Air Force Institute of Technology
AFIT Scholar

3-26-2015

Credible Set Estimation, Analysis, and Applications in Synthetic Aperture Radar Canonical Feature Extraction

Andrew C. Rexford

Follow this and additional works at: https://scholar.afit.edu/etd

CREDIBLE SET ESTIMATION, ANALYSIS, AND APPLICATIONS IN
SYNTHETIC APERTURE RADAR CANONICAL FEATURE EXTRACTION

THESIS

Andrew C. Rexford, 1st Lieutenant, USAF

AFIT-ENG-MS-15-M-033

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

CREDIBLE SET ESTIMATION, ANALYSIS, AND APPLICATIONS IN SYNTHETIC

APERTURE RADAR CANONICAL FEATURE EXTRACTION

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Andrew C. Rexford, B.S.E.E.

1st Lieutenant, USAF

March 2015

AFIT-ENG-MS-15-M-033

CREDIBLE SET ESTIMATION, ANALYSIS, AND APPLICATIONS IN SYNTHETIC

APERTURE RADAR CANONICAL FEATURE EXTRACTION

THESIS

Andrew C. Rexford, B.S.E.E.
1st Lieutenant, USAF

Committee Membership:

Dr. Julie Jackson
Chair

Dr. Christine Schubert Kabban
Member

Lt Col Jeremy Stringer, PhD
Member

Abstract

Traditional estimation schemes such as Maximum A Posterior (MAP) or Maximum Likelihood Estimation (MLE) determine the most likely parameter set associated with received signal data. However, traditional schemes do not retain entire posterior distribution, provide no confidence information associated with the final solution, and often rely on simple sampling methods which induce significant errors. Also, traditional schemes perform inadequately when applied to complex signals which often result in multi-modal parameter sets. Credible Set Estimation (CSE) provides a powerful and flexible alternative to traditional estimation schemes.

CSE provides an estimation solution that accurately computes posterior distributions, retains confidence information, and provides a complete set of credible solutions. Determination of a credible region becomes especially important in Synthetic Aperture Radar (SAR) Automated Target Recognition (ATR) problems where signal complexity leads to multiple potential parameter sets. The presented research provides validation of methods for CSE, extension to high dimension/large observation sets, incorporation of Bayesian methods with previous work on SAR canonical feature extraction, and evaluation of the CSE algorithm. The results in this thesis show that: the CSE implementation of Gaussian-Quadrature techniques reduces computational error of the posterior distribution by up to twelve orders of magnitude, the presented formula for computation of the posterior distribution enables numerical evaluation for large observation sets (greater than 7,300 observations), and the algorithm is capable of producing M-th dimensional parameter estimates when applied to SAR canonical features. As such, CSE provides an ideal estimation scheme for radar, communications and other statistical problems where retaining the entire posterior distribution and associated confidence intervals is desirable.

To the gents in RAIL. I cannot think of a better group of people to share the AFIT experience with.

Acknowledgments

I would especially like to thank my wife for her daily support through all of the long days and late nights. Secondly, I would like to thank my advisor, Dr. Jackson, whose experience and guidance enabled me to succeed with this thesis.

Andrew C. Rexford

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol	Definition
$\mathbf{x}$	signal observation vector
$\boldsymbol{\Theta}$	parameter set
σ	variance
N	number of observations
N_l	integration polynomial order
w_i	Gaussian-Quadrature weights
x_i	Gaussian-Quadrature abscissa points
α	normalization constant
μ	mean parameter
M	number of parameters
c	speed of light
f	frequency (in Hz)
σ_R	Rayleigh distribution shape coefficient
$\boldsymbol{\Lambda}$	radar flight path
$S(\boldsymbol{\Lambda}, \boldsymbol{\Theta})$	canonical feature scene response
$P_(\boldsymbol{\Lambda}, \boldsymbol{\Theta})$	canonical feature polarization response
$M_{\Gamma}(\boldsymbol{\Lambda}, \boldsymbol{\Theta})$	canonical feature complex amplitude response
$\Delta R(\boldsymbol{\Lambda}, \boldsymbol{\Theta})$	canonical feature differential range model
VV	vertical-vertical polarization
HH	horizontal-horizontal polarization
VH	vertical-horizontal polarization
ϕ	flight path azimuth
θ	flight path elevation

Symbol	Definition
A	canonical feature amplitude scale term
X	position on the x-axis
Y	position on the Y-axis
Z	position on the z-axis
L	canonical feature length
H	canonical feature height
r	canonical feature radius

List of Acronyms

Acronym	Definition
3-D	three-dimensional
AFRL	Air Force Research Laboratory
ATR	automatic target recognition
CDF	cumulative distribution function
CSE	Credible Set Estimation
CUDA	Compute Unified Device Architecture
CV	civilian vehicles
DC	direct current
DoD	Department of Defense
DRSC	DoD Supercomputing Resource Center
GPU	graphics processing unit
HPC	high performance computing
HPF	high precision float
LOS	line of sight
MAE	mean absolute error
MAP	maximum a posteriori
MLE	maximum likelihood estimation
MSE	mean squared error
PH	phase history
PDF	probability distribution function
PMF	probability mass function
RAM	random access memory
SAR	synthetic aperture radar

Acronym	Definition
SPLIT	spectrum parted linked image test
TAE	total absolute error
WGN	white Gaussian noise

CREDIBLE SET ESTIMATION, ANALYSIS, AND APPLICATIONS IN SYNTHETIC
APERTURE RADAR CANONICAL FEATURE EXTRACTION

I.   Introduction

Traditional estimation schemes apply maximum a posteriori (MAP) or maximum
likelihood estimation (MLE) to determine the most likely parameter set associated
with received signal data. However, traditional schemes do not retain entire posterior
distribution, provide no confidence information associated with the final solution and
often rely on simple sampling methods which induce significant errors. It is desirable
to implement an estimation scheme that can accurately compute posterior distributions
associated with various signal parameters. The synthetic aperture radar (SAR) automatic
target recognition (ATR) community has interest in knowing the probabilities associated
with the parameters of various feature models [1]. The presented research builds on
the ability to extract probability distributions [2] by covering four areas: validation of
methods for credible set estimation, extension to large dimension/large observation sets,
integration of Bayesian methods with previous work on SAR canonical feature extraction,
and evaluation of the Credible Set Estimation (CSE) algorithm.

1.1   Problem Statement

It is possible to apply CSE to accurately compute posterior distributions associated
with various signal parameters [3]. CSE computes the entire posterior distribution using
Gauss-Quadrature integration, resulting in a high accuracy estimator that retains confidence
information associated with MAP estimates.

Previous work on CSE is limited to small observation vectors (less than 200 observations) and two parameters [2]. It is desirable to extract posterior distributions for features with more parameters and observations; however, numerical precision issues prevent direct computation of the posterior distribution equation. Also, little validation has been performed for the CSE method. Finally, it is desirable to present results for the CSE when applied to more complex SAR data which contains a realistic target. For this thesis, CSE is implemented to demonstrate estimator performance in three scenarios:

1. Simple single parameter estimation,

2. Simple direct current (DC) and sinusoidal signals with two parameters,

3. Complex canonical features with up to five parameters.

1.2   Research Goals, Methodology and Results

This thesis encompasses four areas: validation, extension, incorporation and evaluation. Validation is accomplished by examining methods presented by Rademacher [2] and testing the approach on simple DC and sinusoidal signals. The results of validation are presented in an error analysis study by comparing the CSE method against traditional probability distribution function (PDF) sampling methods. The algorithm presented by Rademacher is extended to higher order data sets with large numbers of observations. Extension to data sets with large observation vectors is accomplished by presenting a formula for the numerical evaluation of the posterior distributions that overcomes numerical precision limitations associated with direct computation of the distribution. Incorporation is accomplished by applying the SAR canonical feature extraction fundamentals presented by Crosser [4] to the algorithm presented by Rademacher, specifically by implementing Crosser's flight path selection techniques to the CSE. Finally, evaluation of the new estimator is performed on simple SAR scenes comprised of 3-D canonical features.

The results in this thesis show the following:

- Validation: application of Gauss-Quadrature numerical integration techniques to Bayesian posterior distribution computation is more effective at providing accurate probability mass function (PMF)s than traditional sampling techniques. The presented method is also more effective at capturing probability, with fewer missed regions, than traditional sampling techniques.

- Extension: The presented formula for the computation of the Bayesian posterior distribution enables numerical evaluation for large observation sets (greater than 6000 observations). Extremely large observation sets will require higher-order polynomials for Gauss-Quadrature integration, resulting in exceptional computational requirements. This thesis presents an equation for the computational complexity in Chapter 4. The computational complexity can be effectively controlled by varying grid interval, integration polynomial and the search space for each parameter.

- Incorporation: The CSE method performs well for "straight-and-level" flight paths. However, adequate flight path selection provides more reliable parameter estimation (within the limitations of SAR) than arbitrarily selected flight paths.

- Evaluation: CSE is able to extract posterior distributions for features with up to five parameters. This thesis presents formulas for the computational complexity required when implementing the CSE. Estimation of features with six or more parameters requires splitting each large distribution into multiple smaller distributions to avoid large computational overhead.

1.3 Thesis Organization

This thesis is organized in the following manner:

- Chapter 2: presents background material from a comprehensive literature review. The material includes information on Bayesian methods, numerical methods, Gaussian-Quadrature integration methods, 3-D canonical features, radar phase history data, flight path selection, and high performance computing (HPC).

- Chapter 3: presents the results of validation and extension. A formula is derived for numerical evaluation of posterior distributions with large observation sets and an updated process for the CSE. The updated formula is then used to validate the CSE method when compared to MAP estimators that utilize simple sampling methods.

- Chapter 4: presents the results of incorporation and validation. The work presented in Chapter 3 is extended to SAR scenes containing 3-D canonical features. The algorithm integrates flight path selection techniques presented by Crosser [4]. The basic CSE method is applied to SAR phase history data to provide posterior distributions for up to five parameters. The results presented in Chapter 4 are considered a simulation since the CSE is applied to test scenarios containing only simple shapes.

- Chapter 5: presents conclusions and future work.

1.4 Key Assumptions

Applicable key assumptions will be presented at the start of each chapter. However, the following list details key assumptions applicable to the research as a whole:

- Received signals contain only a signal of interest and additive white Gaussian noise. This assumption is required by application of the univariate normal distribution as the likelihood function.

- Parameters are considered to be independent. This assumption is required in order to initialize prior distributions without a covariance matrix. Also, the parameters

4

used for three-dimensional (3-D) canonical features (size, location, orientation) are independent by definition.

- Prior distributions are finitely bounded. This assumption makes numerically computing the denominator in the posterior distribution possible.

- Shape type is known a priori. The shape type must be known in advance so the CSE can apply the appropriate signal model based on Jackson's 3-D canonical features [4, 5].

II. Background

This chapter presents background material from a comprehensive literature review. The material includes information on Bayesian methods, numerical methods, Gauss-Quadrature integration methods, 3-D canonical features, radar phase history data, flight path selection, and HPC.

2.1   Bayesian Methods

The goal of the CSE is to provide parameter estimates of the posterior distribution for canonical features. Typical likelihood functions provide the probability that random signal, $\mathbf{x}$, was received given a known set of parameters, $\boldsymbol{\Theta}$. For a signal in additive white Gaussian noise, the likelihood function for $\mathbf{x}$ becomes

$$p(\mathbf{x}|\boldsymbol{\Theta}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n] - f_n(\boldsymbol{\Theta}))^2\right], \tag{2.1}$$

where σ is the variance associated with the noise power, N is the number of observations, $x[n]$ is the received signal, and $f_n(\boldsymbol{\Theta})$ is a signal model based on a parameter set, $\boldsymbol{\Theta}$. One way to estimate the signal parameters $\boldsymbol{\Theta}$ from measurements $\mathbf{x}$ is to find the maximum likelihood estimate: the $\boldsymbol{\Theta}$ that maximizes (2.1) for a given $\mathbf{x}$. However, for ATR, it is informative and useful to find the probability of a set of parameters, based on a received signal. Determining $p(\boldsymbol{\Theta}|\mathbf{x})$ requires the application of Bayes' Rule, as well as prior knowledge of the parameters of interest [6]. Bayes' Rule is defined as

$$p(\boldsymbol{\Theta}|\mathbf{x}) = \frac{p(\mathbf{x}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})}{\int p(\mathbf{x}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})d\boldsymbol{\Theta}}, \tag{2.2}$$

where $p(\boldsymbol{\Theta})$ is the prior distribution. Equation (2.2) provides a solution for the posterior distribution based on unknown parameters, $\boldsymbol{\Theta}$ and a known signal $\mathbf{x}$. It is important to note that the likelihood function as well as the posterior distribution presented above are continuous functions. Since there is often no closed form solution to the

posterior distribution, the posterior distribution will be discretized using Gauss-Quadrature techniques and estimated as a PMF throughout the presented research.

2.1.1 Maximum a Posteriori Estimation.

The MAP estimation approach chooses a parameter set, $\hat{\mathbf{\Theta}}$ that maximizes the posterior distribution, $p(\mathbf{\Theta}|\mathbf{x})$ [6]. The MAP estimate is defined as

$$\hat{\mathbf{\Theta}} = \arg \max_{\mathbf{\Theta}} p(\mathbf{\Theta}|\mathbf{x}). \tag{2.3}$$

In many applications, the MAP estimator is implemented to find the maximum argument of the log-likelihood function, defined as

$$\hat{\mathbf{\Theta}} = \arg \max_{\mathbf{\Theta}} \log p(\mathbf{\Theta}|\mathbf{x}). \tag{2.4}$$

Applying the MAP estimator as a function of log-likelihood significantly reduces the numerical precision required [6]. Typically, any leading constants for $p(\mathbf{\Theta}|\mathbf{x})$ are dropped to reduce precision requirements further. Equation (2.4) is implemented via a grid search over a range of parameters, $\mathbf{\Theta}$. The sampled distribution is then either normalized (such that the distribution sums to one), or left without normalization. The argument of the maximum is simply selected, resulting in the MAP estimate of the parameter set.

2.2 Gaussian Quadrature Integration Methods

Methods presented by Rademacher [2] implement the posterior distribution function as a PMF, as opposed to sampling the PDF and normalizing. In order to compute the PMF, each sub-region in the posterior PDF must be integrated to determine the total probability in that region. Gaussian Quadrature numerical integration techniques can be employed to compute the integral for each sub-region [7–10]. Consider the finitely bounded integral

$$g(x) = \int_a^b f(x)w(x)dx, \tag{2.5}$$

where $f(x)$ is the function to be integrated and $w(x)$ is a weighting function. Using the specific case where the weight function is the unity, $w(x) = 1$, the integral in Equation (2.5)

can be estimated as

$$g(x) \approx \sum_{i=1}^{N_l} w_i f(x_i),$$ (2.6)

where w_i are individual weights, x_i are abscissa points, and i is the order of the integration polynomial. In the specific case presented above, the abscissa points as well as the weights are defined according to the desired polynomial order N_l. Rademacher presents the following process for setting up the numerical integration method [2]:

1. Decide which orthogonal polynomial is most appropriate based on weight function, $w(x)$. In this thesis, we assume a legendre polynomial associated with the weighting function, $w(x) = 1$.

2. Decide on polynomial order N_l.

3. Compute the roots of the N_l-th order orthogonal polynomial. These are the abscissa points x_i.

4. Compute the weight coefficients, w_i.

5. Compute the value of the function at the N abscissa points.

6. Multiply each weight, w_i, with the function value $f(x_i)$ to result in N weighted coefficients.

7. Sum the weighted coefficients together. The result is the integral approximation.

In this thesis, Gauss-Quadrature integration is implemented as part of the CSE posterior distribution estimation scheme. Work by Rademacher [2] provided limited results using the quadrature integration method. This thesis provides validation of the method by comparing the quadrature method against a truth using a distribution with a closed-form integral. Further, the concept is extended to M-dimensions.

2.3   Credible Set Estimation

CSE relies on Bayesian methods in order to compute the posterior distribution of a set of parameters given a received signal [2]. For a signal with additive white Gaussian noise (WGN), the resulting likelihood function is the univariate normal distribution [6], shown in Equation (2.1). The "true" signal is based upon a model of the ideal signal. Classical estimators, such as MAP, seek to extract one, and only one "true" value [6]. MAP estimation is typically accomplished by evaluating the posterior distribution across a set of sampled parameter values, and determining the maximum value. In applications such as ATR, it is more desirable to determine the probability of a parameter *set* given an observed signal, $p(\Theta|\mathbf{x})$ [1]. Bayes' Rule (2.2) is applied to determine the desired posterior distribution.

Retaining the entire posterior distribution is desirable, as often times the same signal can be generated using various sets of parameters (estimating the phase in a sinusoid for example). The result from retaining the posterior distribution is multiple, distinct regions of the likelihood that would be masked if a MAP were used. This thesis extends CSE to multiple dimensions, large numbers of observations and provides validation data to support the implementation.

CSE returns the posterior distribution estimation of Equation (2.1), as well as the "credible region", which contains a user-defined amount of probability mass. The computation for the posterior distribution is then iterated, using finer search intervals until a user-defined search interval is reached. Figure 2.1 contains a flowchart depicting the generalized estimator.
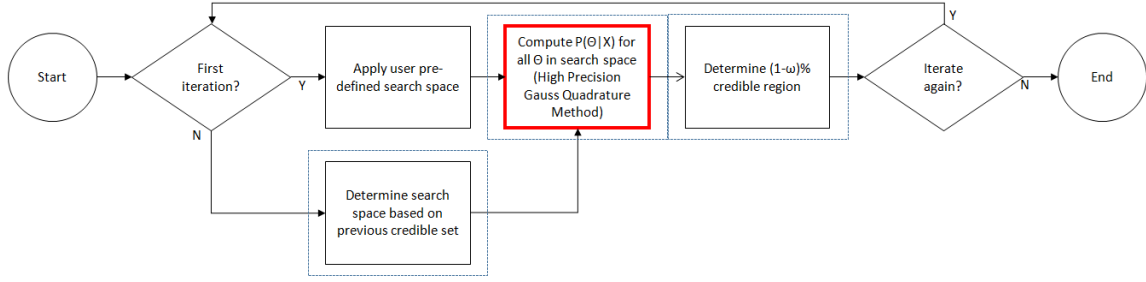
Figure 2.1: Flowchart depicting the CSE implementation. Red box indicates parallel process. Dashed blue box indicates updates to Rademacher's algorithm presented in this thesis.

2.4   3-D Canonical Feature Models

To illustrate application of the CSE on complex signals, we select 3-D canonical features as defined by Jackson [11]. Jackson defines models associated with six canonical shapes: plate, dihedral, trihedral, sphere, cylinder and top-hat, as shown in Figure 2.2. The canonical feature model defines each shape in terms of a polarization response, $P_\beta(\Lambda, \Theta)$, complex amplitude response, $M_\Gamma(\Lambda, \Theta)$, and a differential range model, $\Delta R(\Lambda, \Theta)$. These models are defined according to the radar flightpath, Λ, and parameter set, Θ. The overall scene response can be modeled as a sum of all scatterers

$$S(\Lambda, \Theta) = \sum_m P_\beta(\Lambda, \Theta) M_\Gamma(\Lambda, \Theta) e^{jk\Delta R(\Lambda, \Theta)}. \tag{2.7}$$

Each feature is defined by a parameter set (Θ) consisting of an amplitude scale factor (A), shape location (X, Y, Z), roll, pitch, and yaw (γ, θ, ϕ), and length, height and radius (L, H, r). A vector (Λ) contains the azimuth and elevation angles throughout the flight path. Only a portion of the parameters are applicable to each shape due to the unique geometries, as shown in Table 2.1.

Table 2.2 presents the response definitions for each shape. Note that the models are defined only within the azimuth (θ) and elevation (ϕ) ranges depicted. The differential
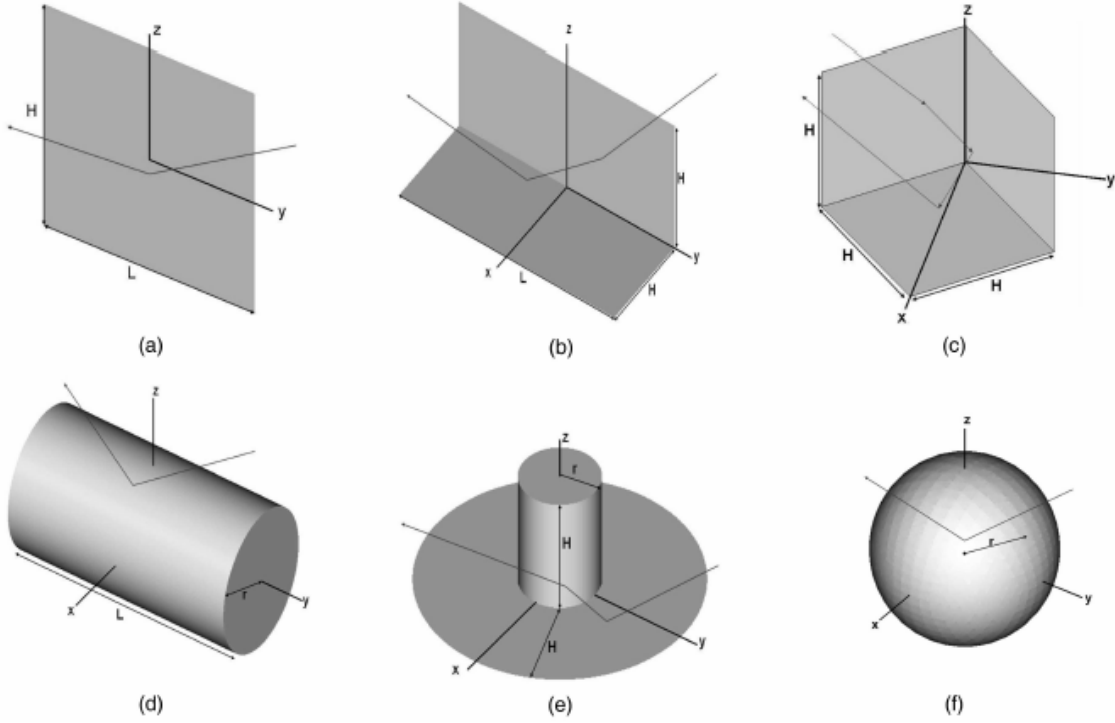
Figure 2.2: Six canonical features as defined by Jackson [5]. The shapes are a (a) plate, (b) dihedral, (c) trihedral, (d) cylinder, (e) top-hat, and (f) sphere.

Table 2.1: Applicable parameters for each shape type [4, 5].

Shape	X	Y	Z	L	H	r	γ	θ	ϕ
plate	X	X	X	X	X		X	X	X
dihedral	X	X	X	X	X		X	X	X
trihedral	X	X	X		X		X	X	X
sphere	X	X	X			X			
cylinder	X	X	X	X		X	X	X	X
top-hat	X	X	X		X	X	X	X	X

range models for each shape are shown in Table 2.3. The polarization response matrix is defined as [12]

$$P(\Lambda; \Theta) = \begin{bmatrix} p_{vv} & p_{vh} \\ p_{hv} & p_{hh} \end{bmatrix}. \tag{2.8}$$

11

The polarization responses for odd and even bounce are defined depending on object rotation around the receiver line of sight (LOS), ζ_r and rotation about the transmitter LOS, ζ_t. The polarization response models are defined as

$$P_{odd} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$
(2.9)

$$P_{even} = \begin{bmatrix} -\cos(\zeta_t + \zeta_r) & \sin(\zeta_t + \zeta_r) \\ \sin(\zeta_t + \zeta_r) & \cos(\zeta_t + \zeta_r) \end{bmatrix}.$$
(2.10)

Table 2.2: Odd or even polarization bounces and amplitude response for each shape type [11]. The magnitude of the amplitude is scaled by the shape size and is given with the parameter A. Note that $\tilde{\phi}$ and $\tilde{\theta}$ are azimuth and elevation angles relative to the reference orientation of each shape (see Figure 2.2) and k is the wave number, defined as $k = \frac{2\pi f}{c}$ where c is the speed of light and f is the transmitting frequency.

Shape	Polarization Bounce	$\mathbf{M_{\Gamma m}(k, \Lambda; \Theta_m)}$	$\mathbf{A}$
plate	odd	$\frac{jk}{\sqrt{\pi}} A \mathrm{sinc}(kL \sin\tilde{\phi}\cos\tilde{\theta})\mathrm{sinc}(kH\sin\tilde{\theta});\ \tilde{\theta} \in [-\frac{\pi}{2}, \frac{\pi}{2}], \tilde{\phi} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$	LH
dihedral	even	$\frac{jk}{\sqrt{\pi}} A \mathrm{sinc}(kL\sin\tilde{\phi}\cos\tilde{\theta}) \times \begin{cases} \sin\tilde{\theta};\ \tilde{\theta} \in [0,\frac{\pi}{4}] \\ \cos\tilde{\theta};\ \tilde{\theta} \in [\frac{\pi}{4},\frac{\pi}{2}] \end{cases};\ \tilde{\phi} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$	2LH
trihedral	odd	$\frac{jk}{\sqrt{\pi}} A \times \begin{cases} \sin(\tilde{\theta}+\frac{\pi}{4}-\tan^{-1}(\frac{1}{\sqrt{2}}));\ \tilde{\theta} \in [0,\tan^{-1}(\frac{1}{\sqrt{2}})] \\ \cos(\tilde{\theta}+\frac{\pi}{4}-\tan^{-1}(\frac{1}{\sqrt{2}}));\ \tilde{\theta} \in [\tan^{-1}(\frac{1}{\sqrt{2}}),\frac{\pi}{2}] \end{cases}$ $\times \begin{cases} -\cos(\tilde{\phi}-\frac{\pi}{4});\ \tilde{\phi} \in [-\frac{\pi}{4},0] \\ \sin(\tilde{\phi}-\frac{\pi}{4});\ \tilde{\phi} \in [0,\frac{\pi}{4}] \end{cases}$	$2\sqrt{3}H^2$
sphere	odd	$A\sqrt{\pi};\ \tilde{\theta} \in [-\frac{\pi}{2}, \frac{\pi}{2}], \tilde{\phi} \in [-\pi, \pi]$	r
cylinder	odd	$A\sqrt{jk\cos\tilde{\phi}}\ \mathrm{sinc}(kL\sin\tilde{\phi}\cos\tilde{\theta});\ \tilde{\theta} \in [-\frac{\pi}{2}, \frac{\pi}{2}], \tilde{\phi} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$	$L\sqrt{r}$
top-hat	even	$A\sqrt{jk} \times \begin{cases} \sin\tilde{\theta};\ \tilde{\theta} \in [0,\frac{\pi}{4}] \\ \cos\tilde{\theta};\ \tilde{\theta} \in [\frac{\pi}{4},\frac{\pi}{2}] \end{cases};\ \tilde{\phi} \in [-\pi, \pi]$	$\sqrt{\frac{8r}{\sqrt{2}}}H$

Table 2.3: 3-D scatterer differential range models. ΔR_0 represents the range to the scattering center location. ϕ_t and θ_t are the bistatic azimuth and elevation angles at the transmitter, and ϕ_r and θ_r are the bistatic azimuth and elevation angles at the receiver.

	ΔR
plate	R_0
dihedral	R_0
trihedral	R_0
sphere	$\Delta R_0 + \left(r(\cos\theta_t + \cos\theta_r)\cos\left(\frac{\phi_t-\phi_r}{2}\right)\cos\left(\frac{\theta_t+\theta_r}{2}\right) + r\sin\left(\frac{\theta_t+\theta_r}{2}\right)(\sin\theta_t + \sin\theta_r) \right)$
cylinder	$\Delta R_0 + \left(r\cos\left(\frac{\theta_t-\theta_r}{2}\right)(\cos\phi_t + \cos\phi_r) \right)$
top-hat	$\Delta R_0 + \left(r\cos\left(\frac{\phi_t-\phi_r}{2}\right)(\cos\theta_t + \cos\theta_r) \right)$

2.5 Canonical Feature Extraction

For the simulation portion of this thesis, the phase history (PH) returns observed from an airborne radar are generated according to the equations presented for Jackson's canonical models [5]. To this point, extracting 3-D canonical features has been performed using a gradient descent [5] and dictionary search methods [4, 13]. Hammond [13] initially performed the search using a coarsely sampled dictionary, in which every possible combination of parameters was used to generate test PH. This test data was then compared against the true data, and the dictionary entry with the highest correlation was selected as the true value.

Follow-on work by Crosser [4] developed a more complete process which provided a detailed algorithm to extract Jackson's 3-D canonical features. Crosser's process, depicted in Figure 2.3 included three primary steps: Radar Data Collection and Image formation, Parameter Estimation and Bounding, and Dictionary Formation and Search. The final step, dictionary formation and search, will be replaced by the methods in this thesis, and as such will not be discussed here.
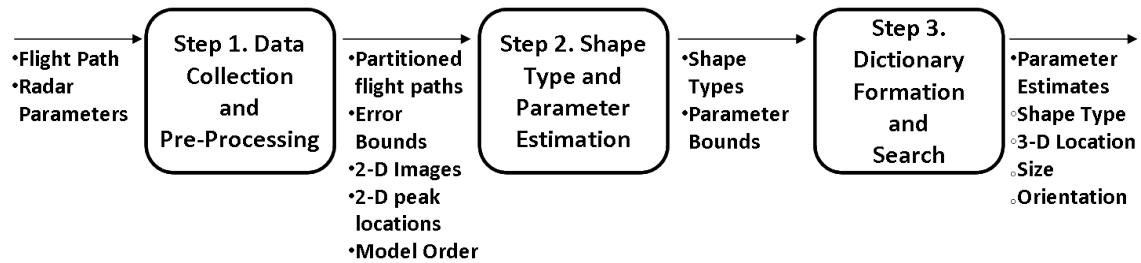
Figure 2.3: Canonical feature extraction algorithm including steps and associated inputs/outputs [4].

2.5.1 Flight Path Selection.

Crosser presented a method to determine adequate flight paths for 3-D canonical feature extraction [4]. The method produces flight paths that consider two key aspects of feature extraction. The first aspect is specular response for the plate, dihedral and cylinder. Because of the narrow specular response, flight paths must overlap. The second aspect is radius estimation. Crosser demonstrates that the radius can only be estimated with flight paths that include diversity in both grazing angle and tilt angle [4]. Crosser presents three flight paths:

- Up-Down: provides sufficient specular diversity to be used with the plate, dihedral, trihedral.

- Disjoint: provides grazing and tilt angle diversity to be used with the sphere and top-hat.

- Two-Pass: balances specular diversity and grazing/tilt angle diversity to be used with the cylinder.

Each flight path consists of two radar passes. Figure 2.4 depict the flight paths as recommended by Crosser [4]. For this thesis, CSE is implemented using Crosser's flight paths.

14

(a) Up-down flight path.

(b) Disjoint flight path.

(c) Two-pass flight path.
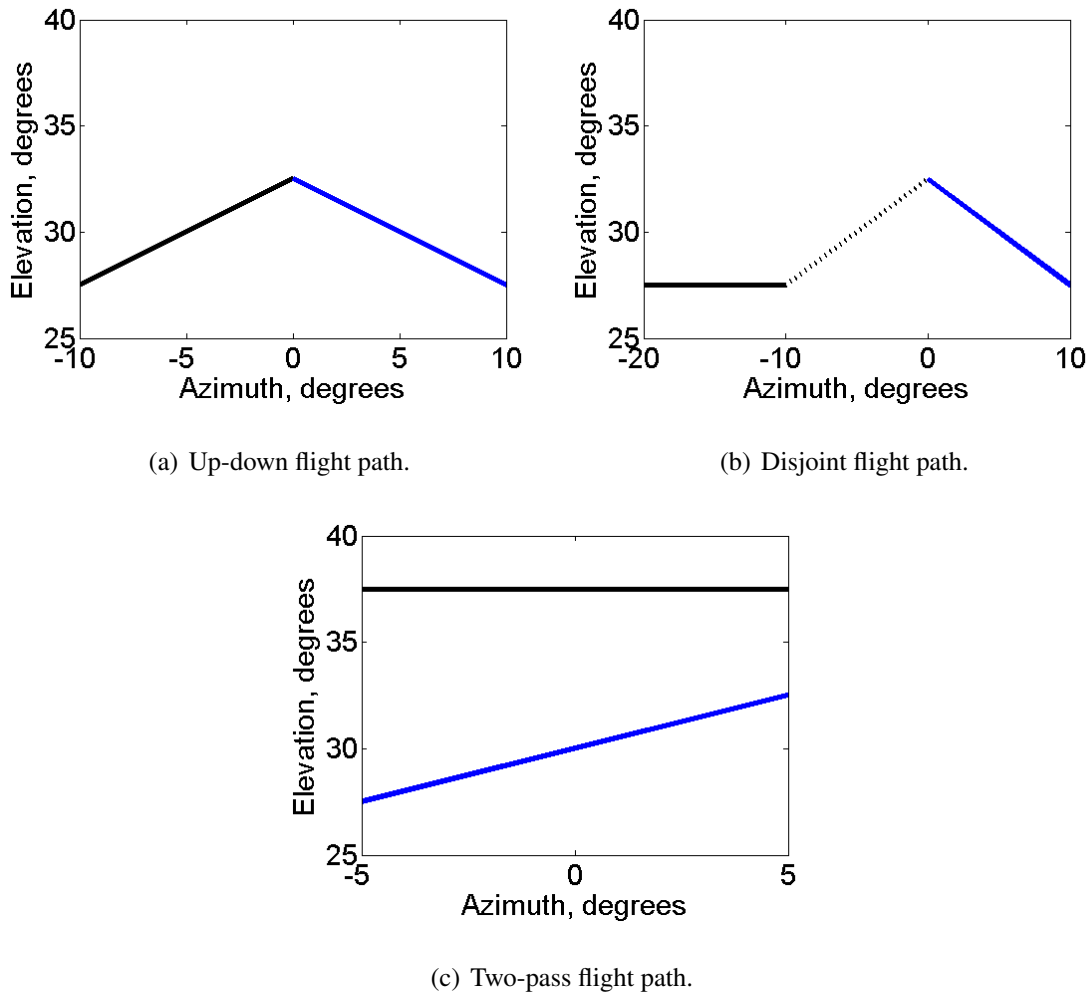
Figure 2.4: Three flight paths shown in azimuth and elevation [4].

2.6 High Performance Computing

Previous work by Rademacher [2] utilized Compute Unified Device Architecture (CUDA) to execute the CSE algorithm on high powered graphics processing unit (GPU). However, extending the CSE to large observation sets and higher dimensions significantly increases memory requirements beyond what current graphics cards are capable of providing. Also, implementing the CSE on a four-core, 3.0 GHz Intel i7 workstation with 32 GB of random access memory (RAM), to perform five parameter estimates using a signal with 6000 observations, and a total of 50 sampling points along each parameter

takes approximately 72 hours. As a result, it is impractical to implement the CSE to run on a standard workstation.

The presented research utilizes Department of Defense (DoD) high performance computing (HPC) modernization program resources, specifically the Air Force Research Laboratory (AFRL) DoD Supercomputing Resource Center (DRSC) Spirit system. Table 2.4 outlines the capabilities of the system [14]. The CSE is implemented in Chapter 4 to execute calculations for each parameter permutation in parallel on the Spirit system, resulting in a computation time of about six hours for the scenario described above.

Table 2.4: Capabilities of the AFRL DRSC Spirit system [14].

	Login Nodes	Compute Nodes
Total Cores — Nodes	128 — 8	73440 — 4590
Cores/Node	16	16
Core Type	Intel Xeon Sandy Bridge	Intel Xeon Sandy Bridge
Core Speed	2.6 GHz	2.6 GHz
Memory/Node	64 GB	32 GB
Accessible Memory/Node	62 GB	30 GB

III. Theory

The intent of this chapter is to validate the estimation model presented by Rademacher [2]. A generalized method, including extraction of PMFs for both single and multiple parameters is explored. Validation of the method is performed by comparing traditional sampling methods and Gaussian quadrature integration. Both graphical and tabular results are presented.

3.1 Key Assumptions and Implementation

The following key assumptions apply to the method presented in this chapter:

- Received signals contain only a signal of interest and additive white Gaussian noise. This assumption is required by application of the univariate normal distribution as the likelihood function.

- Parameters are considered to be independent. This assumption is required in order to initialize prior distributions without a covariance matrix.

- Prior distributions are finitely bounded. This assumption makes numerically computing the denominator in the posterior distribution possible.

3.2 Numerical PMF Computation

Directly evaluating Equation (2.1) presents significant numerical precision issues. For large observation vectors, the exponential term quickly grows, even for small values in the error term $(x[n] - f_n(\Theta))^2$. As a result, the exponential term can quickly exceed the MATLAB limit of e^{710}. To avoid precision issues, it is possible to perform all computations as the log of the PMF. The Log-Sum-Exponential method can then be applied to create an equation MATLAB can compute without precision issues. The equation for the posterior

distribution for a signal in white Gaussian noise is

$$
\begin{aligned}
p(\mathbf{\Theta}|\mathbf{x}) &= \frac{p(X|\mathbf{\Theta})p(\mathbf{\Theta})}{\int p(X|\mathbf{\Theta})p(\mathbf{\Theta})d\mathbf{\Theta}}, \\
&= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})}{\int \frac{1}{(2\pi\sigma^2)^{N/2}} \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}}.
\end{aligned}
\tag{3.1}
$$

The denominator in Equation (3.1) acts as a normalization term, ensuring that the PDF integrates to one. However, for large observations N, MATLAB and other computational tools may compute the denominator to be zero, resulting in divide by zero errors. The numerical precision required can be significantly reduced by removing the leading constant, which may be very large or very small depending on N and σ. To avoid precision issues, Equation (3.1) is rewritten as

$$
\begin{aligned}
p(\mathbf{\Theta}|\mathbf{x}) &= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})}{\int \frac{1}{(2\pi\sigma^2)^{N/2}} \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}}, \\
&= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})}{\frac{1}{(2\pi\sigma^2)^{N/2}} \int \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}}, \\
&= \frac{\exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})}{\int \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}}.
\end{aligned}
\tag{3.2}
$$

It is now possible to employ additional numerical computation techniques. Recall, we will be integrating sub-regions using the Gauss Quadrature rule, shown as

$$
p(a \leq \mathbf{\Theta} \leq b|\mathbf{x}) = \int_{a}^{b} \left[\frac{\exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})}{\int \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}} \right] d\mathbf{\Theta}.
\tag{3.3}
$$

The denominator term is a constant, so it can be removed from the integral, shown as

$$
p(a \leq \mathbf{\Theta} \leq b|\mathbf{x}) = \frac{\int_{a}^{b} \left[\exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta}) \right] d\mathbf{\Theta}}{\int \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}}.
\tag{3.4}
$$

Now the Gauss Quadrature method can be applied to the numerator, shown as

$$
p(a \leq \mathbf{\Theta} \leq b|\mathbf{x}) \approx \frac{\sum_{i}^{N_l} \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}_i))^2}{2\sigma^2})p(\mathbf{\Theta}_i)w_i}{\int \exp(-\frac{\sum_{n=1}^{N}(x[n]-f(\mathbf{\Theta}))^2}{2\sigma^2})p(\mathbf{\Theta})d\mathbf{\Theta}}.
\tag{3.5}
$$

Computing Equation (3.5) directly induces large precision errors as the denominator term will evaluate to very small numbers with large observations N, resulting in divide by zero errors. To avoid precision issues, the Log-Sum-Exponential method can be applied. The method works by determining the common term in a sum of exponentials and then computing the remaining terms, for example

$$
\begin{aligned}
f(x) &= \log(e^{-5000} + e^{-5001}), \\
&= \log(e^{-}5000(e^0 + e^{-1})), \\
&= -5000 + \log(1 + e^{-1}).
\end{aligned} \tag{3.6}
$$

To apply the log-sum-exponential method, Equation (3.5) must be rearranged such that the prior distribution and the weighting function are included in the exponential term, shown as

$$
\begin{aligned}
\log p(a \leq \boldsymbol{\Theta} \leq b | \mathbf{x}) &\approx \log \left[\frac{\sum_i^{N_l} \exp(-\frac{\sum_{n=1}^N (x[n]-f(\boldsymbol{\Theta}_i))^2}{2\sigma^2}) p(\boldsymbol{\Theta}_i) w_i}{\int \exp(-\frac{\sum_{n=1}^N (x[n]-f(\boldsymbol{\Theta}))^2}{2\sigma^2}) p(\boldsymbol{\Theta}) d\boldsymbol{\Theta}} \right], \\
&\approx \log \left[\sum_i^{N_l} \exp\left(-\frac{\sum_{n=1}^N (x[n]-f(\boldsymbol{\Theta}_i))^2}{2\sigma^2} \right) p(\boldsymbol{\Theta}_i) w_i \right] \\
&\quad - \log \left[\int \exp\left(-\frac{\sum_{n=1}^N (x[n]-f(\boldsymbol{\Theta}))^2}{2\sigma^2} \right) p(\boldsymbol{\Theta}) d\boldsymbol{\Theta} \right].
\end{aligned} \tag{3.7}
$$

Notice that Equation (3.7) does not exactly match the form in the Log-Sum-Exponential methods. Applying the equalities $p(\boldsymbol{\Theta}_i) = \exp(\log(p(\boldsymbol{\Theta}_i)))$ and $w_i = \exp(\log(w_i))$ results in

$$
\begin{aligned}
\log \text{numerator}_{\text{term}} &\approx \log \left[\sum_i^{N_l} \exp\left(-\frac{\sum_{n=1}^N (x[n]-f(\boldsymbol{\Theta}_i))^2}{2\sigma^2} \right) \exp(\log(p(\boldsymbol{\Theta}_i))) \exp(\log(w_i)) \right], \\
&\approx \log \left[\sum_i^{N_l} \exp\left(-\frac{\sum_{n=1}^N (x[n]-f(\boldsymbol{\Theta}_i))^2}{2\sigma^2} + \log(p(\boldsymbol{\Theta}_i)) + \log(w_i) \right) \right].
\end{aligned} \tag{3.8}
$$

Numerically, the exponential in Equation (3.8) is computed as

$$
\text{exponential}_{\text{term}} = \exp\left[\begin{bmatrix} \left(\frac{-\sum_{n=0}^{N-1}(x[n]-f(\Theta_1))^2}{2\sigma^2} + \log(p(\Theta_1)) + \log(w_1) - \alpha\right) + \alpha \\ \left(\frac{-\sum_{n=0}^{N-1}(x[n]-f(\Theta_2))^2}{2\sigma^2} + \log(p(\Theta_2)) + \log(w_2) - \alpha\right) + \alpha \\ \vdots \end{bmatrix}\right],
$$

$$
= \exp(\alpha)\exp\left[\begin{bmatrix} \left(\frac{-\sum_{n=1}^{N}(x[n]-f(\Theta_1))^2}{2\sigma^2} + \log(p(\Theta_1)) + \log(w_1) - \alpha\right) \\ \left(\frac{-\sum_{n=1}^{N}(x[n]-f(\Theta_2))^2}{2\sigma^2} + \log(p(\Theta_2)) + \log(w_2) - \alpha\right) \\ \vdots \end{bmatrix}\right], \qquad (3.9)
$$

where α is a common factor between entries. Equation (3.9) is now computed in MATLAB as

$$
\log\text{numerator}_{\text{t}} \approx \log\left[\sum_i^{N_l}\exp(\alpha)\exp\left(\frac{-\sum_{n=1}^{N}(x[n]-f(\Theta_i))^2}{2\sigma^2} + \log(p(\Theta_i)) + \log(w_i) - \alpha\right)\right],
$$

$$
\approx \log\left[\exp(\alpha)\sum_i^{N_l}\exp\left(\frac{-\sum_{n=1}^{N}(x[n]-f(\Theta_i))^2}{2\sigma^2} + \log(p(\Theta_i)) + \log(w_i) - \alpha\right)\right],
$$

$$
\approx \alpha + \log\left[\sum_i^{N_l}\exp\left(\frac{-\sum_{n=1}^{N}(x[n]-f(\Theta_i))^2}{2\sigma^2} + \log(p(\Theta_i)) + \log(w_i) - \alpha\right)\right].
$$

$$(3.10)$$

By pulling the common factor, α, out of the exponential, MATLAB is able to compute each term without overflowing. This makes computing the PMF possible for large numbers of observations. At this point, the numerator term can be computed for each parameter set, and stored in log form. Computation of the denominator term follows the same pattern. Notice that the denominator is just the integral of the original numerator term. If the quadrature rule is applied over sub-regions of the entire integral (which we assume in our implementation), then it is possible to convert the denominator to a summation, which will approximately represent the integral over the entire region. Keep in mind that the regions were calculated above as the log of the numerator. Therefore the exponential of each numerator region must be computed. Here, we apply the quadrature rule, the identities in Equation (3.8) and the Log-Sum-Exponential method to obtain the denominator term,

shown as

$$\text{denominator}_{\text{term}} \approx \log \sum_{ALL} \exp \left[\log \sum_{i} \exp \left(\frac{- \sum_{n=1}^{N}(x[n] - f(\mathbf{\Theta}_i))^2}{2\sigma^2} + \log(p(\mathbf{\Theta}_i)) + \log(w_i) \right) \right],$$

$$\approx \log \sum_{ALL}^{N_l} \exp \left[\log \text{numerator}_{\text{terms}} \right]. \tag{3.11}$$

Note that $\sum_{ALL}$ indicates that all of the numerator terms computed using Equation (3.10) are summed. The presented research assumes that prior distributions are finitely bounded, which makes numerically computing the sum in Equation (3.11) possible. Using Equations (3.10) and (3.11), the PMF estimation of Equation (2.2) is computed directly as

$$\log p(a \leq \mathbf{\Theta} \leq b|\mathbf{x}) \approx \log \text{numerator}_{\text{term}} - \log \sum_{ALL}^{N_l} \exp(\log \text{numerator}_{\text{terms}}). \tag{3.12}$$

Equation (3.12) provides a numerical solution for directly computing the log-posterior distribution that greatly reduces numerical precision issues. It is often more practical to depict the actual posterior distribution, which can simply be accomplished by taking the exponential of Equation (3.12). The posterior distribution can now be approximated as

$$p(a \leq \mathbf{\Theta} \leq b|\mathbf{x}) \approx \exp \left[\log \text{numerator}_{\text{term}} - \log \sum_{ALL} \exp(\log \text{numerator}_{\text{terms}}) \right]. \tag{3.13}$$

Note that extremely small values will be represented as zero due to precision limitations. However, the presented method still overcomes the divide by zero issues presented by direct computation of the denominator.

3.2.1 Computation for the Sampling Case.

Since this chapter provides comparisons between the Gaussian-Quadrature integration method and traditional PMF sampling methods, it is prudent to provide a numerical solution for the sampled PMF case. Instead of integrating over sub-regions of the numerator of the posterior distribution, we directly sample at user-defined intervals. Sampling methods assume that the sampling interval is fine enough to produce a "good enough" solution. Typically, the sampled PMF would then be normalized, such that it sums to one. However,

the posterior distribution used in the presented research already contains a normalization term in the denominator, so additional normalization is not required. We compute the numerator for the sampled case as

$$\log \text{numerator}_{\text{sampled}} \approx \frac{-(x[n] - f(\mathbf{\Theta}_i))^2}{2\sigma^2} + \log(p(\mathbf{\Theta}_i)) + \log(w_i). \qquad (3.14)$$

Note that Equation (3.14) is simply the log of the numerator. We now apply the assumption that the selected sampling interval is "good enough", and following the same process the Gaussian-Quadrature casse, the entire sampled posterior distribution can be computed as

$$\log p(a \leq \mathbf{\Theta} \leq b | \mathbf{x}) \approx \log \text{numerator}_{\text{sampled}} - \log \sum_{ALL} \exp(\log \text{numerator}_{\text{sampled}}), \qquad (3.15)$$

where "ALL" represents all of the sampled locations. Keep in mind that the implementation shown in Equation (3.15) assumes uniform sample spacing. The final computation is the same as performing a rectangular integration estimate.

3.3   Validation of Computation Method

Before proceeding further, it is imperative to provide a validation metric for the computational method defined above. We implement the following validation process:

1. Generate a simple observation vector, with observations drawn directly from a univariate normal distribution with known mean. To ensure that the estimated parameters do not consistently align with the sampling interval, each mean is randomly drawn from a uniform distribution between $a = 2$ and $b = 8$.

2. Apply three methods for PMF computation of the posterior distribution: analytic computation using the closed form cumulative distribution function (CDF) of the univariate normal distribution, computation using Gaussian-Quadrature integration techniques, and computation using traditional sampling methods.

3. Execute a Monte Carlo simulation. The simulation is iterated 10,000 times for each combination of sampling intervals, integration polynomial order, and number of observations. The cases are presented in Table 3.1.

4. Apply mean absolute error (MAE) to determine an average error between the MAP of the posterior distribution and the true parameter value. We also apply the total absolute error (TAE) metric to determine the total error between the true PMF (generated using the CDF) and the quadrature/sampled distributions. The tabular data shows TAE as the expected TAE, $E[TAE]$.

A "true" solution is computed for the posterior distribution to provide a comparison baseline. The true posterior distribution is computed using the error function, which is precise up to double floating point precision (the computational limit for the workstations implementing CSE).

We select a Monte Carlo length of $10,000$ realizations to support the validation process presented above. The simulation contains sufficient realizations once the variance of each metric converges on a single value. Figure 3.1 depicts the variance of the mean squared error (MSE) vs number of realizations (for the $N = 5000$ observations case at 0.01 interval). Note that the variance for each metric converges around 2,000 realizations. Therefore we can conclude that 10,000 realizations used for the validation are more than sufficient.

Table 3.2 depicts sample data from the first 20 realizations of the Monte Carlo simulation for the $N = 1$ observation at 1.0 interval case. Additional data samples are located in the Appendix. Each row in Table 3.2 contains the results of one simulation, including true parameter value (μ), MAP estimates ($\hat{\mu}$), and TAE. Table 3.3 depicts the results of the validation process. Analysis of the results is divided into the following sections: graphical method comparison, MAP error analysis, effects of quadrature order, and effects of sampling interval.

Table 3.1: Computation method validation scenarios. Graphical results for rows with asterisks are presented in this chapter. Additional samples of data from each scenario can be found in the Appendix.

Observations	Interval	Quadrature Order
1	1	3
*1	1	5
1	1	11
1	0.1	5
1	0.01	5
5	1	5
5	0.1	5
5	0.01	5
50	1	5
50	0.1	5
50	0.01	5
500	1	5
500	0.1	5
500	0.01	5
5000	1	5
5000	0.1	5
5000	0.01	3
*5000	0.01	5
5000	0.01	11
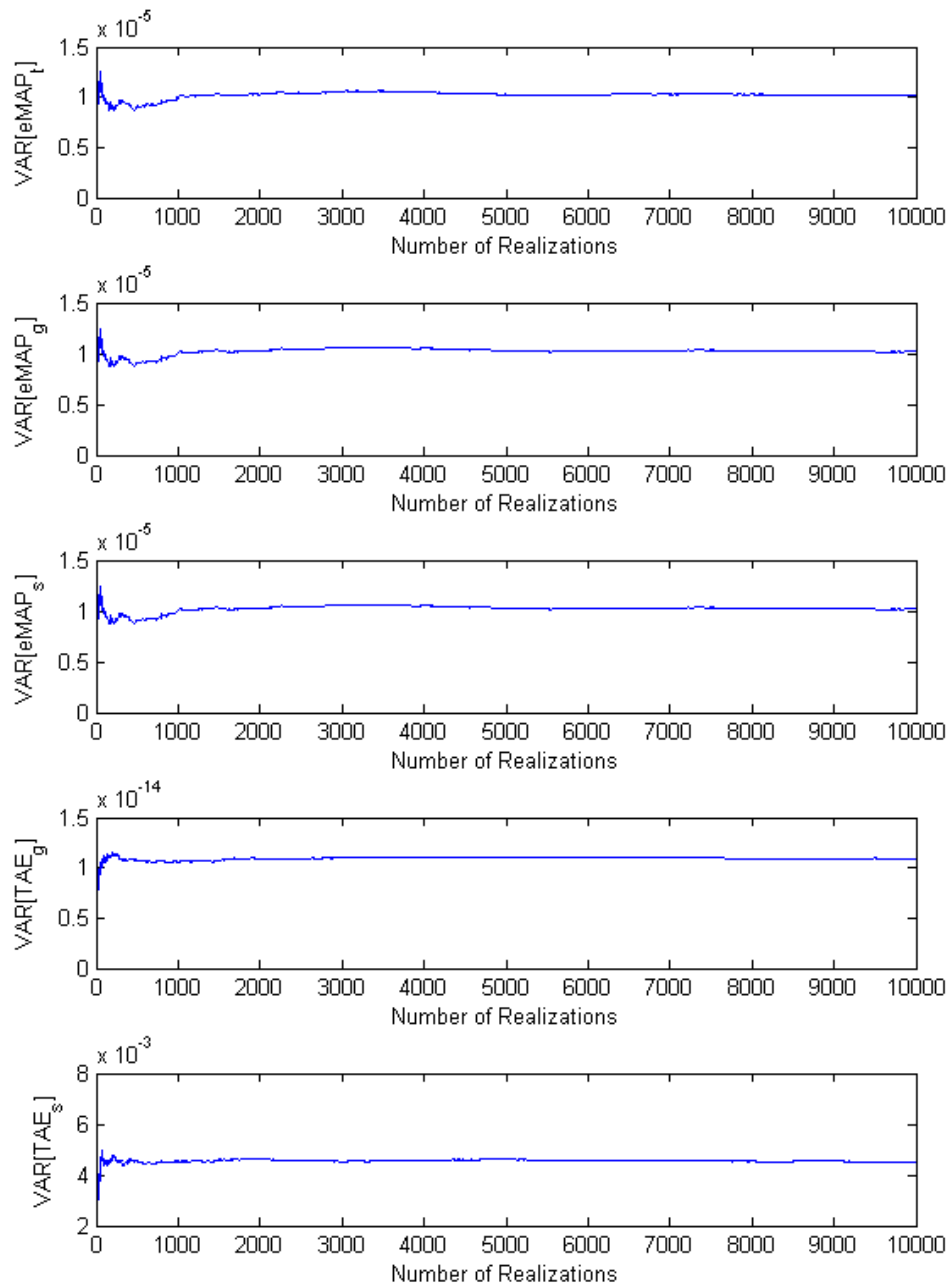
Figure 3.1: Variance versus realizations for the validation simulations. Legendre polynomial order $Nl = 5$, $N = 5000$ observations, 0.01 interval.

Table 3.2: Sample Monte Carlo data from the $N = 1$ observation at 1.0 interval case. Data includes the actual mean, μ, MAP estimates of the mean, $\hat{\mu}$, error between MAP and the true mean, $\epsilon\hat{\mu}$ and TAE. Subscripts e, g, and s indicate error function computation, Gaussian-Quadrature and sampling respectively. Additional samples of data from each scenario can be found in the Appendix.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
6.3861	7	7	7	0.61394	0.61394	0.61394	7.6491e-07	0.25532
6.4953	6	6	6	0.49532	0.49532	0.49532	1.4213e-06	0.26071
6.0176	6	6	6	0.01761	0.01761	0.01761	2.7533e-05	0.2594
6.2469	7	7	7	0.75307	0.75307	0.75307	2.3394e-05	0.30034
5.2109	5	5	5	0.21087	0.21087	0.21087	1.1269e-05	0.22228
6.6955	7	7	7	0.30452	0.30452	0.30452	1.6234e-05	0.20375
7.201	7	7	7	0.201	0.201	0.201	1.5209e-05	0.20778
7.714	8	8	8	0.28603	0.28603	0.28603	1.4878e-05	0.12619
7.671	8	8	8	0.32896	0.32896	0.32896	4.0393e-06	0.03323
6.241	6	6	6	0.24097	0.24097	0.24097	1.243e-05	0.10477
6.6375	7	7	7	0.36249	0.36249	0.36249	1.0067e-05	0.084403
6.9368	7	7	7	0.063236	0.063236	0.063236	2.7704e-05	0.28792
6.7631	7	7	7	0.23688	0.23688	0.23688	2.7608e-05	0.26093
7.6998	7	7	7	0.69984	0.69984	0.69984	4.1897e-06	0.26847
5.6838	6	6	6	0.31624	0.31624	0.31624	1.5742e-05	0.20569
5.2989	6	6	6	0.70112	0.70112	0.70112	1.8191e-05	0.15596
5.7187	6	6	6	0.28126	0.28126	0.28126	2.803e-05	0.27429
5.8683	6	6	6	0.13168	0.13168	0.13168	2.7864e-05	0.26724
5.9294	6	6	6	0.070595	0.070595	0.070595	9.9202e-07	0.25467
5.0678	5	5	5	0.067777	0.067777	0.067777	5.6491e-06	0.24084
5.6006	6	6	6	0.39943	0.39943	0.39943	1.4934e-06	0.011785

Table 3.3: Statistical data for the validation process. Data includes the number of observations, interval, Gauss-Legendre polynomial order (Nl), MAE, and expected TAE ($E[TAE]$). Subscripts e, g, and s indicate error function computation, Gaussian-Quadrature and sampling respectively.

# Obs.	Int.	Nl	MAE_e	MAE_g	MAE_s	$E[TAE_g]$	$E[TAE_s]$
1	1	3	0.34989	0.34989	0.34989	0.0016934	0.22549
1	1	5	0.34421	0.34421	0.34421	1.5266e-05	0.22777
1	1	11	0.34052	0.34052	0.34052	3.2251e-14	0.22786
1	0.1	5	0.25519	0.25519	0.25519	4.4549e-15	0.0040228
1	0.01	5	0.25285	0.25285	0.25285	1.9682e-14	4.0327e-05
5	1	5	0.26829	0.26829	0.26829	0.001384	0.16794
5	0.1	5	0.11494	0.11494	0.11494	1.0357e-11	0.019926
5	0.01	5	0.11377	0.11377	0.11377	5.2196e-14	0.00020162
50	1	5	0.25091	0.25091	0.25091	0.012431	0.066141
50	0.1	5	0.042495	0.042495	0.042495	3.1179e-07	0.17388
50	0.01	5	0.03572	0.03572	0.03572	3.8447e-14	0.0020139
500	1	5	0.24875	0.24875	0.24875	0.016555	0.022144
500	0.1	5	0.027314	0.027314	0.027314	0.0013885	0.1694
500	0.01	5	0.011432	0.011432	0.011432	1.0355e-11	0.019904
5000	1	5	0.24962	0.24962	0.24962	0.0080092	0.0088198
5000	0.1	5	0.025262	0.025262	0.025262	0.0126	0.064866
5000	0.01	3	0.0042946	0.0042946	0.0042946	0.00029287	0.17263
5000	0.01	5	0.004306	0.004306	0.004306	3.1338e-07	0.17319
5000	0.01	11	0.0042444	0.0042444	0.0042444	1.7014e-12	0.17387

3.3.1   *Graphical Method Comparison.*

This subsection provides graphical comparison of the sampled and Gaussian-Quadrature methods. Figures 3.2 and 3.3 depict posterior distributions for the two scenarios with asterisks listed in Table 3.1 (additional results can be found in the Appendix). Notice that the Gaussian-Quadrature integration method provides a significantly more accurate solution than the sampling method when compared against the error function computation. The results in the following subsections indicate a significant increase in accuracy for the posterior distribution when computed using Gaussian-Quadrature integration techniques.
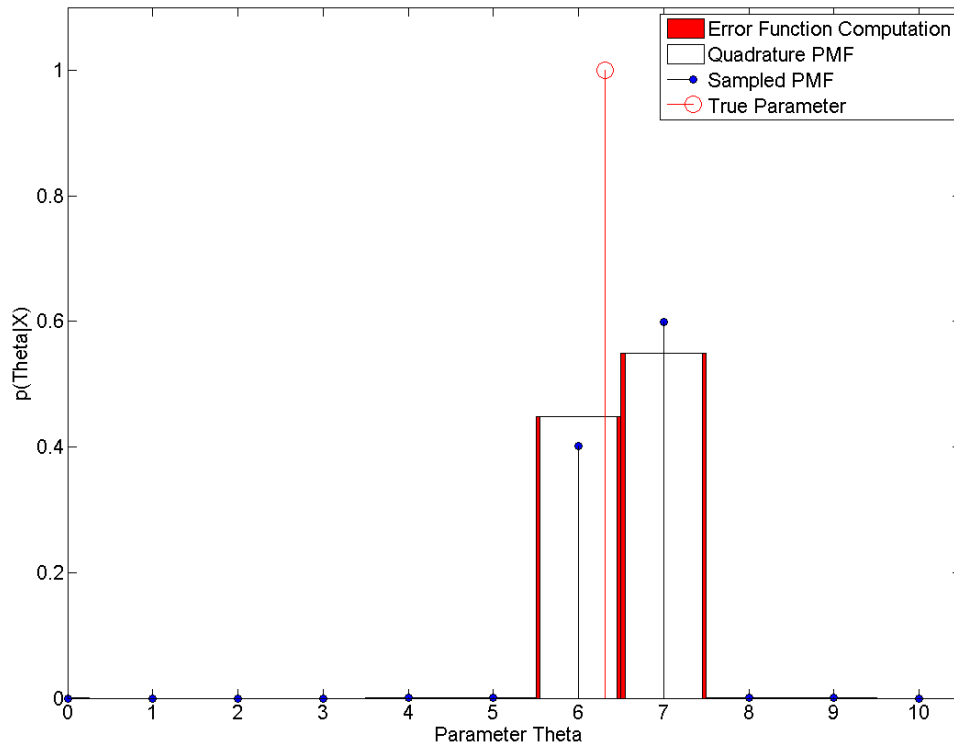
Figure 3.2: Validation results for *N* = 1 observation at 1.0 interval. The red outlines are the distribution as computed with the error function, the white boxes are the distribution generated using the Gaussian-Quadrature method, and the blue/black stems are the distribution generated using the sampling method. Note that the true mean is the actual parameter of the drawn distribution.
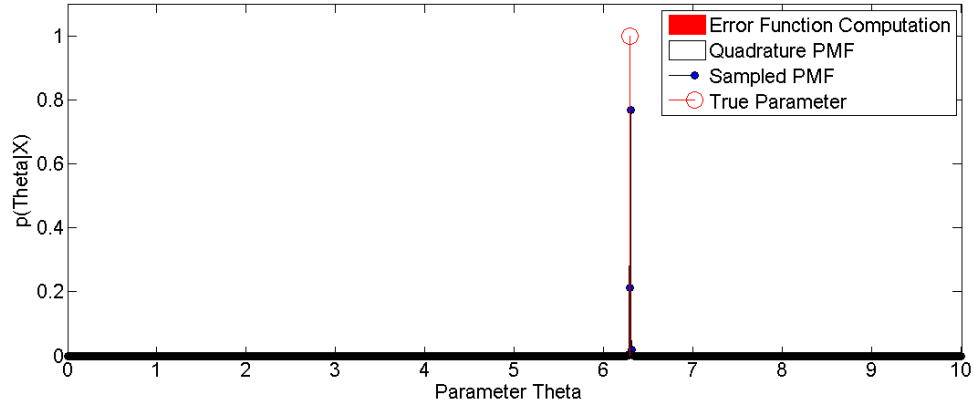
(a) Full distribution.
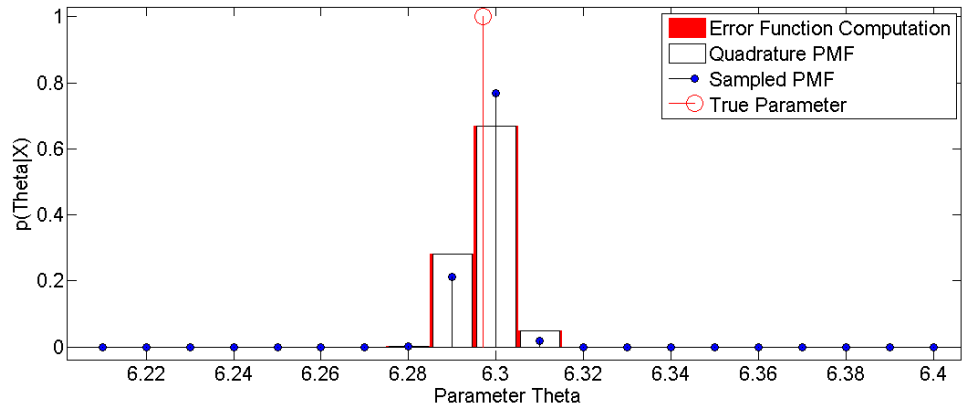
(b) Zoomed region $6.2 \leq \Theta \leq 6.5$.

Figure 3.3: Validation results for $N = 5000$ observation at 0.01 interval. The red outlines are the distribution as computed with the error function, the white boxes are the distribution generated using the Gaussian-Quadrature method, and the blue/black stems are the distribution generated using the sampling method. Note that the true mean is the actual parameter of the drawn distribution.

3.3.2 MAP and Distribution Error Analysis.

Table 3.3 contains various metrics for comparing quadrature and sampled methods. The MAE metric is computed as

$$MAE = \frac{\sum_{n=1}^{N} |\mu - \hat{\mu}_n|}{N}, \tag{3.16}$$

29

where N is the total number of data points, μ is the true parameter value, and $\hat{\mu}$ is the MAP estimate. Table 3.3 also includes TAE, which is computed as

$$TAE = \sum_{ALL\Theta} |p(\Theta|x)_1 - p(\Theta|x)_2|, \tag{3.17}$$

where $p(\Theta|x)_1$ and $p(\Theta|x)_2$ are the two distributions being compared.

The data in Table 3.3 demonstrates that the MAE is identical for each method, showing that using the MAP parameter set determined by the CSE provides the same solution as the sampling method. The similarity in performance is due to the cost function and sampling interval. Recall the cost function for MAP estimators, defined as

$$C(\epsilon) = C(\mu - \hat{\mu}), \tag{3.18}$$

where $\hat{\mu}$ is the closest discrete parameter [6]. Since the MAP is computed from the center of the search bin, the resulting estimates are the same for both the CSE and traditional sampling methods. In terms of a sole MAP estimator, it would not be beneficial to apply Gaussian-Quadrature methods. However, application of CSE is intended for scenarios when retaining the entire distribution is desirable. The next section shows that traditional sampling induces significant errors in the posterior distribution.

3.3.3 Posterior Distribution Error Analysis.

The data in Table 3.3 shows that the expected TAE is significantly lower in all cases for the Gaussian-Quadrature method than the sampling method. The Gaussian-Quadrature method also produces lower error, regardless of the number of observations or sampling interval. On average, the sampling method produces $6.43e10$ times more error than the Gaussian-Quadrature method. In addition, an error of 0.2277 (row 2, Table 3.3) for a posterior distribution is quite significant given that the distribution sums to one.

3.3.4 Effects of Sampling Interval.

The data in Table 3.3 provides some insight into CSE performance as the sampling interval varies. As mentioned previously, MAP estimates provided by the CSE are more

precise as the space is sampled more finely. The precision increases at the same rate as the sampling method because MAP performance depends on sampling interval, as shown in Equation (3.18). Performance in terms of expected TAE demonstrates in some interesting findings. In every case, the CSE implementation using Gaussian-Quadrature techniques results in smaller expected TAE than the sampling method. Note that the CSE implementation produces more accurate posterior distribution estimates as the sampling interval varies from coarse to fine in every case but the $N = 5000$ at 0.1 interval case. However, the expected TAE still increases as the sampling interval varies from coarse to fine. The next section shows that larger quadrature polynomials can overcome the increase in error. However, the sampling method does not share the same accuracy improvement. In most cases, the sampling method performs better with finer sampling interval. However, notice that the error associated with finest sampling interval for the $N = 5000$ observations case greatly increases. Recall Figure 3.3 and note that the general shape of the distribution is much different for the sampling method than the quadrature integration method. As the results of this thesis are presented, it will become apparent that using a finer sampling interval does not result in the same shaped distributions when comparing against quadrature integration. As a result, decreasing the sampling interval typically produces more TAE for the sampling method in cases with very narrow distributions, as is the case with large numbers of observations.

3.3.5 *Effects of Quadrature Order.*

Recall Equation (3.13) and notice that as the number of observations increases, the width of the distribution will decrease due to the summation term in the exponential. Referring to Table 3.3, note that as the number of observations increases, the accuracy of the CSE implementation decreases. While the CSE produces more accurate estimates of the posterior distributions (compared to sampling methods), it is desirable to explore the possibility of maintaining the same accuracy by varying the order of the integration

31

polynomial. Maintaining accuracy becomes increasingly important in the case of SAR canonical models, where the number of observations in phase history can easily exceed 7,300.

Consider the $N = 1$ observation at 1.0 interval case. Notice that the expected TAE decreases significantly as the polynomial order is increased from $N_l = 3$ to $N_l = 11$. Comparing to the sampling case, note that a polyomial of order $N_l = 3$ is likely sufficient to estimate the posterior distribution. Now, consider the $N = 5000$ case, recalling that the increase in number of observations results in a much narrower posterior distribution. Again, increasing the integration polynomial order results in a more accurate representation of the posterior distribution. Note that the increase in accuracy comes at a computational cost as the $N_l = 11$ polynomial results in more than double the computations when compared to $N_l = 5$. It is especially important to consider the computational cost for multi-dimensional distributions as the complexity exponentially increases with the number of computations. In order to balance the computational complexity while still producing accurate posterior distributions, we select an integration polynomial of $Nl = 5$ for the remainder of the results presented in this thesis. Chapter 4 will show that using a higher order integration polynomial makes implementation of CSE computationally impractical.

3.4 Determining the Credible Set

Parameter ambiguities can lead to multiple peaks in the output posterior distribution. The sinusoidal signal leads to multiple peaks of equal probability for the phase estimate, which creates issues for MAP style estimators (there is no intuitive way to pick the maximum peak). In many cases, noise leads to one peak that contains marginally more probability than the others. Traditional MAP estimators would just accept this peak as the solution. However, running the estimator on multiple realizations of the same signal would demonstrate "hopping" parameters, indicating that multiple "credible" solutions exist.

The CSE retains these peaks, as well as the entire PMF. The CSE then determines the region in which the total probability equals a user-defined confidence level. For the purposes of this thesis, a confidence level of 90% will be used.

3.5 Parameter Estimation and Priors

The PMFs presented in this chapter are generated by implementing the CSE in MATLAB. The results presented are two-parameter PMF estimates after applying the CSE on simple test signals (DC ramp and sinusoid). The following subsections will detail the results of applying the CSE to each test signal. Results are generated for posterior PMFs which are discretized at 1, 0.1 and 0.01 intervals. All quadrature distributions were generated using a legendre polynomial of order $N_l = 5$. The validation section, Section 3.3, has shown that these parameters result in a very accurate computation of the posterior distribution. As such, the TAE metrics presented in the following sections directly compare the distributions generated with the quadrature method and the sampling method. Since no closed form solution exists for the DC ramp and sinusoidal signals, the following sections accept the posterior distributions generated by the quadrature method as the correct distributions. This chapter will close with a table containing parameter estimates and associated errors.

3.5.1 Simple DC Ramp.

Here, we demonstrate the CSE on a multiple parameter estimate of the signal

$$x[n] = An + B + w[n], \tag{3.19}$$

where A and B are the signal parameters, and $w[n]$ is a vector containing WGN with variance σ^2. For this case, the resulting likelihood function is

$$p(\mathbf{x}|A, B) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=1}^{N} (x[n] - An - B)^2\right]. \tag{3.20}$$

Bayes' Rule is applied using Equation (3.20) and a uniform prior. The resulting posterior distributions are depicted in Figures 3.4 and 3.5. Notice that a traditional MAP

style estimator would likely result in the correct solution, since the distribution only contains one peak. The credible region for the simple DC signal case corresponds to the MAP estimate of the parameters. Notice the drastic difference in shape between the two distributions, especially at the finest zoom level. Since the quadrature method has been demonstrated to provide more accurate representations of the distributions, we can assume that the sampling method has induced significant error in the distribution. The tabular results section at the end of this chapter further depicts the error.

(a) Posterior distribution and credible set at 1.00 interval.

(b) Posterior distribution and credible set at 0.10 interval.

(c) Posterior distribution and credible set at 0.01 interval.
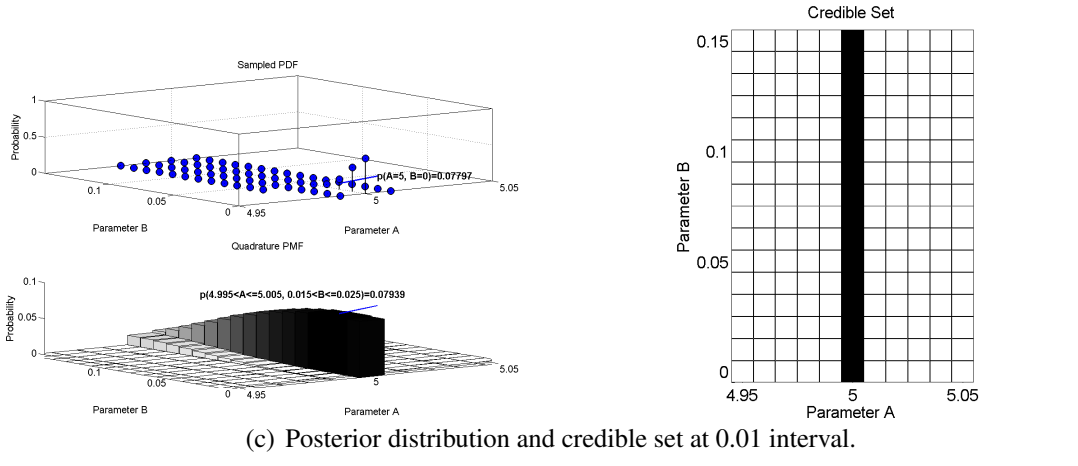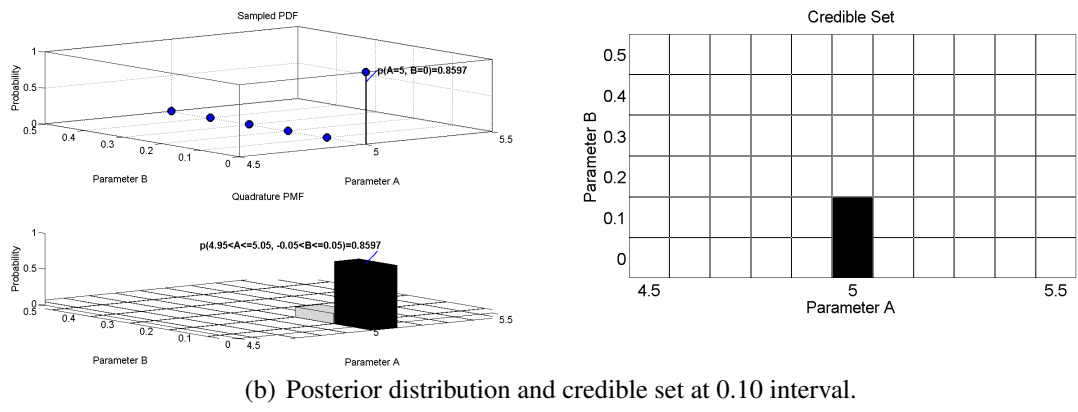
Figure 3.4: Left: PMFs for the multiple parameter estimate of $An + B$ using uniform priors. Parameters used are $A = 5$, $B = 0$ and $N = 50$. Right: Resulting 90% credible regions.

(a) Marginal distribution at 1.00 interval.

(b) Marginal distribution at 0.10 interval.

(c) Marginal distribution at 0.01 interval.

Figure 3.5: Marginal PMFs for the multiple parameter estimate of $An + B$ using uniform priors. Parameters used are $A = 5$, $B = 0$ and $N = 50$.

3.5.2 *Sinusoidal Signal.*

Sinusoidal signals are an interesting case because of the cyclical nature of the phase term. The signal demonstrated here is $x[n] = \cos(2\pi An + B)$, where A is an unknown frequency, and B is an unknown phase offset. A uniform prior distribution (centered on the true value, with width = 2) is initialized, and the PMF is computed according to Equation (3.13). The results are depicted in Figures 3.6 and 3.7. Notice in these figures a traditional MAP style estimator would likely not result in the correct solution. The credible set for the cosine case shows that multiple credible solutions exist. A typical MAP estimator would likely result in "hopping" solutions from realization to realization. Again, notice the drastic difference in shape between the two distributions, especially at the finest zoom level. Since the quadrature method has been demonstrated to provide more accurate representations of the distributions, we can again assume that the sampling method has induced significant error in the distribution.

(a) Posterior distribution and credible set at 1.00 interval.

(b) Posterior distribution and credible set at 0.10 interval.
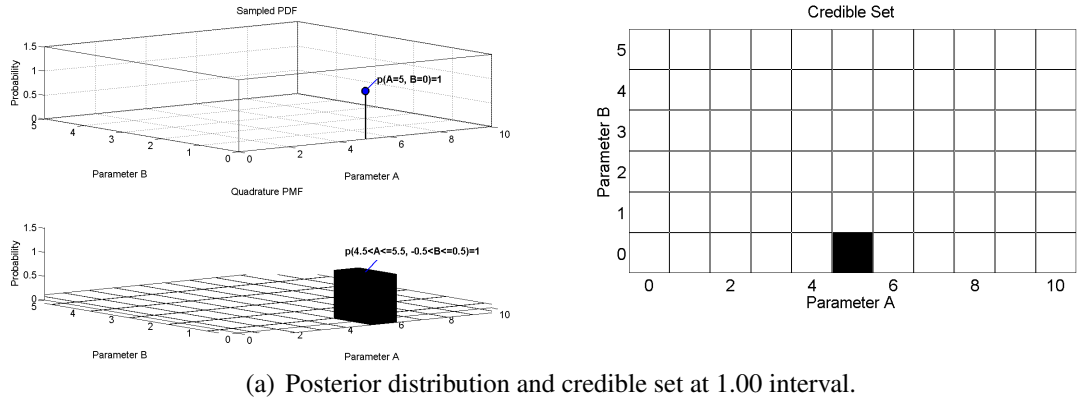

(c) Posterior distribution and credible set at 0.01 interval.
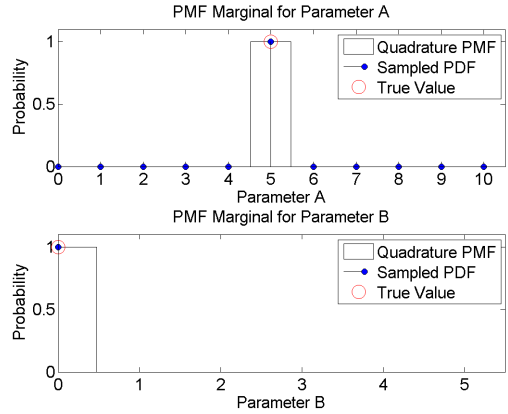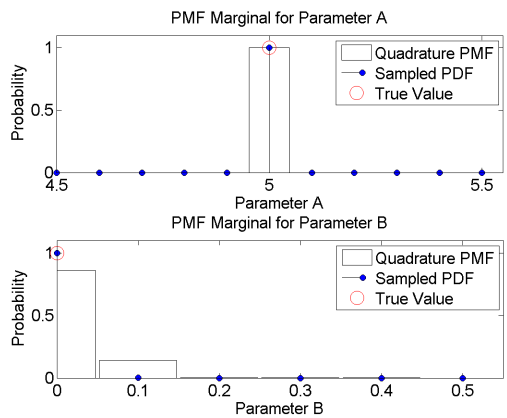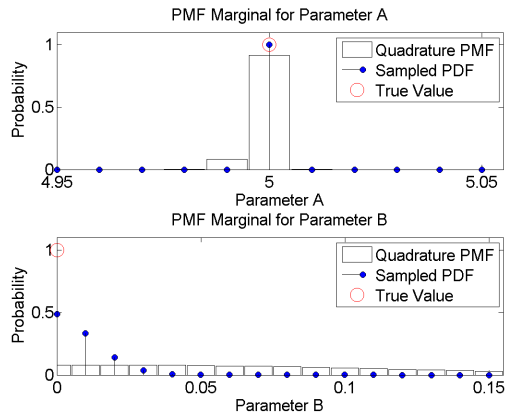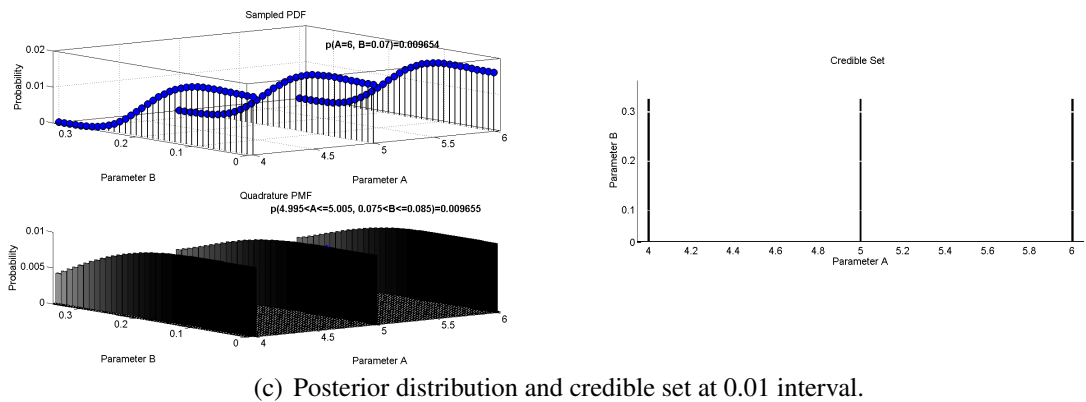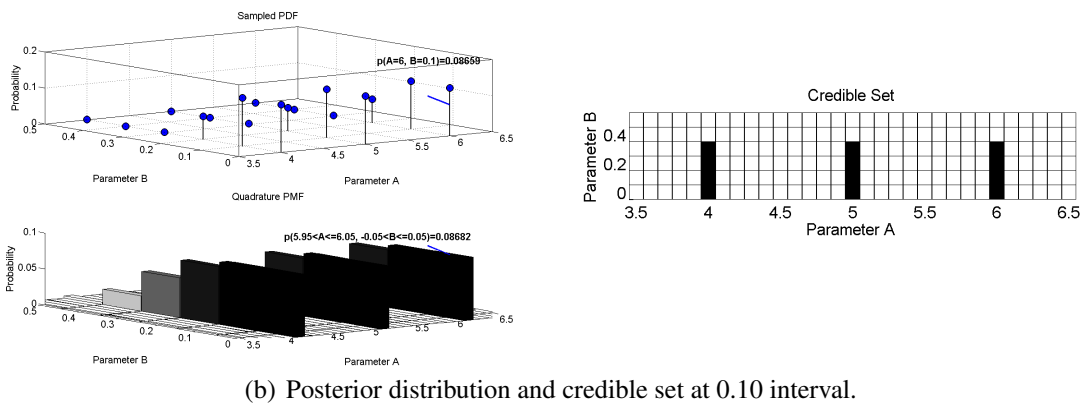
Figure 3.6: Left: PMFs for the multiple parameter estimate of $\cos(2\pi A + B)$ using uniform priors. Parameters used are $A = 5$, $B = 0$ and $N = 50$. Right: Resulting 90% credible regions. Note: interval size decreases from top to bottom.
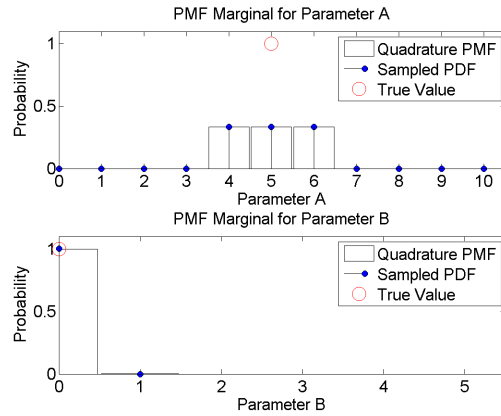
(a) Marginal distribution at 1.00 interval.
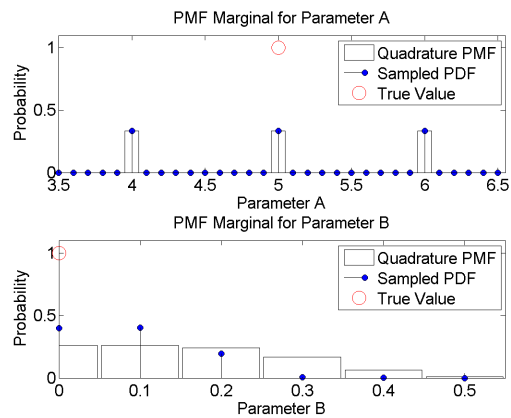
(b) Marginal distribution at 0.10 interval.
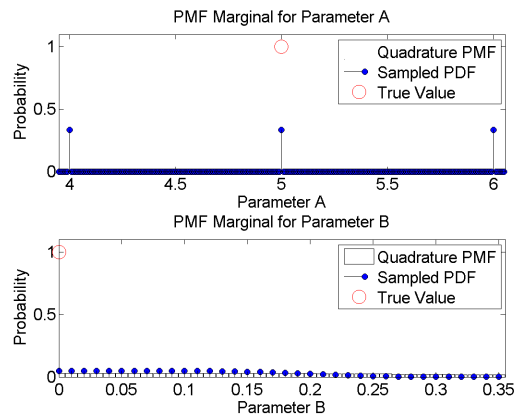
(c) Marginal distribution at 0.01 interval.

Figure 3.7: Marginal PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$ using uniform priors. Parameters used are $A = 5$, $B = 0$ and $N = 50$.

3.6   Selection of Priors

Selection of priors can significantly impact the performance of the estimator. For the purposes of this thesis, three different prior distributions are considered: uniform, Gaussian (normal) and Rayleigh. It is useful to demonstrate prior selection on a sinusoidal signal, as estimating the phase results in a cyclical response in the likelihood function, as demonstrated in Figure 3.6.

3.6.1   Multi-Dimensional Priors.

Prior to presenting the results of prior selection, it is pertinent to define how multi-dimensional priors are intialized in this thesis. For the purposes of this thesis, signal parameters are considered to be independent. That is, the value of each parameter does not rely on the value of another. This assumption significantly simplifies the generation of multi-dimensional priors, and matches well with the canonical feature models used in the next chapter (parameters such as size, orientation and location are independent). As such, multi-dimensional priors are defined as a vector of distributions, shown as

$$f(\mathbf{\Theta}) = f(\mathbf{\Theta}_1, \mathbf{\Theta}_2, \ldots, \mathbf{\Theta}_N), \tag{3.21}$$

where $\mathbf{\Theta}_1, \mathbf{\Theta}_2, \ldots, \mathbf{\Theta}_N$ are prior distributions for each parameter.

3.6.2   Uniform Priors.

The uniform distribution is desirable both because of its simplicity to implement, and because it provides finite integration bounds for the posterior distribution in the denominator of Equation (2.1) [15], and is shown as

$$f(\mathbf{\Theta}) = \frac{1}{B - A}, \tag{3.22}$$
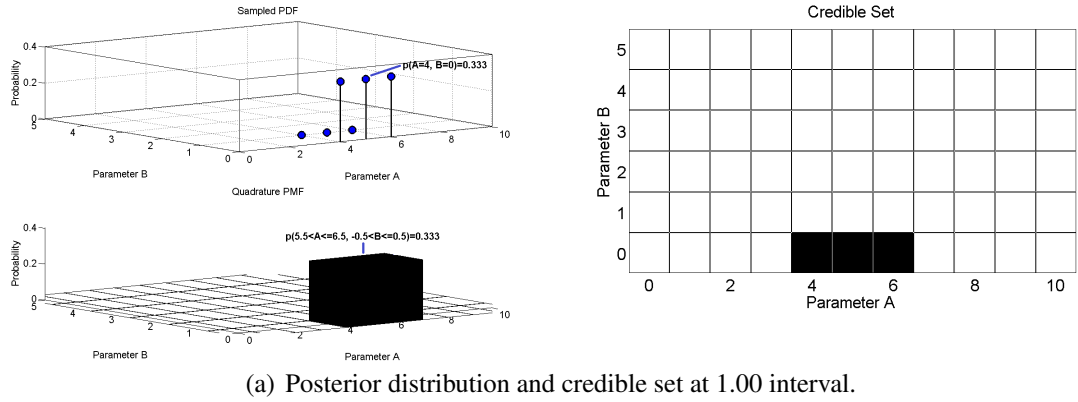
where A and B are the chosen end points. Initializing the uniform prior proves to be an issue, however, as the canonical feature estimators used in this thesis [4, 13] provide only one estimated value. For the purposes of this research, we center the distribution on the

true value using the equation for the mean of the uniform distribution, defined as

$$\mu = \frac{1}{2}(A + B). \tag{3.23}$$

It is clear that infinitely many combinations of A and B can be selected to result in the same mean. As such, initialization of uniform priors requires the user to select the width of the distribution. For the purposes of this thesis, arbitrary values of $A = \mu - 1$ and $B = \mu + 1$ were selected.

Figure 3.6 depicts the resulting PMF when the sinusoidal signal is estimated using uniform priors. The uniform distribution is bounded such that only three peaks remain. An unbounded uniform distribution would have resulted in repeated peaks indefinitely. The uniform distribution is useful in applications where multiple credible regions occur, but should be weighted the same.

3.6.3 Normal Priors.

The Gaussian (normal) distribution is desirable because it is parameterized by two moments, which enables tailoring of the prior to a desired outcome [15]. The Gaussian distribution is defined as

$$f(\Theta) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(\Theta - \mu)^2}{2\sigma^2}}, \tag{3.24}$$

where μ is the mean, and σ is the variance. Similar to the uniform case, only one initialization value is provided by traditional estimators. The variance must be selected by the user, enabling the user to choose whether emphasis should be placed on the likelihood estimate or on the moment estimate for the prior. Selecting small variance places more emphasis on initialization value, whereas large variance emphasizes the likelihood estimate. For the purposes of this thesis, the mean is initialized as the true value, an arbitrary value of variance, $\sigma = 1$ is selected.

Figure 3.8 depicts the resulting PMF when the sinusoidal signal is estimated using Gaussian priors. Figure 3.9 contains the resulting marginal distributions. Notice that five peaks remain (compared to using an unbounded prior distribution), however the peaks

41

reduce in amplitude as the parameters deviate from the true value. Implementing normal priors is desirable in cases where one peak should be emphasized, but it is undesirable to discard other possibilities.
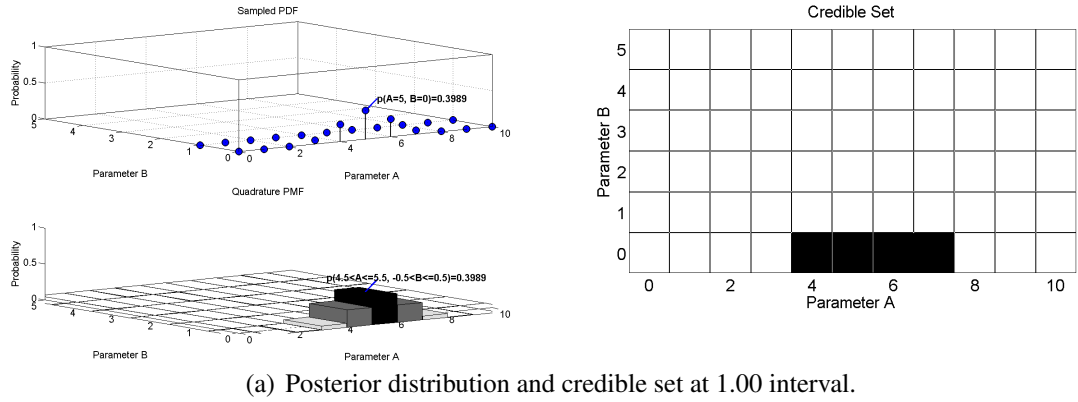
(a) Posterior distribution and credible set at 1.00 interval.
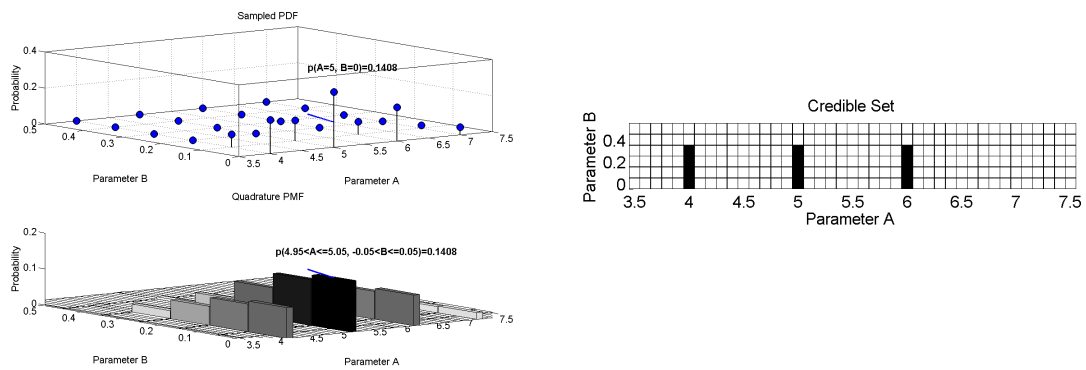
(b) Posterior distribution and credible set at 0.10 interval.
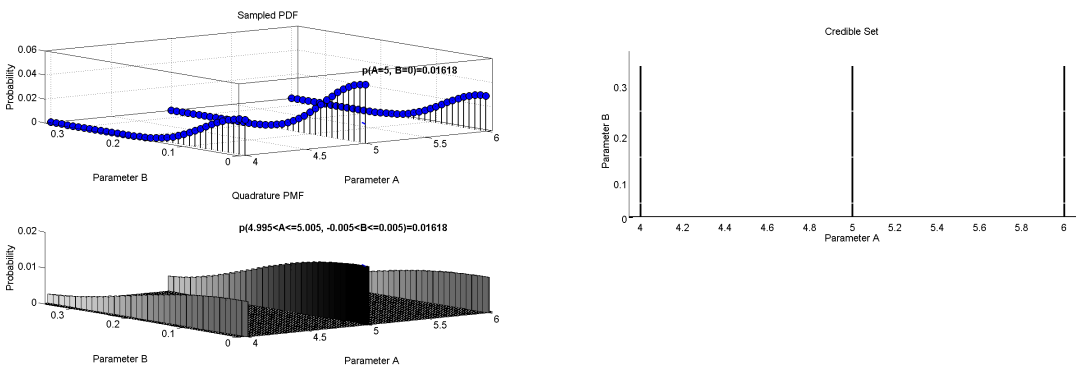
(c) Posterior distribution and credible set at 0.01 interval.

Figure 3.8: Left: PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$ using normal priors. Parameters used are $A = 5$, $B = 0$ and $N = 50$. Right: Resulting credible regions. Note: interval size decreases from top to bottom.
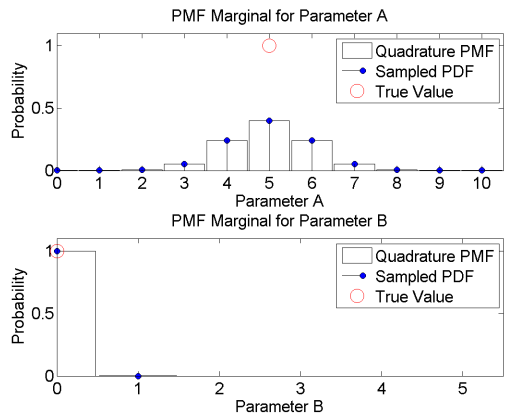
(a) Marginal distribution at 1.00 interval.

(b) Marginal distribution at 0.10 interval.

(c) Marginal distribution at 0.01 interval.

Figure 3.9: Marginal PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$ using normal priors. Parameters used are A=5, B=0 and N=50.

3.6.4 *Rayleigh Priors.*

The Rayleigh distribution is desirable because it is parameterized by a single parameter, which can be traced back to the mean of the distribution [15]. The Rayleigh distribution is defined as

$$f(\Theta) = \frac{\Theta}{\sigma_r^2} e^{-\frac{\Theta^2}{2\sigma_r^2}}, \tag{3.25}$$

where σ_r is the scaling parameter. The Rayleigh distribution also has the unique property in which the variance and mean are coupled to the same parameter, resulting in a narrower distribution as the mean value decreases. The coupling is useful for the CSE, as estimates of the mean value are weighted more in the final PMF when the value is smaller. The coupling will prove useful in size estimates of canonical features, as it is more likely that an object in the scene is small relative to the scene size. One drawback to the Rayleigh distribution is that it is undefined at $\sigma_r = 0$, making the distribution unsuitable for parameters that can contain zero. Figures 3.10 and 3.11 depict the resulting PMF when the sinusoidal signal is estimated using Rayleigh priors. Notice that multiple peaks remain (compared to using an unbounded prior distribution), however the peaks reduce in amplitude as the parameters deviate from the true value. This reduction is desirable in cases where one peak should be emphasized, but it is undesirable to discard other possibilities. Notice that the 90% region still includes two, disjoint regions. Setting a flag in the CSE can force the estimator to treat this large region, as two separate, disjoint regions.

(a) Posterior distribution and credible set at 1.00 interval.
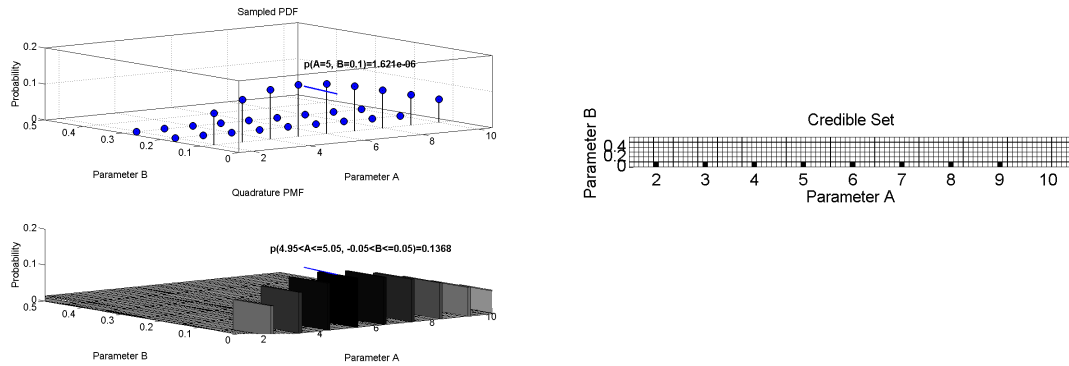

(b) Posterior distribution and credible set at 0.10 interval.
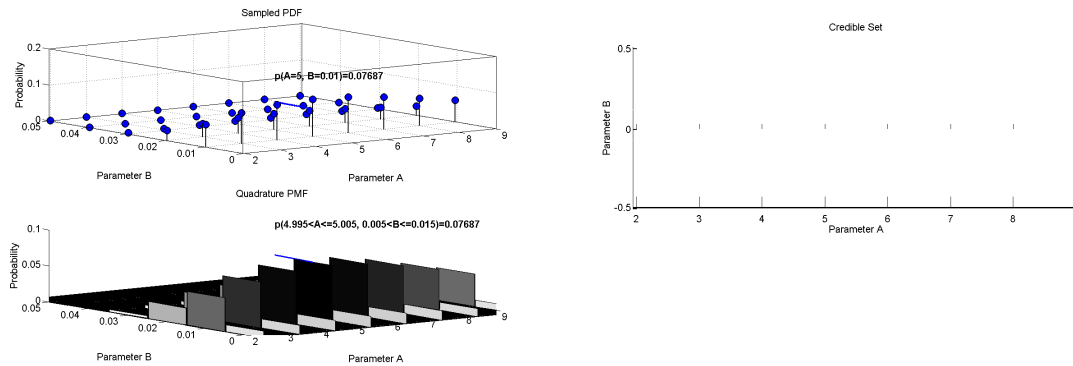
(c) Posterior distribution and credible set at 0.01 interval.

Figure 3.10: Left: PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$ using Rayleigh priors. Parameters used are $A = 5$, $B = 0$ and $N = 50$. Right: Resulting 90% credible regions. Note: interval size decreases from top to bottom.
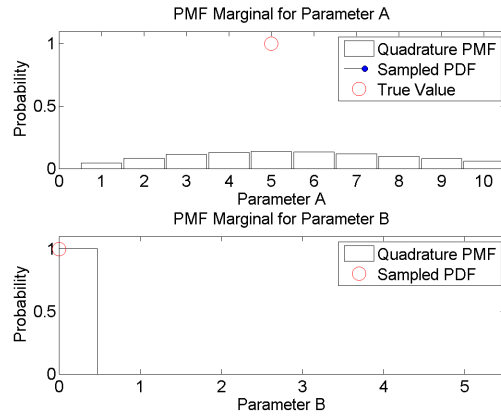
(a) Marginal distribution at 1.00 interval.
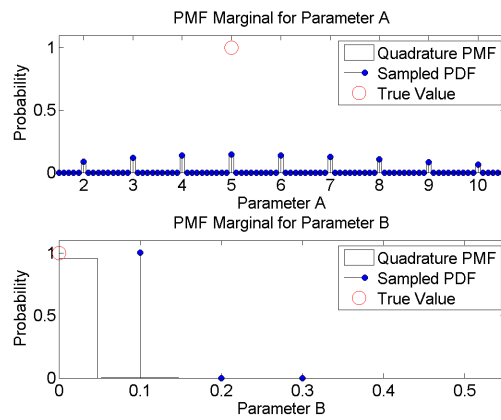
(b) Marginal distribution at 0.10 interval.

(c) Marginal distribution at 0.01 interval.

Figure 3.11: Marginal PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$ using Rayleigh priors. Parameters used are A=5, B=0 and N=50.

3.6.5 Large Observation Vectors.

Applying the CSE with large observation vectors provides higher confidence levels at the cost of increased numerical precision requirements. Computing the PMF as shown in Equation (3.13) resolves the precision requirements issue, thus providing the benefit of multiple observations while minimizing the risk of data overflow. Figures 3.12, 3.13 3.14 and 3.15 show the results of applying the CSE to both the DC signal and sinusoidal signal, using an observation vector with length $N = 5000$. This value was selected as it is of the same order as the observation vector used for the canonical features phase history examples shown in the next chapter. Notice that the PMFs contain significantly narrower peaks than the cases with few observations, which can cause the estimator to miss the region of probability. When this occurs, reducing the interval size or increasing the legendre polynomial order can alleviate the problem. Again, notice the drastic difference in shape between the two distributions, especially at the finest zoom level. We can assume that the sampling method has induced significant error in the distribution.

(a) Posterior distribution and credible set at 1.00 interval.

(b) Posterior distribution and credible set at 0.10 interval.
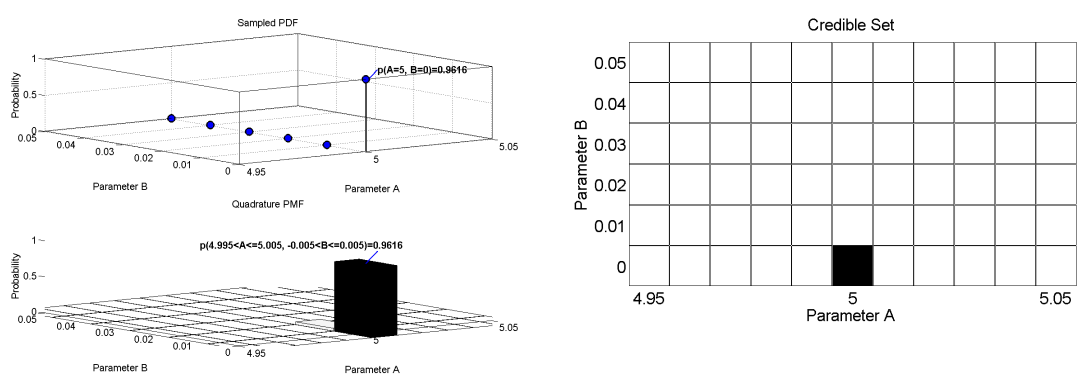
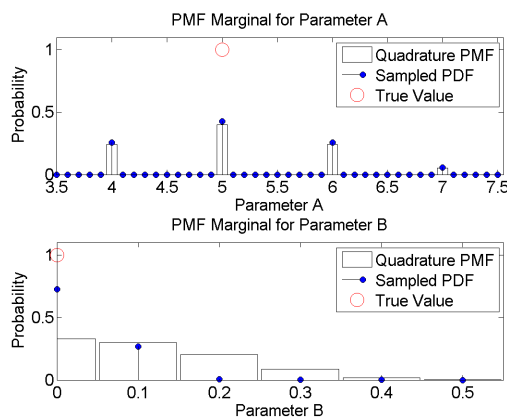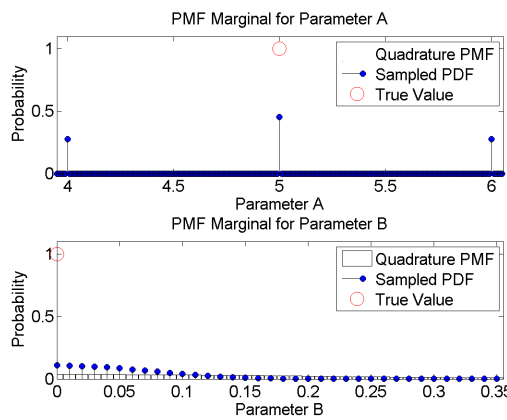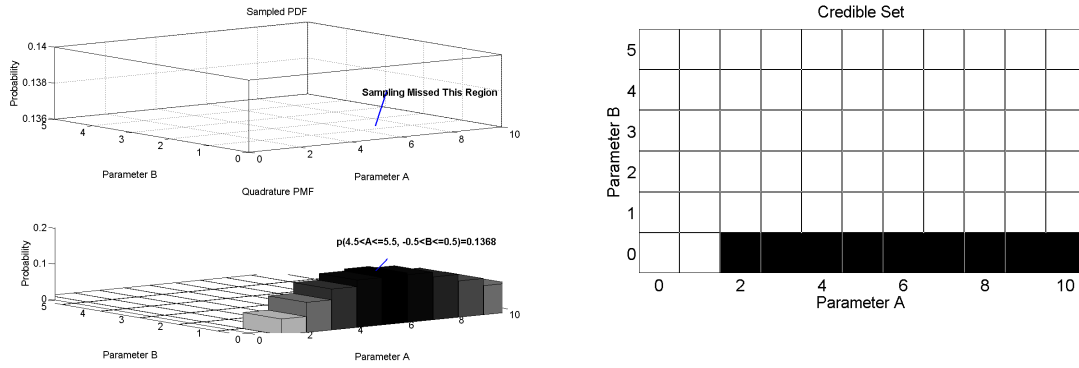(c) Posterior distribution and credible set at 0.01 interval.

Figure 3.12: Left: PMFs for the multiple parameter estimate of *An + B*. Parameters used are *A* = 5, *B* = 0 and *N* = 5000. Right: Resulting 90% credible regions. Note: interval size decreases from top to bottom.
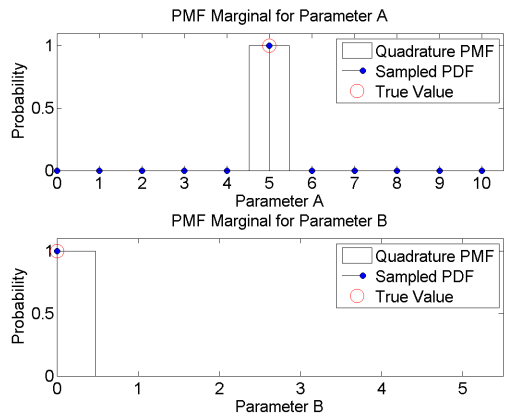
(a) Marginal distribution at 1.00 interval.
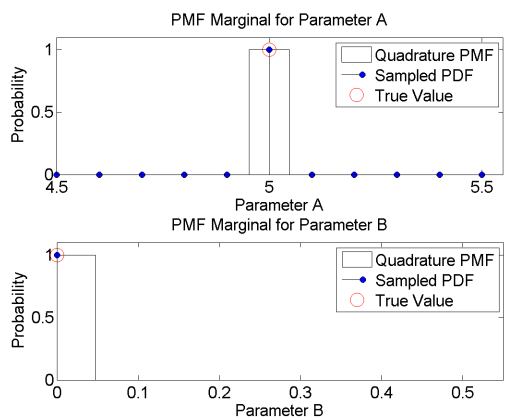
(b) Marginal distribution at 0.10 interval.
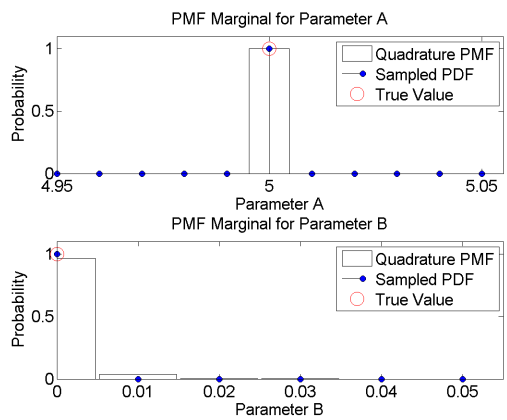
(c) Marginal distribution at 0.01 interval.

Figure 3.13: Marginal PMFs for the multiple parameter estimate of $An + b$. Parameters used are $A = 5$, $B = 0$ and $N = 5000$.

(a) Posterior distribution and credible set at 1.00 interval.

(b) Posterior distribution and credible set at 0.10 interval.

(c) Posterior distribution and credible set at 0.01 interval.
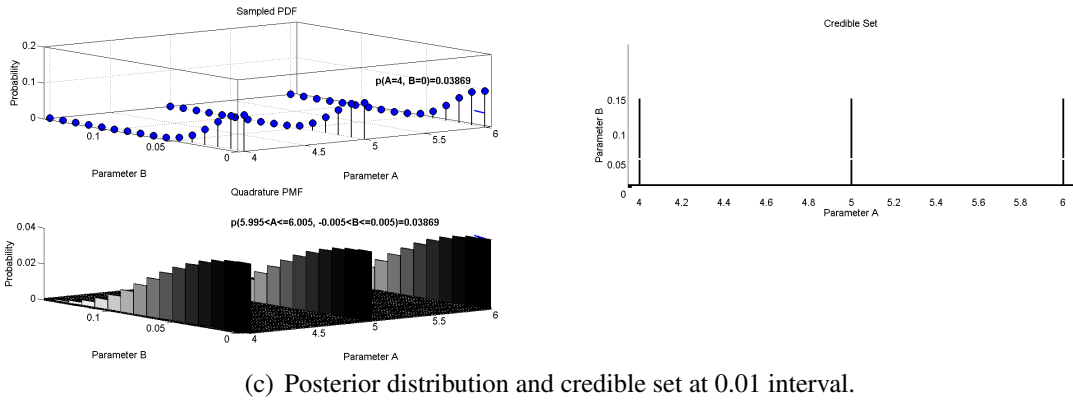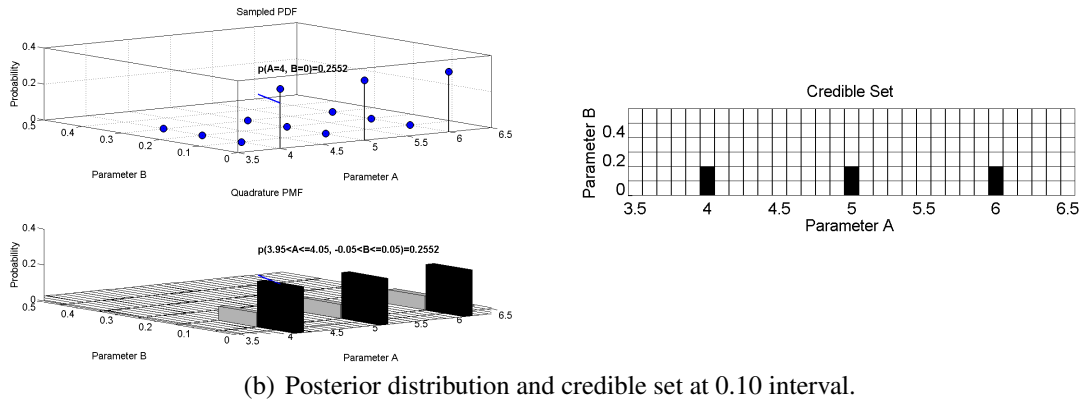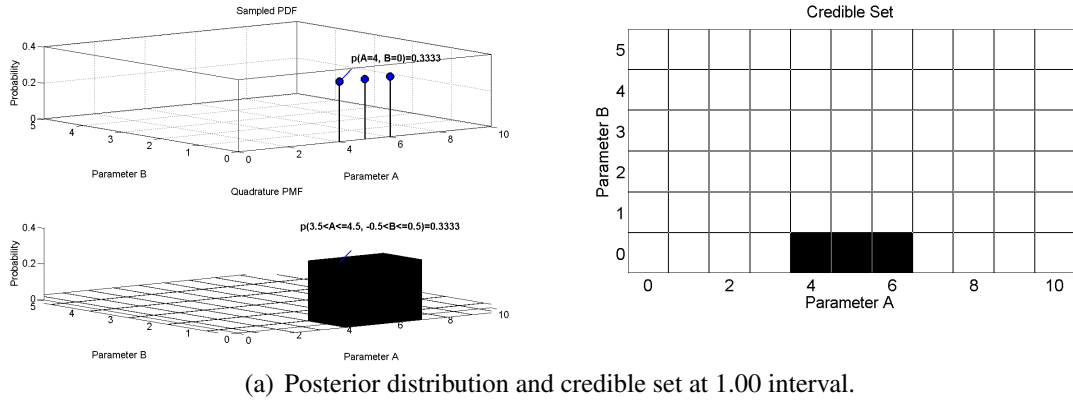
Figure 3.14: Left: PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$. Parameters used are $A = 5$, $B = 0$ and $N = 5000$. Right: Resulting credible regions. Note: interval size decreases from top to bottom.

51
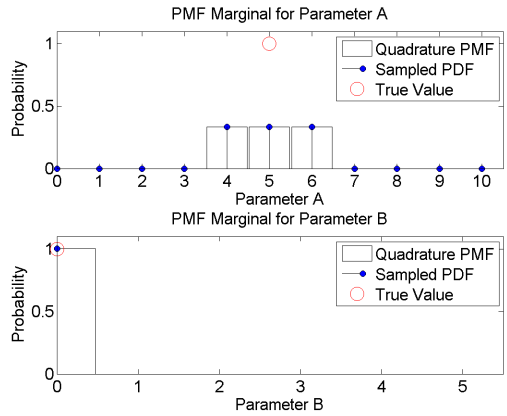
(a) Marginal distribution at 1.00 interval.
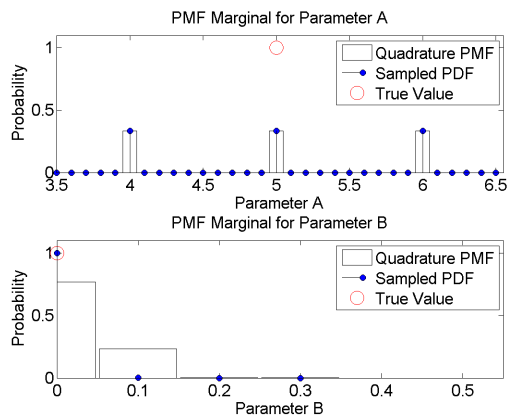
(b) Marginal distribution at 0.10 interval.
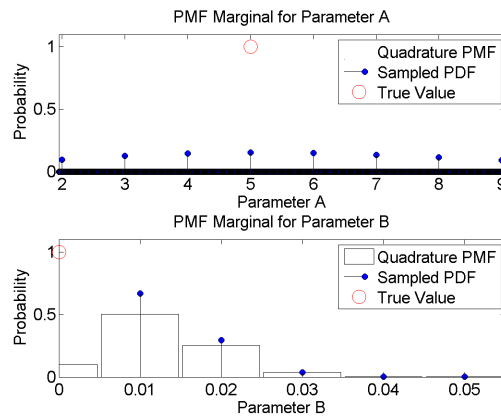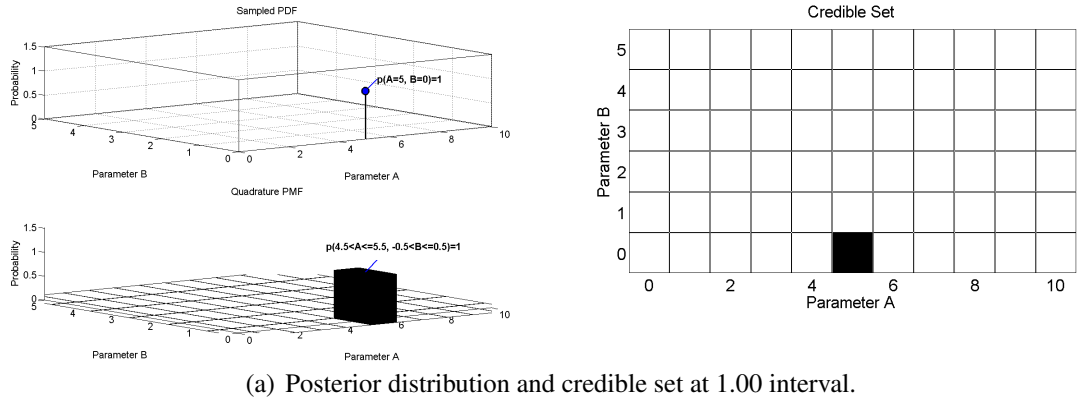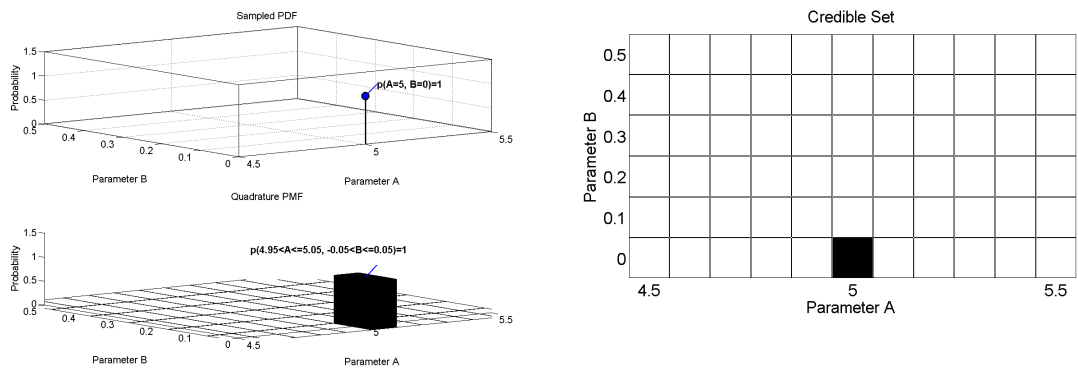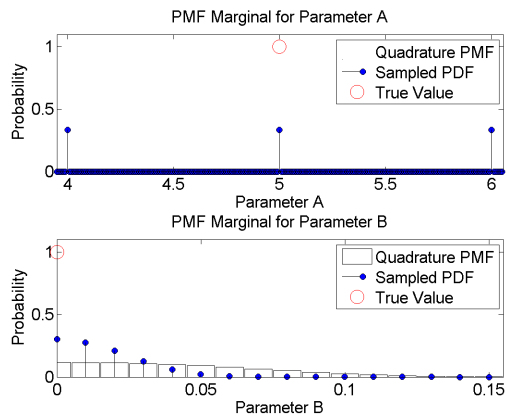

(c) Marginal distribution at 0.01 interval.

Figure 3.15: Marginal PMFs for the multiple parameter estimate of $\cos(2\pi An + B)$. Parameters used are $A = 5$, $B = 0$ and $N = 5000$.

3.6.6 Tabular Results.

Here we present the results which contrast the CSE Gaussian-Quadrature implementation with the traditional sampling method. Here, we use the total absolute error for the entire posterior distribution, computed as $TAE = \sum_{ALL\Theta} |p(\Theta|x)_s - p(\Theta|x)_q|$. We directly compare the quadrature method against the sampling method, under the assumption that the quadrature method has produced a more accurate posterior distribution (following the results of the validation section at the start of this chapter).

Tables 3.4 and 3.5 depict the results of the error analysis study. Notice that the TAE varies greatly depending on the interval. As with the validation results presented in Section 3.3, the TAE reduces in *most* cases as the interval is made finer, from 1.0 to 0.1 to 0.01. However, significant error is still present even at the finest zoom interval (note that in most cases the error is on the order of 10^{-1}). Also note that on a few occasions the TAE is as high as 2 (see $An + B$, Rayleigh, $N = 500$, Interval $= 0.10$) and other cases showing TAE around 1.11 (see cosine, Normal, $N = 500$, Interval $= 0.1$). The maximum TAE for a probability distribution is $TAE = 2$, implying that the sampling method when applied to real signals can result in a completely incorrect representation of the posterior distribution.

Finally, the entries labelled "NaN" are scenarios where the sampling method completely missed the region containing probability. The result is a distribution where the total probability is equal to zero, meaning the sampling interval was too coarse. The additional samples used in Gaussian-Quadrature integration alleviate this issue, enabling the CSE to extract the distribution when the sampling method cannot.

Table 3.4: Tabular results for signal $An + B$. Rows with TAE showing "NaN" are iterations where the sampling method missed the region containing probability.

Signal	Prior	N	Interval	TAE
$An + B$	Uniform	50	1.00	6.97e-34
$An + B$	Uniform	50	0.10	2.81e-01
$An + B$	Uniform	50	0.01	1.44e+00
$An + B$	Uniform	500	1.00	1.42e-320
$An + B$	Uniform	500	0.10	1.20e-03
$An + B$	Uniform	500	0.01	6.05e-01
$An + B$	Uniform	5000	1.00	0.00e+00
$An + B$	Uniform	5000	0.10	2.92e-38
$An + B$	Uniform	5000	0.01	7.68e-02
$An + B$	Normal	50	1.00	2.19e-35
$An + B$	Normal	50	0.10	1.89e-01
$An + B$	Normal	50	0.01	1.20e+00
$An + B$	Normal	500	1.00	0.00e+00
$An + B$	Normal	500	0.10	1.91e-06
$An + B$	Normal	500	0.01	6.08e-01
$An + B$	Normal	5000	1.00	0.00e+00
$An + B$	Normal	5000	0.10	1.12e-36
$An + B$	Normal	5000	0.01	1.26e-01
$An + B$	Rayleigh	50	1.00	NaN
$An + B$	Rayleigh	50	0.10	2.00e+00
$An + B$	Rayleigh	50	0.01	7.37e-01
$An + B$	Rayleigh	500	1.00	NaN
$An + B$	Rayleigh	500	0.10	2.00e+00
$An + B$	Rayleigh	500	0.01	7.03e-01
$An + B$	Rayleigh	5000	1.00	NaN
$An + B$	Rayleigh	5000	0.10	2.00e+00
$An + B$	Rayleigh	5000	0.01	7.71e-01

Table 3.5: Tabular results for signal $\cos(2\pi An + B)$. Rows with TAE showing "NaN" are iterations where the sampling method missed the region containing probability.

Signal	Prior	N	Interval	TAE
$\cos(2\pi An + B)$	Uniform	50	1.00	2.11e-03
$\cos(2\pi An + B)$	Uniform	50	0.10	5.61e-01
$\cos(2\pi An + B)$	Uniform	50	0.01	5.81e-01
$\cos(2\pi An + B)$	Uniform	500	1.00	1.71e-13
$\cos(2\pi An + B)$	Uniform	500	0.10	8.08e-01
$\cos(2\pi An + B)$	Uniform	500	0.01	8.56e-01
$\cos(2\pi An + B)$	Uniform	5000	1.00	1.20e-243
$\cos(2\pi An + B)$	Uniform	5000	0.10	4.69e-01
$\cos(2\pi An + B)$	Uniform	5000	0.01	8.49e-01
$\cos(2\pi An + B)$	Normal	50	1.00	1.45e-04
$\cos(2\pi An + B)$	Normal	50	0.10	7.27e-01
$\cos(2\pi An + B)$	Normal	50	0.01	8.74e-01
$\cos(2\pi An + B)$	Normal	500	1.00	5.78e-12
$\cos(2\pi An + B)$	Normal	500	0.10	1.11e+00
$\cos(2\pi An + B)$	Normal	500	0.01	1.04e+00
$\cos(2\pi An + B)$	Normal	5000	1.00	7.32e-218
$\cos(2\pi An + B)$	Normal	5000	0.10	9.17e-01
$\cos(2\pi An + B)$	Normal	5000	0.01	9.22e-01
$\cos(2\pi An + B)$	Rayleigh	50	1.00	NaN
$\cos(2\pi An + B)$	Rayleigh	50	0.10	1.96e+00
$\cos(2\pi An + B)$	Rayleigh	50	0.01	3.08e-01
$\cos(2\pi An + B)$	Rayleigh	500	1.00	NaN
$\cos(2\pi An + B)$	Rayleigh	500	0.10	1.96e+00
$\cos(2\pi An + B)$	Rayleigh	500	0.01	3.02e-01
$\cos(2\pi An + B)$	Rayleigh	5000	1.00	NaN
$\cos(2\pi An + B)$	Rayleigh	5000	0.10	1.96e+00
$\cos(2\pi An + B)$	Rayleigh	5000	0.01	3.03e-01

IV. PMF Extraction using Canonical Feature Models

Chapter 4 serves as an extension to the theory presented in Chapter 3. The current chapter focuses on application of the CSE to SAR phase history data, specifically focusing on canonical feature models [5]. The focus will be on estimator performance in terms of parameter estimation. The data presented is the result of applying the CSE to SAR scenes containing simulated phase history data of the canonical features.

4.1 Key Assumptions

- Received signals contain only a signal of interest and additive white Gaussian noise. This assumption is required by application of the univariate normal distribution as the likelihood function.

- Parameters are considered to be independent. This assumption is required in order to initialize prior distributions without a covariance matrix. Also, the parameters used for 3-D canonical features (size, location, orientation) are independent.

- Prior distributions are finitely bounded. This assumption makes numerically computing the denominator in the posterior distribution possible.

- Shape type is known a priori. The shape type must be known in advance so the CSE can apply the appropriate signal model based on Jackson's 3-D canonical features [4, 5]. It is possible to estimate shape type using methods such as spectrum parted linked image test (SPLIT) [16]. However, for the purposes of this thesis it is assumed that the shape type is already known.

4.2 Credible Set Estimator and Canonical Features

Applying the CSE to canonical feature models requires a few additions to the basic CSE algorithm presented in the previous chapter. The additions include generation

of canonical feature model phase history, flight path selection, adapting the estimator to representative frequencies, and initialization of canonical feature specific prior distributions.

4.2.1 Computation Method.

The computational method for canonical features follows the same form as the simple signals presented in the previous chapter. The likelihood function in the canonical feature model case is computed as

$$p(\mathbf{x}|\boldsymbol{\Theta}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - S_n(\boldsymbol{\Lambda}, \boldsymbol{\Theta}))^2\right], \tag{4.1}$$

where $S_n(\boldsymbol{\Lambda}, \boldsymbol{\Theta})$ is the canonical feature model containing parameters $\boldsymbol{\Theta}$ computed along flight path $\boldsymbol{\Lambda}$. Computation of Equation (4.1) requires that the shape type be known in advance in order to apply the correct feature model. The remainder of the calculation of the posterior distribution follows the derivation presented in the previous chapter.

4.2.2 Algorithm.

Application of the CSE to canonical features requires the following modifications to the estimator presented in Chapter 3:

- Prior distributions are initialized to match the specific shape and parameters to be estimated (for example, the sphere has a radius, but not length or height). The following sections detail how the prior distributions are initialized in the presented research.

- The shape type is determined in advance. It is possible to use the SPLIT algorithm to determine what shape types are present in a scene [16]. For the purposes of the current chapter, it is assumed that the shape type is known a priori, without using SPLIT.

The algorithm as implemented in this chapter is presented in Figure 4.1.
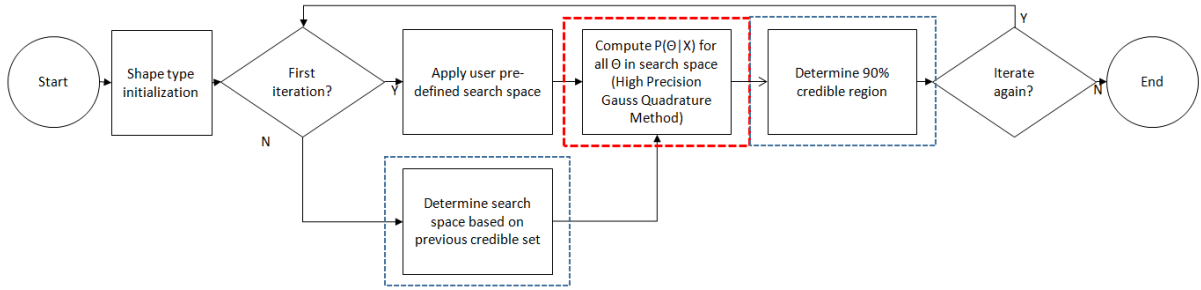
Figure 4.1: Flowchart depicting the CSE implementation for use with canonical features. Red box indicates parallel process. Dashed blue box indicates updates to Rademacher's algorithm.

4.2.3 Simulation Parameters.

Table 4.1 details the parameters used in the presented simulations. We choose the parameters to closely align with previous research as well as to simulate a SAR scene that is representative of real-world radar systems. We implement CSE on a scene using "HH" polarization. It is possible to use all polarization types (HH,VV,HV), however this results in three times as many observations.

Table 4.1: Radar parameters used in the canonical features CSE implementation.

Parameter	Value
Center Frequency (f_0)	300 MHz
Bandwidth (BW)	100 MHz
Flight Path	Varies by shape
Number of Frequency Bins	128
Signal-to-Noise Ratio	30 dB
Polarization	HH
Prior	Uniform($\Theta - 1, \Theta + 1$)

4.2.4   Flight Path Selection.

Crosser [4] presented a method of flight path selection for canonical feature extraction. Shapes such as the top-hat and plate require sufficient azimuth and elevation diversity to determine the correct specular response [4]. Crosser presents three flight paths with sufficient azimuth and elevation diversity to estimate each feature reliably. Each flight path consists of two segments. Figure 2.4 depicts the three flight paths. Table 4.2 depicts the association between features and flight paths. The up-down flight path is used to estimate the plate, dihedral and trihedral as it provides coverage for specular response. The disjoint flight path is used to estimate the sphere and top-hat as it provides the best radius estimate. Finally, the two-pass flight path is used to estimate the cylinder as it provides limited specular response along the azimuth while providing sufficient diversity in elevation for the radius estimate.

Table 4.2: Flight paths used for canonical feature estimation.

	Up-down	Disjoint	Two-pass
Plate	X		
Dihedral	X		
Trihedral	X		
Sphere		X	
Cylinder			X
Top-Hat		X	

4.2.5   Prior Distribution Initialization.

The presented research considers a simple set of uniform prior distributions. The key assumption when initializing the prior distributions is that the M parameters are independent, effectively making the total distribution a M-th dimensional product of the marginal distributions. The application of prior distributions used in this thesis is a simple

case in which all marginal distributions are defined as uniform distributions. Utilizing all uniform distributions is ideal when no prior knowledge of associated probabilities is available. Also, the uniform distribution can be tailored by the user to adjust the finite bounds of the posterior distribution, matching well with the key assumptions of the presented research. For the presented results, the prior distribution is defined as Uniform($\Theta - 1, \Theta + 1$).

4.3 Test Scenes

The presented research focuses on seven test scenes utilizing each of the canonical features. The presented test scenes are selected to maintain continuity with previous publications [4, 5, 13]. Table 4.3 details the test scenes and associated parameters. Two search intervals are selected for each scene: coarse (0.5m) and fine (0.25m).

Table 4.3: Canonical feature test scenes.

Shape	X (m)	Y (m)	Z (m)	L (m)	H (m)	r (m)	Roll (deg)	Pitch (deg)	Yaw (deg)
Plate	-0.30	0	0.30	0.60	0.30	-	0	-30	0
Dihedral 1	0.25	0.50	0	0.60	0.20	-	0	0	0
Dihedral 2	0	1.10	2.50	0.60	0.30	-	-10	0	0
Trihedral	0.75	0	0.50	-	0.18	-	0	0	0
Sphere	-1.50	-1.00	0	-	-	0.75	0	0	0
Cylinder	0.50	2.00	0	0.50	-	1.50	0	0	-5
Top Hat	0.25	-1.50	0	-	0.40	0.60	0	0	0

4.3.1 Computational Complexity.

Extension of the CSE to M-th order parameter estimates results in significant computational requirements. The computation of each parameter permutation for phase history vector of length N includes: generation of test phase history data for that permutation using N absolute values, $2N$ exponential calculations, and $4N$ add-multiplies.

The total computations required for all parameter permutations becomes

$$\#_{computations} = N_l^M \prod_{m=1}^{M} \Theta_i, \qquad (4.2)$$

where N_l is the order of the Legendre polynomial used for Gauss-Quadrature integration, M is the number of parameters to be estimated, and Θ_i is the number of grid points along each dimension to be evaluated. On a 64-bit machine, the resulting memory requirements to store the posterior distribution as a double precision floating point becomes

$$\text{MEMORY} = 16 \prod_{m=1}^{M} \Theta_i \text{bytes}. \qquad (4.3)$$

In order to estimate five parameters, where each parameter is sampled at 50 points requires 5GB of memory to store the posterior distribution. Increasing this to six parameters would require 250GB of memory, making direct computation of the posterior distribution impossible on current computers.

4.4 Results and Discussion

The following subsections detail the results of the CSE when applied to canonical features. The first subsection presents the data in a tabular format, while the following subsections present the data in graphical form. Note, the posterior distributions are presented as marginal distributions in order to visualize the *M* dimensional probability space. For all figures, the bars represent the posterior distribution estimated using CSE, and the red stems represent the true parameter values. Tick marks on the X-axis indicate regions in the search space where zero probability was found. Table 4.1 details the radar parameters used to generate the results in this section.

4.4.1 Tabular Results.

Table 4.4 contains tabular results for the CSE implementation using canonical features. We present the MAP values for each parameter to compare against the true parameters. Analysis and discussion for each scene is discussed in the sections below.

Table 4.4: Tabular results for the canonical feature test scenes. Data includes shape, interval (Int.), true parameter value (X,Y,etc.) and MAP estimate for each parameter ($\hat{X},\hat{Y}$,etc.).

Shape	Int.	$X(m)$	$\hat{X}(m)$	$Y(m)$	$\hat{Y}(m)$	$Z(m)$	$\hat{Z}(m)$	$L(m)$	$\hat{L}(m)$	$H(m)$	$\hat{H}(m)$	$r(m)$	$\hat{r}(m)$
Plate	0.5	-0.30	-0.25	0	-0.25	0.30	0.25	0.60	0.75	0.30	0.25	-	-
	0.25	-0.30	-0.375	0	-0.125	0.30	0.375	0.60	0.625	0.30	0.375	-	-
Dihedral 1	0.5	0.25	0.25	0.50	0.25	0	0.25	0.60	0.25	0.20	0. 25	-	-
	0.25	0.25	0.375	0.50	0.375	0	-0.125	0.60	0.625	0.20	0.375	-	-
Dihedral 2	0.5	0	0.75	1.10	1.25	2.50	1.25	0.60	0.75	0.30	0.25	-	-
	0.25	0	1.125	1.10	1.125	2.50	1.125	0.60	0.625	0.30	0.375	-	-
Trihedral	0.5	0.75	0.75	0	-0.25	0.50	0.25	-	-	0.18	All	-	-
	0.25	0.75	0.625,0.875	0	-0.125	0.50	0.375,0.625	-	-	0.18	0.375	-	-
Sphere	0.5	-1.50	-1.25	-1.00	-1.25	0	0.75	-	-	-	-	0.75	0.75
	0.25	-1.50	-1.625	-1.00	-1.125	0	0.875	-	-	-	-	0.75	0.375
Cylinder	0.5	0.50	-0.25	2.00	2.25	0	0.25	0.50	0.75	-	-	1.50	1.75
	0.25	0.50	0.125	2.00	2.125	0	0.125	0.50	0.625	-	-	1.50	1.625
Top Hat	0.5	0.25	0.25	-1.50	-1.25	0	0.25	-	-	0.40	0.25	0.60	0.75
	0.25	0.25		-1.50		0		-	-	0.40		0.60	

4.4.2 Plate.

The plate scene represents an "ideal" application of the CSE. Notice from Table 4.4 that the MAP estimates all fall within half of a search interval, aligning with the expected bias shown in Chapter 3. Figures 4.2 and 4.3 depict the marginal PMFs for the plate scene. Notice that the MAP estimates for each parameter contain the true value except the length at coarse zoom. The error here is corrected when zoomed into the fine level, shown in Figure 4.3. The initial error likely occurs because the true value falls almost directly between search bins, and noise has induced bias on the estimator. The result is the CSE selecting the search bin associated with $L = 0.25m$. However, note that the true value is still contained in the credible set, as the 90% set contains both the $0.25m$ and $0.75m$ search bins.
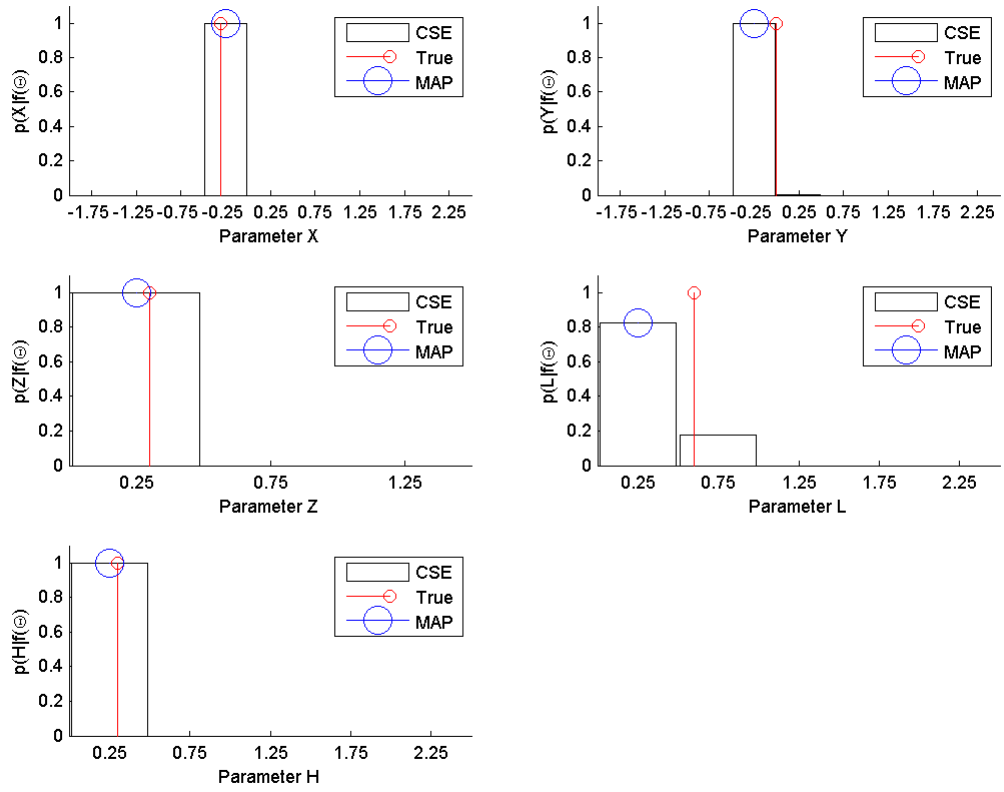
Figure 4.2: Marginal PMFs for the plate scene at coarse zoom.
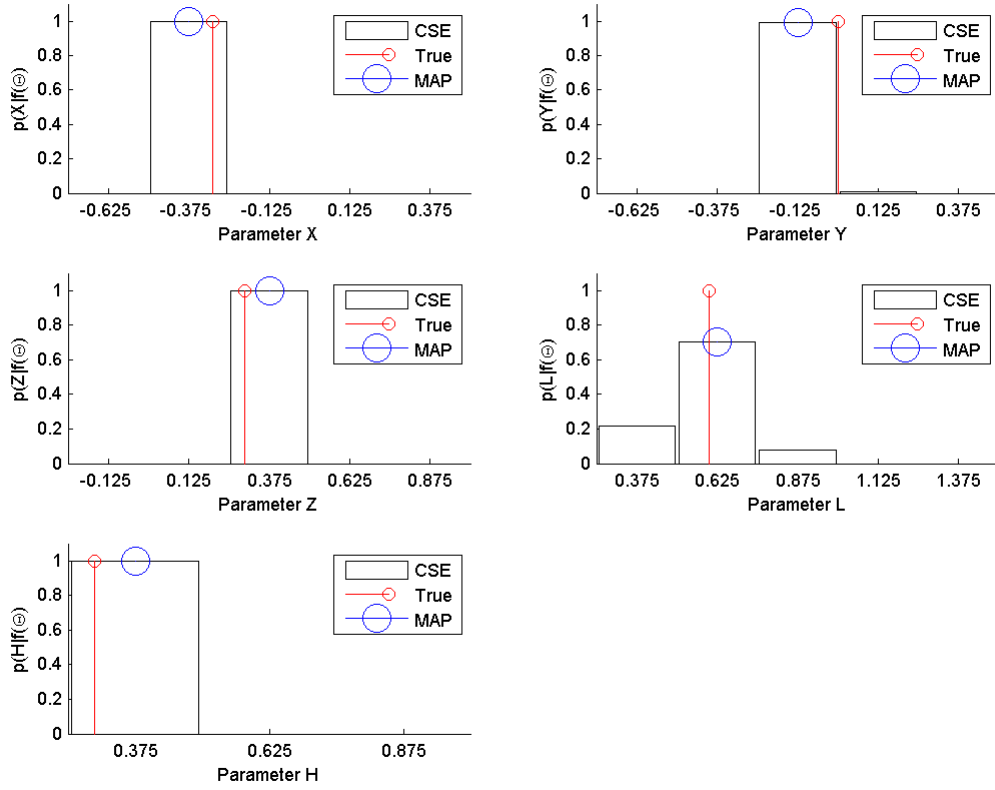
Figure 4.3: Marginal PMFs for the plate scene at fine zoom.

4.4.3 Dihedral 1.

The first dihedral scene presents some interesting results in the application of the CSE. Notice from Table 4.4 that the MAP estimates all fall within half of a search interval, again aligning with the expected bias. Figures 4.4 and 4.5 present the marginal distributions for the first dihedral scene. Note that the marginal distribution for height shows that all values are equally as likely. Recall the amplitude term for the dihedral [5], shown as

$$\mathbf{M_{\Gamma m}}(\mathbf{k}, \mathbf{\Lambda}; \mathbf{\Theta_m}) = \frac{jk}{\sqrt{\pi}} A \text{sinc}(kL \sin\tilde{\phi}\cos\tilde{\theta}) \times \begin{cases} \sin\tilde{\theta}; \tilde{\theta} \in [0, \frac{\pi}{4}] \\ \cos\tilde{\theta}; \tilde{\theta} \in [\frac{\pi}{4}, \frac{\pi}{2}] \end{cases}; \tilde{\phi} \in [-\frac{\pi}{2}, \frac{\pi}{2}], \tag{4.4}$$

where $A = 2LH$. Notice that the height term only exists as a scaling factor for the sinc. For the bistatic case, there is an additional sinc term containing H, which makes estimating height possible (see [5]). As such, it is impossible to directly estimate the height term for a dihedral using the CSE in the monostatic case. Note that the true solution for height is contained in the 90% credible set, which happens to contain the entire search area. A credible set consisting of the entire search area for a parameter does provide some useful insight, as it informs the user that the specific parameter does not impact the rest of the distribution, allowing the user to treat the parameter as a nuissance parameter. Nuissance parameters can be removed from the search space by setting an arbitrary value (since changing the parameter has no effect on the output of the estimator), effectively reducing the dimensionality of the search space. Treating parameters as nuissance parameters is an effective way to reduce computational complexity, if allowable.

Figure 4.4: Marginal PMFs for the first dihedral scene at coarse zoom.
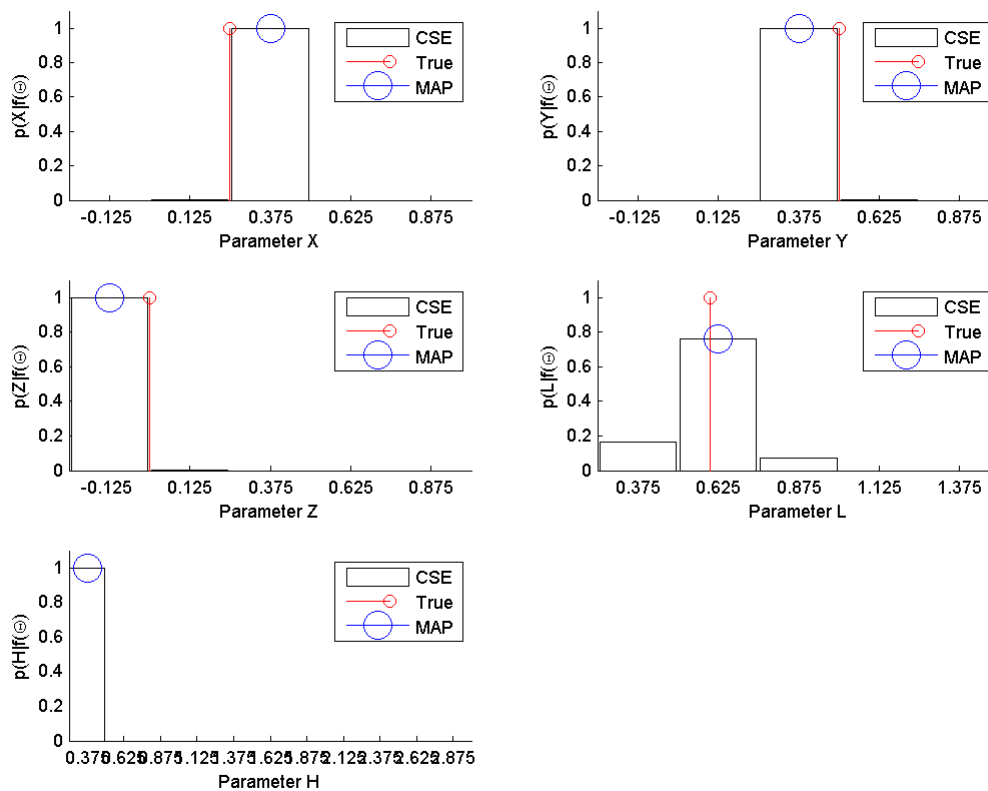
Figure 4.5: Marginal PMFs for the first dihedral scene at fine zoom.

4.4.4 Dihedral 2.

The second dihedral scene is intended to depict a different scattering geometry than the first dihedral. The first dihedral is oriented along the surface plane, resulting in a double bounce scattering response. The second dihedral is oriented at a slight angle (10°), inducing some single bounce scattering response [17]. Notice from Table 4.4 that the MAP estimates all fall within half of a search interval, aligning with the expected bias shown in Chapter 3. Figures 4.6 and 4.7 present the marginal distributions for the first dihedral scene. Note that the slight angle of 10° adds just enough scattering response to estimate the height correctly.
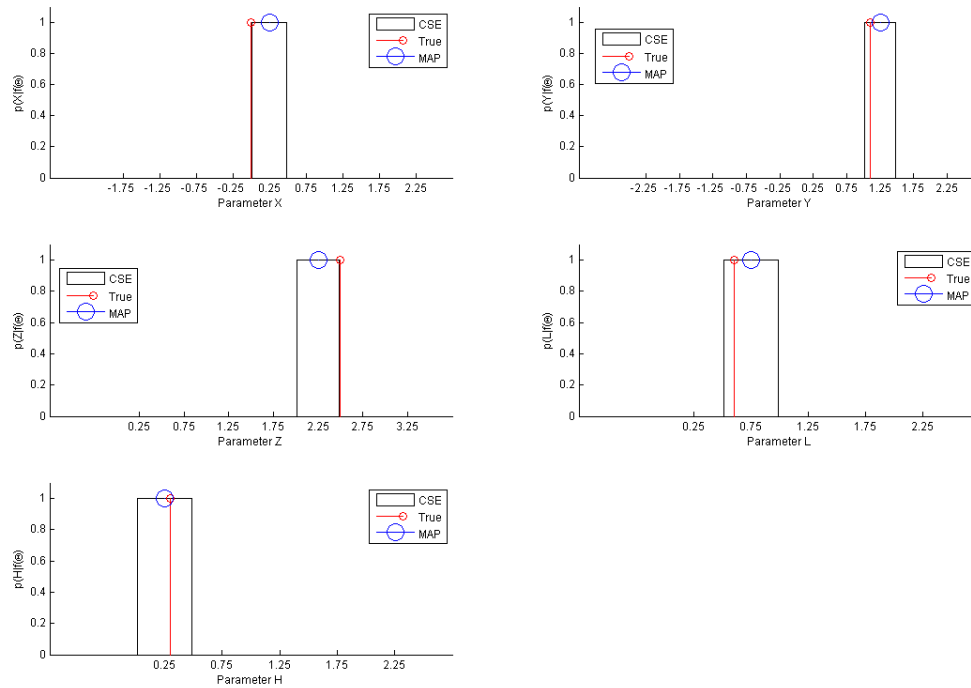
Figure 4.6: Marginal PMFs for the second dihedral scene at coarse zoom.
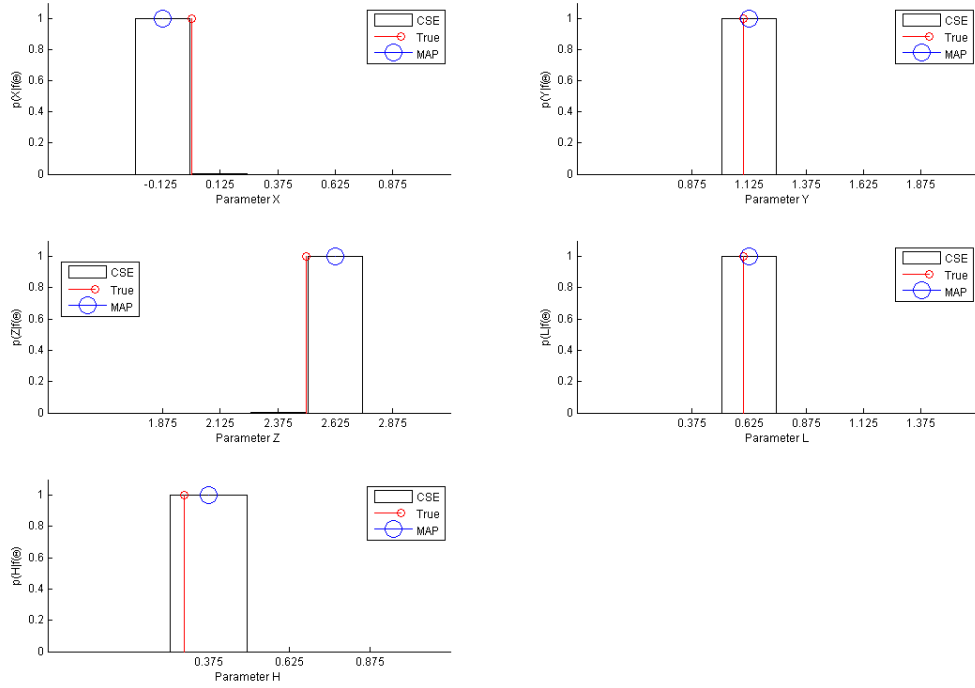
Figure 4.7: Marginal PMFs for the second dihedral scene at fine zoom.

4.4.5 Trihedral.

The results for the trihedral scene indicate favorable performance for the CSE. Notice from Table 4.4 that the MAP estimates all fall within half of a search interval, aligning with the expected bias shown in Chapter 3. Figures 4.8 and 4.9 depict the marginal PMFs for the trihedral scene. Note that the marginal distribution for height shows that all values are equally as likely. The reason for this is the same as the dihedral cases above. Recall the equation for the amplitude response term of a trihedral [5], shown as

$$\mathbf{M}_{\Gamma m}(\mathbf{k}, \mathbf{\Lambda}; \mathbf{\Theta_m}) = \frac{jk}{\sqrt{\pi}} A \times \begin{cases} \sin(\tilde{\theta} + \frac{\pi}{4} - \tan^{-1}(\frac{1}{\sqrt{2}})); \ \tilde{\theta} \in [0, \tan^{-1}(\frac{1}{\sqrt{2}})] \\ \cos(\tilde{\theta} + \frac{\pi}{4} - \tan^{-1}(\frac{1}{\sqrt{2}})); \ \tilde{\theta} \in [\tan^{-1}(\frac{1}{\sqrt{2}}), \frac{\pi}{2}] \end{cases} \\ \times \begin{cases} -\cos(\tilde{\phi} - \frac{\pi}{4}); \ \tilde{\phi} \in [-\frac{\pi}{4}, 0] \\ \sin(\tilde{\phi} - \frac{\pi}{4}); \ \tilde{\phi} \in [0, \frac{\pi}{4}] \end{cases}, \tag{4.5}$$

where $A = 2\sqrt{3}H^2$. Notice that the height term is not present in any of the sine or cosine components for the monostatic case. For the bistatic case, there are additional sinc terms

69

containing *H*, which makes estimating the height possible. As such, it is impossible to directly estimate the height term for a trihedral using the CSE in the monostatic case. Note that the true solution for height is contained in the $\alpha = 90\%$ credible set, which happens to contain the entire search area. As mentioned in the dihedral cases, it is also possible to treat the height term in the trihedral case as a nuissance parameter, reducing the overall dimensionality of the computation to three.
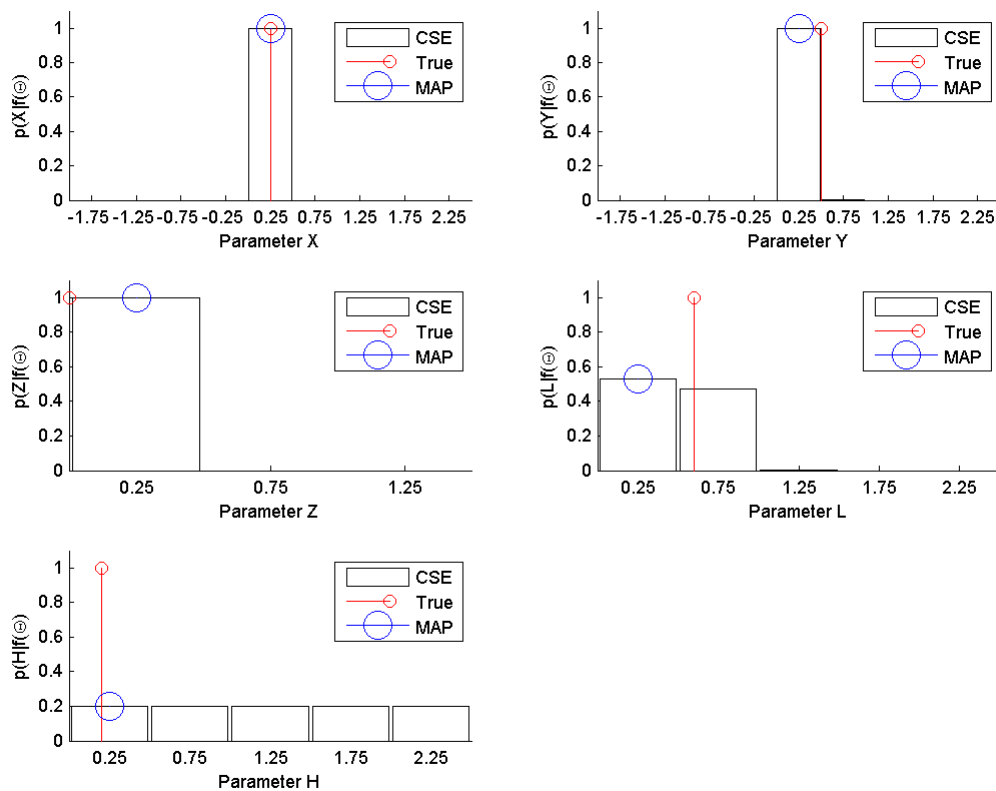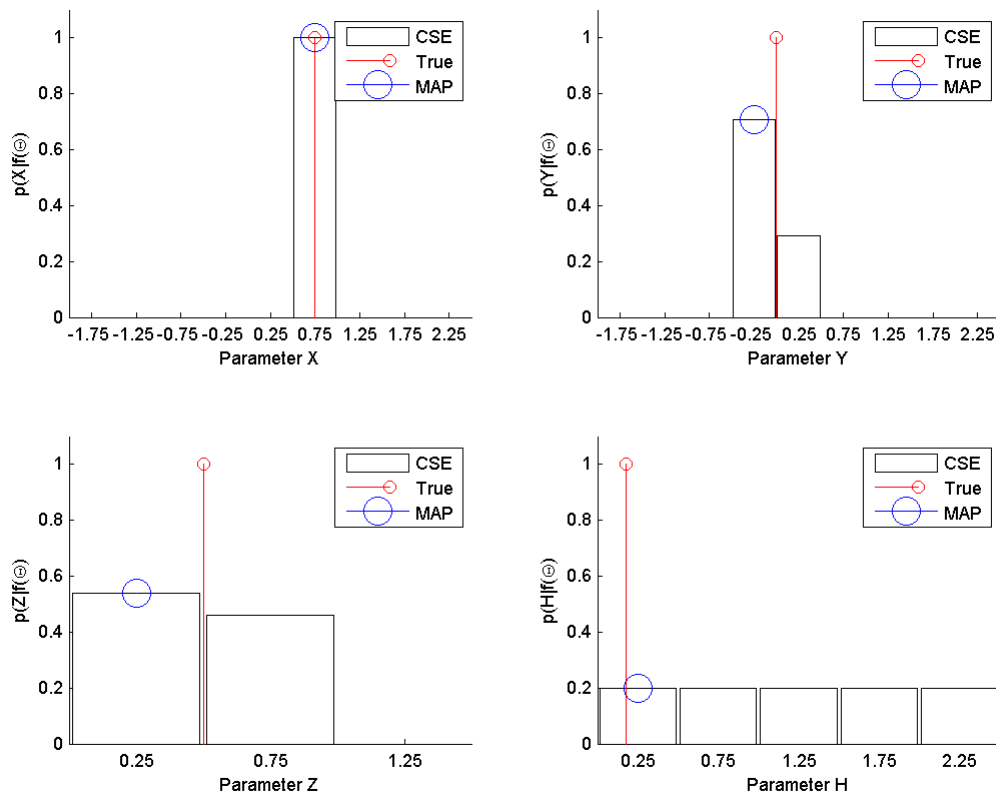

Figure 4.8: Marginal PMFs for the trihedral scene at coarse zoom.
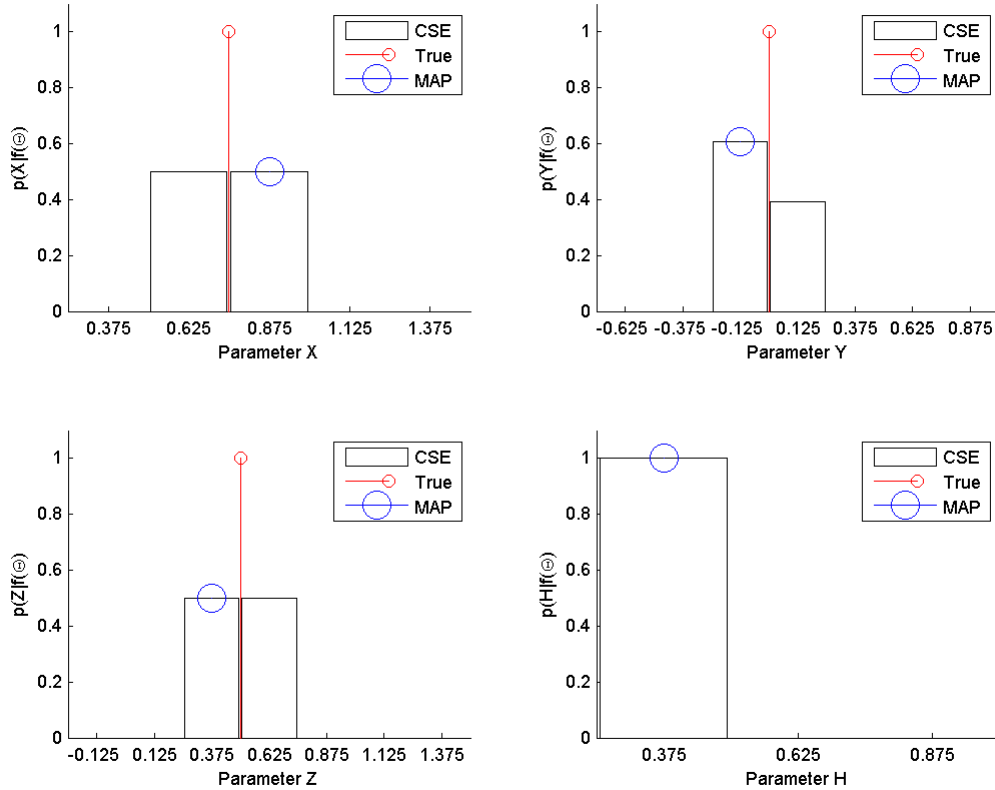
Figure 4.9: Marginal PMFs for the trihedral scene at fine zoom.

4.4.6 Cylinder.

The results for the cylinder scene demonstrate an issue with phase wrapping in the differential range term, $e^{jk\Delta R}$ and inadequate flight path extent. Notice from Table 4.4 that the MAP estimates all fall within half of a search interval, except for the estimate on X position. Figures 4.10 and 4.11 depict the marginal PMFs for the cylinder scene. Notice that all MAP estimates for each parameter contains the true value except the X location. Recall the differential range model for the cylinder [5], shown as

$$\Delta R = \Delta R_0 + \left(r \cos\left(\frac{\theta_t - \theta_r}{2}\right)(\cos\phi_t + \cos\phi_r)\right). \tag{4.6}$$

71

In the monostatic case, the equation reduces to

$$\Delta R = \Delta R_0 + 2r(\cos \phi). \qquad (4.7)$$

Phase wrapping in the differential range term makes location estimation difficult because the values recycle every 2π radians [4]. Also, the estimation error for X position is likely due to inadequate azimuth extent. Recall that layover causes objects above the ground plane to appear closer to the radar on the slant plane [18]. Since the cylinder is positioned $2m$ above the ground plane, it appears closer to the radar (negative x-axis). Recall, the method presented by Crosser partitions the flight path into two segments, generates two images, and uses geometric properties of layover to estimate the scattering center [4]. The CSE application presented in this thesis is incapable of producing a single PMF with two separate flight path segments. The CSE instead treats both segments of the flight path as one longer flight path. As such, the CSE requires more azimuth extent than the dictionary method presented by Crosser [4]. It is important to note that the CSE provides the probability of a parameter set (Θ) given observations ($\mathbf{x}$), and as such the distribution produced is correct for the given observations. However, greater azimuth extent is required in the observations for the MAP estimate and true parameter values to match.

Figure 4.10: Marginal PMFs for the cylinder scene at coarse zoom.

Figure 4.11: Marginal PMFs for the cylinder scene at fine zoom.

4.4.7 *Top-Hat.*

The top-hat scene represents another "ideal" application of the CSE. Notice from Table 4.4 that the MAP estimates all fall within half of a search interval, aligning with the expected bias shown in Chapter 3. Figures 4.12 and 4.13 depict the marginal PMFs for the top-hat scene. Notice that the MAP estimates for each parameter contain the true value. Note that the top-hat *X/Y* location is not affected by the layover since the scatterer is located on the ground plane.
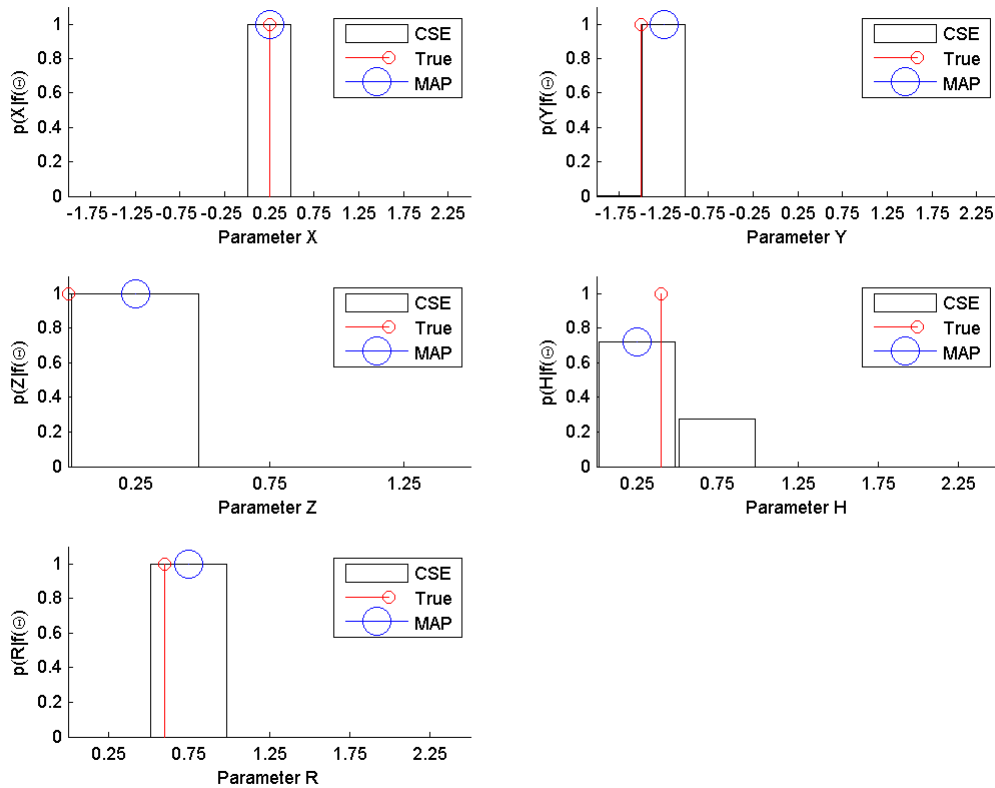
Figure 4.12: Marginal PMFs for the top-hat scene at coarse zoom.
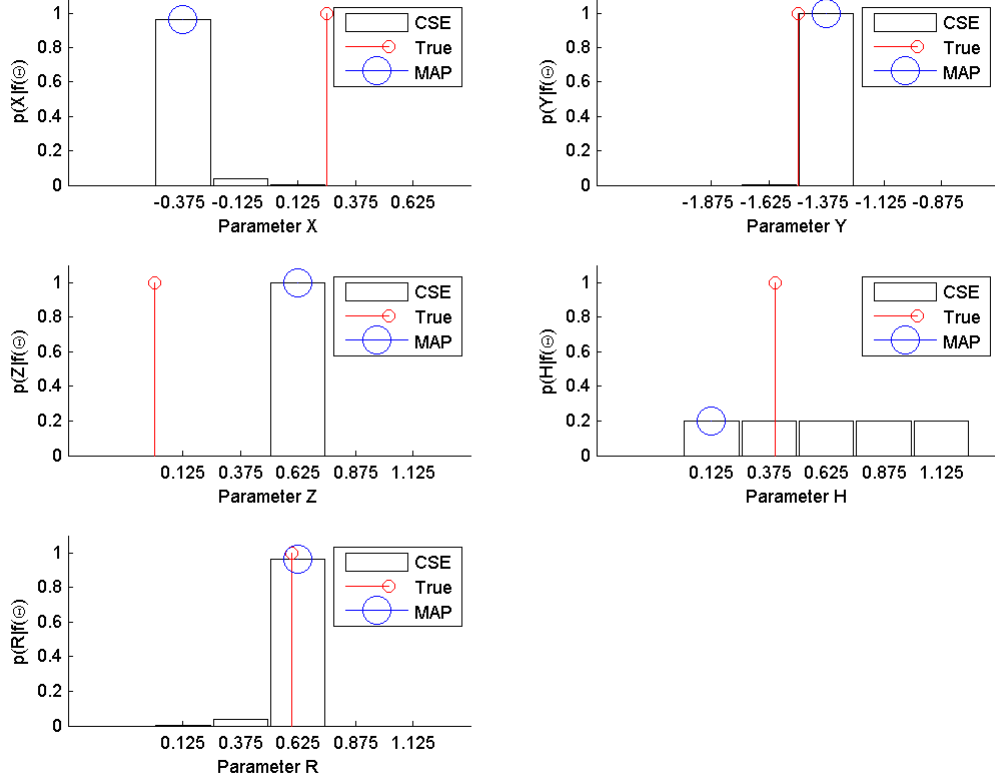
Figure 4.13: Marginal PMFs for the top-hat at scene fine zoom.

4.4.8   Sphere.

The results for the sphere scene demonstrate another issue with phase wrapping in the differential range term, $e^{jk\Delta R}$ and inadequate flight path extent. Notice from Table 4.4 that the MAP estimates for X/Y position fall within half of a search interval, but the Z location and radius estimates are incorrect. Figures 4.14 and 4.15 depict the marginal PMFs for the sphere scene. Notice the discrepencies in the Z location and radius estimates. Recall the differential range model for the sphere [5], shown as

$$\Delta R = \Delta R_0 + \left(r(\cos\theta_t + cos\theta_r)\cos\left(\frac{\phi_t - \phi_r}{2}\right)\cos\left(\frac{\theta_t + \theta_r}{2}\right) + r\sin\left(\frac{\theta_t + \theta_r}{2}\right)(\sin\theta_t + \sin\theta_r)\right).$$

(4.8)

For the monostatic case, the equation reduces to

$$\Delta R = \Delta R_0 + 2 \left[r(\cos\theta)^2 + r(\sin\theta)^2 \right]. \tag{4.9}$$

Notice that the only terms remaining in Equation (4.9) are radius (r) and elevation (θ), both of which effect the estimates of radius and height for a sphere. Crosser generates two images, using two different flight paths, and applies geometric layover properties to estimate both the radius and Z location of a sphere [4]. The CSE treats two segments of a flight path as one longer flight path. Since the radius and Z-location parameters are coupled in the data, the estimates radius and Z location are affected, indicating that the flight path contains inadequate elevation extent to produce a reliable estimate. A flight path with greater elevation extent is required to produce more accurate estimates. Again, it is important to note that the CSE provides the probability of a parameter set ($\mathbf{\Theta}$) given observations ($\mathbf{x}$), and as such the distribution produced is correct for the given observations. However, greater elevation extent is required in the observations for the MAP estimate and true parameter values to match.
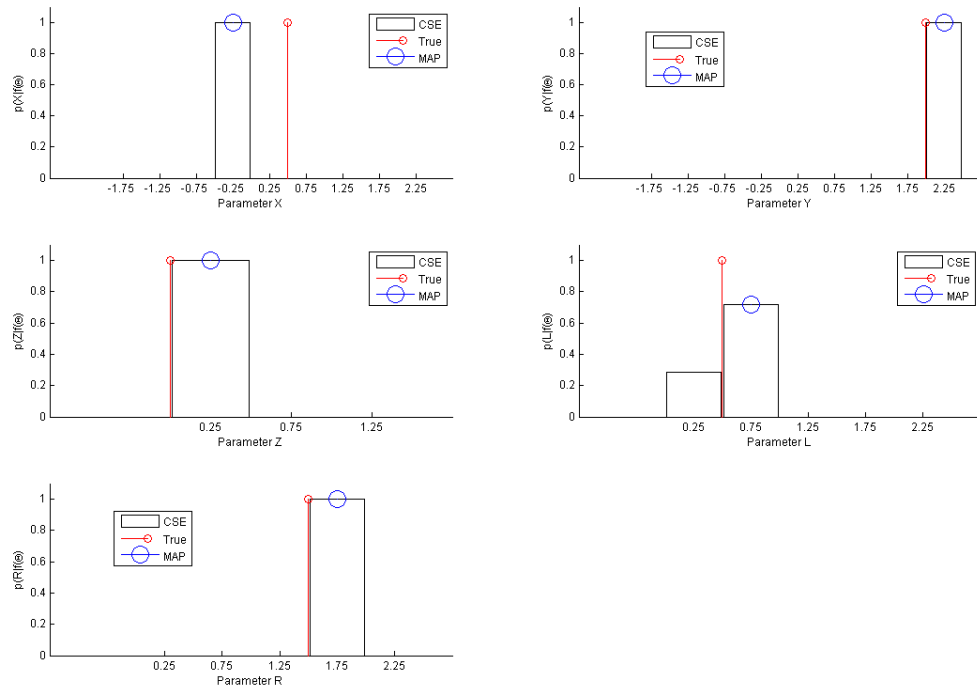
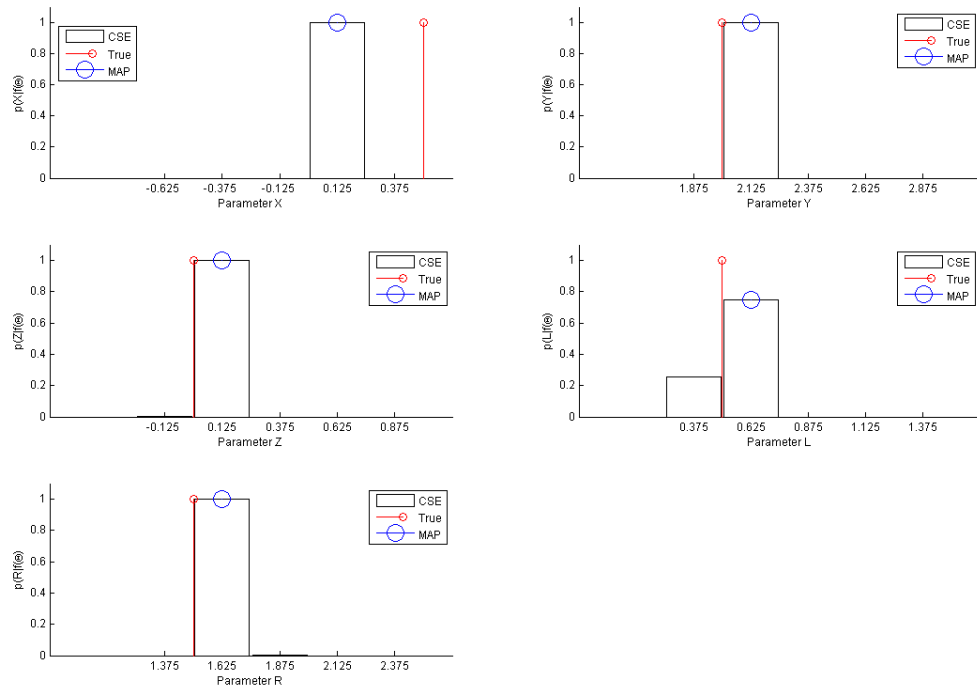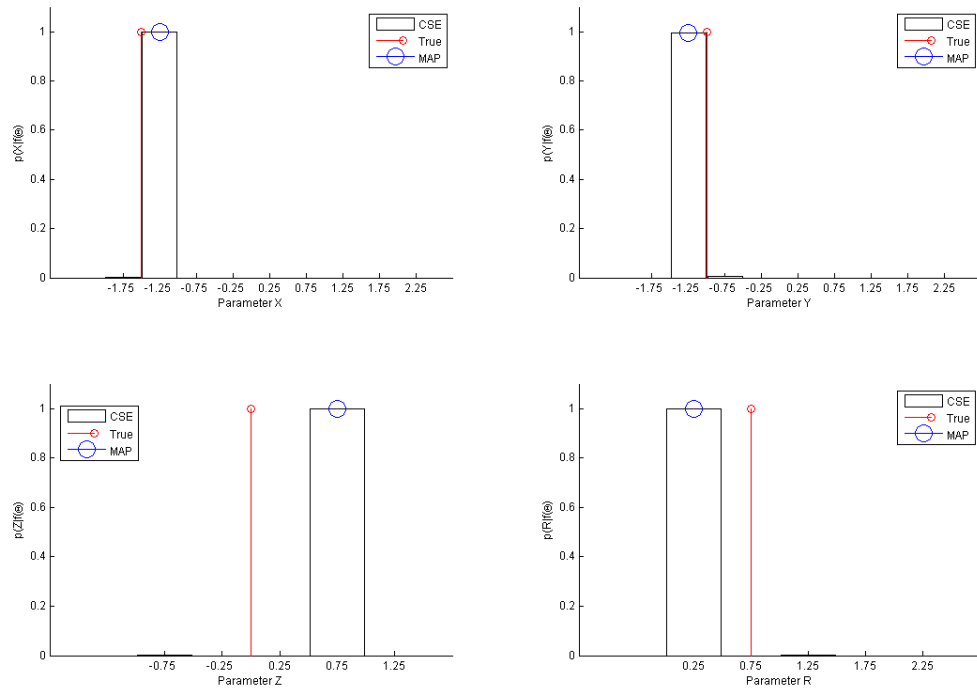Figure 4.14: Marginal PMFs for the sphere scene at coarse zoom.
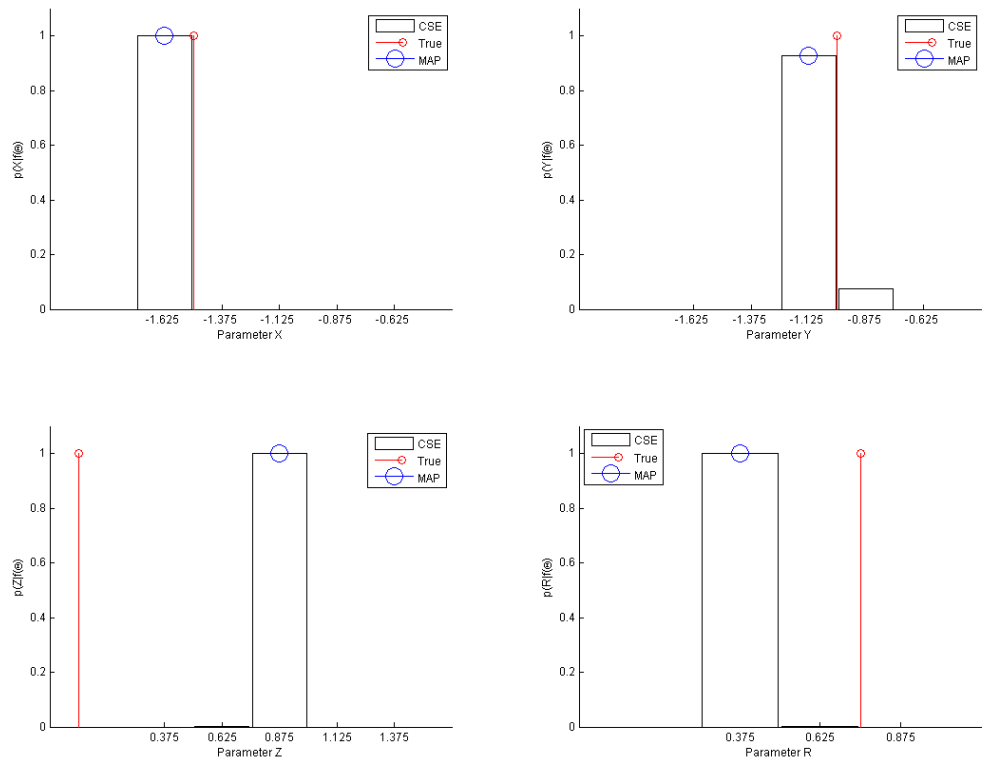
Figure 4.15: Marginal PMFs for the sphere scene at fine zoom.

4.5   Multiple Shapes in One Scene

In order to transition to SAR scenes containing real target data, it is pertinent to apply the estimator to a scene with multiple canonical features. To evaluate this concept, an additional scene containing a dihedral and cylinder are examined. Figure 4.16 depicts a SAR image containing the two shapes. Table 4.5 details the scene parameters. The CSE is applied twice to each scene, first utilizing the two-pass flight path to extract the cylinder, then utilizing the up-down flight path to extract the dihedral. Figures 4.17 and 4.18 depict the marginal PMFs associated with the cylinder and dihedral respectively.

Note that most of the parameter estimates are incorrect. The multiple shape implementation of the CSE presented in this section attempted to apply key assumptions from both the likelihood function and the CLEAN method [19]. Application of the likelihood function assumes the received signal contains only the signal of interest and noise. The CLEAN method assumes that objects separated in the image space are orthogonal in phase history. However, the likelihood and CLEAN assumptions contradict each other, therefore we must accept the assumptions of the likelihood function. In order to apply the CSE on a scene with multiple shapes, some form of data segmentation must be implemented in order to reduce the scene into multiple regions that contain one shape and noise. Chapter 5 proposes an experiment to implement the CSE on multiple shape scenes.

Table 4.5: Scene parameters for the multiple canonical feature test scene. Note that roll, pitch and yaw are fixed at zero degrees.

Shape	X (m)	Y (m)	Z (m)	L (m)	H (m)	r (m)
Dihedral	2.5	2.00	0	0.30	0.20	-
Cylinder	-2.50	-2.00	0	0.50	-	0.50
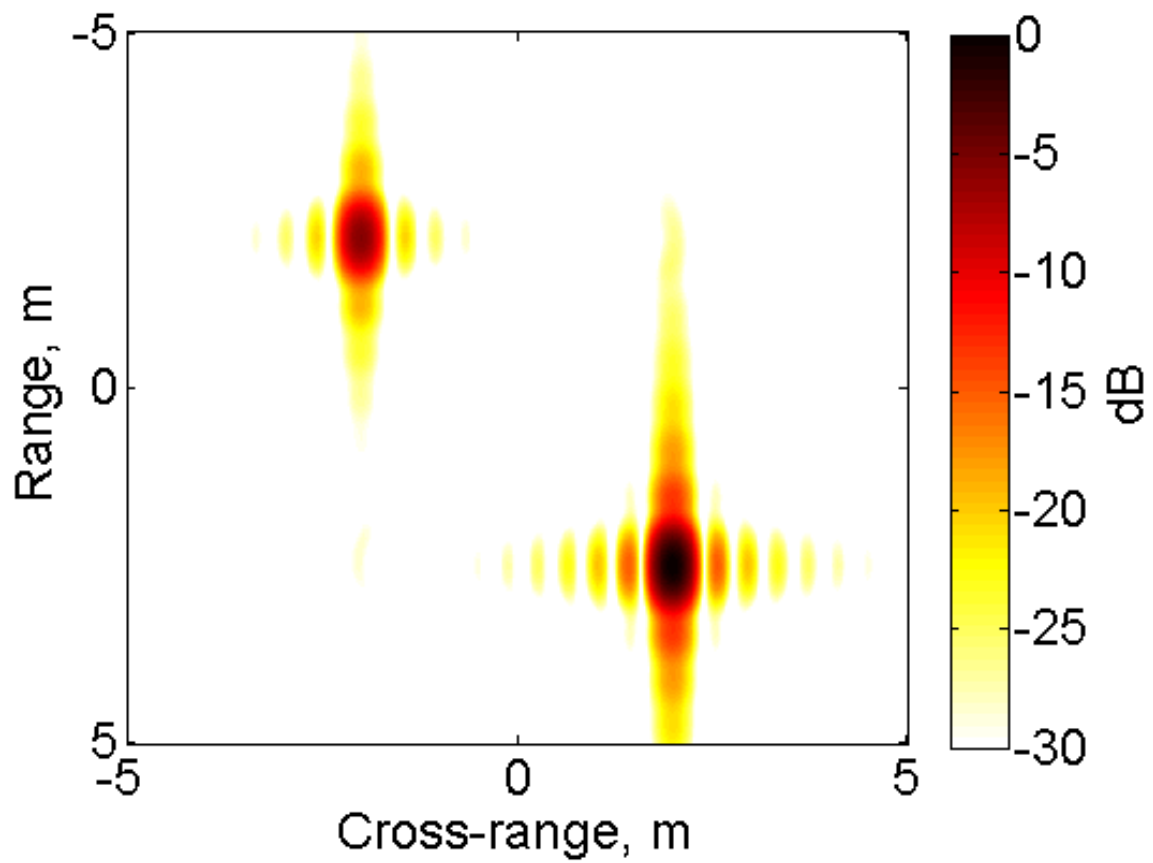
Figure 4.16: SAR image for multiple shape test scene. Note that the two shapes are sufficiently separated
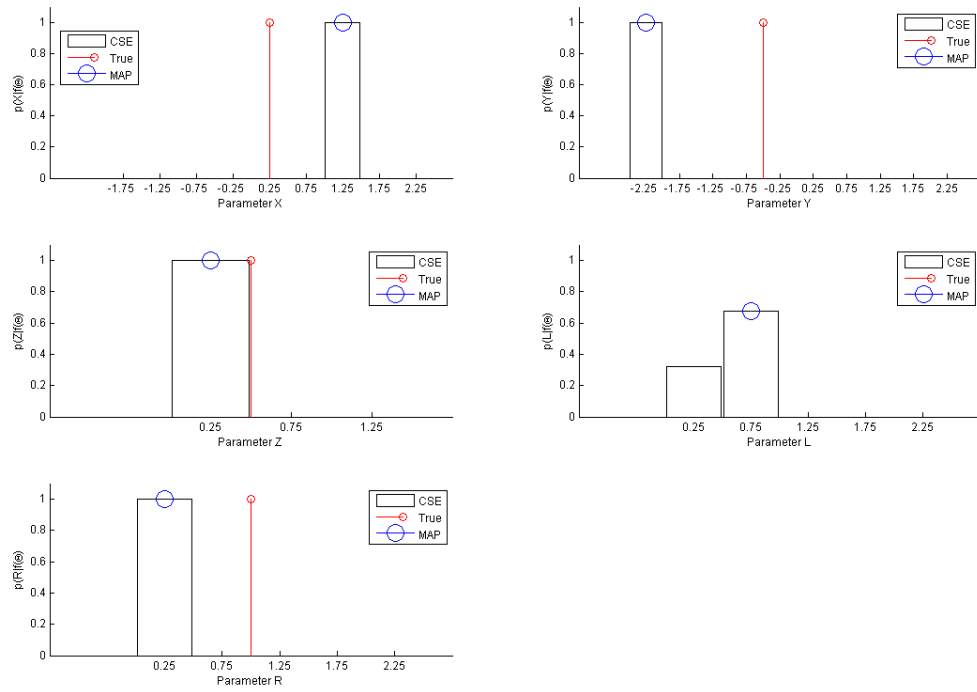
Figure 4.17: Marginal PMFs for the cylinder in the two shape scene at coarse zoom.
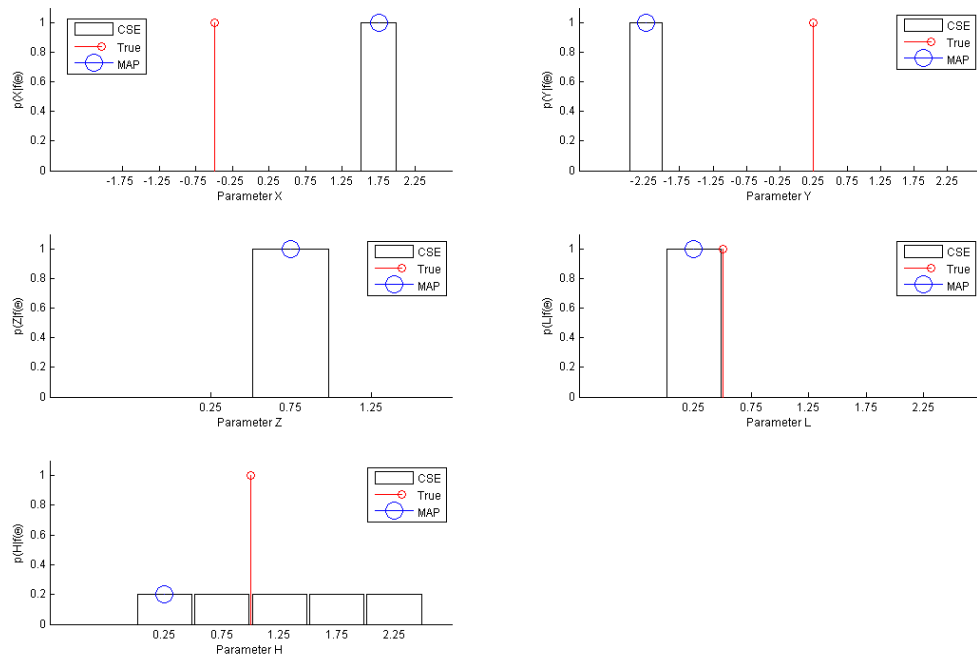
Figure 4.18: Marginal PMFs for the dihedral in the two shape scene at coarse zoom.

V. Conclusions and Future Work

The following sections detail future work and the conclusions from this thesis. We consider two areas in which CSE can be extended:

- Variable Zoom and Sampling: a potential method for reducing the computational complexity of CSE that involves randomly sampling the search space.

- Proposed Experiment: a potential method for applying CSE to SAR data containing more than one canonical feature.

This chapter will close with conclusions and final thoughts about CSE.

5.1 Future Work: Variable Zoom and Sampling

The CSE in its current form is incredibly computationally intensive. The canonical features results generated in Chapter 4 take over eighteen hours to compute using the HPC. Also, we have demonstrated that sampling too coarsely initially severely impacts estimator performance, and is often the cause of missing the true parameter value. One potential solution would be to randomly sample the space. Figure 5.1 depicts the concept.
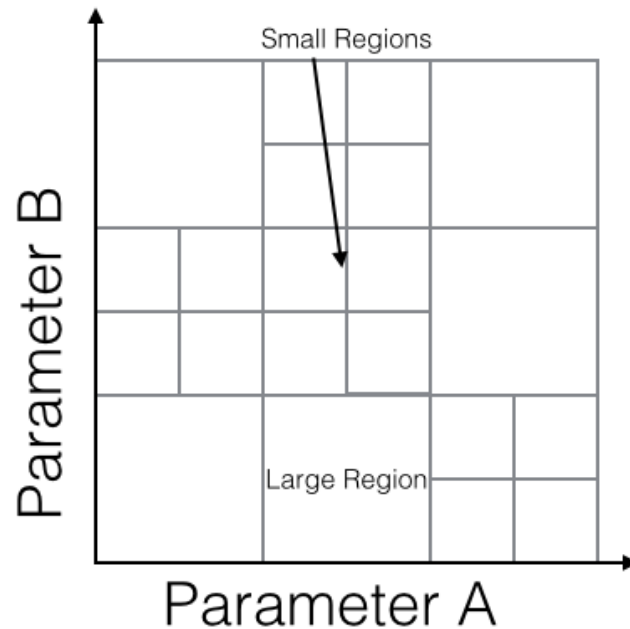
Figure 5.1: Marginal PMFs for the sphere scene at coarse zoom.

5.2 Future Work: Proposed Experiment

This section serves as an "proposed experiment" applying the CSE to more realistic SAR phase history data with multiple shapes.

5.2.1 Key Assumptions.

- Received signals contain only a signal of interest and additive white Gaussian noise. This assumption is required by application of the univariate normal distribution as the likelihood function.

- Parameters are considered to be independent. This assumption is required in order to initialize prior distributions without a covariance matrix.

- Prior distributions are finitely bounded. This assumption makes numerically computing the denominator in the posterior distribution possible.

- Individual canonical features can be segmented from the data. This assumption enables the use of the univariate normal distribution for the likelihood function.

5.2.2 Algorithm.

Application of the CSE to canonical features requires the following modifications to the estimator presented in Chapter 4:

- The SPLIT algorithm is used to determine which canonical features are present in the scene [16].

- The phase history data for each feature is segmented resulting in only the data associated with the feature and noise.

- The CSE as presented in Chapter 4 is used in multiple iterations to provide posterior distributions for each shape present in the scene.

5.2.3 Shape Type Initialization Using SPLIT.

The presented experiment explores the possibility of fitting a set of features to collected phase history data. Work by Jackson and Fuller [5, 16] has shown that vehicles exhibit two dominant types of scatters, namely the cylinder and dihedral. Layover from the roof line of the vehicle often results in a series of cylinders, while the double bounce response of the vehicle side and the ground plane result in a series of dihedrals. Figure 5.2 depicts this concept.
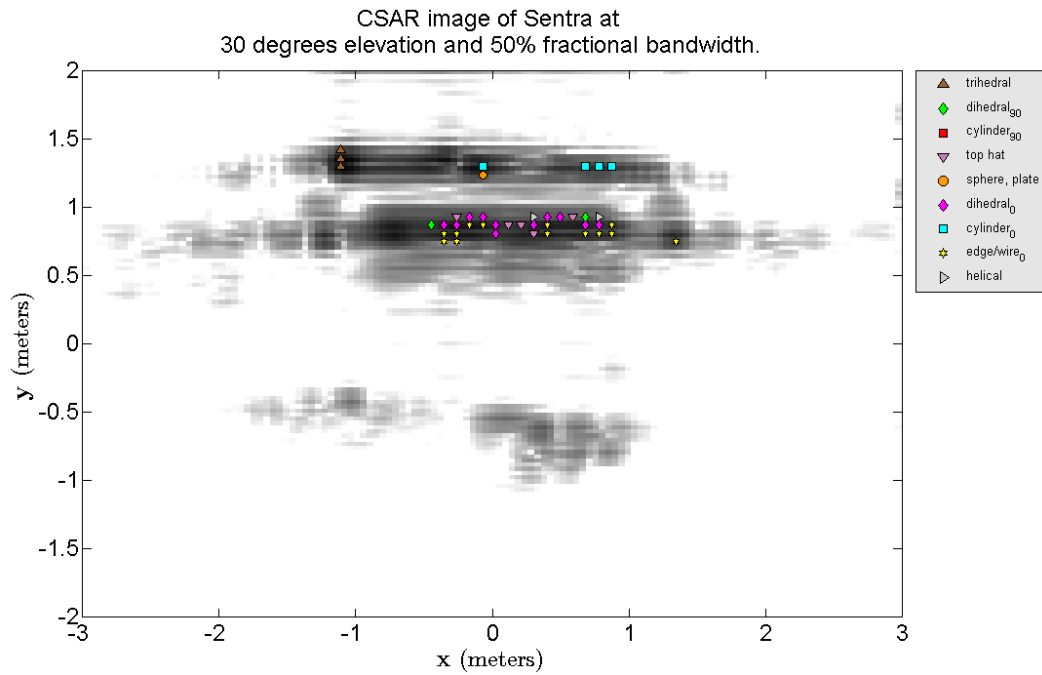
Figure 5.2: SAR image of a Nissan Sentra with canonical features extracted by the SPLIT algorithm.

5.2.4 *Experiment Summary.*

A notional algorithm is presented in Figure 5.3. To ensure reliability, the algorithm should be tested first against a synthetic scene containing multiple canonical features. Once the algorithm is validated, it can then be tested against more realistic data, such as the civilian vehicles (CV) Domes data set [20].
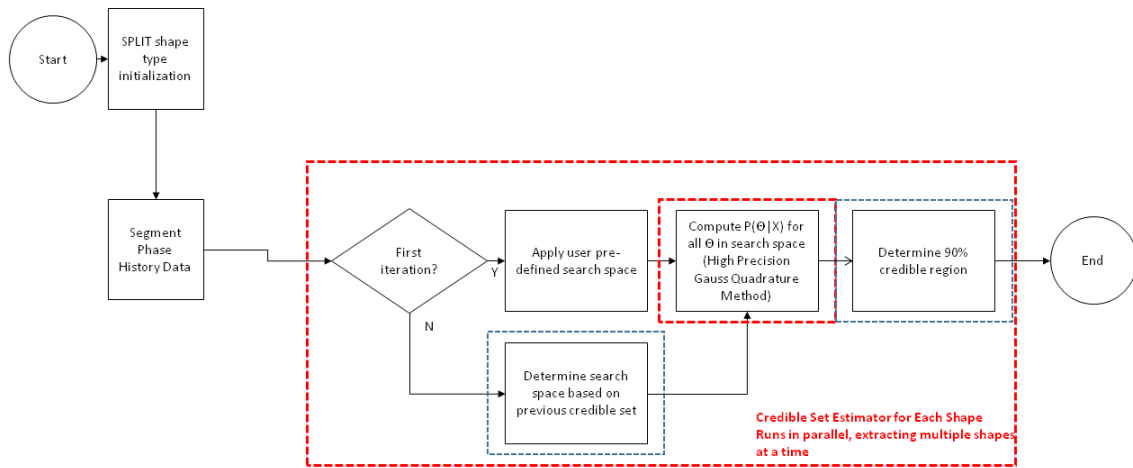
Figure 5.3: Flowchart depicting the CSE implementation for use with CV Domes data. Red box indicates parallel process. Dashed blue box indicates updates to Rademacher's algorithm.

5.3 Conclusions

This research has focused validation of the credible set estimation scheme, extension of the CSE to large observation sets and higher dimensions, incorporation of Crosser's work on flight path selection [4], and evaluation of the new CSE as presented in this thesis. The following sections detail the findings.

5.3.1 Alternative Computation Method.

Prior to deriving the formula presented in Equation (3.12), attempts were made to solve for the posterior distribution using a high precision float (HPF) tool. However, applying the HPF tool to the CSE resulted in expected run times of over two weeks (compared to the 18 hour run times using the formula presented in this research).

5.3.2   *Validation.*

Application of Gauss-Quadrature numerical integration techniques to the computation of the posterior distribution results in significant reductions in TAE. The results presented in this thesis show that the quadrature implementation results in almost negligible TAE (on the order of 10^{-5} or less when using a $Nl = 5$ integration polynomial), which is significantly less than the sampling method (where error is on the order of 10^{-1}). In terms of MAP performance, we determined that the CSE implementation results in exactly the same amount of error as the sampling method, demonstrating that the computation of the entire posterior distribution does not degrade the performance if a sole MAP estimate is needed.

It is important to note that the method presented in this thesis applies to any estimation problem requiring Bayesian methods, not just the simple signals and SAR models used in this thesis. Also, the method is not limited to a specific likelihood function. The method can be implemented for any combination of likelihood and prior distributions.

5.3.3   *Extension.*

The presented formula for the computation of the Bayesian posterior distribution enables numerical evaluation for large observation sets (greater than 6000 observations). The canonical feature examples generated in this thesis contain approximately 7,300 observations. The equation presented in Chapter 3 enables computation of posterior distributions where the $\log p(\mathbf{\Theta}|\mathbf{X})$ is on the order of e^{-730}. The validation section of this thesis has shown that estimates of posterior distributions with more than 5000 observations may require a higher-order integration polynomial, depending on the desired accuracy.

5.3.4   *Incorporation.*

Incorporation of the flight paths presented by Crosser [4] into CSE results in acceptable performance for the extraction of posterior distributions. However, in terms of MAP performance, it is important to note that additional azimuth/elevation extent may

be required in addition to the requirements presented by Crosser. However, it is important to note that the distributions extracted by the CSE are the probability of a parameter set given a set of observations, $p(\mathbf{\Theta}|\mathbf{X})$. The CSE produces accurate representations of the posterior distributions for a set of observations. However, MAP performance decreases (in both the CSE and standard MAP estimators) when estimating canonical feature parameters without sufficient data extent.

5.3.5 *Evaluation.*

The new CSE is able to extract posterior distributions for features with up to five parameters. This thesis presents formulas for the computational complexity required when implementing the CSE. Estimation of features with six or more parameters requires splitting each large distribution into multiple smaller distributions to avoid large computational overhead.

5.3.6 *Closing Thoughts.*

The CSE presented in this thesis is capable of estimating posterior distributions for high-dimensional problems with large numbers of observations. The estimation scheme, as implemented, is ideal for radar, communications and other statistical problems where retaining the entire posterior distribution and associated confidence intervals is desirable. As such, CSE provides a powerful and flexible alternative to other estimation schemes.

Appendix A: Additional Monte Carlo Data

Table A.1: Sample Monte Carlo data from the $N = 1$ observation at 1.0 interval case. Legendre polynomial $Nl = 3$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
5.1974	5	5	5	0.19736	0.19736	0.19736	0.00054046	0.20152
5.7168	5	5	5	0.71679	0.71679	0.71679	0.0027565	0.2965
5.9023	6	6	6	0.097669	0.097669	0.097669	0.0010208	0.22191
5.2568	5	5	5	0.25678	0.25678	0.25678	0.0026833	0.29954
7.6388	8	8	8	0.36121	0.36121	0.36121	0.0016324	0.16905
7.499	7	7	7	0.49904	0.49904	0.49904	0.0019216	0.20035
5.5914	6	6	6	0.40861	0.40861	0.40861	0.00095225	0.21909
6.3891	7	7	7	0.61095	0.61095	0.61095	0.0021419	0.27512
6.8639	7	7	7	0.13607	0.13607	0.13607	0.0022373	0.27983
7.8521	8	8	8	0.1479	0.1479	0.1479	0.0023404	0.24573
6.0783	6	6	6	0.078288	0.078288	0.078288	0.0026343	0.29794
6.4362	7	7	7	0.56379	0.56379	0.56379	0.00064146	0.20737
5.648	6	6	6	0.35205	0.35205	0.35205	0.00078288	0.2126
7.4158	7	7	7	0.41581	0.41581	0.41581	0.0027598	0.29803
7.2786	7	7	7	0.27862	0.27862	0.27862	0.0024848	0.26166
7.8279	8	8	8	0.17206	0.17206	0.17206	0.0024675	0.29086
6.2746	6	6	6	0.27458	0.27458	0.27458	0.0026019	0.2967
7.2369	7	7	7	0.23694	0.23694	0.23694	0.00071949	0.21039
6.0785	6	6	6	0.078539	0.078539	0.078539	0.0019334	0.26477
7.0885	7	7	7	0.088482	0.088482	0.088482	0.00053527	0.20122

Table A.2: Sample Monte Carlo data from the $N = 1$ observation at 0.1 interval case. Legendre polynomial $Nl = 5$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
7.5199	7.8	7.8	7.8	0.28006	0.28006	0.28006	7.6613e-15	0.0040225
7.2978	7.4	7.4	7.4	0.10215	0.10215	0.10215	4.008e-15	0.0040427
5.0329	5.2	5.2	5.2	0.16706	0.16706	0.16706	3.8846e-15	0.0040256
7.4427	7.5	7.5	7.5	0.057273	0.057273	0.057273	4.4507e-15	0.0040459
5.1615	5.1	5.1	5.1	0.061535	0.061535	0.061535	3.9848e-15	0.0040207
5.8954	5.8	5.8	5.8	0.095392	0.095392	0.095392	3.934e-15	0.0040359
7.9866	8.3	8.3	8.3	0.31343	0.31343	0.31343	5.3713e-15	0.0040163
6.8731	6.7	6.7	6.7	0.17311	0.17311	0.17311	3.8728e-15	0.0040325
7.8115	7.4	7.4	7.4	0.41147	0.41147	0.41147	4.0174e-15	0.004046
5.8195	5.6	5.6	5.6	0.21953	0.21953	0.21953	4.0635e-15	0.0040269
7.009	6.9	6.9	6.9	0.10901	0.10901	0.10901	3.6726e-15	0.0040455
5.5013	5.5	5.5	5.5	0.0013448	0.0013448	0.0013448	3.9113e-15	0.004032
5.6221	5.8	5.8	5.8	0.17788	0.17788	0.17788	3.9354e-15	0.0040215
7.8599	7.7	7.7	7.7	0.15987	0.15987	0.15987	7.1573e-15	0.0040465
5.1401	5.1	5.1	5.1	0.040096	0.040096	0.040096	3.9611e-15	0.004015
7.4261	7.9	7.9	7.9	0.47392	0.47392	0.47392	8.6877e-15	0.0040457
5.3837	5.3	5.3	5.3	0.08373	0.08373	0.08373	3.7327e-15	0.0040431
6.8345	6.7	6.7	6.7	0.13447	0.13447	0.13447	4.1361e-15	0.0040119
7.0417	7	7	7	0.041729	0.041729	0.041729	3.879e-15	0.0040432
6.272	6.5	6.5	6.5	0.22803	0.22803	0.22803	3.7855e-15	0.0040469

Table A.3: Sample Monte Carlo data from the $N = 1$ observation at 0.01 interval case. Legendre polynomial $Nl = 5$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
5.824	5.28	5.28	5.28	0.54397	0.54397	0.54397	1.1007e-14	4.033e-05
5.8527	5.75	5.75	5.75	0.10268	0.10268	0.10268	1.0286e-14	4.033e-05
5.3076	5.22	5.22	5.22	0.087555	0.087555	0.087555	1.0082e-14	4.0325e-05
6.1702	6.29	6.29	6.29	0.11977	0.11977	0.11977	1.0013e-14	4.0326e-05
5.798	5.97	5.97	5.97	0.17198	0.17198	0.17198	1.0386e-14	4.0328e-05
6.398	6.32	6.32	6.32	0.077975	0.077975	0.077975	1.0734e-14	4.0325e-05
6.0987	5.77	5.77	5.77	0.32865	0.32865	0.32865	1.04e-14	4.0329e-05
6.1934	6.47	6.47	6.47	0.27662	0.27662	0.27662	1.0245e-14	4.033e-05
7.8447	7.97	7.97	7.97	0.12531	0.12531	0.12531	9.2645e-14	4.0328e-05
7.1108	6.98	6.98	6.98	0.13077	0.13077	0.13077	1.0024e-14	4.0328e-05
7.8357	7.54	7.54	7.54	0.29568	0.29568	0.29568	2.9337e-14	4.0329e-05
6.4631	6.48	6.48	6.48	0.016911	0.016911	0.016911	9.6901e-15	4.0325e-05
5.8921	5.86	5.86	5.86	0.032131	0.032131	0.032131	9.6508e-15	4.0325e-05
5.029	4.97	4.97	4.97	0.058972	0.058972	0.058972	1.0341e-14	4.0329e-05
5.4041	5.86	5.86	5.86	0.45594	0.45594	0.45594	1.0427e-14	4.0325e-05
6.2713	6.19	6.19	6.19	0.081275	0.081275	0.081275	9.7663e-15	4.0329e-05
6.7118	6.36	6.36	6.36	0.35184	0.35184	0.35184	1.0364e-14	4.0324e-05
5.7821	5.6	5.6	5.6	0.18207	0.18207	0.18207	1.0327e-14	4.0326e-05
7.0992	6.95	6.95	6.95	0.14924	0.14924	0.14924	1.0662e-14	4.033e-05
6.943	7.08	7.08	7.08	0.137	0.137	0.137	1.0287e-14	4.033e-05

Table A.4: Sample Monte Carlo data from the $N = 5$ observation at 0.01 interval case. Legendre polynomial $Nl = 5$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
5.4231	5.22	5.22	5.22	0.20309	0.20309	0.20309	3.8002e-14	0.00020168
5.4576	5.61	5.61	5.61	0.15243	0.15243	0.15243	3.8485e-14	0.00020168
6.8517	6.89	6.89	6.89	0.038343	0.038343	0.038343	6.0104e-14	0.00020164
7.3032	7.14	7.14	7.14	0.16322	0.16322	0.16322	5.9795e-14	0.00020164
7.1285	7.07	7.07	7.07	0.058517	0.058517	0.058517	5.9498e-14	0.00020157
7.5282	7.66	7.66	7.66	0.13178	0.13178	0.13178	6.1191e-14	0.00020161
5.7593	5.78	5.78	5.78	0.020705	0.020705	0.020705	3.857e-14	0.00020168
5.2004	5.45	5.45	5.45	0.24961	0.24961	0.24961	3.8295e-14	0.00020161
7.4649	7.15	7.15	7.15	0.31486	0.31486	0.31486	6.0688e-14	0.00020167
5.7116	5.75	5.75	5.75	0.038409	0.038409	0.038409	3.8929e-14	0.0002016
7.8366	7.72	7.72	7.72	0.11658	0.11658	0.11658	6.2241e-14	0.00020167
7.0801	7.1	7.1	7.1	0.019948	0.019948	0.019948	6.0095e-14	0.00020168
6.2169	6.35	6.35	6.35	0.13307	0.13307	0.13307	4.7122e-14	0.00020157
7.1124	7.4	7.4	7.4	0.28758	0.28758	0.28758	6.0388e-14	0.00020168
7.9449	8.42	8.42	8.42	0.47505	0.47505	0.47505	7.3375e-14	0.00020157
6.1209	6.37	6.37	6.37	0.24909	0.24909	0.24909	4.8536e-14	0.00020157
6.3004	6.56	6.56	6.56	0.25965	0.25965	0.25965	5.873e-14	0.00020159
7.0496	6.88	6.88	6.88	0.16962	0.16962	0.16962	6.0606e-14	0.00020157
6.5974	6.52	6.52	6.52	0.077434	0.077434	0.077434	5.6967e-14	0.00020168
7.7345	7.86	7.86	7.86	0.12546	0.12546	0.12546	8.3124e-14	0.00020167

Table A.5: Sample Monte Carlo data from the $N = 50$ observation at 0.01 interval case. Legendre polynomial $Nl = 5$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
6.4188	6.42	6.42	6.42	0.0011693	0.0011693	0.0011693	3.7513e-14	0.002
6.6832	6.74	6.74	6.74	0.056809	0.056809	0.056809	4.0114e-14	0.0020196
6.2456	6.21	6.21	6.21	0.035575	0.035575	0.035575	3.4067e-14	0.0019996
5.8216	5.79	5.79	5.79	0.031608	0.031608	0.031608	8.2565e-15	0.0020209
7.5371	7.47	7.47	7.47	0.06708	0.06708	0.06708	3.1418e-14	0.0020073
5.0667	5.03	5.03	5.03	0.036658	0.036658	0.036658	3.1318e-14	0.002021
5.473	5.48	5.48	5.48	0.0069756	0.0069756	0.0069756	2.8784e-14	0.0020038
7.4218	7.4	7.4	7.4	0.021777	0.021777	0.021777	4.2085e-14	0.0019996
5.7105	5.71	5.71	5.71	0.00046015	0.00046015	0.00046015	3.0524e-14	0.0020025
5.4709	5.45	5.45	5.45	0.020875	0.020875	0.020875	2.8406e-14	0.0020221
6.3879	6.39	6.39	6.39	0.0020747	0.0020747	0.0020747	4.7967e-14	0.0020191
6.9677	6.96	6.96	6.96	0.0077444	0.0077444	0.0077444	2.9653e-14	0.0020179
5.479	5.43	5.43	5.43	0.04904	0.04904	0.04904	3.0703e-14	0.0020136
5.1368	5.15	5.15	5.15	0.013156	0.013156	0.013156	3.7198e-14	0.0020205
7.7127	7.68	7.68	7.68	0.032719	0.032719	0.032719	3.4765e-14	0.0019996
7.2558	7.25	7.25	7.25	0.0058374	0.0058374	0.0058374	3.2324e-14	0.0020214
7.5587	7.52	7.52	7.52	0.038651	0.038651	0.038651	9.5541e-14	0.0020202
5.2537	5.28	5.28	5.28	0.026264	0.026264	0.026264	2.6259e-14	0.0020174
6.4553	6.45	6.45	6.45	0.0053114	0.0053114	0.0053114	3.0998e-14	0.0020065
7.3247	7.27	7.27	7.27	0.054673	0.054673	0.054673	3.5757e-14	0.0020171

Table A.6: Sample Monte Carlo data from the $N = 500$ observation at 0.01 interval case. Legendre polynomial $Nl = 5$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
6.9124	6.94	6.94	6.94	0.027625	0.027625	0.027625	1.0412e-13	0.02001
5.1465	5.14	5.14	5.14	0.0065074	0.0065074	0.0065074	3.2254e-13	0.018997
6.1271	6.14	6.14	6.14	0.012865	0.012865	0.012865	3.1705e-13	0.018958
6.8031	6.77	6.77	6.77	0.033084	0.033084	0.033084	2.1156e-13	0.020661
6.2372	6.25	6.25	6.25	0.012839	0.012839	0.012839	2.9373e-13	0.020302
5.9068	5.89	5.89	5.89	0.016755	0.016755	0.016755	6.662e-14	0.020109
6.0492	6.07	6.07	6.07	0.020762	0.020762	0.020762	3.4181e-13	0.020309
7.0844	7.08	7.08	7.08	0.0043916	0.0043916	0.0043916	1.1158e-13	0.02018
6.6055	6.61	6.61	6.61	0.0045213	0.0045213	0.0045213	1.9463e-13	0.019789
6.0447	6.04	6.04	6.04	0.0046675	0.0046675	0.0046675	1.1907e-13	0.019003
5.0469	5.08	5.08	5.08	0.033112	0.033112	0.033112	6.4135e-14	0.020393
7.1637	7.17	7.17	7.17	0.0063228	0.0063228	0.0063228	2.8886e-13	0.020642
5.0031	5.02	5.02	5.02	0.016948	0.016948	0.016948	1.6936e-13	0.020647
6.6333	6.62	6.62	6.62	0.013318	0.013318	0.013318	6.2661e-13	0.020283
6.9201	6.92	6.92	6.92	0.00014483	0.00014483	0.00014483	3.9256e-13	0.019369
6.1027	6.09	6.09	6.09	0.012729	0.012729	0.012729	2.0749e-13	0.018981
6.8466	6.86	6.86	6.86	0.013402	0.013402	0.013402	4.6883e-14	0.02
6.7115	6.7	6.7	6.7	0.011499	0.011499	0.011499	1.8904e-13	0.020431
7.8673	7.86	7.86	7.86	0.0072952	0.0072952	0.0072952	7.4445e-14	0.01938
5.5569	5.55	5.55	5.55	0.0068763	0.0068763	0.0068763	2.0212e-13	0.02042

Table A.7: Sample Monte Carlo data from the $N = 5000$ observation at 0.1 interval case. Legendre polynomial $Nl = 5$ was used.

μ	$\hat{\mu}_e$	$\hat{\mu}_g$	$\hat{\mu}_s$	$\epsilon\hat{\mu}_e$	$\epsilon\hat{\mu}_g$	$\epsilon\hat{\mu}_s$	TAE_g	TAE_s
7.0573	7.1	7.1	7.1	0.04273	0.04273	0.04273	0.0058767	0.019491
5.1511	5.1	5.1	5.1	0.051142	0.051142	0.051142	0.10874	0.41538
5.6286	5.6	5.6	5.6	0.028576	0.028576	0.028576	0.00020511	0.00024857
6.1293	6.1	6.1	6.1	0.029271	0.029271	0.029271	6.7695e-06	7.0346e-06
5.1751	5.2	5.2	5.2	0.024917	0.024917	0.024917	0.00010277	0.00011651
5.4461	5.4	5.4	5.4	0.04612	0.04612	0.04612	0.10975	0.44679
6.3965	6.4	6.4	6.4	0.0035282	0.0035282	0.0035282	2.8464e-29	2.0298e-101
6.5731	6.6	6.6	6.6	0.026892	0.026892	0.026892	1.7852e-07	1.8147e-07
5.6881	5.7	5.7	5.7	0.011862	0.011862	0.011862	2.1656e-16	2.2204e-16
5.1158	5.1	5.1	5.1	0.015815	0.015815	0.015815	2.1785e-14	2.2204e-14
6.8519	6.8	6.8	6.8	0.05186	0.05186	0.05186	0.013797	0.50455
5.1221	5.1	5.1	5.1	0.022094	0.022094	0.022094	1.4079e-08	1.4256e-08
5.8594	5.9	5.9	5.9	0.040649	0.040649	0.040649	0.036784	0.094867
5.4638	5.5	5.5	5.5	0.036203	0.036203	0.036203	0.00090266	0.0018272
5.0109	5	5	5	0.010883	0.010883	0.010883	1.3525e-23	4.8757e-91
5.1128	5.1	5.1	5.1	0.01277	0.01277	0.01277	3.797e-14	3.8636e-14
5.9968	6	6	6	0.0031862	0.0031862	0.0031862	1.1616e-28	2.6557e-100
7.3982	7.4	7.4	7.4	0.0018186	0.0018186	0.0018186	3.8202e-25	7.1622e-94
6.3252	6.3	6.3	6.3	0.025248	0.025248	0.025248	1.0109e-07	1.0263e-07
5.964	6	6	6	0.035993	0.035993	0.035993	0.00088434	0.0017604

Appendix B: Additional Graphical Validation Results
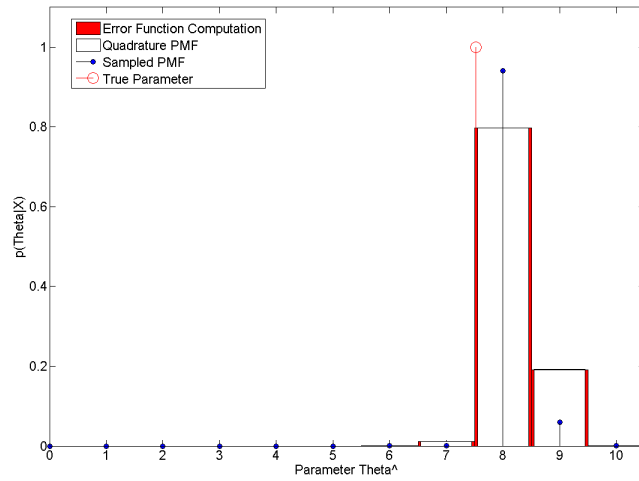

Figure B.1: Validation results for *N* = 1 observation at 1.0 interval. Legendre polynomial of order *Nl* = 3.
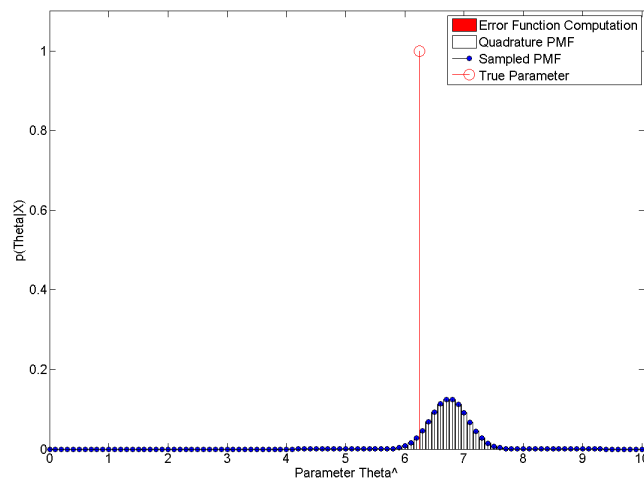

Figure B.2: Validation results for *N* = 1 observation at 0.1 interval. Legendre polynomial of order *Nl* = 5.
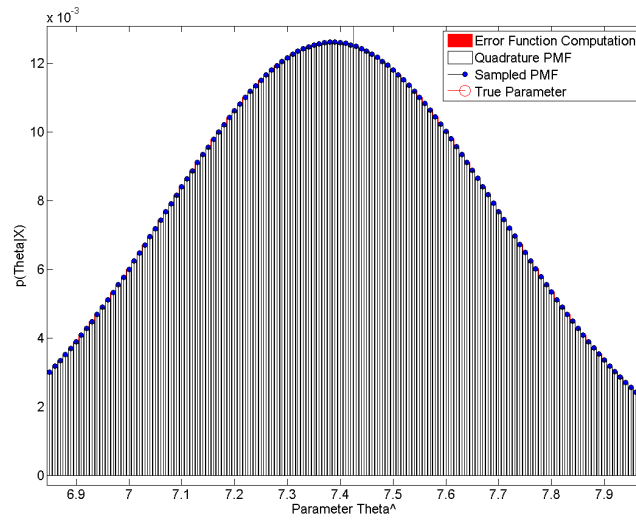
Figure B.3: Validation results for $N = 1$ observation at 0.01 interval. Legendre polynomial of order $Nl = 5$. Figure is zoomed into the region $5.9 \leq \hat{\Theta} \leq 8.1$.

Figure B.4: Validation results for $N = 5$ observations at 0.01 interval. Legendre polynomial of order $Nl = 5$. Figure is zoomed into the region $6.9 \leq \hat{\Theta} \leq 7.6$.
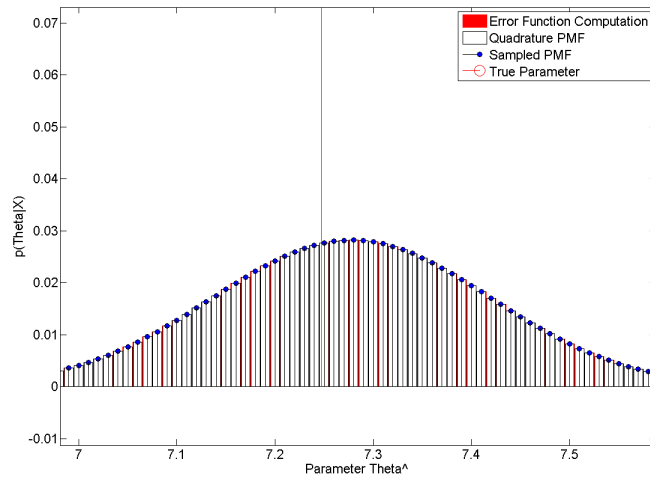
Figure B.5: Validation results for $N = 50$ observations at 0.01 interval. Legendre polynomial of order $Nl = 5$. Figure is zoomed into the region $7.7 \leq \hat{\Theta} \leq 8.2$.
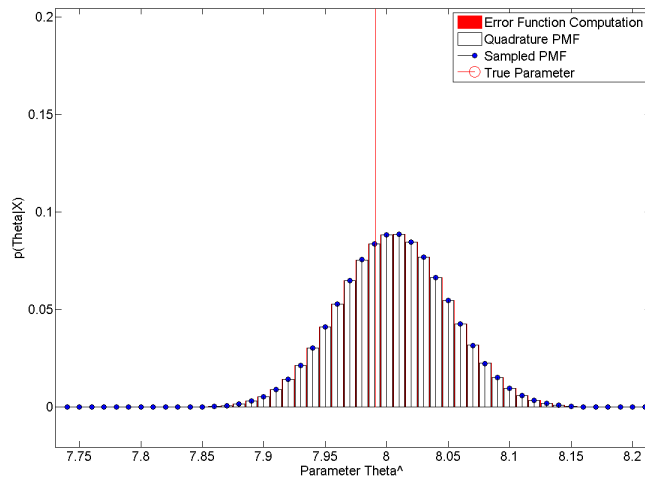

Figure B.6: Validation results for $N = 500$ observations at 0.01 interval. Legendre polynomial of order $Nl = 5$. Figure is zoomed into the region $7.7 \leq \hat{\Theta} \leq 8.1$.
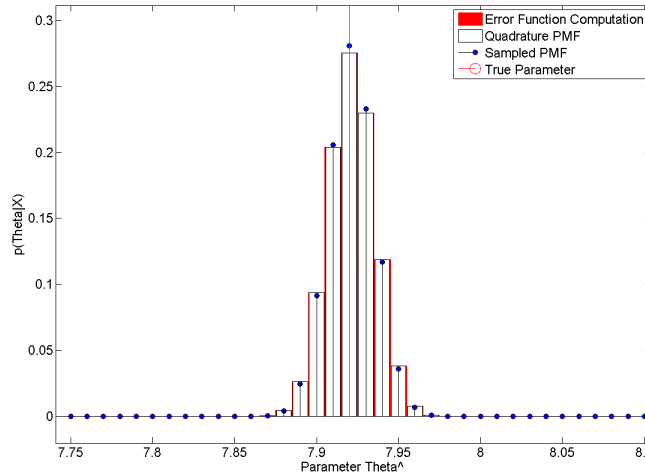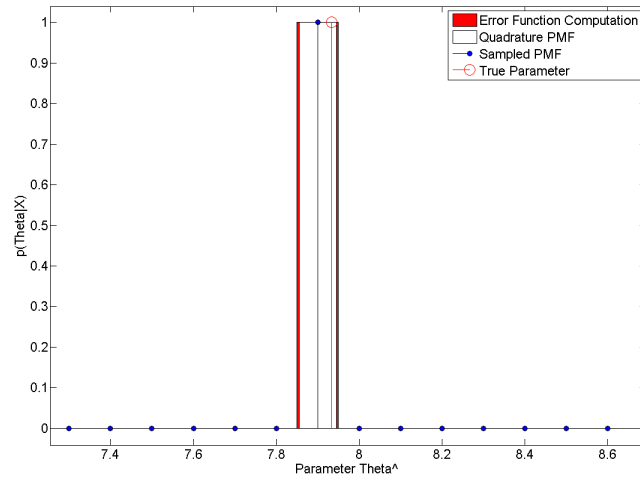
Figure B.7: Validation results for $N = 5000$ observations at 0.1 interval. Legendre polynomial of order $Nl = 5$. Figure is zoomed into the region $7.3 \le \hat{\Theta} \le 8.7$.

Bibliography

[1] T. D. Ross, "Confidence intervals for ATR performance metrics," *Proceedings of SPIE*, vol. 4382, pp. 318–329, 2001.

[2] R. W. Rademacher, "Bayesian methods and confidence intervals for automatic target recognition of SAR canonical shapes," Master's thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 2014.

[3] R. Rademacher, J. A. Jackson, A. Rexford, and C. Schubert Kabban, "Credible set estimation for SAR feature extraction," to appear in IEEE International Radar Conference, May 2015.

[4] M. P. Crosser, "Improved dictionary formation and search for synthetic aperture radar canonical feature extraction," Master's thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 2014.

[5] J. A. Jackson, "Three-dimensional feature models for synthetic aperture radar and experiments in feature extraction," Ph.D. dissertation, The Ohio State University, Columbus, OH, 2009.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing*, ser. Estimation Theory. Prentice Hall, 2013, vol. 1.

[7] G. Evans, *Practical Numerical Integration*. John Wiley & Sons, 1993.

[8] A. Krommer and C. W. Ueberhuber, *Computational Integration*. SIAM, 1998, vol. 53.

[9] P. Davis and P. Rabinowitz, *Methods of Numerical Integration*. Courier Dover Publications, 2007.

[10] G. H. Golub and J. H. Welsch, "Calculation of Gauss quadrature rules," *Mathematics of Computation*, vol. 23, no. 106, pp. 221–230, 1969.

[11] J. Jackson, B. Rigling, and R. L. Moses, "Canonical scattering feature models for 3D and bistatic SAR," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 2, pp. 525–541, April 2010.

[12] J. A. Jackson and R. L. Moses, "Identifiability of 3D attributed scattering features from sparse nonlinear apertures," in *Proceedings of SPIE Algorithms for Synthetic Aperture Radar Imagery XIV*, vol. 6568, 2007.

[13] G. B. Hammond, "SAR canonical feature extraction using molecule dictionaries," Master's thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 2012.

[14] "DoD high performance computing modernization program," http://centers.hpc.mil/, 2015, online; accessed 21-Jan-2015.

[15] G. Casella and R. Berger, *Statistical Inference*, 2nd ed. Belmont, CA: Brooks/Cole Cengage Learning, 2002.

[16] D. F. Fuller, "Phase history decomposition for efficient scatterer classification in SAR imagery," Ph.D. dissertation, Air Force Institute of Technology, Wright-Patterson AFB, OH, 2009.

[17] J. A. Jackson, "Analytic physical optics solution for bistatic, 3D scattering from a dihedral corner reflector," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 3, pp. 1486–1495, March 2012.

[18] C. V. Jakowatz, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-mode Synthetic Aperture Radar: A Signal Processing Approach*. Boston, MA: Kluwer Academic Publishers, 1996.

[19] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Prentice Hall, 2005.

[20] K. Dungan, C. Austin, J. Nehrbass, and L. C. Potter, "Civilian vehicle radar data domes," in *Algorithms for Synthetic Aperture Radar Imagery XVII, Proceedings of SPIE*, E. G. Zelnio and F. D. Garber, Eds., vol. 7699, 2010, p. 76990P.

REPORT DOCUMENTATION PAGE

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704–0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
26–03–2015	Master's Thesis	Oct 2013–Mar 2015

4. TITLE AND SUBTITLE

Credible Set Estimation, Analysis, and Applications in Synthetic Aperture Radar Canonical Feature Extraction

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Rexford, Andrew C., 1st Lieutenant, USAF

5d. PROJECT NUMBER

LRIR12RY19COR

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB, OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENG-MS-15-M-033

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Dr. Michael Kendra
Air Force Office of Scientific Research
875 N. Randolph St. Suite 325
Arlington, VA 22203
(703) 588-0671, michael.kendra@afosr.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFOSR/RTA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Traditional estimation schemes such as Maximum A Posterior (MAP) or Maximum Likelihood Estimation (MLE) determine the most likely parameter set associated with received signal data. However, traditional schemes do not retain entire posterior distribution, provide no confidence information associated with the final solution, and often rely on simple sampling methods which induce significant errors. Also, traditional schemes perform inadequately when applied to complex signals which often result in multi-modal parameter sets. Credible Set Estimation (CSE) provides a powerful and flexible alternative to traditional estimation schemes.

CSE provides an estimation solution that accurately computes posterior distributions, retains confidence information, and provides a complete set of credible solutions. Determination of a credible region becomes especially important in Synthetic Aperture Radar (SAR) Automated Target Recognition (ATR) problems where signal complexity leads to multiple potential parameter sets. The presented research provides validation of methods for CSE, extension to high dimension/large observation sets, incorporation of Bayesian methods with previous work on SAR canonical feature extraction, and evaluation of the CSE algorithm. The results in this thesis show that: the CSE implementation of Gaussian-Quadrature techniques reduces computational error of the posterior distribution by up to twelve orders of magnitude, the presented formula for computation of the posterior distribution enables numerical evaluation for large observation sets (greater than 7,300 observations), and the algorithm is capable of producing M-th dimensional parameter estimates when applied to SAR canonical features. As such, CSE provides an ideal estimation scheme for radar, communications and other statistical problems where retaining the entire posterior distribution and associated confidence intervals is desirable.

15. SUBJECT TERMS

SAR imaging, Canonical Shape parameterization, Bayesian Methods, Confidence Interval, Gaussian Quadrature, Numerical Methods, Parameter Estimation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Julie A. Jackson, AFIT/ENG
U	U	U	UU	123	**19b. TELEPHONE NUMBER** *(include area code)* (937) 255-3636 x4678 julie.jackson@afit.edu