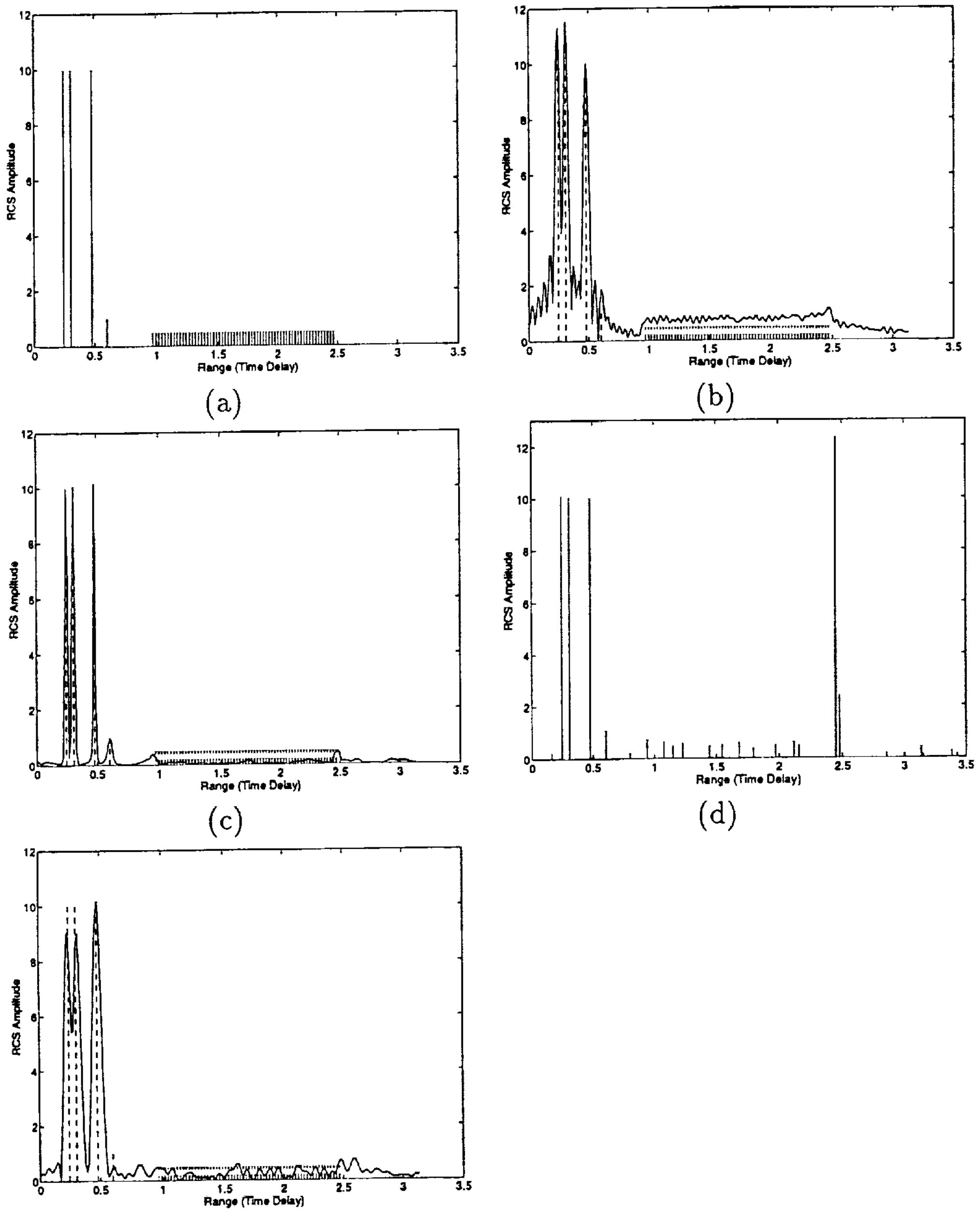
[15] D. J. Andersh, M. Hazlett, S. W. Lee, D. D. Reeves, D. P. Sullivan, and Y. Chu, "XPATCH:, a high-frequency electromagnetic scattering prediction code and environment for complex three-dimensional objects," *IEEE Antennas and Propagation Magazine*, vol. 36, pp. 65–69, February 1994.

[16] J. Li, B. Halder, P. Stoica, M. Viberg, and T. Kailath, "Decoupled maximum likelihood angle estimation for signals with known waveforms," Technical Report No. CTH-TE-8, Chalmers University of Technology, Gothenburg, Sweden, February 1994.

[17] J. Li, B. Halder, P. Stoica, and M. Viberg, "Decoupled maximum likelihood angle estimation for signals with known waveforms," *submitted to IEEE Transactions on Signal Processing.*

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

[19] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York, NY: John Wiley & Sons Inc., 1968.



(a) (b) (c) (d) (e)

Figure 5.1: Range signature estimates (solid lines in (b) - (e)) compared to the true range signature (dashed lines in (b) - (e)) when the white noise variance is 1 and $M = 128$. (a) True range signature. (b) FFT or APES with $\tilde{M} = 1$. (c) APES with $\tilde{M} = 48$. (d) MODE-LSF by assuming 48 complex sinusoids. (e) FFT with Kaiser window and shape parameter 4.

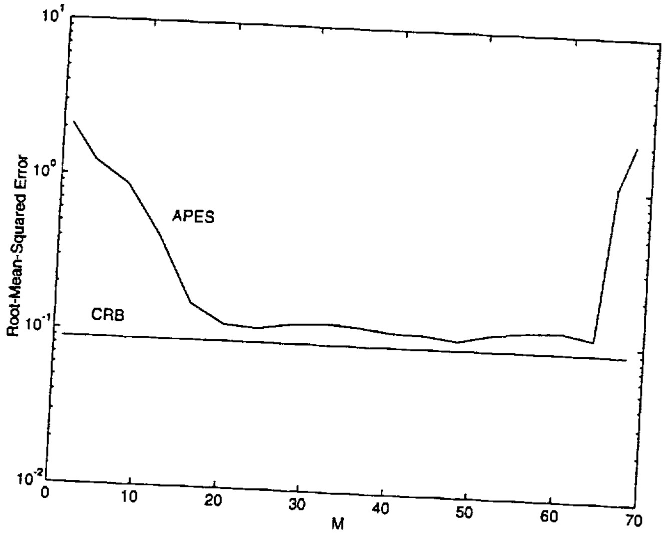
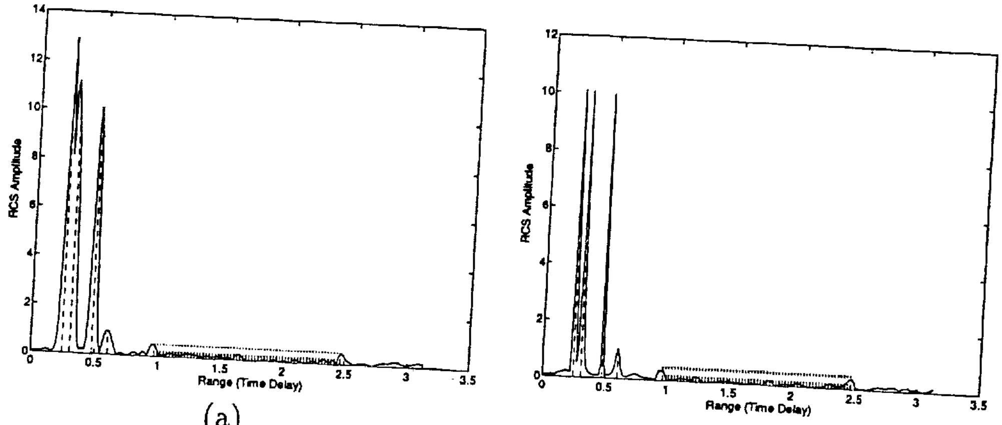


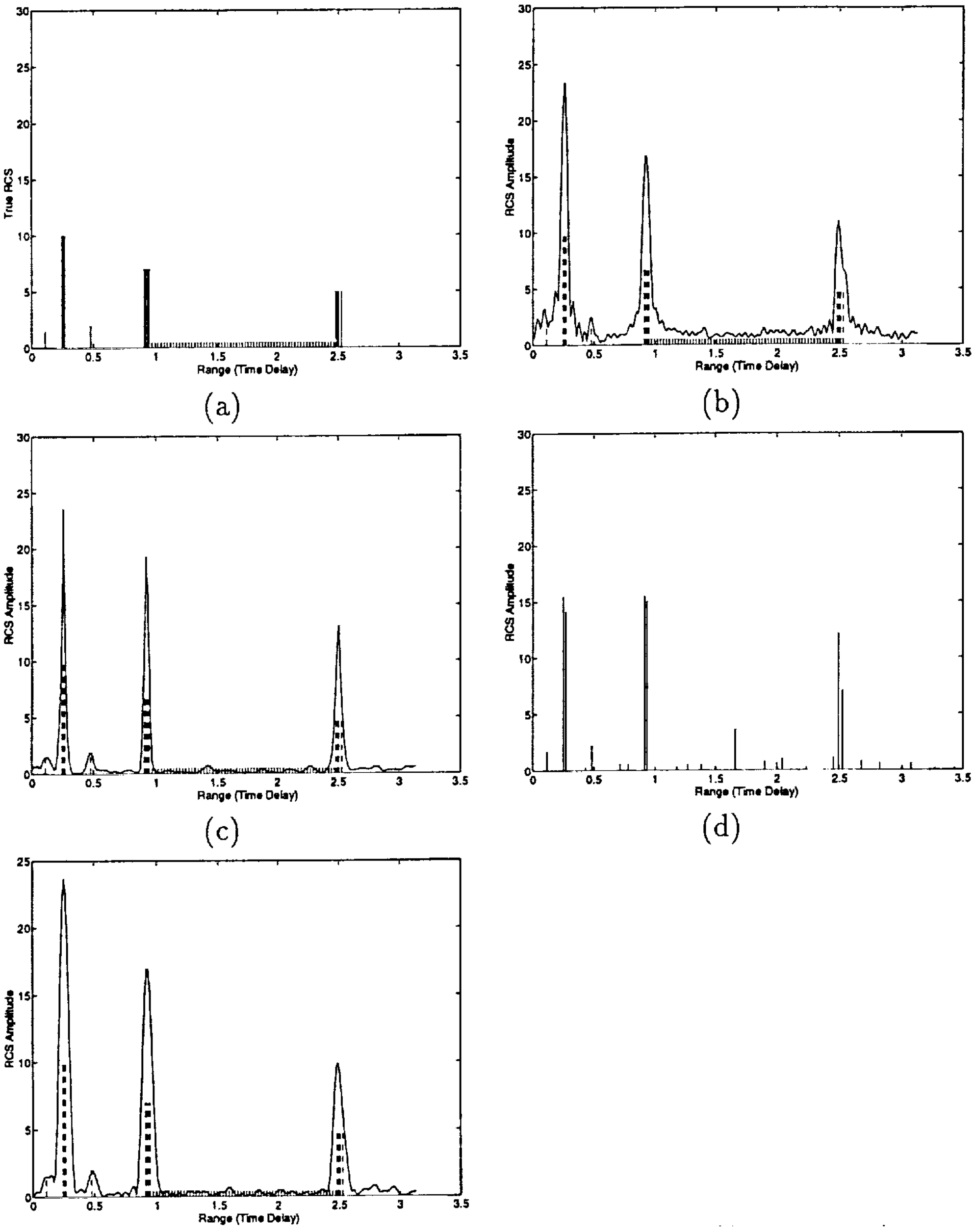
Figure 5.2: Root-Mean-squared errors (MSEs) of $\hat{\gamma}(r = 0.2332)$ obtained with the APES estimator as a function of $\tilde{M}$ for the example in Figure 5.1.



(a)

(b)

Figure 5.3: Range signature estimates (solid lines) compared to the true range signature (dashed lines) when the white noise variance is 1 and $M = 128$. (a) APES with $\tilde{M} = 28$. (b) APES with $\tilde{M} = 64$.

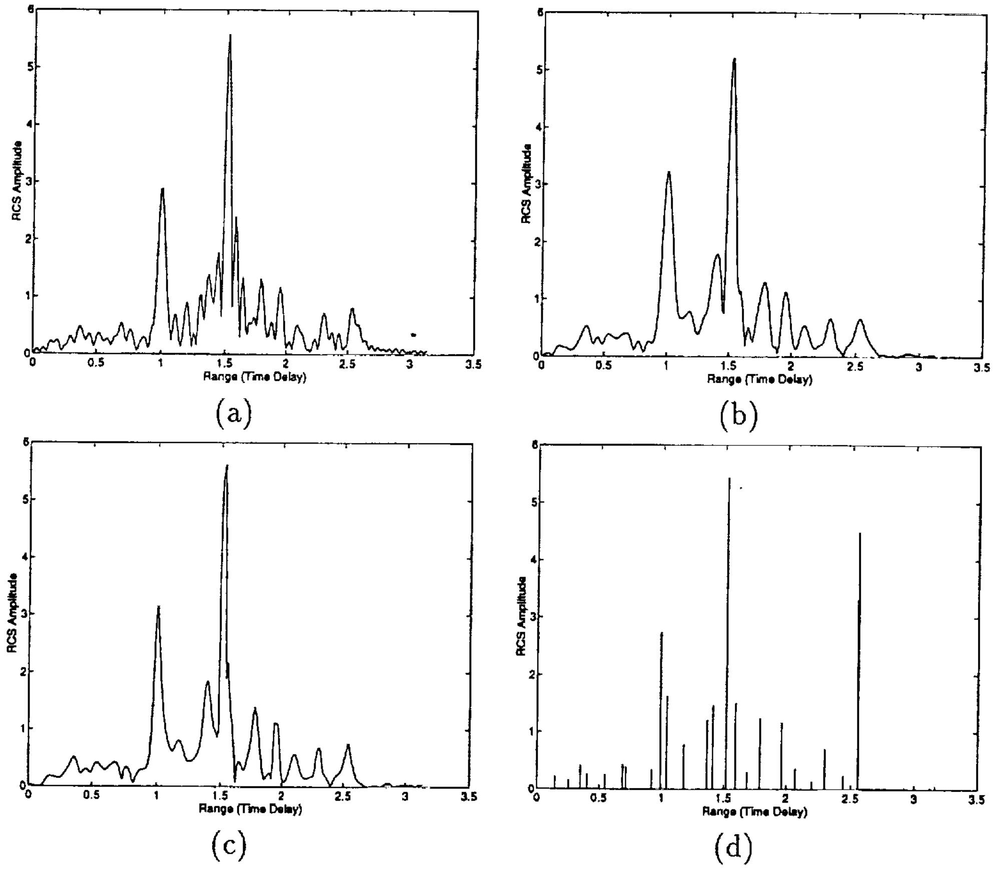


(a) (b)

(c) (d)

(e)

Figure 5.4: Range signature estimates (solid lines in (b) – (e)) compared to the true range signature (dashed lines in (b) – (e)) when the white noise variance is 10 and $M = 128$. (a) True range signature. (b) FFT or APES with $\tilde{M} = 1$. (c) APES with $\tilde{M} = 48$. (d) MODE-LSF by assuming 48 complex sinusoids. (e) FFT with Kaiser window and shape parameter 4.



(a) (b)

(c) (d)

Figure 5.5: Range signature estimates of an airplane measured with a ground-to-air radar. (a) FFT or APES with $\tilde{M} = 1$. (b) FFT with Kaiser window and shape parameter 4. (c) APES with $\tilde{M} = 48$. (d) MODE-LSF by assuming 48 complex sinusoids.

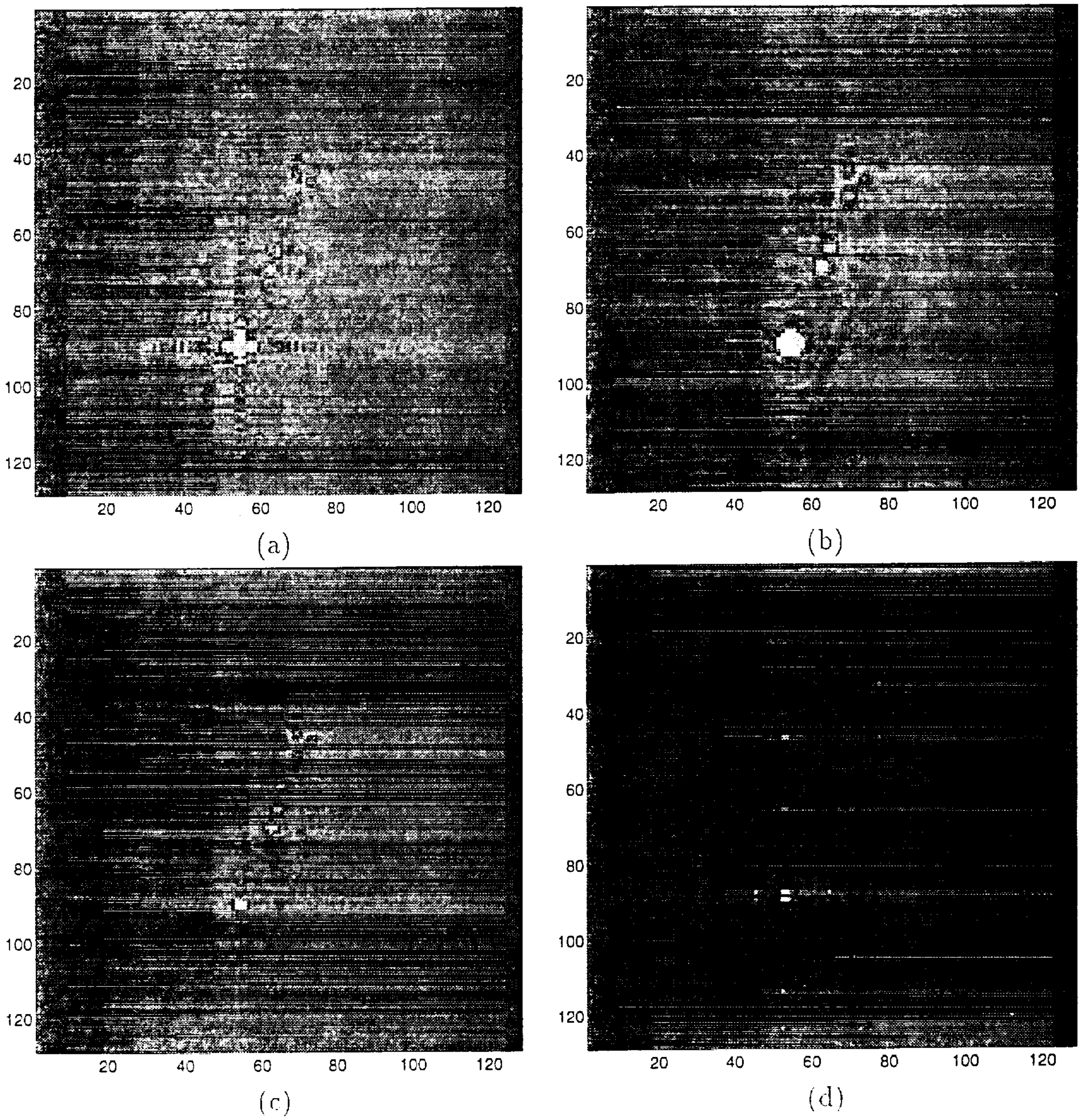
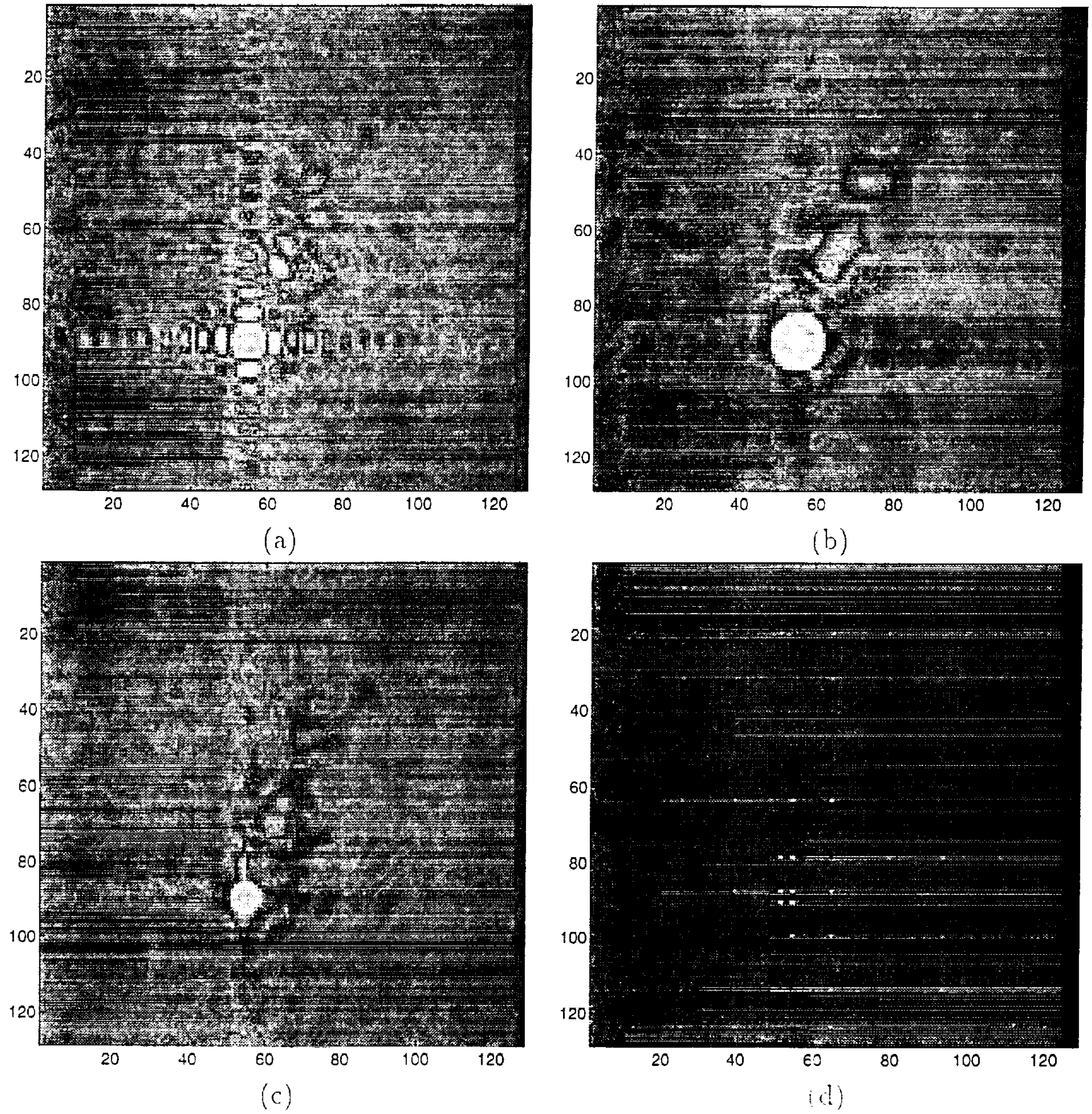


Figure 5.6: SAR images of an XPATCH simulated firetruck when the additive white noise variance is 0.1 and $M_1 = M_2 = 128$. (a) 2-D FFT. (b) 2-D FFT with circularly symmetric Kaiser window and shape parameter 6. (c) 1-D APES with $\bar{M} = 48$. (d) 2D-MODE-LSF by assuming 121 complex sinusoids.



(a)

(b)

(c)

(d)

Figure 5.7: SAR images of an XPATCH simulated firetruck when the additive white noise variance is 0.1 and $M_1 = M_2 = 128$. Further, 1/2 of the bandwidth and 1/2 of the synthetic aperture of that in Figure 5.6 are used. (a) 2-D FFT. (b) 2-D FFT with circularly symmetric Kaiser window and shape parameter 6. (c) 1-D APES with $M = 24$. (d) 2D-MODE-LSF by assuming 121 complex sinusoids.

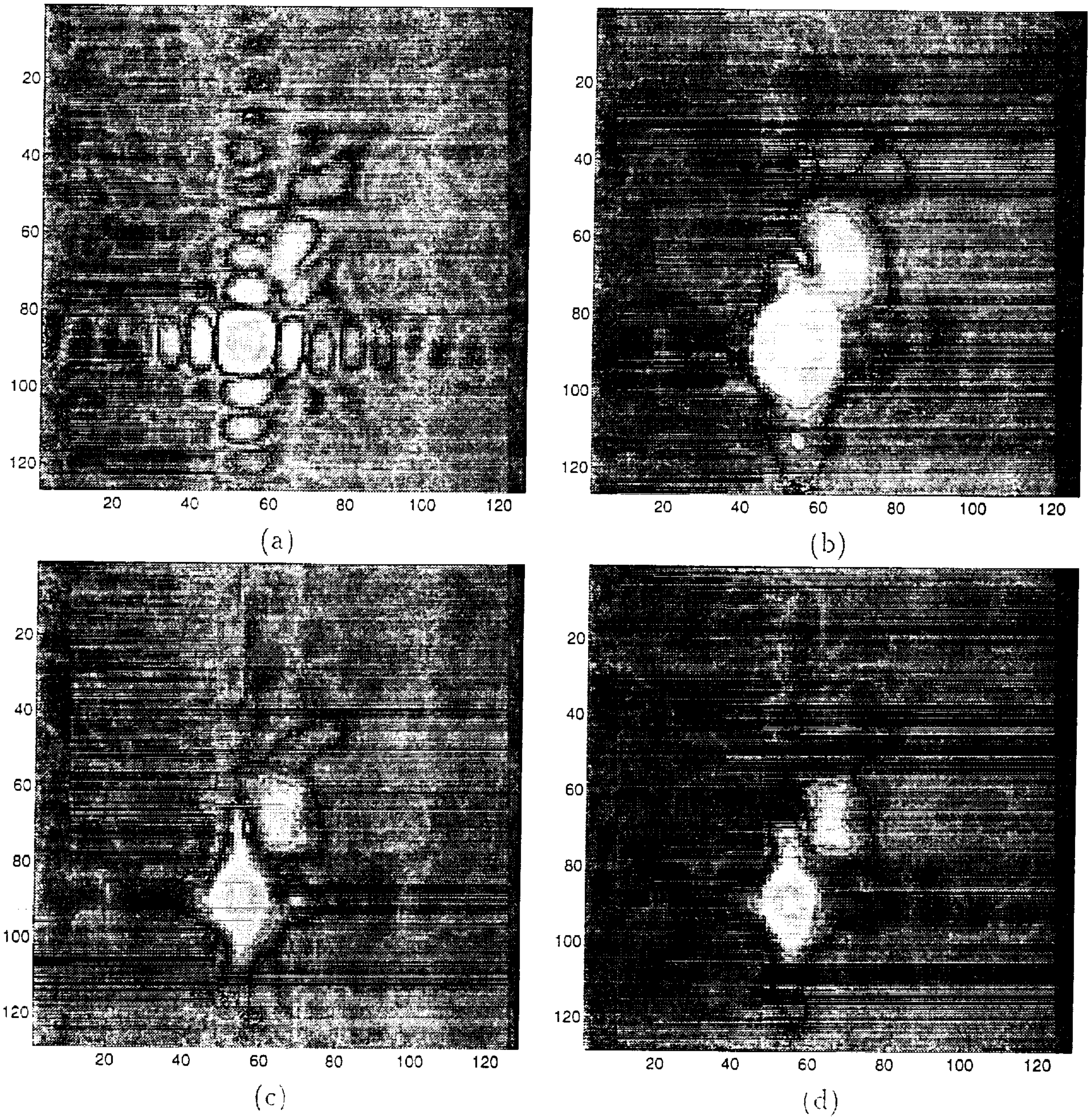


Figure 5.8: SAR images of an XPATCH simulated firetruck when the additive white noise variance is 0.1 and $M_1 = M_2 = 128$. Further, 1/3 of the bandwidth and 1/3 of the synthetic aperture of that in Figure 5.6 are used. (a) 2-D FFT. (b) 2-D FFT with circularly symmetric Kaiser window and shape parameter 6. (c) 1-D APES with $\bar{M} = 16$. (d) 2-D APES with $\bar{M}_1 = \bar{M}_2 = 16$. (e) 2D-MODE-LSF by assuming 121 complex sinusoids.

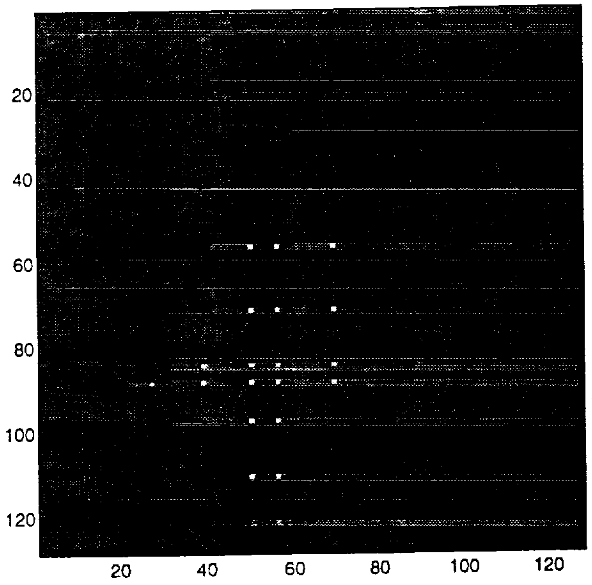


Figure 8: (e)