ANALYTIC METHODS FOR SAR IMAGE FORMATION IN THE PRESENCE OF NOISE AND CLUTTER

By

Hüseyin Çağrı Yanık

A Dissertation Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: ELECTRICAL ENGINEERING

Approved by the
Examining Committee:

Birsen Yazıcı, Dissertation Adviser

Richard J. Radke, Member

John W. Woods, Member

Joyce R. McLaughlin, Member

Rensselaer Polytechnic Institute
Troy, New York

October 2014
(For Graduation December 2014)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my dissertation advisor, Prof. Birsen Yazıcı, for her aspiring guidance, everlasting support, invaluably constructive criticism, patience and encouragement on uncountable occasions during this dissertation and my entire graduate education years. As my mentor, I hope Prof. Yazıcı will always be present in my future work.

I would like to acknowledge the members of the dissertation committee, Prof. Richard J. Radke, Prof. John W. Woods and Prof. Joyce R. McLaughlin for their time, valuable discussions and productive comments on this work.

I would also like to take this opportunity to express my warm thanks to Kaan Duman, Nihat Baysal, Deniz Rende and Yasemin Yeşiltepe, who have been my family in Troy. Many thanks to Oğuzhan Uyar, Cemal Çağatay Bilgin, Başak Öztan, Şahin Cem Geyik and Buğra Çaşkurlu for their great friendship and making the past five years enjoyable.

Last but not least, I would like to express my deepest love and dedicate this dissertation to my father Rıza Tevfik Yanık, my mother Pakize Yanık, and my *little* brother Hamdullah Yanık, who supported me without any doubt my whole life and assured me that I can achieve anything.

ABSTRACT

Synthetic Aperture Radar (SAR) is a valuable imaging modality in civil and environmental monitoring, defense and homeland security applications. This thesis presents novel analytic and computationally efficient SAR image formation methods in the presence of noise and clutter using *a priori* models for target, clutter and noise.

The first part of the thesis presents statistical SAR inversion methods to suppress noise and clutter using spatially-varying quadratic priors. We present a novel class of *non-stationary* stochastic processes, which we refer to as *pseudo-stationary*, to model radar targets and clutter. First, we develop analytic filtered-backprojection- and backprojection-filtering-type SAR inversion methods based on the minimum mean square error criterion when the target and the clutter are pseudo-stationary. Next, we develop an analytic inversion formula based on a best linear unbiased estimation criterion when the clutter is a pseudo-stationary process.

In the second part of the thesis, we investigate non-quadratic prior models to represent target scenes. Specifically, we consider edge-preserving prior models. First, we present a simultaneous analytic, image formation and edge detection method. Then, we formulate the SAR image reconstruction as non-quadratic optimization problems. We solve these optimization problems approximately with sequences of filtered-backprojection operators.

The methods presented in this thesis have the advantages of computational efficiency, applicability to arbitrary imaging geometries and several different SAR modalities.

The methods and algorithms derived in this thesis are extensively tested using high-fidelity simulated data and real SAR data.

CHAPTER 1
INTRODUCTION

1.1 Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is a coherent imaging technique used in a wide range of applications from civil and environmental monitoring to defense and homeland security. SAR has the advantage of providing rich information, long-range, all-weather operation and large area coverage.

A typical SAR system involves a source of electromagnetic (EM) illumination and (a) receiving antenna(s). The movement of the source and/or receiver(s) allows mathematical synthesis of an image from the backscattered EM waves.

SAR can be used in a variety of modes depending on the number and geometric configurations of antennas, transmitted waveforms, center frequency and other operating parameters.

In mono-static SAR configuration, receiving and transmitting antennas are located on the same platform [3].

In a bi-static SAR, receiving and transmitting antennas are sufficiently far apart so that the transmitter and the receiver look directions are different [1]. Recently, bi-static configuration has received increasing attention in the context of passive radar applications.

A multi-static SAR configuration involves three or more sufficiently far apart antennas that are receiving, transmitting or both [4]. With the recent advances in unmanned air vehicle (UAV) and microsatellite technologies [5]–[11], there is a growing interest in multi-static SAR [4], [12]–[15].

Hitchhiker is a novel SAR modality that involves multiple receiving antennas [16]. It takes advantage of ambient EM waves provided by the transmitters of opportunity to perform radar tasks. Hitchhiker SAR may be useful in urban areas where there are many transmitters of opportunity such as GSM, TV, Wi-Fi transmitters.

In addition to these, there are many other SAR modalities that can be derived

by configuring imaging parameters in space, time, polarization state, frequency etc. These modalities include the along-track and cross-track interferometric SAR, polarimetric SAR to mention a few.

1.2 SAR Image Reconstruction Problem

There are analytical and numerical optimization-based SAR inversion techniques. Analytical techniques are widely used and they are computationally efficient. However, they cannot accommodate the measurement noise and clutter explicitly. Moreover, there are limitations to the antenna trajectories. In order some of the assumptions to hold for image formation with some of these techniques, antennas need to traverse short, linear trajectories.

The range-Doppler algorithm is the most commonly used analytical method in SAR image formation [17], [18]. The method is easy to implement and requires calculation of a matched-filter. However, the method has some limitations since the ideal matched-filter depends on the location of the target which is unknown and the calculation of the matched-filter can be computationally burdensome. The chirp-scaling algorithm [19] offers a solution to the computational load by bypassing some of the interpolation needed for the range-Doppler algorithm via scaling the *chirp signal*, which is a special type of transmitted waveform in SAR. There is also ω-k algorithm that formulates the SAR inversion problem by using the wave equation [20], [21]. This algorithm is also referred to as the *wave number domain algorithm* and performs focusing via Stolt interpolation [22] to deal with the wide aperture issues inherent with the range-Doppler and chirp-scaling methods. The polar-format algorithm compensates for the limitations of range-Doppler method by exploiting the fact that the SAR received signal is collected in a spherical geometry (polar coordinate system) [23], [24]. Finally, there are backprojection-based analytic inversion methods [1], [3]. Our research group has been extensively working on backprojection-based analytic inversion methods from a generalized Radon transform (GRT) point of view [1], [4], [16], [25]. [25] was the first work in the literature to incorporate *a priori* statistical information about the target and clutter scenes with an analytic inversion method. We consider SAR collected data as a

GRT of the scene reflectivity function (representing the scatterers on the ground) weighted and projected onto some smooth manifolds. Backprojection-based methods can accommodate arbitrary imaging geometries. In Chapter 2, we describe the SAR image formation from a GRT point of view.

Numerical optimization-based techniques are a second class of SAR inversion methods. They use discrete models and can be statistical or deterministic. Numerical optimization-based techniques consider the measurement noise and clutter present in the scene, however most of the time they are computationally intensive. Numerical optimization-based methods treat SAR image formation operator as a forward matrix and attempt to find inverse matrices to form SAR images. Detailed literature survey on numerical optimization based methods are given in Section 3.1 and Section 6.2.

In this thesis, we wish to develop computationally efficient, backprojection-based, analytic SAR inversion methods that are robust in the presence of noise and clutter and use *a priori* information to accurately characterize SAR target and clutter scenes.

1.3 Challenges in SAR Image Formation

1- Noise and clutter: The measurement noise and clutter are ubiquitous in SAR data. The classical SAR image reconstruction techniques, such as matched filtering, range-Doppler [17] and chirp scaling algorithms [19] do not take into account noise and clutter, explicitly. Noise and clutter reduce the detectability and the recognition of targets.

2- Computational efficiency: SAR images cover large swaths. For example, a typical TerraSAR-X image of New Orleans area is a $14,000 \times 14,000$-pixel image. Therefore, SAR is a computationally demanding, large inversion problem. Fast algorithms are desired to achieve computational efficiency. There are analytical and numerical optimization-based techniques to address these problems. Analytical techniques are widely used and they are computationally efficient. However, they cannot explicitly accommodate noise and clutter. Numerical optimization-based techniques can use statistical or deterministic models. However, most of the time

they are computationally intensive.

3- Use of "accurate" *a priori* information: The spectral estimation-based techniques generally require *a priori* information about clutter and suitable initialization for the convergence of the algorithms [26]. In the knowledge aided space-time adaptive processing (KA-STAP) methods, this prior information is collected when there are no targets in the imaging scene and provided as a clutter covariance matrix to be used as training data. To overcome the requirement on prior information, parametric techniques are proposed [27]–[29] but they are computationally intensive.

4- Non-stationary nature of clutter and target: Often times, in numerical optimization based methods, measurement noise is modeled to be additive Gaussian [30]–[34] without any specification on the statistical nature of the target and the clutter. This induces model-based errors to the reconstructed SAR images due to non-stationary nature of the target and the clutter in reality.

5- Limited data: SAR systems can be used to operate in multiple modes such as imaging and ground-moving target imaging. Such systems are referred to as the *interrupted SAR systems*. Switching between different operation modes may result in missing or undersampled data. Moreover, the amount of data collected for a single acquisition may be very large to transmit and process in real time. Also, in many practical applications the data may be incomplete to apply analytic inversion methods, effectively [35]–[38]. Therefore, it is desirable to develop SAR reconstruction algorithms using limited aperture data.

1.4 The Goal of the Thesis

In this thesis, we wish to develop computationally efficient, analytic image reconstruction methods that are (1) robust in the presence of noise and clutter; (2) based on specific *a priori* information that can accurately characterize SAR target and clutter scenes.

Towards this goal, the thesis develops the following:

1. A novel class of non-stationary processes to model radar targets and clutter.

2. Analytic and computationally efficient inversion methods within the minimum mean square error (MMSE) and best linear unbiased estimator (BLUE) frameworks when target and clutter belong to this class of processes.

3. Analytic, iterative and computationally efficient inversion methods using edge-preserving non-quadratic prior models on radar targets.

1.5 Contributions and Organization of the Thesis

This thesis comprises of two parts. In the first part, we investigate quadratic priors to model radar targets and clutter. In the second part, we use non-quadratic priors for the target scenes.

The contributions of this thesis are as follows:

In Chapter 2:

- We briefly describe the SAR image reconstruction problem from a GRT point of view.

In Chapter 3:

- We develop analytic, computationally efficient solutions to SAR image formation problem in the presence of noise and clutter. The methods described in this thesis are of backprojection-type. Therefore, they can be implemented with the computational complexity of the fast backprojection algorithms [39], [40].

- We develop a novel class of statistically non-stationary processes to model radar targets and clutter. We refer to this new family of processes as the *pseudo-stationary processes* and they are locally stationary.

- We further show that the stochastic processes that can be represented by using pseudo-stationary processes are outputs of time-varying convolution filters.

- By using this new family of processes, we develop an FBP-type, analytic SAR reconstruction method that suppresses noise and clutter based on MMSE criterion. We use MMSE locally which overcomes the over-smoothing effect of MMSE.

- We introduce backprojection-filtering as a novel SAR inversion method.
- Implementation of FBP- and BPF-type noise and clutter suppression methods require the *space-varying spectral density* (SVSD) functions for the target and clutter. We develop an algorithm similar to the one described in [41] to estimate target and clutter SVSD functions directly from the radar data in the same chapter.

In Chapter 4:

- We develop an analytic SAR image reconstruction method that suppresses noise and clutter based on BLUE criterion.

In Chapter 5:

- We describe a novel direct edge-enhancement method that simultaneously reconstructs images and enhances edges along desired directions. The method is also capable of smoothing edges along desired directions.

In Chapter 6:

- We use non-quadratic priors to model radar targets.
- We model SAR image formation problem as an L^p-norm constrained minimization problem. We approximate this problem with a sequence of L^2-norm constrained inversion problems.
- We present iterative reweighted- and shrinkage-type algorithms to solve this non-quadratic optimization problem.
- We use the method described in Chapter 5 to sharpen the reconstructed images.
- Different than other noise/clutter suppression techniques, L^p-type regularization can be applied to limited data.

We demonstrate the performance of the proposed methods with synthetic data formed by MATLAB, real data from Wide Angle SAR experiment [2] and the Civilian Vehicles (CV) dome data set [42], both provided by the Air Force Research Laboratories (AFRL). Chapter 6 concludes the thesis.

PART I

CHAPTER 2
SAR IMAGE RECONSTRUCTION AS A GENERALIZED RADON TRANSFORM

In this chapter, we first derive the SAR received signal model starting from the scalar wave equation. Then, we describe the SAR image formation from a generalized Radon transform point of view. Finally, we derive a filtered-backprojection type filter for SAR in a deterministic setting.

2.1 SAR Received Signal Model

For a bi-static SAR system, where transmitting and receiving antennas are located on different platforms, we model the transmitted EM waves as a time-varying source $j_{TR}(t,\mathbf{x})$ distributed over an aperture. $j_{TR}(t,\mathbf{x})$ is proportional to an effective current density on the transmitting antenna and can include arbitrary waveforms. We assume that the earth's surface is located at $\mathbf{x}=[x_1,x_2,\boldsymbol{\psi}(x_1,x_2)]$ where $\boldsymbol{\psi}:\mathbb{R}^2\rightarrow\mathbb{R}$, is a known smooth function of the ground topography, and scattering takes place in a thin region near the surface. Throughout this thesis, the bold Roman, bold italic, and Roman small letters will denote points in $\mathbb{R}^3$, $\mathbb{R}^2$ and $\mathbb{R}$, respectively, i.e. $\boldsymbol{x}\in\mathbb{R}^2$, $x_3\in\mathbb{R}$, and $\mathbf{x}=[\boldsymbol{x},x_3]\in\mathbb{R}^3$.

The source term satisfies the scalar wave equation given by:

$$\left(\nabla^2-c^{-2}\partial_t^2\right)u^{in}(t,\mathbf{x})=-j_{TR}(t,\mathbf{x}) \tag{2.1}$$

where ∇ denotes the gradient operator, c is the speed of light in free-space, ∂ is the partial derivative operator and $u^{in}(t,\mathbf{x})$ is the incident field at time t and location $\mathbf{x}$. Using the Green's function

$$g_0(t,\mathbf{x})=\frac{\delta(t-|\mathbf{x}|/c)}{4\pi|\mathbf{x}|}=\int\frac{\mathrm{e}^{-\mathrm{i}\omega(t-|\mathbf{x}|/c)}}{8\pi^2|\mathbf{x}|}\mathrm{d}\omega \tag{2.2}$$

where ω is the angular frequency. Using the fact that

$$\left(\nabla^2 - c^{-2}\partial_t^2\right) g_0(t, \mathbf{x}) = -\delta(t, \mathbf{x}) \tag{2.3}$$

we obtain

$$u^{in}(t, \mathbf{x}) = (g_0 * j_{TR})(t, \mathbf{x}) = \int \frac{\mathrm{e}^{-\mathrm{i}\omega(t-|\mathbf{x}-\mathbf{y}|/c)}}{4\pi|\mathbf{x}-\mathbf{y}|} \hat{j}_{TR}(\omega, \mathbf{y}) \mathrm{d}\mathbf{x}\omega\mathrm{d}\mathbf{y}, \tag{2.4}$$

where $*$ denotes the convolution and

$$\hat{j}_{TR}(\omega, \mathbf{y}) = \frac{1}{2\pi} \int j_{TR}(t, \mathbf{y}) \mathrm{e}^{\mathrm{i}t\omega} \mathrm{d}t \tag{2.5}$$

is the Fourier transform of j_{TR}.

For a pulsed radar system it is assumed that the pulses are transmitted at t_i and the antenna position at the i^{th} pulse transmission is $\boldsymbol{\gamma}(t_i)$. Let $s := t_i$ denote the time at which a pulse is transmitted. Since the time scale at which the electromagnetic wave moves is much smaller than the time scale at which the antenna moves, we refer to s as the *slow-time* and t as the *fast-time*. We further assume that the antenna is small compared with the distance to the scatterers. Using the far-field expansion we write

$$|\mathbf{x}-\mathbf{y}| = |(\mathbf{x}-\boldsymbol{\gamma}(s)) + (\boldsymbol{\gamma}(s)-\mathbf{y})| \approx |\mathbf{x}-\boldsymbol{\gamma}(s)| + (\widehat{\mathbf{x}-\gamma(s)}) \cdot (\boldsymbol{\gamma}(s)-\mathbf{y}) \tag{2.6}$$

since $|\boldsymbol{\gamma}(s)-\mathbf{y}| \ll |\mathbf{x}-\boldsymbol{\gamma}(s)|$ where $\widehat{\mathbf{x}-\boldsymbol{\gamma}(s)} = \frac{\mathbf{x}-\gamma(s)}{|\mathbf{x}-\gamma(s)|}$ is the unit vector in the direction $\mathbf{x}-\gamma(s)$. Then,

$$u^{in}(t, \mathbf{x}) \approx \int \frac{\mathrm{e}^{-\mathrm{i}\omega(t-|\mathbf{x}-\gamma(s)|/c)}}{4\pi|\mathbf{x}-\gamma(s)|} \hat{j}_{TR}(\omega, \widehat{\mathbf{x}-\gamma(s)}) \mathrm{d}\omega\mathrm{d}\mathbf{y}, \tag{2.7}$$

and

$$j_{TR}(\omega, \widehat{\mathbf{x}-\gamma(s)}) = \mathrm{e}^{\mathrm{i}\omega(\widehat{\mathbf{x}-\gamma(s)})\cdot\gamma(s)/c} \int \mathrm{e}^{-\mathrm{i}\omega(\widehat{\mathbf{x}-\gamma(s)})\cdot\mathbf{y}/c} \hat{j}_{TR}(\omega, \mathbf{y}) \mathrm{d}\mathbf{y}. \tag{2.8}$$

The Fourier transform (2.8) gives the antenna beam pattern at each fixed frequency in the far-field. It is seen from (2.7) that the field transmitted by the antenna is a superposition of the fixed-frequency point sources that are each shaped by the antenna beam pattern. Using the scalar wave equation for the total field $u^{tot}(t, \mathbf{x})$ we obtain

$$\left(\nabla^2 - c^{-2}(\mathbf{x})\partial_t^2\right) u^{tot}(t, \mathbf{x}) = -j_{TR}(t, \mathbf{x}) \tag{2.9}$$

where $c(\mathbf{x})$ is the speed of the EM wave in the medium. We write total field as

$$u^{tot}(t, \mathbf{x}) = u^{in}(t, \mathbf{x}) + u^{sc}(t, \mathbf{x}) \tag{2.10}$$

and use (2.1)-(2.9) to get

$$\left(\nabla^2 - c^{-2}\partial_t^2\right) u^{sc}(t, \mathbf{x}) = -V(\mathbf{x})\partial_t^2 u^{tot}(t, \mathbf{x}) \tag{2.11}$$

where

$$V(\mathbf{x}) = \frac{1}{c^2} - \frac{1}{c^2(\mathbf{x})}. \tag{2.12}$$

Here, we refer to $V(\boldsymbol{x})$ as the *ground reflectivity function* that contains all the information related to the scattering nature of the ground.

Under the single-scattering (or Born) approximation $u^{tot}(t, \mathbf{x})$ is replaced by $u^{in}(t, \mathbf{x})$ and solving (2.11), we obtain

$$u_B^{sc}(t, \mathbf{x}) = \int g_0(t - \tau, \mathbf{x} - \mathbf{z}) V(\mathbf{z}) \partial_\tau^2 u^{in}(\tau, \mathbf{z}) \mathrm{d}\tau \mathrm{d}\mathbf{z}, \tag{2.13}$$

which for the incident field (2.7) becomes

$$u_B^{sc}(t, \mathbf{x}, s) = \int \frac{\mathrm{e}^{-\mathrm{i}\omega(t - (|\mathbf{x}-\mathbf{z}| + |\mathbf{z}-\gamma(s)|)/c)}}{(4\pi)^2 |\mathbf{x} - \mathbf{z}||\mathbf{z} - \gamma(s)|} \omega^2 \widehat{J_{TR}(\omega, \mathbf{z} - \gamma(s))} V(\mathbf{z}) \mathrm{d}\omega \mathrm{d}\mathbf{y} \mathrm{d}\mathbf{z}. \tag{2.14}$$

2.2 SAR Image Formation as Inversions of GRTs

Under the Born (single scattering) and start-stop approximations, for many SAR modalities [1], [3], [16] the ideal scattered field data can be modeled as follows:

$$\tilde{f}(s,t) \approx \mathcal{F}[V](s,t) := \int \mathrm{e}^{-\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{x})/c)} A(\boldsymbol{x},s,\omega) V(\boldsymbol{x})\, \mathrm{d}\omega \mathrm{d}\boldsymbol{x}, \tag{2.15}$$

where ω is the angular frequency, c denotes the speed of light in free-space, A is a complex amplitude term that includes the antenna beam patterns, transmitted waveform, geometrical spreading factors, etc. $s \in [s_0, s_1] \subset \mathbb{R}$ is the slow-time and $t \in [t_0, t_1] \subset \mathbb{R}$ is the fast-time.

$V(\boldsymbol{x})$ is given by

$$V(\mathbf{x}) = \frac{1}{c^2} - \frac{1}{c^2(\mathbf{x})} \tag{2.16}$$

where $c(\boldsymbol{x})$ is the speed of the EM wave in the scatterer and $V(\boldsymbol{x})$ is the ground reflectivity function. $V(\boldsymbol{x})$ recovered from the backscattered EM waves is called the *SAR image*. $R(s,\boldsymbol{x})$ can be interpreted as the distance between the antennas and the scatterer at $\mathbf{x}$.

Depending on the SAR modality, the antenna range, $R(s,\boldsymbol{x})$, takes different forms.

$$R_m(s,\boldsymbol{x}) = 2|\boldsymbol{\gamma}(s) - \mathbf{x}| \tag{2.17}$$

for mono-static SAR [3] where $\boldsymbol{\gamma}(s)$ denotes the antenna trajectory;

$$R_b(s,\boldsymbol{x}) = |\boldsymbol{\gamma}_T(s) - \mathbf{x}| + |\boldsymbol{\gamma}_R(s) - \mathbf{x}| \tag{2.18}$$

for bi-static SAR [1] where $\boldsymbol{\gamma}_T(s)$ and $\boldsymbol{\gamma}_R(s)$ denote the transmitter and the receiver antenna trajectories, respectively; and

$$R_h(s,\boldsymbol{x}) = |\boldsymbol{\gamma}_i(s) - \mathbf{x}| - |\boldsymbol{\gamma}_j(s) - \mathbf{x}| \tag{2.19}$$

for hitchhiker SAR [16] where $\boldsymbol{\gamma}_i(s)$ and $\boldsymbol{\gamma}_j(s)$ denote the trajectories of the i^{th} and

j^{th} receiver antennas, respectively. Note that $\mathbf{x} = [\boldsymbol{x}, \psi(\boldsymbol{x})]$ in (2.17)-(2.19).

For all SAR modalities, it is assumed that for some m_A, the amplitude term A satisfies the symbol estimate [1], [16]:

$$\sup_{(s,\boldsymbol{x})\in K} \mid \partial_\omega^\alpha \partial_s^\beta \partial_{\boldsymbol{x}}^{\boldsymbol{\rho}} A(\boldsymbol{x}, s, \omega) \mid \leq B_{K,\alpha,\beta,\boldsymbol{\rho}} \, (1 + |\omega|)^{m_A - |\alpha|} \tag{2.20}$$

where K is any compact subset of $\mathbb{R}\times\mathbb{R}^2$, $B_{K,\alpha,\beta,\boldsymbol{\rho}}$ is a $K, \alpha, \beta, \boldsymbol{\rho}$ dependent constant, and $\boldsymbol{\rho}$ is a multi-index. This assumption holds since most SAR modalities involve waveforms that are slowly varying in frequency domain. This assumption makes the "forward" operator $\mathcal{F}$, a *Fourier Integral Operator* (FIO) [43]–[46].

The image of the scene is reconstructed from the data collected from the measurements of the scattered waves. An FIO can be viewed as a GRT where the input function is weighted and projected onto some smooth manifolds. SAR measurement data can be viewed as a GRT of the scene reflectivity function in which the reflectivity is weighted and projected onto some smooth manifolds. These manifolds can be circles (mono-static SAR), ellipsis (bi-static SAR) or hyperbolas (hitchhiker SAR) depending on the imaging geometry assuming a flat topography.

The weighting or filtering is defined by the amplitude term and the smooth manifolds are defined by the phase function of the FIO which is determined by the antenna configuration. If the amplitude function, $A(\boldsymbol{x}, \omega, s)$ equals to 1, which corresponds to an isotropic antenna radiating a delta-like impulse and compensating for the geometric spreading factors, then the FIO simply projects the input function onto the manifolds defined by its phase term as given in (2.17)-(2.19). Figures 2.2 (a)-(c) illustrate the smooth manifolds onto which the scene reflectivity function is projected assuming that the topography is flat.

2.2.1 SAR Image Reconstruction by Using the Deterministic Filtered -Backprojection

Since the forward operator $\mathcal{F}$ in (2.15) is a Fourier integral operator, an image $\tilde{T}$ of the target scene T is formed by another Fourier integral operator $\mathcal{K}$ that is of

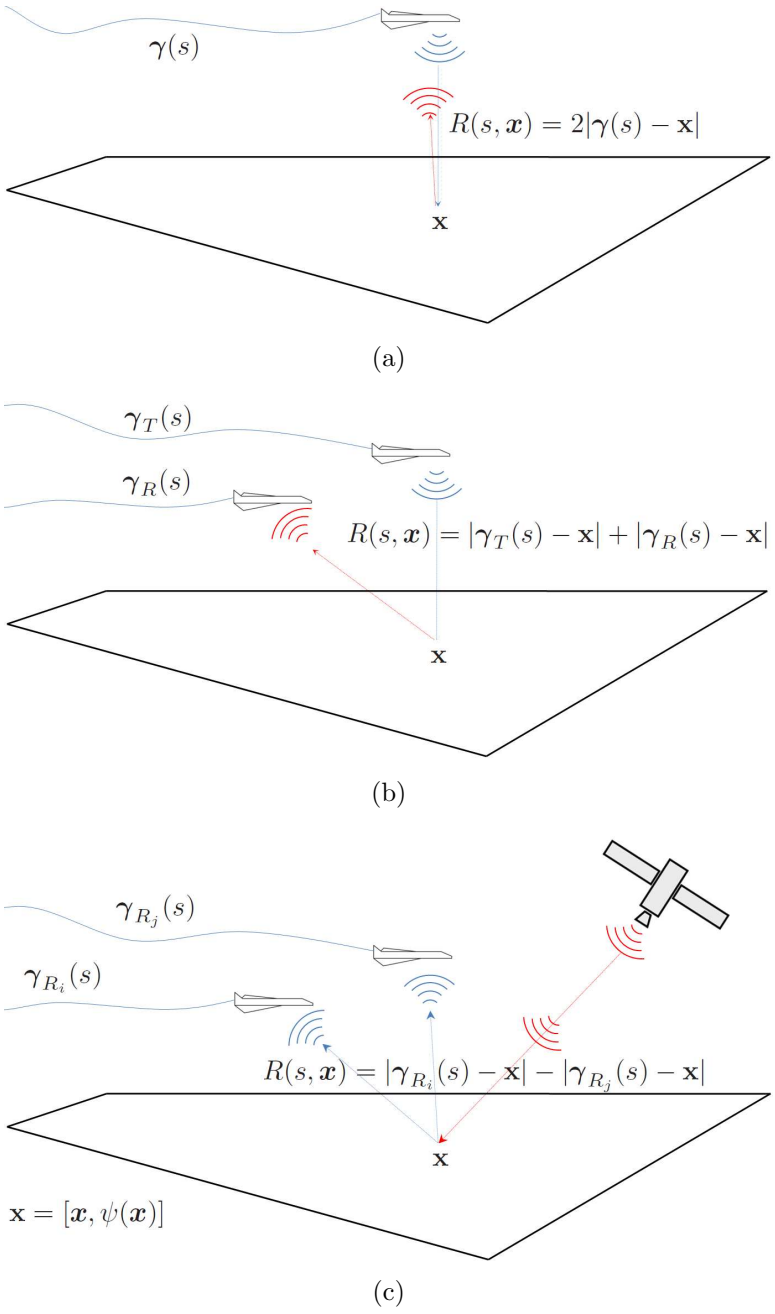


Figure 2.1: Imaging geometries for (a) mono-static SAR where $\boldsymbol{\gamma}(s)$ is the antenna trajectory, (b) bi-static SAR where $\boldsymbol{\gamma}_T(s)$ and $\boldsymbol{\gamma}_R(s)$ are the transmitter and receiver antenna trajectories, (c) hitchhiker SAR where $\boldsymbol{\gamma}_{R_i}(s)$, $\boldsymbol{\gamma}_{R_j}(s)$ are receiver trajectories.

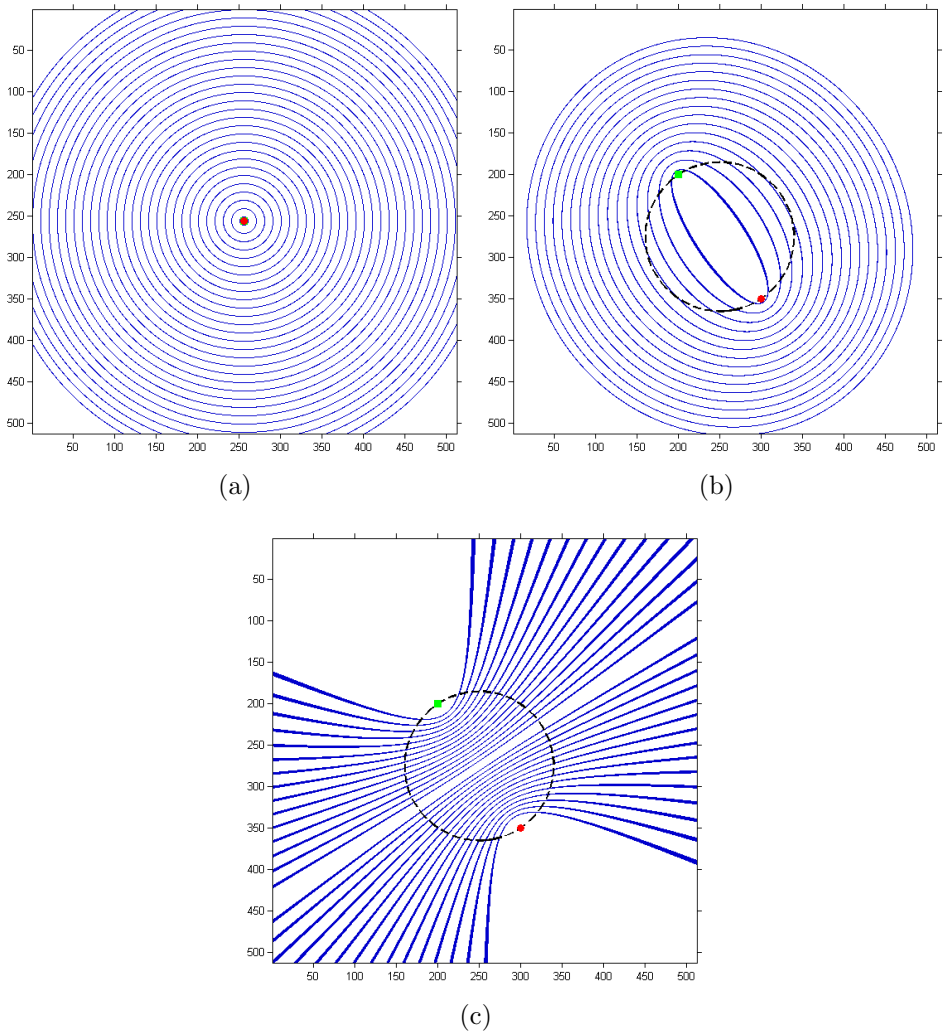


Figure 2.2: Iso-range contours for a fixed pulse. Red circle and green square depict transmitter and receiver, respectively. When intersected with a flat topography, iso-range contours form (a) circles in mono-static SAR (b) ellipsis in bi-static SAR (c) hyperbolas in hitchhiker SAR modalities.

filtered-backprojection (FBP) type as follows [1]:

$$\tilde{T}(\boldsymbol{z}) := \mathcal{KF}[T](\boldsymbol{z}) := \int Q(\boldsymbol{z}, s, \omega) \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{z})/c)} d(s,t)\ \mathrm{d}t\mathrm{d}s\mathrm{d}\omega, \tag{2.21}$$

$$\approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} \eta(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\xi}) A(\boldsymbol{x}, \boldsymbol{\xi}) Q(\boldsymbol{z}, \boldsymbol{\xi}) T(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi} \tag{2.22}$$

where η is the determinant of the Jacobian coming from the change of variables

$$(s,\omega) \to \boldsymbol{\xi} = \frac{\omega}{c}\boldsymbol{\Xi}(s,\boldsymbol{z},\boldsymbol{z}) = \frac{\omega}{c}\nabla_{\boldsymbol{z}} R(s,\boldsymbol{z},\boldsymbol{z}) \tag{2.23}$$

and for $\boldsymbol{x} = \boldsymbol{z}$. The filter is given by

$$Q(\boldsymbol{z},\boldsymbol{\xi}) = \chi_{\Omega}(\boldsymbol{z},\boldsymbol{\xi})\frac{\overline{A(\boldsymbol{z},\boldsymbol{\xi})}}{|A(\boldsymbol{z},\boldsymbol{\xi})|^2 \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})} \tag{2.24}$$

which satisfies a symbol estimate of the form (2.20) where $\chi_{\Omega}(\boldsymbol{z},\boldsymbol{\xi})$ is a smooth function that prevents division by 0. It is important to note here that the most important part of the filter is $\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})$ which is also known as the *Beylkin determinant* [47].

Microlocal analysis shows that FBP filter given in (2.24) reconstructs and preserves the location, orientation, order, and the strength of the edges of the scene [1]. Moreover, this filter can be implemented with fast backprojection algorithms which are computationally efficient.

Some of the abbreviations and notations used in this thesis are shown in Table 2.1 and Table 2.2.

Table 2.1: Table of Notations Part 1.

Symbol	Designation
t	Fast-time variable
s	Slow-time variable
c	Speed of light in free space
$\mathbf{x} = [\boldsymbol{x}, \boldsymbol{\psi}(\boldsymbol{x})]$	Earth's surface
$V(\mathbf{x})$	3D reflectivity function
$\tilde{f}(s,t)$	Ideal radar measurement without noise
$\mathcal{F}$	Forward modeling operator
ω	Angular frequency component
A	Complex amplitude function including the transmitter and receiver antenna beampatterns, geometric spreading factors, etc.
$\boldsymbol{\gamma}(s)$	Flight trajectory of the antenna
$R(\boldsymbol{x}, s)$	Travel distance or range between antennas and the scatterer at $\mathbf{x}$
$T(\boldsymbol{x})$	Scene containing target scatterers
$C(\boldsymbol{x})$	Scene containing clutter scatterers
$d(s,t)$	Received measurement containing noise and clutter
$n(s,t)$	Additive measurement noise
$\mathrm{B}(\boldsymbol{x})$	Zero-mean Wiener process (Brownian motion)
$\delta(\boldsymbol{x})$	Dirac delta function
$\nabla_{\boldsymbol{x}}$	Gradient with respect to $\boldsymbol{x}$
E	Expectation operator
$\mathcal{R}_T(\boldsymbol{x}, \boldsymbol{x}')$	Autocorrelation function of the target
$\mathcal{R}_C(\boldsymbol{x}, \boldsymbol{x}')$	Autocorrelation function of the clutter
$\mathcal{R}_n(\boldsymbol{x}, \boldsymbol{x}')$	Autocorrelation function of the noise

Table 2.2: Table of Notations Part 2.

Symbol	Designation
$\lvert S_T(\boldsymbol{x},\boldsymbol{\zeta})\rvert^2$	Space-varying spectral density function of the target
$\lvert S_C(\boldsymbol{x},\boldsymbol{\zeta})\rvert^2$	Space-varying spectral density function of the clutter
$\lvert S_n(s,\omega)\rvert^2$	Power spectral density function of the noise
$\mathcal{K}$	Filtered-backprojection operator
$\tilde{T}(\boldsymbol{z})$	Reconstructed image of the target scene
$\mathcal{K}_{ns}$, Q_{ns}	FBP operator and reconstruction filter with non-stationary models
$\mathcal{K}_{bp}$, Q_{bp}	BPF operator and reconstruction filter with non-stationary models
$\mathcal{K}_{bl}$, Q_{bl}	imaging filter and reconstruction filter with BLUE criteria
$J_{ns}(Q_{ns})$	Mean square error between target scene, $T(\boldsymbol{x})$ and reconstructed scene with Q_{ns}
$J_{bp}(Q_{bp})$	Mean square error between target scene, $T(\boldsymbol{x})$ and reconstructed scene with Q_{bp}
$J_{bl}(Q_{bl})$	Mean square error between target scene, $T(\boldsymbol{x})$ and reconstructed scene with Q_{bl}
J_{re}	Objective function for iterative reweighted-type reconstruction
J_{sh}	Objective function for iterative shrinkage-type reconstruction
η	Determinant of the Jacobian coming from the change of variables
$\Omega_{\boldsymbol{z}}$	Data collection manifold at $\mathbf{z} = [\boldsymbol{z}, \boldsymbol{\psi}(\boldsymbol{z})] \in \mathbb{R}^3$
$\tilde{\chi}_{\Omega_{\boldsymbol{z}}}$	Characteristic function associated with the data collection manifold $\Omega_{\boldsymbol{z}}$
$\mathcal{I}_{\Omega}$	Bandlimited identity operator
$\tilde{\chi}_{\Omega}$	A smooth function that prevents division by 0

CHAPTER 3
SYNTHETIC APERTURE INVERSION FOR STATISTICALLY NON-STATIONARY TARGET AND CLUTTER SCENES

In this chapter, we present a class of non-stationary stochastic processes that is suitable for modeling radar targets and clutter. This class of processes, which we refer to as *pseudo-stationary*, can be characterized by pseudo-differential operators driven by the Wiener process. A space-varying spectral density (SVSD) function can be defined, characterizing the spectral behavior of a pseudo-stationary process, through the symbol of the underlying pseudo-stationary process. We derive analytic filtered-backprojection (FBP) and backprojection-filtering (BPF)-type formulae for SAR image reconstruction when the target and clutter belong to this class of processes. We present results based on simulated and real SAR data to demonstrate the performances of the reconstruction methods. The inversion formulae are derived based on minimum square error estimation criterion. The underlying FBP and BPF filters depend on the SVSD functions of target and clutter.

3.1 Related Work

The measurement noise and clutter are ubiquitous in SAR data. Classical SAR image reconstruction techniques, such as matched filtering, range-Doppler [17] and chirp-scaling algorithms [19] do not take into account noise and clutter.

A number of different methods has been introduced to reconstruct SAR images in the presence of measurement noise and clutter explicitly. These methods can be roughly categorized into two classes: Numerical optimization-based methods [26]–[29], [48]–[64] and analytic image reconstruction methods [4], [25], [65]–[70].

Portions of this chapter to appear in: H. C. Yanık and B. Yazıcı, "Bi-static synthetic aperture Inversion for non-stationary target and clutter," to be presented at the IEEE Int. Conf. Image Process., Paris, France. Oct. 2014.

Portions of this chapter have been submitted to: H. C. Yanik and B. Yazici, "Synthetic aperture inversion for statistically non-stationary target and clutter scenes," *SIAM J. Imag. Sciences*.

The numerical optimization-based methods use discrete received signal models. They are typically iterative and computationally demanding. These include the feature-enhanced methods [30], [32], [33], [50], [54], [71] and spectral estimation-based methods [53], [55]–[62], [64], [72]. The feature-enhanced methods consider typically quadratic data-likelihood constrained by feature-enhancing prior terms, such as the L^p-norm constraint [50]. The numerical optimization methods employed vary from conjugate gradient to heuristic greedy approaches. Quadratic optimization functionals in conjunction with majorization-minimization or greedy-type approaches are also used in SAR image reconstruction [30], [32], [33], [54].

The spectral estimation-based SAR image reconstruction methods vary from classical techniques [61], [73], [74] to non-parametric, adaptive and compressive sensing based methods [53], [56], [58], [64], [75]. These methods are limited to certain imaging geometries or assumptions, such as linear flight trajectory and linear wavefront curvatures.

In [53], [56], [57], [64] it is assumed that the scene consists of point scatterers and the radar data is a collection of sinusoidal point scatterers. Li *et al.* presents amplitude and phase estimation of a sinusoid algorithm in [56], which is an adaptive filter-bank estimation method and compares its performance with the Capon method [75]. In [57], Bi *et al.* develops super resolution SAR image formation methods that suppress noise and clutter based on parametric data models using fast Fourier transform (FFT). Larsson *et al.* extends the algorithms presented in [56] and [75] to the periodically missing SAR data in [58] by interpolating the data under some conditions. [59] presents a computationally efficient parametric approach for image formation and feature extraction when the radar data has clutter and noise. In [61], Xiao *et al.* extends the algorithms in [74], [76] to 3D SAR image formation problem. [53] and [64] are iterative spectral estimation-based methods. In [64] a weighted least-squares approach is used and [53] is an iterative minimization method similar to maximum a priori estimation. [60] and [62] develop spectral estimation methods for tomographic SAR. An adaptive spectral estimation method is presented in [60] for tomographic SAR using an adaptive Capon estimator. [62] presents a spectral estimation method-based on compressed sensing for super resolution SAR image

formation with L^1-norm minimization using analytic data model.

Space-time adaptive processing (STAP) is another numerical optimization-based method to tackle with the noise and clutter in SAR imaging [77]–[80]. STAP-based methods are widely used specially for moving target imaging in which stationary objects are considered as clutter. STAP methods can be classified as parametric [27]–[29] and knowledge aided STAP (KA-STAP) [26], [48], [49], [51], [52] depending on the availability of prior information about clutter and noise. This prior information is collected when there are no targets in the imaging scene and provided as a clutter covariance matrix to be used as training data in KA-STAP methods. The dependency on prior information on clutter and suitable initialization for the convergence of the algorithm [26] restrict the efficacy of KA-STAP methods.

The second class of SAR image reconstruction methods that take into account noise and clutter is the analytical inversion methods [25], [65]–[68], [70]. Unlike the numerical optimization-based methods, these methods formulate the image reconstruction problem as an inversion of an operator and seek to determine an explicit inverse operator. The resulting methods are often computationally more efficient than the numerical optimization-based methods. Our current work falls into this class of methods.

In [65] and [66] interferometric methods to suppress noise, clutter and phase errors were presented. In [25], we presented an FBP-type SAR inversion method in the presence of noise and clutter based on minimum mean square error (MMSE) criterion. This method was extended and applied to multi-static SAR imaging in [4], SAR imaging in multiple scattering environments in [69] and polarimetric SAR imaging in [70].

The method in [25] relies on the assumption that the target, clutter and noise are statistically stationary. This assumption, however, may not be valid for typical SAR scenes that include electromagnetically heterogenous objects.

In this thesis, we wish to develop computationally efficient, analytic and backprojection-based SAR inversion methods that are robust in the presence of noise and clutter and use *a priori* information to accurately characterize SAR target and clutter scenes.

3.2 Statistical Models for Target, Clutter and Noise

3.2.1 Non-stationary Target Model

In radar applications, the scatterers of interest, such as certain vehicles or buildings, are typically referred to as the *target* and those scatterers that result in unwanted reflections, such as a tree or a lamp-post, are referred to as the *clutter*. Thus, the scene of scatterers V can be decomposed into target T, and clutter C:

$$V(\boldsymbol{x}) = T(\boldsymbol{x}) + C(\boldsymbol{x}). \tag{3.1}$$

Taking into account the measurement noise, we extend (2.15) and model the measured scattered field data by:

$$d(s,t) = \mathcal{F}[T + C](s,t) + n(s,t), \tag{3.2}$$

where $n(s,t)$ denotes the receiver noise.

Let Ω_T be a compact subset of $\mathbb{R}^2$ and $T(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega_T$ be the target scene. Without loss of generality, we assume that $T(\boldsymbol{x})$ is a zero-mean stochastic process for all $\boldsymbol{x} \in \Omega_T$. We assume that $T(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega_T$ can be characterized in the mean square sense as follows:

$$T(\boldsymbol{x}) = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} S_T(\boldsymbol{x},\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \mathrm{dB}(\boldsymbol{x}') \tag{3.3}$$

where $\mathrm{B}(\boldsymbol{x}')$ denotes the zero-mean Wiener process (or Brownian motion).

For S_T, we make the following assumptions:

Assumption 1: S_T is an even function of $\boldsymbol{\xi}$, i.e. $|S_T(\boldsymbol{x},\boldsymbol{\xi})| = |S_T(\boldsymbol{x},-\boldsymbol{\xi})|$.

Assumption 2: S_T is sufficiently rapidly decreasing in $\boldsymbol{\xi}$ so that the variance of $T(\boldsymbol{x})$ satisfies

$$\mathrm{E}[|T(\boldsymbol{x})|^2] = \int |S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 \, \mathrm{d}\boldsymbol{\xi} < +\infty \tag{3.4}$$

for all $\boldsymbol{x} \in \Omega_T$.

Assumption 3: S_T satisfies

$$\sup_{\boldsymbol{x}\in U} \mid \partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \partial_{\boldsymbol{x}}^{\boldsymbol{\beta}} S_T(\boldsymbol{x},\boldsymbol{\xi}) \mid \leq B_{U,\boldsymbol{\alpha},\boldsymbol{\beta}}^{T} (1 + |\boldsymbol{\xi}|)^{m_T - |\boldsymbol{\alpha}|} \tag{3.5}$$

where U is any compact subset of $\mathbb{R}^2$, $B^T_{U,\boldsymbol{\alpha},\boldsymbol{\beta}}$ is a constant that depends on the multi-indices $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and U, and m_T is some real number.

Thus, Assumption 3 makes

$$\tilde{h}(\boldsymbol{x}, \boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} S_T(\boldsymbol{x}, \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \tag{3.6}$$

the kernel of a pseudo-differential operator with symbol S_T satisfying Assumptions 1-3. We refer to stochastic processes defined in (3.3) as the *pseudo-stationary processes*. Note that if $S_T(\boldsymbol{x}, \boldsymbol{\zeta})$ is independent of $\boldsymbol{x}$, (3.3) defines $T(\boldsymbol{x})$ as a stationary process.

Let

$$h_T(\boldsymbol{x}, \boldsymbol{x} - \boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} S_T(\boldsymbol{x}, \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}. \tag{3.7}$$

Then, (3.3) can be alternatively expressed as follows:

$$T(\boldsymbol{x}) = \int h_T(\boldsymbol{x}, \boldsymbol{x} - \boldsymbol{x}')\, \mathrm{dB}(\boldsymbol{x}'). \tag{3.8}$$

(3.8) shows that $T(\boldsymbol{x})$ is generated as the output of a "time-varying convolution filter" driven by $\mathrm{dB}(\boldsymbol{x})$, the zero-mean *"white noise process"*.

We write the autocovariance of $T(\boldsymbol{x})$ as follows:

$$\begin{aligned}\mathcal{R}_T(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{E}[T(\boldsymbol{x})\overline{T(\boldsymbol{x}')}] = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}'')\cdot\boldsymbol{\xi}} S_T(\boldsymbol{x}, \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \mathrm{E}[\mathrm{dB}(\boldsymbol{x}'')\mathrm{dB}(\boldsymbol{x}''')] \\ \times \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{x}''')\cdot\boldsymbol{\xi}'} S_T(\boldsymbol{x}', \boldsymbol{\xi}')\, \mathrm{d}\boldsymbol{\xi}'\end{aligned} \tag{3.9}$$

where E denotes the expectation operator. Since $\mathrm{B}(\boldsymbol{x})$ is an orthogonal increment process satisfying

$$\mathrm{E}[\mathrm{dB}(\boldsymbol{x})\overline{\mathrm{dB}(\boldsymbol{x}')}] = \delta(\boldsymbol{x} - \boldsymbol{x}')\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'. \tag{3.10}$$

We express (3.9) as

$$\mathcal{R}_T(\boldsymbol{x}, \boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} S_T(\boldsymbol{x}, \boldsymbol{\xi})\overline{S_T(\boldsymbol{x}', \boldsymbol{\xi})}\, \mathrm{d}\boldsymbol{\xi} \tag{3.11}$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \Omega_T$.

Let $\mathcal{T}$ be the pseudo-differential operator whose symbol is $S_T(\boldsymbol{x}, \boldsymbol{\xi})$. Then, (3.11) shows that the autocovariance $\mathcal{R}_T(\boldsymbol{x}, \boldsymbol{x}')$ of $T(\boldsymbol{x})$ is the kernel of the operator $\mathcal{T}\mathcal{T}^\dagger$ where $\mathcal{T}^\dagger$ is the L^2-adjoint of $\mathcal{T}$.

We make another interesting observation by rearranging the terms of (3.11). Let $\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{\tau}/2$, $\boldsymbol{x}' \to \boldsymbol{x} - \boldsymbol{\tau}/2$. Then (3.11) becomes

$$\mathcal{R}_T(\boldsymbol{x} + \boldsymbol{\tau}/2, \boldsymbol{x} - \boldsymbol{\tau}/2) = \int \mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{\tau}\cdot\boldsymbol{\xi}} S_T(\boldsymbol{x} + \boldsymbol{\tau}/2, \boldsymbol{\xi}) \overline{S_T(\boldsymbol{x} - \boldsymbol{\tau}/2, \boldsymbol{\xi})} \mathrm{d}\boldsymbol{\xi}. \tag{3.12}$$

Taking the Fourier transform of the right hand side of (3.12) with respect to $\boldsymbol{\tau}$, we obtain the Wigner distribution function [81] of $S_T(\boldsymbol{x}, \boldsymbol{\xi})$ with respect to $\boldsymbol{x}$ integrated over $\boldsymbol{\xi}$.

Similarly, we assume that the clutter, $C(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega_C$ where Ω_C is a compact subset of $\mathbb{R}^2$, and it is a zero-mean, pseudo-stationary process with the following spectral representation:

$$C(\boldsymbol{x}) = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} S_C(\boldsymbol{x}, \boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} \mathrm{dB}(\boldsymbol{x}') \tag{3.13}$$

with the autocovariance function $\mathcal{R}_C$ and the space-varying spectral density function $|S_C(\boldsymbol{x}, \boldsymbol{\xi})|^2$ by making similar assumptions while defining S_T.

3.2.2 Model for the Measurement Noise

We model the additive measurement noise to be a zero-mean, stationary process in fast-time and statistically mutually independent in slow-time. We write

$$S_n(s, \omega) = \sigma(s)\tilde{S}_n(\omega). \tag{3.14}$$

To avoid peculiar behavior, we assume that

$$\int \left|\tilde{S}_n(\omega)\right|^2 \mathrm{d}\omega < +\infty. \tag{3.15}$$

Note that $\sigma(s)$ allows noise process to have a different variance for each slow-time $s \in [s_0, s_1]$. The autocovariance function of noise is denoted by

$$\mathrm{E}[n(s,t)\overline{n(s',t')}] = \sigma^2(s)\delta(s-s')\mathcal{R}_n(t,t') \tag{3.16}$$

where

$$\mathcal{R}_n(t,t') = \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-t')}|\tilde{S}_n(\omega)|^2 \,\mathrm{d}\omega. \tag{3.17}$$

and we refer to $|S_n(\omega)|^2$ as the *power spectral density function of noise*. Finally, without loss of generality; we assume that the target, clutter and noise are mutually statistically independent.

3.3 Filter Design Based on MMSE Criteria

Since the forward operator $\mathcal{F}$ is a Fourier integral operator, we can form an image T_{ns} of the target scene T by another Fourier integral operator $\mathcal{K}_{ns}$ that is of filtered-backprojection-type as follows:

$$T_{ns}(\boldsymbol{z}) := \mathcal{K}_{ns}[d](\boldsymbol{z}) := \int Q_{ns}(\boldsymbol{z}, s, \omega) \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{z})/c)} d(s,t) \mathrm{d}\omega \mathrm{d}t \mathrm{d}s, \tag{3.18}$$

where Q_{ns} is a filter that satisfies a symbol estimate similar to the one in (2.20). Since $\mathcal{K}_{ns}$ is an FBP-type operator, we call T_{ns} the FBP image of T.

We form a second image T_{bp} of T as follows:

$$T_{bp}(\boldsymbol{z}) = \mathcal{K}_{bp}[d](\boldsymbol{z}) := \int Q_{bp}(\boldsymbol{z}, \boldsymbol{x}') \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{x}')/c)} d(s,t) \mathrm{d}\omega \mathrm{d}t \mathrm{d}s \mathrm{d}\boldsymbol{x}' \tag{3.19}$$

where Q_{bp} is the kernel of a pseudo-differential operator with symbol $\hat{Q}_{bp}$, i.e.,

$$Q_{bp}(\boldsymbol{z}, \boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}'} \hat{Q}_{bp}(\boldsymbol{z}, \boldsymbol{\xi}') \mathrm{d}\boldsymbol{\xi}'. \tag{3.20}$$

We assume that $\hat{Q}_{bp}$ satisfies a symbol estimate similar to the one in (2.20). Since $\mathcal{K}_{bp}$ performs backprojection followed by filtering with Q_{bp}, we refer to T_{bp} as the BPF image of T.

Our objective is to design the filters Q_i, for Q_{ns} and Q_{bp} so that the following mean square error (MSE) of the reconstructed images are minimized:

$$\mathcal{J}(Q_i) = \int \mathrm{E}\left[|T_i(\boldsymbol{z}) - T(\boldsymbol{z})|^2\right] \mathrm{d}\boldsymbol{z}. \tag{3.21}$$

The image T_i is related to the target T as follows:

$$T_i = \mathcal{K}_i[d] = \mathcal{K}_i[\mathcal{F}[(T+C)+n]] = \mathcal{K}_i\mathcal{F}[T+C] + \mathcal{K}_i[n]. \tag{3.22}$$

Inserting (3.22) for T_i into (3.21), we obtain

$$\begin{aligned} \mathcal{J}(Q_i) &= \int \mathrm{E}\left[|\mathcal{K}_i\left[\mathcal{F}[T+C]+n\right](\boldsymbol{z}) - T(\boldsymbol{z})|^2\right] \mathrm{d}\boldsymbol{z} \\ &= \mathcal{J}_T(Q_i) + \mathcal{J}_C(Q_i) + \mathcal{J}_n(Q_i), \end{aligned} \tag{3.23}$$

where

$$\mathcal{J}_T(Q_i) = \int \mathrm{E}[|(\mathcal{K}_i\mathcal{F} - \mathcal{I}_\Omega)[T](\boldsymbol{z})|^2]\,\mathrm{d}\boldsymbol{z} \tag{3.24}$$

$$\mathcal{J}_C(Q_i) = \int \mathrm{E}[|\mathcal{K}_i\mathcal{F}[C](\boldsymbol{z})|^2]\,\mathrm{d}\boldsymbol{z} \tag{3.25}$$

$$\mathcal{J}_n(Q_i) = \int \mathrm{E}[|\mathcal{K}_i[n](\boldsymbol{z})|^2]\,\mathrm{d}\boldsymbol{z}. \tag{3.26}$$

$\mathcal{I}_\Omega$ in (3.24) stands for the *bandlimited identity operator* which will be defined later in this section.[1]

Before we determine the filters Q_{ns} and Q_{bp}, we simplify $\mathcal{K}_i\mathcal{F}$ as in [1], [25], [46] and next approximate each of the qualities in (3.24)-(3.26).

Let $f(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega_f$, $\subseteq \mathbb{R}^2$ be a distribution. Under appropriate assumptions on A and Q_{ns} [1], [3], [16], [25], [43]–[45], [67], we write

$$\mathcal{K}_{ns}\mathcal{F}[f](\boldsymbol{z}) = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} Q_{ns}(\boldsymbol{z},\boldsymbol{\xi}) A(\boldsymbol{x},\boldsymbol{\xi}) \eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi}) f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}, \tag{3.27}$$

where

$$(s,\omega) \to \boldsymbol{\xi} = \frac{\omega}{c}\boldsymbol{\Xi}(s,\boldsymbol{x},\boldsymbol{z}). \tag{3.28}$$

and

$$\eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi}) = \left|\frac{\partial(s,\omega)}{\partial\boldsymbol{\xi}}\right|, \tag{3.29}$$

is the Jacobian that comes from the change of variables (3.28). $\boldsymbol{\Xi}$ is given by

$$\frac{\omega}{c}[R(s,\boldsymbol{x}) - R(s,\boldsymbol{z})] = \frac{\omega}{c}(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\Xi}(s,\boldsymbol{x},\boldsymbol{z}), \tag{3.30}$$

where for $\boldsymbol{x} = \boldsymbol{z}$,

$$\boldsymbol{\Xi}(s,\boldsymbol{z},\boldsymbol{z}) = \nabla_{\boldsymbol{z}} R(s,\boldsymbol{z}), \tag{3.31}$$

[1]The mean square errors defined for $\mathcal{K}_{ns}$ and $\mathcal{K}_{bp}$, hence the corresponding functionals $\mathcal{J}$, $\mathcal{J}_T$, $\mathcal{J}_C$ and $\mathcal{J}_n$, defined in (3.21)-(3.26) are not necessarily equal. However, to simplify our notation, we use the same notation for both cases.

and $A(\boldsymbol{x},\boldsymbol{\xi}) = A(\boldsymbol{x}, s(\boldsymbol{\xi}), \omega(\boldsymbol{\xi}))$.

(3.27) shows that $\mathcal{K}_{ns}\mathcal{F}$ as a pseudo-differential operator. The main contributions to $\mathcal{K}_{ns}\mathcal{F}[f]$ comes from $\boldsymbol{x} = \boldsymbol{z}$ [46], [82]. Substituting $\boldsymbol{z}$ for $\boldsymbol{x}$ in $\boldsymbol{\Xi}$, we conclude that the leading-order singularities of $\mathcal{K}_{ns}\mathcal{F}[f]$ is given by

$$\mathcal{K}_{ns}\mathcal{F}[f](\boldsymbol{z}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} Q_{ns}(\boldsymbol{z},\boldsymbol{\xi}) A(\boldsymbol{x},\boldsymbol{\xi}) \eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi}) f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{3.32}$$

Given a flight trajectory and the bandwidth, the best possible image (in the least-squares sense) one could reconstruct would be

$$\mathcal{I}_{\Omega}[f](\boldsymbol{z}) := \int_{\Omega_{\boldsymbol{z}}} \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}, \tag{3.33}$$

where $\Omega_{\boldsymbol{z}}$ is the data collection manifold given by

$$\Omega_{\boldsymbol{z}} = \{\boldsymbol{\xi} = \frac{\omega}{c}\nabla_{\boldsymbol{z}} R(s,\boldsymbol{z}) : A(\boldsymbol{z},s,\omega) \neq 0\}. \tag{3.34}$$

We refer to $\mathcal{I}_{\Omega}$ as the *bandlimited identity operator* and denote its kernel by $\tilde{\chi}_{\Omega}$.

Following the steps in (3.27)-(3.31), we express $\mathcal{K}_{bp}\mathcal{F}$ as follows:

$$\mathcal{K}_{bp}\mathcal{F}[f] \approx \int Q_{bp}(\boldsymbol{z},\boldsymbol{x}') \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} A(\boldsymbol{x}',\boldsymbol{\xi}) \eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi}) f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}'. \tag{3.35}$$

Inserting

$$Q_{bp}(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}'} \hat{Q}_{bp}(\boldsymbol{z},\boldsymbol{\xi}') \mathrm{d}\boldsymbol{\xi}' \tag{3.36}$$

into (3.35), we obtain

$$\mathcal{K}_{bp}\mathcal{F}[f] \approx \int \mathrm{e}^{\mathrm{i}2\pi[(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}'+(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}]} \hat{Q}_{bp}(\boldsymbol{z},\boldsymbol{\xi}') A(\boldsymbol{x}',\boldsymbol{\xi}) \eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi}) f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}'\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}'. \tag{3.37}$$

We now apply the method of the stationary phase in the variables $\boldsymbol{x}'$ and $\boldsymbol{\xi}'$

simultaneously[2] and obtain the critical points at

$$\boldsymbol{\xi}' = \boldsymbol{\xi} \quad \text{and} \quad \boldsymbol{x}' = \boldsymbol{z}. \tag{3.38}$$

The leading-order term of (3.35) is then

$$\mathcal{K}_{bp}\mathcal{F}[f](\boldsymbol{z}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} \hat{Q}_{bp}(\boldsymbol{z},\boldsymbol{\xi}) A(\boldsymbol{z},\boldsymbol{\xi}) \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi}) f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}. \tag{3.39}$$

Note that (3.32) and (3.39) show that the operators $\mathcal{K}_{ns}\mathcal{F}$ and $\mathcal{K}_{bp}\mathcal{F}$ are the same to the leading-order, i.e., $\mathcal{K}_{ns}\mathcal{F}[f]$ and $\mathcal{K}_{bp}\mathcal{F}[f]$ differ only by a smoother function.

Having simplified $\mathcal{K}_{ns}$ and $\mathcal{K}_{bp}$, we now approximate each term given in (3.24)-(3.26).

Lemma 1: Let the images T_{ns} and T_{bp} be formed as in (3.18) and (3.19) where the filters Q_{ns} and Q_{bp} satisfy symbol estimates similar to the one in (2.20). We assume that the geometric conditions on the flight trajectories and the antenna beam patterns satisfy certain conditions such that artifacts are avoided [46], [82].

(i) Then, the leading-order singularities of each term in the mean square error is given by

$$\mathcal{J}_T(Q_{ns}) \approx \int \left| Q_{ns}(\boldsymbol{x},\boldsymbol{\xi}) A(\boldsymbol{x},\boldsymbol{\xi}) \eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi}) - \tilde{\chi}_{\Omega}(\boldsymbol{x},\boldsymbol{\xi}) \right|^2 |S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{3.40}$$

$$\mathcal{J}_C(Q_{ns}) \approx \int \left| Q_{ns}(\boldsymbol{x},\boldsymbol{\xi}) A(\boldsymbol{x},\boldsymbol{\xi}) \eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi}) \right|^2 |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{3.41}$$

$$\mathcal{J}_n(Q_{ns}) \approx \int |Q_{ns}(\boldsymbol{x},\boldsymbol{\xi})|^2 |S_n(\boldsymbol{\xi})|^2 \eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{3.42}$$

where $Q_{ns}(\boldsymbol{x},\boldsymbol{\xi}) = Q_{ns}(\boldsymbol{x}, s(\boldsymbol{\xi}), \omega(\boldsymbol{\xi}))$, $A(\boldsymbol{x},\boldsymbol{\xi}) = A(\boldsymbol{x}, s(\boldsymbol{\xi}), \omega(\boldsymbol{\xi}))$ and $\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})$ is the Jacobian that comes from the change of variables (3.28).

[2] In order to apply the method of the stationary phase the determinant of the Hessian of the phase function must be non-zero. Here, the Hessian of the phase function $\phi_1(\boldsymbol{x},\boldsymbol{\xi}',\boldsymbol{\xi},\boldsymbol{x}') = (\boldsymbol{x}' - \boldsymbol{z})\cdot\boldsymbol{\xi}' + (\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}$ is given by $H(\phi_1) = \begin{bmatrix} \nabla_{\boldsymbol{\xi}'^2}\phi_1 & \nabla_{\boldsymbol{\xi}'\boldsymbol{x}'}\phi_1 \\ \nabla_{\boldsymbol{x}'\boldsymbol{\xi}'}\phi_1 & \nabla_{\boldsymbol{x}'^2}\phi_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}$ where $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\det(H(\phi_1)) = -1$.

(ii) The leading-order singularities of $\mathcal{J}_T(Q_{bp})$, $\mathcal{J}_C(Q_{bp})$ and $\mathcal{J}_n(Q_{bp})$ are given by (3.40)-(3.42) with Q_{ns} replaced by $\hat{Q}_{bp}$ where

$$Q_{bp}(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}'}\hat{Q}_{bp}(\boldsymbol{z},\boldsymbol{\xi}')\mathrm{d}\boldsymbol{\xi}'. \tag{3.43}$$

Note that $\mathcal{J}_{T,C,n}(Q_{bp})$ is equal to $\mathcal{J}_{T,C,n}(Q_{ns})$ up to the leading-order terms. They differ only by a smoother function.

Proof: See Appendix B.

Theorem 1: Let the data d be given by (3.2) and $T_i(\boldsymbol{z})$ for T_{ns} and T_{bp} be as defined in (3.18) and (3.19).

(i) Then, the following filter minimizes the leading-order MSE $\mathcal{J}(Q_{ns})$:

$$Q_{ns}(\boldsymbol{z},\boldsymbol{\xi}) = \frac{\overline{A(\boldsymbol{z},\boldsymbol{\xi})}|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2\tilde{\chi}_\Omega}{|A(\boldsymbol{z},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})[|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2+|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2]+|S_n(\boldsymbol{\xi})|^2} \tag{3.44}$$

where $\boldsymbol{\xi}\in\Omega_{\boldsymbol{z}}$, $\Omega_{\boldsymbol{z}}$ is given by (3.34), $|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2$ and $|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2$ are the SVSD functions of target and clutter defined in (3.11) and (3.13), $S_n(\boldsymbol{\xi})$ is the noise power spectral density function defined in (3.17) and $\tilde{\chi}_\Omega$ is a smooth cut-off function that prevents division by zero.

(ii) The leading-order singularities of the filter Q_{bp} that minimizes the (leading-order) MSE $\mathcal{J}(Q_{bp})$ is given by

$$Q_{bp}(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{z}-\boldsymbol{x}')\cdot\boldsymbol{\zeta}}Q_{ns}(\boldsymbol{z},\boldsymbol{\zeta})\mathrm{d}\boldsymbol{\zeta}. \tag{3.45}$$

(iii) With these choices of filters, the leading-order MSE of $\mathcal{J}(Q_i)$ for Q_{ns} and Q_{bp} is given by

$$\mathcal{J}(Q_i) \approx \int \frac{\alpha(\boldsymbol{x},\boldsymbol{\xi})|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2}{|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2+\alpha(\boldsymbol{x},\boldsymbol{\xi})}\tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{3.46}$$

where

$$\alpha(\boldsymbol{x},\boldsymbol{\xi}) = |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 + \frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})}. \tag{3.47}$$

Proof: See Appendix C.

3.4 The Image Reconstruction Algorithm

The filters derived in the previous section require a priori information, specifically the SVSD functions of target and clutter and noise power spectral density function. In many radar applications, it is assumed that the prior information on clutter can be obtained by collecting radar data in the absence of the target in the scene of interest [26], [48], [49], [51], [52]. Similarly, thermal noise prior information can be obtained in the absence of scattered field data. However, a priori information on the target is often not available. In this section, we describe an algorithm to estimate the target SVSD function and to reconstruct target scene simultaneously using the measured data. Finally, we describe the computational complexity of the algorithm in comparison with the algorithms available in the literature.

3.4.1 The Estimation of the Space-Varying Spectral Density Function

In this subsection, we briefly describe a method introduced in [41, Ch. 11] for the spectral density function estimation of non-stationary processes. The method can be viewed as a straightforward extension of the spectral density function estimation for stationary processes.

Let $\Phi(\boldsymbol{x})$ be a square integrable, compactly supported windowing function and $f(\boldsymbol{x})$ denote a realization of a zero-mean pseudo-stationary process. We define

$$Y(\boldsymbol{x},\boldsymbol{\xi}) = \int_{\Omega_\Phi} \mathrm{e}^{\mathrm{i}\boldsymbol{\xi}\cdot(\boldsymbol{x}-\boldsymbol{y})} f(\boldsymbol{x}-\boldsymbol{y})\Phi(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \tag{3.48}$$

where $\boldsymbol{y} \in \mathbb{R}^2$ and Ω_Φ is the support of the windowing function Φ. It is assumed that the "width" of the windowing function is much smaller than the support of the observations. (See [41, pp. 837].) (3.48) can be viewed as the time-frequency transform of the observations. It was shown that $|Y(\boldsymbol{x},\boldsymbol{\xi})|^2$ is an unbiased estimate of $S_f(\boldsymbol{x},\boldsymbol{\xi})$, the SVSD function of f. This estimate is analogous to the classical periodgram estimate of the spectral density function of stationary processes. Similar to the results of the classical spectral estimation theory, better bias-variance tradeoffs can be achieved by convolving $|Y(\boldsymbol{x},\boldsymbol{\xi})|^2$ with another windowing function. (See [83, pp. 838-839] for details.) A detailed implementation of the method described

in [83] to our model can be found in Appendix E.

3.4.2 The Steps of the Algorithm

We observe that the filter Q_{ns} can be factored into two components as follows:

$$Q_{ns}(\boldsymbol{z},\boldsymbol{\xi}) = Q^1_{ns}(\boldsymbol{z},\boldsymbol{\xi})Q^2_{ns}(\boldsymbol{z},\boldsymbol{\xi}) \tag{3.49}$$

where

$$Q^1_{ns}(\boldsymbol{z},\boldsymbol{\xi}) = \frac{\overline{A(\boldsymbol{z},\boldsymbol{\xi})}\tilde{\chi}_\Omega(\boldsymbol{z},\boldsymbol{\xi})}{|A(\boldsymbol{z},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})}, \tag{3.50}$$

$$Q^2_{ns}(\boldsymbol{z},\boldsymbol{\xi}) = \frac{|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2}{|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2 + |S_C(\boldsymbol{z},\boldsymbol{\xi})|^2 + \frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{z},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})}} \tag{3.51}$$

Q^1_{ns} is the filter derived in [1] under the assumptions that the target is deterministic and the received data do not have noise or clutter components. Clearly, this filter does not involve any target information. The second filter, Q^2_{ns}, is a low pass filter that can be expressed solely by the target-to-clutter ratio (SCR) and target-to-noise ratio (SNR). Dividing the numerator and the denominator by the target SVSD, we obtain

$$Q^2_{ns}(\boldsymbol{z},\boldsymbol{\xi}) = \left[\mathrm{SCR}(\boldsymbol{z},\boldsymbol{\xi})^{-1} + (\mathrm{SNR}(\boldsymbol{z},\boldsymbol{\xi})|A(\boldsymbol{z},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi}))^{-1}\right]^{-1} \tag{3.52}$$

where

$$\mathrm{SCR}(\boldsymbol{z},\boldsymbol{\xi}) = |S_T(\boldsymbol{z},\boldsymbol{\xi})|^2 / \left[|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2 + |S_C(\boldsymbol{z},\boldsymbol{\xi})|^2\right] \tag{3.53}$$

$$\mathrm{SNR}(\boldsymbol{z},\boldsymbol{\xi}) = S_T(\boldsymbol{z},\boldsymbol{\xi})|^2 / |S_n(\boldsymbol{\xi})|^2. \tag{3.54}$$

Q^2_{ns} is the component of the filter Q_1 that suppresses noise and clutter. If a priori information on SCR and SNR is available, then this information can be used to build the filter Q^2_{ns}. Otherwise, the target SVSD function can be estimated from the data itself to build the filter Q^2_{ns}. To facilitate the estimation of the target SVSD, we

define the following images:

$$\tilde{T}_0(\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{x}')/c)} d(s,t) Q(\boldsymbol{x}', s(\boldsymbol{\xi}), t(\boldsymbol{\xi})) \mathrm{d}\omega \mathrm{d}t \mathrm{d}s \tag{3.55}$$

and

$$\tilde{T}(\boldsymbol{z}) = \int Q^2_{bp}(\boldsymbol{z}, \boldsymbol{x}') \tilde{T}_0(\boldsymbol{x}') \mathrm{d}s \mathrm{d}\boldsymbol{x}' \tag{3.56}$$

where

$$Q^2_{bp}(\boldsymbol{z}, \boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{z}-\boldsymbol{x}')\cdot\boldsymbol{\zeta}} Q^2_{ns}(\boldsymbol{z}, \boldsymbol{\zeta}) \mathrm{d}\boldsymbol{\zeta}. \tag{3.57}$$

The first and second order statistics of the image T_{ns} and the image T_{bp} given in (3.18) have the same leading-order singularities and only differ by smoother functions. To build the filter Q^2_{ns}, we use the image $\tilde{T}_0$ to estimate the target SVSD as described in the previous section.

A pseudo-code describing the estimation of the target SVSD and the reconstruction of the images $\tilde{T}_0$ and $\tilde{T}$ is described in Algorithm 1.

3.4.3 The Computational Complexity of the Algorithm

The computational complexity of our method is determined by the following major steps: estimating the SVSD functions, filtering in the Fourier domain using Q^1_{ns}, the backprojection operation and filtering using Q^2_{bp}. Below, we summarize the computational complexity of each of these steps.

We assume that the image to be reconstructed is $N \times N$ and that the measured data have $\mathcal{O}(N)$ samples in both the fast-time and slow-time variables.

1. As described in [1], the filtering in the Fourier domain and backprojecting the filtered data can be computed with $\mathcal{O}(N^2 \log N)$ computational complexity using either the fast-backprojection [39], [40] or the fast FIO calculation algorithms [84].

2. The estimation of the SVSD function requires computing the magnitude of the Fourier transform around each pixel within a window. Assuming that the window size is $m \times m$, the estimation of the SVSD function has the computational

Algorithm 1 SAR image reconstruction algorithm for non-stationary target and clutter.

$\tilde{T}_0 \leftarrow \mathbf{0}_{X_1 \times X_2}$, $\tilde{T} \leftarrow \mathbf{0}_{X_1 \times X_2}$
Define $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{x}' = (x_1', x_2')$
Define $\mathbf{x} = [\boldsymbol{x}, \psi(\boldsymbol{x})]$
%Form $\tilde{T}_0$ using the filter Q_{ns}^1 as in [1]
%Form $\tilde{C}_0$ using the filter Q_{ns}^1 as in [1] in the absence of target
for $x_1, x_2 = 1 \rightarrow X_1, X_2$ **do**
  for $s = 1 \rightarrow N_s$ **do**
    $R \leftarrow |\boldsymbol{\gamma}_i(s) - \mathbf{x}| + |\boldsymbol{\gamma}_j(s) - \mathbf{x}|$
    $1/\eta \leftarrow \omega/c^2 |\Xi_1 \dot{\Xi}_2 - \dot{\Xi}_1 \Xi_2|$
    $D(\omega, s) \leftarrow \mathrm{FFT}\{d(:, s)\}$
    $T_\eta(t, s) \leftarrow \mathrm{IFFT}\{D(\omega, s) R/\eta(\omega, s)\}$
    $\tilde{T}_0(\boldsymbol{x}, s) \leftarrow T_\eta(R/c, s)$, $t = R/c$
    $\tilde{T}_0(\boldsymbol{x}) \leftarrow \tilde{T}_0(\boldsymbol{x}) + \tilde{T}_0(\boldsymbol{x}, s)$
% Form $\tilde{T}$ as in (3.56)
for $x_1, x_2 = 1 \rightarrow X_1, X_2$ **do**
  % Estimate target and clutter SVSD as in [41]
  $T_\Phi(\boldsymbol{x}, \boldsymbol{\xi}) \leftarrow \mathrm{FFT}\{\Phi(\boldsymbol{u}) \tilde{T}_0(\boldsymbol{x} - \boldsymbol{u})\}$ & $S_T \leftarrow |T_\Phi|^2$
  $C_\Phi(\boldsymbol{x}, \boldsymbol{\xi}) \leftarrow \mathrm{FFT}\{\Phi(\boldsymbol{u}) C_\Phi(\boldsymbol{x} - \boldsymbol{u})\}$ & $S_C \leftarrow |C_\Phi|^2$
  $\boldsymbol{\xi} \leftarrow \omega/c[\widehat{(\boldsymbol{\gamma}_i(s) - \mathbf{z})} + \widehat{(\boldsymbol{\gamma}_j(s) - \mathbf{z})}]$
  $Q_{ns}^2(\boldsymbol{x}, \boldsymbol{\xi}) \leftarrow S_T(\boldsymbol{x}, \boldsymbol{\xi}) / (S_T(\boldsymbol{x}, \boldsymbol{\xi}) + S_C(\boldsymbol{x}, \boldsymbol{\xi}) + S_n(\boldsymbol{\xi})/\eta)$
  $Q_{bp}^2(\boldsymbol{x}, \boldsymbol{x}') \leftarrow \mathrm{IFFT}\{Q_{ns}^2\}$
  $\tilde{T}(x_1, x_2) \leftarrow \sum_{\boldsymbol{x}'} \tilde{T}(\boldsymbol{x}') Q(\boldsymbol{x}, \boldsymbol{x}')$
return $\tilde{T}$

complexity of $\mathcal{O}(N^2 m^2 \log m)$. Since $m \ll N$, the computational complexity of this step is $\mathcal{O}(N^2)$.

3. Q_{bp}^2 is the kernel of a pseudo-differential operator. Therefore, the computational complexity of filtering with Q_{bp}^2 is $\mathcal{O}(N^2 \log N)$ when fast-backprojection of fast FIO calculation algorithms are used.

Thus, our algorithm can be implemented with the computational complexity of $\mathcal{O}(N^2 \log N)$.

The feature-preserving reconstruction techniques, such as the one in [50], can be implemented with the computational complexity of $\mathcal{O}(N^4)$-$\mathcal{O}(N^6)$ depending on the optimization method and the nature of the regularization employed. STAP is one of the most widely used noise/clutter suppression techniques in moving target

imaging with a computational complexity of $\mathcal{O}(N^4)$ [85]. Parametric STAP techniques reduced the computational complexity from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^3)$ [29]. Computational complexity of the spectral estimation-based methods are generally governed by matrix inversions. For an N^2-by-N^2 matrix, inversion with Gauss-Jordan elimination is $\mathcal{O}(N^6)$. Recently developed spectral estimation-based methods, such as the ones in [64], [86], [87], have the computational complexity of $\mathcal{O}(N^4)$, $\mathcal{O}(N^4)$, $\mathcal{O}(N^6)$, respectively. In [53], computational complexities of the methods described in [64], [87] were reduced to $\mathcal{O}(N^2 \log N)$-$\mathcal{O}(N^3)$ and $\mathcal{O}(N^2 \log N)$-$\mathcal{O}(N^5)$ depending on the length of the radar data and the frequency sampling. The linear programming and orthogonal matching pursuit based techniques described in [88] have the computational complexity of $\mathcal{O}(N^{7/2})$, $\mathcal{O}(N^3)$. The subspace pursuits based technique presented in [88] has the computational complexity of $\mathcal{O}(N^2 \log N)$-$\mathcal{O}(N^3)$, depending on the sparsity of the underlying data. Thus, our algorithm has a lower computational complexity than those of the algorithms in [50], [64], [86]–[88] and it is $\mathcal{O}(N^2 \log N)$ without "any restriction" on the data.

3.5 Numerical Simulations

We performed numerical simulations to evaluate the performance of our image reconstruction method and compared the results with the methods presented in [1] and [25].

3.5.1 Simulation Setup

We consider a 100m×100m scene as shown in Figure 3.1(a). The scene is discretized into 64×64 pixels. A bi-static antenna system, with $\pi/18$ separation between the transmitter and receiver, traverses a circular trajectory given by $\boldsymbol{\gamma}(s) = [11\cos s, 11\sin s, 6.5]\,\text{km}$. The trajectory is uniformly sampled for $s \in [0, 2\pi]$ at 128 points. We set $A(\boldsymbol{x}, \omega, s) \equiv 1$, which corresponds to an isotropic antenna radiating a delta-like impulse and precompensation of the geometric spreading factors in the data.

The target is the airplane-like figure placed at the center of the scene as shown in Figure 3.1(a). Three rectangles shown in Figure 3.2(a)) form the localized clutter patches. The reflectivity of the clutter patches is generated using the Rayleigh distribution, a commonly used statistical model for clutter in radar data [89]. The additive thermal noise is white Gaussian in both fast-time and slow-time parameters.

3.5.2 Evaluation Method

We reconstructed SAR images using four different image reconstruction methods and compared the results. These methods are as follows:

- Method 1: We used the reconstruction method described in [1]. This method assumes that the scene is deterministic and the measurements are ideal without any noise or clutter components.

- Method 2: We used the method described in [25]. This method assumes that the target and clutter are statistically stationary processes and that their spectral density functions are known *a priori*.

- Method 3: We used the FBP method with filter Q_{ns} given in (3.44) under the assumptions that the target and clutter are pseudo-stationary processes and

that their SVSD functions are known a priori. We determined these SVSD functions from target- and clutter-only images.

- Method 4: We used the algorithm described in Section 5. In this case, the target and clutter SVSD functions were estimated from the images reconstructed by Method 1 described above.

In Method 3 and 4, the SVSD function of target and clutter were estimated using an 11-by-11 Bartlett window [83, Sec. 9].

We generated data at different signal-to-clutter and signal-to-noise ratios. These measured quantities are defined as follows:[3]

$$\text{SNR} = 20\log_{10}\left\{\frac{\frac{1}{N_S N_T}\sum_{i=1}^{N_S}\sum_{j=1}^{N_T}(d(s_i,t_j)-\mu_d)^2}{\sum_{i=1}^{N_S}\sum_{j=1}^{N_T} n^2(s_i,t_j)}\right\}, \tag{3.58}$$

$$\text{SCR} = 20\log_{10}\left\{\frac{\frac{1}{N}\sum_{i=1}^{N}(T(\boldsymbol{x}_i)-\mu_T)^2}{\frac{1}{N}\sum_{i=1}^{N}(C(\boldsymbol{x}_i)-\mu_C)^2}\right\} \tag{3.59}$$

where N_S is the number of slow-time samples, N_T is the number fast-time samples, μ_T is the mean of the target scene, μ_d is the mean of its data. We used the mean-square error (MSE) as a figure of merit to compare the reconstructed images. We define MSE in dB as follows:

$$\text{MSE} = 20\log_{10}\left\{\frac{1}{MN}\sum_{r=1}^{M}\sum_{i=1}^{N}|T(\boldsymbol{x}_i)-\tilde{T}_r(\boldsymbol{x}_i)|^2\right\}, \tag{3.60}$$

where M is the number of realizations used in an experiment, N is the number of pixels and $\tilde{T}_r(\boldsymbol{x})$ is the reconstructed image using the r^{th} realization.

3.5.3 Results

We performed two sets of experiments:

1. In the first set of experiments, the clutter level was kept constant at 10dB SCR, and the noise level was changed from -20dB to 40dB SNR with 4dB

[3]Note that the measured SNR and SCR defined for simulations are related to, but different than the spatially and spectrally resolved SNR and SCR concepts defined in Section 5.

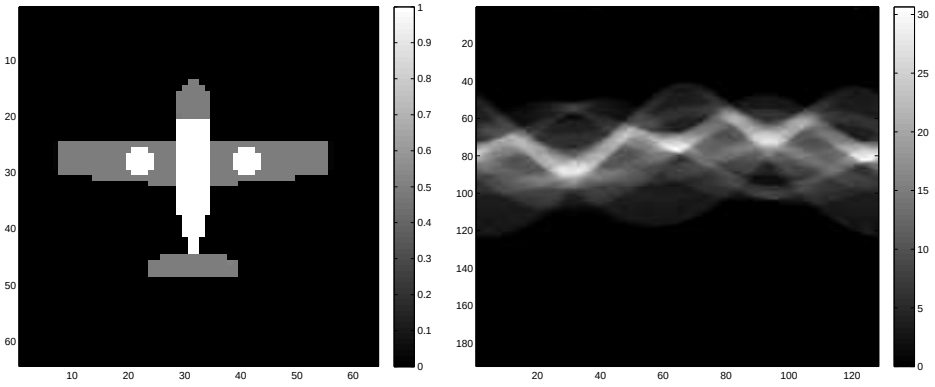
increments. Ten different realizations of noise were produced and added to the clutter contaminated data simulated using (3.2).

2. In the second set of experiments, the noise level was kept constant at 10dB SNR, and the clutter level was changed from -20dB to 40dB SCR with 4dB increments. For each SCR level, ten different realizations of clutter were generated. The measurement data was simulated using (3.2).

We computed the MSE for each SNR and SCR level using the ten reconstructed images, each one corresponding to a different noise or clutter realization. Figure 3.6 and Figure 3.10 show the MSE versus SNR and MSE verssus SCR performances of the four methods described above. We see that for both sets of experiments Method 3 produces the lowest MSE values for all SNR and SCR levels. This is followed by Method 4, Method 2 and Method 1. As SNR (or SCR) increases the difference between the MSE values of Method 1 and Method 2 gets smaller as expected. We also observed that as the SNR (or SCR) value increases, the difference between the MSE values associated with Method 4 and Method 2 becomes larger. This can be explained with better target SVSD estimation at higher SNR levels (or SCR) levels.

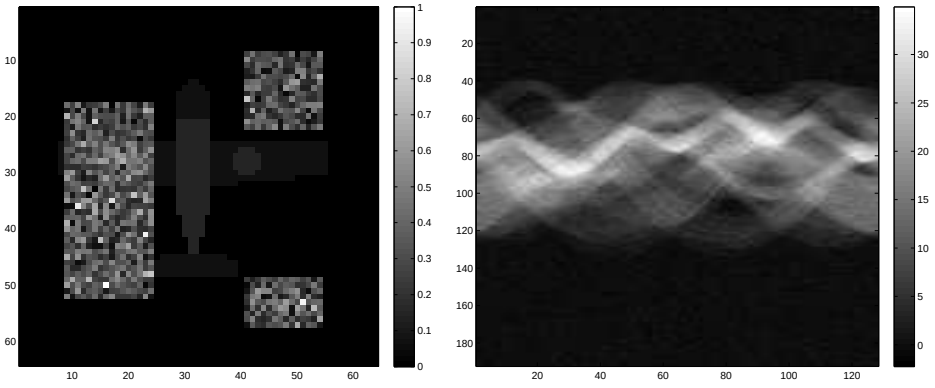
Figures 3.3 - 3.5 show the reconstructed images (averaged over ten realizations) for SNR levels of $-20, 8, 40$dB and 10dB SCR. We see that at -20dB SNR, the method described in [1] cannot recover the scene. At 10dB SNR, however, all four methods can recover the scene, but Method 1 retains singularities induced by noise and clutter and Method 3 produces visually the cleanest image.

Figure 3.7 - 3.9 show the reconstructed images for SCR levels of $-20, 8, 40$dB and 10dB SNR. We see that all four methods recover a visually recognizable scene at all SCR levels. However, at low SCR the clutter patches are visually more prominent than the target in images produced by Method 1, 2 and 4. Only Method 3, which uses SVSD function estimated from a clutter free scene, can suppress clutter and recover target with good contrast. At high SCR levels, the target intensity dominates and therefore all four methods can recover the target scene. Method 3 and 4 perform better than Method 1 and 2 in suppressing the clutter strength.



(a) (b)

Figure 3.1: (a) Target and (b) radar data, where horizontal and vertical lines correspond to slow- and fast-time variables s and t respectively.



(a) (b)

Figure 3.2: (a) Target embedded in clutter and (b) its radar data when SCR is 10dB.

3.5.4 Numerical Simulations with Real Data

As a second set of simulations, we apply the Algorithm 1 to real data provided by Air Force Research Laboratories' Wide Angle SAR experiment [2]. In this experiment, radar data is collected by an X-band system over a parking lot that includes 33 civilian vehicles. The average range to a target is 10km and the radar data provided contains 14243 pulses with 47 frequency bins. Figures 3.11 (a)-(b) show the SAR images reconstructed from wide-angle SAR data [2] by Method 1 and Method 4 for Chevrolet Impala (dataset $PH\_fcara1\_0216.mat$). We observe

that there is strong clutter in the lower-left part of the image reconstructed by FBP. However, this clutter is suppressed when we use the non-stationary reconstruction algorithm.

3.6 Conclusion

In this chapter, we presented novel SAR inversion methods when the data is noisy or clutter is present in the scene. First, we described a novel statistical model for SAR target and clutter. We performed numerical simulations to demonstrate the performance of the method both with synthetic and real data. These simulations also included image reconstruction with three other methods to compare the performance of the method. Numerical simulations show a reduction in the MSE of the images as well as improved image quality compared to other SAR image reconstruction methods.

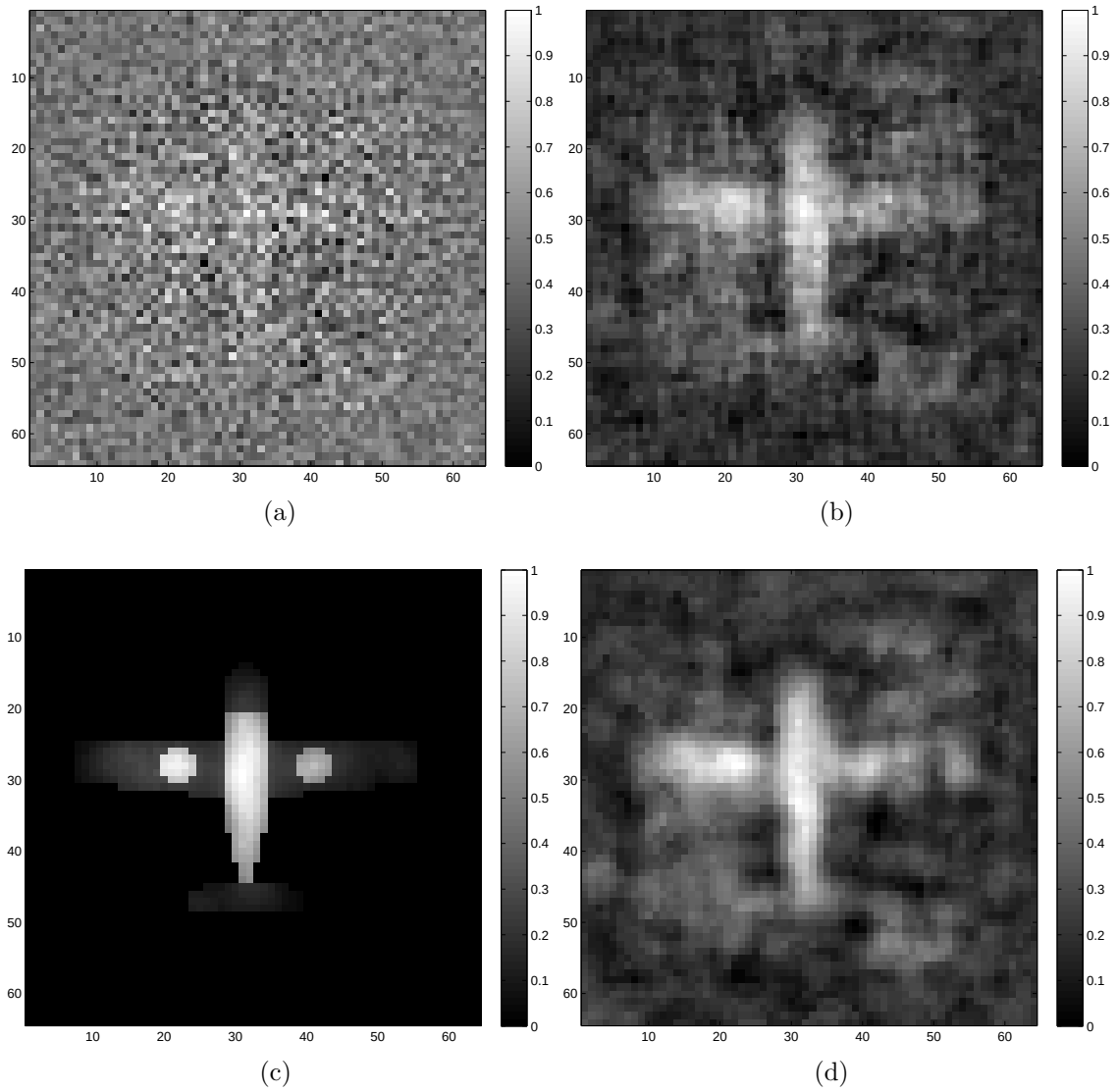


Figure 3.3: Reconstructed images using (a) the method in [1], (b) with stationary target and clutter assumption, (c) non-stationary target and clutter assumption with known SVSD, (d) non-stationary target and clutter assumption with known SVSD when SNR is -20dB and SCR is 10dB.

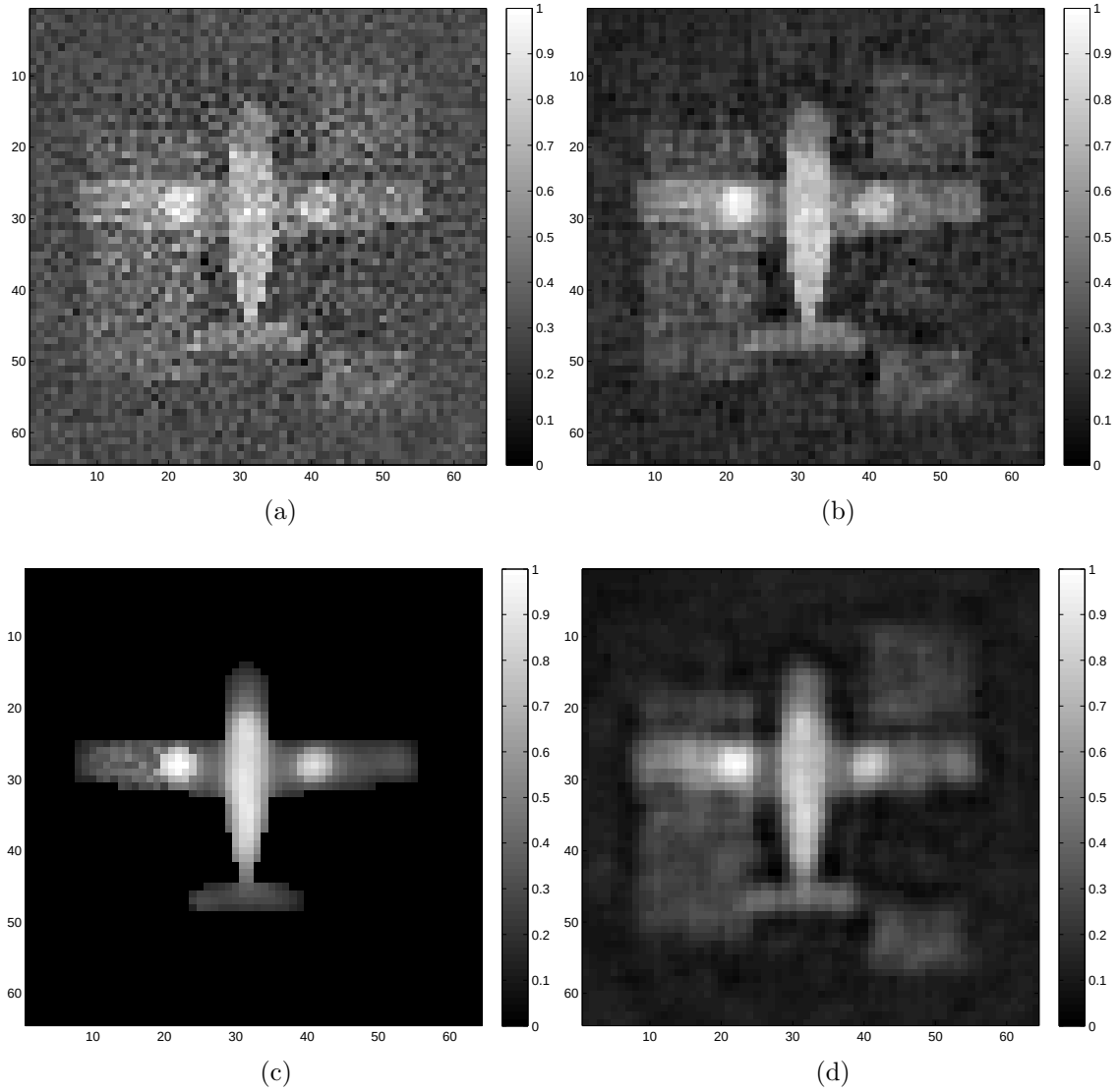


Figure 3.4: Reconstructed images using (a) the method in [1], (b) with stationary target and clutter assumption, (c) non-stationary target and clutter assumption with known SVSD, (d) non-stationary target and clutter assumption with known SVSD when SNR is 8dB and SCR is 10dB.

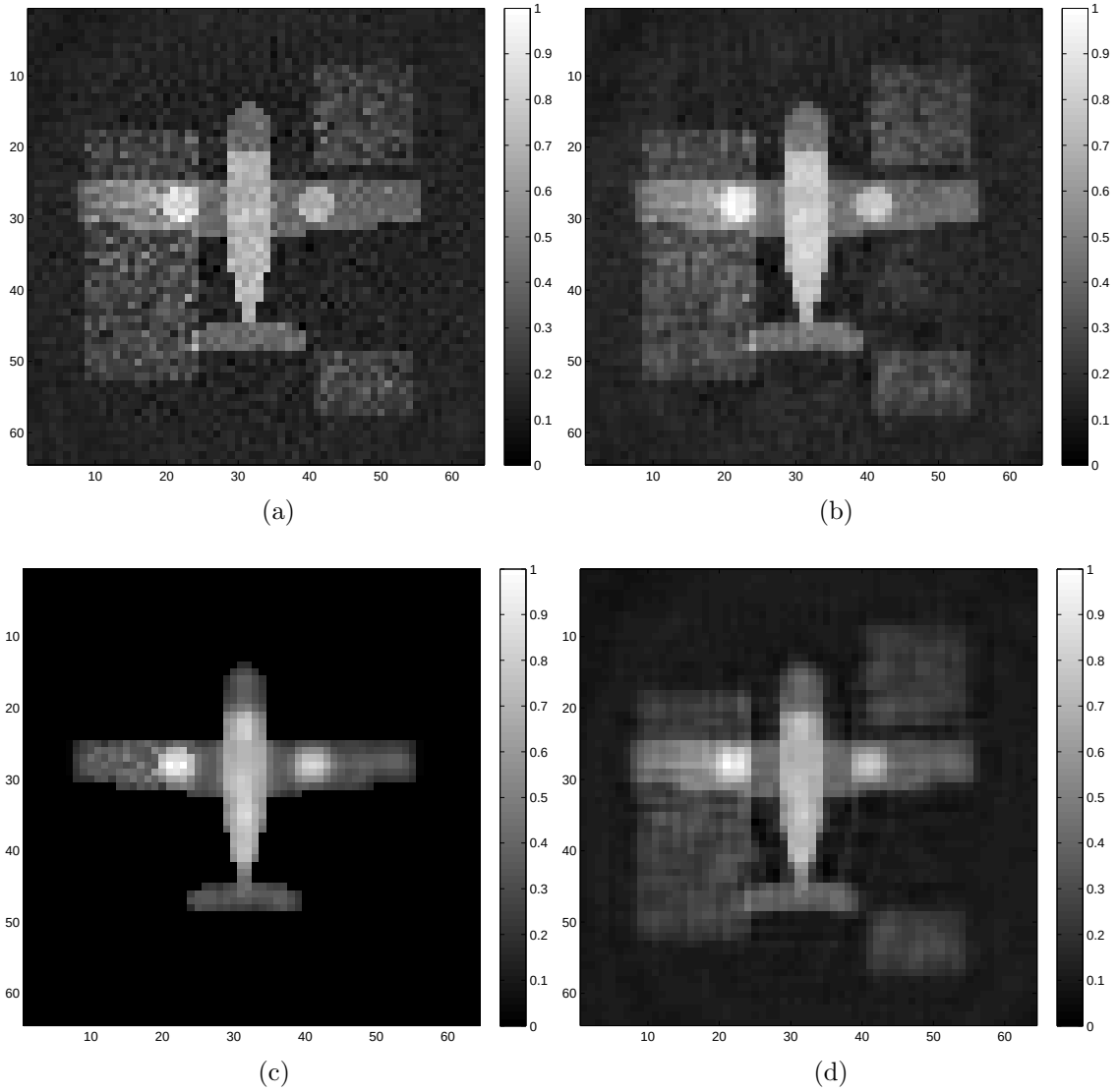


Figure 3.5: Reconstructed images using (a) the method in [1], (b) with stationary target and clutter assumption, (c) non-stationary target and clutter assumption with known SVSD, (d) non-stationary target and clutter assumption with known SVSD when SNR is 40dB and SCR is 10dB.

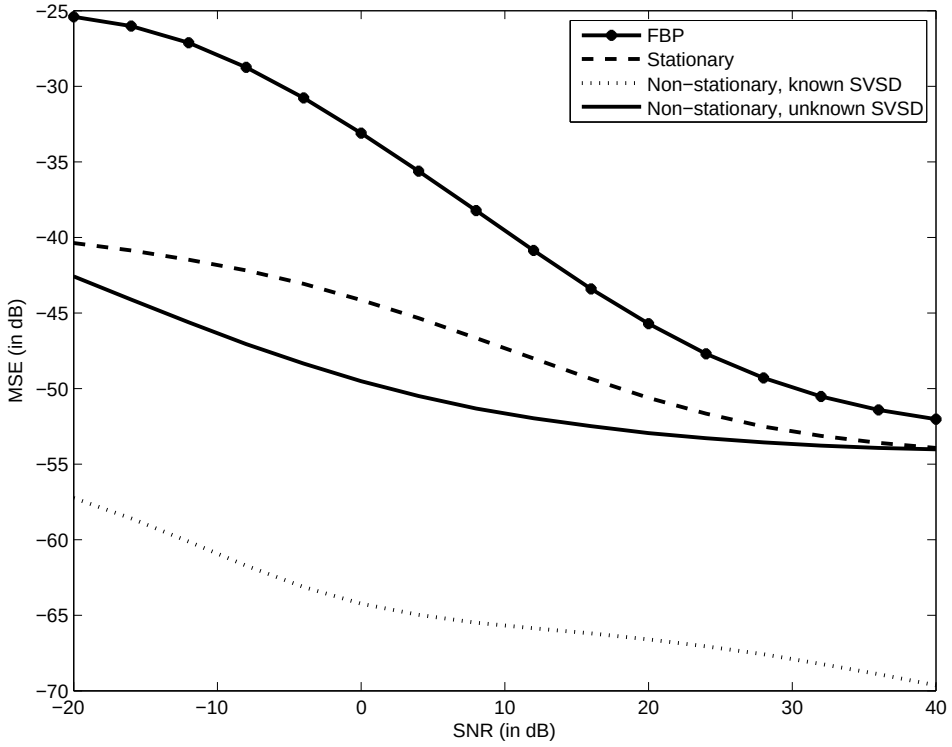


Figure 3.6: MSE (vertical axis, in log scale) versus SNR (horizontal axis) averaged over ten reconstructed images for each SNR level using four different image reconstruction methods.

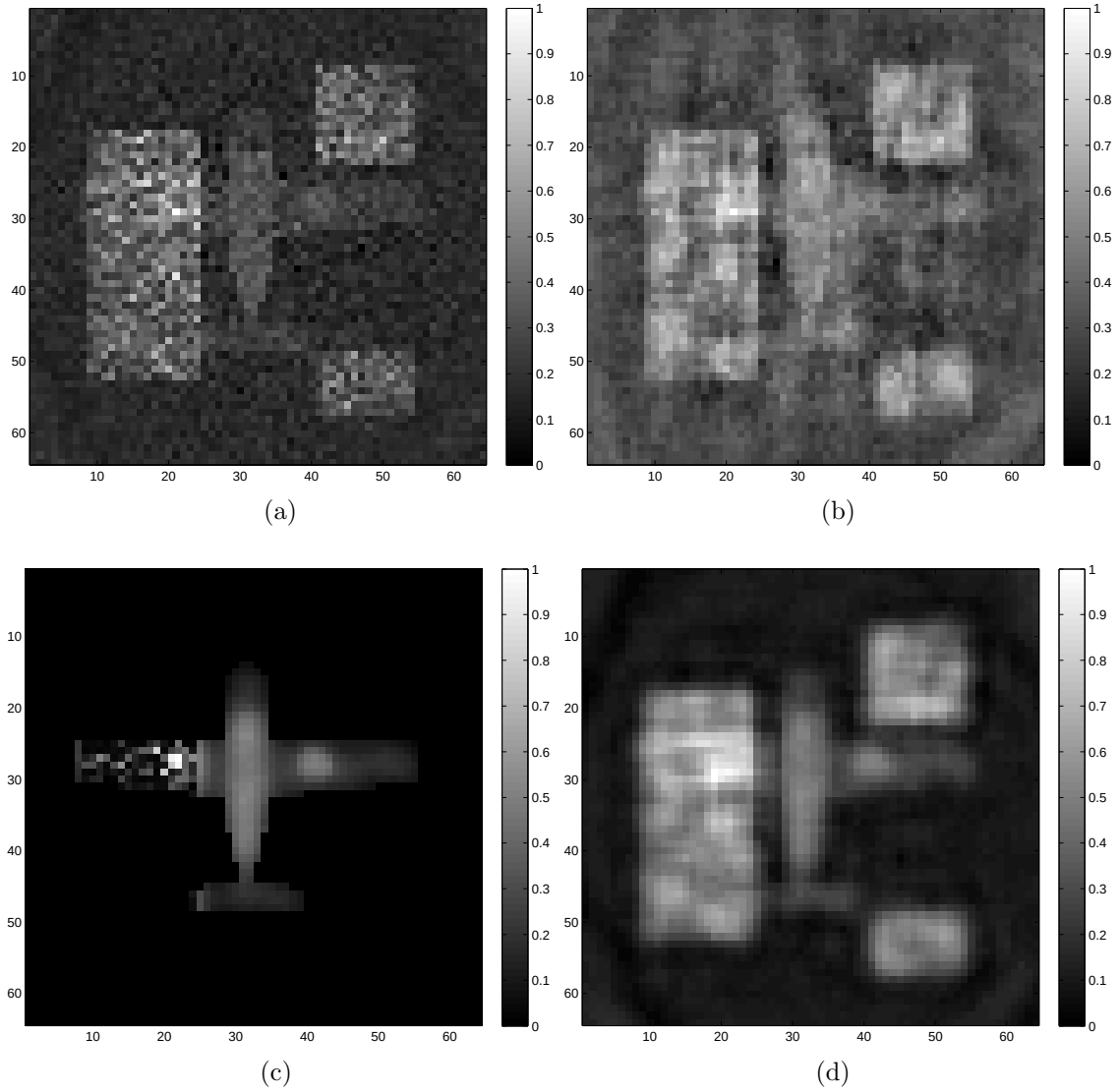


Figure 3.7: Reconstructed images using (a) the method in [1], (b) with stationary target and clutter assumption, (c) non-stationary target and clutter assumption with known SVSD, (d) non-stationary target and clutter assumption with known SVSD when SCR is -20dB and SNR is 10dB.

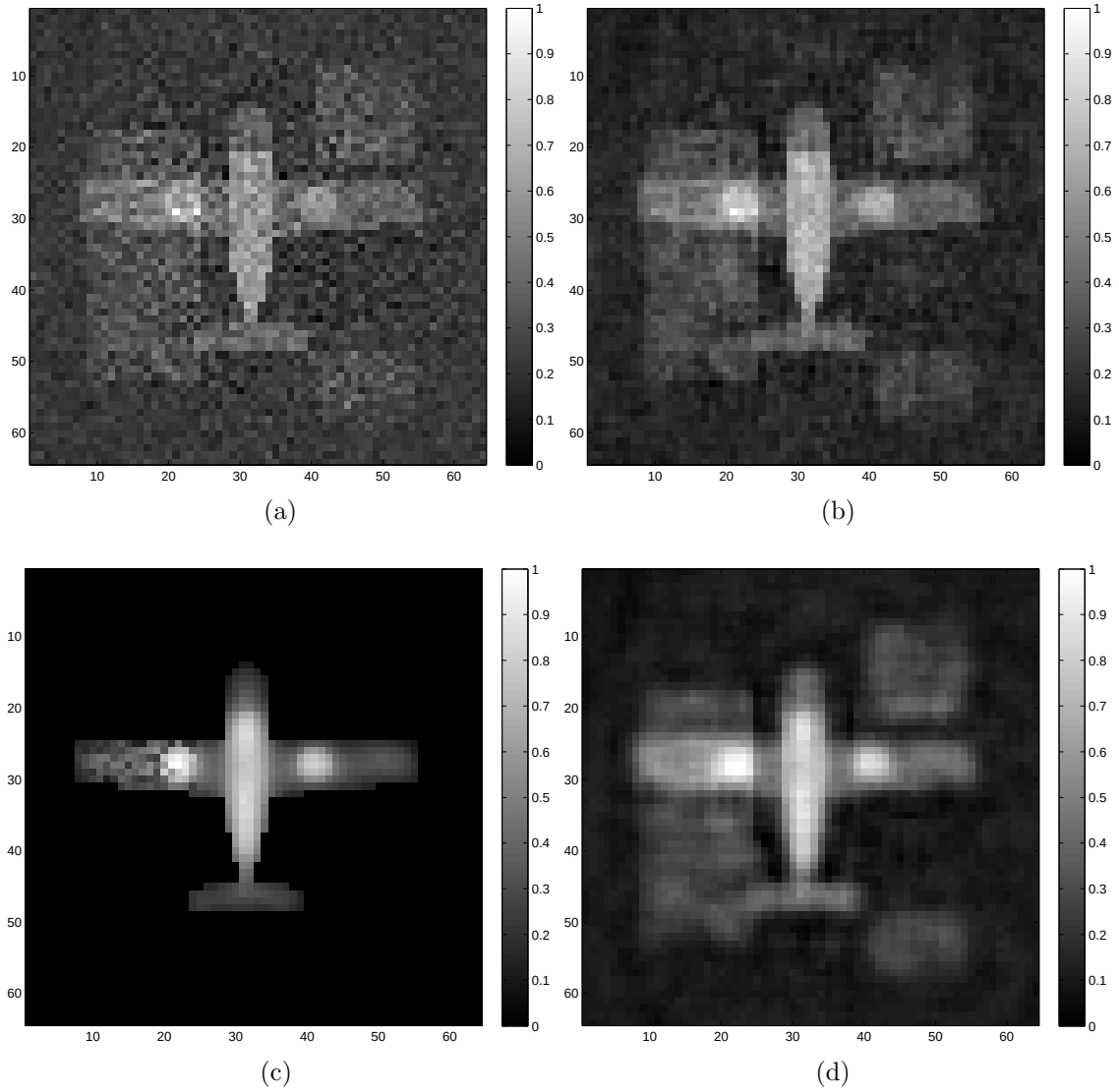


Figure 3.8: Reconstructed images using (a) the method in [1], (b) with stationary target and clutter assumption, (c) non-stationary target and clutter assumption with known SVSD, (d) non-stationary target and clutter assumption with known SVSD when SCR is 8dB and SNR is 10dB.

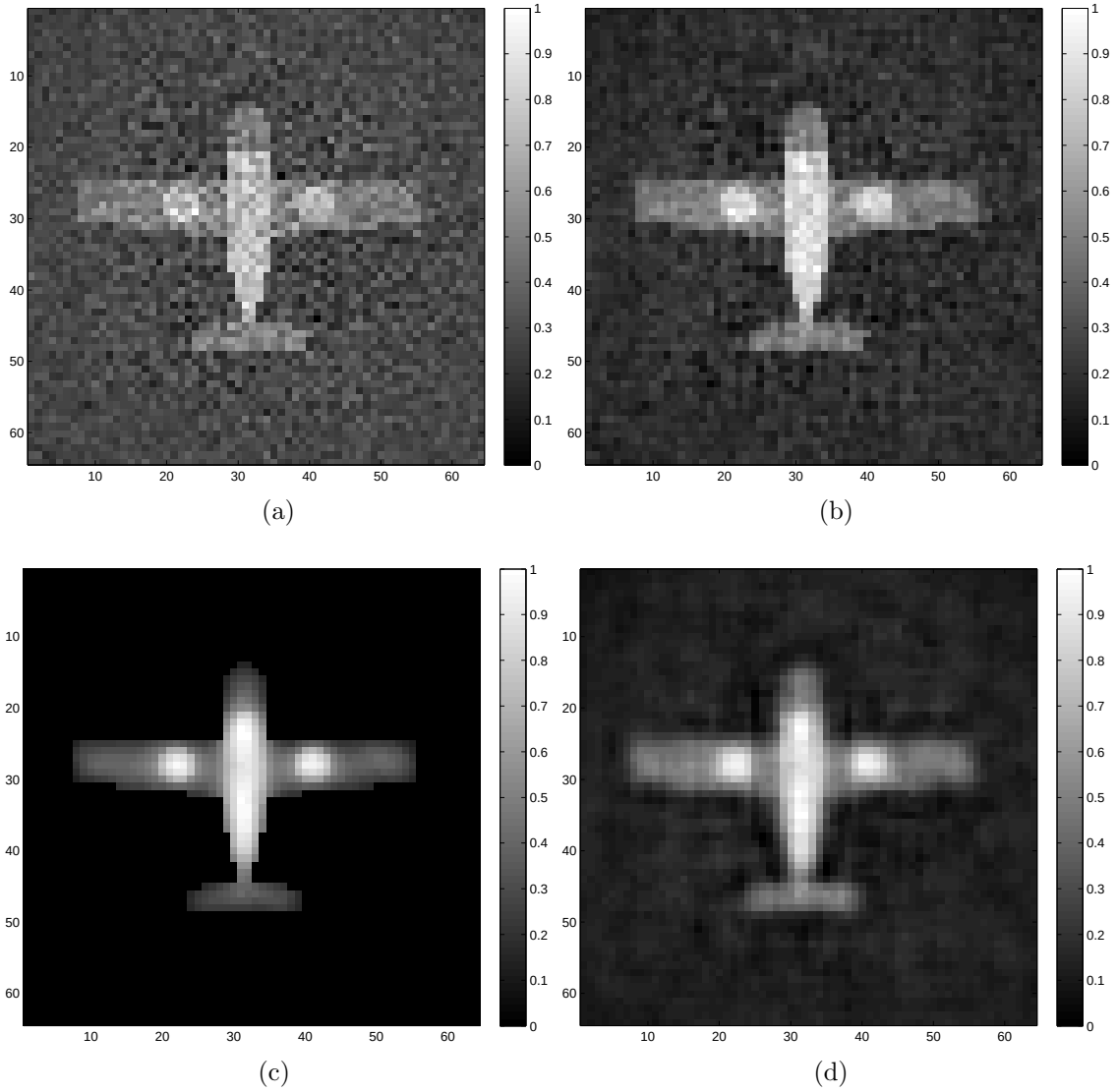


Figure 3.9: Reconstructed images using (a) the method in [1], (b) with stationary target and clutter assumption, (c) non-stationary target and clutter assumption with known SVSD, (d) non-stationary target and clutter assumption with known SVSD when SCR is 40dB and SNR is 10dB.

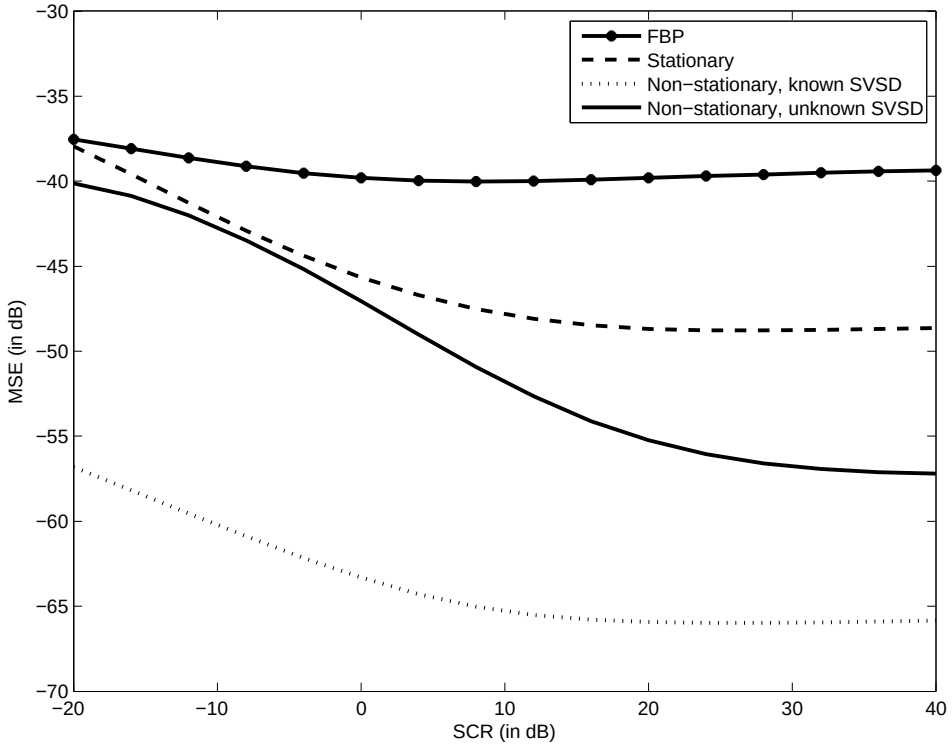
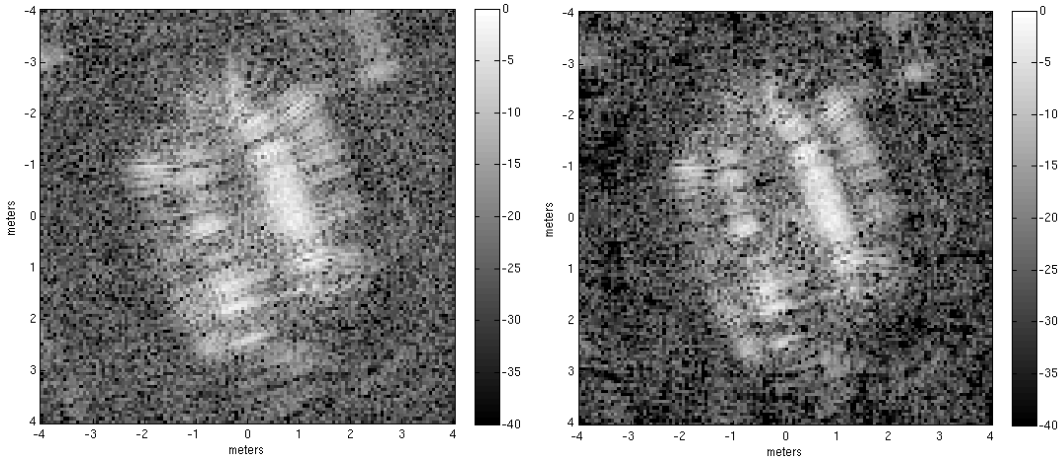


Figure 3.10: MSE (vertical axis, in log scale) versus SCR (horizontal axis) averaged over ten reconstructed images for each SCR level using four different image reconstruction methods.



(a) (b)

Figure 3.11: (a) Real SAR image reconstructed from wide-angle SAR data [2] by Method 1, (b) and Method 4.

CHAPTER 4
AN ANALYTIC SAR INVERSION METHOD BASED ON BEST LINEAR UNBIASED ESTIMATION

In this chapter, we develop an analytic SAR inversion formula based on best linear unbiased estimation criterion. For a brief review of the best linear unbiased estimation method, see Appendix F.

We assume that the clutter is a pseudo-stationary process and additive noise is stationary in fast-time and uncorrelated in slow-time. We formulate the image formation as a constrained optimization problem. We present an approximate solution for the optimization problem using microlocal analysis, the methods of Lagrange multipliers and variational derivatives. The resulting estimator is a composition of two linear operators, one of which is a pseudo-differential operator and the other one is an FIO in the form of a filtered-backprojection operator. The filters associated with these two operators are given in terms of the noise power spectral density function, spatially-varying clutter spectral density function, as well as the Beylkin determinant, antenna beam patterns, transmitted waveform and geometric spreading factors. The inversion formula has the advantages of computational efficiency, applicability to arbitrary imaging geometries and multiple SAR modalities.

4.1 Received Signal Model

4.1.1 Forward Model

Under the start-stop approximation, for many SAR modalities [1], [3], [16], noise-free scattered field data, $\tilde{d}$, can be modeled as follows:

$$\tilde{d}(s,t) \approx \mathcal{F}[V](s,t) := \int \mathrm{e}^{-\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{x})/c)} A(\boldsymbol{x},s,\omega) V(\boldsymbol{x})\, \mathrm{d}\omega \mathrm{d}\boldsymbol{x}, \tag{4.1}$$

where s and t denote slow-time and fast-time, respectively, and V denotes the scene reflectivity function. R denotes the range between the antenna(s) and the scatterer

located at $\boldsymbol{x}$ at slow-time s. For a bi-static configuration, R is given by

$$R(s,\boldsymbol{x}) = |\boldsymbol{\gamma}_T(s) - \mathbf{x}| + |\boldsymbol{\gamma}_R(s) - \mathbf{x}| \tag{4.2}$$

where $\boldsymbol{\gamma}_T(s)$ and $\boldsymbol{\gamma}_R(s)$ denote the transmitter and the receiver trajectories, respectively.

We assume that for some m_A, the amplitude term A satisfies the symbol estimate [1], [16]:

$$\sup_{(s,\boldsymbol{x})\in K} \mid \partial_\omega^\alpha \partial_s^\beta \partial_{\boldsymbol{x}}^{\boldsymbol{\rho}} A(\boldsymbol{x},s,\omega) \mid \leq B_{K,\alpha,\beta,\boldsymbol{\rho}}\ (1+|\omega|)^{m_A-|\alpha|} \tag{4.3}$$

where K is any compact subset of $\mathbb{R}\times\mathbb{R}^2$, $B_{K,\alpha,\beta,\boldsymbol{\rho}}$ is a $K,\alpha,\beta,\boldsymbol{\rho}$ dependent constant, and $\boldsymbol{\rho}$ is a multi-index. Under the assumption in (4.3), $\mathcal{F}$ becomes a Fourier integral operator [43]–[46]. For the rest of this chapter, we refer to $\mathcal{F}$ as the forward operator.

Let $T(\boldsymbol{x})$, $\boldsymbol{x}\in\Omega_T\subseteq\mathbb{R}^2$ denote the target scene. We assume that the target is embedded in clutter and write

$$V(\boldsymbol{x}) = T(\boldsymbol{x}) + C(\boldsymbol{x}). \tag{4.4}$$

Taking into account the additive thermal noise we model the received signal as

$$d(s,t) = \mathcal{F}[V](s,t) + n(s,t) = \mathcal{F}[T+C](s,t) + n(s,t). \tag{4.5}$$

4.1.2 Target, Clutter and Noise Models

Unlike the target models in Chapters 3 and 6, in this chapter, we model target as a deterministic quantity.

We model clutter, $C(\boldsymbol{x})$, $\boldsymbol{x}\in\Omega_C\subseteq\mathbb{R}^2$, as a zero-mean, pseudo-stationary process with the following spectral representation:

$$C(\boldsymbol{x}) = \int \mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\zeta}} S_C(\boldsymbol{x},\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}\mathrm{dB}(\boldsymbol{x}') \tag{4.6}$$

where $\mathrm{B}(\boldsymbol{x}')$ is the Wiener process and $S_C(\boldsymbol{x},\boldsymbol{\zeta})$ satisfies the following symbol esti-

mate

$$\sup_{\boldsymbol{x}\in U} \mid \partial_{\boldsymbol{\zeta}}^{\boldsymbol{\alpha}}\partial_{\boldsymbol{x}}^{\boldsymbol{\beta}} S_C(\boldsymbol{x},\boldsymbol{\zeta}) \mid \leq B^C_{U,\boldsymbol{\alpha},\boldsymbol{\beta}}(1+|\boldsymbol{\zeta}|)^{m_C-|\boldsymbol{\alpha}|} \tag{4.7}$$

where U is any compact subset of $\mathbb{R}^2$, $B^T_{U,\alpha,\beta}$ is a constant that depends on the multi-indices $\boldsymbol{\alpha},\boldsymbol{\beta}$, and U, m_C are real numbers.

The autocovariance function of C is given by

$$\mathrm{E}\left[C(\boldsymbol{x})\overline{C(\tilde{\boldsymbol{x}})}\right] = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\tilde{\boldsymbol{x}})\cdot\boldsymbol{\zeta}} S_C(\boldsymbol{x},\boldsymbol{\zeta})\overline{S_C(\tilde{\boldsymbol{x}},\boldsymbol{\zeta})}\,\mathrm{d}\boldsymbol{\zeta}. \tag{4.8}$$

We assume that the additive measurement noise is a zero-mean, stationary process in fast-time and statistically mutually uncorrelated in slow-time and write

$$S_n(s,\omega) = \sigma(s)\tilde{S}_n(\omega). \tag{4.9}$$

The autocovariance function of the noise is given by

$$\mathrm{E}[n(s,t)\overline{n(s',t')}] = \sigma^2(s)\delta(s-s')\mathcal{R}_n(t,t') \tag{4.10}$$

where

$$\mathcal{R}_n(t,t') = \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-t')}|\tilde{S}_n(\omega)|^2\,\mathrm{d}\omega \tag{4.11}$$

and $|S_n(s,\omega)|^2$ is the power spectral density function of the noise. To avoid peculiar behavior, we assume $\tilde{S}_n$ satisfies the following condition

$$\mathrm{E}[|n(s,t)|^2] = \sigma^2(s)\int \left|\tilde{S}_n(\omega)\right|^2 \mathrm{d}\omega < +\infty. \tag{4.12}$$

Finally, without loss of generality, we assume that the clutter and noise are mutually statistically uncorrelated.

4.2 Image Formation based on BLUE Criterion

4.2.1 Problem Statement

Let $\mathcal{K}_{bl}$ be a linear operator and

$$T_{bl}(\boldsymbol{z}) = \mathcal{K}_{bl}[d](\boldsymbol{z}) \tag{4.13}$$

be an image of the target. We wish to design $\mathcal{K}_{bl}$ so that the following conditions are satisfied:

1. $\mathcal{K}_{bl}$ is a composition of two FIOs, i.e., $\mathcal{K}_{bl} = \mathcal{K}^2_{bl}\mathcal{K}^1_{bl}$. Specifically, $\mathcal{K}^1_{bl}$ is the following filtered-backprojection (FBP) operator:

$$T(\boldsymbol{x}') = \mathcal{K}^1_{bl}[d](\boldsymbol{x}') := \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{x}')/c)} Q^1_{bl}(\boldsymbol{x}', s, \omega) d(s,t)\, \mathrm{d}\omega \mathrm{d}s\mathrm{d}t \tag{4.14}$$

   where Q^1_{bl} satisfies a symbol estimate similar to the one stated in (4.3).

   $\mathcal{K}^2_{bl}$ is a pseudo-differential operator with the following kernel:

$$Q^2_{bl}(\boldsymbol{z}, \boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\zeta}} \hat{Q}^2_{bl}(\boldsymbol{z}, \boldsymbol{\zeta}) \mathrm{d}\boldsymbol{\zeta} \tag{4.15}$$

   where $\hat{Q}^2_{bl}$ satisfies a symbol estimate similar to the one stated in (4.3). Note that Q^1_{bl} is a filter applied prior to the backprojection and Q^2_{bl} is a filter applied after the backprojection operation.

2. $T_{bl}(\boldsymbol{z}) = \mathcal{K}_{bl}[d](\boldsymbol{z})$ is an unbiased estimate of $T(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega_T \subseteq \mathbb{R}^2$ up to the leading-order term of $\mathcal{K}_{bl}$.[4]

3. The total variance of the estimate

$$\int \mathrm{E}\left[|T_{bl}(\boldsymbol{z}) - \mathrm{E}[T_{bl}(\boldsymbol{z})]|^2\right] \mathrm{d}\boldsymbol{z} \tag{4.16}$$

   is minimum among all the linear estimates satisfying Condition 1 and 2 above.

[4]Here the "leading-order" is in the microlocal sense [43]–[46], meaning that the higher order terms are smoother than the leading-order one. In other words, the bias of the estimate is one degree smoother than $\mathrm{E}[T_{bl}]$.

We refer to $T_{bl}(\boldsymbol{z}) = \mathcal{K}_{bl}[d](\boldsymbol{z})$ satisfying Conditions 1-3 as the best linear unbiased estimate (BLUE) of $T(\boldsymbol{z})$ and $\mathcal{K}_{bl}$ as the BLUE operator. Note that we require the estimator to be not only linear, unbiased and minimum variance, but also of a particular form as stated in Condition 1.

Clearly, determining the BLUE operator is equivalent to determining the filters Q_{bl}^1 and Q_{bl}^2 satisfying Conditions 1-3.

4.2.2 The Bias and Total Variance of BLUE

We begin by defining a bias image

$$B(\boldsymbol{z}) = \mathrm{E}[T_{bl}(\boldsymbol{z})] - \mathcal{I}_{\Omega}[T](\boldsymbol{z}) \tag{4.17}$$

where

$$\mathcal{I}_{\Omega}[T](\boldsymbol{z}) := \int_{\Omega_{\boldsymbol{z}}} \chi_{\Omega}(\boldsymbol{z}, \boldsymbol{\xi}) \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} T(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{x} \tag{4.18}$$

and $\chi_{\Omega}(\boldsymbol{z}, \boldsymbol{\xi})$ is a smoothed version of the characteristic function of Ω given by $\Omega = \cup_{\boldsymbol{z}} \Omega_{\boldsymbol{z}}$ with

$$\Omega_{\boldsymbol{z}} = \left\{ \boldsymbol{\xi} \ : \ \boldsymbol{\xi} = \frac{\omega}{c} \nabla_{\boldsymbol{x}} R(s, \boldsymbol{x}) \mid_{\boldsymbol{x}=\boldsymbol{z}}, \ A(\boldsymbol{z}, s, \omega) \neq 0 \right\}. \tag{4.19}$$

The image $\mathcal{I}_{\Omega}[T]$ is the best image (in the least square sense) we can obtain given the bandwidth of the transmitted waveform and the effective length of the synthetic aperture. Ω can be viewed as the set of target Fourier coefficients that can be obtained from our measurements.

Let

$$\mathcal{C}_{T_{bl}}(\boldsymbol{z}, \boldsymbol{z}') := \mathrm{E}\left[(T_{bl}(\boldsymbol{z}) - \mathrm{E}[T_{bl}(\boldsymbol{z})]) \overline{(T_{bl}(\boldsymbol{z}') - \mathrm{E}[T_{bl}(\boldsymbol{z}')])} \right] \tag{4.20}$$

denote the autocorrelation function of $T_{bl}(\boldsymbol{z})$. We next define

$$\mathcal{J}_{bl}^{C}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) = \int \mathrm{E}[|\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z})|^{2}]\mathrm{d}\boldsymbol{z}, \tag{4.21}$$

$$\mathcal{J}_{bl}^{n}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) = \int \mathrm{E}[|\mathcal{K}_{bl}[n](\boldsymbol{z})|^{2}]\mathrm{d}\boldsymbol{z}. \tag{4.22}$$

Lemma 2: We assume that the image T_{bl} is formed by the BLUE operator as outlined in Conditions 1-3, where the filters Q_{bl}^{1} and $\hat{Q}_{bl}^{2}$ satisfy the symbol estimates similar to the one stated in (4.7) and where the data d are given by (4.5). We assume that the geometrical conditions on the flight trajectories and the antenna beam patterns satisfy certain conditions such that artifacts are avoided [82].

(i) Then, the leading-order singularities of the bias image are

$$B(\boldsymbol{z}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} \left(\hat{Q}_{bl}^{2}(\boldsymbol{z},\boldsymbol{\xi})Q_{bl}^{1}(\boldsymbol{z},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi}) - \chi_{\Omega(\boldsymbol{z},\boldsymbol{\xi})}\right) \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{4.23}$$

(ii) The total variance of the BLUE image is given by

$$\int \mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z})\mathrm{d}\boldsymbol{z} \approx \mathcal{J}_{bl}^{C}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) + \mathcal{J}_{bl}^{C}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) \tag{4.24}$$

where the leading-order singularities of each term in (4.24) are

$$\mathcal{J}_{bl}^{C}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) \approx \int |\hat{Q}_{bl}^{2}(\boldsymbol{x},\boldsymbol{\xi})Q_{bl}^{1}(\boldsymbol{x},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})|^{2}|S_{C}(\boldsymbol{x},\boldsymbol{\xi})|^{2}\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{4.25}$$

and

$$\mathcal{J}_{bl}^{n}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) \approx \int |\hat{Q}_{bl}^{2}(\boldsymbol{x},\boldsymbol{\xi})Q_{bl}^{1}(\boldsymbol{x},\boldsymbol{\xi})|^{2}\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})|S_{n}(\boldsymbol{\xi})|^{2}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}. \tag{4.26}$$

Proof:

(i) Let

$$T^1_{bl}(\boldsymbol{z}) := \mathcal{K}^1_{bl}\mathcal{F}[T](\boldsymbol{z}). \tag{4.27}$$

We simplify $\mathcal{K}^1_{bl}\mathcal{F}$ as in [1]. Using (4.1) in (4.14) and carrying out the t integration and one of the ω integrations results in

$$\mathcal{K}^1_{bl}\mathcal{F}[T](\boldsymbol{x}') := \int \mathrm{e}^{\mathrm{i}2\pi\frac{2\omega}{c}(R(s,\boldsymbol{x})-R(s,\boldsymbol{x}'))} Q^1_{bl}(\boldsymbol{x}', s, \omega) A(\boldsymbol{x}, s, \omega) T(\boldsymbol{x}) \mathrm{d}\omega \mathrm{d}s \mathrm{d}\boldsymbol{x} \tag{4.28}$$

Next, we use the identity

$$h(\boldsymbol{x}) - h(\boldsymbol{x}') = (\boldsymbol{x} - \boldsymbol{x}') \cdot \int_0^1 \nabla h(\boldsymbol{x}' + \lambda(\boldsymbol{x} - \boldsymbol{x}')) \mathrm{d}\lambda. \tag{4.29}$$

For $h(\boldsymbol{x}) = \frac{2\omega}{c} R(s, \boldsymbol{x})$ we write

$$\frac{2\omega}{c}\left(R(s,\boldsymbol{x}) - R(s,\boldsymbol{x}')\right) = \frac{\omega}{c}(\boldsymbol{x} - \boldsymbol{x}') \cdot \boldsymbol{\Xi}(s, \boldsymbol{x}, \boldsymbol{x}'), \tag{4.30}$$

where for $\boldsymbol{x} = \boldsymbol{x}'$ we can write

$$\boldsymbol{\Xi}(s, \boldsymbol{x}', \boldsymbol{x}') = \nabla h(\boldsymbol{x}') = \mathbf{L}(\boldsymbol{x}', s) \cdot D\psi(\boldsymbol{x}'), \tag{4.31}$$

and $D\psi(\boldsymbol{x}')$ is given by

$$D\psi(\boldsymbol{x}') = \begin{bmatrix} 1 & 0 & \nabla_{x'_1}\psi(\boldsymbol{x}') \\ 0 & 1 & \nabla_{x'_2}\psi(\boldsymbol{x}') \end{bmatrix} \tag{4.32}$$

and

$$\mathbf{L}(\boldsymbol{x}', s) = \frac{\boldsymbol{\gamma}_T(s) - \mathbf{x}'}{|\boldsymbol{\gamma}_T(s) - \mathbf{x}'|} + \frac{\boldsymbol{\gamma}_R(s) - \mathbf{x}'}{|\boldsymbol{\gamma}_R(s) - \mathbf{x}'|} \tag{4.33}$$

is the bi-static antenna look direction. We make the change of variables

$$(s,\omega) \to \boldsymbol{\xi} = \frac{2\omega}{c}\boldsymbol{\Xi}(s,\boldsymbol{x},\boldsymbol{x}') \tag{4.34}$$

where the geometrical assumptions made for the trajectory makes the change of variables non-singular [82]. Then, (4.28) can be written as

$$\mathcal{K}^1_{bl}\mathcal{F}[T](\boldsymbol{x}') \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} Q^1_{bl}(\boldsymbol{x}',\boldsymbol{\xi}) A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{\xi})T(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi} \tag{4.35}$$

where

$$\eta(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{\xi}) = \left|\frac{\partial(s,\omega)}{\partial\boldsymbol{\xi}}\right| \tag{4.36}$$

is the determinant of the Jacobian coming from the change of variables.

Combining (4.35) with $\mathcal{K}^2_{bl}$ we obtain

$$\begin{aligned}\mathcal{K}_{bl}\mathcal{F}[T](\boldsymbol{z}) = \int \hat{Q}^2_{bl}(\boldsymbol{z},\boldsymbol{\zeta})\mathrm{e}^{\mathrm{i}2\pi[(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}+(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\zeta}]} Q^1_{bl}(\boldsymbol{x}',\boldsymbol{\xi}) \\ \times A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{\xi})T(\boldsymbol{x})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'\mathrm{d}\boldsymbol{\zeta}.\end{aligned} \tag{4.37}$$

We next apply the method of the stationary phase in the variables $\boldsymbol{x}'$ and $\boldsymbol{\zeta}$ simultaneously for the phase function:[5]

$$\phi_1(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{z},\boldsymbol{\zeta},\boldsymbol{\xi}) = (\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\zeta} + (\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}. \tag{4.38}$$

The method of the stationary phase results in the critical points at

$$\nabla_{\boldsymbol{x}'}\phi_1 \to \boldsymbol{\zeta} = \boldsymbol{\xi}, \tag{4.39}$$

$$\nabla_{\boldsymbol{\zeta}}\phi_1 \to \boldsymbol{x}' = \boldsymbol{z}. \tag{4.40}$$

[5] The determinant of the Hessian of ϕ_1 is -1. See Appendix G.1 for the computation of the Hessian of ϕ_1.

The leading-order term of (4.37) is then

$$\mathcal{K}_{bl}\mathcal{F}[T](\boldsymbol{z}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}}\hat{Q}^2_{bl}(\boldsymbol{z},\boldsymbol{\xi})Q^1_{bl}(\boldsymbol{z},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi})T(\boldsymbol{x})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{4.41}$$

Combining (4.41) and (4.18), the leading-order singularities of the bias image is given by

$$\begin{aligned} B(\boldsymbol{z}) =& \mathrm{E}[T_{bl}(\boldsymbol{z})] - \mathcal{I}_\Omega[T](\boldsymbol{z}) && (4.42)\\ \approx& \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}}\left(\hat{Q}^2_{bl}(\boldsymbol{z},\boldsymbol{\xi})Q^1_{bl}(\boldsymbol{z},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi}) - \chi_\Omega(\boldsymbol{z},\boldsymbol{\xi})\right)\\ & \times T(\boldsymbol{x})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. && (4.43)\end{aligned}$$

(ii) The total variance of the BLUE image is given by:

$$\begin{aligned} \int \mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z})\mathrm{d}\boldsymbol{z} &= \int \mathrm{E}\left[|(T_{bl}(\boldsymbol{z}) - \mathrm{E}[T_{bl}(\boldsymbol{z})|^2])\right]\mathrm{d}\boldsymbol{z}, && (4.44)\\ &= \int \mathrm{E}\left[|\mathcal{K}_{bl}\mathcal{F}[T+C](\boldsymbol{z}) + \mathcal{K}_{bl}[n](\boldsymbol{z}) - T(\boldsymbol{z})|^2\right]\mathrm{d}\boldsymbol{z}. && (4.45)\end{aligned}$$

Using the facts that T_{bl} is unbiased up to its leading-order singularities, the noise and clutter are statistically uncorrelated, and the approximation in (4.37), we get

$$\int \mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z})\mathrm{d}\boldsymbol{z} \approx \int \left(\mathrm{E}[|\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z})|^2] + \mathrm{E}[|\mathcal{K}_{bl}[n](\boldsymbol{z})|^2]\right)\mathrm{d}\boldsymbol{z}. \tag{4.46}$$

Using the approximation to $\mathcal{K}_{bl}$ given in (4.41), $\mathcal{J}^C_{bl}$ can be expressed as

$$\begin{aligned} \mathcal{J}^C_{bl}(Q^1_{bl},\hat{Q}^2_{bl}) &= \int \mathrm{E}\left[(\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z}))\overline{(\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z}))}\right]\mathrm{d}\boldsymbol{z},\\ &\approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}}\hat{Q}^2_{bl}(\boldsymbol{z},\boldsymbol{\xi})Q^1_{bl}(\boldsymbol{z},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi})\\ &\quad\times \mathrm{e}^{-\mathrm{i}2\pi(\tilde{\boldsymbol{x}}-\boldsymbol{z})\cdot\tilde{\boldsymbol{\xi}}}\overline{\hat{Q}^2_{bl}(\boldsymbol{z},\tilde{\boldsymbol{\xi}})Q^1_{bl}(\boldsymbol{z},\tilde{\boldsymbol{\xi}})A(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{\xi}})\eta(\tilde{\boldsymbol{x}},\boldsymbol{z},\tilde{\boldsymbol{\xi}})}\\ &\quad\times \mathrm{E}[C(\boldsymbol{x})\overline{C(\tilde{\boldsymbol{x}})}]\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\tilde{\boldsymbol{x}}\mathrm{d}\tilde{\boldsymbol{\xi}}\mathrm{d}\boldsymbol{z}. && (4.47)\end{aligned}$$

The phase function of the integrand in (4.47) is given by

$$\phi_2(\boldsymbol{x},\boldsymbol{\xi},\tilde{\boldsymbol{x}},\tilde{\boldsymbol{\xi}},\boldsymbol{z}) = (\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi} - (\tilde{\boldsymbol{x}}-\boldsymbol{z})\cdot\tilde{\boldsymbol{\xi}}. \tag{4.48}$$

to approximate $\boldsymbol{z}$ and $\tilde{\boldsymbol{\xi}}$ integrations, we apply the method of the stationary phase in the variables $\boldsymbol{z}$ and $\tilde{\boldsymbol{\xi}}$, simultaneously.[6] We obtain the critical points at

$$\nabla_{\boldsymbol{z}}\phi_2 \to \tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}, \tag{4.49}$$

$$\nabla_{\tilde{\boldsymbol{\xi}}}\phi_2 \to \boldsymbol{z} = \tilde{\boldsymbol{x}}. \tag{4.50}$$

Hence, the leading-order term of (4.47) is

$$\begin{aligned}\mathcal{J}_{bl}^{C}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\tilde{\boldsymbol{x}})\cdot\boldsymbol{\xi}}|\hat{Q}_{bl}^{2}(\tilde{\boldsymbol{x}},\boldsymbol{\xi})Q_{bl}^{1}(\tilde{\boldsymbol{x}},\boldsymbol{\xi})|^{2}A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\tilde{\boldsymbol{x}},\boldsymbol{\xi}) \\ \times \overline{A(\tilde{\boldsymbol{x}},\boldsymbol{\xi})\eta(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}},\boldsymbol{\xi})}\mathrm{E}[C(\boldsymbol{x})\overline{C(\hat{\boldsymbol{x}})}]\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\tilde{\boldsymbol{x}}.\end{aligned} \tag{4.51}$$

Inserting (4.8) in (4.51), we obtain

$$\begin{aligned}\mathcal{J}_{bl}^{C}(Q_{bl}^{1},\hat{Q}_{bl}^{2}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\tilde{\boldsymbol{x}})\cdot\boldsymbol{\xi}}|\hat{Q}_{bl}^{2}(\tilde{\boldsymbol{x}},\boldsymbol{\xi})Q_{bl}^{1}(\tilde{\boldsymbol{x}},\boldsymbol{\xi})|^{2}A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\tilde{\boldsymbol{x}},\boldsymbol{\xi}) \\ \times \overline{A(\tilde{\boldsymbol{x}},\boldsymbol{\xi})\eta(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}},\boldsymbol{\xi})}\mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{x}-\tilde{\boldsymbol{x}})\cdot\boldsymbol{\zeta}}S_C(\boldsymbol{x},\boldsymbol{\zeta})\overline{S_C(\tilde{\boldsymbol{x}},\boldsymbol{\zeta})}\,\mathrm{d}\boldsymbol{\zeta}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\tilde{\boldsymbol{x}}.\end{aligned} \tag{4.52}$$

We apply the method of the stationary phase to the phase function one more time in variables $\tilde{\boldsymbol{x}}$ and $\boldsymbol{\zeta}$, simultaneously:[7]

$$\phi_3(\boldsymbol{x},\tilde{\boldsymbol{x}},\boldsymbol{\xi},\boldsymbol{\zeta}) = (\boldsymbol{x}-\tilde{\boldsymbol{x}})\cdot\boldsymbol{\xi} - (\boldsymbol{x}-\tilde{\boldsymbol{x}})\cdot\boldsymbol{\zeta} \tag{4.53}$$

[6]The determinant of the Hessian of ϕ_2 is -1. See Appendix G.2 for the computation of the Hessian of ϕ_2.

[7]The determinant of the Hessian of ϕ_3 is -1. See Appendix G.3 for the computation of the Hessian of ϕ_3.

which results in the critical conditions

$$\nabla_{\tilde{\boldsymbol{x}}}\phi_3 \to \boldsymbol{\zeta} = \boldsymbol{\xi}, \tag{4.54}$$

$$\nabla_{\boldsymbol{\zeta}}\phi_3 \to \tilde{\boldsymbol{x}} = \boldsymbol{x}. \tag{4.55}$$

Hence, the leading-order term of (4.52) is

$$\mathcal{J}_{bl}^{C}(Q_{bl}^1, \hat{Q}_{bl}^2) \approx \int |\hat{Q}_{bl}^2(\boldsymbol{x}, \boldsymbol{\xi}) Q_{bl}^1(\boldsymbol{x}, \boldsymbol{\xi}) A(\boldsymbol{x}, \boldsymbol{\xi}) \eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi})|^2 |S_C(\boldsymbol{x}, \boldsymbol{\xi})|^2 \, \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{4.56}$$

Following steps similar to those in (4.47)-(4.52) and using (4.11), it can be shown that the leading-order term of $\mathcal{J}_{bl}^{n}$ is

$$\mathcal{J}_{bl}^{n}(Q_{bl}^1, \hat{Q}_{bl}^2) \approx \int |\hat{Q}_{bl}^2(\boldsymbol{x}, \boldsymbol{\xi}) Q_{bl}^1(\boldsymbol{x}, \boldsymbol{\xi})|^2 \eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi}) |S_n(\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}. \tag{4.57}$$

4.2.3 Derivation of the Imaging Filters

The leading-order singularities of the bias image is given by

$$B(\boldsymbol{z}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}} \left(\hat{Q}_{bl}^2(\boldsymbol{z}, \boldsymbol{\xi}) Q_{bl}^1(\boldsymbol{z}, \boldsymbol{\xi}) A(\boldsymbol{x}, \boldsymbol{\xi}) \eta(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\xi}) - \chi_{\Omega}(\boldsymbol{z}, \boldsymbol{\xi}) \right) T(\boldsymbol{x}) \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{4.58}$$

The leading-order contribution to (4.58) occurs when $\boldsymbol{x} = \boldsymbol{z}$. Let

$$\mathcal{B}(Q_{bl}^1, Q_{bl}^1) := \hat{Q}_{bl}^2(\boldsymbol{z}, \boldsymbol{\xi}) Q_{bl}^1(\boldsymbol{z}, \boldsymbol{\xi}) A(\boldsymbol{z}, \boldsymbol{\xi}) \eta(\boldsymbol{z}, \boldsymbol{z}, \boldsymbol{\xi}) - \chi_{\Omega}(\boldsymbol{z}, \boldsymbol{\xi}). \tag{4.59}$$

For the leading-order singularities of T_{bl} to be unbiased, the following must hold

$$\mathcal{B}(Q_{bl}^1, \hat{Q}_{bl}^2) = 0 \tag{4.60}$$

for all $\boldsymbol{z} \in \Omega_T$. Our objective is to design the filters Q_{bl}^1 and $\hat{Q}_{bl}^2$ so that the leading-

order singularities of the total variance

$$\int \mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z})\mathrm{d}\boldsymbol{z} \approx \int \left(\mathrm{E}[|\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z})|^2] + \mathrm{E}[|\mathcal{K}_{bl}[n](\boldsymbol{z})|^2]\right)\mathrm{d}\boldsymbol{z} \tag{4.61}$$

is minimized subject to the constraint in (4.60).

We use the method of Lagrange multipliers to address this constrained optimization problem. Let $\lambda(\boldsymbol{z},\boldsymbol{\xi})$ be the Lagrange multiplier associated with the constraint $\mathcal{B}(Q_{bl}^1,\hat{Q}_{bl}^2)=0$. Recall that

$$\int \mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z})\mathrm{d}\boldsymbol{z} \approx \mathcal{J}_{bl}^C(Q_{bl}^1,\hat{Q}_{bl}^2) + \mathcal{J}_{bl}^n(Q_{bl}^1,\hat{Q}_{bl}^2) \tag{4.62}$$

where

$$\mathcal{J}_{bl}^C(Q_{bl}^1,\hat{Q}_{bl}^2) = \int \mathrm{E}[|\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z})|^2]\mathrm{d}\boldsymbol{z}, \tag{4.63}$$

$$\mathcal{J}_{bl}^n(Q_{bl}^1,\hat{Q}_{bl}^2) = \int \mathrm{E}[|\mathcal{K}_{bl}[n](\boldsymbol{z})|^2]\mathrm{d}\boldsymbol{z}. \tag{4.64}$$

To incorporate the constraint (4.60) into the minimization, we define

$$\mathcal{J}_{bl}^{\lambda}(Q_{bl}^1,\hat{Q}_{bl}^2,\lambda) := \int \lambda(\boldsymbol{z},\boldsymbol{\xi})\mathcal{B}(Q_{bl}^1,\hat{Q}_{bl}^2)\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{\xi}. \tag{4.65}$$

We next define the following objective functional:

$$\mathcal{J}_{bl}(Q_{bl}^1,\hat{Q}_{bl}^2,\lambda) = \mathcal{J}_{bl}^C(Q_{bl}^1,\hat{Q}_{bl}^2) + \mathcal{J}_{bl}^n(Q_{bl}^1,\hat{Q}_{bl}^2) - \mathcal{J}_{bl}^{\lambda}(Q_{bl}^1,\hat{Q}_{bl}^2,\lambda). \tag{4.66}$$

Theorem 2: Let the data d be given by (4.5) and T_{bl} be given by (4.13). If the filters Q_{bl}^1 and Q_{bl}^2 in (4.14) and (4.15) are chosen as

$$Q_{bl}^1(\boldsymbol{x}',\boldsymbol{\xi}) = \frac{\overline{A(\boldsymbol{x}',\boldsymbol{\xi})}\eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi})\chi_{\Omega}(\boldsymbol{x}',\boldsymbol{\xi})}{|A(\boldsymbol{x}',\boldsymbol{\xi})\eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi})|^2|S_C(\boldsymbol{x}',\boldsymbol{\xi})|^2 + \eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2} \tag{4.67}$$

and

$$Q_{bl}^2(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}}\frac{|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2 + \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2}\mathrm{d}\boldsymbol{\xi}. \tag{4.68}$$

where $\chi_\Omega(\boldsymbol{x}', \boldsymbol{\xi})$ is a smooth function that prevents division by 0 and $\eta(\boldsymbol{x}', \boldsymbol{x}', \boldsymbol{\xi})$, $S_C(\boldsymbol{z}, \boldsymbol{\xi})$ and $S_n(\boldsymbol{\xi})$ are as defined in (4.36), (4.6), (4.9), respectively. Then, the followings hold:

(i) The leading-order singularities of $\mathrm{E}[T_{bl}(\boldsymbol{z})]$ is equal to the leading order singularities of $T(\boldsymbol{z})$.[8]

(ii) $\mathcal{C}_{T_{bl}}(\boldsymbol{z}, \boldsymbol{z})$ is minimized at its leading-order singularities.

(iii) With these choice of filters, the variance of T_{bl} at its leading-order singularities is given by

$$\mathcal{C}_{T_{bl}}(\boldsymbol{z}, \boldsymbol{z}) \approx \int \frac{|A(\boldsymbol{z}, \boldsymbol{\xi})\eta(\boldsymbol{z}, \boldsymbol{z}, \boldsymbol{\xi})|^2 |S_C(\boldsymbol{z}, \boldsymbol{\xi})|^2 + \eta(\boldsymbol{z}, \boldsymbol{z}, \boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{z}, \boldsymbol{\xi})\eta(\boldsymbol{z}, \boldsymbol{z}, \boldsymbol{\xi})|^2} \mathrm{d}\boldsymbol{\xi}. \quad (4.69)$$

Proof: Our objective is to determine Q^1_{bl}, $\hat{Q}^2_{bl}$, λ that minimizes the following objective functional:

$$\mathcal{J}_{bl}(Q^1_{bl}, \hat{Q}^2_{bl}, \lambda) = \mathcal{J}^C_{bl}(Q^1_{bl}, \hat{Q}^2_{bl}) + \mathcal{J}^n_{bl}(Q^1_{bl}, \hat{Q}^2_{bl}) - \mathcal{J}^\lambda_{bl}(Q^1_{bl}, \hat{Q}^2_{bl}, \lambda). \quad (4.70)$$

(i)-(ii) Having obtained approximations to each functional in (4.70), we study the variational derivative of (4.70) with respect to $\hat{Q}^2_{bl}$, Q^1_{bl} and λ. We evaluate the variational derivative with respect to $\hat{Q}^2_{bl}$ as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}^C_{bl}(\hat{Q}^2_{bl} + \epsilon Q_\epsilon)\Big|_{\epsilon=0} + \frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}^n_{bl}(\hat{Q}^2_{bl} + \epsilon Q_\epsilon)\Big|_{\epsilon=0} - \frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}^\lambda_{bl}(\hat{Q}^2_{bl} + \epsilon Q_\epsilon)\Big|_{\epsilon=0} = 0 \quad (4.71)$$

which is given by

$$\begin{aligned} 0 \;= 2\ \mathrm{Re} \int Q_\epsilon(\boldsymbol{x}, \boldsymbol{\xi}) \Big\{ & \overline{\hat{Q}^2_{bl}(\boldsymbol{x}, \boldsymbol{\xi})} |Q^1_{bl}(\boldsymbol{x}, \boldsymbol{\xi})|^2 \\ & \times \left(|A(\boldsymbol{x}, \boldsymbol{\xi})\eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi})|^2 |S_C(\boldsymbol{x}, \boldsymbol{\xi})|^2 + \eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2 \right) \Big\} \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \\ & - \int \lambda(\boldsymbol{x}, \boldsymbol{\xi}) Q_\epsilon(\boldsymbol{x}, \boldsymbol{\xi}) Q^1_{bl}(\boldsymbol{x}, \boldsymbol{\xi}) A(\boldsymbol{x}, \boldsymbol{\xi}) \eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \end{aligned} \quad (4.72)$$

[8] The bias image $B(\boldsymbol{z})$ is one degree smoother than the $\mathrm{E}[T_{bl}(\boldsymbol{z})]$.

We rewrite (4.72) as follows:

$$\begin{aligned} 0 \;&= 2\ \mathrm{Re} \int Q_\epsilon \left(\overline{\hat{Q}_{bl}^2} |Q_{bl}^1|^2 (|A\eta|^2 |S_C|^2 + \eta |S_n|^2) - \frac{1}{2} \lambda Q_{bl}^1 A \eta \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{\xi} \\ &\quad + \mathrm{Im} \int Q_\epsilon \lambda Q_{bl}^1 A \eta \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{\xi}. \end{aligned} \tag{4.73}$$

(4.73) must hold for all Q_ϵ including real-valued Q_ϵ. Hence, we obtain the following two relations on the filter Q_{bl}^1:

$$\overline{\hat{Q}_{bl}^2} |Q_{bl}^1|^2 (|A\eta|^2 |S_C|^2 + \eta |S_n|^2) - \frac{1}{2} \lambda Q_{bl}^1 A \eta \;=\; 0 \tag{4.74}$$

and

$$\mathrm{Im}\ \lambda Q_{bl}^1 A \eta \;=\; 0. \tag{4.75}$$

The variational derivative with respect to Q_{bl}^1 provides an equation equivalent to (4.73). Therefore, we only use the variational derivative with respect to $\hat{Q}_{bl}^2$ and λ to determine the BLUE filters Q_{bl}^1 and $\hat{Q}_{bl}^2$. Finally, we evaluate the variational derivative of $\mathcal{J}_{bl}$ or $\mathcal{J}_{bl}^\lambda$ with respect to λ and set it equal to 0 as follows:

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\epsilon} \mathcal{J}_{bl}^\lambda (Q_{bl}^1, \hat{Q}_{bl}^2, \lambda + \epsilon \lambda_\epsilon) \Big|_{\epsilon=0} \\ &\quad \approx \int \lambda_\epsilon(\boldsymbol{x}, \boldsymbol{\xi}) \left(\hat{Q}_{bl}^2(\boldsymbol{x}, \boldsymbol{\xi}) Q_{bl}^1(\boldsymbol{x}, \boldsymbol{\xi}) A(\boldsymbol{x}, \boldsymbol{\xi}) \eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi}) - \chi_\Omega(\boldsymbol{x}, \boldsymbol{\xi}) \right) \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{x} = 0. \end{aligned} \tag{4.76}$$

Since (4.76) must be true for all λ_ϵ, we obtain

$$\hat{Q}_{bl}^2(\boldsymbol{x}, \boldsymbol{\xi}) Q_{bl}^1(\boldsymbol{x}, \boldsymbol{\xi}) A(\boldsymbol{x}, \boldsymbol{\xi}) \eta(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{\xi}) = \chi_\Omega(\boldsymbol{x}, \boldsymbol{\xi}) \tag{4.77}$$

where $\chi_\Omega(\boldsymbol{x}, \boldsymbol{\xi})$ is a smooth function that prevents division by 0. Using (4.77)

in (4.74), we obtain

$$\lambda(\boldsymbol{x},\boldsymbol{\xi}) = 2\frac{|A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})|^2|S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 + \eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})|^2}. \tag{4.78}$$

To satisfy the condition in (4.75), we choose the filter Q_{bl}^1

$$Q_{bl}^1(\boldsymbol{x}',\boldsymbol{\xi}) = \frac{\overline{A(\boldsymbol{x}',\boldsymbol{\xi})}\eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi})\chi_\Omega(\boldsymbol{x}',\boldsymbol{\xi})}{|A(\boldsymbol{x}',\boldsymbol{\xi})\eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi})|^2|S_C(\boldsymbol{x}',\boldsymbol{\xi})|^2 + \eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2}, \tag{4.79}$$

and from (4.77) we evaluate Q_{bl}^2

$$Q_{bl}^2(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}}\frac{|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2 + \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2}\mathrm{d}\boldsymbol{\xi}. \tag{4.80}$$

(iii) Recall from (4.61) that the variance of the estimator is given by

$$\mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z}) \approx \mathrm{E}[|\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z})|^2] + \mathrm{E}[|\mathcal{K}_{bl}[n](\boldsymbol{z})|^2]. \tag{4.81}$$

Inserting the approximations to $\mathrm{E}[|\mathcal{K}_{bl}\mathcal{F}[C](\boldsymbol{z})|^2]$ and $\mathrm{E}[|\mathcal{K}_{bl}[n](\boldsymbol{z})|^2]$, we get

$$\begin{aligned}\mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z}) \approx \int &\left(|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2 + \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2\right)\\ &\times |\hat{Q}_{bl}^2(\boldsymbol{z},\boldsymbol{\xi})Q_{bl}^1(\boldsymbol{z},\boldsymbol{\xi})|^2\mathrm{d}\boldsymbol{\xi}.\end{aligned} \tag{4.82}$$

Inserting the expression for $Q_{bl}^1\hat{Q}_{bl}^2$ in (4.77) into (4.82), we obtain the minimum variance

$$\mathcal{C}_{T_{bl}}(\boldsymbol{z},\boldsymbol{z}) \approx \int \frac{|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2 + \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{z},\boldsymbol{\xi})\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|^2}\mathrm{d}\boldsymbol{\xi}. \tag{4.83}$$

4.3 Numerical Simulations

We consider a scene with an airplane-like object in the center as the target. The object has various intensity levels to show that the algorithm reconstructs the target scene with correct intensities at correct locations.

A scene of $100\,\mathrm{m} \times 100\,\mathrm{m}$ is discretized into 64×64 pixels. A bi-static antenna system with $\pi/18$ separation between the transmitter and receiver, flies over the scene on a circular trajectory $\boldsymbol{\gamma}(s) = [11\cos s, 11\sin s, 6.5]\,\mathrm{km}$ uniformly sampled for $s \in [0, 2\pi]$ at 128 points. It is assumed that $A(\boldsymbol{x}, \omega, s) = 1$, which corresponds to an isotropic antenna radiating a delta-like impulse and pre-compensation of geometric spreading factors in the FBP data.

Gaussian noise is added to the projection data to simulate the additive thermal noise. The clutter model described in Section 3.5 with the mask shown in Figure 3.2 is used here as well. The images are reconstructed via deterministic FBP, statistical FBP with stationary target and clutter models; and BLUE criterion at various SNR and SCR levels. For this set of simulations, the BLUE reconstruction method is applied under the stationary clutter assumption, hence we drop the $\boldsymbol{x}$ dependence on the spectral density function in filter (4.67). The target and clutter spectral densities are provided *a priori* to the reconstruct images under stationarity assumption. The reconstructed images are shown in Figures 4.1-4.10.

We performed ten sets of experiments. Figures 4.1-4.5 show the averages of the ten reconstructed images using (a) deterministic FBP, (b) statistical FBP under stationarity assumption, (c) BLUE after applying filter Q_{bl}^1, (d) BLUE after applying filter Q_{bl}^2. We kept SCR constant at 10dB and increased SNR from 0dB to 30dB. Similarly, Figures 4.6-4.10 show the average of the ten reconstructed images using the same reconstruction methods at before. The SNR is kept constant at 10dB and SCR is increased from 0dB to 32dB with 8 dB increments.

The plots in Figure 4.11(a) show the MSE in log scale versus SNR (horizontal axis) at a 10dB SNR and at 5 different SNR levels. The plots in Figure 4.11(b) show the MSE versus SCR (horizontal axis) at a constant SNR level of 10dB averaged over ten reconstructed images. It can be observed from these plots that the performance of the BLUE method is almost as good as the stationary reconstruction method

even though the stationary reconstruction method requires both target and clutter spectral density functions whereas BLUE method only requires the spectral density function of the clutter.

The performances of these methods are similar when the SNR is constant and SCR is changed since the filter (4.67) used in the reconstruction with BLUE includes the spectral density function of the clutter. As compared to the deterministic FBP reconstruction there is an average of 15% improvement in the MSE of the reconstructed images when the BLUE method is used. At high SNR levels, the deterministic FBP method works well. Therefore, the performances of the deterministic FBP and the BLUE are similar at 32dB SNR as shown in Figure 4.11(a).

4.4 Conclusion

In this chapter, we developed a novel analytic inversion method based on the best linear unbiased estimation criterion where target is embedded in clutter and received signal data is corrupted by thermal noise. The BLUE-based method only requires the first and second statistics of the unknown, spectral density of the clutter which can be estimated from the data by using the method described in the previous chapter. The filter successfully suppresses the effect of noise and clutter.

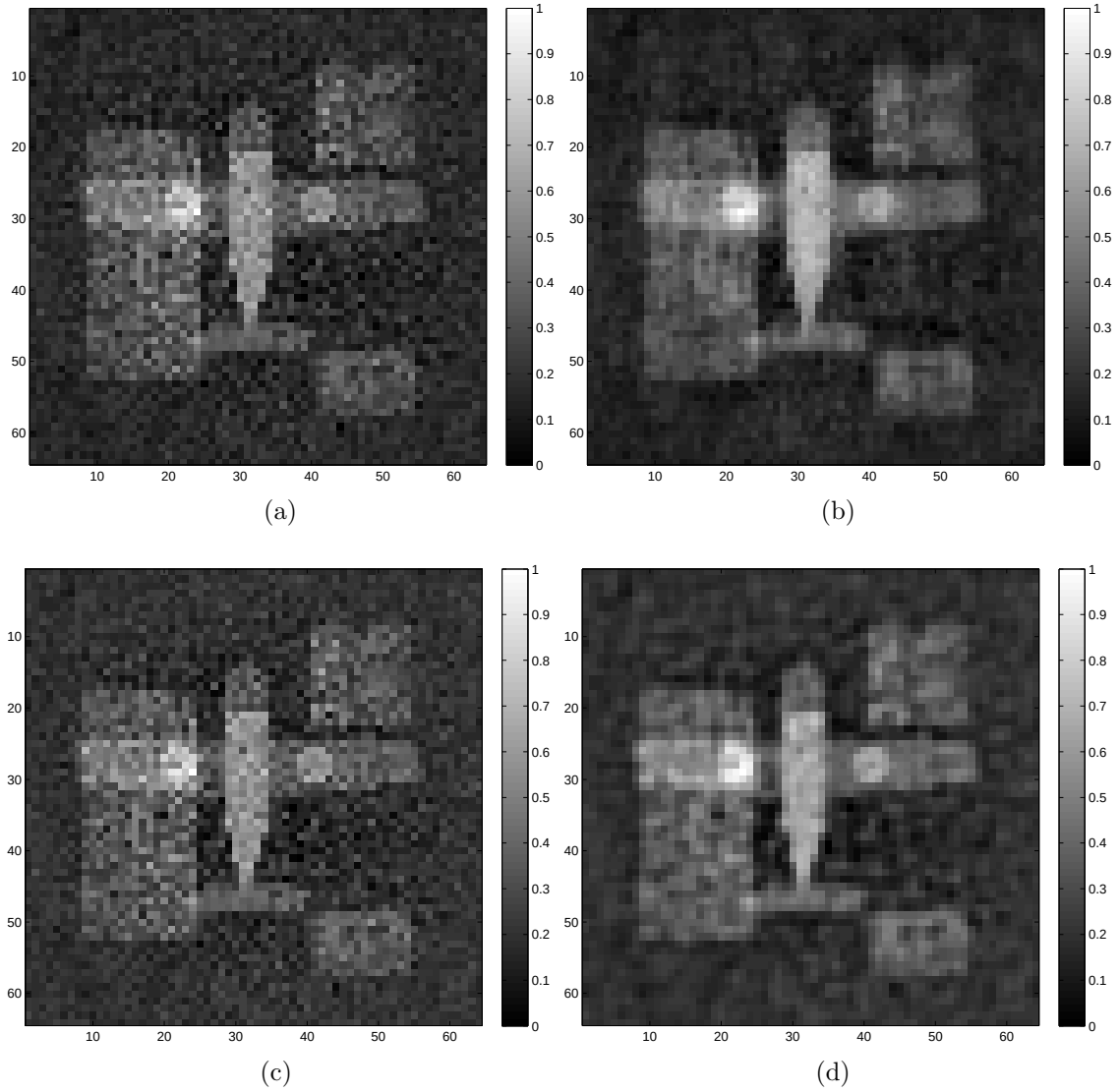


Figure 4.1: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SNR is 10dB and SCR is 0 dB.

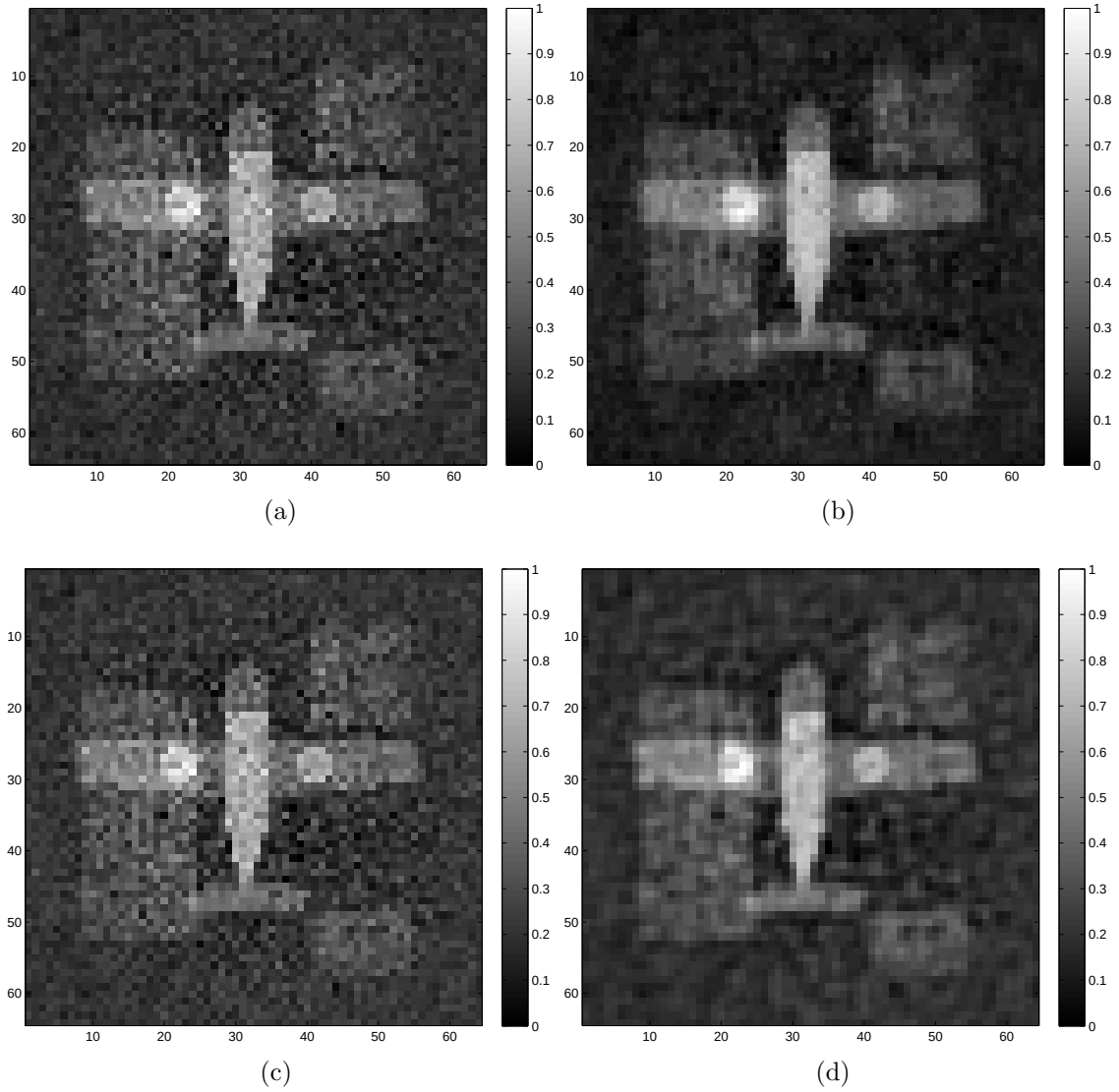


Figure 4.2: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SNR is 10dB and SCR is 8 dB.

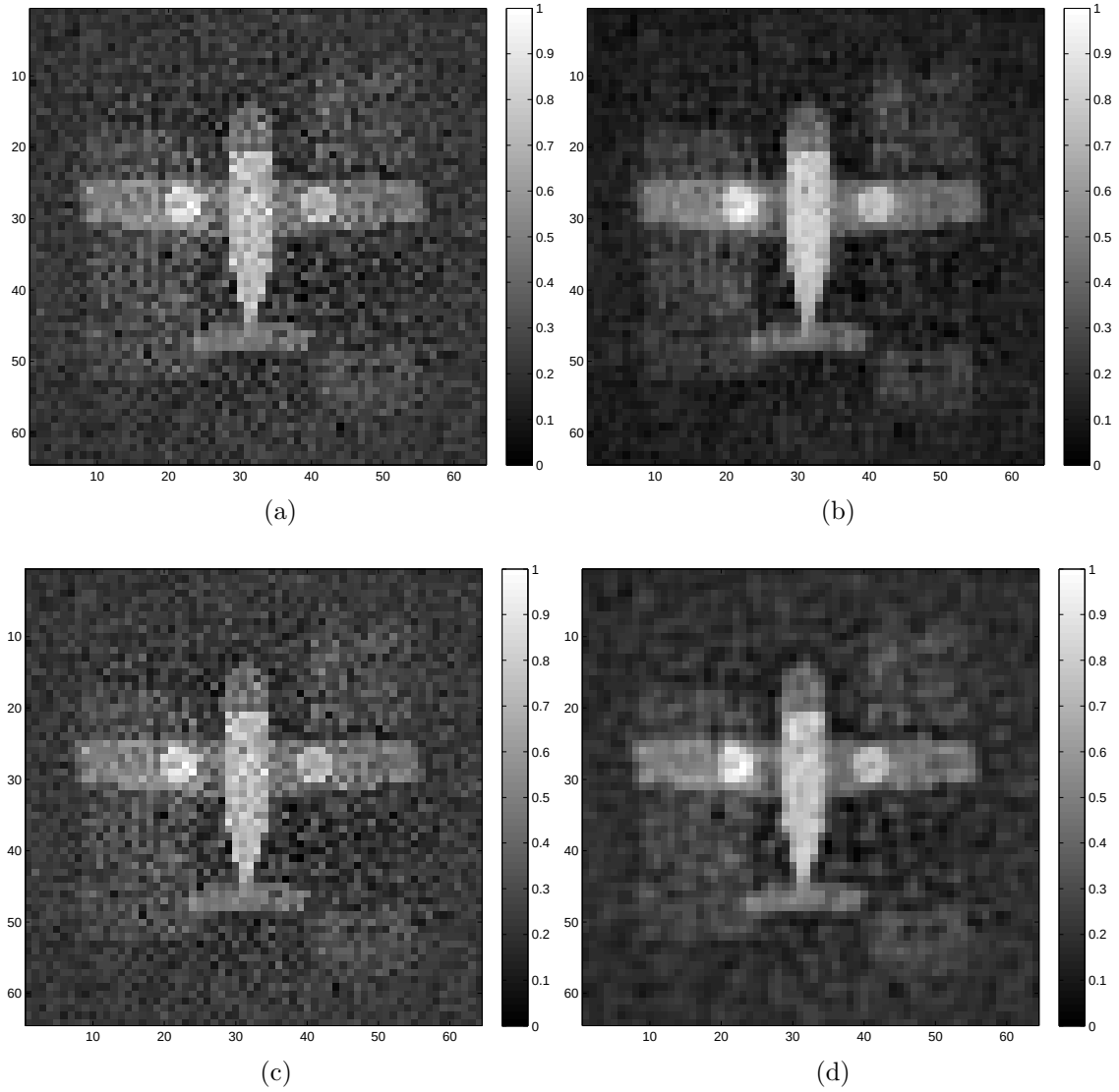


Figure 4.3: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SNR is 10dB and SCR is 16 dB.

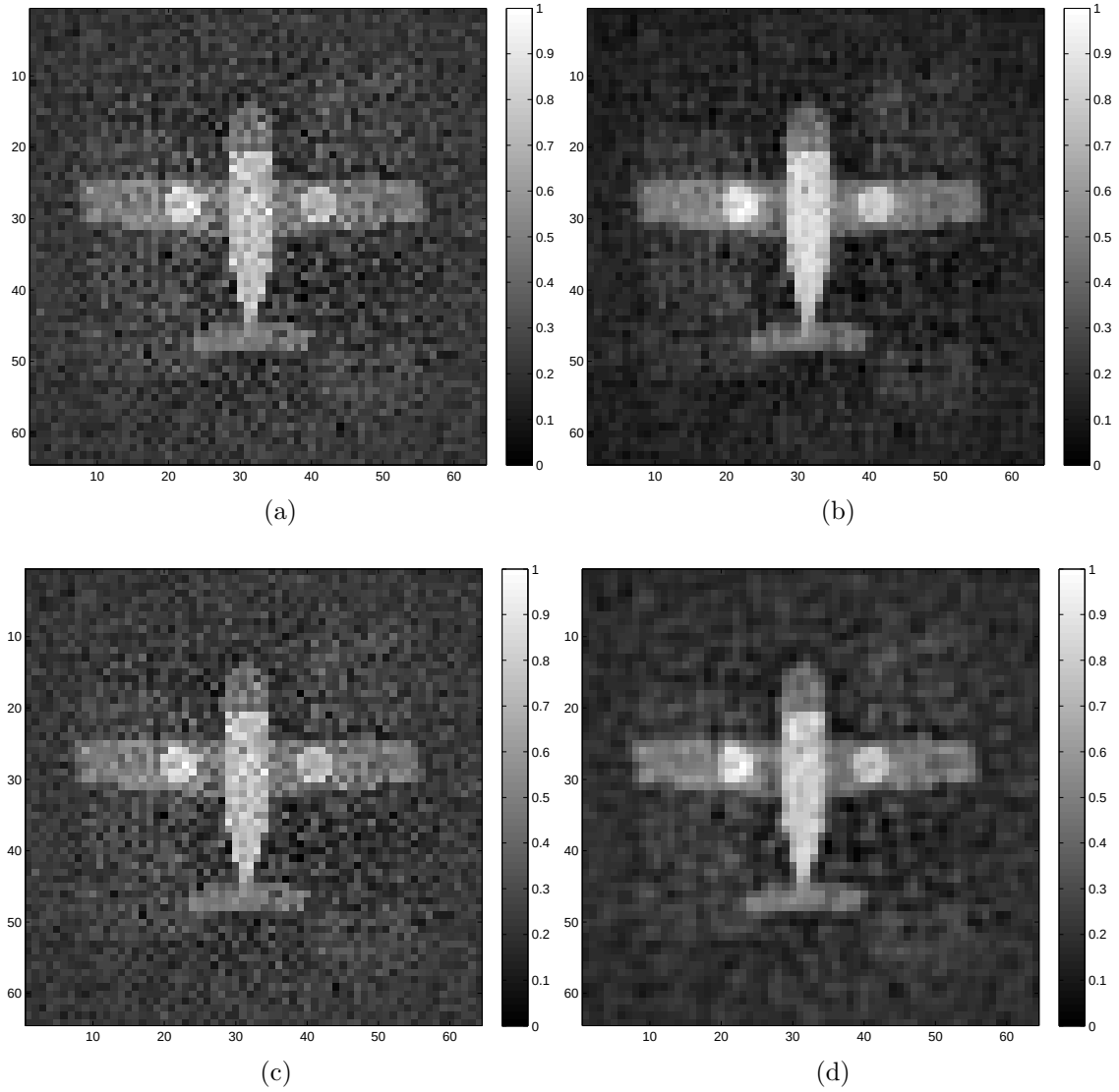


Figure 4.4: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SNR is 10dB and SCR is 24 dB.

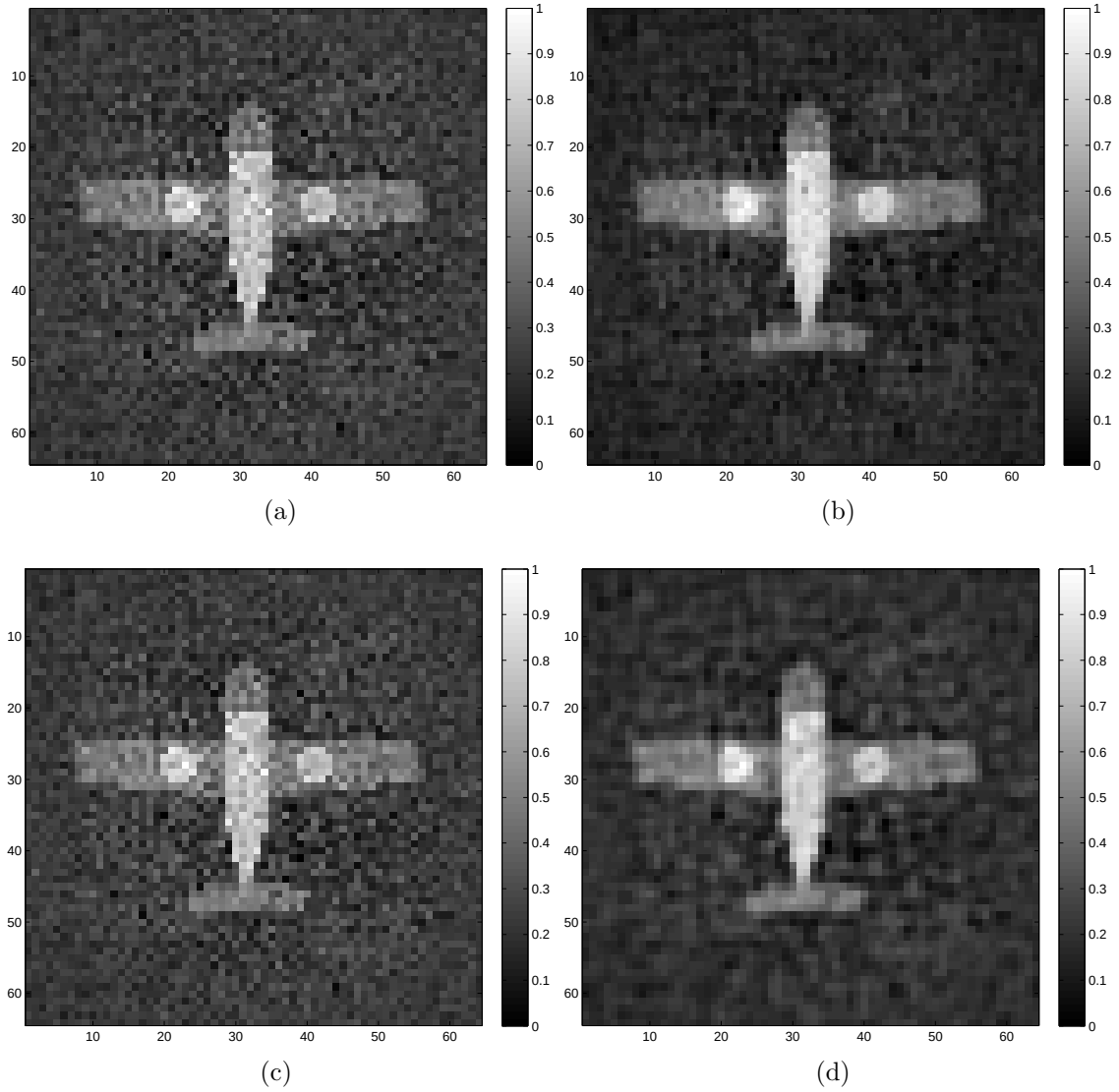


Figure 4.5: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SNR is 10dB and SCR is 32 dB.

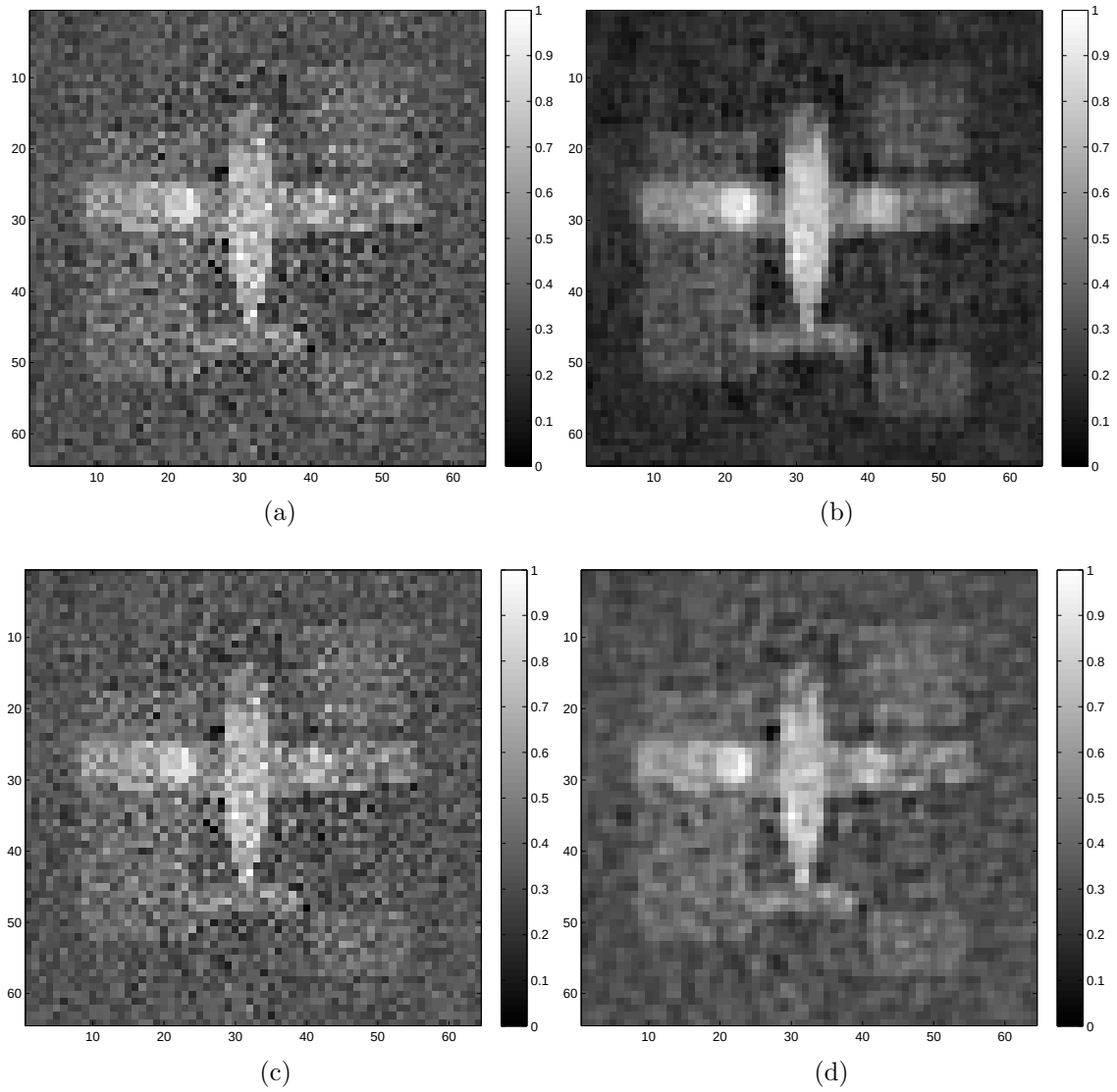


Figure 4.6: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SCR is 10dB and SNR is 0 dB.

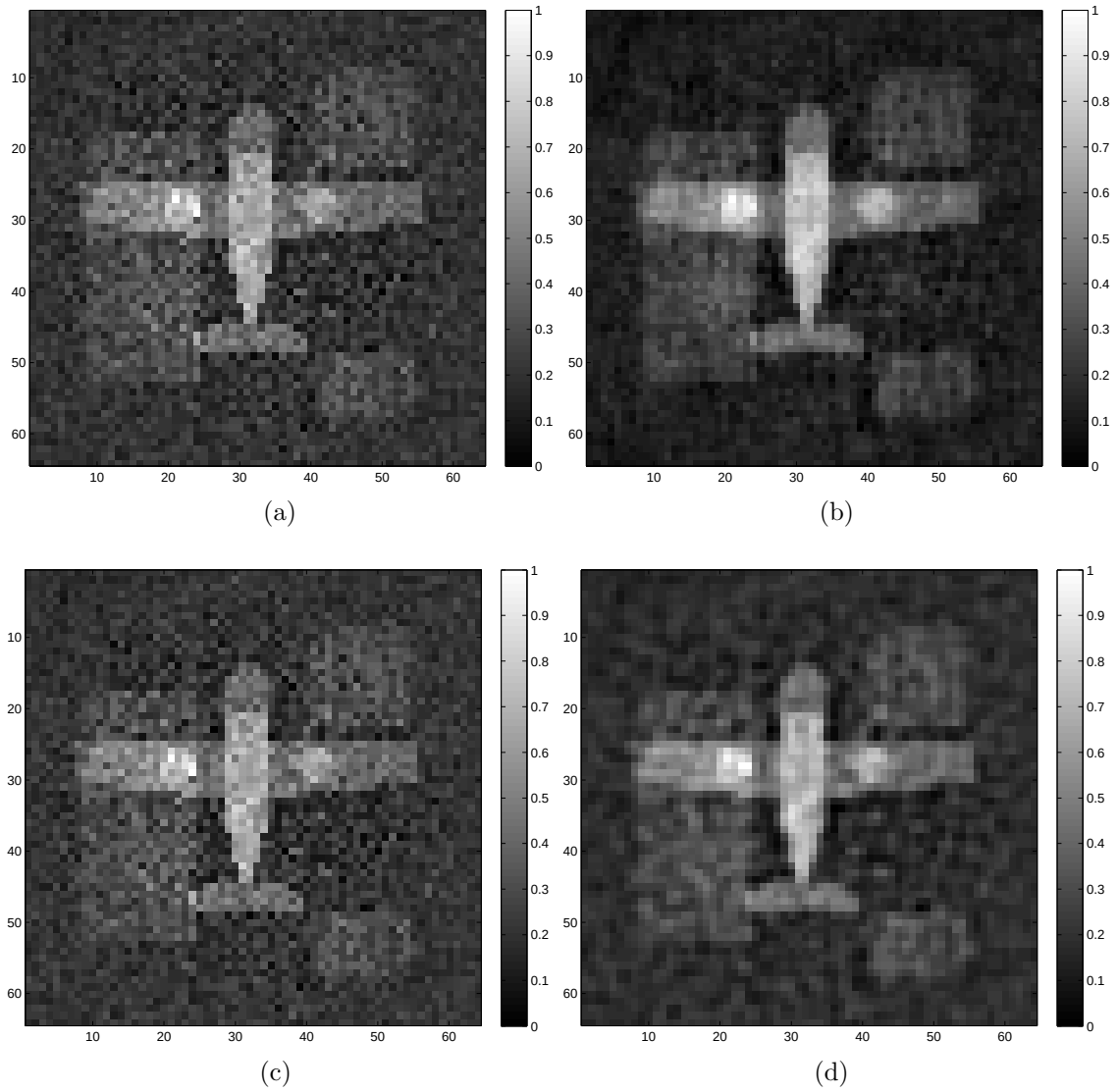


Figure 4.7: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SCR is 10dB and SNR is 8 dB.

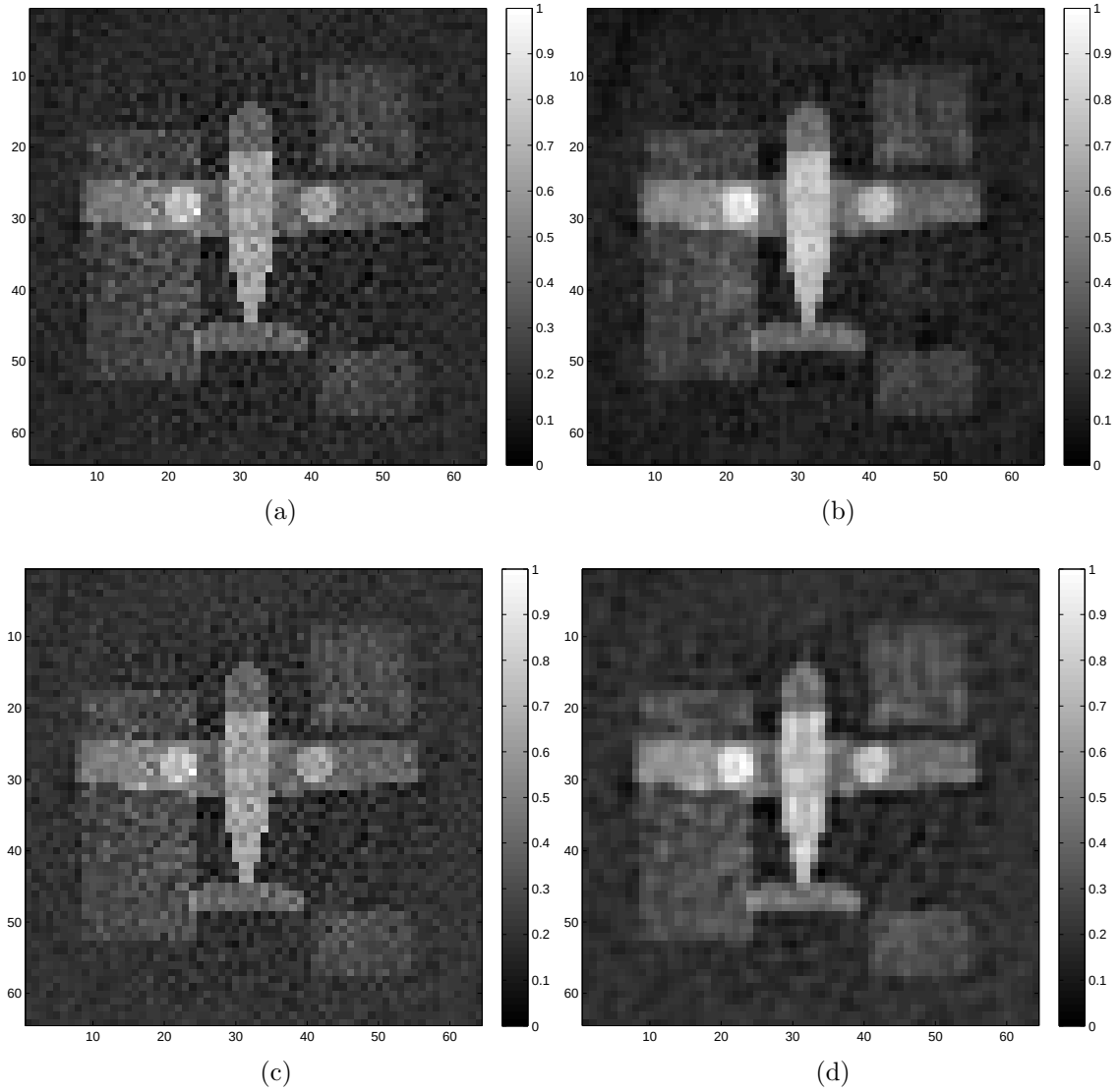


Figure 4.8: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SCR is 10dB and SNR is 16 dB.

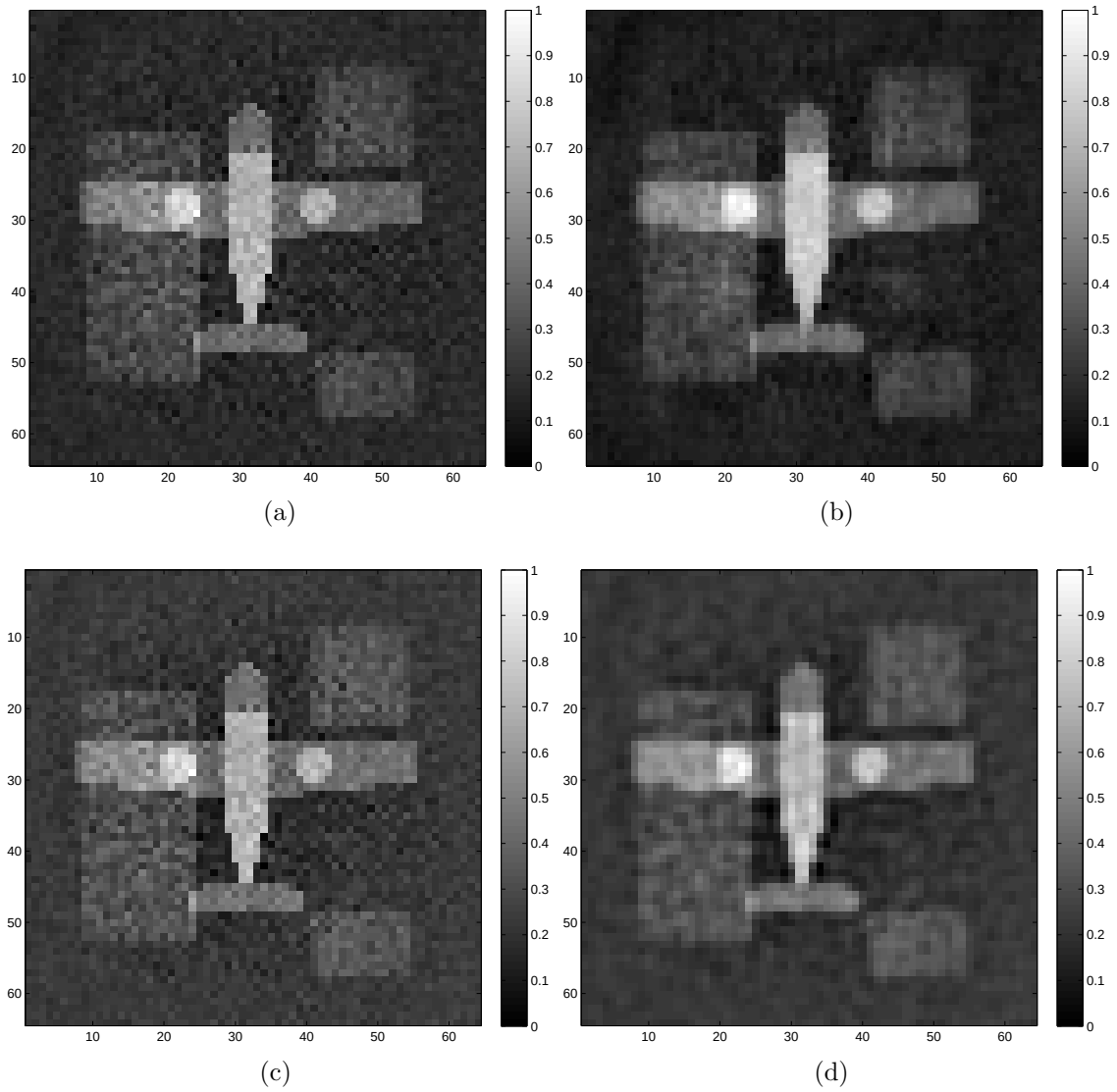


Figure 4.9: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SCR is 10dB and SNR is 24 dB.

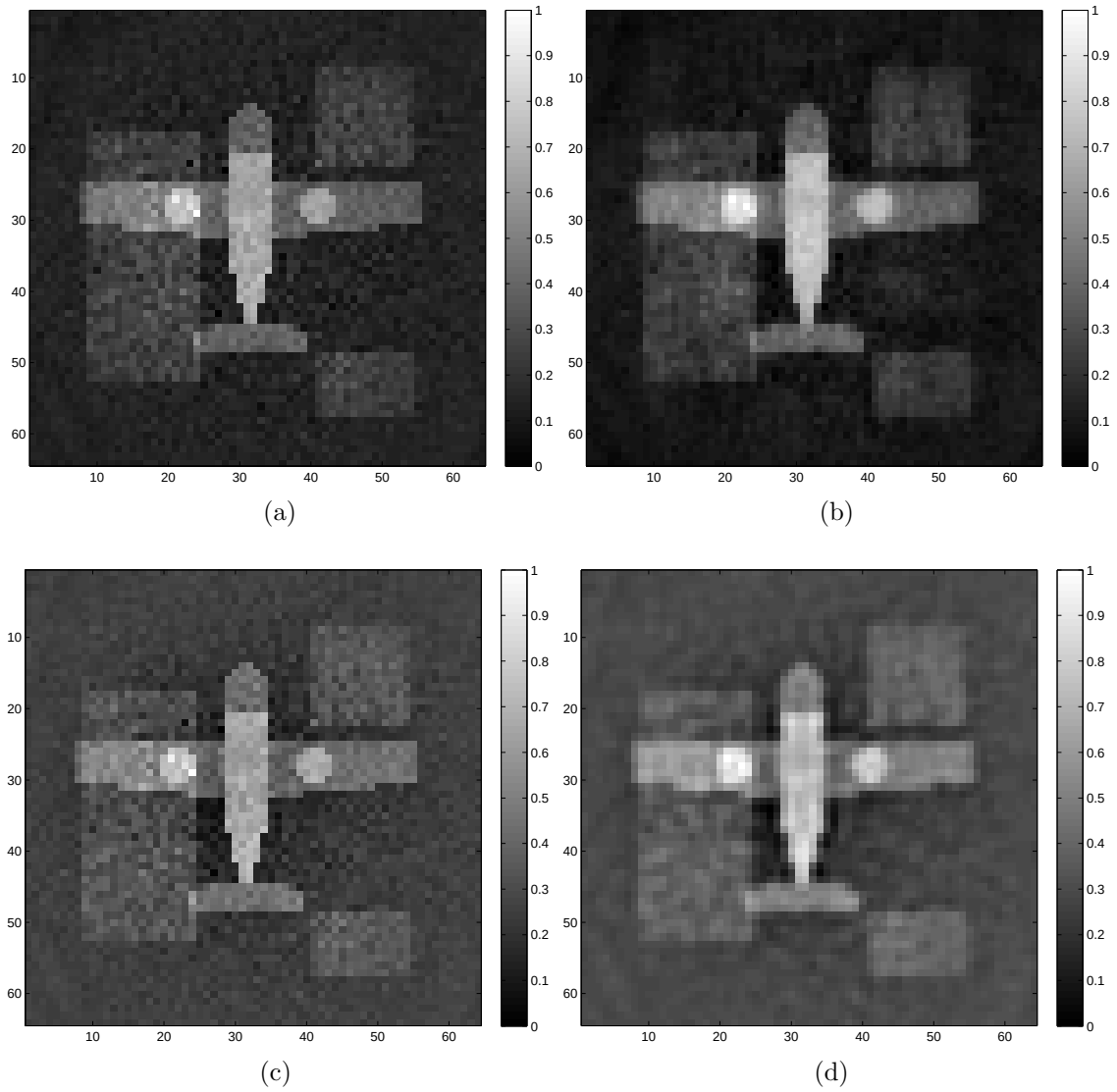
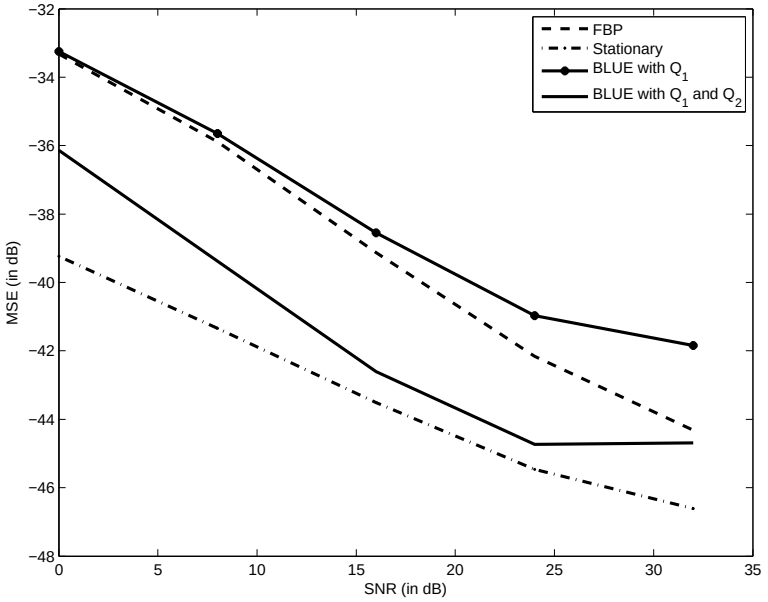
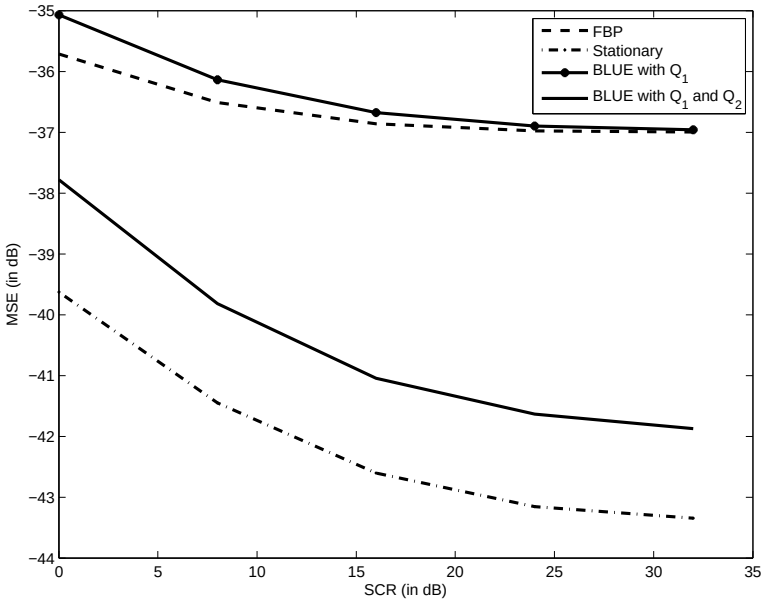


Figure 4.10: The reconstructed images averaged over 10 realizations using (a) the deterministic FBP, (b) MMSE FBP, (c) BLUE after applying filter Q_1, (d) BLUE after applying filter Q_2. The SCR is 10dB and SNR is 32 dB.



(a)



(b)

Figure 4.11: MSE (vertical axis, in dB) versus (a) SNR (horizontal axis) with constant SCR level at 10dB and (b) SCR (horizontal axis) with constant SNR level at 10dB averaged over ten reconstructed images using four different image reconstruction methods.

PART II

CHAPTER 5
FBP-TYPE DIRECT EDGE ENHANCEMENT OF SYNTHETIC APERTURE RADAR IMAGES

In this chapter, we consider a synthetic aperture radar (SAR) system traversing an arbitrary trajectory on a non-flat topography. We present, a novel edge detection method that is applicable directly to SAR received signal. The method first filters the received data, and then backprojects. We design the filter to detect the edges of the scene in different directions at each pixel reconstructed. The method is computationally efficient and may be implemented with the computational complexity of the fast-backprojection algorithms. The method described in this chapter is used in the next chapter to sharpen the reconstructed images. We provide numerical experiments both with the synthetic data and the high-fidelity radar data provided by Air Force Research Laboratories (AFRL) to demonstrate the performance of the method.

5.1 Introduction

SAR images are subjected to automated pattern recognition algorithms that require segmentation of the reconstructed images in many applications. In general, segmentation is an important image processing task to recognize objects or edges in a SAR image, such as houses, vehicles etc. Moreover, segmentation can be addressed by different methods such as clustering and boundary detection-type algorithms [90].

In this chapter, we present a filtered-backprojection (FBP) type image segmentation method that is applied to SAR received signal. The method detects edges directly from the radar data and bypasses the image reconstruction step without any need for image processing. First we model the SAR received signal as a Fourier integral operator (FIO). Next, we describe an FBP-type image reconstruction method to recover and enhance the visible edges of the scene to be imaged. The FBP-type

Portions of this chapter previously appeared as: H. C. Yanik *et al.*, "Computationally efficient FBP-type direct segmentation of synthetic aperture radar images," *in Proc. SPIE, Algorithms for SAR Imagery XVIII*, vol. 8051, 2011, 80510C.

segmentation involves a backprojection step which is given by the L^2-adjoint of the phase of the underlying FIO and a filtering step. The filter is designed so that the resulting point spread function (PSF) of the FBP operator is a differential operator that reconstructs the derivative (or anti-derivative) of the image in desired directions for each pixel reconstructed.

5.2 Forward Modeling and Image Formation

We assume that the earth's surface is located at the position $\mathbf{x} = [\boldsymbol{x}, \boldsymbol{\psi}(\boldsymbol{x})]$, where $\boldsymbol{\psi} : \mathbb{R}^2 \to \mathbb{R}$, is a known smooth function, and scattering takes place in a thin region near the surface. Following [1] and under the single scattering (Born) approximation, we model the received signal $d(s,t)$ as follows:

$$d(s,t) \approx \mathcal{F}[T](s,t) = \int \mathrm{e}^{-\mathrm{i}2\pi\omega(t-R(s,\mathbf{x})/c)} A(\boldsymbol{x},\omega,s) T(\boldsymbol{x}) \,\mathrm{d}\omega \,\mathrm{d}\boldsymbol{x} \tag{5.1}$$

where $s \in \mathbb{R}$ denotes the *slow-time* variable parameterizing the trajectory. $R(s,\mathbf{x})$ is the total travel distance of the electromagnetic waves that takes different forms depending on the imaging modality (see (2.17)-(2.19)), t denotes the *fast-time*, c is the speed of light in vacuum, $T(\boldsymbol{x})$ denotes the surface reflectivity, and A is a complex amplitude function that depends on antenna beam pattern, the transmitted waveform, geometrical spreading factors, etc. [91]

By following similar assumptions as in previous chapters, for some m_A, A satisfies the symbol estimate

$$\sup_{(s,\boldsymbol{x})\in K} |\partial_\omega^\alpha \partial_s^\beta \partial_{x_1}^{\rho_1} \partial_{x_2}^{\rho_2} A(\boldsymbol{x},\omega,s)| \leq B\,(1+\omega^2)^{(m_A-|\alpha|)/2} \tag{5.2}$$

where K is any compact subset of $\mathbb{R}\times\mathbb{R}^2$, and the constant B depends on K, α, β, ρ_1, and ρ_2. This assumption makes the forward operator $\mathcal{F}$ a *Fourier Integral Operator* (FIO) [43]–[45].

The reconstruction of the target reflectivity, $T(\boldsymbol{x})$ requires the inversion of the forward operator, $\mathcal{F}$. Since $\mathcal{F}$ is an FIO, we can form an approximate inverse of $\mathcal{F}$

by another FIO, say $\mathcal{K}$, such that

$$T(\boldsymbol{z}) \approx T(\boldsymbol{z}) = \mathcal{K}\mathcal{F}[T](\boldsymbol{z}) \tag{5.3}$$

Let $\mathcal{L} = \mathcal{K}\mathcal{F}$. We refer to the kernel L of $\mathcal{L}$ as the *point spread function* (PSF) of the system which relates the target scene $T(\boldsymbol{x})$ to the reconstructed scene $T_{ed}(\boldsymbol{z})$. For a perfect reconstruction, L must be a Dirac-delta function. Thus, we choose $\mathcal{K}$ operator such that the kernel of $\mathcal{L}$ becomes as close to Dirac-delta function as possible.

It was shown in [1] that an FBP-type operator $\mathcal{K}$ can be designed to reconstruct an image of the scene

$$\tilde{T}(\boldsymbol{z}) = \mathcal{K}[d](\boldsymbol{z}) := \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\mathbf{z})/c)} Q(\boldsymbol{z},\omega,s) d(s,t)\,\mathrm{d}\omega\,\mathrm{d}s\,\mathrm{d}t. \tag{5.4}$$

Thus, the PSF of the imaging operator $\mathcal{K}$ is given by

$$L(\boldsymbol{z},\boldsymbol{x}) = \int \mathrm{e}^{\mathrm{i}2\pi\omega\phi(\boldsymbol{x},\mathbf{z},s)} Q(\boldsymbol{z},s,\omega) A(\boldsymbol{x},s,\omega)\,\mathrm{d}\omega\,\mathrm{d}s \tag{5.5}$$

where the filter Q_{ed} is chosen such that $L(\boldsymbol{z},\boldsymbol{x}) = \delta(\boldsymbol{z}-\boldsymbol{x})$ and ϕ is given by $\phi(\boldsymbol{x},\boldsymbol{z},s) = R(s,\mathbf{x}) - R(s,\mathbf{z})$. In this chapter, we design a new filter Q_{ed} so that $\mathcal{K}$ reconstructs the edge map of SAR images. The edge map image can be viewed as an approximation to the magnitude of the gradient of the image, i.e. $|\nabla T(\boldsymbol{z})|$.

5.3 Segmentation via Edge Detection

Segmentation is an important image processing operation to extract and locate objects (lines, curves, points etc.) from a given image. These objects are distinctive features of the image and they are utilized to describe, characterize and give meaning to an image.

A basic image segmentation process requires assigning a label to every pixel in an image. Pixels with the same label have similar characteristics. There are many ways to perform image segmentation such as level set methods, graph portioning methods, clustering methods, region growing methods and edge detection [90]. Here,

we describe an edge detection method that can be used for image segmentation tasks for SAR images by applying a differential filter directly to the SAR data at the backprojection step.

Edges can be defined as points at which the transition in pixel intensity is significantly stronger than that of the background at that point. It is known from vector calculus that the gradient vector of a scalar field is a vector field and it points in the direction of the greatest rate of change of the scalar field. Furthermore, magnitude of the vector field corresponds to the greatest rate of change. Thus, edge detection is closely related to the gradient of the image.

A simple method to perform edge detection is applying an operator such as Prewitt, Sobel and then thresholding the image [90]. In [92] an edge detection method is described that contains several steps: First, the image is smoothed with a Gaussian to get rid of the noise and then its gradient is calculated. Depending on the magnitude of the gradient and the angle of the gradient, pixels are labeled, image is thresholded and finally the edge map is formed.

For segmentation of SAR images, the methods described above can be applied after forming the SAR image. Here the goal is to incorporate image formation and edge detection in a single step. The method involves derivation of an edge enhancing (or smoothing) filter and the backprojection operator. The new method reconstructs an approximation to the magnitude of the gradient of the scene, $|\nabla T|$. With this new filter, we can view PSF of the system as an approximation to the derivative (or anti-derivative) of the Dirac-delta function

$$L_{ed}(\boldsymbol{z}, \boldsymbol{x}) \approx \sum_i \alpha_i \delta^{(p_i)}(\boldsymbol{z} - \boldsymbol{x}) \tag{5.6}$$

where $\alpha_i > 0$ and $p_i \in \mathbb{R}$. Here, p_i stands for the p_i^{th} derivative and α_i controls the strength of the edges in the final image.

5.4 Direct Segmentation of the SAR Data

5.4.1 Derivation of an Edge Detection Filter

In this subsection, we derive a new filter Q_{ed} so that the PSF of the imaging system behaves like the derivative of the Dirac-delta function. Thus, reconstructed SAR image will be approximately the edge map of the scene. The filter that results in an approximate Dirac-delta function PSF is given as

$$Q(\boldsymbol{z},\omega,s) = \chi_{\Omega_{\boldsymbol{z}}}(\boldsymbol{\xi}(\omega,s))\frac{\overline{A(\boldsymbol{z},\boldsymbol{\xi}(\omega,s))}}{|A(\boldsymbol{z},\boldsymbol{\xi}(\omega,s))|^2}\,\frac{1}{\eta(\boldsymbol{\xi}(\omega,s),\boldsymbol{z})} \tag{5.7}$$

where $\eta(\boldsymbol{\xi}(\omega,s),\boldsymbol{z})$ is the Jacobian that comes from change of variables

$$(\omega,s)\to\boldsymbol{\xi}=\frac{\omega}{c}\nabla_{\boldsymbol{z}}\,R(s,\boldsymbol{z}). \tag{5.8}$$

With this choice of filter, the PSF of the imaging system is given by

$$L(\boldsymbol{z},\boldsymbol{x}) = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}}\chi_{\Omega_z}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}. \tag{5.9}$$

Here $\chi_{\Omega_{\boldsymbol{z}}}$ is a function that prevents division by 0

$$\Omega_{\boldsymbol{z}} = \{\boldsymbol{\xi}(\omega,s,\boldsymbol{z})\ |A(\boldsymbol{z},s,\omega)\neq 0\}. \tag{5.10}$$

Here we modify the PSF so that when the resulting filter is applied directly to the SAR received data it reconstructs the edge map of the scene. Recall that the relation between actual scene $T(\boldsymbol{x})$ and the reconstructed scene $T_{ed}(\boldsymbol{z})$ is given by

$$T_{ed}(\boldsymbol{z}) = \mathcal{L}[T](\boldsymbol{z}) = \int L_{ed}(\boldsymbol{z},\boldsymbol{x})T(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{5.11}$$

$$= \int_{\Omega_z} \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}}\chi_{\Omega_z}(\boldsymbol{\xi})T(\boldsymbol{x})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{x}, \tag{5.12}$$

$$= \int_{\Omega_z} \mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{z}\cdot\boldsymbol{\xi}}\chi_{\Omega_z}(\boldsymbol{\xi})\hat{T}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{5.13}$$

where $\hat{T}(\boldsymbol{\xi})$ stands for the Fourier transform of $T(\boldsymbol{x})$.

To enhance the edges of the SAR image, an approximate of $|\nabla T(\boldsymbol{z})|$ needs to

be reconstructed. If the gradient operator is applied to both sides of (5.11)

$$\nabla_{\boldsymbol{z}} T(\boldsymbol{z}) \approx \int_{\Omega_z} \nabla_{\boldsymbol{z}} \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{z}\cdot\boldsymbol{\xi}} \chi_{\Omega_z}(\boldsymbol{\xi}) \hat{T}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}, \tag{5.14}$$

$$\approx \int_{\Omega_z} (-\mathrm{i}2\pi\boldsymbol{\xi}) \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{z}\cdot\boldsymbol{\xi}} \chi_{\Omega_z}(\boldsymbol{\xi}) \hat{T}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}. \tag{5.15}$$

The limited aperture directional derivative of T in the direction of the unit vector $\widehat{\boldsymbol{\mu}}$ is defined as follows:

$$\widehat{\boldsymbol{\mu}} \cdot \nabla_{\boldsymbol{z}} T(\boldsymbol{z}) \approx (-\mathrm{i}2\pi) \int_{\Omega_z} \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{z}\cdot\boldsymbol{\xi}}\ \widehat{\boldsymbol{\mu}} \cdot \boldsymbol{\xi}\ \chi_{\Omega_z}(\boldsymbol{\xi}) \hat{T}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}. \tag{5.16}$$

Next, an operator, $\Delta^p_{\widehat{\boldsymbol{\mu}}}$ is defined such that:

$$\Delta^p_{\widehat{\boldsymbol{\mu}}} T(\boldsymbol{z}) := \int_{\Omega_z} \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{z}\cdot\boldsymbol{\xi}} |\widehat{\boldsymbol{\mu}} \cdot \boldsymbol{\xi}|^p \chi_{\Omega_z}(\boldsymbol{\xi}) \hat{T}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}, \quad p \in \mathbb{R}. \tag{5.17}$$

The operator $\Delta^p_{\widehat{\boldsymbol{\mu}}}$ can be interpreted as the p^{th} directional derivative for $p > 0$, or p^{th} integral for $p < 0$. Note that by using this operator, we can represent the magnitude of the gradient of the scene as

$$|\nabla_{\boldsymbol{z}} T(\boldsymbol{z})|^2 = \Delta^2_{\widehat{\boldsymbol{e}}_1} T(\boldsymbol{z}) + \Delta^2_{\widehat{\boldsymbol{e}}_2} T(\boldsymbol{z}) \tag{5.18}$$

where $\widehat{\boldsymbol{e}}_i$, $i = 1, 2$ is the i^{th} column of the 2×2 identity matrix.

With the modified PSF, filter Q_{ed} becomes

$$Q_{ed}(\boldsymbol{\xi}) = \alpha_p |\widehat{\boldsymbol{\mu}}_p \cdot \boldsymbol{\xi}|^p Q(\boldsymbol{\xi}) + \alpha_q |\widehat{\boldsymbol{\mu}}_q \cdot \boldsymbol{\xi}|^q Q(\boldsymbol{\xi}) \tag{5.19}$$

where $p > 0$ and $q < 0$, $\alpha_p,\ \alpha_q \geq 0$ and $\widehat{\boldsymbol{\mu}}_p,\ \widehat{\boldsymbol{\mu}}_q$ are two possibly orthogonal unit directions. The directional derivative of the image is enhanced along $\widehat{\boldsymbol{\mu}}_p$ while it is smoothed in the direction of $\widehat{\boldsymbol{\mu}}_q$. Note that if $p = q = 0$, then $Q_{ed}(\boldsymbol{\xi}) = Q(\boldsymbol{\xi})$; and if $\widehat{\boldsymbol{\mu}}_q = \widehat{\boldsymbol{\mu}}_p^{\perp}$, then the derivative of the image is enhanced in the direction of $\widehat{\boldsymbol{\mu}}_p$ while it is smoothed in the direction perpendicular to $\widehat{\boldsymbol{\mu}}_p$.

More generally,

$$Q_{ed}(\boldsymbol{\xi}) = \sum_i \alpha_i |\widehat{\boldsymbol{\mu}}_{p_i} \cdot \boldsymbol{\xi}|^{p_i} Q(\boldsymbol{\xi}) \tag{5.20}$$

where $\alpha_i > 0$ and $p_i \in \mathbb{R}$ so that edges can be enhanced (or smoothed) in different directions.

5.4.2 Point Spread Function of the Edge Enhanced Reconstruction

With new filter given in (5.20), the PSF of the edge enhancing (smoothing) reconstruction operator becomes

$$L_{ed}(\boldsymbol{z}, \boldsymbol{x}) \approx \sum_i \alpha_{p_i} \int \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\xi} \cdot (\boldsymbol{z} - \boldsymbol{x})} |\widehat{\boldsymbol{\mu}}_i \cdot \boldsymbol{\xi}|^{p_i} \, \mathrm{d}\boldsymbol{\xi}. \tag{5.21}$$

Furthermore, with this choice of PSF, the location and orientation of the edges are preserved.

- If $m = \arg\max_{\{i\}} p_i$, and $p_m > 0$, the strength of the edges is increased (enhanced) by an order of p_m in the direction of $\boldsymbol{\mu}_m$.

- If $n = \arg\min_{\{i\}} p_i$ and $p_n < 0$, the strength of the edges is decreased (smoothed) by an order of p_n in the direction of $\boldsymbol{\mu}_n$.

Thus, with this choice of filter, we can control the directions along which the edges will be enhanced (or suppressed).

5.5 Numerical Experiments

We perform two sets of simulations to demonstrate the performance of the edge-enhanced reconstruction. First set is a synthetic data set simulated by using MATLAB. In the second set of simulations, we use the Civilian Vehicles (CV) dome data set provided by AFRL [42].

5.5.1 Synthetic Data Simulations

We consider two scenes for synthetic data simulations. In the first scene, we have a 4.9km $\times$ 4.9km square target and a 4.9km $\times$ 9.8km rectangular target in a

22km $\times$ 22km scene. Their centers are located at $[-4.1, -0.7]$km and $[4.5, -4.1]$km, respectively. In the second scene, we consider a 12.8km $\times$ 12.8km square target with its center located at $[2.2, -0.4]$km, and there is a 1 pixel-wide (86m) wire-like line that makes a 26° angle with the horizontal axis inside the square. The scene is discretized with 256×256 pixels.

$A(\boldsymbol{x}, \omega, s)$ is assumed to be 1, which corresponds to an isotropic antenna radiating a delta-like impulse and compensating for the geometric spreading factors in the data. Then, using the definition of Dirac-delta function and its homogeneity property,

$$d(s,t) \approx c \int \delta\left(ct - 2|\mathbf{x} - \boldsymbol{\gamma}(s)|\right) T(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \tag{5.22}$$

where we use a discrete version of (5.22) to generate the simulation data. In these experiments, we consider a circular flight trajectory $\boldsymbol{\gamma}(s) = [11 + 22\cos s, 11 + 22\sin s, 6.5]$ km uniformly sampled for $s \in [0, 2\pi]$ at 512 points. (See Figure 5.1)

Figures 5.2,5.3,5.4,5.5,5.6 show reconstructed (part a) and thresholded (part b) images with different values for p, α and $\boldsymbol{\mu}$. p denotes the order of derivative (or anti-derivative) operation, α controls the strength of the edges in the final image and $\boldsymbol{\mu}$ is the direction along which edges will be enhanced (or smoothed).

To threshold the reconstructed image, we use Algorithm 2.

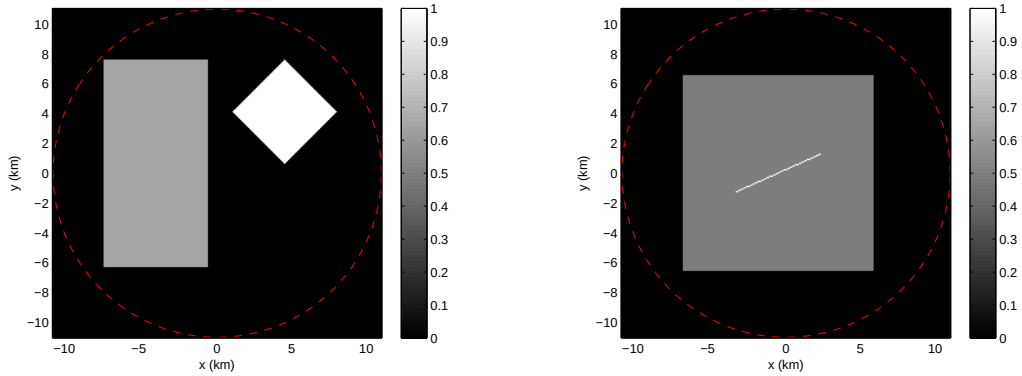
Algorithm 2 An iterative thresholding method for the reconstructed images [90].

Choose an initial threshold, α_0 for image $I(x, y)$
$\alpha_{new} \leftarrow \alpha_0$ and $\alpha_{prev} \leftarrow 0$
while $|\alpha_{new} - \alpha_{prev}| > \epsilon$ **do**
  $\alpha_{prev} \leftarrow \alpha_{new}$
  $I_1 = \{I(x, y) : I(x, y) > \alpha_{new}\}$
  $I_2 = \{I(x, y) : I(x, y) \leq \alpha_{new}\}$
  $m_1 \leftarrow$ mean of pixel values of $I_1(x, y)$ and $m_2 \leftarrow$mean of pixel values of $I_2(x, y)$
  $\alpha_{new} \leftarrow \frac{m_1 + m_2}{2}$

In Figures 5.2 and 5.3, edges are enhanced in all directions. In Figure 5.4, edges along y-direction are enhanced with $\boldsymbol{\mu} = [0, 1]$. In Figure 5.5, edges along x-direction are enhanced with $\boldsymbol{\mu} = [1, 0]$. Finally in Figure 5.6, edges along x and y-directions are enhanced with $\boldsymbol{\mu}_1 = [0, 1]$ and $\boldsymbol{\mu}_2 = [1, 0]$. Note that, in Figures

5.4, 5.5, 5.6 wire is not reconstructed since its direction is not along $\boldsymbol{\mu}$'s used in the examples.

The numerical simulations show that the edges are enhanced (or smoothed) in desired directions and the edges are reconstructed directly from the SAR data successfully.



(a) (b)

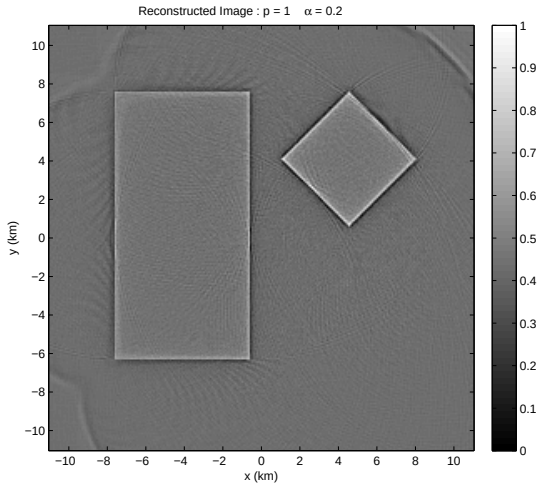
Figure 5.1: Imaging scenes used to generate the radar data.

5.5.2 Civilian Vehicles Dome Data Set

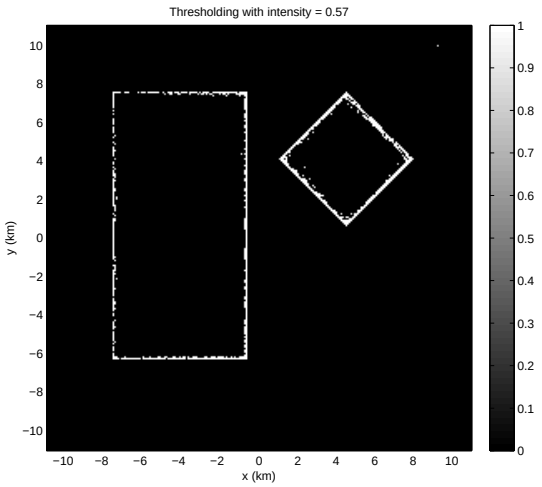
We use CV dome data set as a second set of simulations to demonstrate the performance of the method [42]. In CV dome data set, SAR data is simulated for various civilian vehicles by using high-fidelity electromagnetic wave simulating codes. Cars are modeled by using realistic CAD models and electromagnetic properties of different parts of the cars (glass, metallic, plastic etc.). The radar data provided includes 5760 pulses with 512 frequency bins. The results are shown in Figures 5.7-5.10. We see that the edges are enhanced along desired directions.

5.6 Conclusion

In this chapter, we presented a novel edge detection method applicable directly to SAR received signal. The method first filters the received data, and then back-projects. The filter detects the edges of the scene in different directions at each pixel



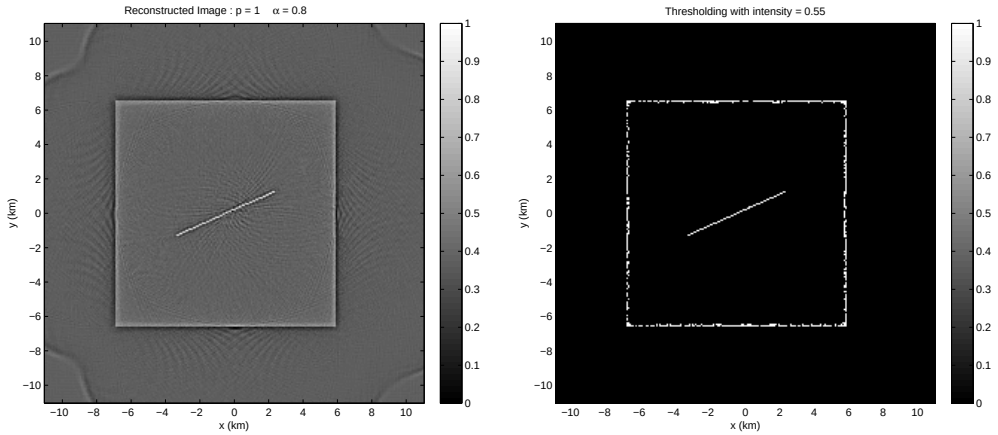
(a) Reconstructed image for the first scene.



(b) Reconstructed image after thresholding for the first scene.

Figure 5.2: Enhancement of edges in all directions for the first scene.

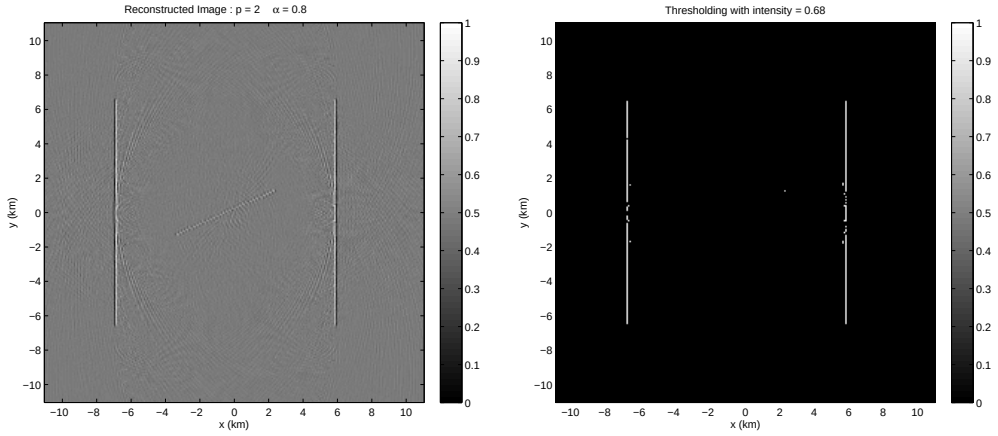
reconstructed. The method is computationally efficient and may be implemented with the computational complexity of the fast-backprojection algorithms.



(a) Reconstructed image for the second scene.

(b) Reconstructed image after thresholding for the second scene.

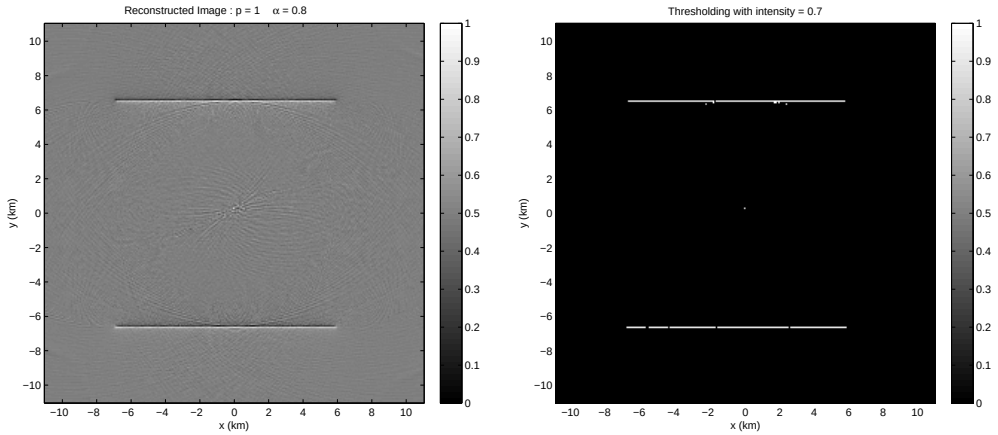
Figure 5.3: Enhancement of edges in all directions for the second scene.



(a) Reconstructed image for the second scene.

(b) Reconstructed image after thresholding for the second scene.

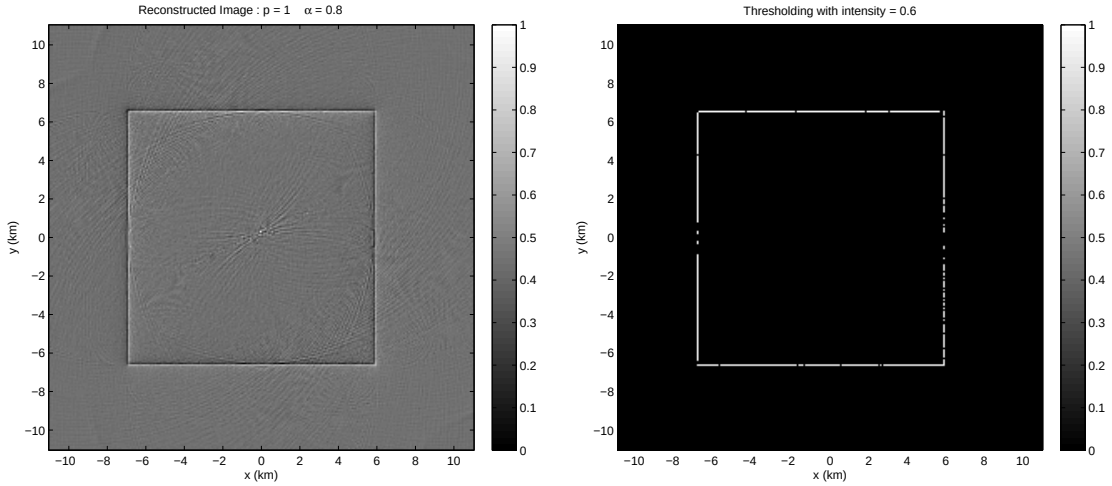
Figure 5.4: Enhancement of edges in x-direction for scene 2 with $\boldsymbol{\mu} = [1, 0]$.



(a) Reconstructed image for the second scene.

(b) Reconstructed image after thresholding for the second scene.

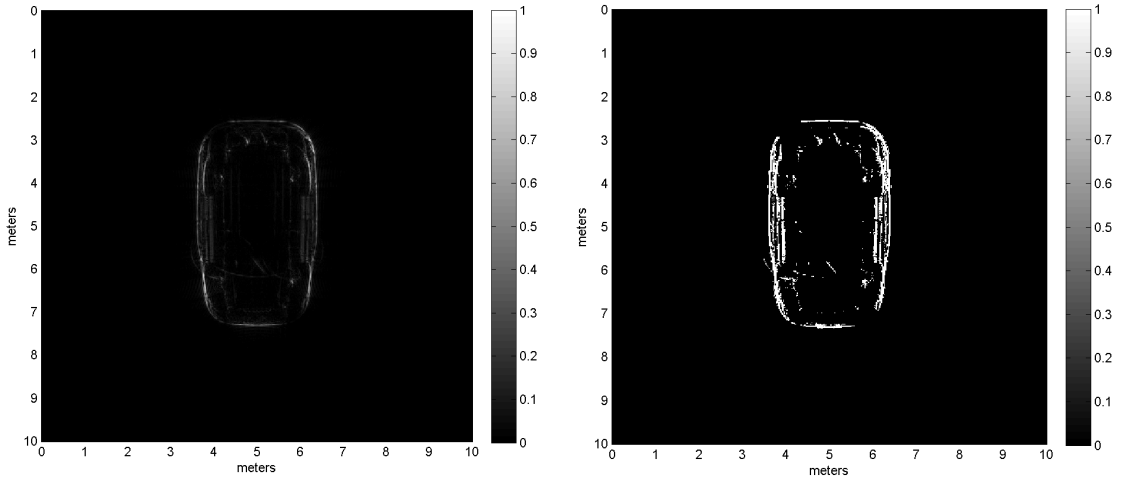
Figure 5.5: Enhancement of edges in y-direction for scene 2 with $\boldsymbol{\mu} = [0, 1]$.



(a) Reconstructed image for the second scene.

(b) Reconstructed image after thresholding for the second scene.

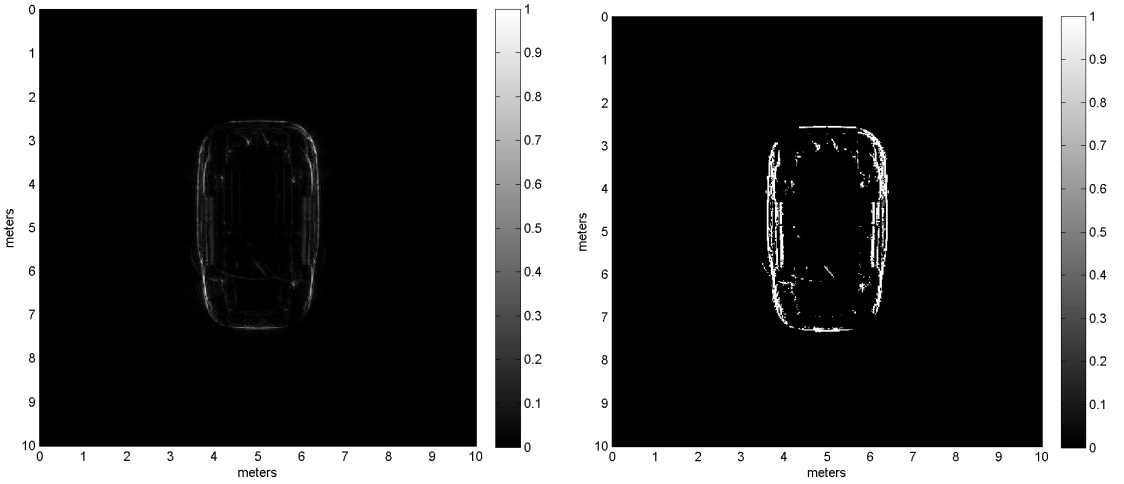
Figure 5.6: Enhancement of edges in x and y-directions for scene 2 with $\boldsymbol{\mu}_1 = [1, 0]$ and $\boldsymbol{\mu}_2 = [0, 1]$.



(a) Reconstructed image for the Jeep data.

(b) Reconstructed image after thresholding for the Jeep data.

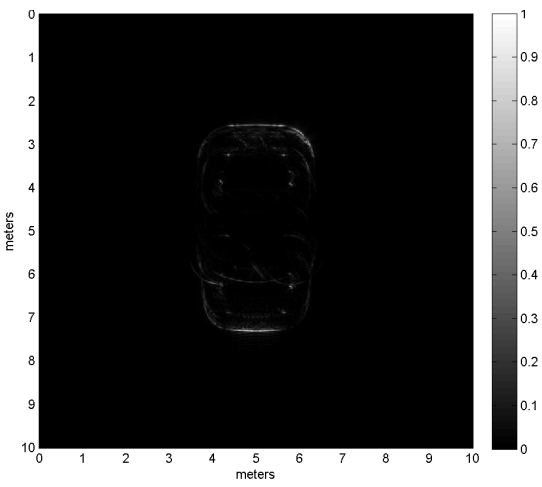
Figure 5.7: Original image reconstructed with FBP.



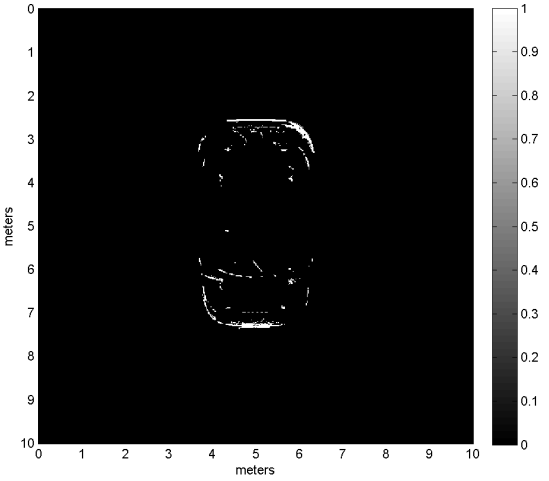
(a) Reconstructed image for the first scene.

(b) Reconstructed image after thresholding for the Jeep data.

Figure 5.8: Enhancement of edges in all directions.

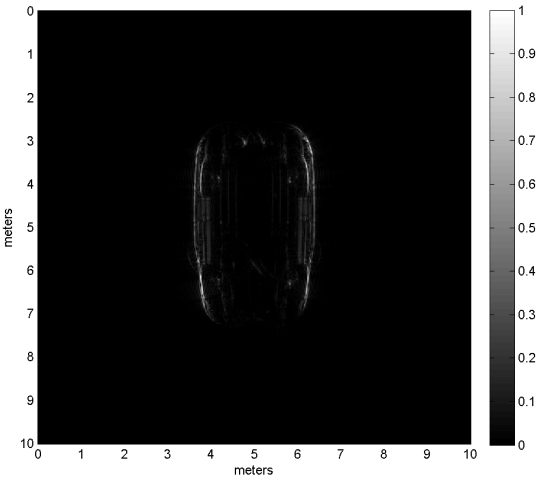


(a) Reconstructed image for the first scene.

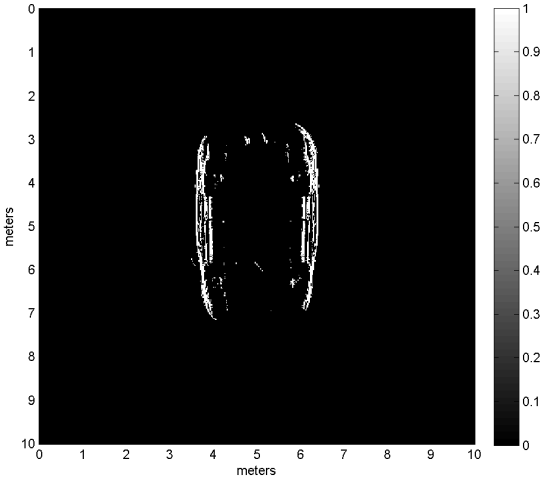


(b) Reconstructed image after thresholding for the Jeep data.

Figure 5.9: Enhancement of edges along x_1 direction, i.e. $\mu = [1, 0]$.



(a) Reconstructed image for the Jeep data.



(b) Reconstructed image after thresholding for the Jeep data.

Figure 5.10: Enhancement of edges along x_2 direction, i.e. $\mu = [0, 1]$.

CHAPTER 6
ITERATIVE ANALYTIC SAR INVERSION WITH L^p-TYPE REGULARIZATION

In this chapter, we formulate the SAR image formation problem with non-quadratic edge-preserving prior models and address the resulting optimization problems iteratively by sequences of FBP operators. These algorithms are computationally efficient with the complexity equal to that of the fast-backprojection. Additionally, they are applicable to the so-called interrupted SAR systems [93], [94] where data samples are limited and measured data is corrupted by noise.

First, we develop an iterative reweighted-type SAR image formation algorithm. SAR image reconstruction problem is modeled as an L^p-norm constrained minimization problem. Then, we approximate this problem by a sequence of L^2-norm constrained quadratic minimization problems as inversions of the underlying GRT. We approximate the solution of each optimization problem iteratively with a sequence of analytic, FBP operators. At every iteration, we use the solution obtained in the previous iteration to determine the filter of the next FBP operator.

Next, we develop an iterative shrinkage-type algorithm that is applicable to limited data. The algorithm includes iterative applications of forward projection and backprojection operators.

The methods that are developed in this chapter have the following advantages: (1) They can be implemented efficiently using the fast-backprojection algorithms. (2) They do not require *a priori* information. (3) They are applicable to a wide range of image reconstruction problems including other SAR modalities, X-ray computed tomography and acoustics where GRTs naturally arise. (4) They are particularly suitable for SAR reconstruction with limited aperture.

We demonstrate the performance of the methods using both synthetic data and the Civilian Vehicles (CV) dome data set provided by the Air Force Research Laboratory (AFRL) [42] and compare its performance to those of the other state-of-the-art sparse signal recovery techniques.

6.1 Introduction

SAR image reconstruction is a challenging problem. From a real world application point of view, the challenge is threefold: (1) The amount of data collected for a single acquisition may be very large to transmit and process in real time. As an example, in Air Force Research Laboratories' (AFRL) wide-angle SAR experiment, 2.718Gb/seconds data was collected over 647 seconds. This resulted in terabytes of data [2]. Therefore, it is desirable to reduce the data needed for image reconstruction. (2) Measured data in many practical applications may be incomplete [35]–[38]. (3) Finally, there is a growing interest in SAR systems that are capable of operating in multiple modes. These systems are referred to as the *interrupted SAR systems* and have various operation modes such as imaging, target recognition and tracking [93]–[97]. Switching between different operation modes may result in missing or incomplete data.

All these limitations require the development of high fidelity SAR image reconstruction with under-sampled or missing data. In addition to the undersampled data, the noise is always present in the measurements. To address this, the sparse signal recovery techniques offer a powerful approach and have been extensively used in the SAR imaging community [98]–[102]. In [103], [104], it was shown that given the knowledge about sparsity of a signal, i.e. the number of non-zero elements, the signal may be reconstructed with fewer samples than the Nyquist theorem requires.

6.2 Related Work

Many solutions with performance guarantees are developed and presented in recent years for sparse signal recovery. These methods can be broadly categorized into four classes: (1) L^p-norm constrained optimization algorithms [30], [50]; (2) greedy pursuits-type algorithms [88], [101], [105]–[112]; (3) iterative reweighted-type algorithms (IRtA) [34], [113]–[115]; and (4) iterative shrinkage-type algorithms (IStA) [116]–[128].

Within L^p-norm constrained optimization framework, Cetin *et al.* used L^p-norm constraint for $0 < p \leq 1$ in [50] with a non-quadratic data fidelity term to suppress noise and a total variation term to enhance features in the image. Kragh

et al. applied an iterative majorization-minimization scheme to a quadratic cost function in [30].

A second class of algorithms for sparse signal recovery is the greedy-pursuits-type algorithms. The basis pursuits (BP) [101], [111], [112], matching pursuits (MP) [105] and orthogonal matching pursuits (OMP) [109] are widely used within this class of algorithms. The BP algorithm decomposes a signal into a superposition of dictionary elements by means of L^1-norm minimization [101], [111], [112]. Mallat *et al.* introduced the MP algorithm in [105]. Matching pursuits are used to decompose signals into a linear expansion of waveforms that are selected from a redundant dictionary of Gabor functions to give adaptive signal representations. Many greedy algorithms are derived from the MP method. In [106], Burns *et al.* decomposed radar signals using a dictionary of functions. In [33] an MP method was used with a nonlinear iterative algorithm. In [32], another MP-based algorithm used a tree search algorithm. In [109], Tropp used OMP to solve the sparse signal recovery problem and unified the prior BP-based methods and extended them to OMP.

Another set of algorithms for the sparse signal recovery techniques is the iterative reweighted-type algorithms. In [113], iterative reweighted least squares minimization procedures are used while forming the SAR image with a sequence of L^2-norm based minimizations to solve an L^1-norm minimization problem. Bayesian estimation techniques are also used within this framework [31], [53]. In [53], Vu *et al.* proposed an approach similar to maximum *a priori* estimation with iterative minimization procedures. In [115], Candes *et al.* solved a sequence of weighted L^1-norm minimization problems where the weights used for the next iteration are calculated from the current solution with fewer measurements. In [114] algorithms for optimal basis selection are derived by minimizing various diversity measures such as L^p-norm, Gaussian and Shannon entropies. It was shown that the Gaussian entropy minimization is equivalent to the L^0-norm minimization problem. In [54], the SAR image reconstruction problem was formulated as an L^p-norm constrained quadratic inversion problem approximated with a sequence of L^2-norm constrained problems analytically.

The last set of algorithms belongs to the family of iterative shrinkage-type al-

gorithms, first introduced by Donoho *et al.* in [116]–[118]. Iterative shrinkage-type algorithms can be considered as gradient-based algorithms for solving sparse signal recovery problems via iterative multiplications with the underlying linear transformation matrices [126]. In radar imaging, this corresponds to iterative applications of forward projection and backprojection operators. There are different derivations of the iterative shrinkage-type of algorithms based on thresholding [126], expectation maximization [121] and majorization-minimization [119], [120].

In this chapter, we develop novel iterative reweighted-type and shrinkage-type algorithms to suppress noise that is applicable to limited data. For the rest of the thesis, L^p-norm will correspond to the norms for $0 < p \leq 1$.

6.3 Sparse Signal Recovery Problem

The goal of the sparse signal recovery techniques is to solve an underdetermined (or ill-conditioned) system of linear equations given by

$$\mathbf{FT} + \mathbf{n} = \mathbf{d} \tag{6.1}$$

where $\mathbf{n}$ is the measurement noise, $\mathbf{F} \in \mathbb{R}^{n \times m}$, $n < m$ is the linear transformation matrix , $\mathbf{T} \in \mathbb{R}^m$ is the "sparse" input signal to be recovered from the measurements $\mathbf{d} \in \mathbb{R}^n$.

The objective of the sparse signal recovery techniques is to find $\mathbf{T}$ which solves the equation given in (6.1) and has the minimum number of non-zero elements, i.e. L^0-norm. Therefore, the sparse recovery approach to solving (6.1) is given by

$$\min |\mathbf{T}|_0 \quad \text{subject to} \quad |\mathbf{d} - \mathbf{FT}|_2^2 \leq \epsilon \tag{6.2}$$

where $|\mathbf{n}|_2^2 < \epsilon$. This is an NP-hard problem [129].

As an alternative, under the conditions given in Appendix H, we can replace L^0-norm with L^p-norm and define

$$\min |\mathbf{T}|_1 \quad \text{subject to} \quad |\mathbf{d} - \mathbf{FT}|_2^2 \leq \epsilon \tag{6.3}$$

for $p = 1$. (6.3) is a *convex relaxation* of (6.2) [109].

Another approach to address the problem stated in (6.3) consists of the following [130]: (1) the minimization of an objective functional that includes a quadratic error term to suppress noise, and (2) a sparsity inducing regularization term to represent $\mathbf{T}$. This approach results in the following optimization problem:

$$\min_{\mathbf{T}} |\mathbf{d} - \mathbf{F}\mathbf{T}|_2^2 + \lambda \mathbf{1}^T \rho(\mathbf{T}) \tag{6.4}$$

where $\mathbf{1} \in \mathbb{R}^n$ is a vector of ones; λ is the regularization parameter that determines the trade-off between the sparsity and the residue $\mathbf{d} - \mathbf{F}\mathbf{T}$; and $\rho, \mathbb{R}^m \to \mathbb{R}$ is a sparsity inducing scalar-valued potential function that operates pointwise on its argument.

6.4 Received Signal, Target, and Noise Models

As described in detail in the previous chapters, we can model the measurements in many SAR systems as follows:

$$\tilde{d}(s,t) = \mathcal{F}[T](s,t) := \int \mathrm{e}^{\mathrm{i}\Phi(\omega,\boldsymbol{x},s,t)} A(\omega,\boldsymbol{x},s,t) T(\boldsymbol{x}) \mathrm{d}\omega \mathrm{d}\boldsymbol{x} \tag{6.5}$$

where $T(\boldsymbol{x})$ is the target to be recovered from the measurements $\tilde{d}(s,t)$, A is the amplitude function of the forward operator $\mathcal{F}$, and

$$\Phi(\omega,\boldsymbol{x},s,t) = 2\pi\omega(t - R(s,\boldsymbol{x})/c) \tag{6.6}$$

is the phase function. $R(s,\boldsymbol{x})$ is a range function which takes different forms as given in (2.17)-(2.19) depending on the SAR modality.

We assume that the measurements are corrupted by zero-mean Gaussian noise $n(s,t)$ and model the measured data as

$$d(s,t) = \mathcal{F}[T](s,t) + n(s,t). \tag{6.7}$$

The autocovariance function of the noise is given by

$$\mathrm{E}[n(s,t)\overline{n(s',t')}] = \sigma^2(s)\delta(s-s')\mathcal{R}_n(t,t') \tag{6.8}$$

where

$$\mathcal{R}_n(t,t') = \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-t')}|\tilde{S}_n(\omega)|^2\,\mathrm{d}\omega. \tag{6.9}$$

We refer to $|S_n(\omega)|^2$ as the power spectral density function of noise.

From a Bayesian estimation point of view, (6.4) can be considered as the maximum *a posteriori* (MAP) estimation of the unknown T with an appropriate prior from the noisy measurements d [131]. Let $p_{\mathrm{T}}(T|\mathrm{d}=d),\ p_{\mathrm{d}}(d|\mathrm{T}=T),\ p_{\mathrm{T}}(T),\ p_{\mathrm{d}}(d)$ denote the *a posteriori* distribution of the target given the measurements, the distribution of the measurements given the target, the distribution of the target and the distribution of the measurements, respectively. Using the Bayes' Theorem, we write

$$p_{\mathrm{T}}(T|\mathrm{d}=d) = \frac{p_{\mathrm{d}}(d|\mathrm{T}=T)p_{\mathrm{T}}(T)}{p_{\mathrm{d}}(d)} \tag{6.10}$$

Let $L(T|d)$ denote the log-likelihood of the *a posteriori* distribution of the target scene given the measurements. Using (6.10), we can write

$$L(T|\mathrm{d}=d) = L(d|\mathrm{T}=T) + \log(p_{\mathrm{T}}(T)) - \log(p_{\mathrm{d}}(d)). \tag{6.11}$$

Using the MAP estimation, we can find an estimate for T as follows:

$$T_{MAP} = \max_T L(T|\mathrm{d}=d) \tag{6.12}$$

$$= \max_T L(\mathrm{d}=d|T) + \log(p_{\mathrm{T}}(T)) \tag{6.13}$$

$$= \max_T L(\mathrm{d}=d,T), \tag{6.14}$$

i.e. the MAP estimate maximizes the joint distribution as well [31]. When the prior distribution is Gaussian then (6.4) corresponds to an MMSE estimate. However, MMSE estimates result in over-blurred images if applied globally. We overcome this shortcoming by applying MMSE criteria locally in Chapter 3. Here, we wish to

use sparsity promoting priors for p_T to preserve edges of the reconstructed images. The Markov random fields (MRF) offer a comprehensive framework to design such priors [131], [132]. The probability distribution of an MRF can be expressed using the Gibbs distribution as follows [131], [133]:

$$p_T(T) = \frac{1}{Z} \exp\left(- \sum_{c \in \mathcal{C}} \rho_c(T_c) \right) \tag{6.15}$$

where Z is a normalization constant, ρ_c is any function of a local group of points c among all local groups of points $\mathcal{C}$. Finally, under some conditions, *Hammersley-Clifford* theorem states that a random field is an MRF if and only if its probability distribution can be represented using the Gibbs distribution [134], [135].

There are many possible sparsity inducing potential functions for the choice of ρ that are convex and non-convex. The ones that are investigated in this thesis are listed in Table 6.1 [31], [136]–[143].

Table 6.1: Table showing the sparsity inducing potential functions, $\rho(f)$ for an input function f investigated in the thesis.

Convex functions	$\frac{f^2}{1+f^2}$	f^2	$\log(1+f^2)$	$\frac{\lvert f\rvert}{1+\lvert f\rvert}$	
Non-convex functions	$\lvert f\rvert$	$\log\cosh(f)$	$\min(\lvert f\rvert^2, 2f-1)$	$\lvert f\rvert^p$	$\frac{\lvert f\rvert^p}{1+f^{q-p}}$

6.5 Analytic SAR Image Formation with Sparsity Promoting Prior Models

6.5.1 Problem Definition

In Chapter 3, we describe a novel SAR inversion method based on locally defined MMSE criterion. The method results in an FBP-type analytic reconstruction technique in which the FBP filter requires second-order statistics of the target, clutter and the noise. Furthermore, we estimate these parameters from the data as described in Appendix E.

The MMSE criterion function inherently relies on the Gaussianity of the underlying processes of the target, clutter and noise, hence the L^2-norm. If applied globally, the L^2-norm minimization over-smooths the images resulting in loss of features in reconstructed images. In this chapter, we develop an analytic inversion method with the L^p-norm ($p > 0$) constraint. Furthermore, we approximate the L^p-norm constrained inversion problem with a sequence of L^2-norm constrained inversion problems.

There are many sparse signal recovery techniques to solve the problem (6.4). Using the target model described in the previous section and the prior $\exp(-\rho(T))$, we can address a solution to this problem by considering the following optimization problem:

$$\mathcal{J}_{re}(T) = \lambda \int \rho(T(\boldsymbol{z}))\mathrm{d}\boldsymbol{z} + \frac{1}{2}|d - \mathcal{F}[T]|_2^2 \tag{6.16}$$

where ρ is defined in the previous section and can be one of the functions listed in Table 6.1. We will find solutions to solve (6.16) by using the techniques described in [120], [123], [144]. It should be noted here that we can easily replace $\rho(T(\boldsymbol{z}))$ with $\rho(\nabla T(\boldsymbol{z}))$ to find a total-variations estimate for the target scene.

6.5.2 Approximate Solutions for the Optimization Function

6.5.2.1 Iterative Reweighted-type Analytic Reconstruction

For $\rho(T) = |T|_{p,\alpha}^p$, we write

$$T \sim \exp(-|T|_{p,\alpha}^p) \tag{6.17}$$

where

$$|T|_{p,\alpha}^p = \int \frac{|T(\boldsymbol{z})|^p}{\alpha(\boldsymbol{z})}\mathrm{d}\boldsymbol{z} < +\infty, \quad \alpha(\boldsymbol{z}) > 0. \tag{6.18}$$

(6.17) can be considered as a MAP estimate of the target T, under a sparsity inducing prior, i.e. L^p-norm. Similarly we can model a total variations estimate of

the target by using the following model:

$$\nabla T \sim \exp(-|\nabla T|_{p,\alpha}^{p}). \tag{6.19}$$

Using (6.17), the MAP estimation of T is equivalent to the optimization problem given by

$$\mathrm{P}_1: \quad \min\{\lambda|T|_{p,\alpha}^{p} + |d - \mathcal{F}[T]|_{2,\mathcal{R}_n}^{2}\}\ . \tag{6.20}$$

In order to solve this optimization problem efficiently, we will convert P_1 to a sequence of quadratic problems P_2 and solve each problem analytically. This will result in an analytic and fast solution to L^p-norm constrained problem described in (6.20). To do that, we will adapt the methodology described in [144] into our problem formulation.

Let $T_k,\ k = 0, 1, \ldots$ be the solution of the P_2 problem at the k^{th} iteration. Then, we define an operator $\mathcal{I}_{\alpha,k-1}^{q}$ such that

$$\mathcal{I}_{\alpha,k-1}^{q} T(\boldsymbol{z}) = \frac{T(\boldsymbol{z})}{\alpha^{1/2}(\boldsymbol{z})|T_{k-1}(\boldsymbol{z})|^{q}}. \tag{6.21}$$

Instead of applying this operator on T, we apply it on an edge-enhanced version of T to enhance features in the image:

$$\mathcal{P}[T](\boldsymbol{z}) := \int_{\Omega_{\boldsymbol{z}}} \mathrm{e}^{-\mathrm{i}\boldsymbol{\xi}\cdot(\boldsymbol{z}-\boldsymbol{x})}(1 + |\boldsymbol{\xi}|)T(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi}, \tag{6.22}$$

$$\mathcal{P}[T](\boldsymbol{z}) := \int_{\Omega_{\boldsymbol{z}}} \mathrm{e}^{-\mathrm{i}\boldsymbol{\xi}\cdot(\boldsymbol{z}-\boldsymbol{x})}Q^{e}(\boldsymbol{z},\boldsymbol{\xi})d(s,t)\mathrm{d}s\mathrm{d}t\mathrm{d}\boldsymbol{x} \tag{6.23}$$

where

$$|\nabla T(\boldsymbol{z})| \approx \int_{\Omega_{\boldsymbol{z}}} \mathrm{e}^{-\mathrm{i}\boldsymbol{\xi}\cdot(\boldsymbol{z}-\boldsymbol{x})}|\boldsymbol{\xi}|T(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi} \tag{6.24}$$

as given in [145] and

$$Q^{e}(\boldsymbol{z},\boldsymbol{\xi}) = (1 + |\boldsymbol{\xi}|)Q(\boldsymbol{z},\boldsymbol{\xi}) \tag{6.25}$$

acts as an unsharp masking filter.

By setting $p = 2 - 2q$ for (6.18)

$$\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z}) = \frac{\mathcal{P}[T](\boldsymbol{z})}{\alpha^{1/2}(\boldsymbol{z})|\mathcal{P}[T_{k-1}](\boldsymbol{z})|^{1-p/2}}, \tag{6.26}$$

$$|\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z})|_2^2 = \int \left| \frac{\mathcal{P}[T](\boldsymbol{z})|\mathcal{P}[T_{k-1}](\boldsymbol{z})|^{p/2}}{\alpha^{1/2}(\boldsymbol{z})\mathcal{P}[T_{k-1}](\boldsymbol{z})} \right|^2 \mathrm{d}\boldsymbol{z}, \tag{6.27}$$

$$|\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z})|_2^2 = \int \left| \frac{\mathcal{P}[T](\boldsymbol{z})}{\alpha^{1/2}(\boldsymbol{z})|\mathcal{P}[T_{k-1}](\boldsymbol{z})|} \right|^2 |\mathcal{P}[T_{k-1}](\boldsymbol{z})|^p \mathrm{d}\boldsymbol{z}. \tag{6.28}$$

For $|\mathcal{P}[T(\boldsymbol{z}) - \mathcal{P}[T_{k-1}](\boldsymbol{z})| < \epsilon$ where ϵ is a predetermined sufficiently small threshold, we assume that $\frac{|\mathcal{P}[T](\boldsymbol{z})|^2}{|\mathcal{P}[T_{k-1}](\boldsymbol{z})|} \approx 1$. Therefore,

$$|\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z})|_2^2 \approx \int \frac{|\mathcal{P}[T](\boldsymbol{z})|^p}{\alpha(\boldsymbol{z})} \mathrm{d}\boldsymbol{z} \quad \rightarrow \quad |\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z})|_2^2 \approx |\mathcal{P}[T]|^p_{p,\alpha}. \tag{6.29}$$

and the modified optimization problem is given by

$$\mathrm{P}_2 : \quad \min\{\lambda |\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z})|_2^2 + |g - \mathcal{F}[T]|^2_{2,\mathcal{R}_n}\} \tag{6.30}$$

This way we convert the optimization problem given by P_1 which includes an L^p-norm to P_2 which only includes an L^2-norm. It should be noted here that T in (6.30) can be effectively replaced with $\rho(T)$ and $\rho(\nabla T)$.

A pseudocode for SAR image reconstruction with limited data using IRtA is given in Algorithm 3.

Algorithm 3 A pseudocode for IRtA.

1: Initialize $f_0 \equiv 1$ on its support and $k = 1$ where $\rho(\mathcal{P}[T](\boldsymbol{z}))$ is the edge-enhanced FBP image formed from the noisy data.
2: **while** $|\mathcal{P}[T_k] - \mathcal{P}[T_{k-1}]|_2^2 > \epsilon$ **do**
3: Calculate $\mathcal{I}^q_{\alpha,k-1}T(\boldsymbol{z}) = \frac{\rho(\mathcal{P}[T](\boldsymbol{z}))}{|\rho(\mathcal{P}[T_{k-1}](\boldsymbol{z}))|^q}$ for $q = 1/2$.
4: Derive the edge-enhancing FBP filter $Q_k = \min_{Q_k} J(Q_k)$ using (3.44) and [145].
5: Find $\mathcal{P}[T_k] = \mathcal{B}\mathcal{Q}^e_k[g]$.
6: Update $\mathcal{I}^q_{\alpha,k-1} \rightarrow \mathcal{I}^q_{\alpha,k}$
7: $k = k + 1$
8: **return** $\tilde{T}_k$

6.5.2.2 Iterative Shrinkage-type Analytic Reconstruction

Iterative shrinkage-type algorithms (IStA) are another set of algorithms that are used to solve the optimization problem given in (6.4) [116]–[128]. We can consider IStA as gradient-based algorithms for solving (6.4) via iterative multiplications with the matrix $\mathbf{F}$ and $\mathbf{F}^\dagger$ [126]. In radar imaging, this corresponds to repeated forward and backprojections.

In the literature there are different derivations of IStA based on thresholding [126], majorization-minimization [119], [120] and expectation maximization algorithms [121]. In this section, we approach to solve the problem given in (6.4) by using the proximal point and surrogate functions [120], and the fixed-point strategy [123]. We define the following functional:

$$\mathcal{J}_s(T, T_0) = \frac{c}{2}\left(|T - T_0|_2^2 - |\mathcal{B}^\dagger T - \mathcal{B}^\dagger T_0|_2^2 - |\mathcal{B}^\dagger(\mathcal{A} - \mathcal{I})T|_2^2\right) \tag{6.31}$$

to solve (6.4) where $\mathcal{B}$ is the backprojection operator (without any filtering), $\mathcal{B}^\dagger$ is its adjoint, $\mathcal{A}$ is another operator whose kernel is the amplitude function A in (2.15) and $c > 0$ is a constant chosen such that $\mathcal{J}_s(T, T_0)$ is strictly convex.

After adding $\mathcal{J}_s$ to the objective function given in (6.16) and making some simplifications, we form a new objective function:

$$\tilde{\mathcal{J}}_{sh}(T) = \mathcal{J}_{re}(T) + \mathcal{J}_s(T, T_0) = \langle \mathcal{F}^\dagger d - \mathcal{B}\mathcal{B}^\dagger T_0 + cT_0, f\rangle + |T|_2^2 + \lambda \int \rho(T)\mathrm{d}\boldsymbol{z} + c_1 \tag{6.32}$$

where $\mathcal{F}$ is the forward operator given in (2.15) and $\mathcal{F}^\dagger$ is its adjoint.

Dividing (6.32) by c and adding $|h_0|^2$ where

$$h_0 = \frac{1}{c}(\mathcal{F}^\dagger d - \mathcal{B}\mathcal{B}^\dagger T_0) + T_0 \tag{6.33}$$

we obtain

$$\mathcal{J}_{sh}(T) = \frac{1}{2}|T - h_0|^2 + \frac{\lambda}{c}\int \rho(T(\boldsymbol{z}))\mathrm{d}\boldsymbol{z} + c_2 \tag{6.34}$$

$$= \int \left\{ \frac{1}{2}|T(\boldsymbol{z}) - h_0(\boldsymbol{z})|^2 + \frac{\lambda}{c}\rho(T(\boldsymbol{z})) \right\} \mathrm{d}\boldsymbol{z} + c_2. \tag{6.35}$$

Similar to the derivation of the optimal filter (3.44), a T_{opt} for which (6.34) is minimized needs to be derived. To do that, the variation of $\mathcal{J}_{sh}$ is calculated with respect to T by fixing some T_ϵ and considering the variation, $T + \epsilon T_\epsilon$ for some small ϵ:

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{sh}(T + \epsilon T_\epsilon)\Big|_{\epsilon=0} = 0. \tag{6.36}$$

Since the integrand in (6.34) is non-negative, $\mathcal{J}_{sh}$ is minimized whenever the integrand is minimum. Therefore,

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\left\{ \frac{1}{2}|T + \epsilon T_\epsilon - h_0|^2 + \frac{\lambda}{c}\rho(T + \epsilon f_\epsilon) \right\}\Big|_{\epsilon=0} \tag{6.37}$$

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\left\{ \frac{1}{2}|T - h_0|^2 + \frac{1}{2}\langle T - h_0, \epsilon T_\epsilon\rangle + \frac{1}{2}|\epsilon T_\epsilon|^2 + \frac{\lambda}{c}\rho(T + \epsilon T_\epsilon) \right\}\Big|_{\epsilon=0} \tag{6.38}$$

$$0 = \left\{ \frac{1}{2}\langle T - h_0, T_\epsilon\rangle + \epsilon T_\epsilon + \frac{\lambda}{c}\langle \rho'(T + \epsilon T_\epsilon), T_\epsilon\rangle \right\}\Big|_{\epsilon=0} \tag{6.39}$$

$$0 = \frac{1}{2}(T - h_0) + \frac{\lambda}{c}\rho'(T) \tag{6.40}$$

$$h_0 = T_{opt} + \frac{2\lambda}{c}\rho'(T_{opt}). \tag{6.41}$$

where ρ' stands for the gradient of the sparsity inducing potential function ρ. Possible choices for ρ are given in Table 6.1. A pseudocode for SAR image reconstruction with IStA is given in Algorithm 4.

Algorithm 4 A pseudocode for IStA.

Initialize $T_0 \equiv 0$, $\tilde{d} = \mathcal{A}^\dagger d$ on its support.
while $|\mathcal{P}[T_k] - \mathcal{P}[T_{k-1}]|_2^2 > \epsilon$ **do**
    Calculate the residual: $r_k = \tilde{d} - \mathcal{B}^\dagger \mathcal{P}[T_k]$
    Backproject the residual and calculate h_k: $h_k = \mathcal{B} r_k + \mathcal{P}[T_k]$
    Solve T_{k+1} in $h_k = T_{k+1} + \frac{2\lambda}{c}\rho'(\mathcal{P}[T_{k+1}])$.
    Update the residual: $r_{k+1} = \tilde{d} - \mathcal{B}^\dagger \mathcal{P}[T_{k+1}])$
    $k = k + 1$
return T_{k+1}

6.6 Numerical Simulations

We consider two data sets to demonstrate the performance of the proposed methods: A data set formed by using MATLAB and the Civilian Vehicle (CV) Dome data set provided by AFRL [42]. For the first dataset, we perform two sets of experiments:

1. **Undersampled System:** For a 22km×22km scene, we descritize the scene into 128×128 pixels. We place a 5.16km×5.16km (30 pixels-by-30 pixels) square at the scene center. A mono-static antenna traverses a circular trajectory given by $\boldsymbol{\gamma}(s) = [11\cos s, 11\sin s, 6.5]\,\text{km}$. The trajectory is uniformly sampled for $s \in [0, 2\pi]$ at 32 points. This is a much sparser sampling than required. We sample the fast-time at 384 points. This corresponds to:

   - $384 \times 32 - by - 1$ vector to represent the measurement data,
   - $384 \times 32 - by - 128 \times 128$ matrix to represent the forward operator and
   - $128 \times 128 - by - 1$ vector to represent the imaging scene.

   Since 384×32 is less than 128×128, the inverse problem is underdetermined.

2. **Regularly Sampled System:** For a 22km×22km scene, we descritize the scene into 64×64 pixels. We place a 7.9km×7.9km (23 pixels-by-23 pixels) square at the scene center. A mono-static antenna traverses a circular trajectory given by $\boldsymbol{\gamma}(s) = [11\cos s, 11\sin s, 6.5]\,\text{km}$. The trajectory is uniformly sampled for $s \in [0, 2\pi]$ at 128 points. The fast-time is sampled with 543 points. This corresponds to

   - $543 \times 128 - by - 1$ vector to represent the measurement data,
   - $543 \times 128 - by - 64 \times 64$ matrix to represent the forward operator and
   - $64 \times 64 - by - 1$ vector to represent the imaging scene.

6.6.1 Numerical Simulations for IRtA with Synthetic Data

In these simulations, $A(\boldsymbol{x}, \omega, s)$ is set to 1, which corresponds to the assumption of an isotropic antenna radiating a Dirac-delta-like impulse and the compensation

of the geometric spreading in the measured data. We add Gaussian noise to the projection data to simulate the thermal noise at an SNR level of 0dB. (See (3.54) for a definition of SNR.) Then, we reconstruct the images with the limited data using the Algorithm 3. We use the (3.44) with a known spectral density function for the target. We use ten different potential functions given in Table 6.1 and evaluate the performance of the algorithm for each potential function by considering the following figures of merit:

1. Mean-square-error (MSE) between original target scene and the reconstructed target image calculated by

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^{N} \sum_{r=1}^{M} \left(T(\boldsymbol{x}_i) - \tilde{T}_r(\boldsymbol{x}_i) \right)^2, \tag{6.42}$$

   where M is the number of realizations of the experiment, N is the number of imaging grids $\hat{T}_r(\boldsymbol{x})$ is the reconstructed image in the r^{th} realization.

2. The relative L^2-norm difference, which we will refer to as the Δ-norm, between the images that are reconstructed in consecutive iterations, calculated by

$$\Delta\text{-norm} = \frac{1}{MN \sum_{i=1}^{N} \sum_{r=1}^{M} \left(\tilde{T}_{r-1}(\boldsymbol{x}_i) \right)^2} \sum_{i=1}^{N} \sum_{r=1}^{M} \left(\tilde{T}_r(\boldsymbol{x}_i) - \tilde{T}_{r-1}(\boldsymbol{x}_i) \right)^2. \tag{6.43}$$

3. Correlation of the range profiles of the reconstructed images with the range profile of the true target scene (See Figure 6.3) calculated via Pearson's correlation coefficient [146]:

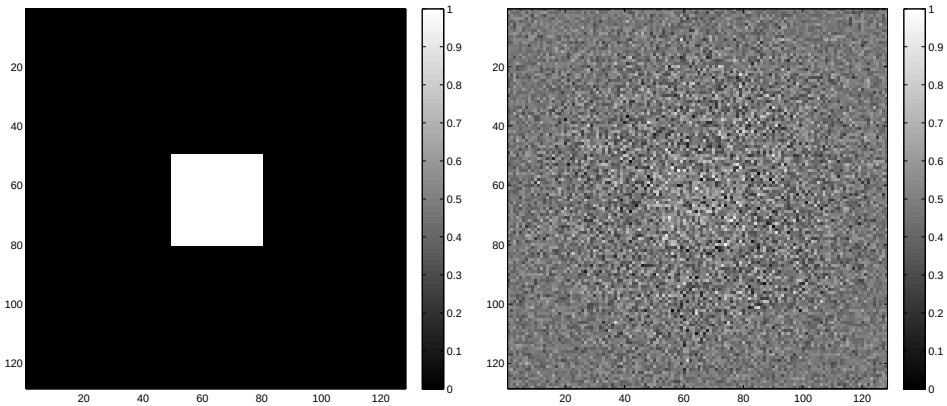
$$\text{CORR} = \sum_{r=1}^{M} \frac{\text{cov}(\tilde{T}_r^{rp}, T^{rp})}{M \sigma_{\tilde{T}_r^{rp}}, \sigma_{T^{rp}}} \tag{6.44}$$

   where $\tilde{T}_r^{rp}$, T^{rp} denote the range profiles of the reconstructed image at the r^{th} realization and the true target scene, respectively. "cov" denotes the covariance and σ denotes the standard deviation.

Furthermore, Δ-norm is used to decide when the algorithm converges [50].

6.6.1.1 Numerical Simulations for the Under-Sampled Data

As a first set of experiments, we apply IRtA on the under-sampled data. Figure 6.1(a) shows the original target scene used for these simulations and Figure 6.1(b) shows the FBP image when the SNR is 0dB. We perform these simulations ten times for each potential function. It should be noted here that the algorithm converges at different iterations for each potential function. Therefore, algorithm is run thirty times for each objective function. Figure 6.2(a) shows the reconstructed image using the deterministic FBP, which is the initialization to the algorithms. Figures 6.2(b)-(f) show the reconstructed images using IRtA after the first, second, third, tenth and the twentieth iterations, respectively. The numerical values for the convergence of the algorithm for each potential function are reported in detail in Subsection 6.7.



(a) (b)

Figure 6.1: (a) The original and (b) the reconstructed scene using the deterministic FBP with a single realization. SNR is set to 0dB and the Gaussian noise is added to the data to simulate the measurement noise.

When the SNR level is 0dB, i.e. when the noise power is comparable to the signal power, the image reconstructed with FBP almost has no information about the target. The target scene is vaguely visible and the background is noisy. However, right after the first iteration (see Figure 6.2(b)), the target is visible even though the initialization to the algorithm is the FBP image given in Figure 6.2(a) for all possible potential functions. After several iterations, noise is suppressed around the target and the features are more visible. Moreover, the noise is suppressed more in

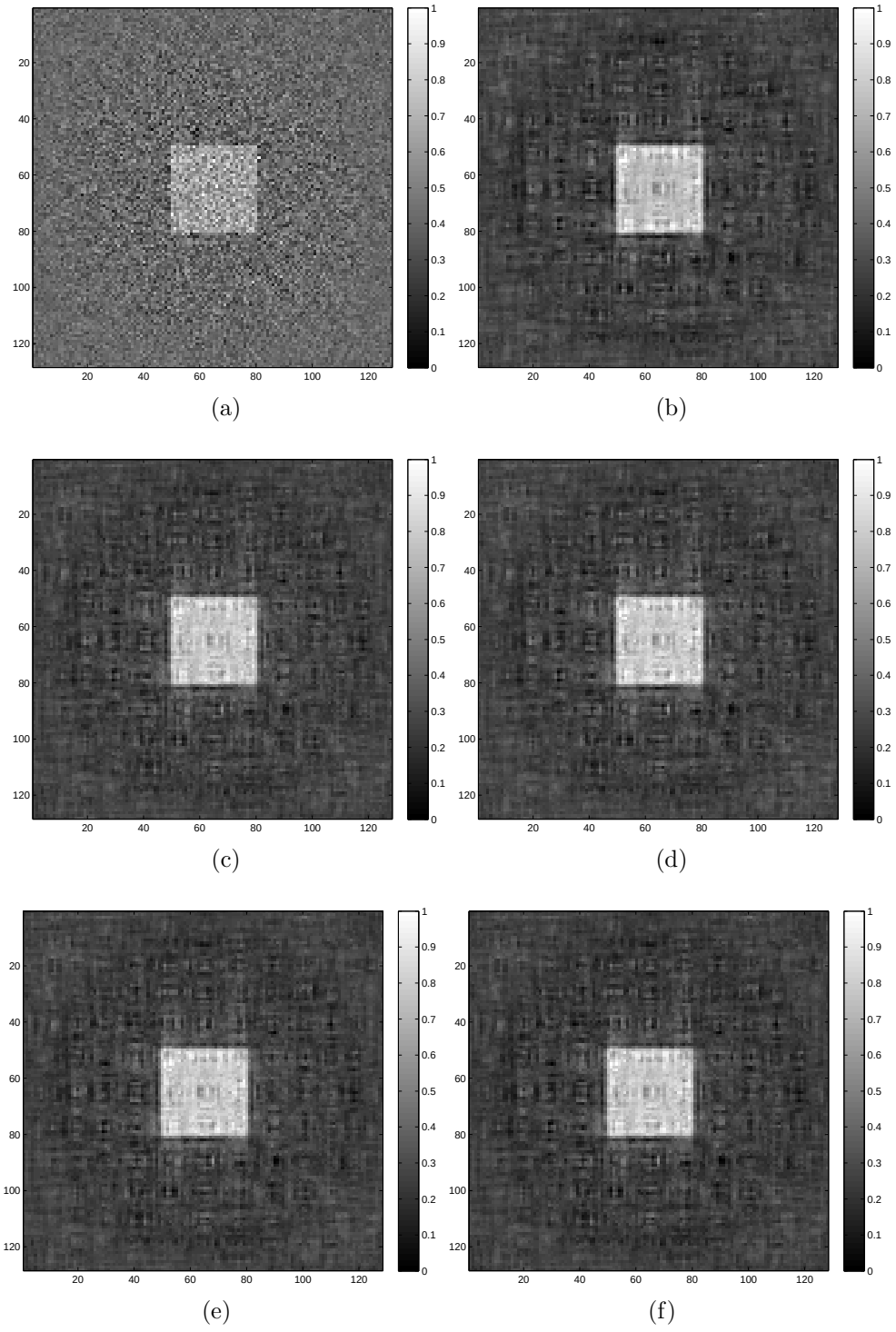


Figure 6.2: The reconstructed images (a) using the deterministic FBP (initialization of the algorithm); and IRtA after (b) the first, (c) second, (d) third, (e) tenth, and (f) the twentieth iterations when the SNR is 0dB and $\rho(f) = \frac{f^2}{1+f^2}$ for the target in Figure 6.1.

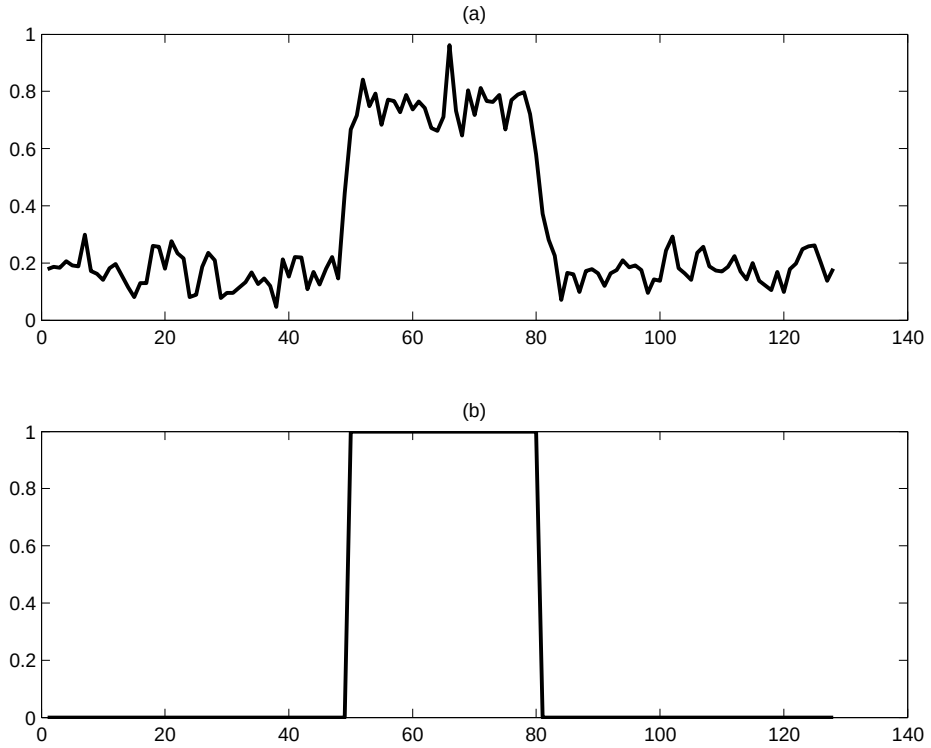


Figure 6.3: The range profiles of the (a) reconstructed image at the tenth iteration, and (b) the true target scene in Figure 6.1(a).

the background.

Figure 6.4 shows the plots for the MSE, Δ-norm and correlation for the first potential function $\rho(f) = \frac{f^2}{f^2+1}$. For the other objective functions, the patterns for the plots are similar therefore they are not shown here. Figures 6.5 (a)-(c) show the values of these metrics for all potential functions at every iteration as an image grid. It is seen from the image given in Figure 6.5(b) that the Δ-norm, which is the relative L^2-norm difference between images that are reconstructed in consecutive iterations, drops sharply after the first iteration and remains almost constant after the second iteration. Similarly, we see in Figure 6.5(a) that the MSE between the true and the reconstructed target images decreases sharply after the first iteration and then remains constant. Finally, we show the correlation of the range profiles in Figure 6.5 and observe that it increases right after the first iteration, then remains almost constant for all of the potential functions.

Considering the results with different potential functions, potential function

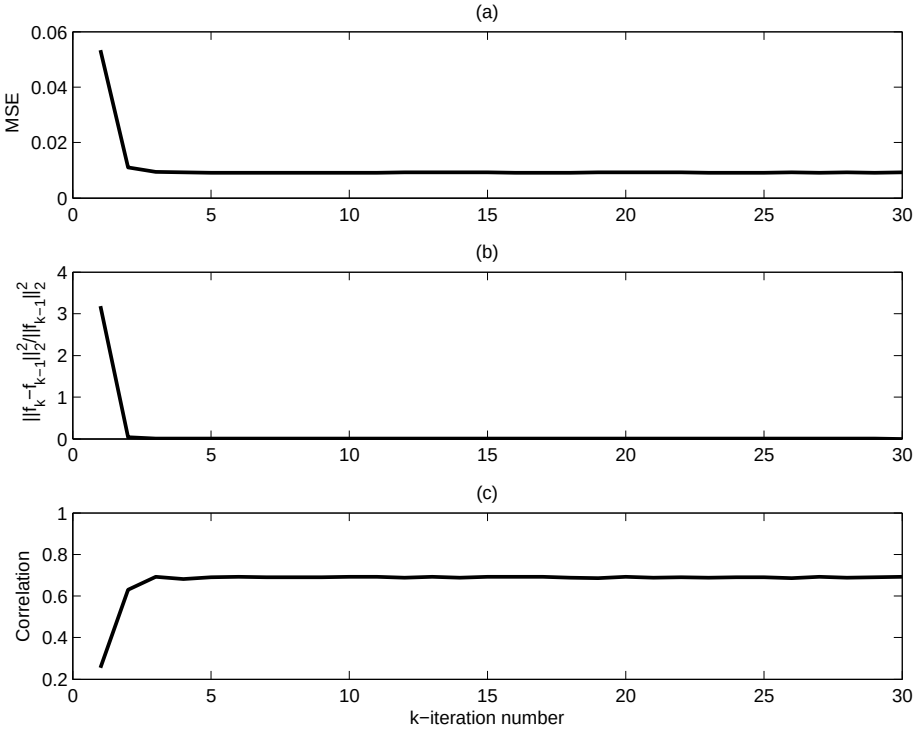


Figure 6.4: The plots showing (a) the MSE between the original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles when $\rho(f) = \frac{f^2}{1+f^2}$ for IRtA and the scene in Figure 6.1(a).

$\rho(f) = \min(f^2, 2f-1)$ results in the reconstructed images with the minimum MSE, Δ-norm and the maximum contrast while potential function $\frac{f^p}{1+f^{2-p}}$ gives the results with the maximum MSE, Δ-norm and the minimum contrast. Tables 6.2, 6.3, 6.4 show the values for MSE, Δ-norm, correlation respectively for the first five iterations and each potential function. The rapid decrease in MSE, Δ-norm and the increase in correlation is clearly seen in these tables.

6.6.1.2 Numerical Simulations for the Regularly Sampled Data

As a second set of experiments, we apply IRtA on regularly sampled data. We perform these simulations ten times.

The algorithm converges at different iterations for each potential function. Therefore, the algorithm is run thirty times for each potential function. Figure

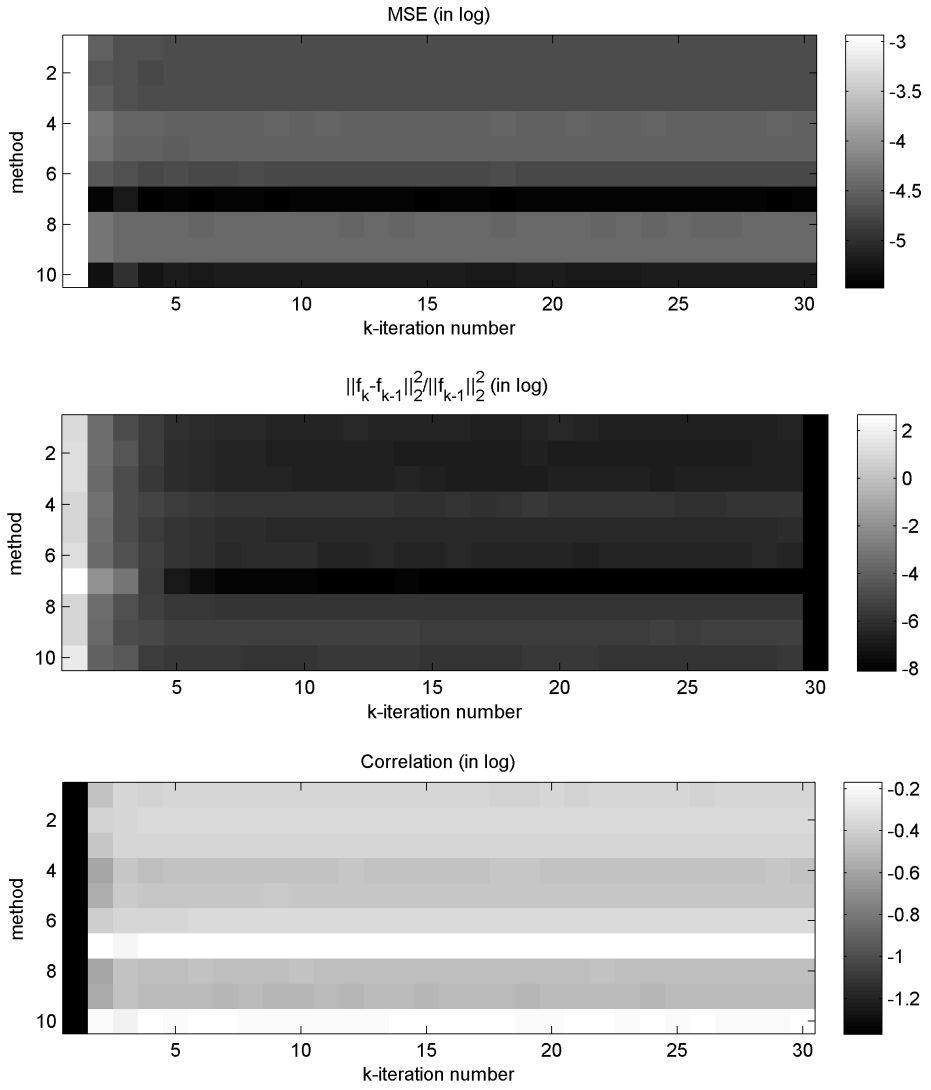


Figure 6.5: Images showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles for the target scene in Figure 6.1(a) with all potential functions in the order of $\frac{f^2}{1+f^2}$, f^2, $\log(1+f^2)$, $\frac{f}{1+f}$, f, $\log$ $\mathbf{cosh}(T)$, $\min(f^2, 2f-1)$, $f^{1/2}$, $\frac{f^{1/2}}{1+f^{2-1/2}}$ for IRtA.

Table 6.2: The MSE between the true and the reconstructed target image for the first five iterations using IRtA for the target scene in Figure 6.1(a). The minimum values are shown in bold fonts.

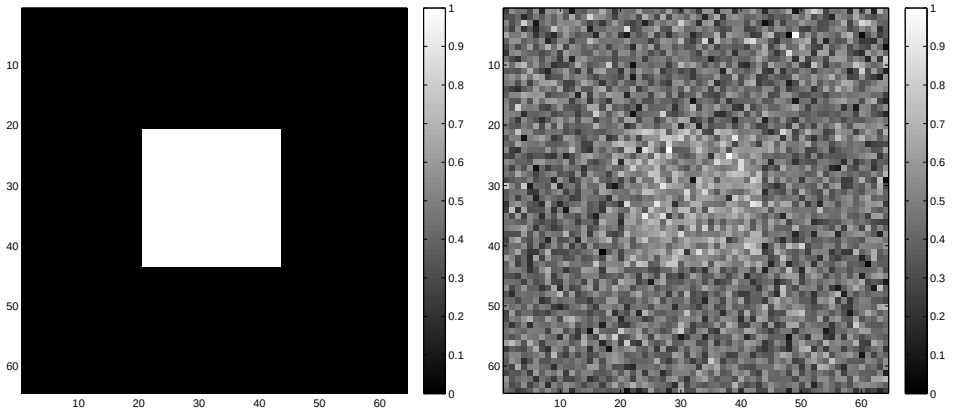
Potential Function	Iteration				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	5.331e-02	1.104e-02	9.486e-03	9.312e-03	9.122e-03
f^2	5.331e-02	9.996e-03	9.632e-03	8.892e-03	8.984e-03
$\log(f^2+1)$	5.331e-02	1.048e-02	9.509e-03	9.091e-03	9.015e-03
$\frac{f}{f+1}$	5.331e-02	1.373e-02	1.136e-02	1.138e-02	1.127e-02
f	5.331e-02	1.281e-02	1.111e-02	1.096e-02	1.084e-02
$\log(\cosh(f))$	5.331e-02	1.022e-02	9.420e-03	8.877e-03	8.967e-03
$\min(f^2, 2f-1)$	**5.331e-02**	4.495e-03	**5.358e-03**	**4.184e-03**	**4.442e-03**
f^p	5.331e-02	1.366e-02	1.203e-02	1.178e-02	1.182e-02
$\frac{f^p}{1+f^{2-p}}$	5.331e-02	1.332e-02	1.202e-02	1.214e-02	1.189e-02
∇f	5.331e-02	**4.961e-03**	6.886e-03	5.288e-03	5.715e-03

Table 6.3: The relative L^2-norm difference between the images reconstructed at every iteration for the first five iterations using IRtA for the target scene in Figure 6.1(a). The minimum values are shown in bold fonts.

Potential Function	Iteration				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	3.183	3.363e-02	7.702e-03	4.151e-03	2.550e-03
f^2	3.544	3.087e-02	1.044e-02	4.382e-03	2.281e-03
$\log(f^2+1)$	3.360	2.812e-02	**7.530e-03**	**3.659e-03**	2.215e-03
$\frac{f}{f+1}$	2.461	3.777e-02	8.400e-03	5.392e-03	4.241e-03
f	2.659	3.159e-02	8.015e-03	4.189e-03	2.812e-03
$\log(\cosh(f))$	3.451	2.743e-02	9.949e-03	5.036e-03	2.894e-03
$\min(f^2, 2f-1)$	1.454e+01	1.460e-01	4.344e-02	4.473e-03	**9.142e-04**
f^p	**2.454**	2.919e-02	8.875e-03	5.232e-03	3.617e-03
$\frac{f^p}{1+f^{2-p}}$	2.531	2.579e-02	8.626e-03	6.501e-03	4.966e-03
∇f	6.172	**1.857e-02**	1.243e-02	3.893e-03	3.680e-03

Table 6.4: The correlation of the true and the reconstructed target image range profiles for the first five iterations, using IRtA for the target scene in Figure 6.1(a). The maximum values are shown in bold fonts.

Potential Function	**Iteration**				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	2.542e-01	6.295e-01	6.923e-01	6.824e-01	6.902e-01
f^2	2.542e-01	6.755e-01	6.975e-01	7.015e-01	7.007e-01
$\log(f^2+1)$	2.542e-01	6.472e-01	6.955e-01	6.882e-01	6.947e-01
$\frac{f}{f+1}$	2.542e-01	5.542e-01	6.421e-01	6.251e-01	6.302e-01
f	2.542e-01	5.762e-01	6.505e-01	6.425e-01	6.471e-01
$\log(\cosh(f))$	2.542e-01	6.644e-01	6.984e-01	6.944e-01	6.976e-01
$\min(f^2, 2f-1)$	**2.542e-01**	**8.410e-01**	**8.045e-01**	**8.445e-01**	**8.316e-01**
f^p	2.542e-01	5.498e-01	6.350e-01	6.194e-01	6.202e-01
$\frac{f^p}{1+f^{2-p}}$	2.542e-01	5.599e-01	6.307e-01	6.042e-01	6.086e-01
∇f	2.542e-01	8.237e-01	7.911e-01	8.331e-01	8.240e-01



(a) (b)

Figure 6.6: (a) The original and (b) the reconstructed scene using the deterministic FBP with a single realization. SNR is set to 0dB and Gaussian noise is added to data to simulate the additive thermal noise.

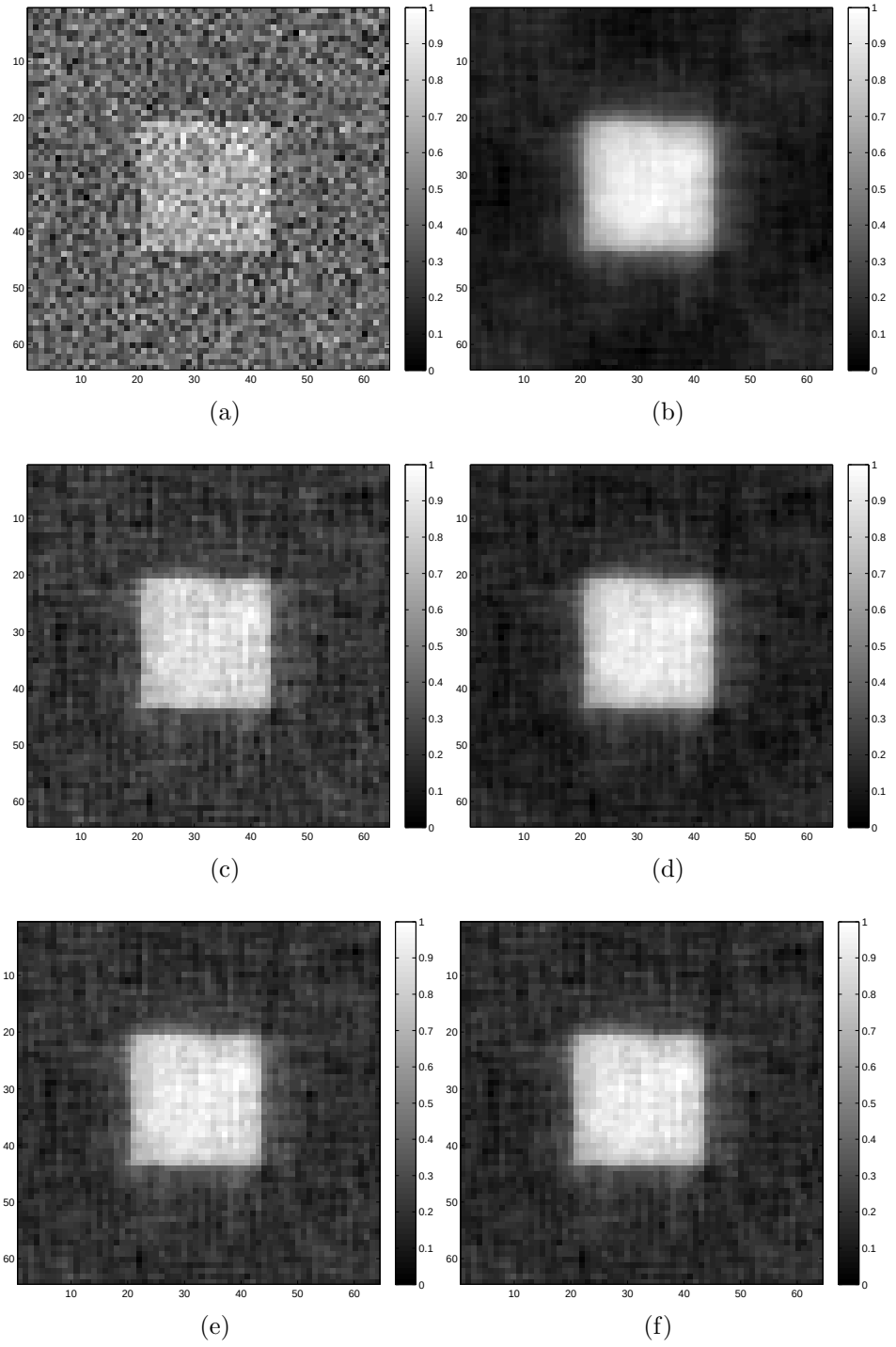


Figure 6.7: The reconstructed images (a) using the deterministic FBP (initialization of the algorithm); and IRtA (b) after the first, (c) second, (d) third, (e) tenth, and (f) twentieth iterations when the SNR is 0dB and $\rho(f) = \frac{f^2}{1+f^2}$ for the target in Figure 6.6.

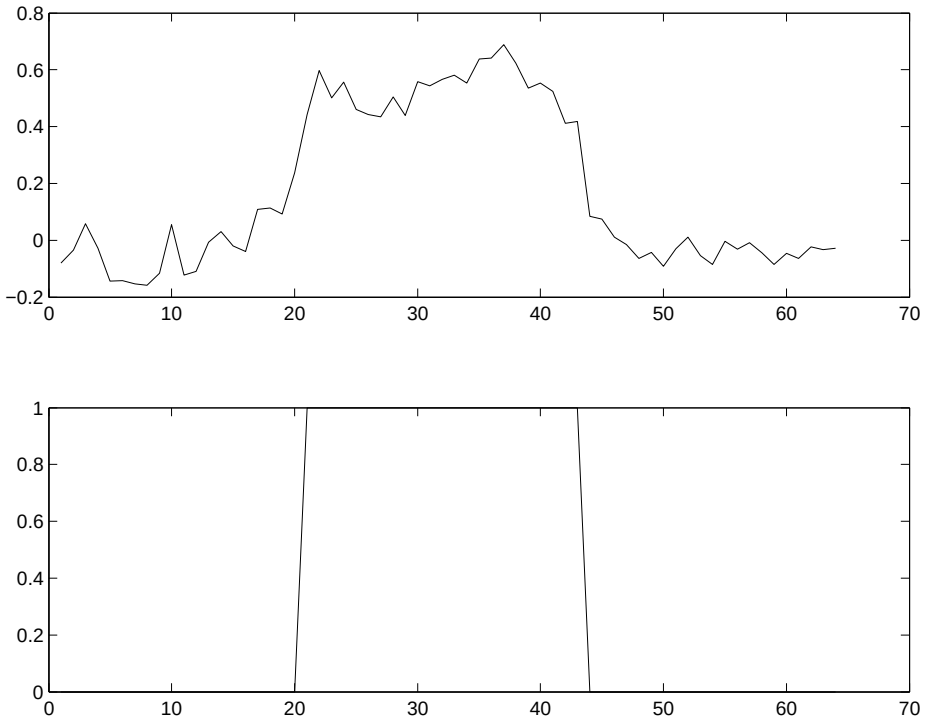


Figure 6.8: The range profiles of the (a) reconstructed image at the tenth iteration, and (b) the true target scene in Figure 6.6(a).

6.7(a) shows the reconstructed image using the deterministic FBP, which is the initialization to the algorithms. Figures 6.7(b)-(f) show the reconstructed images using IRtA after the first, second, third, tenth and the twentieth iterations, respectively when the target scene is Figure 6.6(a) for the potential function $\rho(f) = \frac{f^2}{f^2+1}$.

When the SNR level is 0dB, again the image reconstructed with FBP almost has no information about the target. The target scene is not visible and noisy. After the first iteration (see Figure 6.7(b)), the target is visible even though the initialization to the algorithm is the FBP image given in Figure 6.7(a). After several iterations, noise is suppressed around the target and the features are more visible. Moreover, the noise is suppressed more in the background.

In assessing the quality of the reconstructed images, we consider three figures of merit: MSE, Δ-norm and the correlation of the range profiles with a rectangular window defined earlier.

Figure 6.9 shows the plots for the MSE, Δ-norm and correlation for the first

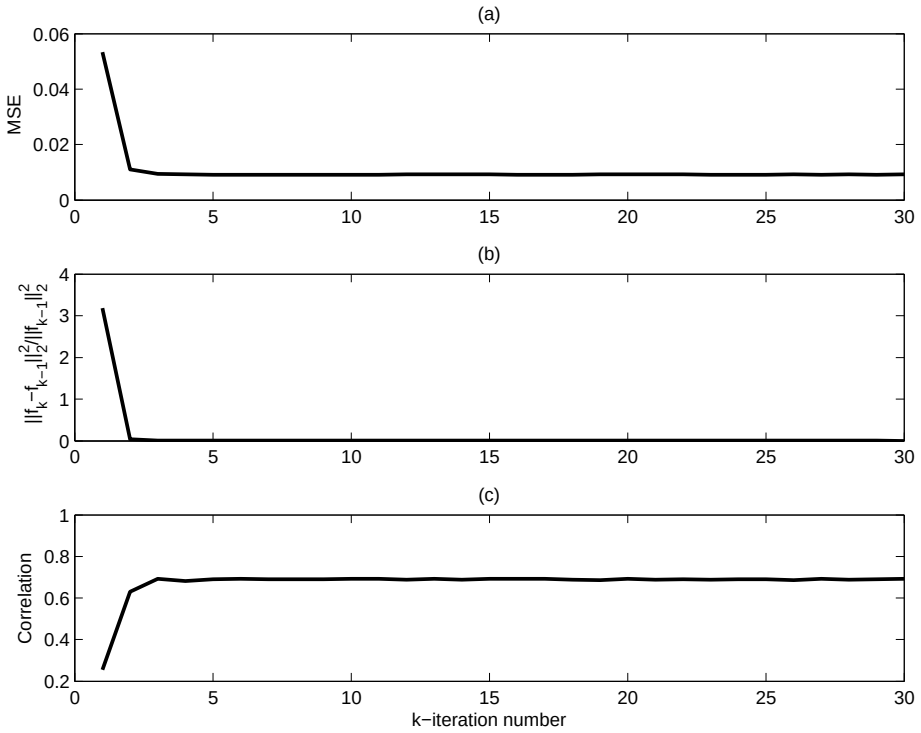


Figure 6.9: The plots showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles when $\rho(f) = \frac{f^2}{1+f^2}$ for IRtA and the scene in Figure 6.6(a).

potential function $\rho(f) = \frac{f^2}{f^2+1}$. For the other objective functions, the patterns for the plots are similar therefore they are not shown here. Figures 6.10 (a)-(c) show the values of these metrics for all potential functions at every iteration as an image grid. We observe from Figure 6.10(b) that the value of the Δ-norm, drops sharply after the first iteration and remains almost constant after the second iteration. Similarly, we see in Figure 6.10(a) that the MSE between the original target scene and the reconstructed image decreases sharply after the first iteration and remains almost constant for higher iteration numbers. Finally, we show the correlation of the range profiles in Figure 6.10. We observe that the correlation increases right after the first iteration, then remains almost constant for higher number of iterations for all of the potential functions.

Comparing the results for different potential functions, unlike the simulations

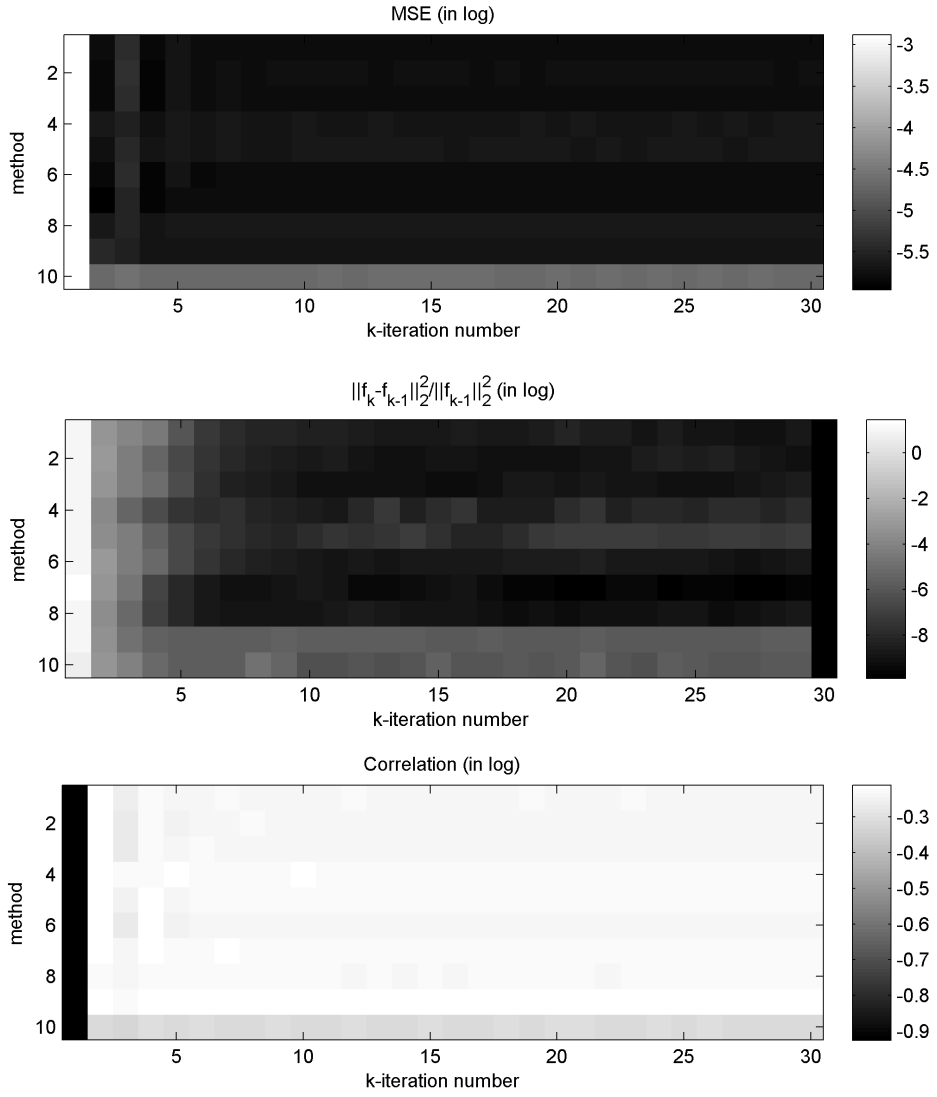


Figure 6.10: Images showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles for the target scene in Figure 6.6(a) with all potential functions in the order of $\frac{f^2}{1+f^2}$, f^2, $\log(1+f^2)$, $\frac{f}{1+f}$, f, $\log$ cosh(T), $\min(f^2, 2f-1)$, $f^{1/2}$, $\frac{f^{1/2}}{1+f^{2-1/2}}$ for IRtA.

performed in the previous subsection, we conclude that any one of the potential functions does not result in better reconstructed images than the rest for all metrics. Tables 6.5, 6.6, 6.7 show the values for MSE, Δ-norm, correlation, respectively for the first five iterations and each potential function. $\min(f^2, 2f-1)$ and $\frac{f^p}{1+f^{2-p}}$ perform slightly better considering the correlation values. The rapid decrease in the MSE, Δ-norm and the increase in the correlation are clearly seen in these results.

Table 6.5: The MSE between the true and the reconstructed target image for the first five iterations using IRtA for the target scene in Figure 6.6(a). The minimum values are shown in bold fonts.

Potential	**Iteration**				
Function	1	2	3	4	5
$\frac{f^2}{f^2+1}$	5.632e-02	3.021e-03	4.474e-03	2.909e-03	3.359e-03
f^2	5.632e-02	2.906e-03	4.656e-03	2.787e-03	3.429e-03
$\log(f^2+1)$	5.632e-02	2.969e-03	4.572e-03	2.832e-03	3.366e-03
$\frac{f}{f+1}$	5.632e-02	3.460e-03	3.854e-03	3.244e-03	3.521e-03
f	5.632e-02	3.206e-03	4.196e-03	3.415e-03	3.536e-03
$\log(\cosh(f))$	5.632e-02	2.955e-03	4.509e-03	**2.739e-03**	3.338e-03
$\min(f^2, 2f-1)$	**5.632e-02**	**2.580e-03**	3.988e-03	2.829e-03	**3.117e-03**
f^p	5.632e-02	3.595e-03	4.138e-03	3.361e-03	3.557e-03
$\frac{f^p}{1+f^{2-p}}$	5.632e-02	4.241e-03	**3.797e-03**	3.389e-03	3.402e-03
∇f	5.632e-02	9.468e-03	1.005e-02	9.062e-03	9.434e-03

Table 6.6: The relative L^2-norm difference between the images reconstructed at every iteration for the first five iterations using IRtA for the target scene in Figure 6.6(a). The minimum values are shown in bold fonts.

Potential Function	Iteration				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	2.913	4.035e-02	2.000e-02	1.180e-02	2.492e-03
f^2	2.917	4.396e-02	1.295e-02	4.340e-03	1.286e-03
$\log(f^2+1)$	2.908	4.122e-02	1.353e-02	6.509e-03	1.551e-03
$\frac{f}{f+1}$	2.855	**2.108e-02**	**4.640e-03**	1.497e-03	5.400e-04
f	2.873	2.705e-02	1.300e-02	4.265e-03	1.476e-03
$\log(\cosh(f))$	2.921	4.592e-02	1.323e-02	5.652e-03	1.414e-03
$\min(f^2, 2f-1)$	4.239	3.744e-02	9.151e-03	1.181e-03	3.339e-04
f^p	2.819	2.648e-02	5.627e-03	**9.382e-04**	**3.165e-04**
$\frac{f^p}{1+f^{2-p}}$	2.850	3.310e-02	8.373e-03	3.649e-03	3.136e-03
∇f	**1.805**	3.952e-02	1.536e-02	5.372e-03	3.037e-03

Table 6.7: The correlation of the true and the reconstructed target image range profiles for the first five iterations, using IRtA for the target scene in Figure 6.6(a). The maximum values are shown in bold fonts.

Potential Function	Iteration				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	3.964e-01	8.021e-01	7.691e-01	7.976e-01	7.872e-01
f^2	3.964e-01	8.026e-01	7.600e-01	8.002e-01	7.793e-01
$\log(f^2+1)$	3.964e-01	8.017e-01	7.640e-01	7.988e-01	7.850e-01
$\frac{f}{f+1}$	3.964e-01	8.014e-01	7.930e-01	7.982e-01	8.042e-01
f	3.964e-01	8.021e-01	7.803e-01	8.035e-01	7.904e-01
$\log(\cosh(f))$	3.964e-01	8.024e-01	7.605e-01	8.012e-01	7.826e-01
$\min(f^2, 2f-1)$	**3.964e-01**	**8.098e-01**	7.907e-01	8.032e-01	7.989e-01
f^p	3.964e-01	8.005e-01	7.846e-01	7.928e-01	7.921e-01
$\frac{f^p}{1+f^{2-p}}$	3.964e-01	8.045e-01	**7.935e-01**	**8.058e-01**	**8.064e-01**
∇f	3.964e-01	7.325e-01	7.210e-01	7.372e-01	7.286e-01

6.6.2 Numerical Simulations for IStA with Synthetic Data

6.6.2.1 Numerical Simulations for the Under-Sampled Data

For this set of simulations, we use the target in 6.1(a). We apply the Algorithm 4 to suppress the noise and to reconstruct images using limited data. We use 10 different potential functions given in Subsection 6.5.1 with IStA.

As noted in the previous subsection, the algorithm converges at different iterations for each objective function and the algorithms is run thirty times for each potential function. Figure 6.11(a) shows the reconstructed image using the deterministic FBP, which is the initialization to the algorithms. Figures 6.11(b)-(f) show the reconstructed images using IStA after the first, second, third, tenth and the twentieth iterations, respectively for the potential function $\rho(f) = \frac{f^2}{f^2+1}$. When the SNR level is 0dB the target is not visible in the image reconstructed with deterministic FBP. After the first iteration algorithm performs the following operation:

$$f_1 \approx T_0 + \mathcal{B}(d - \mathcal{B}^\dagger T_0) - \frac{2\lambda}{c}\rho'(T_0) \tag{6.45}$$

where f_0 corresponds to the FBP image. The image is blurred since the algorithm reconstructs images via backprojection without filtering. However, right after the second iteration the shape of the target is revealed, but somewhat smooth. As more iterations, are performed the level of smoothness decreases and the structure of the target becomes more visible.

Figure 6.12 shows the MSE, Δ-norm and correlation plots for the potential function $\rho(f) = \frac{f^2}{f^2+1}$. For the other objective functions, the patterns for the plots are similar, therefore they are not shown here. Figures 6.13 (a)-(c) show the values of these metrics for all potential functions at every iteration on an image grid. The plots in Figure 6.12(a) show that the relative L^2-norm difference between the images reconstructed in consecutive iterations drops sharply after the second iteration. However, the MSE and Δ-norm fluctuate. Note that the algorithm attempts to reduce the residual. However, when the residual is large, the algorithm attempts to smooth the image aggressively in the next iteration. After that, the algorithm attempts to compensate for the over-smoothing by using the original noisy radar

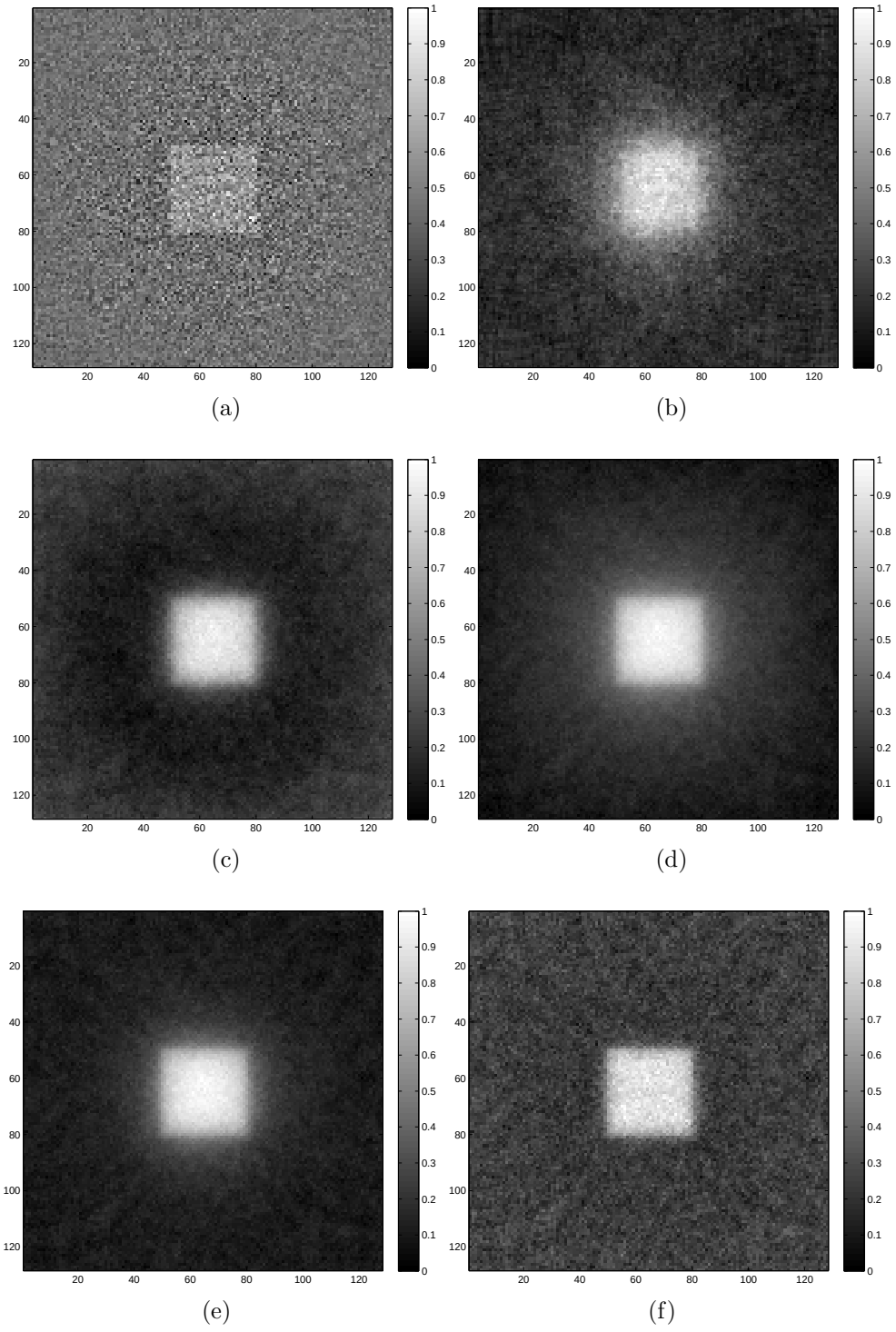


Figure 6.11: The reconstructed images (a) using the deterministic FBP (initialization of the algorithm); and IStA (b) after the first, (c) second, (d) third, (e) tenth, and (f) twentieth iterations when the SNR is 0dB and $\rho(f) = \frac{f^2}{1+f^2}$ for the target in Figure 6.1.

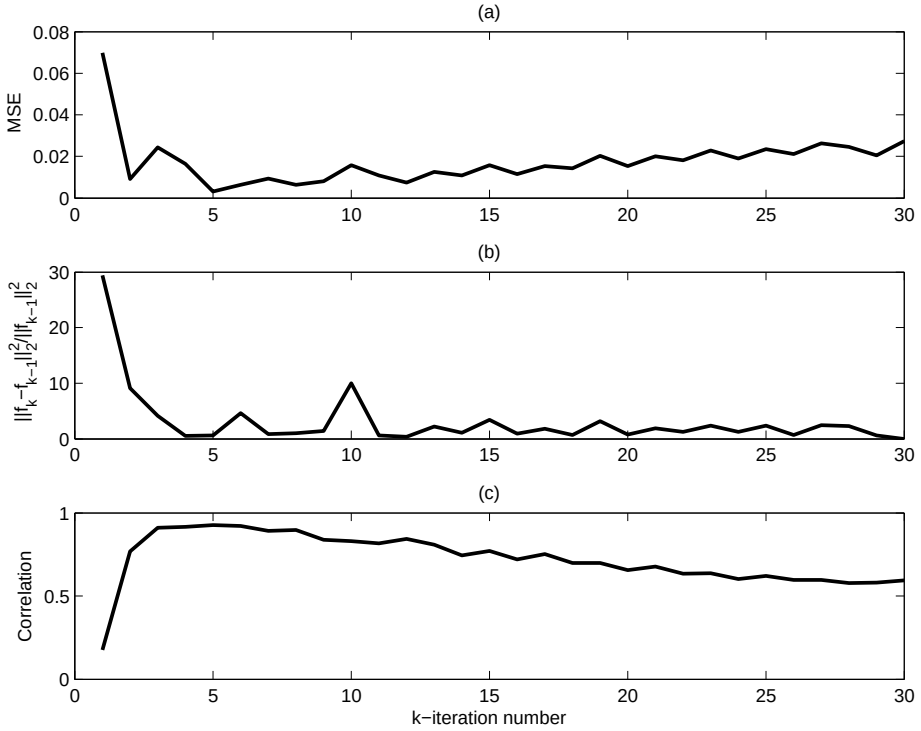


Figure 6.12: The plots showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles when $\rho(f) = \frac{f^2}{1+f^2}$ for IStA and the scene in Figure 6.1(a).

data. This increases the MSE as compared to that of the previous iteration. Hence, MSE and Δ-norm values fluctuate from one iteration to the next.

The algorithm is run for thirty iterations in order to compare its performance with that of the iterative reweighted-type reconstruction algorithm. The algorithm converges around the fifth iteration almost for all the objective functions. Additional iterations result in degradation in the image quality as well as the loss in the correlation.

Unlike IRtA, the type of the potential function used in the algorithm does not affect performance as can be seen in Figures 6.13. Nevertheless, potential function given by $\rho(f) = \min(f^2, 2f - 1)$, results in reconstructed images with the minimum MSE and the maximum contrast. Tables 6.8, 6.9, 6.10 show the values of the MSE, Δ-norm and the correlation, respectively for the first five iterations and for each

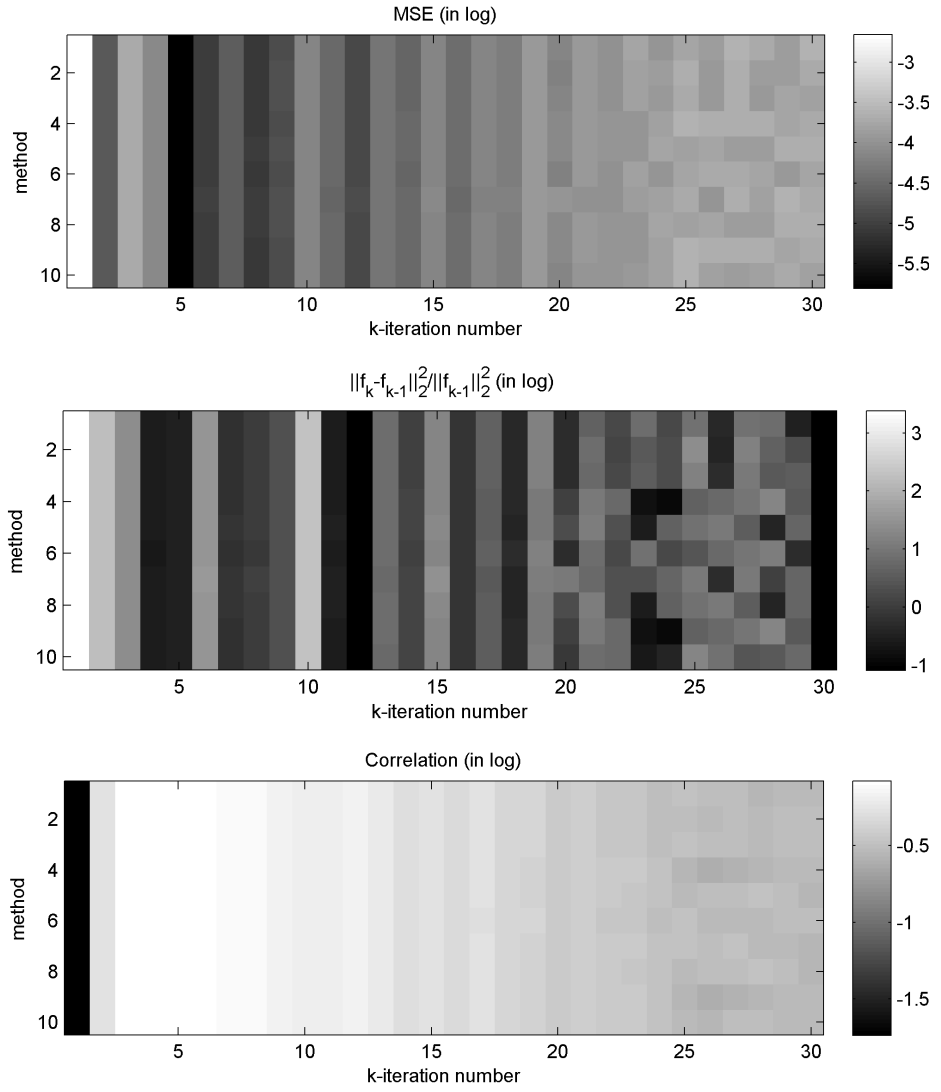


Figure 6.13: Images showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles for the target scene in Figure 6.1(a) with all potential functions in the order of $\frac{f^2}{1+f^2}$, f^2, $\log(1+f^2)$, $\frac{f}{1+f}$, f, $\log$ $\mathbf{cosh}(T)$, $\min(f^2, 2f-1)$, $f^{1/2}$, $\frac{f^{1/2}}{1+f^{2-1/2}}$ for IStA.

potential function.

Table 6.8: The MSE between the true and the reconstructed target image for the first five iterations using IStA for the target scene in Figure 6.1(a). The minimum values are shown in bold fonts.

Potential	**Iteration**				
Function	1	2	3	4	5
$\frac{f^2}{f^2+1}$	6.992e-02	9.129e-03	2.427e-02	1.638e-02	3.140e-03
f^2	6.992e-02	9.131e-03	2.431e-02	1.637e-02	3.134e-03
$\log(f^2+1)$	6.992e-02	9.130e-03	2.429e-02	1.637e-02	3.137e-03
$\frac{f}{f+1}$	6.992e-02	9.123e-03	2.431e-02	1.632e-02	3.112e-03
f	6.992e-02	9.123e-03	2.442e-02	1.626e-02	3.086e-03
$\log(\cosh(f))$	6.992e-02	9.130e-03	**2.422e-02**	1.641e-02	3.155e-03
$\min(f^2, 2f-1)$	**6.992e-02**	**9.115e-03**	2.469e-02	**1.608e-02**	**3.011e-03**
f^p	6.992e-02	9.123e-03	2.442e-02	1.626e-02	3.086e-03
$\frac{f^p}{1+f^{2-p}}$	6.992e-02	9.123e-03	2.431e-02	1.632e-02	3.112e-03
∇f	6.992e-02	9.126e-03	2.428e-02	1.635e-02	3.131e-03

6.6.2.2 Numerical Simulations for the Regularly Sampled Data

As a second set of experiments, we apply the IStA on regularly sampled data. Algorithm is run thirty times for each potential function. Figure 6.14(a) shows the reconstructed image using the deterministic FBP, which is the initialization to the algorithms. Figures 6.14(b)-(f) show the reconstructed images using IStA after the first, second, third, tenth and the twentieth iterations, respectively for the potential function $\rho(f) = \frac{f^2}{f^2+1}$. After the first iteration, the image is blurred since the algorithm reconstructs images via backprojection without filtering. However, right after the second iteration the shape of the target is revealed but somewhat smooth. More iterations reduce the over-smoothing and the high frequency content of the target becomes more visible.

Figure 6.15 shows the plots for the MSE, Δ-norm and the correlation for the potential function $\rho(f) = \frac{f^2}{f^2+1}$. For the other potential functions, the patterns for the plots are similar, therefore they are not shown here. Figures 6.16 (a)-(c) show the

Table 6.9: The relative L^2-norm difference between the images reconstructed at every iteration for the first five iterations using IStA for the target scene in Figure 6.1(a). The minimum values are shown in bold fonts.

Potential	**Iteration**				
Function	1	2	3	4	5
$\frac{f^2}{f^2+1}$	2.944e+01	9.113	4.098	5.516e-01	6.000e-01
f^2	2.942e+01	9.145	4.088	5.526e-01	5.999e-01
$\log(f^2+1)$	2.943e+01	9.127	4.094	5.520e-01	6.000e-01
$\frac{f}{f+1}$	2.938e+01	9.155	4.075	5.555e-01	6.031e-01
f	2.930e+01	9.247	4.041	5.598e-01	6.046e-01
$\log(\cosh(f))$	2.947e+01	**9.071**	4.115	**5.494e-01**	**5.988e-01**
$\min(f^2, 2f-1)$	**2.908e+01**	9.490	**3.947**	5.732e-01	6.085e-01
f^p	2.930e+01	9.247	4.041	5.598e-01	6.046e-01
$\frac{f^p}{1+f^{2-p}}$	2.938e+01	9.155	4.075	5.555e-01	6.031e-01
∇f	2.942e+01	9.119	4.090	5.526e-01	6.005e-01

Table 6.10: The correlation of the true and the reconstructed target image range profiles for the first five iterations, using IStA for the target scene in Figure 6.1(a). The maximum values are shown in bold fonts.

Potential	**Iteration**				
Function	1	2	3	4	5
$\frac{f^2}{f^2+1}$	1.760e-01	7.679e-01	9.099e-01	9.163e-01	9.258e-01
f^2	1.760e-01	7.679e-01	9.099e-01	9.164e-01	9.259e-01
$\log(f^2+1)$	1.760e-01	7.679e-01	9.099e-01	9.163e-01	9.259e-01
$\frac{f}{f+1}$	1.760e-01	7.680e-01	9.099e-01	9.163e-01	9.258e-01
f	1.760e-01	7.679e-01	9.098e-01	9.163e-01	9.258e-01
$\log(\cosh(f))$	1.760e-01	**7.680e-01**	**9.099e-01**	9.163e-01	9.258e-01
$\min(f^2, 2f-1)$	**1.760e-01**	7.678e-01	9.096e-01	**9.163e-01**	**9.258e-01**
f^p	1.760e-01	7.679e-01	9.098e-01	9.163e-01	9.258e-01
$\frac{f^p}{1+f^{2-p}}$	1.760e-01	7.680e-01	9.099e-01	9.163e-01	9.258e-01
∇f	1.760e-01	7.679e-01	9.101e-01	9.163e-01	9.258e-01

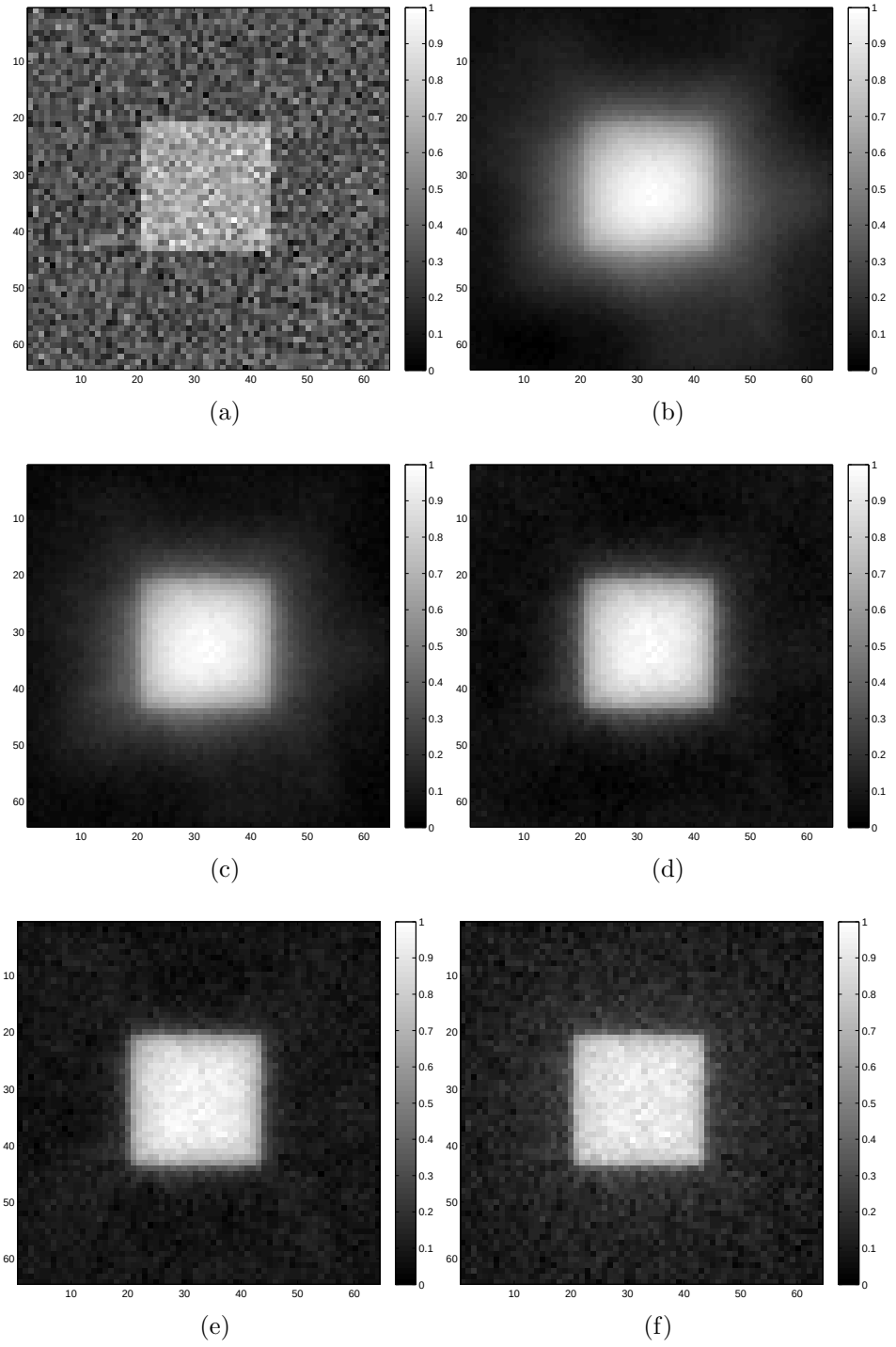


Figure 6.14: The reconstructed images (a) using the deterministic FBP (initialization of the algorithm); and IStA (b) after the first, (c) second, (d) third, (e) tenth, and (f) twentieth iterations when the SNR is 0dB and $\rho(f) = \frac{f^2}{1+f^2}$ for the target in Figure 6.6.

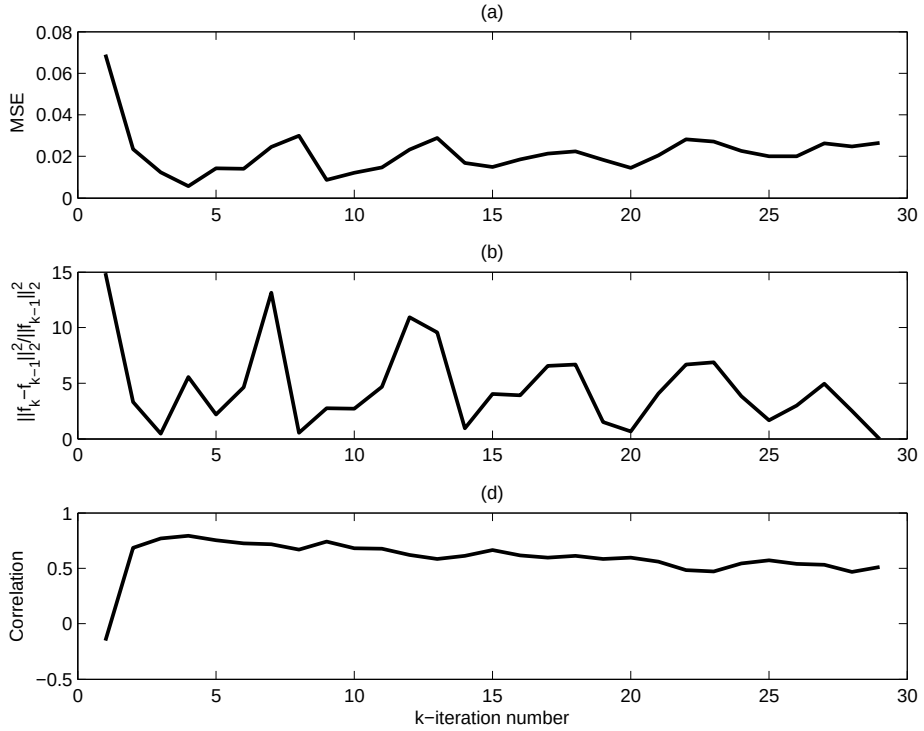


Figure 6.15: The plots showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles when $\rho(f) = \frac{f^2}{1+f^2}$ for IStA and the scene in Figure 6.6(a).

values of these metrics for all potential functions at every iteration on an image grid. We see from the plot in Figure 6.15(a) that the relative L^2-norm difference between the reconstructed images in consecutive iterations drops sharply after the second iteration. Similar to the performance of the IStA using under-sampled data, the MSE and Δ-norm fluctuate. The same explanation for IStA using under-sampled data applies here as well.

Like in Section 6.6.1.2, the nature of the potential function used in the algorithm does not significantly change the quality of the reconstructed images as can be seen in Figures 6.16. Nevertheless, the potential function $\rho(f) = \log(\cosh(f))$, results in the reconstructed images with the minimum MSE. Similar to IRtA with the regularly sampled data, iterative algorithm described here gives especially better results than regular FBP-type reconstruction when the data is limited. Tables 6.11,

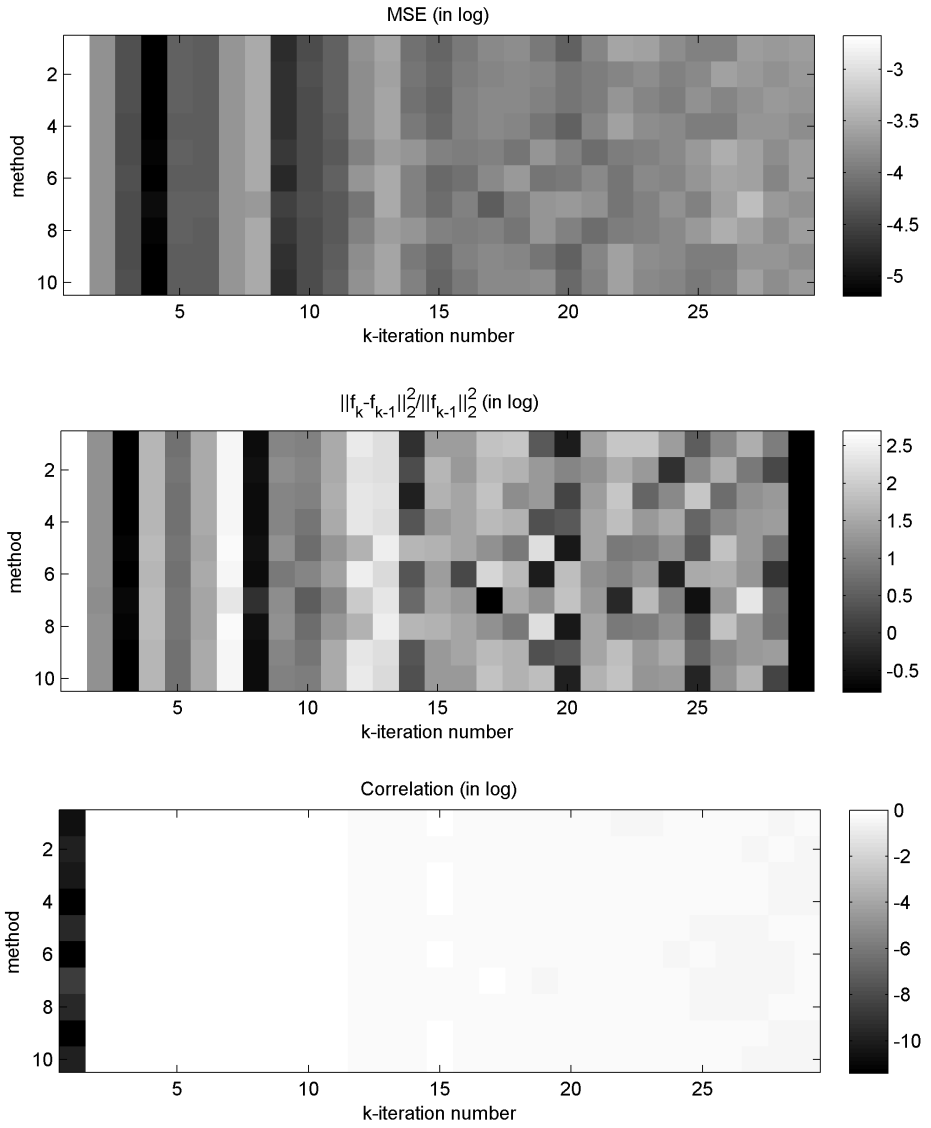


Figure 6.16: Images showing (a) the MSE between original target scene and the reconstructed target image for each iteration, (b) Δ-norm, (c) correlation of the range profiles for the target scene in Figure 6.6(a) with all potential functions in the order of $\frac{f^2}{1+f^2}$, f^2, $\log(1+f^2)$, $\frac{f}{1+f}$, f, $\log$ $\mathbf{cosh}(T)$, $\min(f^2, 2f-1)$, $f^{1/2}$, $\frac{f^{1/2}}{1+f^{2-1/2}}$ for IStA.

6.12, 6.13 show the values for the MSE, Δ-norm and the correlation, respectively for the first five iterations and each potential function.

Table 6.11: The MSE between the true and the reconstructed target image for the first five iterations using IStA for the target scene in Figure 6.6(a). The minimum values are shown in bold fonts.

Potential	**Iteration**				
Function	1	2	3	4	5
$\frac{f^2}{f^2+1}$	6.889e-02	2.350e-02	1.226e-02	5.636e-03	1.433e-02
f^2	6.890e-02	2.351e-02	1.225e-02	5.709e-03	1.442e-02
$\log(f^2+1)$	6.889e-02	2.350e-02	1.226e-02	5.667e-03	1.437e-02
$\frac{f}{f+1}$	6.890e-02	2.350e-02	1.216e-02	5.733e-03	1.427e-02
f	6.891e-02	2.351e-02	1.209e-02	5.922e-03	1.439e-02
$\log(\cosh(f))$	**6.889e-02**	**2.350e-02**	1.231e-02	**5.562e-03**	1.429e-02
$\min(f^2, 2f-1)$	6.894e-02	2.354e-02	**1.184e-02**	6.492e-03	1.457e-02
f^p	6.891e-02	2.351e-02	1.209e-02	5.922e-03	1.439e-02
$\frac{f^p}{1+f^{2-p}}$	6.890e-02	2.350e-02	1.216e-02	5.733e-03	1.427e-02
∇f	6.890e-02	2.349e-02	1.222e-02	5.660e-03	**1.425e-02**

Table 6.12: The relative L^2-norm difference between the images reconstructed at every iteration for the first five iterations using IStA for the target scene in Figure 6.6(a). The minimum values are shown in bold fonts.

Potential Function	**Iteration**				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	1.491e+01	3.304	4.656e-01	5.535	2.176
f^2	1.490e+01	3.305	4.704e-01	5.584	2.248
$\log(f^2+1)$	1.491e+01	3.304	4.677e-01	5.556	2.207
$\frac{f}{f+1}$	1.490e+01	3.276	4.709e-01	5.525	2.136
f	1.488e+01	3.257	4.826e-01	5.598	2.213
$\log(\cosh(f))$	1.492e+01	3.317	**4.608e-01**	5.510	2.154
$\min(f^2, 2f-1)$	**1.482e+01**	**3.184**	5.139e-01	5.706	2.277
f^p	1.488e+01	3.257	4.826e-01	5.598	2.213
$\frac{f^p}{1+f^{2-p}}$	1.490e+01	3.276	4.709e-01	5.525	2.136
∇f	1.490e+01	3.290	4.664e-01	**5.500**	**2.131**

Table 6.13: The correlation of the true and the reconstructed target image range profiles for the first five iterations, using IStA for the target scene in Figure 6.6(a). The maximum values are shown in bold fonts.

Potential Function	**Iteration**				
	1	2	3	4	5
$\frac{f^2}{f^2+1}$	-1.539e-01	6.847e-01	7.687e-01	7.927e-01	7.538e-01
f^2	-1.539e-01	6.846e-01	**7.700e-01**	**7.950e-01**	7.565e-01
$\log(f^2+1)$	-1.539e-01	6.847e-01	7.692e-01	7.937e-01	7.550e-01
$\frac{f}{f+1}$	-1.539e-01	6.848e-01	7.674e-01	7.906e-01	7.519e-01
f	-1.538e-01	6.847e-01	7.686e-01	7.928e-01	7.552e-01
$\log(\cosh(f))$	-1.539e-01	6.847e-01	7.686e-01	7.924e-01	7.531e-01
$\min(f^2, 2f-1)$	**-1.537e-01**	6.845e-01	7.696e-01	7.948e-01	**7.588e-01**
f^p	-1.538e-01	6.847e-01	7.686e-01	7.928e-01	7.552e-01
$\frac{f^p}{1+f^{2-p}}$	-1.539e-01	**6.848e-01**	7.674e-01	7.906e-01	7.519e-01
∇f	-1.539e-01	6.847e-01	7.676e-01	7.911e-01	7.522e-01

6.6.3 Numerical Simulations for L^p-norm Regularization with CV Dome Data Set

We perform a second set of simulations to demonstrate the performance of the methods using the CV dome data set [42]. This data set was generated by AFRL. The SAR data were simulated for various civilian vehicles by using high-fidelity electromagnetic wave simulation codes. The cars were modeled by using realistic CAD models and the electromagnetic properties of the different parts of the cars (glass, metallic, plastic etc.). The radar data provided includes 5760 pulses with 512 frequency bins. Note that this dataset does not include measurement noise. Therefore, Gaussian noise is added to the data at a 20dB SNR. To show that the algorithms work well with limited data, we use only the 10% of the data. Therefore, the size of the data is 576×512. The results are provided in Figures 6.18(a)-(d) using IRtA, 6.20(a)-(d) and 6.21(a)-(b) using IStA.

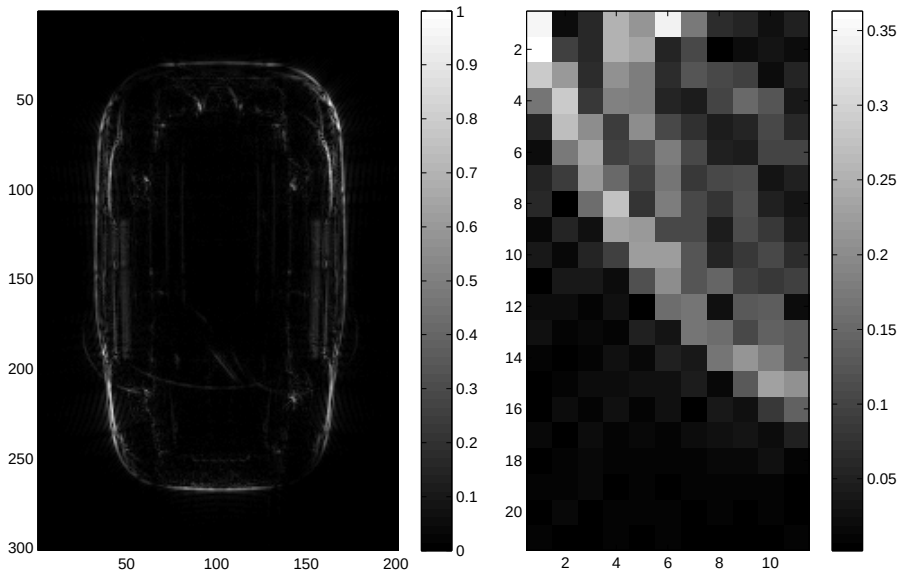


Figure 6.17: (a) The reconstructed Jeep image via FBP using the entire CV dome data set and (b) the left-bottom corner of the car that is used in correlation calculations.

Figure 6.17(a) shows the original target scene reconstructed using the deterministic FBP and the full aperture data without any noise. Figure 6.18(a) shows

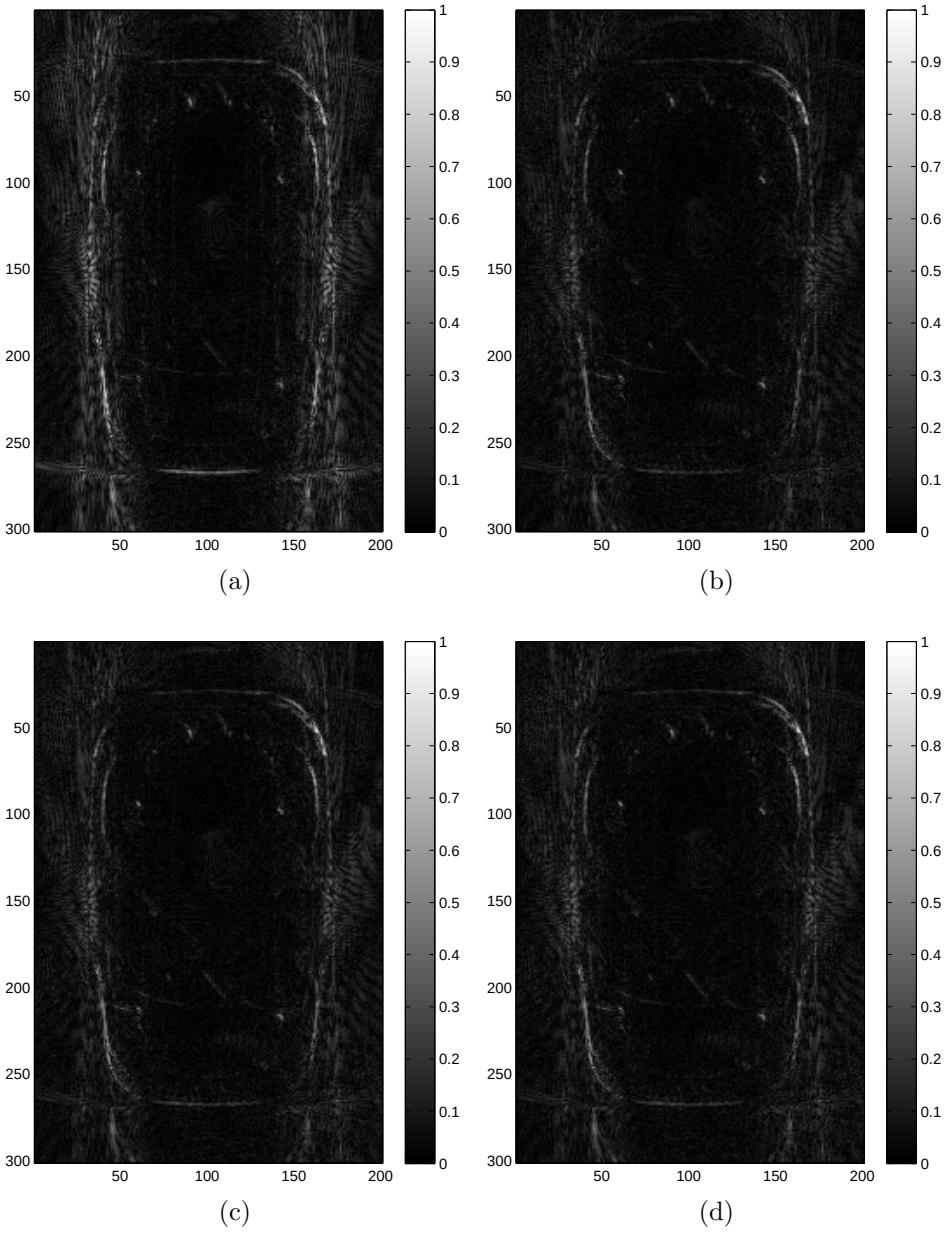


Figure 6.18: The reconstructed SAR images for the CV dome data set using (a) the deterministic FBP; and IRtA after (b) the first, (c) second, and (d) the tenth iterations, respectively when the SNR is 20dB and $\rho(f) = \nabla f$.

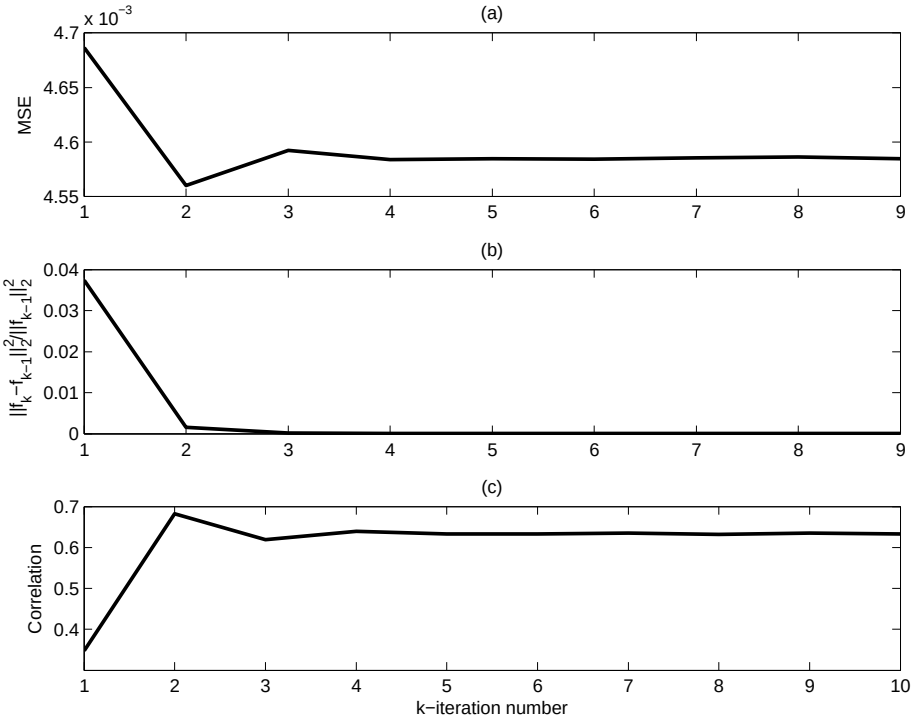


Figure 6.19: (a) The MSE between the original target scene and the reconstructed target image for each iteration, (b) the Δ-norm, (c) the correlation of the reconstructed image patch shown in Figure 6.17(b) using IRtA and the CV dome data set when the SNR is 20dB and $\rho(f) = \nabla f$.

the reconstructed image using the deterministic FBP. Figures 6.18(b)-(d) shows the reconstructed images after the first, second and the fifth iterations, respectively at 20dB SNR using IRtA. The image reconstructed using the deterministic FBP is noisy and contains limited aperture artifacts. These artifacts are the strongest around the corners of the vehicle. However, right after the first iteration, the target becomes visible. We can clearly see the artifact suppression around the corners of the car.

The plots in Figures 6.19(a)-(c) show the MSE between the true and the reconstructed target image, the relative L^2-norm difference between the images reconstructed in consecutive iterations, the correlation of the reconstructed image with the patch shown in Figure 6.17(b) when $\rho(f) = \nabla f$ using IRtA and the CV dome data set. It is seen that the Δ-norm drops sharply after the first iteration and

remains almost constant after the fourth iteration. Similarly, as in Figure 6.19(a), the MSE between the original target scene and the reconstructed image sharply decreases after the first iteration, but then fluctuates at every iteration. As can be seen in Table 6.14, the algorithm converges after the third iteration (see Figure 6.23(a)) and additional iterations may result in increase in the MSE.

Table 6.14: The MSE, Δ-norm and the correlation values using IRtA and the CV dome data set for 10 iterations. The tolerance for the IRtA is set to Δ-norm less than $1e-3$. The row in bold fonts shows the iteration step at which the algorithm converges.

Iteration	MSE	Δ-norm	Correlation
1	4.175e-03	3.734e-02	3.480e-01
2	4.687e-03	1.556e-03	6.823e-01
3	**4.560e-03**	**1.533e-04**	**6.187e-01**
4	4.593e-03	3.697e-05	6.393e-01
5	4.584e-03	2.410e-05	6.330e-01
6	4.585e-03	2.058e-05	6.326e-01
7	4.585e-03	1.910e-05	6.347e-01
8	4.586e-03	1.820e-05	6.324e-01
9	4.586e-03	1.899e-05	6.346e-01

It should be noted here that the data generation code for the CV dome data set is not publicly available. Therefore, we assume that $A(\omega, \boldsymbol{x}, s) = 1$. This assumption ignores the electromagnetic properties of the car which results in an initial increase in the MSE as in Figure 6.22. Figures 6.20(a)-(d) and Figures 6.21(a)-(b) show the reconstructed image with the deterministic FBP, the reconstructed images after the first, second, tenth, twentieth and the thirtieth iterations, respectively at 20dB SNR using IStA. Similar to the IRtA case, the image reconstructed using the deterministic FBP is noisy especially due to limited aperture artifacts. These artifacts are the strongest around the corners of the vehicle. After the first iteration of IStA, noise and limited aperture artifacts are greatly suppressed. However, image is blurred due to backprojection without filtering. As more iterations are performed, the target becomes more visible and the artifact removal is more clearly seen around the corners

of the car.

Plots in Figures 6.22(a)-(c) show the MSE between the original target scene and the reconstructed target image, Δ-norm, and the correlation of the reconstructed image patch shown in Figure 6.17(b) when $\rho(f) = \nabla f$ for IStA using CV radar domes dataset at every iteration. We see that Δ-norm drops sharply after the first iteration and remains almost constant after the third iteration. As shown in 6.19(a), MSE between the original target scene and the reconstructed image first increase then drops sharply. As stated earlier, not using the real data generation code for the CV Radar Domes may cause this initial increase in the MSE. However, after the second iteration where forward projection data is now calculated by the same code, MSE drops consistently. As can be seen from Table 6.15, the algorithm converges after the eighteenth iteration (see Figure 6.23(b)) and additional applications of the algorithm results in a decrease in the correlation.

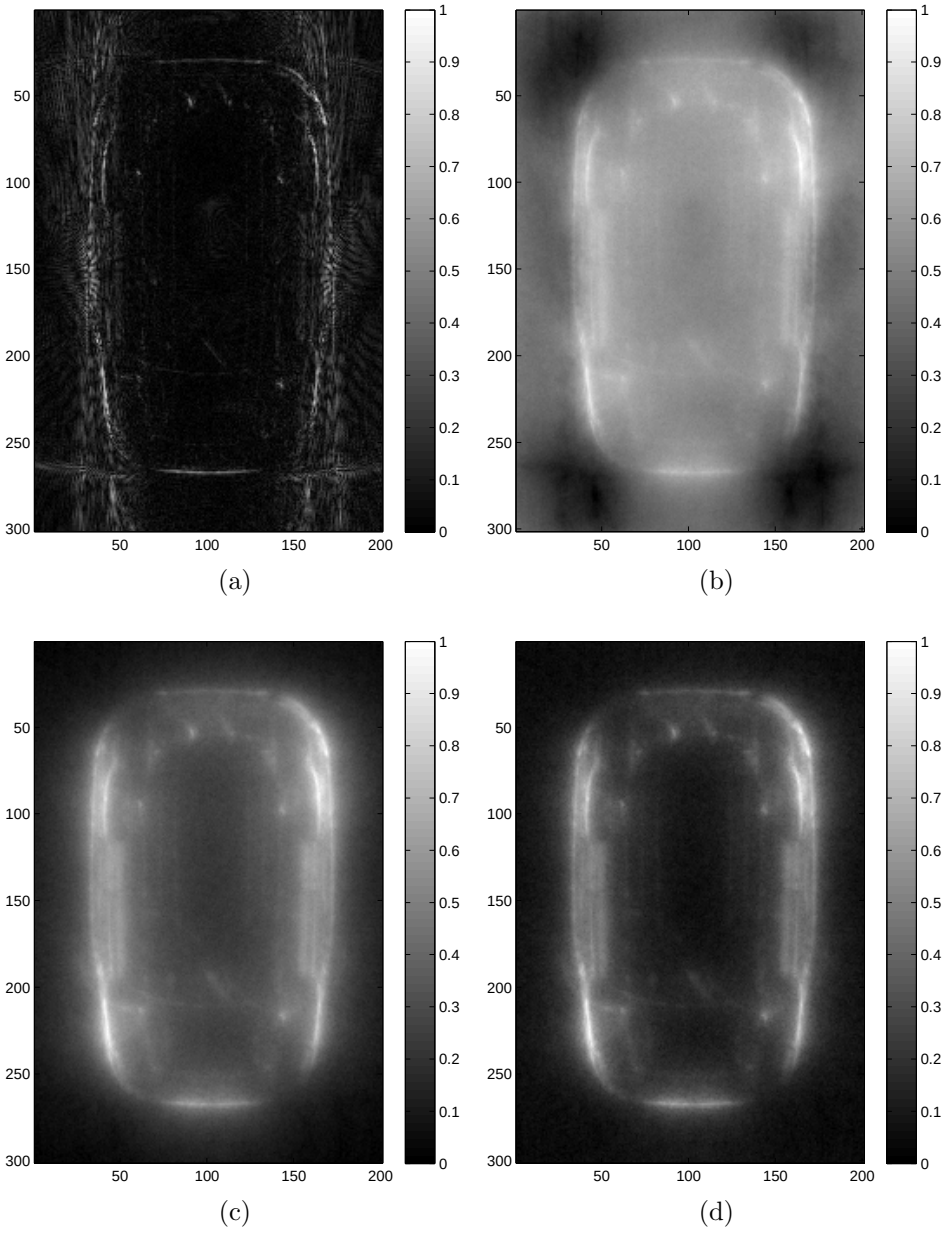
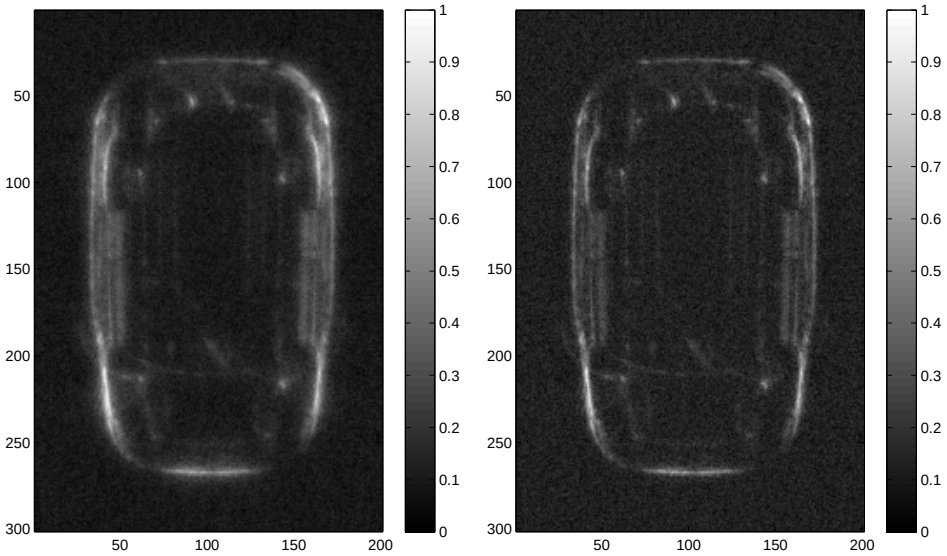


Figure 6.20: The reconstructed SAR images for the CV dome data set using (a) the deterministic FBP; and IStA after (b) the first, (c) second, and (d) the tenth iterations, respectively when the SNR is 20dB and $\rho(f) = \nabla f$.



(a) (b)

Figure 6.21: The reconstructed SAR images for the CV dome data set using IStA after (a) the twentieth, and (b) the thirtieth iterations, respectively when the SNR is 20dB and $\rho(f) = \nabla f$.

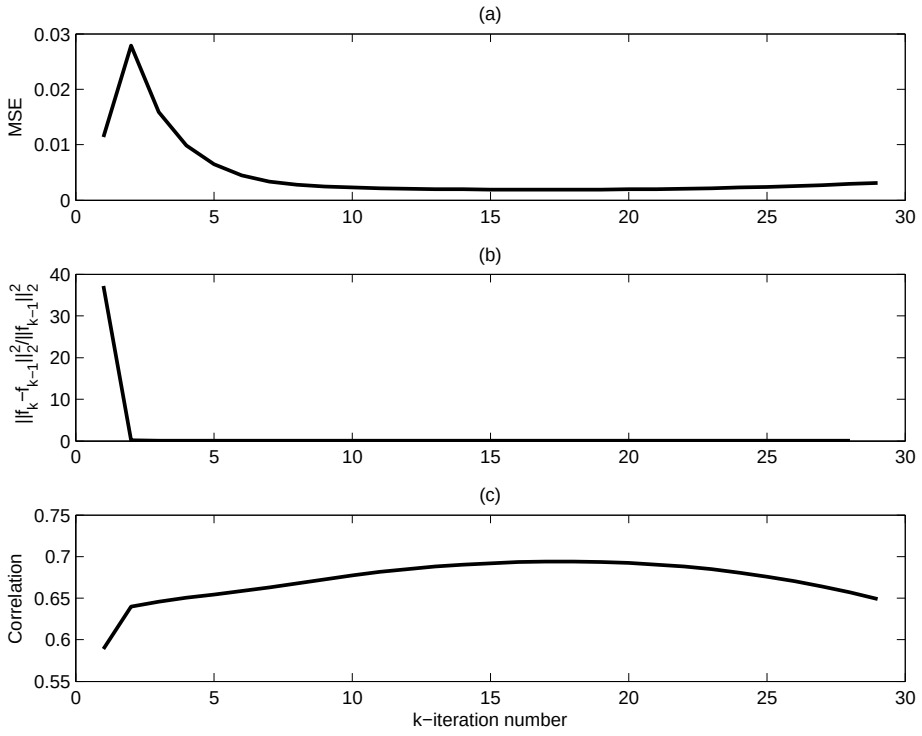


Figure 6.22: (a) The MSE between the original target scene and the reconstructed target image for each iteration, (b) the Δ-norm, (c) the correlation of the reconstructed image patch shown in Figure 6.17(b) using IStA and the CV dome data set when the SNR is 20dB and $\rho(f) = \nabla f$.

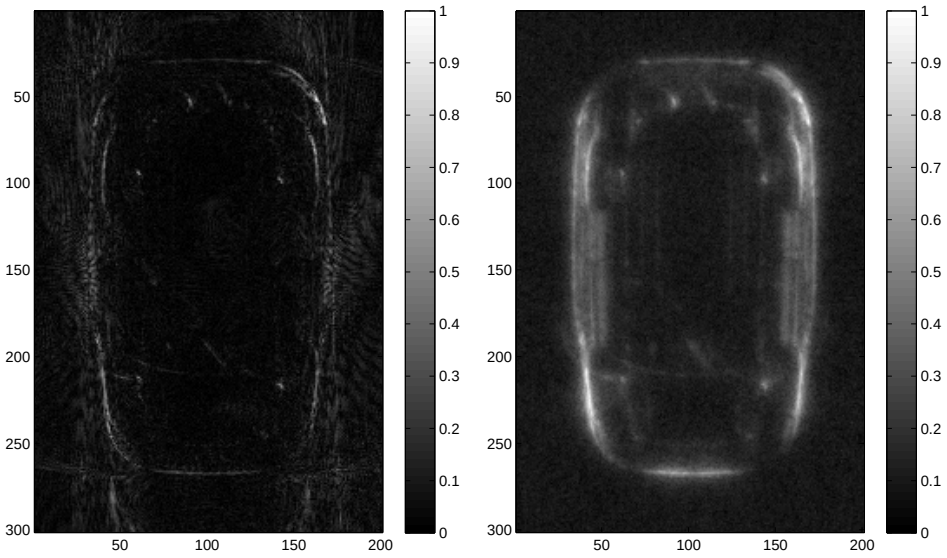


Figure 6.23: Images reconstructed using (a) IRtA and (b) IStA at the convergence step of the algorithms with CV dome data set. IRtA converged at the third iteration and IStA converged at the eighteenth iteration with the tolerance level of 1e-03 when the SNR is 20dB and $\rho(f) = \nabla f$.

Table 6.15: The MSE, Δ-norm and the correlation values using IStA and the CV dome data set for 20 iterations. The tolerance for the IStA is set to Δ-norm less than $1e-3$. The row in bold fonts shows the iteration step at which the algorithm converges.

Iteration	MSE	Δ-norm	Correlation
1	4.175e-03	4.233e+00	3.480e-01
2	1.141e-02	3.718e+01	5.887e-01
3	2.791e-02	1.471e-01	6.397e-01
4	1.592e-02	1.037e-01	6.460e-01
5	9.853e-03	7.726e-02	6.503e-01
6	6.452e-03	5.861e-02	6.542e-01
7	4.469e-03	4.128e-02	6.584e-01
8	3.349e-03	2.388e-02	6.629e-01
9	2.763e-03	1.093e-02	6.678e-01
10	2.464e-03	5.042e-03	6.728e-01
11	2.289e-03	3.039e-03	6.774e-01
12	2.166e-03	2.216e-03	6.815e-01
13	2.074e-03	1.765e-03	6.850e-01
14	2.003e-03	1.489e-03	6.879e-01
15	1.951e-03	1.306e-03	6.903e-01
16	1.915e-03	1.173e-03	6.921e-01
17	1.895e-03	1.072e-03	6.933e-01
18	**1.888e-03**	**9.939e-04**	**6.939e-01**
19	1.896e-03	9.357e-04	6.940e-01
20	1.919e-03	8.949e-04	6.934e-01

6.7 Comparison of IRtA and IStA with Other Sparse Recovery Techniques

In this subsection, we compare the performance of the novel algorithms presented in this thesis with those of other sparse recovery techniques and present numerical results. We consider the following sparse signal recovery algorithms: basis pursuits (BP) [112], iterative reweighted least squares (IRWLS) [147], iterative soft thresholding (IST) [147], least angle regression (LARS) [148], least absolute shrinkage and selection operator (LASSO) [149], polytope faces pursuit (PFP) [150], least squares and stagewise orthogonal matching pursuits (StOMP) [151]. We also consider two SAR reconstruction methods: backprojection and FBP [1]. We use the algorithmic implementations provided by SparseLab package developed by Stanford University [152] in these simulations. For IStA, the convergence is declared when Δ-norm drops below 1e-3. For all other methods, this threshold is set to 1e-4.

We consider a 22km$\times$22km scene descritized into 64×64 pixels. As compared to the numerical simulations in previous sections, the SNR level is set to 30dB since most of the implementations in SparseLab package failed to converge at lower SNR levels. There is a square of size 5.16km$\times$5.16km (15 pixels-by-15 pixels) at the scene center. A mono-static antenna flies over the scene on a circular trajectory given by $\boldsymbol{\gamma}(s) = [11\cos s, 11\sin s, 6.5]$ km. The trajectory is uniformly sampled for $s \in [0, 2\pi]$ at 16 points which is much less than the typical number of aperture points. The fast-time is sampled at 192 points. The computer used for these simulations had an Intel Model X3460 8-core CPU clocked at 2.80GHz and 16 GB RAM. All algorithms are implemented in MATLAB.

The reconstructed images for a single realization using different algorithms are shown in Figures 6.24(a)-(f) and Figures 6.25(a)-(f).

As can be seen from the Figures 6.24(a)-(f) and Figures 6.25(a)-(f), IRtA gives visually the best results followed by the IStA. MSE and the correlation values also concur this with the lowest two MSE values and the highest two correlation values. The fastest method is the backprojection. However, due to the absence of a filter correcting for the intensities of the target, the reconstructed image is greatly blurred (See Figure 6.25(c)). There is more structure in the reconstructed image

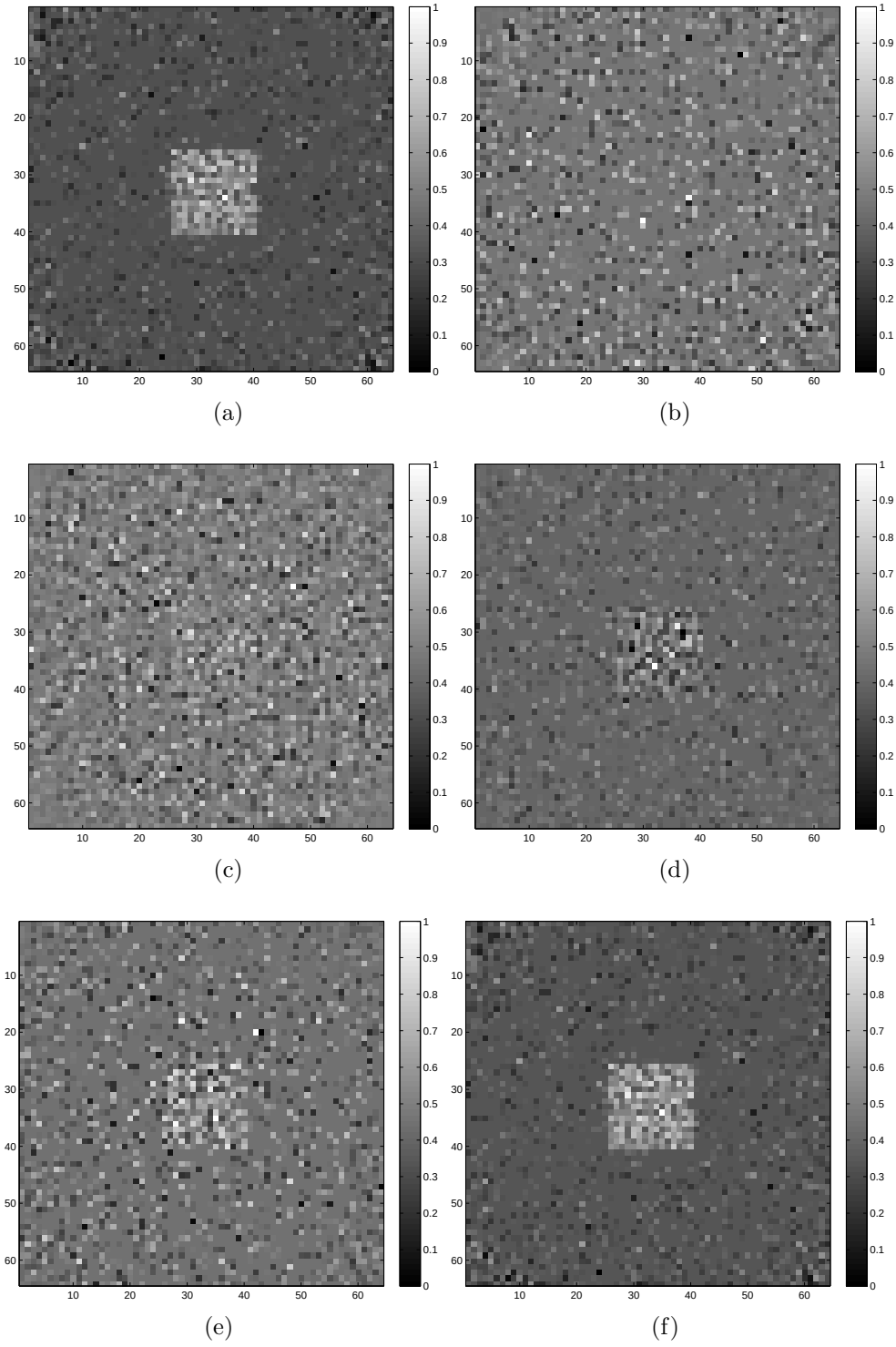


Figure 6.24: The reconstructed images using the methods (a) BP, (b) IRWLS, (c) IST, (d) LARS, (e) LASSO, and (f) PFP at 30dB SNR.

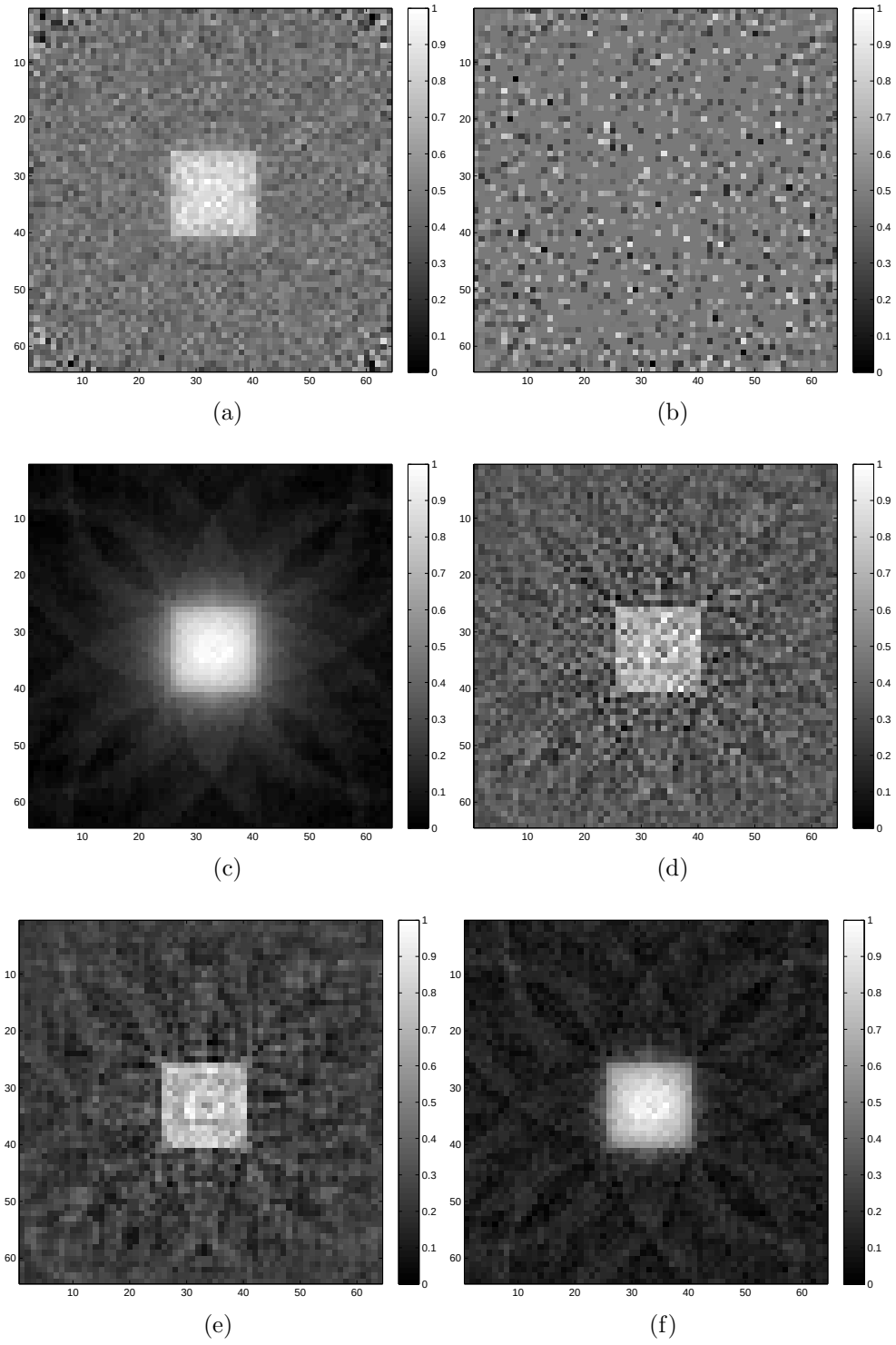


Figure 6.25: The reconstructed images with the methods (a) least squares, (b) stagewise OMP, (c) backprojection, (d) FBP, (e) IRtA, and (f) IStA at 30dB SNR.

when FBP is used as shown in Figure 6.25(c). The image is less blurry, nevertheless it is noisy due to the high-pass nature of the FBP filter, i.e. $\eta \sim \boldsymbol{\xi} \sim \omega$ term in (2.24). Although, the image reconstructed using the deterministic FBP is visually comparable to the images formed by IRtA and IStA, it should be noted that the images reconstructed with the deterministic FBP becomes too noisy when SNR is low. (See Figure 6.1(b) for an image formed by FBP when SNR is 0dB.) Out of eight methods that are used from SparseLab package only BP, LARS, LASSO and PFP produced images that have visible target scenes. However, their computation time is 10 to 100 times greater than that of the IRtA. Only the least squares method, which uses MATLAB's fast implementation of singular value decomposition, has comparable computation time and the image quality. Note that both IRtA and IStA can be implemented with the computational complexity of fast-backprojection algorithms which can further reduce the computation time.

As we can see in Figures 6.25(c)-(f), the artifacts in the reconstructed images have arc-like morphology. These artifacts are direct result of the backprojection based image formation. The backprojection operator is designed according to the imaging geometry. Due to the limited aperture data and the full circular trajectory, the reconstructed images have arc-like artifacts. In addition to the limited aperture effects, there are also artifacts due to noise.

Finally, as we can see from the Table 6.17, the convergence rate of the algorithm greatly depends on the potential function used for IRtA. However, the choice of potential function does not effect the convergence for IStA. Although, IRtA converges five times slower than IStA, it results in lower MSE and higher correlation.

6.7.1 Computational Complexity

Assuming that there are $\mathcal{O}(N)$ samples in the fast-time, slow-time variables and in both directions (x_1, x_2) of the imaging domain, the computational complexity of the IRtA is determined by the following three major steps:

1. Algorithm is initialized by forming an edge-enhanced image of the scene. Using fast-backprojection algorithms, computational complexity of this step is given by $\mathcal{O}(N^2 \log N)$ [40].

Table 6.16: The MSE, correlation of the range profiles, computation time and the iteration steps at which the algorithms converge for various sparse reconstruction techniques. For IRtA and IStA, average values over all possible potential functions are shown. The computer used for these simulations has an Intel Model X3460 8-core CPU clocked at 2.80GHz and 16 GB RAM. All algorithms are implemented in MATLAB.

Method	MSE	Correlation	Time (sec)	Convergence
BP [112]	1.519e-02	3.668e-01	7.909e+01	N/A
IRWLS [147]	1.188e-01	6.058e-02	2.282e+03	N/A
IST [147]	1.156e-01	2.474e-01	6.940e+02	N/A
LARS [148]	8.459e-02	1.408e-01	6.926e+01	1857
LASSO [149]	8.618e-02	3.344e-01	1.161e+03	10624
PFP [150]	1.543e-02	3.658e-01	1.775e+04	30720
Least Squares	1.077e-02	4.031e-01	3.518e+01	N/A
Stagewise OMP [151]	1.247e-01	1.243e-03	2.545e+01	N/A
Backprojection	1.578e-02	4.416e-01	2.408e-02	N/A
FBP	1.586e-02	4.040e-01	2.944e-01	N/A
Reweighted (Average)	4.957e-03	5.151e-01	1.631e+02	18.6
Shrinkage (Average)	2.571e-03	4.517e-01	3.195e+01	9.8

2. Then, 2D FFT of the image formed in the previous step is calculated. This step has a computational complexity of $\mathcal{O}(N^2 \log N)$.

3. We use the filter given in (3.44) for solving each L^2-norm constrained inversion problem. The computational complexity of this step is given by $\mathcal{O}(N^2 \log N)$ at every iteration, if fast backprojection algorithms are used [40].

Assuming that there are k iterations, computational complexity of this method is given by $\mathcal{O}(kN^2 \log N)$. Since, $k \ll N$ the computational complexity of the overall method is given by $\mathcal{O}(N^2 \log N)$.

The computational complexity of the IStA is determined by the following three major steps:

1. Algorithm is initialized by forming an edge-enhanced image of the scene. Using the fast-backprojection algorithms, the computational complexity of this

Table 6.17: The MSE, correlation, computation time and the iteration steps at which the algorithms converge for IRtA and IStA with different potential functions. The computer used for these simulations has an Intel Model X3460 8-core CPU clocked at 2.80GHz and 16 GB RAM. All algorithms are implemented in MATLAB.

Method	MSE	Correlation	Time (sec)	Convergence
IRtA-$\frac{f^2}{f^2+1}$	4.615e-03	5.133e-01	1.432e+02	18
IRtA-f^2	5.342e-03	5.137e-01	1.100e+02	12
IRtA-$\log(f^2+1)$	4.898e-03	5.103e-01	1.066e+02	12
IRtA-$\frac{f}{f+1}$	4.611e-03	5.195e-01	2.317e+02	25
IRtA-f	4.950e-03	5.175e-01	1.494e+02	18
IRtA-$\log(\cosh(f))$	4.741e-03	5.158e-01	9.883e+01	12
IRtA-$\min(f^2, 2f-1)$	4.761e-03	5.058e-01	1.739e+02	19
IRtA-f^p	5.248e-03	5.152e-01	2.697e+02	29
IRtA-$\frac{f^p}{1+f^{2-p}}$	5.041e-03	5.168e-01	9.761e+01	12
IRtA-∇f	5.357e-03	5.232e-01	2.497e+02	29
IStA-$\frac{f^2}{f^2+1}$	2.567e-03	4.511e-01	2.634e+00	10
IStA-f^2	2.586e-03	4.505e-01	3.339e+00	10
IStA-$\log(f^2+1)$	2.573e-03	4.509e-01	3.327e+00	10
IStA-$\frac{f}{f+1}$	2.569e-03	4.522e-01	3.305e+00	10
IStA-f	2.582e-03	4.518e-01	3.296e+00	10
IStA-$\log(\cosh(f))$	2.569e-03	4.510e-01	3.389e+00	10
IStA-$\min(f^2, 2f-1)$	2.548e-03	4.526e-01	2.645e+00	8
IStA-f^p	2.582e-03	4.518e-01	3.430e+00	10
IStA-$\frac{f^p}{1+f^{2-p}}$	2.569e-03	4.522e-01	3.300e+00	10
IStA-∇f	2.560e-03	4.525e-01	3.287e+00	10

method is given by $\mathcal{O}(N^2 \log N)$ [40].

2. This image is forward projected with the computational complexity of the fast backprojection algorithms which is $\mathcal{O}(N^2 \log N)$.

3. An image is formed via backprojection with the computational complexity of $\mathcal{O}(N^2 \log N)$.

Assuming that there are k iterations, the computational complexity of IStA is given by $\mathcal{O}(kN^2 \log N)$. Since $k \ll N$ the computational complexity of the overall method is given by $\mathcal{O}(N^2 \log N)$.

In comparison, the edge-preserving algebraic reconstruction method described in [50] has a computational complexity of $\mathcal{O}(N^4)$. The linear programming and the OMP based techniques described in [88] have the computational complexities of $\mathcal{O}(N^{7/2})$, $\mathcal{O}(N^3)$, respectively. Finally, the subspace pursuits technique presented in [88] has the computational complexity between $\mathcal{O}(N^2 \log N)$ and $\mathcal{O}(N^3)$ depending on the sparsity of the measurements.

6.8 Conclusion

In this chapter, we describe novel, computationally efficient algorithms for SAR image reconstruction suitable for the limited (or sparse) aperture SAR data corrupted by noise. First, an iterative reweighted-type algorithm is described in which L^p-norm constrained SAR image reconstruction problem is approximated with L^2-norm constrained inversion problems iteratively. At every iteration, the solution obtained in the previous iteration is used to determine the filter of the new FBP operator. To increase the sharpness of the reconstructed target scene, we use the edge enhancement filter described in the previous chapter. Then, we introduce an iterative shrinkage-type algorithm which includes repeated applications of forward projection and backprojection operators. We demonstrate the performances of the proposed methods with both synthetic data and the CV dome data set for different scenerios. Finally, we compare the performances of the new methods with those of other sparse signal recovery methods.

CHAPTER 7
CONCLUSION

Measurement noise is ubiquitous in SAR data and targets are often embedded in clutter. There are analytical and numerical optimization-based techniques to address these challenges. Analytic techniques are widely used because of their computationally efficiency. However, they do not address noise and clutter explicitly. There are also numerical optimization-based techniques. These methods use discrete statistical or deterministic models. However, these methods are known to be computationally intensive. In this thesis, we develop computationally efficient analytical solutions for SAR image formation in the presence of noise and clutter.

First, we describe a statistical SAR inversion method to suppress noise and clutter. We present, a novel class of non-stationary stochastic processes to represent radar targets and clutter. We refer to this class of non-stationary processes as pseudo-stationary. They can be considered as locally stationary processes driven by the Wiener process. We develop analytic filtered-backprojection- and backprojection-filtering-type methods based on the minimum mean square error criterion for SAR image reconstruction when the target and clutter scenes belong to this class of non-stationary processes. The optimal filter that suppresses noise and clutter depends on the space-varying spectral densities of the target and clutter. We describe a method to estimate these quantities directly from the SAR data during the image reconstruction process.

We investigate the performance of this method theoretically. We apply these methods to both synthetic data and the real SAR data. Next, we develop an analytic inversion formula based on a best linear unbiased estimation criterion when the clutter is a pseudo-stationary process. We assume that the target is deterministic and demonstrate the performance of the method with numerical simulations.

We investigate non-quadratic prior models to represent target scenes in the second part of the thesis. Specifically, we consider edge preserving prior models that is also applicable to SAR imaging with limited data or sparse aperture. We

formulate the image reconstruction as non-quadratic optimization problems and describe two methods for SAR inversion. First, we present an iterative reweighted-type algorithm. The optimization problem is approximated with a sequence of L^2-norm constrained problems and solved iteratively with a sequence of FBP operators. At every iteration, we use the solution obtained in the previous step to determine the filter of the FBP operator. The second method is a shrinkage-type method. The algorithm is gradient-based and consists of iterative applications of forward projection and backprojection operators.

We also present a novel edge enhancement method that is applicable directly to the SAR received signal. The method enhances the edges in desired directions in the reconstructed image. This novel filter simultaneously reconstructs SAR images and enhances edges of the objects in the reconstructed image. We use this method to sharpen the reconstructed images for IRtA and IStA. We demonstrate the performances of the edge enhancement and the sparse recovery methods both with the synthetic data and the CV dome data set. We also compare the computational complexities and the computation times of these algorithms with those of other state-of-the-art sparse signal recovery techniques.

Noise/clutter suppression techniques are essential for SAR imaging. Moreover, it is also desirable to reconstruct SAR images with limited data and sparse aperture. One of the biggest challenges for real-time SAR imaging is the immense computation power required to form images and the necessity of real-time operation in many applications. Therefore, it becomes a necessity to develop computationally efficient, fast-backprojection algorithms and to use parallel programming techniques.

While we develop IRtA and IStA with noise as the only disturbance term, we can easily extend these methods to include clutter. Additionally, the optimal filter designed for the non-stationary target and clutter, and edge enhancement methods can be modified further for object classification purposes in automatic target recognition applications. Another topic to look into for the methods developed in this thesis may be developing algorithms that do not require additional data to estimate clutter space-varying spectral density functions.

Finally, all the methods described in this thesis can be easily applied to other

imaging modalities such as X-Ray computed tomography, sonar and geophysical imaging in which forward operator is modeled as FIOs.

BIBLIOGRAPHY

[1] C. E. Yarman *et al.*, "Bistatic synthetic aperture radar imaging for arbitrary flight trajectories," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 84–93, Jan. 2008.

[2] K. E. Dungan *et al.*, "Wide angle SAR data for target discrimination research," in *Proc. SPIE, Algorithms for SAR Imagery XIX*, vol. 8394, May 2012, p. 83940M.

[3] C. J. Nolan and M. Cheney, "Synthetic aperture inversion for arbitrary flight paths and nonflat topography," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1035–1043, Sept. 2003.

[4] V. Krishnan *et al.*, "Multistatic synthetic aperture radar image formation," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1290–1306, May 2010.

[5] S. Hensley *et al.*, "The UAVSAR instrument: Description and first results," in *Proc. IEEE Radar Conf.*, 2008, pp. 1–6.

[6] T. Yue-sheng *et al.*, "A multi-mode UAV SAR real-time signal processor system," in *Proc. IEEE 2nd Asian-Pacific SAR Conf.*, 2009, pp. 947–950.

[7] H. Essen *et al.*, "SUMATRA, a W-band SAR for UAV application," in *Proc. IEEE Int. Radar Conf.*, 2009, pp. 1–4.

[8] W. Wang *et al.*, "Waveform-diversity-based millimeter-wave UAV SAR remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 691–700, Mar. 2009.

[9] M. Edwards *et al.*, "MicroASAR: A small, robust LFM-CW SAR for operation on UAVs and small aircraft," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 5, 2008, pp. V–514.

[10] J. Muoz *et al.*, "INTA's developments for UAS and small platforms: QUASAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 4, 2009, pp. IV–999.

[11] E. Peterson *et al.*, "A feasibility assessment for low-cost InSAR formation-flying microsatellites," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2847–2858, Aug. 2009.

[12] J. Coker and A. Tewfik, "Performance synthesis of UAV trajectories in multistatic SAR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 848–863, Apr. 2011.

[13] P. Zhan *et al.*, "Adaptive mobile sensor positioning for multi-static target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 120–132, Feb. 2010.

[14] J. Coker and A. Tewfik, "Compressed sensing and multistatic SAR," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2009, pp. 1097–1100.

[15] J. Coker, "Signal processing algorithms for simplified-hardware multistatic synthetic aperture radar," Ph.D. dissertation, Dept. of Electr. & Comput. Eng., Univ. Minnesota, Minneapolis, MN, 2008.

[16] C. E. Yarman and B. Yazici, "Synthetic aperture hitchhiker imaging," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2156–2173, Nov. 2008.

[17] I. Cumming and J. Bennett, "Digital processing of SEASAT SAR data," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, 1979, pp. 710–718.

[18] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar*, New York, NY: John Wiley & Sons, 1991.

[19] R. Raney *et al.*, "Precision SAR processing using chirp scaling," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 786–799, July 1994.

[20] C. Cafforio *et al.*, "SAR data focusing using seismic migration techniques," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 2, pp. 194–207, Mar. 1991.

[21] R. Bamler, "A comparison of range-Doppler and wavenumber domain SAR focusing algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 4, pp. 706–713, July 1992.

[22] R. Stolt, "Migration by Fourier transform," *Geophysics*, vol. 43, no. 1, pp. 23–48, Feb. 1978.

[23] J. L. Walker, "Range-Doppler imaging of rotating objects," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES–16, no. 1, pp. 23–52, Jan. 1980.

[24] A. W. Doerry. (2012, May). Basics of Polar-Format Algorithm for processing synthetic aperture radar images. Sandia Nat. Laboratories, Albuquerque, NM and Livermore, CA. [Online]. Available: http://prod.sandia.gov/techlib/access-control.cgi/2012/123369.pdf (Retrieved on July 20, 2014).

[25] B. Yazici *et al.*, "Synthetic-aperture inversion in the presence of noise and clutter," *Inverse Problems*, vol. 22, no. 5, p. 1705, Oct. 2006.

[26] R. Fa *et al.*, "Knowledge-aided STAP algorithm using convex combination of inverse covariance matrices for heterogenous clutter," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2010, pp. 2742–2745.

[27] J. Roman *et al.*, "Parametric adaptive matched filter for airborne radar applications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 677–692, Apr. 2000.

[28] K. Sohn *et al.*, "Parametric GLRT for multichannel adaptive signal detection," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5351–5360, Nov. 2007.

[29] P. Wang *et al.*, "A new parametric GLRT for multichannel adaptive signal detection," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 317–325, Jan. 2010.

[30] T. Kragh and A. Kharbouch, "Monotonic iterative algorithms for SAR image restoration," in *Proc. IEEE Int. Conf. Image Process.*, 2006, pp. 645–648.

[31] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, Jul. 1993.

[32] R. Bose *et al.*, "Sequence CLEAN: A modified deconvolution technique for microwave images of contiguous targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 1, pp. 89–97, Jan. 2002.

[33] M. McClure and L. Carin, "Matching pursuits with a wave-based dictionary," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 2912–2927, Dec. 1997.

[34] A. Brito *et al.*, "Sar image superresolution via 2-D adaptive extrapolation," *Multidimensional Syst. and Signal Process.*, vol. 14, no. 1–3, pp. 83–104, Jan. 2003.

[35] M. Pinheiro and M. Rodriguez-Cassola, "Reconstruction methods of missing SAR data: analysis in the frame of TanDEM-X synchronization link," in *Proc. 9th European Conf. SAR*, 2012, pp. 742–745.

[36] W. Rao *et al.*, "ISAR imaging of maneuvering targets with missing data via matching pursuit," in *Proc. IEEE Radar Conf.*, 2011, pp. 124–128.

[37] Y. Wang *et al.*, "Nonparametric spectral analysis with missing data via the EM algorithm," *Digit. Signal Process.*, vol. 15, no. 2, pp. 191–206, Mar. 2005.

[38] Y. Wang *et al.*, "Spectral Analysis of Signals: The Missing Data Case," *Synthesis Lectures on Signal Process. Series*, vol. 1, no. 1, pp. 1–102, Jan. 2006.

[39] S. Nilsson, "Application of fast backprojection techniques for some inverse problems of integral geometry," Ph.D. dissertation, Dept. of Mathematics, Linköping University, Linköping, Sweden, 1997.

[40] L. M. Ulander *et al.*, "Synthetic-aperture radar processing using fast factorized back-projection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 3, pp. 760–776, July 2003.

[41] M. Priestley, *Spectral Analysis and Time Series*, London, United Kingdom: Academic press, 1981.

[42] K. E. Dungan *et al.*, "Civilian vehicle radar data domes," in *Proc. SPIE, Algorithms for SAR Imagery XVII*, vol. 7699, 2010, p. 76990P.

[43] J. Duistermaat, *Fourier Integral Operators*, vol. 130, New York, NY: Birkhauser, 1996.

[44] J. Treves, *Introduction to Pseudodifferential and Fourier Integral Operators*, vol. 2, New York, NY: Springer, 1980.

[45] A. Grigis and J. Sjöstrand, *Microlocal Analysis for Differential Operators: An Introduction*, vol. 196, ser. London Mathematical Soc. Lecture Note Series, New York, NY: Cambridge University Press, 1994.

[46] B. Yazıcı and V. Krishnan. (2010, Mar. 15). *Microlocal Analysis in Imaging.* [Online]. Available: http: //www.ecse.rpi.edu/~yazici/ICASSPTutorial/microlocal_ICASSP10.html (Retrieved on July 10, 2014).

[47] G. Beylkin, "Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform," *J. Math. Phys.*, vol. 26, no. 1, pp. 99–108, Jan. 1985.

[48] Y. Abramovich *et al.*, "Two-dimensional multivariate parametric models for radar applications-part II: Maximum-entropy extensions for hermitian-block matrices," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5527–5539, Nov. 2008.

[49] J. Bergin *et al.*, "Improved clutter mitigation performance using knowledge-aided space-time adaptive processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 997 –1009, July 2006.

[50] M. Çetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Trans. Image Process.*, vol. 10, no. 4, pp. 623–631, Apr. 2001.

[51] W. Melvin and J. Guerci, "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 983–996, July 2006.

[52] P. Stoica *et al.*, "On using a priori knowledge in space-time adaptive processing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2598–2602, June 2008.

[53] D. Vu *et al.*, "Nonparametric missing sample spectral analysis and its applications to interrupted SAR," *IEEE J. Sel. Topics in Signal Process.*, vol. 6, no. 1, pp. 1–14, Feb. 2012.

[54] B. Yazici *et al.*, "Synthetic aperture inversion with sparsity constraints," in *Proc. IEEE Int. Conf. Electromagn. in Advanced Appl.*, 2011, pp. 1404–1407.

[55] S. R. DeGraaf, "SAR imaging via modern 2-D spectral estimation methods," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 729–761, May 1998.

[56] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1469–1484, June 1996.

[57] Z. Bi *et al.*, "Super resolution SAR imaging via parametric spectral estimation methods," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 1, pp. 267–281, Jan. 1999.

[58] E. G. Larsson and J. Li, "Spectral analysis of periodically gapped data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 3, pp. 1089–1097, July 2003.

[59] R. Wu *et al.*, "SAR image formation via semiparametric spectral estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 4, pp. 1318–1333, Oct. 1999.

[60] F. Lombardini and A. Reigber, "Adaptive spectral estimation for multibaseline SAR tomography with airborne L-band data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 3, 2003, pp. III–2014.

[61] S. Xiao and D. C. Munson Jr., "Spotlight-mode SAR imaging of a three-dimensional scene using spectral estimation techniques," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 2, 1998, pp. 642–644.

[62] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, Jan. 2012.

[63] Z. Liu *et al.*, "Efficient implementation of Capon and APES for spectral estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 4, pp. 1314–1319, Oct. 1998.

[64] M. Xue *et al.*, "IAA spectral estimation: fast implementation using the Gohberg–Semencul factorization," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3251–3261, July 2011.

[65] L. Borcea *et al.*, "Adaptive interferometric imaging in clutter and optimal illumination," *Inverse Problems*, vol. 22, no. 4, p. 1405, Aug. 2006.

[66] J. Garnier and K. Sølna, "Coherent interferometric imaging for synthetic aperture radar in the presence of noise," *Inverse Problems*, vol. 24, no. 5, p. 055001, Oct. 2008.

[67] J. Swoboda *et al.*, "Bistatic synthetic aperture radar imaging for arbitrary trajectories in the presence of noise and clutter," in *Proc. SPIE, Airborne Intell., Surveillance, Reconnaissance Syst. and Appl. VI*, vol. 7307, 2009, p. 73070D.

[68] K. Voccola *et al.*, "On the relationship between the generalized likelihood ratio test and backprojection for synthetic aperture radar imaging," in *Proc. SPIE, Automat. Target Recognition XIX*, vol. 7335, 2009, p. 73350I.

[69] V. Krishnan and B. Yazici, "Synthetic aperture radar imaging exploiting multiple scattering," *Inverse Problems*, vol. 27, no. 5, p. 055004, May 2011.

[70] K. Voccola *et al.*, "Polarimetric synthetic-aperture inversion for extended targets in clutter," *Inverse Problems*, vol. 29, no. 5, p. 054003, May 2013.

[71] M. Cetin *et al.*, "Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 27–40, July 2014.

[72] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*, vol. 1, Upper Saddle River, NJ: Prentice Hall, 1997.

[73] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP–34, no. 3, pp. 276–280, Mar. 1986.

[74] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.

[75] J. Capon, "Maximum-Likelihood Spectral Estimation," in *Nonlinear Methods of Spectral Analysis*, New York, NY: Springer-Verlag, 1983, pp. 155–179.

[76] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 281–295, Feb. 1996.

[77] J. Ward, "Space-time adaptive processing for airborne radar," in *Inst. Elect. Eng. Colloq. STAP*, 1998, p. 2.

[78] A. Farina and P. Lombardo, "Space-time techniques for SAR," in *Appl. Space-Time Adaptive Process.*, London, United Kingdom: Inst. Elect. Eng., 2004, pp. 73–122.

[79] D. Cerutti-Maori *et al.*, "Experimental verification of SAR-GMTI improvement through antenna switching," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 2066–2075, Apr. 2010.

[80] W. L. Melvin, "A STAP overview," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 19, no. 1, pp. 19–35, Jan. 2004.

[81] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physical Rev.*, vol. 40, no. 5, p. 749, June 1932.

[82] V. Krishnan and E. Quinto, "Microlocal aspects of common offset synthetic aperture radar imaging," *Inverse Problems and Imag.*, vol. 5, no. 3, pp. 659–674, Aug. 2011.

[83] M. Priestley, "Evolutionary spectra and non-stationary processes," *J. Roy. Statistical Soc. Series B (Methodological)*, vol. 27, no. 2, pp. 204–237, 1965.

[84] E. Candes *et al.*, "Fast computation of Fourier integral operators," *SIAM J. Sci. Comput.*, vol. 29, no. 6, pp. 2464–2493, Oct. 2007.

[85] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES–2, no. 2, pp. 115–127, Mar. 1986.

[86] T. Yardibi *et al.*, "Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 425–443, Jan. 2010.

[87] X. Tan *et al.*, "Sparse learning via iterative minimization with application to MIMO radar imaging," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1088–1101, Mar. 2011.

[88] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[89] D. Shnidman, "Generalized radar clutter model," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 3, pp. 857–865, July 1999.

[90] R. González and R. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 2008.

[91] C. J. Nolan and M. Cheney, "Synthetic aperture inversion," *Inverse Problems*, vol. 18, no. 1, p. 221, Feb. 2002.

[92] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI–8, no. 6, pp. 679–698, Nov. 1986.

[93] J. Bruder and R. Schneible, "Interrupted SAR waveforms for high interrupt ratios," in *Proc. IET Int. Conf. Radar Syst.*, 2007, pp. 1–5.

[94] J. Salzman *et al.*, "Interrupted synthetic aperture radar (SAR)," *IEEE Trans. Aerosp. Electron. Syst. Mag.*, vol. 17, no. 5, pp. 33–39, May 2002.

[95] M. Ferrara *et al.*, "Resolution optimization with irregularly sampled Fourier data," *Inverse Problems*, vol. 29, no. 5, p. 054007, May 2013.

[96] J. N. Ash, "Joint imaging and change detection for robust exploitation in interrupted SAR environments," in *Proc. SPIE, Algorithms for SAR Imagery XX*, vol. 8746, 2013, p. 87460J.

[97] I. Stojanovic *et al.*, "Reconstruction of interrupted SAR imagery for persistent surveillance change detection," in *Proc. SPIE, Algorithms for SAR Imagery XIX*, vol. 8394, 2012, p. 839408.

[98] L. C. Potter *et al.*, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006–1020, June 2010.

[99] J. H. Ender, "On compressive sensing applied to radar," *Signal Process.*, vol. 90, no. 5, pp. 1402–1414, May 2010.

[100] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2275–2284, June 2009.

[101] R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *Proc. IEEE Radar Conf.*, 2007, pp. 128–133.

[102] M. Herman and T. Strohmer, "Compressed sensing radar," in *Proc. IEEE Radar Conf.*, 2008, pp. 1–6.

[103] E. J. Candes *et al.*, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure and Appl. Mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[104] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[105] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[106] J. Burns *et al.*, "Adaptive decomposition in electromagnetics," in *Proc. IEEE Antennas and Propag. Soc. Int. Symp.*, vol. 3, 1997, pp. 1984–1987.

[107] P. S. Huggins and S. W. Zucker, "Greedy basis pursuit," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3760–3772, July 2007.

[108] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. and Comput. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.

[109] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[110] J. A. Tropp *et al.*, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.

[111] S. S. Chen *et al.*, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, Aug. 1998.

[112] S. Chen and D. Donoho, "Basis pursuit," in *Rec. 28th Asilomar Signals, Syst. and Computers Conf.*, vol. 1, 1994, pp. 41–44.

[113] I. Daubechies *et al.*, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure and Appl. Mathematics*, vol. 63, no. 1, pp. 1–38, Jan. 2009.

[114] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 187–200, Jan. 1999.

[115] E. J. Candes *et al.*, "Enhancing sparsity by reweighted l-1 minimization," *J. Fourier Anal. and Appl.*, vol. 14, no. 5-6, pp. 877–905, Dec. 2008.

[116] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sept. 1994.

[117] D. L. Donoho *et al.*, "Wavelet shrinkage: asymptopia?" *J. Roy. Statistical Soc. Series B (Methodological)*, vol. 57, no. 2, pp. 301–369, 1995.

[118] D. L. Donoho and I. M. Johnstone, "Minimax estimation via wavelet shrinkage," *The Ann. Statist.*, vol. 26, no. 3, pp. 879–921, June 1998.

[119] M. A. Figueiredo *et al.*, "Majorization–minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2980–2991, Dec. 2007.

[120] I. Daubechies *et al.*, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure and Appl. Mathematics*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.

[121] M. A. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, Aug. 2003.

[122] M. A. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, 2005, pp. II–782.

[123] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. and Appl.*, vol. 14, no. 5-6, pp. 629–654, Dec. 2008.

[124] H. Guo *et al.*, "Speckle reduction via wavelet shrinkage with application to synthetic-aperture-radar-based automatic target detection/reduction (ATD/R)," in *Proc. SPIE, Wavelet Appl. in Signal and Image Process. II*, vol. 2303, 1994, pp. 333–344.

[125] A. Achim *et al.*, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 8, pp. 1773–1784, Aug. 2003.

[126] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sciences*, vol. 2, no. 1, pp. 183–202, Mar. 2009.

[127] A. Pizurica *et al.*, "Despeckling SAR images using wavelets and a new class of adaptive shrinkage estimators," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, 2001, pp. 233–236.

[128] D. Wei *et al.*, "Simultaneous speckle reduction and data compression using best wavelet packet bases with application to synthetic aperture radar (SAR) based ATD/R," in *Proc. SPIE, Wavelet Appl. II*, vol. 2941, 1995, pp. 1131–1141.

[129] G. Davis *et al.*, "Adaptive greedy approximations," *Constructive Approximation*, vol. 13, no. 1, pp. 57–98, 1997.

[130] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, New York, NY: Springer-Verlag, 2010.

[131] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, July 1993.

[132] C. A. Bouman *et al.*, (1995, Oct. 23-26). *Markov Random Fields and Stochastic Image Models*. [Online]. Available: http://www.cis.temple.edu/

~latecki/Courses/RobotFall08/Papers/MRFBauman.pdf (Retrieved on October 23, 2014).

[133] R. Chellappa and A. Jain, *Markov Random Fields: Theory and Application*, vol. 1, Boston, MA: Academic Press, 1993.

[134] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Statistical Soc. Series B (Methodological)*, vol. 36, no. 2, pp. 192–236, 1974.

[135] R. Kindermann *et al.*, *Markov Random Fields and Their Applications*, vol. 1, Providence, RI: Amer. Math. Soc., 1980.

[136] A. Blake, "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 1, pp. 2–12, Jan. 1989.

[137] J. Besag, "Towards Bayesian image analysis," *J. Appl. Stat.*, vol. 20, no. 5-6, pp. 107–119, 1993.

[138] R. Stevenson and E. Delp, "Fitting curves with discontinuities," in *Proc. 1st Int. Workshop on Robust Comput. Vision*, 1990, pp. 127–136.

[139] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imag.*, vol. 9, no. 1, pp. 84–93, Mar. 1990.

[140] S. Ganan and D. McClure, "Bayesian image analysis: an application to single photon emission tomography," in *Proc. Statistical Comput. Section*, 1985, pp. 12–18.

[141] J.-B. Thibault *et al.*, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med. Physics*, vol. 34, no. 11, pp. 4526–4544, Nov. 2007.

[142] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.

[143] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-d bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. Med. Imag.*, vol. 8, no. 2, pp. 194–202, June 1989.

[144] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.

[145] H. C. Yanik *et al.*, "Computationally efficient FBP-type direct segmentation of synthetic aperture radar images," in *Proc. SPIE, Algorithms for SAR Imagery XVIII*, vol. 8051, 2011, p. 80510C.

[146] K. Pearson, "Mathematical contributions to the theory of evolution.–on a form of spurious correlation which may arise when indices are used in the measurement of organs," *Proc. Roy. Soc. London*, vol. 60, no. 359–367, pp. 489–498, Jan. 1896.

[147] T. Hastie *et al.*, *The Elements of Statistical Learning*, vol. 2, no. 1, New York, NY: Springer, 2009.

[148] B. Efron *et al.*, "Least angle regression," *The Ann. Statist.*, vol. 32, no. 2, pp. 407–499, Apr. 2004.

[149] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Statistical Soc. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[150] M. D. Plumbley, "On polar polytopes and the recovery of sparse representations," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3188–3195, Sept. 2007.

[151] D. L. Donoho *et al.*, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012.

[152] Stanford University. (2007, May 26). *SparseLAB*. [Online]. Available: http://sparselab.stanford.edu/ (Retrieved on July 20, 2014).

[153] Y. S. Abu-Mostafa *et al.*, *Learning from Data*, AMLBook.com, 2012.

[154] S. M. Kay, *Fundamentals of Statistical Processing*, vol. 1, Upper Saddle River, NJ: Prentice Hall, 1993.

[155] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via L1 minimization," *Proc. Nat. Academy Sciences*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.

[156] E. J. Candès *et al.*, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[157] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

APPENDIX A
METHOD OF THE STATIONARY PHASE

Let u be a smooth function of compact support in $\mathbb{R}^n$ and let ϕ be a real-valued function with only non-degenerate critical points. A point $\boldsymbol{x}_0 \in \mathbb{R}^n$ is called a non-degenerate critical point if $\nabla\phi(\boldsymbol{x}_0) = 0$ where

$$\nabla\phi(\boldsymbol{x}_0) = \begin{bmatrix} \nabla_{x_1}\phi \\ \nabla_{x_2}\phi \end{bmatrix}. \tag{A.1}$$

and the Hessian matrix of ϕ has non-zero determinant. The stationary phase theorem states that as $\lambda \to \infty$

$$\begin{aligned} \int \mathrm{e}^{\mathrm{i}\lambda\phi(\boldsymbol{x})} u(\boldsymbol{x}) \mathrm{d}^n\boldsymbol{x} = \left(\frac{2\pi}{\lambda}\right)^{n/2} \sum_{\{\boldsymbol{x}_0 : \nabla\phi(\boldsymbol{x}_0)=0\}} u(\boldsymbol{x}_0) \frac{\mathrm{e}^{\mathrm{i}\lambda\phi(\boldsymbol{x}_0)} \mathrm{e}^{\mathrm{i}(\pi/4)\mathrm{sgn}(H_{\boldsymbol{x}}\phi(\boldsymbol{x}_0))}}{\sqrt{|\det H_{\boldsymbol{x}}\phi(\boldsymbol{x}_0)|}} \\ + \mathcal{O}(\lambda^{-n/2-1}) \end{aligned} \tag{A.2}$$

where $H_{\boldsymbol{x}}(\phi(\boldsymbol{x}_0))$ denotes the Hessian of ϕ given by:

$$H_{\boldsymbol{x}}(\phi(\boldsymbol{x}_0)) = \begin{bmatrix} \nabla_{x_1^2}\phi & \nabla_{x_1,x_2}\phi \\ \nabla_{x_2,x_1}\phi & \nabla_{x_2^2}\phi \end{bmatrix}. \tag{A.3}$$

APPENDIX B
PROOF OF THE LEMMA 1

(i) We write

$$\mathcal{J}_{ns}^{T}(Q) = \int \mathrm{E}[|(\mathcal{K}_{ns}\mathcal{F} - \mathcal{I}_{\Omega})[T](\boldsymbol{z})|^2]\,\mathrm{d}\boldsymbol{z}. \tag{B.1}$$

By using (3.33),

$$\begin{aligned}\mathcal{J}_{ns}^{T}(Q_{ns}) = \int & \mathrm{e}^{\mathrm{i}2\pi[(\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi}-(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}']}[Q_{ns}(\boldsymbol{z},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi}) - \tilde{\chi}_{\Omega}(\boldsymbol{z},\boldsymbol{\xi})] \\ & \times \overline{[Q_{ns}(\boldsymbol{z},\boldsymbol{\xi}')A(\boldsymbol{x}',\boldsymbol{\xi}')\eta(\boldsymbol{x}',\boldsymbol{z},\boldsymbol{\xi}') - \tilde{\chi}_{\Omega}(\boldsymbol{z},\boldsymbol{\xi}')]} \\ & \times \mathcal{R}_{T}(\boldsymbol{x},\boldsymbol{x}')\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'\mathrm{d}\boldsymbol{\xi}'\mathrm{d}\boldsymbol{z},\end{aligned} \tag{B.2}$$

where $\tilde{\chi}_{\Omega}(\boldsymbol{z},\boldsymbol{\xi})$ is a function that prevents division by 0. We apply the method of the stationary phase in the variables $\boldsymbol{z}$, $\boldsymbol{\xi}'$ for the phase

$$\phi_4(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\xi},\boldsymbol{x}',\boldsymbol{\xi}') = (\boldsymbol{x}-\boldsymbol{z})\cdot\boldsymbol{\xi} - (\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}' \tag{B.3}$$

We check the Hessian:

$$H_{\boldsymbol{z},\boldsymbol{\xi}'}(\phi_4) = \begin{bmatrix} \nabla_{\boldsymbol{z}^2}\phi_4 & \nabla_{\boldsymbol{z},\boldsymbol{\xi}'}\phi_4 \\ \nabla_{\boldsymbol{\xi}',\boldsymbol{z}}\phi_4 & \nabla_{\boldsymbol{\xi}'^2}\phi_4 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}. \tag{B.4}$$

$\det(H_{\boldsymbol{z},\boldsymbol{\xi}'}(\phi_4)) = -1$ where $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Therefore, we get the critical points

$$\nabla_{\boldsymbol{z}}\phi_4 \to \boldsymbol{\xi}' = \boldsymbol{\xi}, \tag{B.5}$$

$$\nabla_{\boldsymbol{\xi}'}\phi_4 \to \boldsymbol{z} = \boldsymbol{x}'. \tag{B.6}$$

The leading-order term of B.2 can be written as

$$\mathcal{J}_{ns}^{T}(Q_{ns}) \approx \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}}[Q_{ns}(\boldsymbol{x}',\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{\xi}) - \tilde{\chi}_{\Omega}(\boldsymbol{x}',\boldsymbol{\xi})]\mathcal{R}_{T}(\boldsymbol{x},\boldsymbol{x}') \times \overline{[Q_{ns}(\boldsymbol{x}',\boldsymbol{\xi}')A(\boldsymbol{x}',\boldsymbol{\xi}')\eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi}') - \tilde{\chi}_{\Omega}(\boldsymbol{x}',\boldsymbol{\xi}')]}\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'. \tag{B.7}$$

Writing

$$\mathcal{R}_{T}(\boldsymbol{x},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\zeta}} S_{T}(\boldsymbol{x},\boldsymbol{\zeta})\overline{S_{T}(\boldsymbol{x}',\boldsymbol{\zeta})}\,\mathrm{d}\boldsymbol{\zeta} \tag{B.8}$$

we obtain

$$\mathcal{J}_{ns}^{T}(Q_{ns}) \approx \int \mathrm{e}^{\mathrm{i}2\pi[(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}+(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\zeta}]}[Q_{ns}(\boldsymbol{x}',\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{\xi}) - \tilde{\chi}_{\Omega}(\boldsymbol{x}',\boldsymbol{\xi})] \times \overline{[Q_{ns}(\boldsymbol{x}',\boldsymbol{\xi}')A(\boldsymbol{x}',\boldsymbol{\xi}')\eta(\boldsymbol{x}',\boldsymbol{x}',\boldsymbol{\xi}') - \tilde{\chi}_{\Omega}(\boldsymbol{x}',\boldsymbol{\xi}')]} \times S_{T}(\boldsymbol{x},\boldsymbol{\zeta})\overline{S_{T}(\boldsymbol{x}',\boldsymbol{\zeta})}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\zeta}. \tag{B.9}$$

We apply the method of the stationary phase one more time in the variables $\boldsymbol{x}'$, $\boldsymbol{\zeta}$ for the phase

$$\phi_{5}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{x}',\boldsymbol{\zeta}) = (\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi} + (\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\zeta} \tag{B.10}$$

We check the Hessian:

$$H_{\boldsymbol{x}',\boldsymbol{\zeta}}(\phi_{5}) = \begin{bmatrix} \nabla_{\boldsymbol{x}'^{2}}\phi_{5} & \nabla_{\boldsymbol{x}',\boldsymbol{\zeta}}\phi_{5} \\ \nabla_{\boldsymbol{\zeta},\boldsymbol{x}'}\phi_{5} & \nabla_{\boldsymbol{\zeta}^{2}}\phi_{5} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}. \tag{B.11}$$

$\det H_{\boldsymbol{z},\boldsymbol{\xi}'}(\phi_{5}) = -1$, therefore we get the critical points at

$$\nabla_{\boldsymbol{x}'}\phi_{5} \to \boldsymbol{\zeta} = \boldsymbol{\xi}, \tag{B.12}$$

$$\nabla_{\boldsymbol{\zeta}}\phi_{5} \to \boldsymbol{x}' = \boldsymbol{x}. \tag{B.13}$$

The leading order term of (B.9) is then

$$\mathcal{J}_{ns}^{T}(Q_{ns}) \approx \int \left|Q_{ns}(\boldsymbol{x},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi}) - \tilde{\chi}_{\Omega}(\boldsymbol{x},\boldsymbol{\xi})\right|^2 |S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{B.14}$$

To obtain an approximation to $\mathcal{J}_{ns}^{C}(Q_{ns})$, we follow the same steps (B.1)-(B.14) and get

$$\mathcal{J}_{ns}^{C}(Q_{ns}) \;\approx\; \int \left|Q_{ns}(\boldsymbol{x},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})\right|^2 |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{B.15}$$

For $\mathcal{J}_{ns}^{n}(Q_{ns})$, we first write

$$\begin{aligned}\mathcal{J}_{ns}^{n}(Q_{ns}) &:= \int \mathrm{E}[\mathcal{K}_{ns}[n](\boldsymbol{z})\overline{\mathcal{K}_{ns}[n](\boldsymbol{z}')}]\,\mathrm{d}\boldsymbol{z} \\ &= \int \mathrm{e}^{\mathrm{i}2\pi\omega(t-R(s,\boldsymbol{z})/c)} Q_{ns}(\boldsymbol{z},s,\omega)\sigma^2(s)\mathcal{R}_n(t,t') \\ &\quad \times \mathrm{e}^{-\mathrm{i}2\pi\omega'(t'-R(s,\boldsymbol{z}')/c)}\overline{Q_{ns}(\boldsymbol{z}',s,\omega)}\mathrm{d}\omega\mathrm{d}\omega'\mathrm{d}t\mathrm{d}s\mathrm{d}t'\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{z}'\end{aligned} \tag{B.16}$$

Inserting (3.17) into (B.16) and carrying out the integrations over t, t', ω', gives

$$\begin{aligned}\mathcal{J}_{ns}^{n}(Q_{ns}) = \int & \mathrm{e}^{\mathrm{i}2\pi\omega(R(s,\boldsymbol{z}')-R(s,\boldsymbol{z}))/c} Q_{ns}(\boldsymbol{z},s,\omega)\overline{Q_{ns}(\boldsymbol{z}',s,\omega)} \\ & \times |S_n(s,\omega)|^2 \mathrm{d}\omega\mathrm{d}s\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{z}'.\end{aligned} \tag{B.17}$$

In (B.17), we apply the same change of variables as in (3.28). This transforms (B.17) to

$$\mathcal{J}_{ns}^{n}(Q_{ns}) \approx \int e^{\mathrm{i}2\pi(\boldsymbol{z}'-\boldsymbol{z})\cdot\boldsymbol{\xi}} Q_{ns}(\boldsymbol{z},\boldsymbol{\xi})\overline{Q_{ns}(\boldsymbol{x}',\boldsymbol{\xi})}\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})|S_n(\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{z}'. \tag{B.18}$$

By following a similar argument in [25], if the noise power spectral density decays slowly for large $\boldsymbol{\xi}$, then $\mathcal{J}_n(Q)$ has singularities at $\boldsymbol{x}' = \boldsymbol{x}$ and we

obtain

$$\mathcal{J}_{ns}^{n}(Q_{ns}) \approx \int |Q_{ns}(\boldsymbol{z},\boldsymbol{\xi})||S_n(\boldsymbol{\xi})|^2 \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \tag{B.19}$$

(ii) Using the approximation obtained in (3.39) for $\mathcal{K}_{bp}\mathcal{F}$ and following the steps in (B.1)-(B.14), the leading-order singularities of $\mathcal{J}_{bp}^{T}(Q_{bp})$, $\mathcal{J}_{bp}^{C}(Q_{bp})$ and $\mathcal{J}_{bp}^{n}(Q_{bp})$ are given by (B.14), (B.15) and (B.19) with Q_{ns} replaced by $\hat{Q}_{bp}$ where $\hat{Q}_{bp}$ is defined by (3.43)

$$Q_{bp}(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{z})\cdot\boldsymbol{\xi}'} \hat{Q}_{bp}(\boldsymbol{z},\boldsymbol{\xi}')\mathrm{d}\boldsymbol{\xi}'. \tag{B.20}$$

APPENDIX C
PROOF OF THE THEOREM 1

(i) We have

$$\mathcal{J}_{ns}(Q') = \mathcal{J}_{ns}^{T}(Q') + \mathcal{J}_{ns}^{C}(Q') + \mathcal{J}_{ns}^{n}(Q'). \tag{C.1}$$

Let us fix a function Q_ϵ and for some small ϵ consider a variation of the form $Q' = Q_{ns} + \epsilon Q_\epsilon$. This is substituted into (3.24) - (3.26) and differentiated with respect to ϵ and then ϵ is set to 0. This results in

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{ns}^{T}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0} + \frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{ns}^{C}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0} + \frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{ns}^{n}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0}. \tag{C.2}$$

From (3.40), the first term on the right side of (C.2) is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{ns}^{T}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0} \approx & \int Q_\epsilon A\eta |S_T|^2 \overline{(Q_{ns}A\eta - \tilde{\chi})}\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \\ & + \int (Q_{ns}A\eta - \tilde{\chi})|S_T|^2 \overline{Q_\epsilon A\eta}\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\end{aligned} \tag{C.3}$$

which can be written as

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{ns}^{T}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0} \approx 2\mathrm{Re}\int \overline{Q_\epsilon A\eta}|S_T|^2(Q_{ns}A\eta - \tilde{\chi})\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}. \tag{C.4}$$

Similarly,

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}_{ns}^{C}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0} \approx 2\mathrm{Re}\int \overline{Q_\epsilon A\eta}|S_C|^2 Q_{ns}A\eta\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{C.5}$$

and

$$\frac{d}{d\epsilon}\mathcal{J}_{ns}^{n}(Q_{ns}+\epsilon Q_\epsilon)\bigg|_{\epsilon=0} \approx 2\mathrm{Re}\int |S_n|^2 \overline{Q_\epsilon} Q_{ns}\eta\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \tag{C.6}$$

Recall that

$$\mathcal{J}_{ns}(Q_{ns}) = \mathcal{J}_{ns}^{T}(Q_{ns}) + \mathcal{J}_{ns}^{C}(Q_{ns}) + \mathcal{J}_{ns}^{n}(Q_{ns}). \tag{C.7}$$

Thus, using (C.4), (C.5) and (C.6),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}(Q_{ns}+\epsilon Q_{\epsilon})\bigg|_{\epsilon=0} \approx 2\mathrm{Re}\int \overline{Q_{\epsilon}}\bigg(\int \overline{A}\eta\left(Q_{ns}A\eta(|S_T|^2+|S_C|^2)-|S_T|^2\tilde{\chi}\right)+ \\ \times |S_n|^2 Q_{ns}\eta\bigg)\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{C.8}$$

In order for the right hand side of (C.8) to be 0 for all Q_{ϵ},the second integrand inside the parenthesis needs to be zero.

$$\int \overline{A}\eta\left(Q_{ns}A\eta(|S_T|^2+|S_C|^2)-|S_T|^2\tilde{\chi}\right)+|S_n|^2Q_{ns}\eta\ \ \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}=0 \tag{C.9}$$

which reduces to

$$\overline{A}\eta(Q_{ns}A\eta-\tilde{\chi}_{\Omega})|S_T|^2+\overline{A}\eta Q_{ns}A\eta|S_C|^2+Q_{ns}\eta|S_n|^2=0. \tag{C.10}$$

From (C.10) the filter Q_{ns} can be explicitly written as follows:

$$Q_{ns}(\boldsymbol{z},\boldsymbol{\xi}) = Q_{ns}^{1}(\boldsymbol{z},\boldsymbol{\xi})Q_{ns}^{2}(\boldsymbol{z},\boldsymbol{\xi}) \tag{C.11}$$

where

$$Q_{ns}^{1}(\boldsymbol{z},\boldsymbol{\xi}) = \frac{\overline{A(\boldsymbol{z},\boldsymbol{\xi})}\tilde{\chi}_{\Omega}}{|A(\boldsymbol{z},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})}, \tag{C.12}$$

$$Q_{ns}^{2}(\boldsymbol{z},\boldsymbol{\xi}) = \frac{|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2}{|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2+|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2+\frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{z},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})}} \tag{C.13}$$

where Q_{ns}^{1} is the same filter derived in [1]. $\Omega_{\boldsymbol{z}}$ is the data collection manifold given by (3.34) $|S_T(\boldsymbol{z},\boldsymbol{\xi})|^2$ and $|S_C(\boldsymbol{z},\boldsymbol{\xi})|^2$ are the SVSD functions defined in (3.11) and (3.13). $|S_n(\boldsymbol{\xi})|^2$ is the noise power spectral density function defined in (3.17).

(ii) Using the approximation obtained in (3.39) for $\mathcal{K}_{bp}\mathcal{F}$ and following the steps in (C.1)-(C.10), the leading-order singularities of the filter Q_{bp} that minimizes the (leading-order) MSE $\mathcal{J}_{bp}(\hat{Q}_{bp})$ is given by

$$Q_{bp}(\boldsymbol{z},\boldsymbol{x}') = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{z}-\boldsymbol{x}')\cdot\boldsymbol{\zeta}} Q_{ns}(\boldsymbol{z},\boldsymbol{\zeta})\mathrm{d}\boldsymbol{\zeta}. \tag{C.14}$$

(iii) We calculate the MSE between the true and the reconstructed target scenes with filter given in (3.44) as follows

$$\begin{aligned}\mathcal{J}(Q_{ns}) =& \mathcal{J}_T(Q_{ns}) + \mathcal{J}_C(Q_{ns}) + \mathcal{J}_n(Q_{ns}) \qquad (C.15)\\ \approx& \int \left|Q_{ns}(\boldsymbol{x},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi}) - \tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi})\right|^2 |S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\\ &+ \int \left|Q_{ns}(\boldsymbol{x},\boldsymbol{\xi})A(\boldsymbol{x},\boldsymbol{\xi})\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})\right|^2 |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\\ &+ \int |Q_{ns}(\boldsymbol{z},\boldsymbol{\xi})|^2 |S_n(\boldsymbol{\xi})|^2 \eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \qquad (C.16)\end{aligned}$$

Inserting $Q_{ns} = Q_{ns}^1 Q_{ns}^2$ into (C.16), we write

$$\begin{aligned}\mathcal{J}_{ns}(Q_{ns}) \approx& \int \left|Q_{ns}^2(\boldsymbol{x},\boldsymbol{\xi}) - \tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi})\right|^2 |S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\\ &+ \int \left|Q_{ns}^2(\boldsymbol{x},\boldsymbol{\xi})\right|^2 |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}\\ &+ \int |Q_{ns}^2(\boldsymbol{z},\boldsymbol{\xi})|^2 \frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \qquad (C.17)\end{aligned}$$

We expand the first term in (C.17) and write

$$\begin{aligned}\mathcal{J}_{ns}(Q_{ns}) \approx \int \Bigg\{ &|Q_{ns}^2(\boldsymbol{x},\boldsymbol{\xi})|^2 \left(|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 + |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 + \frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2\eta(\boldsymbol{z},\boldsymbol{z},\boldsymbol{\xi})} \right)\\ &+ \left(\left|\tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi})\right|^2 - 2Q_{ns}^2(\boldsymbol{x},\boldsymbol{\xi})\tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi}) \right) |S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 \Bigg\} \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{C.18}$$

Inserting (C.13) into (C.18), we obtain

$$\mathcal{J}(Q_{ns}) = \mathcal{J}(Q_{ns}^1 Q_{ns}^2) = \int \frac{\alpha(\boldsymbol{x},\boldsymbol{\xi})|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2}{|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 + \alpha(\boldsymbol{x},\boldsymbol{\xi})} \tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{C.19}$$

where

$$\alpha(\boldsymbol{x},\boldsymbol{\xi}) = |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 + \frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})}. \tag{C.20}$$

We can obtain the MSE for BPF method by replacing Q_{ns} with $\hat{Q}_{bp}$ which results in the same MSE up to the leading-order.

We investigate the statistics of the reconstructed images in detail in Appendix D.

APPENDIX D
STATISTICS OF THE RECONSTRUCTED IMAGES

In this section, we investigate the deviation of the reconstructed target scene from the true scene and calculate a compensation term. In Appendix C it was shown that

$$\mathcal{J}(Q_{ns}) = \mathcal{J}(Q_{ns}^1 Q_{ns}^2) = \int \frac{\alpha(\boldsymbol{x},\boldsymbol{\xi})|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2}{|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 + \alpha(\boldsymbol{x},\boldsymbol{\xi})} \tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{x}, \tag{D.1}$$

and inserting only Q_{ns}^1 into $\mathcal{J}(Q_{ns})$

$$\mathcal{J}_{ns}^1(Q_{ns}^1) = \int \alpha(\boldsymbol{x},\boldsymbol{\xi}) \tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{x}. \tag{D.2}$$

are the MSE between true and the reconstructed target scenes with (C.12) and (C.13), respectively where

$$\alpha(\boldsymbol{x},\boldsymbol{\xi}) = |S_C(\boldsymbol{x},\boldsymbol{\xi})|^2 + \frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2 \eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})}. \tag{D.3}$$

We assume that the Fourier transform of the windowing function $\hat{\Phi}$ is 1. It should be noted here that $T_{ns}(\boldsymbol{z})$ is an MMSE for $T(\boldsymbol{x})$ and (C.19) is the variance of this estimator.

Equations (C.19) and (D.2) show that an increase in clutter or noise level results in a greater error for filter Q_{ns}^1 compared to Q_{ns}.

It is important to note that these error terms are valid with the assumption that the full knowledge of the target and clutter SVSD functions are available. However, as it is mentioned in earlier chapters, we form an image of the target/clutter scenes to estimate the target and the clutter SVSD functions. Therefore, the variance of the MMSE for the non-stationary reconstruction method described in Chap-

ter 3 is:

$$\mathcal{J}_{ns}(Q_{ns}^1, Q_{ns}^2) = \int \mathrm{E}\left[|T_{ns}(\boldsymbol{z}) - T(\boldsymbol{z})|^2\right] \mathrm{d}\boldsymbol{z} \tag{D.4}$$

$$= \int \mathrm{E}\left[|\mathcal{K}_{ns}\mathcal{F}[T_1 + C_1](\boldsymbol{z}) + \mathcal{K}_{ns}[n](\boldsymbol{z}) - T(\boldsymbol{z})|^2\right] \mathrm{d}\boldsymbol{z} \tag{D.5}$$

$$\begin{aligned} = \int \mathrm{E}\big[&|\mathcal{K}_{ns}\mathcal{F}[T_1](\boldsymbol{z}) - T(\boldsymbol{z})|^2 + |\mathcal{K}_{ns}\mathcal{F}[C_1](\boldsymbol{z})|^2 + |\mathcal{K}_{ns}[n](\boldsymbol{z})|^2 \\ &- 2\mathrm{Re}\{\mathcal{K}\mathcal{F}(\mathcal{K}\mathcal{F})^\dagger[T_1\bar{C}_1]\} - 2\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_{ns}^\dagger[T_1\bar{n}]\} \\ &- 2\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_{ns}^\dagger[C_1\bar{n}]\}\big]\mathrm{d}\boldsymbol{z} \end{aligned} \tag{D.6}$$

where $T_1(\boldsymbol{z}) = \mathcal{K}_1\mathcal{F}[T+C] + \mathcal{K}_1[n]$ and $C_1(\boldsymbol{z}) = \mathcal{K}_1\mathcal{F}[C] + \mathcal{K}_1[n]$ and $\mathcal{K}_1$ is the FBP operator with filter Q_{ns}^1. Now we write

$$\begin{aligned} \int \mathrm{E}\left[|\mathcal{K}_{ns}\mathcal{F}[T_1](\boldsymbol{z}) - T(\boldsymbol{z})|^2\right] \mathrm{d}\boldsymbol{z} = \int \mathrm{E}\big[&\mathcal{K}_{ns}\mathcal{F}T_1\bar{T}_1\mathcal{F}^\dagger\mathcal{K}_{ns}^\dagger \\ &- 2\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}[T_1\bar{T}]\} + T\bar{T}\big]\mathrm{d}\boldsymbol{z}. \end{aligned} \tag{D.7}$$

For the first term of (D.7) we write:

$$\int \mathrm{E}\left[\mathcal{K}_{ns}\mathcal{F}T_1\bar{T}_1\mathcal{F}^\dagger\mathcal{K}_{ns}^\dagger\right] \mathrm{d}\boldsymbol{z} = \int |Q_{ns}Q_1|A\eta|^2|^2(|S_T|^2 + |S_C|^2 + |S_n|^2/|A|^2\eta)\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \tag{D.8}$$

For the second term of (D.7) we have,

$$\int \mathrm{E}\left[\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}[T_1\bar{T}]\}\right] \mathrm{d}\boldsymbol{z}$$

$$= \int \mathrm{E}\left[\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_1\mathcal{F}[T\bar{T}] + \mathcal{K}_{ns}\mathcal{F}\mathcal{K}_1\mathcal{F}[C\bar{T}] + \mathcal{K}_{ns}\mathcal{F}\mathcal{K}_1[n\bar{T}]\}\right] \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \tag{D.9}$$

$$= \int \mathrm{Re}\{Q_{ns}\bar{Q}_1|A\eta|^2|S_T|^2\}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \tag{D.10}$$

since the target is uncorrelated with noise and clutter. For the fourth term of (D.4)

we write

$$\int \mathrm{E}\left[\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}(\mathcal{K}_{ns}\mathcal{F})^{\dagger}[T_1\bar{C}_1]\}\right]\mathrm{d}\boldsymbol{z} = \int |Q_{ns}Q_1|^2|A\eta|^4(|S_C|^2+|S_n|^2/|A|^2\eta)\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}, \tag{D.11}$$

for the fifth term of (D.4)

$$\int \mathrm{E}\left[\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_{ns}^{\dagger}[T_1\bar{n}]\}\right]\mathrm{d}\boldsymbol{z} = \int \mathrm{Re}\{|Q_{ns}|^2\bar{Q}_1\eta|A\eta|^2|S_n|^2\}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \tag{D.12}$$

and for the sixth term of (D.4) we have

$$\int \mathrm{E}\left[\mathrm{Re}\{\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_{ns}^{\dagger}[C_1\bar{n}]\}\right]\mathrm{d}\boldsymbol{z} = \int \mathrm{Re}\{|Q_{ns}|^2\bar{Q}_1\eta|A\eta|^2|S_n|^2\}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \tag{D.13}$$

Similarly, we write

$$\int \mathrm{E}\left[|\mathcal{K}_{ns}\mathcal{F}[C_1](\boldsymbol{z})|^2\right]\mathrm{d}\boldsymbol{z} = \int \mathrm{E}\left[|\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_1\mathcal{F}[C]+\mathcal{K}_{ns}\mathcal{F}\mathcal{K}_1[n]|^2\right]\mathrm{d}\boldsymbol{z} \tag{D.14}$$

$$= \int |Q_{ns}Q_1|^2|A\eta|^4(|S_C|^2+|S_n|^2/|A|^2\bar{\eta})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \tag{D.15}$$

and finally we write

$$\int \mathrm{E}\left[|\mathcal{K}_{ns}[n](\boldsymbol{z})|^2\right]\mathrm{d}\boldsymbol{z} \approx \int |Q_{ns}|^2|S_n|^2\eta\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \tag{D.16}$$

Thus, $\mathcal{J}_{ns}$ is given by

$$\begin{aligned}\mathcal{J}_{ns}(Q_{ns}^1,Q_{ns}^2) \approx & \int |Q_{ns}Q_1|A\eta|^2|^2(|S_T|^2+|S_C|^2+|S_n|^2/|A|^2\eta)\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \\ & -2\int \mathrm{Re}\{Q_{ns}\bar{Q}_1|A\eta|^2|S_T|^2\}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}+\int |S_T|^2\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \\ & +\int |Q_{ns}Q_1|A\eta|^2|^2(|S_C|^2+|S_n|^2/\bar{\eta})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}+\int |Q|^2|S_n|^2\eta\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \\ & -2\int \mathrm{Re}\{|Q_{ns}Q_1|A\eta|^2|^2(|S_C|^2+|S_n|^2/|A|^2\bar{\eta})\}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z} \\ & -4\int \mathrm{Re}\{|Q_{ns}|^2\bar{Q}_1\eta|A\eta|^2|S_n|^2\}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{z}. \end{aligned} \tag{D.17}$$

Therefore, we have

$$\mathcal{J}_{ns}(Q_{ns}) \approx \int \frac{|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2(|\alpha(\boldsymbol{x},\boldsymbol{\xi})|^2 - 3\frac{|S_T(\boldsymbol{x},\boldsymbol{\xi})S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})})}{(|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2 + \alpha(\boldsymbol{x},\boldsymbol{\xi}))^2}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{D.18}$$

where $T_{ns}(\boldsymbol{z})$ is the image of the target scene reconstructed by (C.13) by using the SVSD functions formed by (C.12).

Integrand inside (D.18) can be considered as the SVSD function of the error term due to not using the true target and clutter SVSD functions, i.e.

$$T_{ns}(\boldsymbol{z}) = T(\boldsymbol{z}) + \varepsilon_{ns}(\boldsymbol{z}) \quad \rightarrow \quad |S_{T_{ns}}(\boldsymbol{\xi})|^2 = |S_T(\boldsymbol{\xi})|^2 + |S_{\varepsilon_{ns}}(\boldsymbol{\xi})|^2. \tag{D.19}$$

The error given in (C.19) can be reduced by compensating for the estimated target SVSD function with $|S_{\varepsilon_{ns}}(\boldsymbol{\xi})|^2$. Thus, the best possible image that can be reconstructed after compensating for the error has the variance

$$\int \frac{\alpha(\boldsymbol{x},\boldsymbol{\xi})|S_T(\boldsymbol{x},\boldsymbol{\xi})|^4}{|S_T(\boldsymbol{x},\boldsymbol{\xi})|^4 + \frac{\alpha(\boldsymbol{x},\boldsymbol{\xi})(|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2+\alpha(\boldsymbol{x},\boldsymbol{\xi}))^2}{|S_T(\boldsymbol{x},\boldsymbol{\xi})|^2+2\alpha(\boldsymbol{x},\boldsymbol{\xi})+3\frac{|S_n(\boldsymbol{\xi})|^2}{|A(\boldsymbol{x},\boldsymbol{\xi})|^2\eta(\boldsymbol{x},\boldsymbol{x},\boldsymbol{\xi})}}\tilde{\chi}_\Omega(\boldsymbol{x},\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{x} \tag{D.20}$$

where α is given by (D.3), $\tilde{\chi}_\Omega$ is a smooth function that prevents division by 0, $|S_n|^2$ is the noise spectral density function, and $|S_T|^2$, $|S_C|^2$, are SVSD functions for target and clutter, respectively.

APPENDIX E
ESTIMATION OF THE SVSD FUNCTIONS

In this section, we briefly describe a method introduced in [41], [83] for the spectral density function estimation of non-stationary processes. The method can be viewed as a straightforward extension of the spectral density function estimation for stationary processes.

Theorem 3: Let $\Phi(\boldsymbol{x})$ be a square integrable, compactly supported windowing function satisfying the conditions in [83, eq. (7.5), eq. (7.6)] and $f(\boldsymbol{x})$ denote a realization of a zero-mean pseudo-stationary process with the SVSD function $|S_f(\boldsymbol{x}, \boldsymbol{\xi})|^2$. Let β_f and β_ϕ denote the maximum intervals over which f and Φ can be considered as stationary, respectively [83, eq. (7.3)]. We define:

$$U_{\boldsymbol{\xi}_0}(\boldsymbol{x}) = \int_{\Omega_\Phi} \Phi(\boldsymbol{u}) f(\boldsymbol{x} - \boldsymbol{u}) \mathrm{e}^{-\mathrm{i}\boldsymbol{\xi}_0 \cdot (\boldsymbol{x} - \boldsymbol{u})} \mathrm{d}\boldsymbol{u} \tag{E.1}$$

where

$$\beta_\phi \ll \beta_f \ll \Omega_\Phi \tag{E.2}$$

and $\boldsymbol{u} \in \mathbb{R}^2$, Ω_Φ is the support of the windowing function and $\boldsymbol{\xi}_0$ is any constant frequency. We assume that for each pixel $\boldsymbol{x}_0$, the neighborhood of $f(\boldsymbol{x}_0)$, $\boldsymbol{x}_0 \in \Omega_\Phi$ is stationary, i.e. $f(\boldsymbol{x})$ is locally stationary for each $\boldsymbol{x}$. Then, $U(\boldsymbol{x})$ represents the Fourier transform of every locally stationary neighborhood of $f(\boldsymbol{x})$. Following [41] we write

$$\int \mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x} + \tau) \mathrm{e}^{\mathrm{i}2\pi\boldsymbol{\zeta}\cdot\tau} \mathrm{d}\tau \approx |\hat{\Phi}(\boldsymbol{\zeta})|^2 |S_f(\boldsymbol{x}, \boldsymbol{\zeta})|^2, \tag{E.3}$$

i.e. Fourier transform of $\mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x} + \tau) = \mathbf{E}\left[U_{\boldsymbol{\xi}_0}(\boldsymbol{x}) \overline{U_{\boldsymbol{\xi}_0}(\boldsymbol{x} + \tau)}\right]$ is an estimator for the space-varying spectral density function of $f(\boldsymbol{x})$ where $\hat{\Phi}(\boldsymbol{\xi})$ is the Fourier transform of $\Phi(\boldsymbol{x})$ normalized to 1.

Proof: To prove (E.3), we adapt the estimation method for semi-stationary pro-

cesses presented in [41] into our model. Let $f(\boldsymbol{x})$ denote a zero-mean pseudo-stationary process. $f(\boldsymbol{x})$ can be expressed in the mean square sense as

$$f(\boldsymbol{x}) = \int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} S_f(\boldsymbol{x},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\mathrm{dB}(\boldsymbol{x}'). \tag{E.4}$$

Using (3.8), (3.48) can be written as

$$U_{\boldsymbol{\xi}_0}(\boldsymbol{x}) = \int \Phi(\boldsymbol{u})\mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}'-\boldsymbol{u})\cdot\boldsymbol{\xi}_0}\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}'-\boldsymbol{u})\cdot\boldsymbol{\xi}} S_f(\boldsymbol{x}-\boldsymbol{u},\boldsymbol{\xi})\mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{x}'\cdot\boldsymbol{\xi}_0}\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\mathbf{B}(\boldsymbol{x}'). \tag{E.5}$$

Next, $\Psi_{\boldsymbol{x},\lambda}(\theta)$ is defined such that

$$\Psi_{\boldsymbol{x},\lambda}(\theta) = \int \Phi(\boldsymbol{u})\frac{S_f(\boldsymbol{x}-\boldsymbol{u},\lambda)}{S_f(\boldsymbol{x},\lambda)}\mathrm{e}^{-\mathrm{i}2\pi\theta\cdot\boldsymbol{u}}\mathrm{d}\boldsymbol{u}. \tag{E.6}$$

We rearrange the terms in (E.5)

$$\begin{aligned} U_{\boldsymbol{\xi}_0}(\boldsymbol{x}) = \int \Phi(\boldsymbol{u})\frac{S_f(\boldsymbol{x}-\boldsymbol{u},\boldsymbol{\xi})}{S_f(\boldsymbol{x},\boldsymbol{\xi})}\mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{\xi}-\boldsymbol{\xi}_0)\cdot\boldsymbol{u}}\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{\xi}-\boldsymbol{\xi}_0)\cdot\boldsymbol{u}} \\ \times S_f(\boldsymbol{x},\boldsymbol{\xi})\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}'-\boldsymbol{u})\cdot(\boldsymbol{\xi}-\boldsymbol{\xi}_0)}\mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{x}'\cdot\boldsymbol{\xi}_0}\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\mathbf{B}(\boldsymbol{x}'). \end{aligned} \tag{E.7}$$

We make the change of variables $\boldsymbol{\xi} = \boldsymbol{\xi} + \boldsymbol{\xi}_0$ and use the definition (E.6)

$$U_{\boldsymbol{\xi}_0}(\boldsymbol{x}) = \int \Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) S_f(\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0)\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}}\mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{x}'\cdot\boldsymbol{\xi}_0}\mathrm{d}\boldsymbol{\xi}\mathrm{d}\mathbf{B}(\boldsymbol{x}'). \tag{E.8}$$

Next, we calculate $\mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x},\boldsymbol{x}') = \mathbf{E}[U_{\boldsymbol{\xi}_0}(\boldsymbol{x})\overline{U_{\boldsymbol{\xi}_0}(\boldsymbol{x}')}]$ as follows:

$$\begin{aligned} \mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x},\boldsymbol{x}') =& \mathbf{E}[U_{\boldsymbol{\xi}_0}(\boldsymbol{x})\overline{U_{\boldsymbol{\xi}_0}(\boldsymbol{x}')}] \\ =& \int \Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) S_f(\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0)\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}'')\cdot\boldsymbol{\xi}}\mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{x}''\cdot\boldsymbol{\xi}_0}\overline{\Psi_{\boldsymbol{x}',\boldsymbol{\xi}'+\boldsymbol{\xi}_0}(\boldsymbol{\xi})} \\ & \times \overline{S_f(\boldsymbol{x}',\boldsymbol{\xi}'+\boldsymbol{\xi}_0)}\mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{x}''')\cdot\boldsymbol{\xi}'}\mathrm{e}^{\mathrm{i}2\pi\boldsymbol{x}'''\cdot\boldsymbol{\xi}_0}\mathbf{E}[\mathrm{d}\mathbf{B}(\boldsymbol{x}'')\overline{\mathrm{d}\mathbf{B}(\boldsymbol{x}''')}]\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\xi}'. \end{aligned} \tag{E.9}$$

Using $\mathbf{E}[\mathrm{d}\mathbf{B}(\boldsymbol{x}')\mathrm{d}\mathbf{B}(\boldsymbol{x}'')] = \delta(\boldsymbol{x}'' - \boldsymbol{x}''')\mathrm{d}\boldsymbol{x}'''\mathrm{d}\boldsymbol{x}''$, (E.9) can be written as

$$\begin{aligned}\mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x}') = \int \Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) S_f(\boldsymbol{x}, \boldsymbol{\xi}+\boldsymbol{\xi}_0) \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}'')\cdot\boldsymbol{\xi}} \mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{x}''\cdot\boldsymbol{\xi}_0} \overline{\Psi_{\boldsymbol{x}',\boldsymbol{\xi}'+\boldsymbol{\xi}_0}(\boldsymbol{\xi})} \\ \times \overline{S_f(\boldsymbol{x}', \boldsymbol{\xi}'+\boldsymbol{\xi}_0)} \mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{x}'-\boldsymbol{x}''')\cdot\boldsymbol{\xi}'} \mathrm{e}^{\mathrm{i}2\pi\boldsymbol{x}'''\cdot\boldsymbol{\xi}_0} \delta(\boldsymbol{x}''-\boldsymbol{x}''')\mathrm{d}\boldsymbol{x}''\mathrm{d}\boldsymbol{x}'''\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\xi}' \end{aligned} \tag{E.10}$$

$$\begin{aligned}= \int \Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) S_f(\boldsymbol{x}, \boldsymbol{\xi}+\boldsymbol{\xi}_0) \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}\cdot\boldsymbol{\xi}-\boldsymbol{x}'\cdot\boldsymbol{\xi}')} \overline{\Psi_{\boldsymbol{x}',\boldsymbol{\xi}'+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) S_f(\boldsymbol{x}', \boldsymbol{\xi}'+\boldsymbol{\xi}_0)} \\ \times \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{\xi}-\boldsymbol{\xi}')\cdot\boldsymbol{x}''} \mathrm{d}\boldsymbol{x}''\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\xi}' \end{aligned} \tag{E.11}$$

$\int \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{\xi}-\boldsymbol{\xi}')\cdot\boldsymbol{x}''}\mathrm{d}\boldsymbol{x}''\mathrm{d}\boldsymbol{\xi}' = \delta(\boldsymbol{\xi}-\boldsymbol{\xi}')\mathrm{d}\boldsymbol{\xi}'$, thus

$$\mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x}') = \int \Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) \overline{\Psi_{\boldsymbol{x}',\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi})} S_f(\boldsymbol{x}, \boldsymbol{\xi}+\boldsymbol{\xi}_0) \overline{S_f(\boldsymbol{x}', \boldsymbol{\xi}+\boldsymbol{\xi}_0)} \mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{\xi}} \mathrm{d}\boldsymbol{\xi} \tag{E.12}$$

Now we let $\boldsymbol{x}' = \boldsymbol{x} + \tau$ and calculate the Fourier transform

$$\begin{aligned}\int \mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x}+\tau) \mathrm{e}^{\mathrm{i}2\pi\boldsymbol{\zeta}\cdot\tau} \mathrm{d}\tau = \int \Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi}) \overline{\Psi_{\boldsymbol{x}+\tau,\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi})} \mathrm{e}^{\mathrm{i}2\pi\tau\cdot(\boldsymbol{\zeta}-\boldsymbol{\xi})} \\ \times S_f(\boldsymbol{x}, \boldsymbol{\xi}+\boldsymbol{\xi}_0) \overline{S_f(\boldsymbol{x}+\tau, \boldsymbol{\xi}+\boldsymbol{\xi}_0)} \mathrm{d}\boldsymbol{\xi}\mathrm{d}\tau. \end{aligned} \tag{E.13}$$

By using $\Psi_{\boldsymbol{x},\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi})\overline{\Psi_{\boldsymbol{x}+\tau,\boldsymbol{\xi}+\boldsymbol{\xi}_0}(\boldsymbol{\xi})} \approx |\hat{\Phi}(\boldsymbol{\xi})|^2$ and following a similar argument in [25], we assume that $S_f(\boldsymbol{x}, \boldsymbol{\xi}+\boldsymbol{\xi}_0)$ decays slowly for large $\boldsymbol{\xi}$. Then, the integration in (E.13) has singularities at $\boldsymbol{\xi} = \boldsymbol{\zeta}$ and we write

$$\int \mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x}+\tau) \mathrm{e}^{\mathrm{i}2\pi\boldsymbol{\zeta}\cdot\tau} \mathrm{d}\tau \approx |\hat{\Phi}(\boldsymbol{\zeta})|^2 |S_f(\boldsymbol{x}, \boldsymbol{\zeta}+\boldsymbol{\xi}_0)|^2, \tag{E.14}$$

i.e. Fourier transform of $\mathcal{R}_{U_{\boldsymbol{\xi}_0}}(\boldsymbol{x}, \boldsymbol{x})$ is an (approximately) unbiased estimator for space-varying spectral density function of the target for frequencies around $\boldsymbol{\xi}_0$ for each pixel $\boldsymbol{x}$.

APPENDIX F
A SHORT REVIEW ON BLUE

Let $\mathbf{x} = [x[0], x[1], \ldots, x[N-1]]$ denote a dataset corresponding to measurements of the original signal $\boldsymbol{s} = [s[0], s[1], \ldots, s[N-1]]$ given by

$$x[n] = \theta s[n] + w[n] \tag{F.1}$$

where w is the disturbance term.

Let $p(\mathbf{x}; \theta)$ denote the probability distribution function (PDF) of this random variable depending on an unknown variable θ. Finally, let $\hat{\theta}$ denote an estimator for the parameter θ.

Bias and variance of an estimator generally gives a good idea whether an estimator is a good one or not. The bias and the variance of the estimator $\hat{\theta}$ are defined as follows:

$$\text{BIAS}(\hat{\theta}) := \text{E}[\hat{\theta}] - \theta, \tag{F.2}$$

$$\text{VAR}(\hat{\theta}) := \text{E}\left[\|\hat{\theta} - \text{E}[\hat{\theta}]\|_2^2\right]. \tag{F.3}$$

The mean-square-error (MSE) is a widely-used figure of merit to determine the performance of an estimator and it is defined as

$$\text{MSE}(\hat{\theta}) := \text{E}\left[\|\hat{\theta} - \theta\|_2^2\right]. \tag{F.4}$$

MSE can be expanded as

$$\text{MSE}(\hat{\theta}) = \text{E}\left[\|\hat{\theta} - \theta\|_2^2\right] \tag{F.5}$$

$$= \|\text{BIAS}(\hat{\theta})\|_2^2 + \text{VAR}(\hat{\theta}) \tag{F.6}$$

which shows the well-known trade-off between the bias and the variance of an estimator. (F.6) shows that having a zero-bias, zero-variance estimator is not possible

almost always. Therefore, one needs to design the estimator by considering the fact that if the estimator has zero-bias it will have large variance and vice verse. From a machine learning point of view, (F.6) can be interpreted as the fact that a training dataset can be underfit (low variance, large bias) or overfit (low bias, large variance) [153].

Having low-bias results in more generalizable estimators. Moreover, often times only the bias depends on the unknown parameters. Therefore, it is usually more desirable to design estimators that has zero bias and the minimum variance. This is also equivalent to finding an unbiased estimator that minimizes the MSE among all other estimators.

An estimator $\hat{\theta}$ is said to be a minimum variance unbiased estimator (MVUE) if [154]

- $\mathrm{E}[\hat{\theta}] = \theta$, $\forall \theta \in \Theta$,
- If $\mathrm{E}[\tilde{\hat{\theta}}] = \theta$, $\forall \theta \in \Theta$, then $\mathrm{VAR}(\hat{\theta}) \leq \mathrm{VAR}(\tilde{\hat{\theta}})$, $\forall \theta \in \Theta$.

where Θ is the set that includes all the estimators θ. As specified in the definition of MVUE, one needs to look through all possible estimators to find the minimum variance estimator. Therefore, it may not be possible to find the MVUE as PDF or the model may not be known completely. As a solution to this problem, a suboptimal estimator may be restricted to be a linear one. Then the linear estimator, that has the minimum variance is called the best linear unbiased estimator (BLUE).

As stated earlier there are three conditions for an estimator to be BLUE. For measurements $\boldsymbol{x}$:

1. Estimator needs to be linear:

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n]. \tag{F.7}$$

   provided that $\sum_{n=0}^{N-1} a_n = 1$

2. Bias needs to be 0:

$$\mathrm{E}[\hat{\theta}] = \sum_{n=0}^{N-1} a_n \mathrm{E}[x[n]] = \theta. \tag{F.8}$$

3. Estimator must be a solution to the problem

$$\min_{\theta} \mathrm{VAR}(\hat{\theta}) = \min_{\theta} \mathrm{E}\left[\left(\sum_{n=0}^{N-1} a_n x[n] - \mathrm{E}\left[\sum_{n=0}^{N-1} a_n x[n]\right]\right)^2\right]. \tag{F.9}$$

Using $\mathbf{a} = [a_0, a_1, \ldots, a_{N-1}]^T$, (F.9) can be written as

$$\begin{aligned}
\mathrm{VAR}(\hat{\theta}) =& \mathrm{E}[(\mathbf{a}^T\mathbf{x} - \mathbf{a}^T\mathrm{E}[\mathbf{x}])^2] && \text{(F.10)}\\
=& \mathrm{E}[\mathbf{a}^T(\mathbf{x} - \mathbf{a}^T\mathrm{E}[\mathbf{x}])(\mathbf{x} - \mathbf{a}^T\mathrm{E}[\mathbf{x}])^T\mathbf{a}] && \text{(F.11)}\\
=& \mathbf{a}^T\mathbf{C}\mathbf{a} && \text{(F.12)}
\end{aligned}$$

where $\mathbf{C} = \mathrm{E}[(\mathbf{x} - \mathbf{a}^T\mathrm{E}[\mathbf{x}])(\mathbf{x} - \mathbf{a}^T\mathrm{E}[\mathbf{x}])^T]$. Therefore, the problem given in (F.9) becomes identical to finding $\mathbf{a}$ that minimizes $\mathbf{a}^T\mathbf{C}\mathbf{a}$. To minimize variance of the estimator $\mathbf{a}^T\mathbf{C}\mathbf{a}$ subject to the constraint $\mathbf{a}^T\mathbf{s}$ we use Lagrange multipliers

$$J(\mathbf{a}, \lambda) = \mathbf{a}^T\mathbf{C}\mathbf{a} - \lambda(\mathbf{a}^T\mathbf{s} - 1). \tag{F.13}$$

We calculate the derivative with respect to $\mathbf{a}$ and set it equal to zero:

$$\frac{\partial}{\partial \mathbf{a}} J(\mathbf{a}, \lambda) = 0 = 2\mathbf{C}\mathbf{a}_* + \lambda\mathbf{s} \rightarrow \mathbf{a}_* = \lambda\frac{\mathbf{C}^{-1}\mathbf{s}}{2} \tag{F.14}$$

where $\mathbf{a}_*$ is a vector containing the optimal weights for the estimator. Using the constraint equation

$$\mathbf{a}^T\mathbf{s} = 1 \rightarrow \lambda\frac{\mathbf{s}^T\mathbf{C}^{-1}}{2}\mathbf{s} = 1 \tag{F.15}$$

and λ is given by

$$\lambda = \frac{2}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}. \tag{F.16}$$

Therefore, the BLUE is given by [154]

$$\hat{\theta} = \frac{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} \tag{F.17}$$

with the minimum variance

$$\mathrm{VAR}(\hat{\theta}) = \frac{1}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} \tag{F.18}$$

and the bias

$$\mathrm{E}[\hat{\theta}] = \frac{\mathbf{s}^T\mathbf{C}^{-1}\mathrm{E}[\mathbf{x}]}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} = \frac{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} = \theta. \tag{F.19}$$

Let $\hat{\theta} = q_2\mathbf{q}_1\mathbf{x}$, then we have

$$\mathbf{q}_1 = \mathbf{s}^T\mathbf{C}^{-1} \tag{F.20}$$

$$q_2 = \frac{1}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} = \frac{\lambda}{2}. \tag{F.21}$$

With vector parameters, if the data model is given by

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \tag{F.22}$$

then the BLUE is

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x} \tag{F.23}$$

with the minimum variance

$$\mathcal{C}_{\hat{\boldsymbol{\theta}}} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1} \tag{F.24}$$

where $\mathbf{C}$ is the covariance matrix of the noise. This is also known as the Gauss-

Markov Theorem [154] and if $\mathbf{w}$ is Gaussian then BLUE is MVUE.

Similar to the scalar case, if we let $\hat{\boldsymbol{\theta}} = \mathbf{Q}_2\mathbf{Q}_1\mathbf{x}$

$$\mathbf{Q}_1 = \mathbf{H}^T\mathbf{C}^{-1} \tag{F.25}$$

$$\mathbf{Q}_2 = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}. \tag{F.26}$$

If MVUE is linear then BLUE given in (F.17) is indeed the MVUE and it only requires the first and second order statistics of the given data.

APPENDIX G
CALCULATION OF THE HESSIANS FOR CHAPTER 4

G.1 Hessian Matrix for (4.38)

Hessian for the phase function

$$\phi_1(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{z}, \boldsymbol{\zeta}, \boldsymbol{\xi}) = (\boldsymbol{x}' - \boldsymbol{z}) \cdot \boldsymbol{\zeta} + (\boldsymbol{x} - \boldsymbol{x}') \cdot \boldsymbol{\xi} \tag{G.1}$$

is evaluated as follows

$$H_{\boldsymbol{x}',\boldsymbol{\zeta}}(\phi_1) = \begin{bmatrix} \nabla^2_{\boldsymbol{x}'^2}\phi_1 & \nabla^2_{\boldsymbol{x}',\boldsymbol{\zeta}}\phi_1 \\ \nabla^2_{\boldsymbol{\zeta},\boldsymbol{x}'}\phi_1 & \nabla^2_{\boldsymbol{\zeta}^2}\phi_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \tag{G.2}$$

$\det(H_{\boldsymbol{x}',\boldsymbol{\zeta}}(\phi_1)) = -1$ where $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

G.2 Hessian Matrix for (4.48)

Hessian for the phase function

$$\phi_2(\boldsymbol{x}, \boldsymbol{\xi}, \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{\xi}}, \boldsymbol{z}) = (\boldsymbol{x} - \boldsymbol{z}) \cdot \boldsymbol{\xi} - (\tilde{\boldsymbol{x}} - \boldsymbol{z}) \cdot \tilde{\boldsymbol{\xi}} \tag{G.3}$$

is evaluated as follows

$$H_{\boldsymbol{z},\tilde{\boldsymbol{\xi}}}(\phi_2) = \begin{bmatrix} \nabla^2_{\boldsymbol{z}^2}\phi_2 & \nabla^2_{\boldsymbol{z},\tilde{\boldsymbol{\xi}}}\phi_2 \\ \nabla^2_{\tilde{\boldsymbol{\xi}},\boldsymbol{z}}\phi_2 & \nabla^2_{\tilde{\boldsymbol{\xi}}^2}\phi_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \tag{G.4}$$

and $\det(H_{\boldsymbol{z},\tilde{\boldsymbol{\xi}}}(\phi_2)) = -1$.

G.3 Hessian Matrix for (4.53)

Hessian for the phase function

$$\phi_3(\boldsymbol{x}, \tilde{\boldsymbol{x}}, \boldsymbol{\xi}, \boldsymbol{\zeta}) = (\boldsymbol{x} - \tilde{\boldsymbol{x}}) \cdot \boldsymbol{\xi} - (\boldsymbol{x} - \tilde{\boldsymbol{x}}) \cdot \boldsymbol{\zeta} \tag{G.5}$$

is evaluated as follows

$$H_{\tilde{\boldsymbol{x}},\boldsymbol{\zeta}}(\phi_3) = \begin{bmatrix} \nabla^2_{\tilde{\boldsymbol{x}}^2}\phi_3 & \nabla^2_{\tilde{\boldsymbol{x}},\boldsymbol{\zeta}}\phi_3 \\ \nabla^2_{\boldsymbol{\zeta},\tilde{\boldsymbol{x}}}\phi_3 & \nabla^2_{\boldsymbol{\zeta}^2}\phi_3 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \tag{G.6}$$

and $\det(H_{\tilde{\boldsymbol{x}},\boldsymbol{\zeta}}(\phi_3)) = -1$.

APPENDIX H
CONDITIONS FOR THE SPARSE SIGNAL RECOVERY

Under some conditions, we can replace L^0-norm with L^p-norm and define

$$\min |\mathbf{T}|_1 \quad \text{subject to} \quad |\mathbf{d} - \mathbf{F}\mathbf{T}|_2^2 \leq \epsilon \tag{H.1}$$

where p is set to 1. (6.3) is a *convex relaxation* of (6.2) [109]. In order to guarantee the sparsest solution with S nonzero entries, $\mathbf{F}$ must satisfy the following conditions:

1. *The mutual coherence* of $\mathbf{F}$ defined by $\mu(\mathbf{F}) = \max_{i \neq j} |\mathbf{h}_i^H \mathbf{h}_j|$ should satisfy

$$|\mathbf{T}|_0 \leq \frac{1}{2}(1 + \mu^{-1}(\mathbf{F})) \tag{H.2}$$

   where the superscript H denotes the Hermitian and $\mathbf{h}_i$ is the i^{th} column of $\mathbf{F}$ [155].

2. *The restricted isometric constant* (RIC) of $\mathbf{F}$, [156], [157] denoted by $\delta_S \in (0, 1)$, must satisfy for every submatrix $\mathbf{F}_S$ composed from any S columns of $\mathbf{F}$

$$(1 - \delta_s)|\mathbf{T}|_2^2 \leq |\mathbf{F}_S \mathbf{T}|_2^2 \leq (1 + \delta_S)|\mathbf{T}|_2^2 \tag{H.3}$$

where the RIC and the mutual coherence are related by [155]–[157]

$$\delta_S \leq (S - 1)\mu(\mathbf{F}). \tag{H.4}$$