Computational Methods for Microwave Medical Imaging

A Thesis Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy

by

Qianqian Fang
Thayer School of Engineering
Dartmouth College
Hanover, New Hampshire
DATE: December 2004

Examining Committee:

Paul M. Meaney, Chairman

Keith D. Paulsen

William Lotko

Eric L. Miller

Dean of Graduate Studies

Qianqian Fang, Author

Abstract

Medical imaging methods have become increasingly important in diagnosing diseases and assisting therapeutic treatment. In particular, early detection of breast cancer is considered as a critical factor in reducing the mortality rate of women. Within the various alternative breast imaging modalities being investigated to improve breast cancer detection, microwave imaging is attractive due to the high dielectric property contrast between the cancerous and the normal tissue and has received significant interest over the last decade. The investigation into two-dimensional microwave imaging at the Thayer School of Engineering, Dartmouth College, began in the early 1990's where the first clinical microwave imaging system was brought online at the Dartmouth-Hitchcock Medical Center (DHMC) in 1999.

Although the two dimensional microwave imaging has shown great promise, the image quality is essentially compromised by the various approximations associated with operating in 2D. In this thesis, we focus on the theoretical aspects of the nonlinear tomographic image reconstruction problem with particular emphasis on developing efficient numerical algorithms for 3D microwave imaging. An incremental approach was devised to assess this progress. The concept of the dual-mesh was generalized and served as an organizing theme from which the computational efficiency of various forward field modelling methods were investigated. These methods included the 2D finite element coupled with boundary element methods and the 2D FDTD method with its ex-

tension to 3D space. Significant effort was spent on optimizing the 3D forward model in order to reconstruct images efficiently. Additional reconstruction techniques such as the adjoint method, the nodal adjoint approximation as well as a multiple-frequency dispersion reconstruction algorithm were developed to enhance both the speed and quality of the recovered images. An in-depth analysis of the Jacobian matrix was performed in the context of investigating various important factors including the resolution limit and the impact of system parameters on image quality. Additionally, a mathematical theory encompassing the properties of the phase unwrapping integral and its use with respect to our log-magnitude/phase form (LMPF) imaging algorithm was developed and discussed with particular attention to microwave scattering nulls.

Acknowledgements

I am indebted to many people toward the completion of my Ph.D. study. It is difficult to imagine that I could finish this journey without their guidance and support. Instead, because of these generous helps, I found this journey full of surprises and joyance.

I own my greatest gratitude to my advisor, professor Paul Meaney, for his patient guidance all through my Ph.D. study. Particularly, he patiently read through every paragraph of this over-sized volume, gave me many valuable suggestions in thesis writing, and helped me correct numerous English mistakes throughout the manuscript. His down-to-earth research attitude and step-by-step approach impressed me deeply.

This work would not have been possible without the help from professor Keith Paulsen. My Ph.D. research benefited greatly from his insightful suggestions on computational methods. Moreover, he is a great leader of this twenty-researchers' group, a knowledgable teacher and a tireless person. His devotion to scientific research sets an excellent paradigm for me to follow.

My gratitude also goes to many other professors for various reasons. I would like to thank professor Brian Pogue for teaching me Biomedical Imaging (ENGS167), being one of the courses that I had truly enjoyed at Thayer School. I also want to thank professor Eric Miller at Northeastern University and professor Eugene Demidenko, for their great suggestions on my thesis proposal and stimulating discussions on the statistical aspect of the image reconstruction problem. Professor William Lotko has been

on the committee for all the important events of my Ph.D study: qualification, thesis proposal and defense. I appreciate many of his valuable and constructive advices. I would also thank professor Vladimir Chernov in the Mathematics Department at Dartmouth for his informative discussions on topology when I worked on the topic of phase unwrapping. Moreover, I am deeply indebted to professor Anyao Li, professor Zaiping Nie and professor Quanzhi Xu back in China who initially guided me in the area of microwave engineering, computational electromagnetics and mathematical modelling, respectively.

I also wish to thank my wonderful team members of MIS group for their friendship and help. They are Margaret Fanning, Dun Li, Sarah Pendergrass, Colleen Fox, Shireen Geimer, Timothy Raynolds and Navin Yagnamurthy. Thanks are due also to many of my fiends at Dartmouth, especially my officemates: Xiaomei Song, Qing Feng, Heng Xu, Chao Sheng, Nirmal Soni, Subhadra Srinivasan, Kyung Park, Zhiliang Fan, Xiaodong Zhou, Xiang Li and Yijin He. The warm discussions between us on countless research/non-research topics are part of my enjoyable memory at Thayer School. Moreover, I wish to express my loving thanks to my girlfriend Yinghua Shen.

At Dartmouth-Hitchcock Medical Center (DHMC), I would like to thank Christine Kogel for coordinating all patient exams. I also want to express my greatest respect to all woman volunteers who participated the exam.

My Ph.D. study is supported by NIH/NCI grand P01-CA80139 and Herbert Darling'69 Fellowship. I deeply appreciate the providers of the scholarship for giving me this precious opportunity of top quality education and research.

Finally I would like to thank my parents and grandma-in-law. It is their endless love, support and encouragement that stimulate me to walk through the 22 years of education and to advance further. I dedicate this thesis to them.

Table of Contents

List of Figures

List of Tables

List of Symbols

$\vec{E}(\vec{r},t), \vec{H}(\vec{r},t), \vec{J}(\vec{r},t)$	time-domain electromagnetic field vectors
$\vec{E}(\vec{r}), \vec{H}(\vec{r}), \vec{J}(\vec{r})$	electromagnetic field phasor vectors ($\exp(j\omega t)$)
E_x, E_y, E_z	electromagnetic field scalar components
ϵ, σ	permittivity and conductivity,respectively
μ	magnetic permeability or linear attenuation coefficient
A, B	matrices
$a_{i,j}$	the (i,j)-th element of a matrix
U, V	orthogonal matrices
$\mathbf{v}, \mathbf{u}$	vectors
$\hat{n}, \hat{x}$	unit vectors
$\tilde{x}, \tilde{S}$	variables after certain transformation
f, g	functions
$\mathcal{F}, \mathcal{G}$	functionals
$\mathscr{F}$	Fourier transform
$\mathscr{R}$	Radon transform
ω	the angular frequency
$\mathbb{R}, \mathbb{C}$	real set and complex set, respectively
$\Delta x, \Delta t$	spatial and temporal step sizes, respectively
$\Omega, \partial\Omega$	2D or 3D domain and it's boundary
Γ	paths or curves
$\mathbf{u}_i, \mathbf{v}_i, \sigma_i$	left,right singular vectors and singular value, respectively
$N(\mu, \sigma)$	normal distribution with mean μ and variance σ
$\underset{x}{\operatorname{argmin}} \ \|f(x)\|$	argument optimization problem
J, H, G	Jacobian, Hessian and Gauss Hessian matrix, respectively
I	an identity matrix or real interval [0, 1]
Λ	a diagonal matrix

A^+	the pseudo-inverse of matrix A
α, λ, β	scalar constants
ϕ, φ	basis functions
$\boldsymbol{\beta}, \boldsymbol{\eta}$	random variables or functions
$\hat{\boldsymbol{\beta}}$	the estimator of variable $\boldsymbol{\beta}$
j	imaginary unit, $j = \sqrt{-1}$
i, j, k	spatial indices for 3D array
$O(N^2)$	Big O notation for computational complexity analysis
N_x, N_y, N_z, N_t	spatial dimensions of 3D mesh and time step number
$a_{(i)}, a^{(i)}$	quantities at the i-th iteration
$\{a_i\}, \{a_i\}_{i=1}^N$	a sequence
$\nu : \{1, \cdots, N\} \rightarrow (s, r)$	a mapping from a 1D index array to pair sequence
$\nu : s, \nu : r$	components of the ν-th element in a 1D indexed pair array
$W : X \rightarrow Y$	a mapping from space X to Y
$W^{-1}(\cdot)$	the complete inverse image (pre-image) of map W
$\mathrm{Lk}(\cdot, \cdot)$	linking number
$\mathrm{Ind}_\gamma(0)$	winding number of curve γ
$\langle \cdot, \cdot \rangle$	inner product between two vectors or functions
$\| \cdot \|_2$	l^2-norm
$\delta(\vec{r} - \vec{r}_s)$	Dirac delta function
$s_\xi, \xi = x, y, z$	coordinate stretching coefficients

List of Acronyms

Medical imaging

MWI	microwave imaging
CMI	confocal microwave imaging
CT	computed tomography
MRI	magnetic resonance imaging
US	ultrasound
EIT	electrical impedance tomography
NIR	near infrared
PET	position emission tomography
SPECT	single photon emission computed tomography
DOT	diffusion optical tomography
SAR	synthetic aperture radar

Computational electromagnetics

FDTD	finite difference-time domain method
MoM	method of moment
FE	finite element
HE	hybrid element
BE	boundary element
FMM	fast multipole method
MLFMA	multi-level fast multipole algorithm
ADI	alternating-direction implicit (FDTD)

BIM,DBIM	born iterative method and distorted BIM
Ty(2,4)	implicit 2-order in time, 4-order in space FDTD method
ABC	absorbing boundary condition
RBC	radiation boundary condition
PML	perfectly matched layer
GPML	generalized PML
UPML	uniaxial PML
CFLN	Courant-Friedrichs-Lewy number
MOT	marching-on-time
PEC	perfect electrical conductor

Others

LM	Levenberg-Marquardt method
GN	Gauss-Newton method
GCV	generalized cross-validation
LMPF	log-magnitude phase form
MFDR	multiple-frequency dispersion reconstruction
FFT	fast Fourier transform
R.V.	random variable
LHS,RHS	left-hand-side, right-hand-side
EM	electromagnetic
TE,TM	transverse electrical field and transverse magnetic field
EFIE,MFIE,CFIE	electrical/magnetic/complex field integral equations
SVD	singular value decomposition
TSVD	truncated singular value decomposition
TLS,TTLS	total least-square and truncated TLS
OLS	ordinary least-square
WLS	weighted least-square
MLE	maximum likelihood estimator
MVE	minimum variance estimator
MAP	maximum *a posteriori* estimator
s.t.	subject to
SNR	signal to noise ratio
DAQ	data acquisition

PDE	partial differential equations
RMS	root-mean-square
QMR	quasi-minimum residual method

Part I

Preliminaries

Chapter 1

Introduction to microwave imaging

1.1 Introduction

It is an indisputable fact that both the science and technology have undergone a revolutionary transformation over the last century. The creation of computers has greatly reshaped both theoretical and experimental science in that computations can bridge the gap between theories and experiments. On one side they extend the capabilities and scope of theoretical models into more practical regimes. Large and complex problems which cannot be solved by analytical mathematics can now be easily solved using numerical techniques. Additionally, low-cost and easily implemented simulations readily facilitate experimental procedures. With the bridging effect of computational capabilities, the boundaries between theoretical and experimental science becomes even more vague and their combination facilitates most of the new technological advances.

Among those benefiting from the emergence of computational science, medical imaging is the primary interest of this thesis. Since the discovery of a "new kind of ray" by Wilhelm Röntgen over a century ago, medical imaging methods have undergone explosive proliferation. It is extensively used by researchers in medicine and biology to

unveil countless mysteries of life, or used in clinics by medical practitioners to monitor normal biological activities, diagnose diseases, or guide the treatment of diseases. The integration with computational technologies dramatically improves the viability and performance of many existing imaging modalities. Meanwhile, a large number of novel imaging methods based on computational models, such as CT, MRI, PET, SPECT etc, have been introduced and developed which occupy essential positions in the modern medical imaging armament. Recent trends in medical imaging include imaging methods that utilize nonlinear physical processes or mathematical models, multi-modality imaging, integration of imaging facilities with therapeutic devices along with the growing need to be able to efficiently manage imaging data [68].

Our studies in medical imaging focus specifically on microwave imaging. From a methodology point of view, microwave imaging is a natural extension of traditional wave-based diffraction imaging methods. The creation of this technology is not an isolated event but the direct consequence of a number of related sequential accomplishments. Studies on the dielectric properties of human tissues [171], the advances in nonlinear optimization and inverse problems combined with cheaply available computing power are among the necessary requirements for facilitating the development of this technique. The fusion of these technologies not only created microwave imaging, but a series of imaging methods exploiting nonlinear processes in contrast to traditional linearized imaging modalities. Although linear phenomena have unquestionable advantages due to their simplicity in both modelling and processing, their simplicities are often conditional, i.e. they are approximate versions of more general and sophisticated non-linear processes. The arrival of the computer age significantly altered the fate of nonlinear models and provided powerful tools to describe and explore more “complicated” phenomena. As a result, linear assumptions can now be largely discarded

revealing a more realistic view of the problem. Computational methods for non-linear models not only provide a better understanding of intricate processes, but also enable the utilization of these processes to facilitate the daily life.

Given the various issues that have directly or indirectly contributed to the evolution of microwave imaging, the remainder of this chapter will discuss a range of topics that illustrate the importance of this work and summarize related previous work. In the following part of this chapter, we first present a brief introduction to cancer biology, focused especially on breast cancer, since this is the primary application of our microwave imaging efforts. A history and current status summary of microwave imaging is presented followed by the hardware and algorithmic development of tomographic microwave imaging at Dartmouth College. In the last section of this chapter, we will list the specific goals of this work along with the paths to meet these goals. Chapter 2 focuses on the mathematical aspects of the problem where 1) tomographic imaging based on linear methods, 2) nonlinear optimization techniques, 3) image reconstruction from a statistical perspective, 4) the inverse problem and 5) regularization techniques are introduced.

1.2 Cell biology of cancer and introduction to breast cancer

Based on the definition given by the American Cancer Society (ACS), cancer refers to "a group of diseases characterized by uncontrolled growth and spread of abnormal cells" [181, 180]. Cancer is the second leading cause of death of the United States and the world after cardiovascular diseases (CVD). Every year, roughly 7.1 million people die of cancer, accounting for 12.6% of all global mortalities. Cancer can develop al-

most anywhere in a human body, such as the skin, marrow, bone, brain, breast, colon, liver and lung. Cancer can also strike people at any age. Among the various cancers, roughly 76% are diagnosed for people over 55. However, in all of these cancer cases, only about 10% are genetically related and approximately 1/3 of the deaths can be completely avoided with appropriate diet and healthier living styles (such as not smoking and drinking) [180]. Studies on cell biology reveal important mechanisms for the development of cancer.

1.2.1 Biology of cancer

For human beings and most other living organisms, the cell is the fundamental functional unit for life activities. At any given moment, amazingly sophisticated bio-chemical or physical processes occur within the cells. These processes produce energy storage molecules, assemble proteins, translate genetic information with the cells continuously undergoing creation and self-destruction (apoptosis) all the time. Almost all of these biochemical reactions are catalyzed with the help of enzymes. An enzyme is a special type of protein which can make new proteins or molecules without consuming itself [109].

Cells differentiate into numerous types (there are roughly 200 types in the human body) to fulfill a diverse range of functions. However, most of them share very similar working mechanisms. Simply speaking, a cell is responsible for maintaining a long string of genetic information and manifestation. DNA (deoxyribonucleic acid) is where the genetic information is stored (for a virus, the RNA - ribonucleic acid - stores this information). A single strain of a double helix ladder encodes almost all the secrets of life [109].

The DNA is only one part of the cell structure required to conduct its normal life

cycle. A typical process of fabricating proteins or enzymes begins with the translation of the DNA information to messenger RNA (mRNA) by the RNA polymerase (an enzyme) which is subsequently sent to ribosome. The ribosome reads the sequence from the mRNA and assembles the required amino acids (roughly 20 types) one at a time and binds them into a functional molecule, i.e. protein.

A normal healthy cell has a typical life cycle [109, 21] which is depicted in Figure 1.1. After the cell has been formed, it enters a relatively long and stable period, the G1-phase. During this period, the cell performs various functions, synthesizes proteins and grows rapidly. Following the G1 phase, the cell enters the S-phase where it begins to duplicate its DNA and prepares for the next cell division. A short G2-phase then allows the cell to assemble the enzymes needed for the following division stage. The actual cell division occurs in the M-phase where a single cell divides into two identical cells. Healthy cells constantly "signal" each other to control the speed of division. These signals can be carried by special proteins (including enzymes), hydrophobic molecules or various ions. Eventually, depending on signals from other cells and the environment, a cell stops growing and diverges into a self-destruction phase - apoptosis.

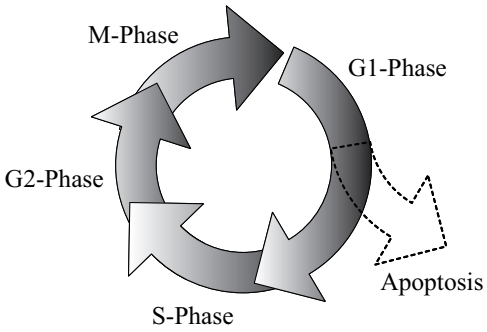


Figure 1.1: Life cycle of normal cells

Under abnormal conditions, some cells lose the ability to respond to these signals. The result of these un-regulated fast growing cells is a mass of functionless tissue - the

tumor (neoplasm). A fully grown tumor imposes pressure to the surrounding tissues or organs and disrupts their normal activities. More seriously, some tumors can release the malfunctioned cells through blood vasculature and lymph systems and spread them to other parts of the body (metastasize). Tumors that can metastasize are called malignant tumors or cancers, while those that cannot are benign tumors. The reasons for the cell's inability to respond to the environmental cell signals may be attributed to mutations induced by chemical compounds (carcinogens), high power radiation, infections by virus or bacteria, or even inherited genetic defects [180].

1.2.2 Breast cancer

Some cancer rates correlate to various population demographics, which can result from distinct cultural and environmental differences. For example, liver cancer is more frequently diagnosed in China than in the United States, while breast cancer has a high incidence rate for women in the US but a relatively low rate in China.

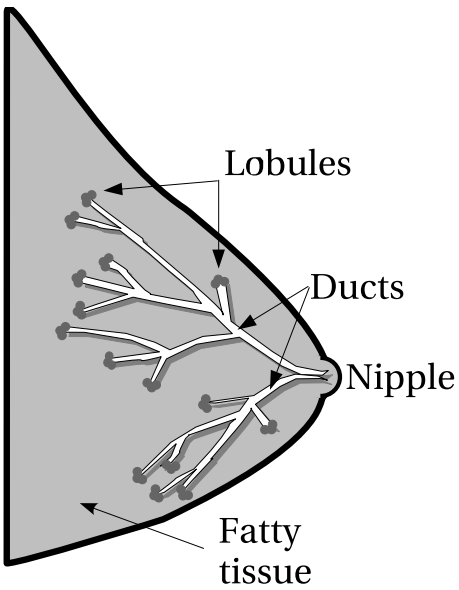


Figure 1.2: Healthy breast anatomy.

Breast cancer is the second leading cause of women mortalities in the US. There are several types of breast tumors and most of them start in the duct and lobular tissues of the breast (Figure 1.2). Common benign breast tumors include fibrosis, the growth of scar-like tissue, and cysts, which are abnormal liquid-filled sacs. Early stages of malignant tumors often appear as ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS). The most common malignant breast tumors include infiltrating (invasive) ductal carcinoma (IDC) and infiltrating (invasive) lobular carcinoma (ILC) which begin growing inside the epithelium of the ducts and milk-producing glandular tissue and spread to surrounding fatty tissue and other parts of the body. Among all breast cancer cases in the US, IDC accounts for roughly 80% while ILC accounts for only 5% of the total. Other uncommon breast cancers include inflammatory breast cancer (1%-3%), medullary carcinoma (5%), and tubular tumor (2%) [181].

It has been reported that detection of breast cancer in early stages is essential for reducing the breast cancer mortality rate [82]. The gold standard for confirming the presence of breast cancer is biopsy which requires the removal of tissue from a patient's breast. Among the various noninvasive means of breast cancer diagnosis, mammography is recommended. Other choices include breast ultrasound, MRI and PET. Of these modalities, mammography uses ionizing radiation which is a potential threat to health with increased dose. Additionally, the compression in mammography is often uncomfortable for the patient. The expense for MRI or PET is quite high and, therefore, are less frequently used. A comprehensive list of modalities used for breast imaging can be found in [136].

1.3 Microwave imaging

This section is intended to provide a general introduction and literature survey for microwave imaging. The overview subsection covers the fundamental rationale of this modality. In the second subsection, we discuss the history and the evolutionary path of this technique especially with respect to the diversity of systems displayed or studied. The following hardware subsection focuses on the imaging systems developed at Dartmouth College and demonstrates the basic components and principles used for data acquisition. In the last subsection, the general procedure for the image reconstruction is outlined which sets the stage for the latest developments comprising this thesis.

1.3.1 Overview

Microwave imaging (MWI) is an active wave-based non-invasive imaging method. First, microwave imaging uses the scattering phenomena of microwave signals as the mechanism for imaging the biological body in contrast to particle-based imaging methods such as PET, SPECT and nuclear medicine (radionuclides, etc.). Second, in microwave imaging, there is no need to deliver the imaging device to the interior of the body via surgery since the microwave signal can penetrate the body and is essentially a noninvasive imaging modality.

Microwave radiation comprises a fraction of the electromagnetic waves spectrum with frequencies ranging from approximately 1 GHz to 30 GHz (UHF is generally considered the frequency range just below it) [100]. It can be used to penetrate the body and retrieve structural and functional information of the tissues via the scattered signals. The physical quantities being imaged in microwave imaging are the dielectric properties, i.e. the permittivity, ϵ, and the conductivity, σ, of the tissues. There is a third

property, the magnetic permeability, μ. Fortunately, most of the biological tissues are non-magnetic, implying that the tissue permeability is identical to that of free-space.

The presence of inhomogeneities in the dielectric properties effects the propagation patterns of the microwave signal throughout the tissue by altering its amplitude, phase or polarization and results in distortions of the microwave field. From an alternative view, the distortions of the field encode the spatial distribution of these dielectric properties. The distorted fields can be measured by microwave detectors, i.e. antennas and receiver electronics, to allow the extraction of the structural information with the help of sophisticated reconstruction algorithms.

The dielectric properties reflect the macroscopic electrical property characteristics of the tissue, implying that they are bulk representations of numerous microscopic physical or bio-chemical processes. In general, the value of the permittivity is related to the molecule dipole moment per volume [166], while the conductivity is related to the free-path length and speed of the electrons inside the material. The value of the dielectric properties can be used as indicators for the microscopic environment of the cellular or molecules processes. When the biological tissues undergo physiological changes, such as those due to the presence of diseases, or those induced by external stimulations, or even by variations in the environmental temperature, the microscopic processes can deviate from their normal state and impact the overall dielectric properties. By monitoring the variations of the dielectric properties with respect to those for the healthy tissues, one may be able to diagnose abnormalities or use the information for treatment of the disease. This is the basic rationale for microwave medical imaging.

The dielectric properties of human tissues have been studied for more than 100 years [59]. The properties undergo significant changes over a wide frequency spectrum with several dielectric relaxation stages impacting the property the most. During 1950s,

Schwan and his collaborators conducted a series of experiments and published a number of papers tabulating their results [173]. More studies were performed during 1980s and 1990s. Published reports included studies by Stuchly and Stuchly [186], Pethig [156], Durney *et al.* [44], Foster and Schwan [55] as well as a summary by Duck [43]. In 1996, Gabriel and Gabriel [59, 171, 60] published their measurements of more than 20 types of human tissues over the frequency band from 10 Hz to 20 GHz. A sample plot of these curves drawn from the data available at their website [88] is shown in Figure 1.3.

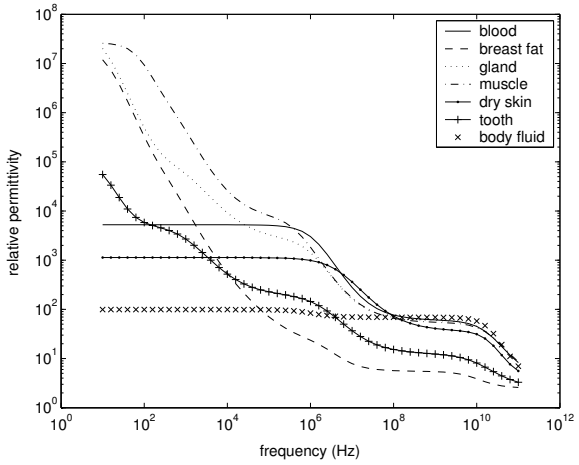
Several observations can be drawn form the curves in Figure 1.3.

1. The property dispersion of biological tissues is not a simple linear relationship with respect to frequency. Instead, a staircase shape feature is observed which can be explained with a Cole-Cole multiple relaxation mechanism;

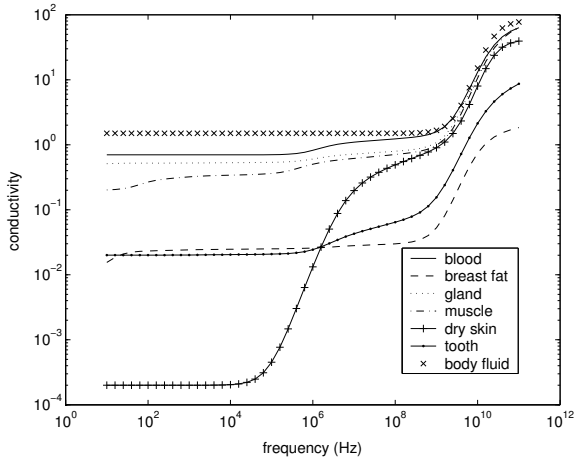
2. Different tissues may have significantly different permittivity and conductivity properties. The lowest dielectric values are found in bone, fatty tissue and lung. In contrast, blood and muscle have much higher permittivities and conductivities due to the abundance of water and free ions.

In terms of the dielectric properties of female breast tissue, Chaudhary et al. [29], Surowiec *et al.* [187], Campbell *et al.* [22] and Joines *et al.* [94] performed *ex vivo* measurements over various frequency bands. The following plot (Figure 1.4) shows the permittivity and conductivity curves for normal breast tissue as well as for malignant breast tissues.

It is interesting to note that there is a significant variation in the reported dielectric properties for overlapping frequency ranges of the different studies. This is probably due to the heterogeneous nature of the breast (adipose and fibroglandular tissue) and the



(a)



(b)

Figure 1.3: Measured dispersion curves of selected human tissues.

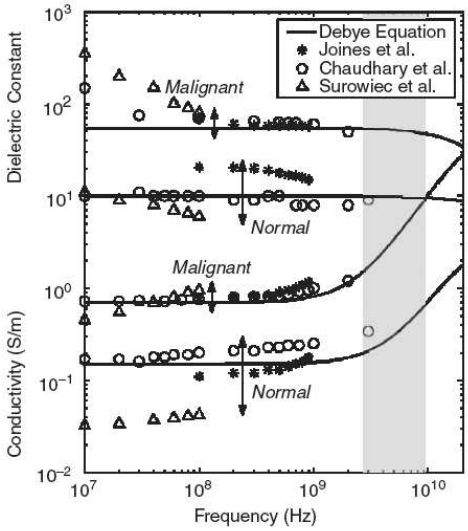


Figure 1.4: The dielectric properties of malignant vs. normal breast tissue (reproduced from [53])

associated variations in water content that drive their properties. It is quite significant that the curves for the breast tumor tissues deviate from those of the normal breast tissue especially in the microwave frequency band. A similar behavior of the malignant to normal tissue was also found between ischemic versus normal heart muscle [178] and normal bone versus leukemic marrow [37]. These discoveries have led to the studies utilizing microwaves to detect tumors by reconstructing dielectric profiles.

The experiments of Gabriel and Gabriel [171] and many other studies utilized *ex vivo* tissue, meaning the tissues used in the measurements were excised from the body but measured as freshly as possible (typically within 24-48 hours after death for human tissues in Gabriel's study). Recent *in vivo* breast tissue measurements at Dartmouth College have demonstrated that the dielectric properties of the *in vivo* tissue, tissue in the living entity, are noticeably different from the *ex vivo* measurements reported by previous researchers and are most likely due to the loss of blood [54].

The high contrast demonstrated in the previous examples provides significant ra-

tionale for the advantages of microwave imaging in breast cancer detection. However, there are other advantages. As previously mentioned, microwave imaging uses non-ionizing radiation which is significantly safer than ionizing radiation, i.e. X-rays used in mammography. The low illumination power levels used in microwave imaging also make regular screening possible. Finally, no compression is needed in microwave imaging making the exams more comfortable than mammography.

Another advantage of microwave imaging is low cost. The expense for building a microwave imaging system could potentially be far less than that for CT and MRI. With the widespread use of microwaves in everyday life, the manufacturing costs of essential components have decreased dramatically. These microwave applications include cell phones (frequency: CDMA 1.880 GHz-1.990 GHz, TDMA: 824.04 MHz-893.7 MHz), wireless networks (802.11a: 5 GHz, 802.11b: 2.45 GHz), microwave ovens (2.450 GHz) and GPS tracking systems (L1: 1.57542 GHz, L2: 1.2276 GHz) [116]. The boom of microwave related technologies has greatly stimulated the growth of microwave component and system design and manufacturing industry. These components are becoming smaller, cheaper and more powerful.

Not only can we recover tissue dielectric properties from MWI, other physical or biological properties that have strong correlations to dielectric properties can also be inferred from the reconstructed dielectric images. At Dartmouth College, the thermal dependence of the dielectric properties have been studied by Meaney *et al.*[27, 123, 154, 153] since 1993. The studies utilize the almost linear dependence of the tissue conductivity with temperature. By reconstructing the dielectric property images in near real-time, the variations of the temperature profiles can be retrieved from the pre-calibrated dielectric-temperature relationship. Temperature monitoring is especially necessary in hyperthermia. In hyperthermia, the malignant tissue is heated to facilitate cancer cell

death in conjunction with radiation treatment. Precise control of the temperature is of great importance for both efficient treatment and minimizing the damage to normal tissue. Non-invasive microwave thermometry is a good candidate for use in conjunction with hyperthermia treatment.

1.3.2 History

Initial efforts in developing an active microwave imaging system can be traced back to a paper by Jacobi, Larsen and Hast in 1979 [89]. Inspired by studies in military underwater telecommunication techniques, they suggested a microwave measurement system by submerging an antenna into water which was shown to be promising for exploring bio-systems. They demonstrated that significant improvements could be achieved in the impedance characteristics of the antennas, energy coupling efficiency as well as the antenna aperture size by simply immersing the antenna system into a high permittivity medium – water. The paper outlined the attractive blueprint of an active noninvasive microwave interrogation system which could provide resolution better than decimetric wavelength. Shortly after the publication of that paper, they reported the successful implementation of such a system in imaging biological tissue – an isolated canine kidney [103]. Considerable interests in studying microwave tomography were stimulated by the promising results demonstrated by Jacobi and Larsen. These early studies included algorithmic studies by Maini *et al.* [119], Rao *et al.* [163], Ermert *et al.* [46], Adams *et al.* [1], Pichot *et al.* [157], as well as attempts at fabricating actual microwave imaging systems, of which the quasi-real-time system using a planner antenna array developed by Bolomey *et al.* [13, 14] and the cylindrical array system by Broquetas *et al.* [20] are of particular importance.

The proliferation of studies into microwave imaging continued until the beginning

of the 1990s when two new advances accelerated the developmental process. The first advance was in the development of iterative reconstruction algorithms. Efficient algorithms were proposed for solving the nonlinear reconstruction problems which were previously solved by limited linear algorithms. The distorted Born iterative method (DBIM) developed by Chew *et al.* [34] and Newton-Kantorovich method (or Gauss-Newton method) developed by Joachimowicz [169, 93] greatly improved the overall image quality (these two methods were later shown to be equivalent [167]). Interestingly, with these nonlinear algorithms, the spatial resolution of the image was limited by the signal to noise ratio (SNR) rather than the half-wavelength diffraction limit (see Section 2.2). With high quality measurement data, microwave imaging can reconstruct objects with diameters down to $1/7$ - $1/10$ of the wavelength [133, 134, 32] which is also referred to as the super-resolution. On the other hand, more potential applications of microwave imaging were discovered and several research groups became involved in these efforts. A number of prototype systems were developed for various purposes, which included the imaging system built by Azaro *et al.* [2, 152] as well as Otto and Chew *et al.* [149] for algorithmic research, a tissue blood content monitoring system by Hawley (1991), the multiple-frequency microwave thermometry system by Bardati *et al.* [4], the CP-MCT (Chirp-Pulse Microwave CT) systems for subsurface thermal imaging by Miyakawa *et al.* [131], the focal plane imaging system by Goldsmith *et al.* [64], the resonant dielectric sensor for localization of breast tumors by Preece *et al.* [160] and Pothecary *et al.* [159], and the 20-channel monolithic dipole imaging array by Hsia *et al.* [85]. In terms of developing efficient imaging methods for breast cancer diagnosis, Meaney *et al.* at Dartmouth College (USA) developed a tomographic imaging system which was initially designed for non-invasive thermometry [127]. Over the course of three major updates, the system became the first laboratory microwave imag-

ing system available for clinical use in breast cancer imaging [122]. Meanwhile, Bond, Hagness and Fear *et al.* are developing a reflection-based breast tumor imaging systems utilizing confocal microwave imaging (CMI) at the University of Wisconsin Madison (USA) and the University of Calgary (Canada) [73, 74, 15, 52, 53]. Liu *et al.* at Duke University (USA) is also developing a tomographic breast imaging system [115]. Parallel efforts were also reported in microwave cardiac imaging by Semonov *et al* - 2D [175] and 3D [176] imaging systems for the heart as well as whole body imaging.

These systems varied significantly in their configurations due to the distinct scopes of their applications. These differences include antenna selection, operating frequencies, wave forms, source/receiver number and spatial arrangement. Waveguide antennas were used in the systems developed by Jacobi *et al.* [89, 103], Bolomey *et al.* [13], Semenov *et al* [174, 178, 177], Mallorqui *et al.* [120] and Miyakawa *et al.* [131, 132], while simple monopole antennas were extensively exploited by Meaney *et al.* [122, 123] and Li *et al* [111]. Patch antennas were reported for the imaging systems by Hsia *et al.* [85] and Hagness *et al.* [74] and spiral antennas for passive heat monitoring by Jacobsen and Stauffer [90]. In terms of operating frequencies and wave form, most research groups used time-harmonic waves ranging from 300 MHz to 3 GHz for tomographic imaging purposes while broad-band pulse signals were found in CMI system by Hagness *et al.* [74] , as well as chirp signals used by Miyakawa *et al.* [132] in their CP-MCT system (the pulse data was usually synthetically extracted from multiple sets of time-harmonic data).

Numerous experiments have been successfully conducted in simulations, phantoms along with *ex vivo* and *in vivo* measurements on small animals and patients to demonstrate the viability and performance of microwave imaging. In simulations, 2D reconstructions have been reported by Joachimowicz *et al*. [93], Caorsi *et al.* [24], Semenov

et al. [175] and Meaney *et al.* [127], among others. Phantom and *ex vivo* image reconstructions have also been reported by Semenov *et al*. [175, 176] and Meaney *et al.* [124, 125]. Preliminary *in vivo* experiments were presented by Semenov *et al.* [177] for imaging a canine heart and by Meaney *et al.* for thermal imaging the torso of small pigs [123] and the human breast [122].

In parallel, significant efforts have been made in the development of efficient reconstruction algorithms. While microwave scattering is a three-dimensional vectorial phenomenon, 2D reconstruction algorithms are quite prevalent mostly due to their simplicity. In early 2D tomographic implementations, diffraction approaches were applied to linearize the reconstruction problem utilizing primarily Born and Rytov approximations. These were appealing at the time because images could be produced efficiently, given the limited computational power available. They were shown to be effective when the scattering objects were electrically small or when the contrast with the background was minimal, which is generally not the case for imaging of biological tissue. Methods for solving nonlinear partial differential equations (PDE's) such as finite element (FE) and finite-difference time-domain(FDTD) methods appear to be more appropriate models for the EM scattered fields given the significant advances in improved computational capabilities.

As reconstruction strategies advanced, iterative methods based on integral equations such as the distorted Born iterative method (DBIM), Born iterative method (BIM) [34, 31] and local shape function (LSF) [149] have been implemented along with approaches based on differential equations such as the Newton-Kantorovich [93] method. The nonlinear scattering nature of microwave signals as well as the ill-posedness of the inverse problem present significant challenges in the development of appropriate algorithms. Depending on the number of unknowns and the size of the reconstruction problem, the

computational time for solution convergence varied from a few minutes [115, 51] to hours or even days [177]. Therefore, many of the current microwave imaging systems used off-line computations due to the extensive computer load. Various attempts were made to reduce the reconstruction problem complexity by taking into account different approximations and simplifications, such as the dual-mesh scheme [155], conformal mesh reconstruction [110], adjoint technique [50] and frequency-hopping reconstruction algorithms [33]. The development of global optimization algorithms such as neural-network (NN) technique, simulated annealing (SA) algorithm), genetic algorithms (GA) have also been applied to microwave image reconstruction [10, 24, 25]. It has been recognized that the computational speed of the forward problem in iterative approaches is the most time consuming part of the problem [189, 50]. Techniques such as iterative block solvers and the adjoint method [18] will be essential as 3D imaging approaches are developed.

1.3.3 Microwave imaging system at Dartmouth College

Led by Professors Keith Paulsen and Paul Meaney, the microwave imaging group at the Thayer School of Engineering, Dartmouth College (USA), started the study on microwave imaging in the early 1990s. After years of preliminary studies on material, antenna design and reconstruction algorithms, the first laboratory-scale prototype system was fabricated in 1995. A four-detector monopole antenna array was used for the receivers. Modulated continuous-wave (CW) signals were transmitted from water-loaded waveguide antennas operating over the frequency between 300 MHz and 1.1 GHz. A superheterodyne technique was used as the scheme to extract the phase and amplitude information from the high frequency signals (see Section 1.3.4). Multiple phantom measurement data were successfully acquired and associated images reconstructed by

a Gauss-Newton reconstruction algorithm previously developed by Paulsen *et al.* (for more details on this algorithm please refer to Chapter 3).

In the year 1999, a new system was built for breast imaging based on the experience learned from the prototype system and was referred to as the "first generation" system. A picture of the system is shown in Figure 1.5. In this system, functionalities of both transmitting and receiving signals were combined into each antenna channel (i.e. the transceiver); the state of transmitting or receiving was dictated by an electronic switching network. All 16 monopole antennas could be used to transmit microwave signals with the 9 antennas on the opposite half circle of the array used to receive scattered field signals. Thus, the total amount of single plane measurement data for this system was $16 \times 9 = 144$ data points (due to the reciprocity relationship, the amount of independent measurement data was half of that total, i.e. 72). Additionally, the antenna array could be moved up and down manually via a hydraulic jack. This facilitates the collection of multiple planar data for 3D objects. The speed of the electronics were also improved significantly. For a typical session, this system required roughly 30 minutes for seven planes of data on two breasts while the prototype system needed several hours to acquire the same amount of data.

With this first generation system, a number of phantom, small animal and patient studies were conducted. The phantoms included both solid and liquid cylinders with inclusions of various sizes. Solid spherical phantoms were also measured to study the 3D effect of the measurement system. The dielectric properties of these phantoms range from $\epsilon_r = 5$ and $\sigma = 0.1$ S/m for bone-fat phantoms to $\epsilon_r = 50$ and $\sigma = 1.3$ S/m for agar cylinders. A series of *in vivo* animal experiments were performed for demonstrating the clinical promise of microwaves in non-invasive thermal monitoring during hyperthermia treatment. A living piglet with a hot water tube surgically inserted through its

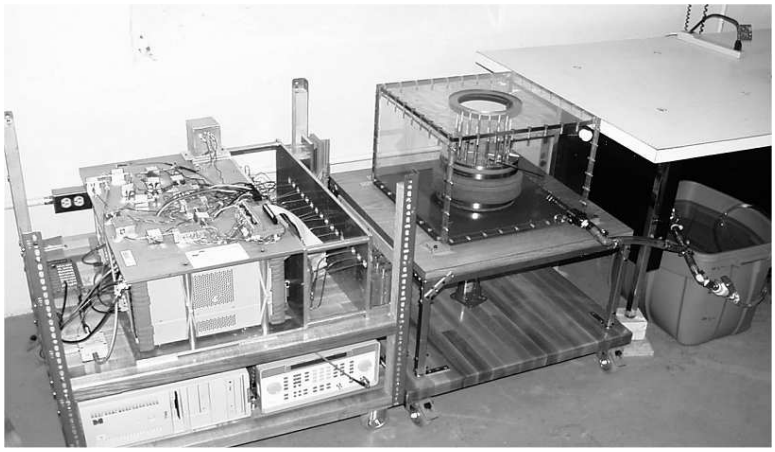


Figure 1.5: Photograph of the "first generation" microwave imaging system at Dartmouth College

abdominal cavity was imaged in a series of experiments. The inserted tube supplied saline at different temperatures with the measurement data collected at each [123]. The reconstructed images showed very consistent variation of the dielectric properties with respect to temperature. The anatomy of the piglet was also successfully reconstructed (see Section 3.7.3 for more details).

In these phantom and animal studies, 0.9% saline was primarily used as the coupling medium instead of pure water as originally suggested by Jacobi and Larsen. The lossiness of the saline reduced the reflections from the illumination tank walls while the signal intensity was still well above the noise floor even with the added attenuation. Starting in 2001, several other possible coupling media were also investigated including mixtures of glycerin and water which has shown great promise [128].

In 2002, an entirely new system was fabricated primarily for the purpose of breast imaging and was referred to as the "second generation" system. A number of major upgrades made the new system attractive and efficient. These improvements are summarized in Table 1.1. The first major difference was the implementation of the parallel data acquisition (DAQ) strategy. This strategy allowed for the collection of measure-

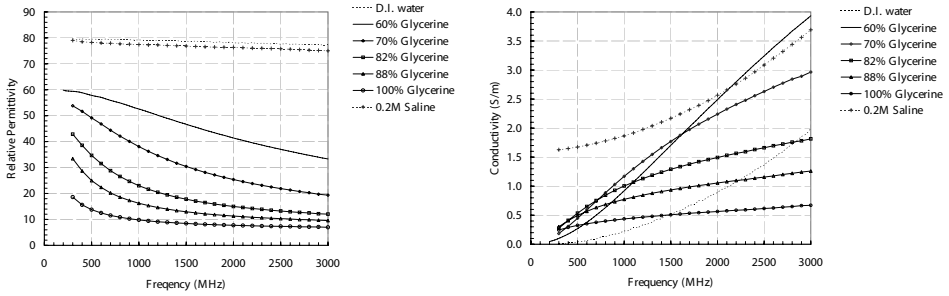
Table 1.1: Differences between the first and second generation imaging systems at Dartmouth College

Differences	1nd Generation	2st Generation
Detection scheme	serial	parallel
Coupling medium	saline	glycerin:water
Measurement time	12.5 minutes for 12,096 data points	10 minutes for 33,600 data points
Operating frequency	300 MHz-1.1 GHz	500 MHz-3 GHz
Antenna positioning	manual	automatic
Receiver number	9	15
Tank volume	64 liters	33 liters

ments from multiple receivers simultaneously. A high speed 16-bit DAQ board was integrated with a variable gain amplifier and controlled by a computer. The improvements in the data acquisition method and hardware resulted in the dramatic reduction in the measurement time while acquiring significantly more data. The upper end of the system operating frequency range was also extended from 1.1 GHz to 3 GHz. The capability of collecting higher frequency data may provide more information about the target and also the opportunity to study image resolution and quality with respect to frequency.

Another major modification was the coupling medium. Glycerin/water solutions replaced the high-attenuation saline as the coupling medium. The advantages provided by the glycerin solution included 1) less loss than saline producing less attenuation of the microwave signals which improved the measurement SNR and 2) the complete solubility of glycerin in water provided a means of tuning the coupling medium relative permittivity over a range of 10 to 80 by simply adjusting the solution concentration. The second feature is especially attractive for breast imaging in clinical applications. Figure 1.6 shows the permittivity and conductivity dispersion curves for glycerin/water solutions as well as those for distilled water and saline. It is clear that for higher water concentrations, the medium is lossier with much higher permittivity values. Interest-

ingly, it has been found through our own patient studies that the average dielectric properties for breast tissue varies significantly with respect to the age and breast density of women. In general, breasts of younger women have more water content and higher density which correlate with higher values in dielectric properties. With increase in age, the percentage of fatty tissue increases and the density of the breast decreases along with the dielectric properties. The adjustable dielectric properties of the glycerin solutions offer the capability of matching the impedance not only between the medium and the antennas but also between the medium and breast. Finally, the glycerin solutions are biostatic, making them safe for patient examines.



(a) (b)

Figure 1.6: The dielectric properties of different coupling media [128]

Besides the improvements in the DAQ system and coupling medium, automatic controls of several routine tasks were more extensively integrated into the system for improved efficiency. For improved management of the coupling medium, a separate fluid management system (FMS) was built to pump the liquid back and forth between the medium reservoir and the illumination tank for re-use of the liquid for multiple exams. An ultraviolet (UV) light sterilizer was also installed for improved patient safety. In addition, a computer-controlled high-precision linear actuator was utilized to control the vertical positions of the antenna array. A photograph of this system is shown in

Figure 1.7.

Figure 1.7: Photograph of the second generation microwave imaging system at Dartmouth College

Beginning in year 1999 (also including exams using the first generation system), the Dartmouth imaging system began clinical trials at Dartmouth Hitchcock Medical Center (DHMC). So far, more than 200 patients have participated in the imaging studies. These studies included a normal breast study (24 patients), abnormal breast study (malignant tumor cases - 8 , benign abnormality cases - 29), menstrual cycle study (8 women - 4 sessions each) and a MWI/MRI co-registration study (6 patients). In addition, an on-going blinded study has currently enrolled over 120 patients. The images reconstructed from these studies have demonstrated impressive correlations in comparison with the clinical information and results from other Dartmouth alternative imaging modalities (near infrared imaging - NIR, and electrical impedance tomography - EIT).

1.3.4 Principle of data acquisition in microwave imaging

The basic components of the microwave imaging DAQ system are illustrated in the diagram in Figure 1.8. In general, the system is divided into three major functional

modules: the microwave source, the antennas and the acquisition and processing unit.

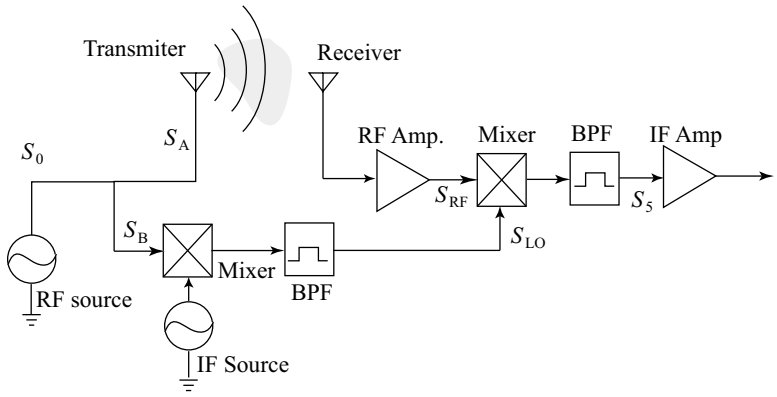


Figure 1.8: Diagram of the data acquisition (DAQ) scheme of the microwave imaging system at Dartmouth.

The synthesized microwave source provides high quality electromagnetic waves at different frequencies. The generated signals are coupled into a wave guiding system and undergo a series of processes, such as amplification, mixing, filtering and dividing. At the end of the wave guiding system are the antennas. The antennas are the interfaces between the microwave circuitry and open space. Electromagnetic waves are radiated into space carrying the signal information. The radiation capability of an antenna depends on its geometrical shape and material. In general there are two types of antennas, the isotropic antenna and the directional antenna. An isotropic antenna radiates microwave signals equally in all directions while a directional antenna transmits the microwave signals in specific directions with significantly reduced power level in the others. Roughly speaking, the larger the surface area of the antenna, the more directional the radiation pattern will be. Parabolic reflection antennas, open-ended waveguides, horn antennas and large antenna arrays can have very narrow radiation beams in specific directions, while simple radiating elements such as wire antennas are used for non-directional radiation in such applications as the cell phone base stations, wireless networks.

The transmitted microwave interacts with the target in the space by means of reflection, attenuation and diffraction. The scattered EM field is then collected by the receiver antennas through electromagnetic induction. Due to the reciprocity, the induction sensitivity at different directions of a receiver antenna is exactly the same as the radiation capability of the antenna in those directions. The receiver antennas convert the spatial wave into high frequency signals which are further amplified and filtered to extract the desired information.

It is quite difficult to process the current signals directly at microwave frequency. Practically speaking, most commercially available DAQ boards only operate up to several hundred kiloherz. A superheterodyne method [161] is a commonly used receiver scheme to convert a high-frequency signal into a lower-frequency signal while preserving the amplitude and phase information. The following is a short introduction to this scheme.

Assume that the signal from the RF source (S_0) has frequency ω_{RF} (which is a microwave frequency). Signal S_0 is split into two identical components, S_A and S_B, which have the same frequency ω_{RF}. S_A is sent to a transmitting antenna and radiated into space, while the other signal, S_B, is multiplied by an intermediate frequency (IF) signal with frequency ω_{IF} (in our case, $\omega_{IF} = 2$ kHz). The resulting signal can be expressed as

$$\begin{aligned} S_{MIX} &= S_B \times S_{IF} \\ &= A_B \cos\omega_{RF}t \times A_{IF} \cos\omega_{IF}t \\ &= \tfrac{1}{2}(A_B A_{IF})(\cos(\omega_{RF}+\omega_{IF})t + \cos(\omega_{RF}-\omega_{IF})t) \end{aligned} \tag{1.1}$$

where A_B and A_{IF} are the amplitudes of S_B and S_{IF}, respectively. S_{MIX} contains two frequencies, the lower side band is selected by a band pass filter (BPF) as the local oscillator (LO) signal and sent to the receiver side. (Note that in practical situations,

some of the original S_B leaks through the mixer and also needs to be filtered). The LO signal is denoted as

$$S_{LO} = A_{LO} \cos(\omega_{RF} - \omega_{IF})t \tag{1.2}$$

The transmitted signal S_A interacts with the unknown target and then is received by a receiver antenna where the received signal is denoted as S_{RF} in the form of

$$S_{RF} = p \cos(\omega_{RF} t + \phi) \tag{1.3}$$

where p is the amplitude and ϕ is the phase. Then S_{RF} is mixed with S_{LO} which results in

$$\begin{aligned} S_{RF} \times S_{LO} &= p \cos(\omega_{RF} t + \phi) \times A_{LO} \cos(\omega_{RF} - \omega_{IF})t \\ &= S_1 + S_2 \end{aligned} \tag{1.4}$$

where

$$\begin{aligned} S_1 &= \tfrac{pA_{LO}}{2} \cos((2\omega_{RF} - \omega_{IF})t + \phi) \\ S_2 &= \tfrac{pA_{LO}}{2} \cos(\omega_{IF} t + \phi) \end{aligned} \tag{1.5}$$

Note again that there is some leakage of the RF and LO frequencies which must be filtered.

In the two signals, S_2 is centered at ω_{IF} which is a much lower frequency than the original microwave frequency ω_{RF} while both the amplitude and phase information are contained in S_2. A Fourier transform is subsequently applied to S_2 to yield the complex form representation $\tilde{S}_2$ by

$$\tilde{S}_2(\omega_{IF}) = \int_0^{2\pi} S_2(t) \exp(-j\omega_{IF} t) dt \tag{1.6}$$

from which the amplitude and phase information can be extracted.

1.3.5 Overview of the image reconstruction algorithm

The EM field components utilized in the image reconstruction algorithms are spatial distributions with their associated time dependence. If the field is excited by a time-harmonic source, the time dependence is then replaced by frequency. In this case, the problem is referred to as a "frequency domain" problem, otherwise, a "time domain" problem. Additionally, the distributions of the fields also depend on the dielectric properties of the propagating medium. The relationship between the dielectric properties and the fields are Maxwell's equation (see Chapter 3 for more details). When the dielectric property changes, so does the EM field distribution. From that point of view, the EM field components are functions of the dielectrics. Since the dielectric properties themselves are functions of space and frequency, it is appropriate to say the fields are functionals (functions with another function as the independent variable) of the dielectrics.

Assuming the electric field $\vec{E}$ is measured at the receiver, the measured field can be expressed in the functional form as $\mathcal{E}(\vec{r}, \omega, \epsilon(\vec{r}, \omega), \sigma(\vec{r}, \omega))$ where functions $\epsilon(\vec{r}, \omega)$ and $\sigma(\vec{r}, \omega)$ are the unknown distributions of permittivity and conductivity, respectively. $\vec{r}$ is the spatial coordinate and ω is the angular frequency. Very often, the relationship between the dielectric properties and the electric field can be explicitly expressed by an integral equation as [100, 31, 87]

$$\begin{aligned} \mathcal{E}(\vec{r}, \omega, k^2(\vec{r}, \omega)) &= \mathcal{E}_{inc}(\vec{r}, \omega, k^2(\vec{r}, \omega)) + \mathcal{E}_{sca}(\vec{r}, \omega, k^2(\vec{r}, \omega)) \\ &= \mathcal{E}_{inc}(\vec{r}, \omega, k^2(\vec{r}, \omega)) + \int_{\Omega} G(\vec{r}, \vec{r}', \omega) k^2(\vec{r}', \omega) \mathcal{E}(\vec{r}', \omega, k^2(\vec{r}', \omega)) d\vec{r}' \end{aligned} \tag{1.7}$$

where $k^2(\vec{r}, \omega)$ is referred to as the square complex wave number defined by (assuming

$\exp(j\omega t)$ time dependence)

$$k^2(\vec{r},\omega) = \omega^2\mu\epsilon(\vec{r},\omega) - j\omega\mu\sigma(\vec{r},\omega) \tag{1.8}$$

Ω is the spatial domain, $G(\vec{r},\vec{r}',\omega)$ is the Green's function, $\mathcal{E}_{inc}$ and $\mathcal{E}_{sca}$ are the incidence and scattered fields, respectively. Note that the incidence field $\mathcal{E}_{inc}$ is typically known for most scattering problems. Equation (1.7) is called a Fredholm integral equation of the second kind [79] because the field functional, $\mathcal{E}$, appears on both sides of the equation. To evaluate the electric field from (1.7), one common treatment is to apply the Born approximation [31, 95], which assumes that the scattered field is weak enough so that the total field, $\mathcal{E}$, can be approximated by the incident field, $\mathcal{E}_{inc}$. In this case, (1.7) can be re-written as

$$\mathcal{E}(\vec{r},\omega,k^2(\vec{r},\omega)) = \mathcal{E}_{inc}(\vec{r},\omega,k^2(\vec{r},\omega)) + \int_\Omega G(\vec{r},\vec{r}',\omega)k^2(\vec{r}',\omega)\mathcal{E}_{inc}(\vec{r}',\omega,k^2(\vec{r}',\omega))d\vec{r}' \tag{1.9}$$

which is essentially a Fredholm integral equation of the first kind because the unknown $\mathcal{E}$ appears only on the left side of the equation. Equation (1.9) represents a linear relationship in the functional space which can be solved easily with either analytical or numerical techniques.

Since the measurement data are known only at discrete receiver locations, for instance, $\{\vec{r}_i\}_{i=1}^{N_r} = \{\vec{r}_1,\vec{r}_2,\cdots,\vec{r}_{N_r}\}$ where N_r is the number of measurement data, the purpose of the image reconstruction is to recover the two functions $\epsilon(\vec{r},\omega)$ and $\sigma(\vec{r},\omega)$ such that the electric fields $\mathcal{E}(\vec{r},\omega,\epsilon(\vec{r},\omega),\sigma(\vec{r},\omega))$ computed from the forward model match that of the actual measurements at the receiver sites $\{\vec{r}_i\}_{i=1}^{N_r}$. Written in mathematical

form, we need to solve for $\epsilon(\vec{r}, \omega)$ and $\sigma(\vec{r}, \omega)$ from the following equation:

$$\mathcal{E}(\{\vec{r}_i\}, \omega, \epsilon(\vec{r}, \omega), \sigma(\vec{r}, \omega)) - \mathcal{E}^M(\{\vec{r}_i\}, \omega) = 0 \tag{1.10}$$

where $\mathcal{E}^M$ is the measured E field.

This is essentially a root-finding problem in the functional space. There are multiple difficulties in solving this equation: 1) the relationship between the dielectric properties and the fields are not linear unless certain approximations are applied such as the Born or Rytovs [95, 31] approximation; 2) the solutions to this equation may be non-unique due to the nature of the reconstruction (inversion of a compact operator, refer to [104, 138]); 3) the measurement data are contaminated by noise such that exact equality is not possible and 4) the evaluation of the solution is natively sensitive to the random noise presented in the measurement (ill-posedness, see Section 2.4). Attempts to obtain an analytical solution are impractical given these difficulties and we must resort to numerical techniques. To deal with the first three difficulties, one needs to convert the root-finding problem into a nonlinear optimization problem (Section 2.3); and regularization techniques should also be implemented to account for the fourth difficulty (Section 2.4).

Written in mathematical form, the optimization problem can be formulated as follows

$$\min_{\epsilon(\vec{r},\omega),\sigma(\vec{r},\omega)} \mathcal{F}\left(\mathcal{E}(\{\vec{r}_i\}, \omega, \epsilon(\vec{r}, \omega), \sigma(\vec{r}, \omega)), \mathcal{E}^M(\{\vec{r}_i\}, \omega)\right) \tag{1.11}$$

where functional $\mathcal{F}(\cdot, \cdot)$ is referred to as the norm functional which measures the differences between the measured field and the predicted measurement computed from the model. A simple and quite popular selection for $\mathcal{F}$ is to use the sum-of-square function, i.e., for the discrete measurement problems, the above optimization problem can

be written as

$$\min_{\epsilon(\vec{r},\omega),\sigma(\vec{r},\omega)} \|\mathcal{E}(\{\vec{r}_i\},\omega,\epsilon(\vec{r},\omega),\sigma(\vec{r},\omega)) - \mathcal{E}^M(\{\vec{r}_i\},\omega))\|^2 \tag{1.12}$$

If the noise probability model is some form other than Gaussian in the optimization (1.12), $\mathcal{F}(\cdot,\cdot)$ can be chosen as the likelihood function or other probability functions.

Recall that this optimization problem is sensitive to noise or ill-posedness. To recover a meaningful solution from the optimization problem (1.12), the object function in (1.11) needs to be altered or constrained in order to reduce or remove the ill-posedness. This is generally referred to as a regularization technique. A general form to problem (1.11) with the consideration of regularization is written as

$$\min_{\epsilon(\vec{r},\omega),\sigma(\vec{r},\omega)} \mathcal{F}(\mathcal{E},\mathcal{E}^M) + \mathcal{G}(\mathcal{E},\mathcal{E}^M) \tag{1.13}$$

where $\mathcal{G}(\cdot,\cdot)$ is called the smoothing norm [77] which is used to leverage the balance between ill-posedness and the measurement information. A further discussion on the selection of regularization and the execution of the optimization can be found in Section 2.4.

From a signal and system perspective, the reconstruction problem is also referred to as the inverse problem, in contrast to the forward problem. In general, there is no clear-cut boundary between the two-types of problems. Roughly speaking, we call the problems of estimating the properties of an unknown system from known input and the measurement as the inverse problem, while computing the output from a known inputs and system properties is called the forward problem. Mapping these concepts into microwave image reconstruction problem, the dielectric distributions are the properties of the system if we consider the structure of the target as a system. With a known target,

i.e., given the dielectric property distributions, the evaluation of the scattered field at the known source by solving the forward model (in microwave imaging, the forward model is Maxwell's equation or the wave equation as explained in Section 3.2.2) is called the forward problem whereas the optimization problem (1.13) is the inverse problem. As one may have noticed, the evaluation of the inversion requires that the field values be solved for at the receiver sites from the model, i.e. the evaluation of the forward problem.

1.4 Hypotheses and aims

In short, the primary goal of this study is to develop efficient reconstruction algorithms for microwave imaging. By "efficient", we refer to two criteria, i.e. the speed of the reconstruction and the quality of the resultant images. In this section, we outline a road map and the specific goals to achieve this general aim.

The underlying hypotheses of developing these algorithms are based on previous studies in both theory and hardware developed at Dartmouth College. These hypotheses include:

- Improved model match leads to a reduction of image artifacts and improved convergence behavior;
- Improved understanding of the measurement data independence is important to the hardware design and reconstruction strategy;
- Image reconstruction benefits from an enriched data set;
- Combination of low and high frequency measurement data can improve the robustness and image quality of the reconstructions.

As a result, we proposed four specific aims to incrementally test these hypotheses:

Aim 1: Development of a viable 3D microwave image reconstruction algorithm.

Facilitated by the dual-mesh technique (refer to Section 3.3 for more details), this goal will be accomplished in two steps (a) developing the 3D forward/2D reconstruction (3D/2D) technique (i.e. computing the forward field solution in 3D at each iteration while assuming a 2D parameter discretization for the reconstruction task), (b) developing the 3D forward/3D reconstruction (3D/3D). The 3D/3D study will implement both 3D scalar (3Ds/3D) and 3D vector (3Dv/3D) forward field calculations in the reconstructions. The effectiveness of these algorithms will be assessed for both simulation and measured data.

To recover representative electrical property profiles from measurement data, matching of the numerical model to the physical situation is one of the most important factors to ensure the correct interpretation of the data and convergence. Modeling of 3D electromagnetic wave propagation in a complex medium is generally computationally expensive even with modern computational resources. Using 2D methods to model this inherently 3D phenomenon may save significant computational time; however, it may impose excessive simplifications which can introduce image artifacts. A viable 3D image reconstruction method will be developed which possesses the following characteristics:

- computational feasibility and scalability between model accuracy and efficiency;
- compatibility with previously developed reconstruction techniques such as the log-magnitude phase-form reconstruction [151], multi-spectrum technique, etc.

Utilizing the dual-mesh approach as a fundamental framework, we have devised a strategy for developing a series of 2D and 3D algorithms in Chapter 3 through 5.

This series includes (a) the 2D forward/2D algorithm (Chapter 3), (b) the 3D scalar forward/2D algorithm (Chapter 4), (c) the 3D scalar forward/3D algorithm and (d) the 3D vector forward/3D algorithm (Chapter 5). The development of these approaches provides a solid framework for exploring the validity of the first hypothesis. The assessment of the relationship between model complexity and efficiency will be valuable not only for microwave imaging but for other non-linear imaging methodologies.

Aim 2: Image reconstruction incorporating multiple frequency measurement data to enhance image quality.

Hypotheses 2, 3 and 4 generally imply that increasing the amount of measurement data could possibly reduce the ill-posedness of non-linear inverse problems when utilizing approaches such as the Gauss-Newton iterative technique. However, for the microwave imaging situation, it is not always clear what is the optimal way to increase the measurement data. For instance, simply increasing the number of antennas surrounding the imaging target may not improve matters significantly because the new data may be largely dependent on existing data depending on how electrically small the spacings between antennas are. Increasing the amount of data by utilizing measurements collected over a range of frequencies may provide considerably more linearly independent data. Again, caution should be used to utilize data for frequencies spaced far enough apart that they increase the amount of new independent data.

In addition, complementary behavior when utilizing the low and high frequency measurement data in the reconstruction process is consistently observed (the lower frequency measurement data provides for more stable convergence behavior, but the final images are quite smooth due to wavelength limitations; the higher frequency measurement data provides for improved spatial resolution; however, the reconstructions in these cases more readily converge to local minima). Consequently, we are developing a

method that can exploit the positive aspects of both to enable stable image reconstructions while still retaining good spatial resolution.

The property dispersion relationships for the imaging target and background medium must also be considered for field modelling at multiple frequencies. Finally, the impact from multiple-frequency measurement and least-square convergence will be assessed in this study. The efforts in developing such a multiple-frequency dispersion reconstruction algorithm can be found in Chapter 6.

Aim 3: Singular value decomposition (SVD) will be exploited to construct appropriate evaluation metrics that can be used to optimize important hardware parameters such as operating frequency, antenna count, antenna distribution and background contrast with respect to overall improvements in image quality.

An important goal in analyzing the measurement data independence and efficiency is to provide strategic suggestions for improvements to the hardware system and associated algorithms to provide the best image quality at reasonable cost.

The SVD of the imaging operator can provide detailed insight into the reconstruction process, and is intimately related to the image resolution limit, degree of ill-posedness, noise level and other important factors in image reconstructions. This analysis has been applied to the Radon transform (linear reconstruction) by Davison [40] and Caponnetto and Bertero [26], and the inverse Born approximation by Brander and DeFacio [19]. The analogous study we propose for non-linear diffraction imaging modalities such as ours will be extremely useful in providing a theoretical basis for analyzing image formation theory and evaluating the effectiveness of such systems.

In this study (Chapter 7), we will focus on simple tomographic 2D and 3D imaging approaches for noiseless and noisy cases. We will investigate the possibility of deriving the imaging operator SVD analytically. Where this is not feasible, we will compute

the numerical SVD. Suitable metrics will be constructed to evaluate the efficiency of different combinations of system parameters.

Aim 4: We will implement a number of general techniques for image quality and computational efficiency improvements including:

- **Time domain forward field solution to reduce computation time - especially in the multi-frequency approach;**
- **Adjoint method for improving the computation efficiency of the Jacobian matrix;**
- **Utilizing row/column weighting for improving numerical accuracy in solving least-square problems;**
- **Understanding the behavior of the unwrapped phase for better implementation of the reconstruction algorithms utilizing the phase measurement.**

In contrast with the first three fundamental aims, aim 4 is comprised of several detailed but quite significant techniques which either provide support functions for the preceding aims or serve as key components for future in-depth studies. For example, the dual-mesh adjoint method (Section 3.3) provides a fundamental computational efficiency enhancement for all dual-mesh based algorithms in Aim 1. In addition, the row/column weighting techniques (Section 6.2.3) provide the a mechanism for fine-tuning the multiple-spectrum reconstructions. Chapter 8, 9 and 10 were devoted to understanding the properties of the phase unwrapping integral and the influence of the presence of phase singularities. This study not only has theoretical significance in mathematics and topological electromagnetics, but also is important in explaining the behavior of the previously developed algorithm in processing the microwave imaging measurement data.

Chapter 2

Mathematical foundations for microwave imaging

As discussed in the previous chapter, the image reconstruction in microwave imaging has now been formulated as a mathematical problem, i.e. an ill-posed nonlinear optimization problem. Moreover, it has also been pointed out that there are several difficulties associated with the solution process including nonlinearity, non-uniqueness, noisy measurement data and ill-posedness. This chapter focuses on the mathematical theories that address these difficulties. First, the general mathematical relationships for the linearized solution in tomographic imaging are presented. These linearized approaches are extensively used in X-ray tomography and were exploited in the early studies of microwave imaging. The philosophy of non-linear optimization and several typical optimization methods are introduced with the emphasis on gradient-based iterative algorithms. In studying the effects of measurement noise, we devote a section to the statistical theory of parameter estimation where various estimators and their statistical meanings are explored. Finally, we explore the issue of ill-posedness and possible remedies especially in the linear case. The differences between the Tikhonov

regularization and the Levenberg-Marquardt method have also been clarified to avoid confusion.

2.1 Non-diffracting source tomography

We begin this chapter by considering the simplest tomographic imaging scenario: parallel beam X-ray CT [95, 31]. In this imaging method, X-ray beams are assumed to travel along straight lines through the target since the X-ray photon energy is so high that these beams are essentially undiffracted. Attenuation is the primary phenomenon for the X-ray beam which is proportional to the linear attenuation coefficient $\mu(\vec{r})$ of the target. Therefore, the image reconstruction problem in X-ray CT amounts to recovering $\mu(\vec{r})$ from the attenuated X-ray beams.

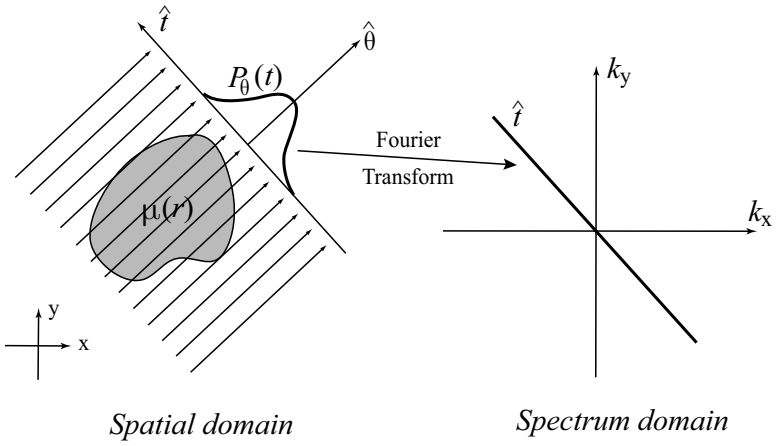


Figure 2.1: Illustration of the Fourier Slice Theorem

The mathematical principle behind the non-diffracting source tomography is referred as the "Fourier slice theorem" [95] which states: for a given 2D function $\mu(\vec{r})$ where $\vec{r}$ is a 2D vector, i.e. $\vec{r} \in \mathbb{R}^2$, the projection of $\mu(\vec{r})$ along the direction of vector $\hat{\theta}$ is denoted as

$$P_\theta(t) = \int_{\vec{r}\cdot\hat{t}=t} \mu(\vec{r}) \cdot d\vec{r} \tag{2.1}$$

where t is the coordinate along $\hat{t}$ which is the perpendicular direction to $\hat{\theta}$ (as shown in Figure 2.1). The 2D Fourier transform of $\mu(\vec{r})$ is denoted by $M(\vec{k}) = \mathscr{F}(\mu(\vec{r})) = \int \mu(\vec{r}) \exp\left(-j\vec{k}\cdot\vec{r}\right) d\vec{r}$ (note that $\mathscr{F}(f)$ is the Fourier transform of function f). Then

$$M_\theta(k_t) = \mathscr{F}(P_\theta(t)) \tag{2.2}$$

where $M_\theta(k_t)$ is defined by

$$M_\theta(k_t) = \int_{\vec{k}\cdot\vec{t}=k_t} \mu(\vec{r}) \exp\left(-j\vec{k}\cdot\vec{r}\right) d\vec{r} \tag{2.3}$$

Figure 2.1 illustrates the relationship between the the projection and the image of the object in spectrum space. By taking measurements of the projected X-ray attenuations, i.e., $P_\theta(t)$, at various angles, the image of $M(\vec{r})$ is filled in spectral space as the result of performing the transforms in (2.1). Consequently, an inverse Fourier transform must be applied to $M(\vec{r})$ to recover the original μ distribution from its spectral representation. Considering the previous steps, the recovered image can be written as

$$\begin{aligned} \mu(\vec{r}) &= \mathscr{F}^{-1}\left(M(\vec{k})\right) \\ &= \int_0^\pi \left(\int_{-\infty}^{\infty} M_\theta(k_t) dk_t\right) d\theta \\ &= \int_0^\pi \left(\int_{-\infty}^{\infty} \mathscr{F}(P_\theta(t))\, dk_t\right) d\theta \end{aligned} \tag{2.4}$$

Relationship (2.1) and (2.4) are referred as the Radon and inverse Radon transforms [31], respectively. These transforms constitute the most fundamental relationships for non-diffracting source projection imaging.

It is interesting to note that the integrations in the Radon transform pair represent linear relationships between the function μ and the projection $P_\theta(t)$ in the functional space, i.e., if we denote the transform on the right-hand-sides (RHS's) of (2.1) and

(2.4) as $\mathscr{R}$ and $\mathscr{R}^{-1}$, respectively, then the following relationships are satisfied

$$\begin{aligned} \mathscr{R}\left(\alpha\mu(\vec{r})\right) &= \alpha\mathscr{R}\left(\mu(\vec{r})\right) \\ \mathscr{R}\left(\mu_1(\vec{r})+\mu_2(\vec{r})\right) &= \mathscr{R}\left(\mu_1(\vec{r})\right)+\mathscr{R}\left(\mu_2(\vec{r})\right) \end{aligned} \tag{2.5}$$

where α is a scalar and μ_1 ad μ_2 are arbitrary property distributions. Similar results can be obtained for the inverse Radon transform $\mathscr{R}^{-1}$. Thus, the discretization forms of these equations lead to linear matrix equations. This linear relationship helps make the X-ray CT so attractive because of the simplicity in determining the solutions.

2.2 Diffracting source tomography

In diffracting source tomography, the response at a single detector is not only related to the properties along the straight line between the source and the detector, but to all of the properties over the target space [95]. This section discusses a simplified case: the far field measurement system, where both sources and receivers are electrically far away from the object. Because of this assumption, the incident and scattered waves can be treated as plane waves. Moreover, a weak scatterer assumption is also assumed which implies that the total scattered field can be approximated by the unknown incident field as in Equation (1.9).The following derivation [31] is provided to demonstrate the differences between the non-diffracting source and diffracting source imaging.

A diagram of such an imaging configuration is shown in Figure 2.2. An object is placed near the origin O. A point source is placed at location $\vec{r}_T$ and the detector is at $\vec{r}_R$. Based on the previous assumptions, $|\vec{r}_T| \gg |\vec{r}'|$ and $|\vec{r}_R| \gg |\vec{r}'|$ where $\vec{r}'$ is located inside the object. The scattered field at a far field point $\vec{r}$ from the object is represented

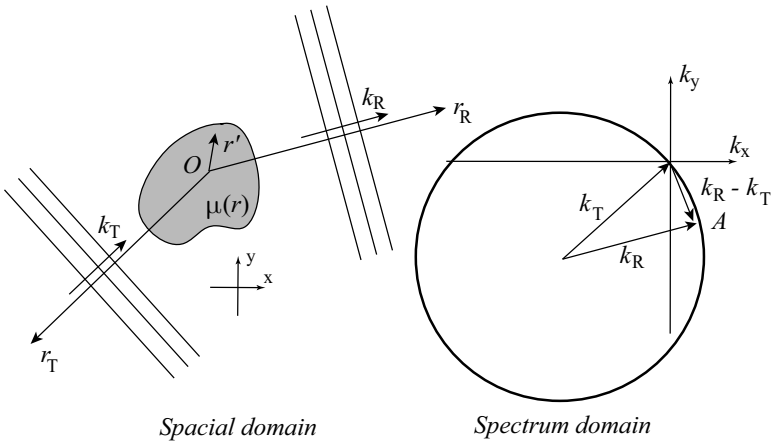


Figure 2.2: Illustration for far-field diffracting source tomography

by equation (1.9) where the 2D Green's function is expanded as

$$g(\vec{r},\vec{r'}) = \frac{j}{4}H_0^{(1)}(k_0|\vec{r}-\vec{r'}|) \tag{2.6}$$

with $H_0^{(1)}$ being the 0-th order Hankel function of the first kind. From the far field assumption, $g(\vec{r},\vec{r'})$ is subsequently approximated by a plane wave as

$$g(\vec{r},\vec{r'}) \approx \frac{j}{4}\sqrt{\frac{2}{j\pi k_0 r}}\exp(jk_0(r_R - \hat{r}\cdot\vec{r'})) \tag{2.7}$$

by performing the asymptotic expansion of the Hankel function. In addition, the incident field at location $\vec{r'}$ generated by the point source located at $\vec{r}_T$ is also a 2D Green's function in the form of

$$\phi_{inc}(\vec{r'},\vec{r}_T) = \frac{j}{4}H_0^{(1)}(k_0|\vec{r'}-\vec{r}_T|) \tag{2.8}$$

Therefore, the plane wave approximation of $\phi_{inc}(\vec{r'},\vec{r}_T)$ produces a result similar to that in (2.7), i.e.

$$\phi_{inc}(\vec{r'},\vec{r}_T) \approx \frac{j}{4}\sqrt{\frac{2}{j\pi k_0 r}}\exp(jk_0(r' - \hat{r}'\cdot\vec{r}_T)) \tag{2.9}$$

Combining (2.7) and (2.9) into (1.9), one can evaluate the scattered field at receiver $\vec{r}_R$ by

$$\phi_{sca}(\vec{r}_R) \approx \frac{j}{8\pi\sqrt{r_T r_R}} \exp(jk_0(r_T + r_R)) M(\vec{k}_R - \vec{k}_T) \tag{2.10}$$

where $M(\vec{k})$ is the Fourier transform of the scatterer $\mu(\vec{r})$, $\vec{k}_R = k_0\vec{r}_R$ and $\vec{k}_T = k_0\vec{r}_T$.

Equation (2.10) represents the relationship between the measurement and the image in Fourier space. For a given source/receiver pair, the measurement $\phi_{sca}(\vec{r}_R)$ is associated with a point in the spectrum space located at $\vec{k}_R - \vec{k}_T$ (denoted by point A in Figure 2.2). If the receiver is moved around the object while fixing the transmitter, the locus of point A becomes a circle centered at $-\vec{k}_T$ with radius $|\vec{k}_T| = k_0$. If the movement of the transmitter around the object is incorporated simultaneously, the set of circles sweep a circular area in the spectrum space which has radius $2k_0$ (as shown in Figure 2.3).

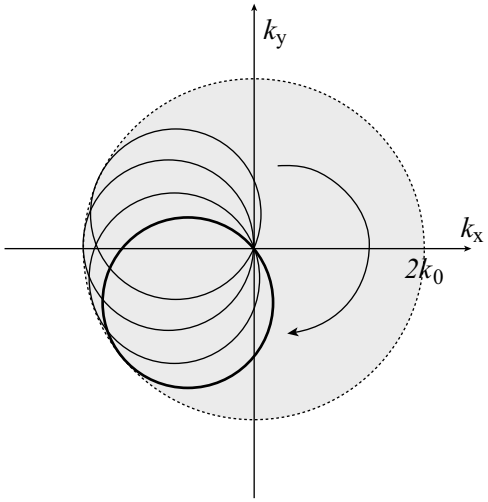


Figure 2.3: Area swept by varying the directions of $\vec{k}_R$ and $\vec{k}_T$

From plot 2.3, several conclusions can be drawn for linearly-approximated diffraction tomography. First, this imaging method is unable to distinguish features whose spatial spectrum is greater than twice that of the incident wave number k_0. This is to

say, if the Fourier transform of two images have identical distributions inside the circle $r = 2k_0$ in $\mathscr{F}$ space but different outside the circle, the images reconstructed from their projections will be identical. This implies non-unique solutions for a given measurement data set in diffraction tomography. The spatial frequency $2k_0$ is referred as the diffraction limit of the imaging system which is consistant with the sampling theorem due to Shannon [148]. Secondly, as the frequency of the incident wave increases toward infinity, i.e. $k_0 \to \infty$, the curvature of the circle approaches zero and the length of the arc extends to infinity. The circle is therefore deformed into a straight line passing through the origin. In this situation, the diffraction relationship (2.10) becomes the relationship in the Fourier slice theorem. This implies that non-diffracting source tomography is a special case of diffracting source tomography at the high frequency limit.

In the above derivations, the term "weak scatterer" was used. Once again, the consequence of this assumption is the linear functional relationship between the measurement and the image similar to that in the non-diffracting source case. Therefore, the image can be efficiently reconstructed by solving linear equations. However, when the scatterer has significant contrast to the background, the assumption is not valid and the relationship between the object and measurement is no longer linear and nonlinear methods need to be used to solve for the image.

Another assumption employed was the far field assumption. If the measurements are acquired in the near field, the integration in (2.10) becomes far more complicated. The detailed derivations can be found in [95].

2.3 Nonlinear optimization

Linearity from a mathematical standpoint implies two things, additivity and homogeneity, i.e.

$$\begin{aligned} f(\alpha \mathbf{x}) &= \alpha f(\mathbf{x}) \\ f(\mathbf{x}+\mathbf{y}) &= f(\mathbf{x}) + f(\mathbf{y}) \end{aligned} \tag{2.11}$$

The function f in (2.11) refers to an operator in a very general sense. The operant $\mathbf{x}$ can be a number, a vector, a matrix, a function, or any other meaningful object. Loosely speaking, linearity means that we can decompose a complex problem into small subproblems which are relatively simple to solve and independent of each other. Unfortunately, as we approach the more realistic nature of the world, more complexities and nonlinearities are introduced. Understanding these nonlinearities is a great challenge for the advancement of both modern science and technology. As a result, nonlinear mathematics has undergone swift growth in recent decades. New disciplines have been created including nonlinear dynamics, chaos, fractal theory, self-organization theory, catastrophe theory and so on. Nonlinear physics and chemistry have also been benefited from the boom of mathematical tools. The invention and development of computers along with advanced numerical algorithms has dramatically increased investigations into nonlinear phenomena. The combination of numerical techniques and nonlinear science inspired a large variety of methods for real-world applications among which nonlinear optimization is one of the most important. Nontechnically speaking, the problem of optimization is a search for the "most appropriate" solution among all possibilities. The technical term referring to "appropriateness" is called the object function; the term for "all possibilities" is the feasible space [91, 164].

An optimization can be mathematically expressed in the following form [91]

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_1(\mathbf{x}) = 0 \\ & g_2(\mathbf{x}) \geq 0 \end{aligned} \tag{2.12}$$

where f is a scalar-valued function. The "s.t." statement denotes the constraints or conditions of the optimization which may not always appear in optimization problems. Optimizations without the "s.t." statement are referred to as non-constrained optimizations, while the remainder are constrained optimizations. The presence of a constraint limits the boundaries of the feasible space within the parameter space. In other words, it provides *a priori* information concerning possible solutions to the problem.

Optimization is a process of searching. Based on the manner of the search, optimization methods are classified into direct and indirect methods. Direct methods enumerate the possible solutions and make comparisons of the corresponding object function values. Common direct methods include exhaustive search, the simplex method, random method, the Fibonacci search and more [91]. These methods employ relatively simple searching patterns which allow the evaluation of the object function in large numbers. However, direct methods are only efficient when there are small number of variables or in cases with countable inputs. For continuous problems with complicated object functions, the implementation of these methods is unpractical. In the last few decades, a much more powerful direct optimization method, the genetic algorithm (GA), was proposed for searching the global optimum for complicated object functions [84]. This algorithm exploits analogous concepts of "genetic mutations" and "genetic inheritances" from biological systems. This strategy results in greatly improved efficiency in the direct search process. GA has been successfully used in many applications with microwave imaging being one of them [25].

Indirect methods are most often referred to as gradient-based optimization techniques [91, 164]. By taking advantage of the known analytical form of the object function, these methods approach the optimum solution in an iterative manner guided by the analytical or numerical gradient information. Steepest decent (SD) method, Newton's method, Gauss-Newton (GN) method, Levenberg-Marquardt (LM) method, conjugate gradient (CG) method and many of their variants belong to this category.

The steepest-descent method is possibly the most straightforward among all of the indirect methods. At each iteration, the SD method seeks the negative gradient direction of the object function and moves the current parameter estimate along that direction attempting to reach the minimum. If we assume the object function is a multi-variable function denoted by f and all its variables consist of a vector $\mathbf{x}$, then at the i-th iteration, the update equation of the SD method can be written as:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - \delta^{(i)} \left(\nabla f(\mathbf{x}^{(i)}) \right) \tag{2.13}$$

where $\nabla f(\mathbf{x}^{(i)})$ is the gradient vector of the object function at point $\mathbf{x}^{(i)}$ defined by

$$\nabla f(\mathbf{x}^{(i)}) = \left\{ \frac{df}{dx_\tau} \right\}_{\tau=1}^{n} \tag{2.14}$$

where n is the length of vector $\mathbf{x}$. The step size $\delta^{(i)}$ is determined by a one-dimensional search denoted by

$$\delta^{(i)} = \arg\min_{\delta} f(\mathbf{x}^{(i)} - \delta \left(\nabla f(\mathbf{x}^{(i)}) \right)) \tag{2.15}$$

The advantages of the SD method include 1) simplicity of implementation, and 2) robustness for poor initial estimates. However, the SD method suffers from several disadvantages including slow convergence, implicitness in determining the step size and

oscillation during convergence.

The Newton's method follows a similar path but involves the second order derivative of the object function. At the i-th iteration, assuming the Taylor's expansion of the function f at $\mathbf{x}^{(i)}$ is expressed as

$$f(\mathbf{x}) = f(\mathbf{x}^{(i)}) + \nabla f(\mathbf{x}^{(i)})(\mathbf{x} - \mathbf{x}^{(i)}) + H_f(\mathbf{x}^{(i)})(\mathbf{x} - \mathbf{x}^{(i)})^2 + \cdots \tag{2.16}$$

where $H_f(\mathbf{x}^{(i)})$ is referred as the Hessian matrix denoting the second order derivative defined by

$$H_f(\mathbf{x}) = \begin{pmatrix} \frac{df^2}{d^2x_1} & \frac{df^2}{dx_1dx_2} & \cdots & \frac{df^2}{dx_1dx_n} \\ \frac{df^2}{dx_2dx_1} & \frac{df^2}{d^2x_2} & \cdots & \frac{df^2}{dx_2dx_n} \\ \frac{df^2}{dx_ndx_1} & \frac{df^2}{dx_ndx_2} & \cdots & \frac{df^2}{d^2x_n} \end{pmatrix} \tag{2.17}$$

The minimization of f can be achieved by evaluating the stationary point from equation (2.16) where

$$\frac{df}{d\mathbf{x}} = \mathbf{0} \tag{2.18}$$

Substituting (2.16) into (2.18) and truncating the Taylor series, the update equation for the Newton's method can finally be expressed as

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - \left(H_f(\mathbf{x}^{(i)})\right)^{-1} \nabla f(\mathbf{x}^{(i)}) \tag{2.19}$$

Newton's method converges faster than the steepest-decent method because of the contribution from the higher order information. However, the price paid is the reduction in robustness, i.e. it is more sensitive to poor initial estimates than the steepest-decent method. An additional drawback is the requirement for computing the Hessian matrix H which can be a significant issue in some applications where the analytical form of the object function is not available or has a complicated form.

For a specific subset of optimization problems, a useful approach is the least-square problems:

$$\min_{\mathbf{x}} f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|^2 = \mathbf{g}^T(\mathbf{x})\mathbf{g}(\mathbf{x}) \tag{2.20}$$

where $\mathbf{g} = \{g_i(\mathbf{x})\}_{i=1}^m$ is a vector of the scalar functions of $\mathbf{x}$. The Gauss-Newton (GN) method is the most frequently used. In this case, the Newton's method is modified by replacing the Hessian matrix, $H_f(\mathbf{x}^{(i)})$ in (2.19), by multiplication of two first order derivatives (the Jacobian matrix) of function $\mathbf{g}$ defined by

$$J_{\mathbf{g}}(\mathbf{x}) = \begin{pmatrix} \frac{dg_1}{dx_1} & \frac{dg_1}{dx_2} & \cdots & \frac{dg_1}{dx_n} \\ \frac{dg_2}{dx_1} & \frac{dg_2}{dx_2} & \cdots & \frac{dg_2}{dx_n} \\ \frac{dg_m}{dx_1} & \frac{dg_m}{dx_2} & \cdots & \frac{dg_m}{dx_n} \end{pmatrix} \tag{2.21}$$

In this case, the "Gauss-Hessian" matrix G has form $G = \left(J_{\mathbf{g}}(\mathbf{x}^{(i)})^T J_{\mathbf{g}}(\mathbf{x}^{(i)})\right)$. The update equation can now be written as

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - \left(J_{\mathbf{g}}(\mathbf{x}^{(i)})^T J_{\mathbf{g}}(\mathbf{x}^{(i)})\right)^{-1} J_{\mathbf{g}}(\mathbf{x}^{(i)})^T \mathbf{g}(\mathbf{x}^{(i)}) \tag{2.22}$$

or

$$\left(J_{\mathbf{g}}(\mathbf{x}^{(i)})^T J_{\mathbf{g}}(\mathbf{x}^{(i)})\right)(\Delta\mathbf{x})^{(i)} = -J_{\mathbf{g}}(\mathbf{x}^{(i)})^T \mathbf{g}(\mathbf{x}^{(i)}) \tag{2.23}$$

where $\Delta\mathbf{x}^{(i)} = \mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}$. Equation (2.23) is referred as the normal equation [67]. In the GN method, one only needs to compute $J_{\mathbf{g}}(\mathbf{x}^{(i)})$ instead of H_f which provides a significant computation time saving.

Notice that the inversion of the Gauss-Hessian matrix in (2.22) requires that 1) the Jacobian matrix $J \in \mathbb{C}^{m\times n}$ is a "skinny" matrix, i.e. the number of unknowns, n, is fewer than the number of constraints, m, and 2) the Jacobian matrix is full-ranked, i.e. $\text{rank}(J) = n$. In this case, the set of least-square problem (2.20) is referred as

over-determined. If the Jacobian matrix is a full-ranked "fat" matrix, i.e. $m < n$ and $\text{rank}(J) = m$, then, (2.22) and (2.23) are not valid. In this case, the update equation should be written as

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - J_{\mathbf{g}}(\mathbf{x}^{(i)}) \left(J_{\mathbf{g}}(\mathbf{x}^{(i)}) J_{\mathbf{g}}(\mathbf{x}^{(i)})^T \right)^{-1} \mathbf{g}(\mathbf{x}^{(i)}) \tag{2.24}$$

Equation (2.24) is referred as the Gauss-Newton update equation for under-determined least square problems.

Both the Newton's method and the Gauss-Newton's method exhibit oscillatory features during their convergence and are both sensitive to the quality of their initial estimates. Levenberg [108] and Marquardt [121] proposed a hybrid technique, the Levenberg-Marquardt (LM) method. They introduced a steering factor λ to switch between the GN direction and SD direction. The update equation in LM is written as

$$\left(J_{\mathbf{g}}(\mathbf{x}^{(i)})^T J_{\mathbf{g}}(\mathbf{x}^{(i)}) + \lambda I \right) (\Delta \mathbf{x})^{(i)} = -J_{\mathbf{g}}(\mathbf{x}^{(i)})^T \mathbf{g}(\mathbf{x}) \tag{2.25}$$

When $\lambda \to 0$, equation (2.25) reduces to (2.23) and the LM method becomes the GN method. When $\lambda \to \infty$, the $J^T J$ term is omitted and the RHS provides the steepest-decent direction of the object function f. The value of λ for each iteration is chosen in the following manner: λ is initialized with a large value. Thus, during the first few iterations, the LM method exhibits the robustness of the SD method so that the initial guess for $\mathbf{x}$ can be chosen with less caution. After each iteration, if the object function decreases, i.e. $f(\mathbf{x}^{(i+1)}) < f(\mathbf{x}^{(i)})$, λ is reduced by a given ratio (for example $\lambda^{(i+1)} = \lambda^{(i)}/2$) to accelerate the convergence. If the object function rises, λ is increased.

2.4 Ill-posedness and regularization

As we have explained in Section 1.3, the inverse problem from a system point of view is to estimate the unknown system parameter from a known input and output. However, a large number of inverse problems exhibit inherent difficulties and the direct evaluation of these problems often produces meaningless solutions. Quite often, one of the following results can occur in the evaluation of the goal of the inverse problem: 1) the solution does not exist, 2) the solution is not unique, or 3) solving for the solution is not stable, i.e. a tiny perturbation in the input results in large differences in the solution. If any of the above results occurs, the inverse problem is said to be ill-posed (in the Hadamard sense) [71, 138]. A more thorough and strict description of ill-posedness can be found in the monograph by Lavrent'et [104].

2.4.1 Linear ill-posed problems

Even linear problems can be ill-posed. A linear ill-posed problem is characterized by the matrix equation

$$A\mathbf{x} = \mathbf{b} \tag{2.26}$$

where the evaluation of the solution $\mathbf{x}$ is sensitive to the noise in $\mathbf{b}$. The sensitivity is characterized by the singular value decomposition (SVD) of the A matrix. For an arbitrary complex valued matrix $A \in \mathbb{C}^{m\times n}$, the SVD of A is given by

$$A = U\Sigma V^T \tag{2.27}$$

where $U = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\} \in \mathbb{C}^{m\times n}$ and $V = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\} \in \mathbb{C}^{n\times n}$ are both column orthogonal matrices, i.e. $U^H U = V^H V = I$. Σ is a diagonal matrix with non-negative

values arranged in non-increasing order, i.e. $\Sigma = \text{diag}\left(\{\sigma_i\}_{i=1}^n\right)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are referred as the i-th left and right singular vectors, respectively, while σ_i is the i-th singular value. The sequence $\{\sigma_1, \sigma_2, \cdots, \sigma_n\}$ is referred as the singular spectrum of matrix A.

Similar to the eigenvalue decomposition for square matrices, the SVD of an arbitrary matrix reveals detailed structure of the corresponding linear transform [67, 77]. With the help of the SVD, a matrix can be expanded into the summation of a number of matrices

$$A = \sum_{i=1}^{n} \sigma_i A_i \tag{2.28}$$

where $A_i = \mathbf{u}_i \mathbf{v}_i^T$. As a result, a linear transformation is equivalent to a sequence of linear transformations with a decreasing magnitude in contribution (notice that coefficients σ_i are non-increasing in (2.28)). The solution to (2.26) is then written as

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i \tag{2.29}$$

Relationship (2.29) is referred as the singular value expansion (SVE) of the solution from which multiple results can be observed. First, the solution is now a summation of n components instead of simply $A^{-1}\mathbf{b}$ or $A^{+}\mathbf{b}$ (A^{+} denotes the pseudo-inversion of A matrix). Secondly, for each component of the series, the singular value σ_i is in the denominator which means a very small value in σ_i can potentially have a large impact on the solution. Thirdly, since each term, $\frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i}$ is a scalar coefficient which leaves the solution $\mathbf{x}$ as essentially a linear combination of $\mathbf{v}_i$'s. If $\mathbf{x}$ represents an image, then the right singular vectors $\mathbf{v}_i$ ($i = 1, \cdots, n$) are the orthogonal bases comprising the image.

The spectrum $\{\sigma_i\}$ provides a natural measurement for the sensitivity of the solution $\mathbf{x}$ with respect to the input $\mathbf{b}$ given the SVE form (2.29) of the solution. Based on

the shape of the spectrum $\{\sigma_i\}$, matrix equations (2.26) can be classified into one of the following categories: a full-rank and well-posed matrix equation, an exact rank-deficient matrix equation or an (discretized) ill-posed matrix equation. The typical singular spectra for these problems are plotted in Figure 2.4 [77]. From the figure, we can see an ill-posed problem is characterized by gradually vanishing spectrum which indicates the constrained equations have increasing amounts of redundancy. An ill-

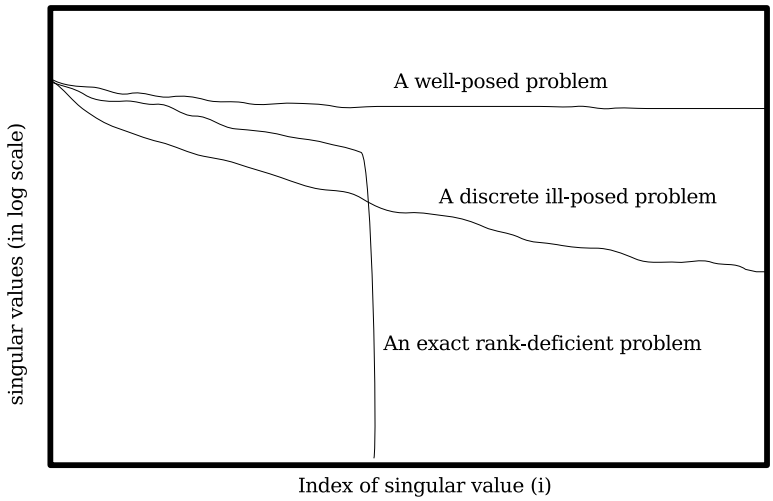


Figure 2.4: Classification of the linear equations based on the distributions of the singular spectra.

posed problem must be converted into a well-posed problem in order to be solved. The technique for this conversion is called "regularization". For a linear ill-posed problem, the following methods are among the most commonly used: 1) the truncated SVD (TSVD), 2) Tikhonov regularization and 3) the truncated total least square (TTLS). A short description of each method is provided below.

As can be observed from equation (2.29), complications arise from the terms with small singular values at the end of the series. Therefore, the most straightforward way is to eliminate these components directly from the summation by setting a truncation level. Let $0 \leq k \leq n$ be the number of terms one wants to include in the summation, the

solution

$$\mathbf{x}_k = \sum_{i=1}^{k} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i \tag{2.30}$$

is referred as the TSVD solution to the problem (2.26).

Very often, equation (2.30) is rewritten in the following form

$$\mathbf{x}_k = \sum_{i=1}^{n} f(\sigma_i) \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i \tag{2.31}$$

where f is called a filtering function. For the TSVD, the filtering function is simply an ideal low-pass filter characterized by

$$f(\sigma_i) = \begin{cases} 1 & i \leq k \\ 0 & i > k \end{cases} \tag{2.32}$$

(the assumption is that the large singular values correspond to the low-frequency components in the solution as shall be illustrated in Chapter 7).

To solve a linear ill-posed problem using the TSVD, the SVD of the LHS matrix A must be computed. This is not realistic when A is very large since the most efficient algorithm for computing the SVD of a general matrix has computational complexity $O(N^3)$ [67, 12]. Tikhonov proposed an alternative algorithm to mitigate the ill-posedness without performing the SVD of A, called the Tikhonov regularization [191]. The solution to (2.26) with Tikhonov regularization is characterized by solving the following optimization problem:

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \|L(\mathbf{x} - \mathbf{x}^*)\|_2^2 \tag{2.33}$$

This object function is similar to the regularized form in functional space as demonstrated in (1.13), where the smoothing norm is achieved by utilizing the positive def-

inite matrix L and *a priori* solution $\mathbf{x}^*$. λ is a scalar referred to as the regularization parameter. The evaluation of the minimization problem (2.33) can be performed by taking the derivative of the object function and setting it equal to zero yielding

$$A^H(A\mathbf{x} - \mathbf{b}) + \lambda L^H L(\mathbf{x} - \mathbf{x}^*) = \mathbf{0} \tag{2.34}$$

and consequently

$$(A^H A + \lambda L^H L)\mathbf{x} = A^H \mathbf{b} + \lambda L^H \mathbf{x}^* \tag{2.35}$$

If no prior information is available, $\mathbf{x}^*$ is typically set to zero. Matrix L can be any arbitrary positive definite matrix. If the statistical properties of the estimated parameter are available, setting $L^T L$ to the inverse of the covariance matrix of random variable $\mathbf{x}$ can provide a good estimator (refer to Section 2.5 for more details). Otherwise, the L matrix is often chosen to be the identity matrix I which statistically implies that the parameters under estimation have constant variances and are independent of each other. In this case, (2.35) becomes

$$(A^H A + \lambda I)\mathbf{x} = A^H \mathbf{b} \tag{2.36}$$

Equation (2.36) is called the standard form of the Tikhonov regularization, which has been shown to be equivalent to the solution in the form of (2.31) with the filtering function defined by [77]

$$f(\sigma_i) = \frac{\sigma_i^2}{\sigma_i^2 + \lambda} \tag{2.37}$$

Application of this filter clearly avoids the sensitivity problem caused by small singular values (noting that when $\sigma_i \to 0$, $f(\sigma_i) \to 0$) while preserving the information corresponding to the large singular values (when $\sigma_i \gg \lambda$, $f(\sigma_i) \to 1$). Moreover, the computational complexity for solving equation (2.36) is only on the order of $O(N^2)$

which provides a significant time saving for problems with large numbers of unknowns.

Notice that when the linear problem is ill-posed, the rank of the matrix A is characterized by its "effective numerical rank", r_{eff}, as defined in [77]. In many cases, r_{eff} is smaller than both the number of unknowns and measurements. Therefore, the matrices $A^H A$ and AA^H are both "numerically" rank deficient. However, the matrices $A^H A + \lambda I$ and $AA^H + \lambda' I$ could both have full-rank for appropriate values of λ and λ'. In this case, the solution can be computed by either

$$\mathbf{x} = (A^H A + \lambda I)^{-1} A^H \mathbf{b} \tag{2.38}$$

or

$$\mathbf{x} = A^H (AA^H + \lambda' I)^{-1} \mathbf{b} \tag{2.39}$$

and the results have only nominal difference. Considering the computational efficiency, when A is a "skinny" matrix, i.e. the number of rows is greater than that of the columns, the solution computed by (2.38) will be significantly faster than by (2.39); alternatively, when A is a "fat" matrix, (2.39) is faster than (2.38) (In the reconstruction algorithms presented in the second part of the thesis, we always choose the faster scheme by looking at the dimension of the matrix ahead of the solution).

A third method for regularizing linear ill-posed problems is the truncated total least square (TTLS) method. The concept of the total least-square (TLS) was developed by Golub and Van Loan in 1980 [66, 67]. The TLS method can produce a robust solution when the A matrix and $\mathbf{b}$ are both contaminated with noise whereas traditional methods only consider the noise in $\mathbf{b}$. The evaluation of the TTLS solution for (2.26) requires

computation of the SVD of the extended matrix $\bar{A}$ defined by

$$\bar{A} = [A, \mathbf{b}] \tag{2.40}$$

essentially appending column vector $\mathbf{b}$ to A. Matrix $\bar{A}$ has dimension $m \times (n+1)$. Given the SVD of $\bar{A} = U\Sigma V^T$, the solution to (2.26) can be expressed as

$$\mathbf{x}_{TTLS} = -V_{12} \cdot V_{22}^T \|V_{22}\|_2^{-2} \tag{2.41}$$

where V_{12} and V_{22} are the sub-matrices of V defined by

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \tag{2.42}$$

Despite the robustness of this method, the evaluation of the solution can be very slow when A is a large matrix due to the computation time in evaluating the SVD.

Notice that all three of these methods require the determination of a regularization parameter, either λ or k. The study of optimal regularization parameters is an active area of study in ill-posed problems. Thus far, there are several of methods widely used for this purpose. These include: the discrepancy principle due to Morozov [137, 192], the L-curve method proposed by Hansen [105, 77], the GCV method proposed by Wahba [199] and the quasi-optimal criteria proposed by Hanke and Raus [75]. Detailed studies and comparisons of these criteria can be found in [77] and [197].

2.4.2 Nonlinear ill-posed problems

When the relationship between the unknown and the output is a nonlinear equation, ill-posedness can also be present by exhibiting one of the symptoms described at the

beginning of this section. In these situations, the regularization methods in the previous subsection cannot be applied directly. Here we shall demonstrate the regularization for ill-posed nonlinear problems using the Tikhonov regularization as an example.

Similar to (2.33), the object function in the non-linear cases can be written as

$$\min_{\mathbf{x}} \|\mathbf{f}(\mathbf{x}) - \mathbf{b}\|_2^2 + \lambda \|L(\mathbf{x} - \mathbf{x}^*)\|_2^2 \tag{2.43}$$

where function $\mathbf{f}(\mathbf{x})$ is a vector of nonlinear functions in $\mathbf{x}$, i.e. $\mathbf{f}(\mathbf{x}) = \{f_1(\mathbf{x}, f_2(\mathbf{x}), \cdots, f_m(\mathbf{x}))\}$. Applying Taylor's expansion to $\mathbf{f}(\mathbf{x})$ at a given value of the parameter $\mathbf{x}_0$, $\mathbf{f}$ can be written as

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \left.\frac{d\mathbf{f}}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0) + \cdots \tag{2.44}$$

By truncating the series after the second order term, a linear relationship with respect to $\mathbf{x}$ is obtained. Denoting $\left.\frac{d\mathbf{f}}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}$ as J and substituting the truncated series into (2.43) produces

$$\min_{\mathbf{x}} \|\mathbf{f}(\mathbf{x}_0) + J(\mathbf{x} - \mathbf{x}_0) - \mathbf{b}\|_2^2 + \lambda \|L(\mathbf{x} - \mathbf{x}^*)\|_2^2 \tag{2.45}$$

Taking the derivative to the object function in (2.45) and setting it to zero yields

$$J^T \left(\mathbf{f}(\mathbf{x}_0) - \mathbf{b} + J\left(\mathbf{x} - \mathbf{x}_0\right)\right) + \lambda L^H L \left(\mathbf{x} - \mathbf{x}^*\right) = \mathbf{0} \tag{2.46}$$

which is expanded to produce

$$(J^T J + \lambda L^H L)\mathbf{x} = J^T \left(\mathbf{f}(\mathbf{b} - \mathbf{x}_0)\right) + J^T J \mathbf{x}_0 + L^H L \mathbf{x}^* \tag{2.47}$$

By setting $L = I$ and $\mathbf{x}^* = \mathbf{0}$ for the standard form, we get

$$(J^T J + \lambda I)\Delta\mathbf{x} = J^T \Delta\mathbf{f} \tag{2.48}$$

where $\mathbf{\Delta x} = \mathbf{x} - \mathbf{x}_0$ and $\mathbf{\Delta f} = \mathbf{b} - \mathbf{f}(\mathbf{x_0})$.

Although (2.48) is similar to (2.36), the unknown solved for in the former is $\mathbf{\Delta x}$ while $\mathbf{x}$ is solved for in the latter, which is an important distinction.

2.4.3 Differences between the Levenberg-Marquardt method and nonlinear Tikhonov regularization

Another similarity can be observed between (2.36)/(2.48) and the Levenberg-Marquardt update equation (2.25). Although these two concepts share the very similar underlying methodologies, i.e. dynamically adjusting the radius of the trust region to improve the robustness of the solution [91], the differences between these two concepts should be clarified to avoid future confusions.

First, although both methods can be viewed as optimization processes, the criteria in the optimizations are different. The LM method is essentially an improved Gauss-Newton method which minimizes the traditional sum-of-square object function. However, in ill-posed problems, the direct optimization of the sum-of-square function is impossible which requires one to modify the object function by adding the regularization term. In this case, the Tikhonov regularization differs from the LM method by having a modified object function. Secondly, they have different purposes. The LM method is intended to conqueror the nonlinearity in a root finding problem or a minimization problem. Tikhonov regularization, on the other hand, is a method which is targeted at relieving the ill-posedness, which may occur in both linear and nonlinear problems. Thirdly, the criteria for selecting and manipulating the parameter λ are different. In the LM method, the value of λ is selected to reduce the oscillations during the convergence. At the beginning of the convergence, λ can be chosen to be arbitrarily large, and is reduced towards zero at each iteration for a well-behaved process. However,

the parameter in the Tikhonov regularization can not be chosen arbitrarily. Generally speaking, the regularization parameter λ can be optimally determined by the L-curve method, generalized cross-validation (GCV) method or others listed in the previous subsection.

2.5 Nonlinear parameter estimation

Before continuing to the statistical aspects of this problem, a broad summary of the previous sections is presented here. In order to reconstruct the spatial distribution (or image) of the unknown properties, one needs to solve an ill-posed nonlinear optimization problem which is best addressed using iterative gradient-based approaches. The object function contains both the measurement and the computed fields from the nonlinear forward model. Beginning with an initial property estimate, the forward model needs to be evaluated to produce the "predicted" measurement. An update of the property is then computed from the difference between the predicted and actual measurements utilizing the gradient information of the forward model. This process is repeated until the satisfactory match is found between the predicted and actual measurements.

It is interesting to note that the reconstruction problem solution described above only requires a deterministic forward model and some measurement data. However, this is simplistic in that both the model or measurements could be uncertain and is actually the case for most real-world imaging systems. The measurement readings from instruments are usually contaminated by noise; the property to be imaged is also subject to random fluctuations due to environmental changes. To accommodate these variations, we need to apply the statistical theory of parameter estimation.

The following quotation in the foreword of "Computational Methods for Inverse

Problems" by Vogel (foreword by H. T. Banks) [197] illustrated some historical relationships between deterministic and stochastic approaches in the parameter estimation:

> "It is not surprising that there is a large mathematical literature on inverse problem methods. What might be surprising is that this literature is significantly divided along deterministic/nondeterministic lines. Methods abound in the statistics literature, where generally the models are assumed quite simple (and often even analytically known!) and the emphasis is on treating statistical aspects of fitting models to data. On the other hand, the applied mathematical literature has a plethora of increasing complex parameterized models (nonlinear ordinary differential equations, partial differential equations, and delay equations) which are treated theoretically and computationally in a deterministic framework with little or no attention to inherent uncertainty in either the modelled mechanisms or the data used to validate the model."

Given the randomness in the measurements, the best parameter values estimated from the deterministic framework may vary when supplied with different data sets. One important task for parameter estimation is to provide a "fair" estimate for all possible measurements [7].

The selection of the "best" estimate once again becomes an optimization problem. To reformulate this problem in the statistical sense, let's assume that the possible measurements at the receivers are a vector of m random variables and is denoted by $\boldsymbol{\eta}$; the parameter is is represented by another variable vector of length p denoted by $\boldsymbol{\beta}$; and the relationship (or model) between these two variables can be expressed by a nonlinear function

$$\boldsymbol{g}(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\tau}) = 0 \tag{2.49}$$

where $\boldsymbol{g}$ represents a group of relationships that connect $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$, and symbol $\boldsymbol{\tau}$ denotes the non-stochastic parameters or constants. In many cases, $\boldsymbol{\eta}$ can be expressed explicitly as

$$\boldsymbol{\eta} = \boldsymbol{f}(\boldsymbol{\beta}) \tag{2.50}$$

where both $\boldsymbol{f}($ and $\boldsymbol{g}($ are referred to as models. In this case, the model is assumed to be a nonlinear deterministic relationship to avoid the more complex case where it is uncertain. Taylor expansion of $\boldsymbol{f}(\boldsymbol{\beta})$ provides similar results to that in the deterministic problem (2.16) as

$$\boldsymbol{f}(\boldsymbol{\beta}) = \boldsymbol{f}(\boldsymbol{\beta}_0) + (\boldsymbol{\nabla}\boldsymbol{f}(\boldsymbol{\beta}_0))^T(\boldsymbol{\beta} - \boldsymbol{\beta}_0) + \cdots \tag{2.51}$$

The task of this parameter estimation problem is to find an estimator $\hat{\boldsymbol{\beta}}$ to the true value of the parameter $\boldsymbol{\beta}$ which minimizes a sum-of-square function [7]

$$S = (\mathbf{Y} - \boldsymbol{f}(\boldsymbol{\beta}))^T W(\mathbf{Y} - \boldsymbol{f}(\boldsymbol{\beta})) + (\boldsymbol{\mu} - \boldsymbol{\beta})^T U(\boldsymbol{\mu} - \boldsymbol{\beta}) \tag{2.52}$$

where $\mathbf{Y}$ is the actual measurement vector (a sample or realization of random variable $\boldsymbol{\eta}$), W and U are symmetric weighting matrices and $\boldsymbol{\mu}$ is the *a priori* estimation of the parameter.

The minimization of the sum-of-square object function S involves the non-linear optimization techniques discussed in Section 2.3. The Gauss-Newton method is selected for its efficiency in solving least-square problems. Following the derivations from (2.20) to (2.23), the update equation of the nonlinear estimation problem can be written as

$$\mathbf{b}^{(i+1)} = \mathbf{b}^{(i)} + \left(J^T W J + U\right)^{-1}\left(J^T W\left(\mathbf{Y} - \boldsymbol{f}(\mathbf{b}^{(i)})\right) + U(\boldsymbol{\mu} - \mathbf{b}^{(i)})\right) \tag{2.53}$$

where $J = \nabla \boldsymbol{f}(\boldsymbol{\beta}_0)$. When iteration $i \to \infty$, vector $\mathbf{b}^{(i)}$ becomes the estimator $\hat{\boldsymbol{\beta}}$.

Depending on the knowledge of the randomness in the measurement $\mathbf{Y}$ and parameter $\boldsymbol{\beta}$, multiple cases can be considered [7, 202]:

case 1: if there is no assumption concerning the statistical properties of the measurements and the parameters, the most reliable approach is to use the ordinary least-square (OLS) estimator $\hat{\boldsymbol{\beta}}_{OLS}$, which can be computed by letting $W = I, U = 0$ and $\boldsymbol{\mu} = \mathbf{0}$ in (2.53). The resulting update equation is (assuming J has full rank and overdetermined)

$$\mathbf{b}^{(i+1)} = \mathbf{b}^{(i)} + (J^T J)^{-1}(J^T(\mathbf{Y} - \boldsymbol{f}(\mathbf{b}^{(i)}))) \tag{2.54}$$

which is identical to the update equation previously derived in the deterministic case (equation (2.22)).

The computation of $\hat{\boldsymbol{\beta}}_{OLS}$ dose not involve any statistical properties of the measurement and parameter. However, the appropriateness of this estimator does depend on these properties. If the measurement noise can be characterized as: 1) additive noise, i.e. $\mathbf{Y} = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\epsilon}$, 2) the noise has zero mean, i.e. $E(\boldsymbol{\epsilon}) = \mathbf{0}$ where $E(\boldsymbol{\epsilon})$ is the expectation of $\boldsymbol{\epsilon}$, 3) parameter $\boldsymbol{\beta}$ is nonstochastic, and 4) there is no *a priori* information about $\boldsymbol{\beta}$, then $\hat{\boldsymbol{\beta}}_{OLS}$ is an unbiased estimator to the true parameter $\boldsymbol{\beta}$, i.e. $E(\hat{\boldsymbol{\beta}}_{OLS}) = \boldsymbol{\beta}$. If the following conditions are also met: 5) the noise is independent, i.e. the covariance matrix $\mathrm{cov}(\boldsymbol{\beta})$ is a diagonal matrix and 6) the noise $\boldsymbol{\epsilon}$ has constant variance, i.e. $V(Y_i|\boldsymbol{\beta}) = \sigma^2 I$, then $\hat{\boldsymbol{\beta}}_{OLS}$ becomes a minimum variance estimator (MVE) whose variance approaches the Cramer-Rao bound based on the Gauss-Markov Theorem [7, 3, 202]. Otherwise, the OLS estimator to the parameter may not be the optimal choice.

case 2: if it is known that the measurement noise satisfies assumptions 1-5 in case 1 plus 7) the covariance matrix of the noise is known as $\Omega = \mathrm{cov}(\boldsymbol{\epsilon})$, then by letting $W = \Omega^{-1}$, $U = 0$ and $\boldsymbol{\mu} = \mathbf{0}$, the iterative update scheme in (2.53) leads to the weighted

least square (WLS) estimator, $\hat{\boldsymbol{\beta}}_{WLS}$. Under these assumptions, $\hat{\boldsymbol{\beta}}_{WLS}$ is an MVE to the true parameter.

case 3: in addition to all assumptions in case 2, if 8) the noise is characterized by a normal distribution, then the WLS estimator in case 2 becomes identical to the maximum likelihood (ML) estimator $\hat{\boldsymbol{\beta}}_{ML}$.

case 4: if the noise satisfies all conditions in case 3 except that *a priori* information is available, and $\boldsymbol{\beta}$ is a random variable with mean $\boldsymbol{\mu}_\beta$ and covariance Ω_β, then by letting $W = \Omega^{-1}$, $\boldsymbol{\mu} = \boldsymbol{\mu}_\beta$, $U = \Omega_\beta^{-1}$, equation (2.53) produces the maximum *a posteriori* (MAP) estimator to the parameter $\boldsymbol{\beta}$.

Note that among all four estimators, only the OLS estimator does not require any statistical assumption in order to compute the solution. This can be both good and bad. The benefits primarily relate to the simplicity of the computations. The drawbacks include the fact that the accuracy of the estimation depends completely on how many of the assumptions are met in the real applications. The more general approach for deriving the update equation given the statistical characteristics of the measurements and parameters is to construct the likelihood function or *a posteriori* probability function based on the analytical form of the forward model. For this approach, the measurement noise must be characterized including its mean, variance and distribution. In Appendix B, we illustrate an example of characterizing the statistical properties of the measurement noise from the microwave imaging system developed at Dartmouth College to justify the reconstruction algorithms selected in the remainder of the thesis.

2.6 Summary

In the first two chapters, the conceptual introduction for the key ingredients of building microwave imaging systems and algorithms were compiled and a road map for developing efficient reconstruction algorithms was outlined. The physical nature of microwave imaging was investigated by discussing the general framework of medical imaging modalities including a literature review profiling the evolution of this technology. In the second chapter, we discussed various mathematical tools to rigorously formulate the image reconstruction problems. These tools include nonlinear optimization, parameter estimation theory, the inverse problem and the regularization for ill-posedness.

From these discussions, a couple of preliminary conclusions can be drawn:

1. microwave imaging is a wave-based active dielectric property imaging method. Note that in this thesis, microwave imaging is synonymously used with microwave near-field tomographic imaging,

2. the exploration of nonlinear physical phenomena is a great challenge for modern imaging techniques and provides new opportunities for high contrast functional imaging,

3. the dielectric properties and measured field qualities are related by a nonlinear model - Maxwell's equations,

4. the iterative image reconstruction process requires the evaluation of the scattered field for the known structure, i.e. the forward problem, as well as the evaluation of the parameter update, i.e. the inverse problem, during each iteration.

Part II

Image reconstruction algorithms

Chapter 3

Dual-mesh based 2D reconstruction algorithms

The primary focus of this part of the thesis are the image reconstruction algorithms themselves and the details of their implementations. The algorithms here are specifically those developed for the near-field tomographic imaging systems at Dartmouth College. Compared to other existing microwave imaging algorithms, these methods feature utilizing nonlinear methods in both forward field modellings and parameter reconstructions.

3.1 Introduction

In this chapter, we discuss the general framework of the nonlinear image reconstruction and associated 2D forward solution algorithms. The concept of the dual-mesh is introduced as an organizing theme for the algorithms sequentially presented in this part of the thesis. The previous efforts and new developments in 2D image reconstruction will be discussed including the 2D scalar forward field reconstruction algorithm utilizing

the finite element method coupled with boundary element method and the 2D forward solution using the FDTD method. A number of important issues are investigated including the coupling between FE and BE methods, the absorbing boundary condition, excitation implementations, stability condition and numerical dispersions in the FDTD method as well as a computational efficiency comparison between 2D FE-based and FDTD methods.

3.2 Regularized Gauss-Newton iterative reconstruction

As we discussed in the first part of the thesis, especially in Chapter 2, the nonlinear and ill-posed nature of our problem are non-negligible issues and must be considered in the development of these algorithms. The high-contrast of the object to the background for the breast imaging problem, one of the supposed advantages of microwave imaging, renders the linearized algorithms, as discussed in Section 2.1 and 2.2, less favorable due to their "weak scatterer" assumptions. Iterative algorithms are more appropriate in this scenario among which the Gauss-Newton based iterative update scheme (equation (2.53) in Section 2.5) is quite promising due to its generality. In this section, we will discuss each term in the update equation by considering the actual application and also investigate the validity of the associated statistical assumptions.

3.2.1 Forward equations

The measurement data from our data acquisition system are the electric fields. The parameters to be estimated are the dielectric properties, i.e. permittivity, ϵ, and conductivity, σ. The field values and properties are related by the curl relationships in

Maxwell's equations [100, 80, 87], i.e.

$$\begin{aligned} \nabla \times \vec{E}(\vec{r},t) &= -\frac{\partial \vec{B}(\vec{r},t)}{\partial t} - \vec{M}(\vec{r},t) \\ \nabla \times \vec{H}(\vec{r},t) &= \frac{\partial \vec{D}(\vec{r},t)}{\partial t} + \vec{J}_i(\vec{r},t) + \vec{J}(\vec{r},t) \end{aligned} \tag{3.1}$$

where $\vec{E}, \vec{H}, \vec{D}, \vec{B}, \vec{J}_i, \vec{J}$ and $\vec{M}$ are the electric field (V/m), magnetic field (A/m), electric flux (C/m^2), magnetic flux (Wb/m^2), induced current density (A/m^2), source current density (A/m^2) and magnetic current density (Wb/(sm^2)), respectively. The constitutive relationships are given by

$$\begin{aligned} \vec{D}(\vec{r},t) &= \bar{\bar{\epsilon}}(\vec{r},t)\vec{E}(\vec{r},t) \\ \vec{B}(\vec{r},t) &= \bar{\bar{\mu}}(\vec{r},t)\vec{H}(\vec{r},t) \\ \vec{J}_i(\vec{r},t) &= \bar{\bar{\sigma}}(\vec{r},t)\vec{E}(\vec{r},t) \\ \vec{M}(\vec{r},t) &= \bar{\bar{\sigma}}^*(\vec{r},t)\vec{H}(\vec{r},t) \end{aligned} \tag{3.2}$$

where $\bar{\bar{\epsilon}}(\vec{r},t)$, $\bar{\bar{\mu}}(\vec{r},t)$, $\bar{\bar{\sigma}}(\vec{r},t)$ and $\bar{\bar{\sigma}}^*(\vec{r},t)$ are the permittivity, permeability, conductivity and magnetic conductivity tensors, respectively. The magnetic current density is a fictitious term introduced for mathematical symmetry in the equations. We will neglect this term in all the subsequent derivations except in Section 3.6 and 5.1.3 where the field in an artificial medium is analyzed. Because our targets are biological tissues which are generally 1) isotropic, i.e. the dielectric tensors become scalars, 2) nonmagnetic, i.e. $\mu(\vec{r},t) = \mu_0$ where μ_0 is the free space permeability, 3) electrically lossy, i.e. $\sigma(\vec{r},t) \neq 0$, and 4) stationary, i.e. the dielectric properties are not functions of time, equation (3.1) can be subsequently expressed as

$$\begin{aligned} \nabla \times \vec{E}(\vec{r},t) &= -\mu_0 \frac{\partial \vec{H}(\vec{r},t)}{\partial t} \\ \nabla \times \vec{H}(\vec{r},t) &= \epsilon(\vec{r}) \frac{\partial \vec{E}(\vec{r},t)}{\partial t} + \sigma(\vec{r})\vec{E}(\vec{r},t) + \vec{J}(\vec{r},t) \end{aligned} \tag{3.3}$$

Assuming the field is time-harmonic, i.e. waves at a single frequency, the complex notation can be introduced to simplify the mathematical derivations. The electric field

component E_ξ ($\xi = x, y, z$) can be rewritten as

$$\begin{aligned} E_\xi(\vec{r}, t) &= |E_{\xi 0}(\vec{r})| \cos(\omega t + \phi_{\xi 0}) \\ &= \Re\mathrm{e}(|E_{\xi 0}(\vec{r})| e^{j\phi_{\xi 0}} e^{j\omega t}) \end{aligned} \tag{3.4}$$

where the complex number $|E_{\xi 0}(\vec{r})|e^{j\phi_{\xi 0}}$ is denoted as $E_\xi(\vec{r})$ which is called the phasor. Similar conversions can be carried for all field vectors in (3.3). Inserting the phasor representations back to 3.3 for all field vectors, and cancelling the time dependence $e^{j\omega t}$ on both sides of the equation, we obtain a relationship of the complex-valued phasors as

$$\begin{aligned} \nabla \times \vec{E}(\vec{r}) &= -j\omega\mu_0 \vec{H}(\vec{r}) \\ \nabla \times \vec{H}(\vec{r}) &= j\omega \left(\epsilon(\vec{r}) - j\tfrac{\sigma(\vec{r})}{\omega}\right) \vec{E}(\vec{r}) + \vec{J}(\vec{r}) \end{aligned} \tag{3.5}$$

Equation (3.5) is referred as the frequency domain representation of (3.3). Solving for $\vec{H}(\vec{r})$ from the first equation in (3.5) and substituting into the second equation, we get

$$\nabla \times \nabla \times \vec{E}(\vec{r}) - \omega^2\mu_0 \left(\epsilon(\vec{r}) - j\frac{\sigma(\vec{r})}{\omega}\right) \vec{E}(\vec{r}) = -j\omega\mu_0 \vec{J}(\vec{r}) \tag{3.6}$$

Applying the vector identity

$$\nabla \times \nabla \times \vec{U} = \nabla\nabla \cdot \vec{U} - \nabla^2 \vec{U} \tag{3.7}$$

and defining the squared complex wave number as

$$k^2(\vec{r}) = \omega^2\mu_0 \left(\epsilon(\vec{r}) - j\frac{\sigma(\vec{r})}{\omega}\right) \tag{3.8}$$

(3.6) can be rewritten as

$$\nabla\nabla \cdot \vec{E}(\vec{r}) - \nabla^2 \vec{E}(\vec{r}) - k^2(\vec{r})\vec{E}(\vec{r}) = -j\omega\mu_0 \vec{J}(\vec{r}) \tag{3.9}$$

From Gauss's law

$$\nabla \cdot \vec{D}(\vec{r}) = 0 \tag{3.10}$$

which can be expanded to give

$$\nabla \cdot \vec{E}(\vec{r}) = -\frac{1}{\epsilon(\vec{r})}\vec{E}(\vec{r}) \cdot \nabla\epsilon(\vec{r}) \tag{3.11}$$

together with the charge conservation law [100]:

$$\nabla \times \vec{J}(\vec{r}) = -\frac{\partial\rho}{\partial t} = 0 \tag{3.12}$$

are used to produce

$$\nabla^2\vec{E}(\vec{r}) + k^2(\vec{r})\vec{E}(\vec{r}) + \nabla\left(\frac{\vec{E}(\vec{r}) \cdot \nabla k^2(\vec{r})}{k^2(\vec{r})}\right) = j\omega\mu_0\vec{J}(\vec{r}) \tag{3.13}$$

Equation (3.13) or (3.6) is the vector-form wave equation [100, 31] in the 3D space which defines the relationship between the frequency-domain electric field and the dielectrics (embodied in $k^2(\vec{r})$). In another word, equation (3.13) and (3.6) are the forward model in microwave imaging. $k^2(\vec{r})$ stores the unknown permittivity and conductivity distributions in its real and imaginary parts, respectively. Once $k^2(\vec{r})$ has been reconstructed, the permittivity and conductivity distributions can be obtained simultaneously. It is important to recognize that the relationship between the measurements in $\vec{E}(\vec{r})$ and unknowns, $k^2(\vec{r})$, is a nonlinear relationship.

In practice, the measurements can not be a continuous function, and neither can the solution of the dielectric property distribution. To evaluate the forward field from (3.13) and reconstruct k^2, both $\vec{E}(\vec{r})$ and $k^2(\vec{r})$ must be discretized. This can be achieved by implementing the computational algorithm discussed in Section 3.2.2. These numerical

methods typically yield a solution in a matrix equation representation in the form of

$$A(\mathbf{k}^2)\mathbf{E} = \mathbf{b} \tag{3.14}$$

where $\mathbf{E} = \{\vec{E}(\vec{r}_1), \vec{E}(\vec{r}_2), \cdots, \vec{E}(\vec{r}_N)\}$ is the discretized field vector over spatial points $\{\vec{r}_i\}_{i=1}^{N}$, $A(\mathbf{k}^2)$ is an $N \times N$ matrix and $\mathbf{k}^2 = \{k^2(\vec{p}_1), k^2(\vec{p}_2), \cdots, k^2(\vec{p}_P)\}$ denotes the complex dielectric parameters over point set $\{\vec{p}_i\}_{i=1}^{P}$. It should be mentioned that even though (3.14) is a matrix equation, the relationship between vectors $\mathbf{E}$ and $\mathbf{k}^2$ is still nonlinear since $\mathbf{E}$ can be conceptually expressed as

$$\mathbf{E} = A^{-1}(\mathbf{k}^2)\mathbf{b} \tag{3.15}$$

3.2.2 Computational methods for evaluating forward problems

Computational electromagnetics (CEM) has evolved into a fast gowning discipline which incorporates the emerging computational methodologies into the study of electromagnetism. Many of the methods in CEM are particularly designed for modelling the EM phenomena due to the distinctive structure of the governing mathematical model. With large computer clusters or vector machines, modern CEM methods can solve problems with huge number of unknowns with good accuracy and within acceptable times. A summary of the most popular computational methods is illustrated in a tree structure in Appendix F.

For evaluating the EM fields requires solving the full Maxwell's equations. However, when the EM wave wavelength is much smaller than the dimension of the structure with which it interacts, high-frequency approximations can be applied to simplify the computation. As a result, the classical ray-optics, beam optics or the particle model

of photons become the more efficient solutions to the problem. A number of methods were developed for high-frequency EM field approximations. These methods include geometric optics (GO), physical optics (PO) [179], geometric theory of diffraction (GTD) [98, 78], physical theory of diffraction (PTD) [195], uniform geometric theory of diffraction (UTD) [198], shooting-and-bouncing ray (SBR) [113], complex ray (CR) along with other techniques. These high-frequency methods were applied primarily for solving the electrically-large object scattering problems.

For cases where the previous small wavelength assumption is not valid, one can not treat the EM wave as an optical ray. Maxwell's equations must be solved directly. With discretization and linearization, these relationships can be converted from continuous representations into matrix equations. Based on differences in the continuous model, we can divide these methods into two major categories: the integration and differentiation based methods.

Gauss's theorem, Stoke's theorem and, more importantly, the Green's identities relate the differential operators with appropriate boundaries integrations. Many useful relationships can be derived from the combinations of these identities. For example, the Helen formula, the Stratton-Chu formula, the electric field integral equation (EFIE), magnetic field integral equation (MFIE) and complex field integral equation (CFIE) are all among the most frequently used integral relationships for evaluating the field scattering problems [87].

These surface integral equations can be further discretized into matrix forms under various approaches. The variational principle is one of them [92, 185]. The implementation of this principle results in an efficient integral based computational method, the method of moment (MoM), which was first developed by Harrington [81] and is still extensively used in many applications. The fast multipole method (FMM) and multi-level

fast multipole algorithm (MLFMA) are relatively new developments in the integration based algorithms, which were first studied by Rokhlin and Greengard [170, 69] and Lu, Song and Chew [117, 183, 184] respectively. Using a binary tree hierarchy structure, these algorithms provide significant improvement in speed and efficiency in modelling large structure scattering problems. The computational complexities for these two methods for 3D scattering problems are $O(N^2)$ and $O(N \log N)$, respectively.

The benefits of using integration based methods include:

1. The reduction in the problem dimension: for surface integral equations, the unknowns are only located on the boundaries, i.e. for the 3D scattering problem, only 2D surface integrations are needed; for 2D scattering problem, only line integrations are required. Thus, the problem size is dramatically reduced;

2. no special processing is needed for unbounded radiation problems: the unbounded assumption is naturally implied in the integral relation.

However, one major drawback for these integration based methods is the dense nature of the resulting matrix equation. It imposes significant computational difficulties when solving the huge dense linear system produced by volume integral equations when modelling inhomogeneities.

The parallel branch to the integration based methods are the differentiation based methods since they are derived directly or indirectly from the partial differential equations (PDE). The finite element (FE) [185, 92, 162] and finite-difference time-domain (FDTD) methods [203, 188, 189, 190] are among the most popular. The finite element method results from the implementation of the variational principle in differential equations, similar to MoM in integration-based methods; whereas the FDTD method is based on direct discretization of the differential operators into difference operators. A

Table 3.1: Major differences between FE and FDTD method in their generic forms

method name	FE	FDTD
mathematical foundation	variational principle	difference representation of derivative
applied equations	not specific (Helmholtz equation in many EM problems)	Maxwell's equation
mesh	flexible, conformal to arbitrary boundaries	less flexible, orthogonal grids
problem type	frequency domain	time domain
relationship	implicit	explicit
solving methods	sparse matrix solver	explicit update

list of the major differences between the FE and FDTD methods is provided in Table 3.1. In this chapter and the following chapters, these two methods are the major numerical algorithms for modelling the forward fields in microwave image reconstructions.

The advantages of the differentiation-based methods lie in the fact that they are 1) simple and straightforward to formulate, 2) yield sparse matrices, and 3) easily model problems with inhomogeneities. The associated major disadvantages stem from 1) the differentiation-based methods requiring the discretization of the whole computational domain using a volumetric mesh resulting in a large unknown, 2) for unbounded radiation problems, the mesh needs to be terminated by absorbing boundary conditions (ABC) which adds to the modelling complexity.

The research on high-efficient absorbing boundary conditions is an active field of research in CEM. Traditional ABC's are mostly mathematical boundary conditions, i.e. they impose special equations at the boundary nodes which are derived from the PDE's. Comprehensive reviews of the traditional ABC's can be found in [162, 189]. A group of traditional ABC's are based on one-way wave principle by decomposing the wave operator at the boundary into waves propagating along opposite directions and eliminating the outgoing waves. This principle is simple to implement; however, it is

selective with respect to frequency and wave incidence angle. For wide band signals or near field radiation problems, unsatisfactory high reflection levels can be observed from these RBC's. Another technique is to utilize the integral-based methods to handle the unbounded radiation outside the mesh while using differential-based methods to model the inhomogeneities in the target domain. The combination of the differentiation and integration-based methods was used by Paulsen and Meaney [127] for developing the first image reconstruction algorithm at Dartmouth. This method will be discussed in details in Section 3.4.

In contrast to mathematical boundary conditions, a material boundary condition refers to an ABC that retains the basic PDE equations while utilizing special dielectric properties to attenuate the outgoing waves. One of the most attractive ABC's recently developed for terminating an FDTD grid is the perfectly matched layer (PML) ABC. [8, 9]. By using a layer of specially designed artificial materials instead of mathematical processes, the PML ABC can efficiently attenuate the outgoing wave and theoretically produce "perfect" transmission for signals at arbitrary frequency and incidence angles. Only nominal numerical reflections are observed due to the discretization of the PML material. Due to the excellent performance of the PML, this ABC has been extended to cases with lossy medium and situations utilizing cylindrical/spherical coordinate and non-conformal meshes [35]. There are also PML ABC for finite element methods [130]. In Section 4 and 5, the PML ABC in coordination with the FDTD method to compute the forward field distribution for microwave imaging.

3.2.3 Gauss-Newton method

Basic formulation

Assuming the electric field measurement, denoted as $\mathbf{E}^{\text{meas}}$, are recorded at N_r receivers located at $\{\vec{r}_i\}_{i=1}^{N_r}$, from the discussions in Section 2.5, to reconstruct the unknown parameters, a sum-of-square optimization problem is formulated as

$$\min_{\mathbf{k}^2} \; S = (\mathbf{E}^{\text{meas}} - \Lambda\mathbf{E})^T W(\mathbf{E}^{\text{meas}} - \Lambda\mathbf{E}) + (\mathbf{k}^{2*} - \mathbf{k}^2)^T U(\mathbf{k}^{2*} - \mathbf{k}^2) \tag{3.16}$$

where Λ is a sampling matrix provided by the selected basis functions in the discretization scheme. W and U are weighting matrices and $\mathbf{k}^{2*}$ is the *a priori* solution to $\mathbf{k}^2$. The corresponding Gauss-Newton update equation is given by

$$\mathbf{k}^2_{(i+1)} = \mathbf{k}^2_{(i)} + \left(J^T_{(i)} W J_{(i)} + U\right)^{-1} \left(J^T_{(i)} W \left(\mathbf{E}^{\text{meas}} - \Lambda\mathbf{E}_{(i)}\right) + U(\mathbf{k}^{2*} - \mathbf{k}^2_{(i)})\right) \tag{3.17}$$

where $J_{(i)} = \frac{d\mathbf{E}}{d\mathbf{k}}\Big|_{\mathbf{k}=\mathbf{k}_{(i)}}$ is the Jacobian matrix at the i-th iteration representing the sensitivity between the field distribution with respect to perturbations of the dielectric properties (more details about the Jacobian matrix construction can be found in Section 5.1.2 and Chapter 7).

As mentioned earlier, the selections of W, U and $\mathbf{k}^{2*}$ depend on the statistical properties of measurement noise and the parameter $\mathbf{k}$. For illustrative purposes, the OLS (ordinary least-square) estimator is used for demonstrating the basic methodologies. The statistical assumptions of the OLS estimator can be found in Section 2.5. We have characterized the actual statistical properties of the measurement data for our current imaging system and the findings are summarized in Appendix B.

The update scheme (3.17) now becomes

$$\mathbf{k}^2_{(i+1)} = \mathbf{k}^2_{(i)} + \left(J^T_{(i)} J_{(i)}\right)^{-1} \left(J^T_{(i)} \left(\mathbf{E}^{\text{meas}} - \Lambda \mathbf{E}_{(i)}\right)\right) \tag{3.18}$$

or

$$G_{(i)} \Delta \mathbf{k}^2_{(i)} = J^T_{(i)} \Delta \mathbf{E}_{(i)} \tag{3.19}$$

where $G_{(i)} = J^T_{(i)} J_{(i)}$ is the Gauss-Hessian matrix, and $\Delta \mathbf{k}^2_{(i)} = \mathbf{k}^2_{(i+1)} - \mathbf{k}^2_{(i)}$ and $\Delta \mathbf{E}_{(i)} = \mathbf{E}^{\text{meas}} - \Lambda \mathbf{E}_{(i)}$ are the parameter update and electric field misfit, respectively.

The reconstruction problem is inherently ill-posed and manifests itself as the ill-conditioning of matrix G. Regularization techniques are subsequently required for solving the linear equation (3.19). Assuming the Tikhonov regularization is chosen, the final form of the update equation is written as

$$\left(G_{(i)} + \lambda_{(i)} I\right) \Delta \mathbf{k}^2_{(i)} = J^T_{(i)} \Delta \mathbf{E}_{(i)} \tag{3.20}$$

where $\lambda_{(i)}$ is the scalar regularization parameter and I is the identity matrix. Among the various techniques for selecting the regularization parameter λ_i, the empirical method developed by Hogunin *et al.* [86] and Jaochimowicz *et al.* [93] is simple to compute and has consistently demonstrated good performance in a large number of reconstructions with our application. In this method, the value of $\lambda_{(i)}$ is computed by the following equation [56]

$$\lambda_{(i)} = \alpha \frac{tr(G_{(i)})}{P} \left(\frac{\|\Delta \mathbf{E}_{(i)}\|}{\|\Delta \mathbf{E}_{(1)}\|} \right)^2 \tag{3.21}$$

where α is a user supplied constant, P is the number of the unknowns, $tr(G_{(i)})$ is the trace of matrix $G_{(i)}$, i.e. the summation of the diagonal elements, and $\|\Delta \mathbf{E}_{(1)}\|$ is the $L2$ norm of the field misfit error at the first iteration. In most of the following reconstructions,

we will use (3.21) to determine λ at each iteration.

Modifications to the update equation

Several empirical techniques were discussed and implemented in [151, 129]. These techniques include the log-magnitude/phase form (LMPF) reconstruction algorithm, spatial filter technique and parameter pre-scaling technique. Other modified Gauss-Newton methods were also reported in the literature such as the damped Gauss-Newton method and the L-matrix regularization method [76].

(a). The log-magnitude/phase form reconstruction

Equation (3.19) is called as the "normal equation" of

$$J_{(i)}\Delta\mathbf{k}^2_{(i)} = \Delta\mathbf{E}_{(i)} \tag{3.22}$$

and the solutions of these two equations are essentially identical.

For the LMPF algorithm, (3.22) is first rewritten in a real form as

$$\begin{pmatrix} \Re\mathrm{e}(J_{(i)}) & -\Im\mathrm{m}(J_{(i)}) \\ \Im\mathrm{m}(J_{(i)}) & \Re\mathrm{e}(J_{(i)}) \end{pmatrix}\begin{pmatrix} \Re\mathrm{e}(\Delta\mathbf{k}^2_{(i)}) \\ \Im\mathrm{m}(\Delta\mathbf{k}^2_{(i)}) \end{pmatrix} = \begin{pmatrix} \Re\mathrm{e}(\Delta\mathbf{E}_{(i)}) \\ \Im\mathrm{m}(\Delta\mathbf{E}_{(i)}) \end{pmatrix} \tag{3.23}$$

Subsequently, (3.23) is transformed into

$$\begin{pmatrix} J_1 & J_2 \\ J_3 & J_4 \end{pmatrix}\begin{pmatrix} \Re\mathrm{e}(\Delta\mathbf{k}^2_{(i)}) \\ \Im\mathrm{m}(\Delta\mathbf{k}^2_{(i)}) \end{pmatrix} = \begin{pmatrix} \Delta\Gamma(\mathbf{E}_{(i)}) \\ \Delta\Phi(\mathbf{E}_{(i)}) \end{pmatrix} \tag{3.24}$$

where $\Gamma(\cdot)$ represents the log-magnitude of a complex variable and $\Phi(\cdot)$ represents its

unwrapped phase. The submatrices on the LHS of (3.24) are defined by

$$\begin{aligned} J_1 &= \frac{\Re\mathrm{e}(E_{(i)})\Re\mathrm{e}(J_{(i)})+\Im\mathrm{m}(E_{(i)})\Im\mathrm{m}(J_{(i)})}{(\Re\mathrm{e}(E_{(i)}))^2+(\Im\mathrm{m}(E_{(i)}))^2} \\ J_2 &= \frac{-\Re\mathrm{e}(E_{(i)})\Im\mathrm{m}(J_{(i)})+\Im\mathrm{m}(E_{(i)})\Re\mathrm{e}(J_{(i)})}{(\Re\mathrm{e}(E_{(i)}))^2+(\Im\mathrm{m}(E_{(i)}))^2} \\ J_3 &= \frac{\Re\mathrm{e}(E_{(i)})\Im\mathrm{m}(J_{(i)})-\Im\mathrm{m}(E_{(i)})\Re\mathrm{e}(J_{(i)})}{(\Re\mathrm{e}(E_{(i)}))^2+(\Im\mathrm{m}(E_{(i)}))^2} \\ J_4 &= \frac{\Re\mathrm{e}(E_{(i)})\Re\mathrm{e}(J_{(i)})+\Im\mathrm{m}(E_{(i)})\Im\mathrm{m}(J_{(i)})}{(\Re\mathrm{e}(E_{(i)}))^2+(\Im\mathrm{m}(E_{(i)}))^2} \end{aligned} \tag{3.25}$$

which were derived in [151]. Equation (3.24) is referred as the log-magnitude phase form of the Gauss-Newton update equation. One may notice that $J_1 = J_4$ and $J_2 = -J_3$; therefore, (3.25) can be further shorten as a complex equation, i.e.

$$J'_{(i)}\Delta\mathbf{k}^2_{(i)} = \Delta\mathbf{E}'_{(i)} \tag{3.26}$$

where $J'_{(i)} = J_1 + jJ_3$ and $\Delta\mathbf{E}'_{(i)} = \Delta\Gamma(\mathbf{E}_{(i)}) + j\Delta\Phi(\mathbf{E}_{(i)})$. This method has demonstrated improved performance in various simulation and measurement data reconstructions, especially when the object is large and the contrast is high. However, the statistical significance of this method is still under investigation.

(b). Spatial filter technique

A spatial filter technique refers to the nodal averaging process during each Gauss-Newton iteration. The averaged value of each parameter can be expressed as

$$\theta_\tau = (1-\alpha)\theta_\tau + \frac{\alpha}{N_\tau}\sum_{i\in R_\tau}\theta_i \tag{3.27}$$

where θ_τ denotes the τ-th parameter, α is a scalar quantity referred to the averaging factor, $\sum_{i\in R_\tau}$ denotes the summation over the neighboring nodes of the τ-th parameter node and N_τ is the total number of the neighboring nodes. The application of this averaging scheme results in a smoother image which is qualitatively similar to the results when choosing larger regularization parameters.

(c). The damped Gauss-Newton method

The damped Gauss-Newton method refers to the following update equation

$$\mathbf{k}^2_{(i+1)} = \mathbf{k}^2_{(i)} + s\left(J^T_{(i)}J_{(i)}\right)^{-1}\left(J^T_{(i)}\left(\mathbf{E}^{\text{meas}} - \Lambda\mathbf{E}_{(i)}\right)\right) \tag{3.28}$$

where s is an empirical constant referred as the damping coefficient. Using a smaller s value can reduce the oscillatory behavior during the Gauss-Newton iterative method.

(d). The parameter pre-scaling method

For the parameter pre-scaling method, the real form update equation prior to regularization, i.e. (3.23), is modified as [129]

$$\begin{pmatrix} q\Re\text{e}(J_{(i)}) & -\Im\text{m}(J_{(i)}) \\ q\Im\text{m}(J_{(i)}) & \Re\text{e}(J_{(i)}) \end{pmatrix}\begin{pmatrix} \frac{1}{q}\Re\text{e}(\Delta\mathbf{k}^2_{(i)}) \\ \Im\text{m}(\Delta\mathbf{k}^2_{(i)}) \end{pmatrix} = \begin{pmatrix} \Re\text{e}(\Delta\mathbf{E}_{(i)}) \\ \Im\text{m}(\Delta\mathbf{E}_{(i)}) \end{pmatrix} \tag{3.29}$$

where q is an empirical scalar term to scale the real and imaginary parts of the solution. This method is essentially a special case of the general concept of matrix weighting discussed in Section 2.5 and 6.2.3.

(e). The L matrix regularization

The L matrix method is a special case of Tikhonov regularization which penalizes the spatial derivative of the solution instead of its absolution value. The identity matrix in (3.20) is replaced by an L matrix as

$$\left(G_{(i)} + \lambda_{(i)}L\right)\Delta\mathbf{k}^2_{(i)} = J^T_{(i)}\Delta\mathbf{E}_{(i)} \tag{3.30}$$

where L is typically chosen as the discrete difference operator as demonstrated in [76]. The high-frequency oscillatory modes of the images are filtered by applying this regularization which results in improved stability with the reconstruction being less sensitive

to noise.

Summary

To summarize this subsection, we have outlined a regularized Gauss-Newton algorithm for microwave imaging. The update equation (3.20) is simple and efficient for the cases where the statistical assumptions are satisfied. Otherwise, the more complete form, i.e. (3.17), should be used instead where the extra terms and coefficients need to be determined from the statistics of the measurements.

3.2.4 Flow chart of the regularized Gauss-Newton method

Based on the analysis in the last subsection, a flow chart is drawn to illustrate the detailed computational steps for a complete reconstruction (Figure 3.1). From the flow chart, four key steps can be identified: (a) evaluation of the forward field solution, (b) construction of the Jacobian matrix, (c) determination of the regularization parameter and (d) updating the reconstruction parameters. The accuracy and efficiency of these procedures are essential for a successful reconstruction. The remainder of this chapter and the following two chapters will focus on improvements in forward modelling and Jacobian matrix construction with other minor issues being mentioned briefly.

3.3 The dual-mesh scheme

In the previous section, the detail of the discretization process of equation (3.13) were omitted for simplification and are discussed in more details here. Spatial basis functions are introduced to approximate the continuous function by a finite summation. Assuming $\{\phi_i(\vec{r})\}_{i=1}^{\infty}$ is a complete orthogonal basis function set, an arbitrary spatial function

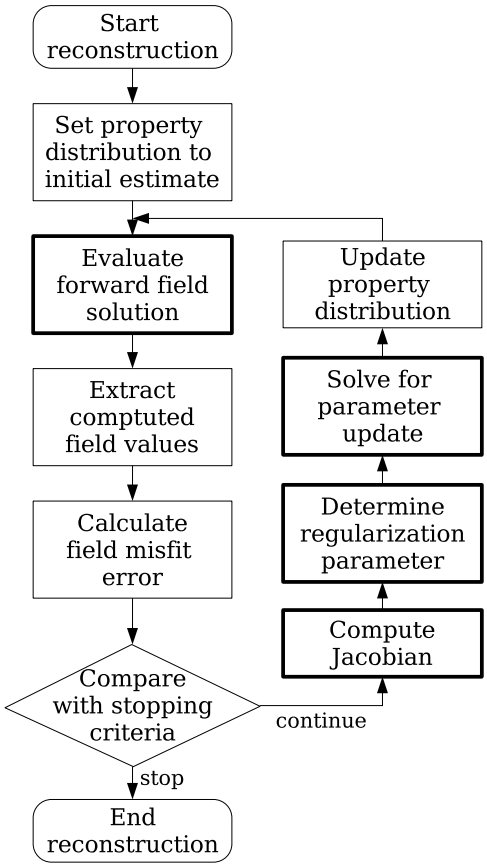


Figure 3.1: Flow chart for illustrating reconstructions utilizing the regularized Gauss-Newton method

$f(\vec{r})$ can be expanded in terms of ϕ_i as

$$f(\vec{r}) = \mathbf{f} \cdot \mathbf{\Phi} \tag{3.31}$$

where $\mathbf{f} = \{f_i\}_{i=1}^{\infty}$ and $\mathbf{\Phi} = \{\phi_i(\vec{r})\}_{i=1}^{\infty}$ are two infinite vectors containing function coefficients and basis functions, respectively. Since an infinite vector can not be manipulated, a truncation is needed to approximate the sum with a finite number of terms. Thus (3.31) becomes

$$f(\vec{r}) \approx \mathbf{f}^N \cdot \mathbf{\Phi}^N \tag{3.32}$$

where vectors $\mathbf{f}^N$ and $\mathbf{\Phi}^N$ contain only the first N terms of their pre-truncated versions.

Returning to the microwave image reconstruction problem, the discretization process must be applied to both the forward fields and property parameters. The more general approach is to utilize two separate meshes for the field and parameter distributions (i.e. a dual-mesh pair) [155, 39], but for convenience, some investigators often utilize a single mesh. There are several reasons why the dual-mesh approach is advantageous. First, the spatial domains for modelling the forward field and dielectric inhomogeneity distributions may be significantly different in size. The forward field is typically evaluated in a physically larger domain containing not only the target but all of the transmitters, receivers and surrounding structures. Using identical mesh structure for both field and parameter representation could result in an uneconomic use of memory and unwanted redundancy in the computations. Second, for most forward modelling methods as listed in Section 3.2.2, a minimum mesh density or spatial sampling rate per a given wavelength is typically required to assure accuracy of the forward solution. On the other hand, the effective density for the parameter mesh is related to the spatial variation of the actual dielectric distribution and the amount of data available

(see Chapter 7). For most microwave imaging cases, the density of the forward mesh is much higher than that of the parameter mesh. High density in reconstruction parameter mesh will cause difficulties when solving the update equation (3.19) since it is a dense matrix equation instead of the sparse matrices used in forward equation.

The implementation utilizing separate spatial basis function sets, i.e. the dual-mesh scheme, is fairly simple and can be summarized by

$$\begin{aligned} \vec{E}(\vec{r}) &\approx \vec{\mathbf{E}} \cdot \mathbf{\Phi} \\ k^2(\vec{r}) &\approx \mathbf{k^2} \cdot \mathbf{\Psi} \end{aligned} \tag{3.33}$$

where $\mathbf{\Phi} = \{\phi_i(\vec{r})\}_{i=1}^{N}$ and $\mathbf{\Psi} = \{\varphi_i(\vec{r})\}_{i=1}^{P}$ are the truncated basis set for fields and parameters, respectively, with N being the length of the discretized field vector and P being that of the parameter vector. The implementation of a spatial basis function results in a mesh which includes a set of discrete nodes, the connections between nodes and the interpolation functions in each individual element. The mesh for representing the forward field is referred as the forward mesh, while that for the dielectric properties is the parameter mesh. Vector $\vec{\mathbf{E}}$ is the field defined over the forward mesh whereas vector $\mathbf{k^2}$ is the parameter defined on the parameter mesh.

When solving the forward field equation (3.13), the electric properties must be known at the forward mesh nodes. To accommodate this, an interpolation process is performed to transfer the parameters defined on the reconstruction mesh to the forward mesh. Similarly, when constructing the Jacobian matrix, the field values need to be interpolated from the forward mesh to the parameter node locations. These interpolations can be mathematical expressed as expansions between the two basis function sets, i.e.

$$\begin{aligned} \mathbf{\Phi} &= A_f \mathbf{\Psi} \\ \mathbf{\Psi} &= A_p \mathbf{\Phi} \end{aligned} \tag{3.34}$$

where the matrix A_f represents the linear interpolation from the parameter basis vector $\mathbf{\Psi}$ to the forward basis vector $\mathbf{\Phi}$, and A_p performs the mapping in the reversed direction. The dimension of A_f is $N \times P$ whereas that of A_p is $P \times N$, and both can be pre-computed and stored for a given dual-mesh pair. Combining (3.34) and (3.33), the interpolated field over the reconstruction mesh $\mathbf{E}_p$ and the interpolated parameter $\mathbf{k}^2_f$ over the forward mesh can be expressed as

$$\begin{aligned} \vec{\mathbf{E}}_p &= A_f^T \vec{\mathbf{E}} \\ \mathbf{k^2}_f &= A_p^T \mathbf{k^2} \end{aligned} \tag{3.35}$$

The forward and parameter basis functions and the bilateral mapping relationship are illustrated in Figure 3.2.

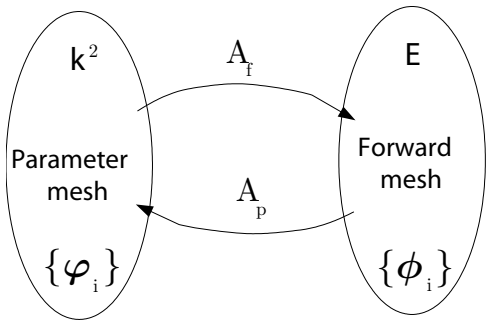


Figure 3.2: Dual-mesh mapping between the forward and parameter meshes

The implementation of the dual-mesh representation allows for the forward solution and parameter distribution to be mapped accurately between meshes and effectively decouples the forward and inverse phases of the reconstruction process. Different forward solution or reconstruction algorithms can be easily substituted into the reconstruction process with only nominal perturbations to the overall reconstruction. Utilizing the dual-mesh scheme, we have studied a number of forward methods with incrementally increased model complexities. The forward solution methods and the corresponding reconstruction algorithms used are summarized in Table 3.2. Among them, the 2D

Table 3.2: Dual-mesh based algorithms

forward mesh	2D		3D	
field representation	scalar	vector(TM)	scalar	vector
2D reconstruction mesh	2Ds/2D	$2Ds^{FDTD}/2D$	3Ds/2D	-
3D reconstruction mesh	-	-	3Ds/3D	3Dv/3D

methods are discussed in the following sections of this chapter, whereas the 3D methods will be discussed in Chapter 4 and 5.

3.4 2D scalar forward field coupled with 2D parameter reconstruction

Two-dimensional scalar forward field coupled with 2D parameter reconstruction, referred as 2Ds/2D method, is the first algorithm developed at Dartmouth College for microwave imaging. This algorithm was introduced by Paulsen *et al.* in the early 1990's. In the reconstruction, the forward field distribution is formulated as a 2D problem under the following assumptions:

1. the scattering dielectric profile is a 2D distribution, i.e. no variations along z-direction, or $\frac{dk^2(\vec{r})}{dz} = 0$;

2. the source is an infinitely long line source parallel to z-axis;

3. and consequently the propagating wave is assumed to be a transverse magnetic (TM) wave where the $\vec{E}$ vector is parallel to z-axis, i.e. $E_x(\vec{r}) = E_y(\vec{r}) = 0$.

The third term in (3.13) is discarded, i.e.

$$\vec{E}(\vec{r}) \cdot \nabla k^2(\vec{r}) = E_x(\vec{r})\frac{dk^2(\vec{r})}{dx} + E_y(\vec{r})\frac{dk^2(\vec{r})}{dy} + E_z(\vec{r})\frac{dk^2(\vec{r})}{dz} = 0 \tag{3.36}$$

Consequently, the x, y and z components in the forward equation (3.13) can be decoupled, and only the E_z component is nonzero. The resultant equation is written as

$$\nabla^2 E_z(\vec{r}) + k^2(\vec{r}) E_z(\vec{r}) = j\omega\mu_0 J_z(\vec{r}) \tag{3.37}$$

Equation 3.37 is the frequency domain scalar forward model (Helmholtz equation) which must be solved in the 2D scalar reconstructions.

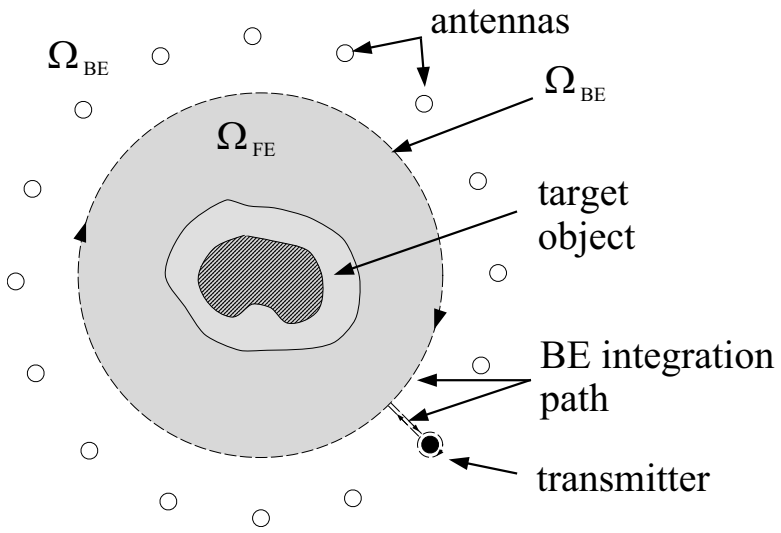


Figure 3.3: The geometric configuration for forward field modelling utilizing FE and BE methods.

Considering the experimental settings for microwave tomographic imaging (as shown in Figure 3.3), the target object is generally located within an imaging zone and surrounded by a circular antenna array. The homogeneous background medium fills the remaining space. Within the imaging zone, the field variations due to the target inhomogeneity can be conveniently modelled using the finite element (FE) method; whereas in the exterior region, the boundary element (BE) method is ideally suited to account for the unbounded radiation. As a result, a hybrid method combining the FE and BE methods was devised to model the forward field distribution in the 2D scalar problem.

The FE region as well as the integral path for BE problem are shown in Figure 3.3. Note that the imaging zone Ω_{FE} and the complementary space Ω_{BE} share an identical boundary $\partial\Omega_{BE}$ (here we assume Ω_{BE}, Ω_{FE} and $\partial\Omega_{BE}$ are the discretized geometries whose boundaries are comprised of straight line segments). We will discuss the field modelling with Ω_{FE} and Ω_{BE} separately in the following subsections.

3.4.1 Finite element region

Within the FE zone, the Helmholtz equation (3.37) is linearized into a matrix relationship by the variational principle. The procedure of applying the variational principle consists of two steps. First, the continuous form scalar functions, $E_z(\vec{r})$ and $k^2(\vec{r})$, are expanded by a truncated spatial basis function set as described in the previous section. The discretized field and parameter distributions are represented by (3.33). Substituting the expanded expressions into (3.37), we get

$$\sum_{i=1}^{N}\left(E_{zi}\nabla^2\phi_i(\vec{r})\right)+\sum_{l=1}^{P}\sum_{i=1}^{N}\left(k_l^2\varphi_l(\vec{r})E_{zi}\phi_i(\vec{r})\right)=j\omega\mu_0 J_z(\vec{r}) \tag{3.38}$$

where N is the number of the forward nodes and P is that of the parameter nodes. Second, integrating (3.38) with another set of orthogonal functions, referred as the weighting functions, $\{w_\ell(\vec{r})\}_{\ell=1}^{N}$, (also defined on the forward mesh) results in

$$\sum_{i=1}^{N}\left(E_{zi}\left\langle\nabla^2\phi_i(\vec{r}),w_\ell(\vec{r})\right\rangle\right)+\sum_{l=1}^{P}\sum_{i=1}^{N}\left(k_l^2E_{zi}\left\langle\varphi_l(\vec{r})\phi_i(\vec{r}),w_\ell(\vec{r})\right\rangle\right)=j\omega\mu_0\left\langle J_z(\vec{r}),w_\ell(\vec{r})\right\rangle \tag{3.39}$$

where $\langle\cdot,\cdot\rangle$ is the notation for inner product between two functions defined by

$$\langle f(\vec{r}),g(\vec{r})\rangle=\int_\Omega f(\vec{r})g(\vec{r})d\vec{r} \tag{3.40}$$

with Ω being the domain where real functions f and g are defined. There are numerous choices for the weighting function, $w(\vec{r})$. One of the most popular is the Galerkin method which uses the same basis function ϕ used to discretize the forward solution, i.e. $w_i(\vec{r}) = \phi_i(\vec{r})$. Applying the Galerkin method in combination with the Green's first identity:

$$\langle \nabla^2 \phi_i(\vec{r}), \phi_\ell(\vec{r}) \rangle = -\langle \nabla \phi_i(\vec{r}), \nabla \phi_\ell(\vec{r}) \rangle + \int_{\partial\Omega} \phi_\ell(\vec{r}) \nabla \phi_i(\vec{r}) \cdot d\vec{r} \tag{3.41}$$

equation (3.39) becomes

$$\begin{array}{rcl} \sum_{i=1}^{N} \left(E_{zi} \langle \nabla \phi_i(\vec{r}), \nabla \phi_\ell(\vec{r}) \rangle \right) & + & \sum_{i=1}^{N} \left(E_{zi} \int_{\partial\Omega_{BE}} \phi_\ell(\vec{r}) \nabla \phi_i(\vec{r}) \cdot d\vec{r} \right) \\ & - & \sum_{l=1}^{P} \sum_{i=1}^{N} \left(k_l^2 E_{zi} \langle \varphi_l(\vec{r}) \phi_i(\vec{r}), \phi_\ell(\vec{r}) \rangle \right) = -j\omega\mu_0 \langle J_z(\vec{r}), \phi_\ell(\vec{r}) \rangle \end{array} \tag{3.42}$$

For all $\phi_\ell(\vec{r})$, $\ell = 1, \cdots, N$, (3.39) comprises a simultaneous system of equations with N equations and N unknowns. If k_i^2 is known (in the configuration in Figure 3.3, there is no source located inside Ω_{FE}, thus $J_z(\vec{r}) = 0$), this equation system becomes a linear equation for E_{zi} similar to (3.14):

$$A(\mathbf{k}^2)\mathbf{E}_z = B\boldsymbol{\nabla}\mathbf{E} \tag{3.43}$$

where $A(\mathbf{k}^2)$ is the forward FE matrix containing parameter $\mathbf{k}^2$, the RHS matrix B and vector $\boldsymbol{\nabla}\mathbf{E}$ represent the discretized integration term in (3.42) evaluated along the mesh surface.

Equation (3.43) by itself is insufficient to uniquely solve for the field without applying boundary conditions. For convenience, we rearrange the matrix equation (3.43) to group the field values at interior nodes and boundary nodes separately as

$$\begin{pmatrix} A_{II} & A_{IB} \\ A_{BI} & A_{BB} \end{pmatrix} \begin{pmatrix} \mathbf{E}_I \\ \mathbf{E}_B \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & B_B \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\nabla}\mathbf{E}_B \end{pmatrix} \tag{3.44}$$

where $\mathbf{E}_I$ is the vector containing E_z fields at the interior nodes, $\mathbf{E}_B$ is the corresponding vector for the boundary nodes. A_{II}, A_{IB}, A_{BI} and A_{BB} are the associated partitions of the A matrix. The surface integration on the RHS involves only field gradients, denoted by $\boldsymbol{\nabla}\mathbf{E}_B$, at mesh boundaries. In order to solve for the E_z on the FE mesh, vector $\boldsymbol{\nabla}\mathbf{E}_B$ must be obtained. To accomplish this without the use of approximate boundary conditions, we apply the boundary element method for the surrounding region.

3.4.2 Boundary element region

There are several important distinctions between the BE and FE regions, Ω_{BE} and Ω_{FE}, respectively. First, Ω_{BE} consists of a homogeneous background medium; therefore, $k^2(\vec{r})$ can be replaced by a constant k_{bk}^2. Second, the source is located inside Ω_{BE}. We assume that the source is a point source with normalizing amplitude located at $\vec{r}_s$, i.e.

$$J_z(\vec{r}) = \frac{-1}{j\omega\mu_0}\delta(\vec{r} - \vec{r}_s) \tag{3.45}$$

where $\delta(\cdot)$ is a Dirac delta function. Inserting (3.45) into (3.37), the differential equation for the BE region looks like

$$\nabla^2 E_z(\vec{r}) \; + \; k_{bk}^2 E_z(\vec{r}) = -\delta(\vec{r} - \vec{r}_s) \tag{3.46}$$

The solution to (3.46) is the 2D Green's function:

$$g(\vec{r}) = \frac{j}{4}H_0^{(1)}(k_{bk}|\vec{r} - \vec{r}_s|) \tag{3.47}$$

Applying Green's second identity

$$\int_{\Omega} u(\vec{r})\nabla^2 v(\vec{r}) - v(\vec{r})\nabla^2 u(\vec{r})d(\vec{r}) = \int_{\partial\Omega} u(\vec{r})\nabla v(\vec{r}) - v(\vec{r})\nabla u(\vec{r})d(\vec{r}) \tag{3.48}$$

and letting $u(\vec{r}) = E_z(\vec{r})$, $v(\vec{r}) = g(\vec{r})$ and $\Omega = \Omega_{BE}$, we obtain an expression for the electric field at any given point $\vec{r} \in \Omega_{BE}$ in terms of a boundary integration [127]:

$$\alpha E_z(\vec{r}) = \int_{\partial\Omega_{BE}} E_z(\vec{r})\nabla g(\vec{r}) - g(\vec{r})\nabla E_z(\vec{r})d(\vec{r}) \tag{3.49}$$

where α is a scalar constant defined by

$$\alpha = \begin{cases} 1/2 & \vec{r} \in \partial\Omega_{BE} \\ 1 & \vec{r} \notin \partial\Omega_{BE} \end{cases} \tag{3.50}$$

In order to couple this with the FE representation, the boundary $\partial\Omega_{BE}$ is discretized in exactly the same manner as the boundary of the finite element mesh, i.e. the nodes and the basis function are precisely matched along the boundary. Sequentially moving $\vec{r}$ over each of a boundary nodes, the version of equation (3.49) provides a matrix equation relating the field value and the field gradient at the boundary nodes. The BE equation is then given by

$$D\boldsymbol{\nabla}\mathbf{E}_B = C\mathbf{E}_B \tag{3.51}$$

Assuming the boundary node number is N_B and the interior node number is N_I, equation (3.44) provides $N_I + N_B$ linear equations while (3.51) provides another N_B equations for a total of $N_I + 2N_B$ constraints. In association with this, there are $N_I + 2N_B$ unknowns related to vectors $\mathbf{E}_I$, $\mathbf{E}_B$ and $\boldsymbol{\nabla}\mathbf{E}_B$. Therefore, we expect to be able to compute a unique solution to the coupled equation set of (3.44) and (3.51). Note that the FE equation has a sparse LHS matrix while in BE problem, matrices C and D are

both dense and well-posed. Because the number of boundary nodes is typically much smaller than that for the interior nodes, the inversion of D can be pre-computed (since we assume that the surrounding medium properties do not change) to solve for vector $\nabla \mathbf{E}_B$ as

$$\nabla \mathbf{E}_B = D^{-1} C \mathbf{E}_B \tag{3.52}$$

Substituting (3.52) into the FE matrix equation (3.13) produces an integrated system of equations,

$$\begin{pmatrix} A_{II} & A_{IB} \\ A_{BI} & A_{BB} - B_B D^{-1} C \end{pmatrix} \begin{pmatrix} \mathbf{E}_I \\ \mathbf{E}_B \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{3.53}$$

from which the forward field vectors $\mathbf{E}_I$ and $\mathbf{E}_B$ associated with a single transmitter can be computed. For the reconstruction problem, this equation needs to be solved for all transmitters at each iteration.

3.5 Building the Jacobian matrix

In the iterative scheme as we outlined above, the Jacobian matrix must be constructed based on the forward field distribution computed by the matrix equations in the last section. Paulsen *et al.* used a method referred as the sensitivity equation [112, 127]. In order to construct the Jacobian matrix, the derivative of equation (3.14) is computed with respect to the τ-th parameter, k_τ^2,

$$A(\mathbf{k}^2)\left(\frac{d\mathbf{E}}{dk_\tau^2}\right) = -\left(\frac{dA(\mathbf{k}^2)}{dk_\tau^2}\right)\mathbf{E} \tag{3.54}$$

Note that $\mathbf{b}$ is not a function of k_τ^2. Because the forward solution is also computed at each iteration, both E and $A(\mathbf{k}^2)$ are already known. Therefore, only $\left(\frac{dA(\mathbf{k}^2)}{dk_\tau^2}\right)$ needs to be constructed to compute $\left(\frac{d\mathbf{E}}{dk_\tau^2}\right)$ which constitutes the terms of the Jacobian matrix, J.

From (3.42), the actual forms of $\left(\frac{dA(\mathbf{k}^2)}{dk_\tau^2}\right)$ is

$$\frac{da_{i,\ell}}{dk_\tau^2} = \langle \phi_i(\vec{r})\phi_\ell(\vec{r})\varphi_\tau(\vec{r})\rangle \tag{3.55}$$

where $a_{i,\ell}$ is an element of the A matrix, and the RHS represents the integration of the product of three basis functions over the space where they all exist.

Given the forward field vector $\mathbf{E}$, one needs to solve $\left(\frac{d\mathbf{E}}{dk_\tau^2}\right)$ from matrix equation (3.54) for all parameter nodes τ and all sources to construct the full Jacobian matrix. If there are P nodes in the reconstruction mesh and N_s sources in the problem, the total number of solutions needed to build the Jacobian matrix from (3.54) is $N_s \times P$. Each is essentially a back-substitution for the factored A matrix. Including the forward field solution, the total number of matrix back-substitutions is $N_s(P + 1)$ for each iteration. Experiments have shown that with the 2D scalar reconstruction algorithm above, the evaluation of a single iteration requires more than an hour on an SGI workstation with a typical problem size ($P = 144$ and $N_s = 16$), for which the time needed for building Jacobian matrix consumed more than 90% of the total computation time. While this approach works, this example demonstrates the unsatisfactory efficiency of the sensitivity equation method. More efficient algorithms for building the Jacobian matrix will be discussed in Chapters 4 and 5.

3.6 2D FDTD forward field solution coupled with 2D parameter reconstruction

Several difficulties were observed during the application of the 2D scalar reconstruction technique introduced above. The need for constructing and solving the boundary

element equation (3.49) complicates both the programming and optimization of the algorithm. The interaction of the BE and FE equations also significantly increases the bandwidth of the final matrix equation, subsequently increasing the computation time. More importantly, the sensitivity equation method for constructing the Jacobian matrix severely limits the overall computational efficiency. In terms of improving the latter point, we have developed the adjoint method [30] and its fast approximation which will be discussed in detail in Chapters 4 and 5. In this section, we will also implemented another forward solution technique, the FDTD method, to investigate the possibility of improving the efficiency of the field modelling (Note that in the 2D FDTD method, the TM vector field is essentially equivalent to the scalar model derived for the FE/BE hybrid method in the frequency domain. To distinguish these two algorithms, we denotes the former as 2Ds(FDTD) while the latter as 2Ds).

The FDTD method is attractive for the following reasons:

1. the modelling technique along with the implementation of the absorbing boundary conditions are conceptually straightforward and intuitively simple to program,

2. in general, the performance of the perfectly matched layer (PML) absorbing boundary condition (ABC) is significantly better than other existing approximate boundary conditions,

3. the marching-on-time (MOT) feature of the FDTD method provides various possibilities for accelerating the computation,

4. in both 2D and 3D, the FDTD method possesses computational advantages in total floating-point operations compared with the previous 2D scalar technique (Illustrated in Section 3.6.1).

In the remainder of this section, the basic formulations of the 2D FDTD method and the implementation of the generalized PML (GPML) ABC will be discussed. Additionally, an analysis of the computational efficiency as well as miscellaneous implementation details such as applying source excitations, amplitude/phase detection schemes and analysis of the stability and numerical dispersion will be investigated.

3.6.1 2D FDTD method

The FDTD method is derived directly from the discretized form of Maxwell's curl equations (3.1). Due to matching in grid topologies with the curl relationships of the electromagnetic field, the FDTD method exhibits significant superiority in modelling the EM wave phenomena.

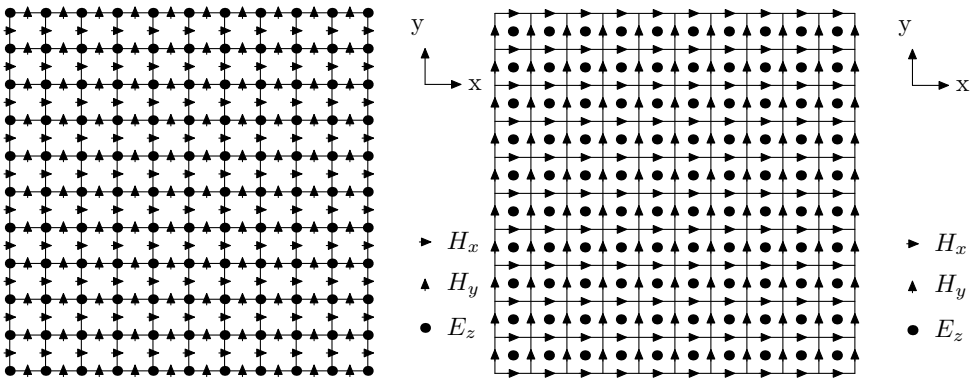
FDTD update scheme

For the 2D FDTD method, we make the same assumptions as those used for the 2D scalar method (Section 3.4), i.e. a 2D dielectric property distribution along with a TM wave with a z-oriented E field. In this case, the curl equations (3.1) can be expanded as

$$\begin{aligned} \mu(x,y)\frac{\partial H_x(x,y)}{\partial t} &= -\frac{\partial E_z(x,y)}{\partial y} \\ \mu(x,y)\frac{\partial H_y(x,y)}{\partial t} &= \frac{\partial E_z(x,y)}{\partial x} \\ \epsilon(x,y)\frac{\partial E_z(x,y)}{\partial t} + \sigma(x,y)E_z(x,y) &= \left(\frac{\partial H_y(x,y)}{\partial x} - \frac{\partial H_x(x,y)}{\partial y}\right) \end{aligned} \tag{3.56}$$

where $\mu(x,y)$, $\epsilon(x,y)$ and $\sigma(x,y)$ are the 2D permeability, permittivity and electrical conductivity distributions, respectively. In the microwave imaging case, $\mu(x,y) = \mu_0$ due to the non-magnetic nature of biological tissues.

In 1966, Yee devised a special staggered grid [203], referred as the Yee's grid, for discretizing Maxwell's equations in the form of (3.1) or (3.56). In 2D rectangular space, this grid has two variants as illustrated in Figure 3.4.



(a) (b)

Figure 3.4: Two-dimensional FDTD meshes: (a) E-grid, (b) H-grid

In Figure 3.4, (a) is referred as the E-grid while (b) is referred as the H-grid. Despite the apparent differences, they actually denote the same spatial arrangement except that each starts from location (1/2,1/2) of the other grid (nodal spacing set to 1). For consistency, we will only use the E-grid.

Given the indices shown in the figure, the $\vec{H}$ vectors have one index located at half grid spacing denoted by $i \pm 1/2$ or $j \pm 1/2$ (i is the x-index, j is the y-index). Applying the difference representation for all temporal and spatial differential operators in (3.56), i.e.

$$\frac{\partial f}{\partial \xi} = \frac{f_{\xi+\Delta\xi/2} - f_{\xi-\Delta\xi/2}}{\Delta\xi} \tag{3.57}$$

where ξ can be x, y or t. Δx and Δy represent the grid sizes in x and y directions, respectively, and Δt is the time step size. The difference form of $\frac{\partial}{\partial t}$ involves the half time steps $t + \Delta t/2$ and $t - \Delta t/2$. Therefore, the derivative-free terms in (3.56) are replaced by averaged values on these two time steps, i.e.

$$f(t) = \frac{f^{t+\Delta t/2} + f^{t-\Delta t/2}}{2} \tag{3.58}$$

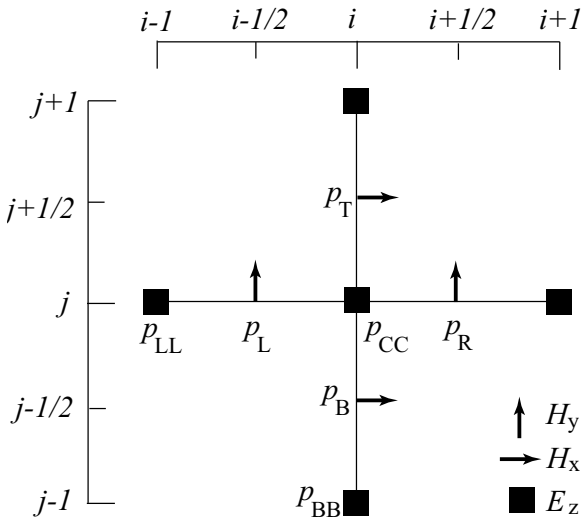


Figure 3.5: Illustration of the vectors around node (i, j)

Thus, the discretized form of equation (3.56) looks like

$$\begin{aligned} \mu(p_B)\frac{H_x^{n+\frac{1}{2}}(p_B)-H_x^{n-1/2}(p_B)}{\Delta t} &= -\frac{E_z^n(p_{CC})-E_z^n(p_{BB})}{\Delta y} \\ \mu(p_L)\frac{H_y^{n+\frac{1}{2}}(p_L)-H_y^{n-1/2}(p_L)}{\Delta t} &= \frac{E_z^n(p_{CC})-E_z^n(p_{LL})}{\Delta x} \\ \epsilon(p_{CC})\frac{E_z^{n+1}(p_{CC})-E_z^n(p_{CC})}{\Delta t} + \sigma(p_{CC})\frac{E_z^{n+1}(p_{CC})+E_z^n(p_{CC})}{2} &= \\ \frac{H_y^{n+\frac{1}{2}}(p_R)-H_y^{n+\frac{1}{2}}(p_L)}{\Delta x} &- \frac{H_x^{n+\frac{1}{2}}(p_T)-H_x^{n+\frac{1}{2}}(p_B)}{\Delta y} \end{aligned} \tag{3.59}$$

where $p_L, p_R, p_T, p_B, p_{CC}, p_{LL}$ and p_{BB} are all symbols of spatial points marked on Figure 3.5.

Moving the terms from the latest time step to the LHS of the equation, i.e. terms at time step $n + 1/2$ in the first two equations and $n + 1$ for the third equation, and moving the remainder to the RHS, we subsequently obtain the explicit update equations for H_x,

H_y and E_z as

$$\begin{aligned} H_x^{n+\frac{1}{2}}(p_B) &= cAHx(p_B)H_x^{n-\frac{1}{2}}(p_B) + cBHx(p_B)\left(\frac{E_z^n(p_{BB})-E_z^n(p_{CC})}{\Delta y}\right) \\ H_y^{n+\frac{1}{2}}(p_L) &= cAHy(p_L)H_x^{n-\frac{1}{2}}(p_L) + cBHy(p_L)\left(\frac{E_z^n(p_{CC})-E_z^n(p_{LL})}{\Delta x}\right) \\ E_z^{n+1}(p_{CC}) &= cAEz(p_{CC})E_z^n(p_{CC}) \\ &\quad + cBEz(p_{CC})\left(\frac{H_y^{n+\frac{1}{2}}(p_R)-H_y^{n+\frac{1}{2}}(p_L)}{\Delta x} - \frac{H_x^{n+\frac{1}{2}}(p_T)-H_x^{n+\frac{1}{2}}(p_B)}{\Delta y}\right) \end{aligned} \tag{3.60}$$

where $cAHx, cBHx$ etc. are the coefficients defined by

$$\begin{aligned} cAHx(p) &= 1 \\ cBHx(p) &= \frac{\Delta t}{\mu(p)} \\ cAHy(p) &= 1 \\ cBHy(p) &= \frac{\Delta t}{\mu(p)} \\ cAEz(p) &= \frac{2\epsilon(p)-\sigma(p)\Delta t}{2\epsilon(p)+\sigma(p)\Delta t} \\ cBEz(p) &= \frac{2\Delta t}{2\epsilon(p)+\sigma(p)\Delta t} \end{aligned} \tag{3.61}$$

Assuming the 2D FDTD mesh stretches throughout space, if we are given the E_z field at $t = 0$, evaluation of the first two equations in (3.60) produces the magnetic fields at $t = \Delta t/2$. The third equation provides the electric fields at $t = \Delta t$. Repeating this update scheme, we obtain the complete history of all field quantities for all time steps.

Unfortunately, for real computations, the mesh size is finite and we must apply boundary conditions to simulate unbounded radiation similar to the previous case. In the following section, we will examine an efficient material-based ABC, the generalized PML absorbing boundary condition.

Generalized PML absorbing boundary condition

One of the most successful material-based ABC's is the perfectly matched layer (PML) technique initially proposed by Berenger in 1994 [8]. The original work focused on the matching the domains filled with lossless media such as free space in 2D. Shortly

after the publication of the first PML paper, this concept was extended to 3D Cartesian space by Berenger [9] and to other orthogonal coordinate systems by Chew, Weedon [35] and Rappaport [165]. For lossy media, Fang and Wu developed a generalized PML (GPML) technique [48, 49] as well as Gedney's uniaxial PML (UPML) technique [61, 62] among others. Considering the lossy nature of biological tissues, we chose to implement the GPML in the 2D FDTD forward computation and the UPML in the 3D FDTD. In this section we will describe the formulations of the 2D GPML while the development of the PML for the 3D FDTD problem can be found in Chapter 5.

The stretched coordinate notation introduced by Chew and Weedon [35] greatly simplifies the derivation and extension of the PML. With this notation, the frequency domain curl equations (3.5) for isotropic lossy media (magnetic current term is also considered since the PML medium is an artificial material) can be written as

$$\begin{aligned} \nabla_s \times \vec{E}(\vec{r}) &= -j\omega\mu(\vec{r})\vec{H}(\vec{r}) - \sigma^*(\vec{r})H(\vec{r}) \\ \nabla_s \times \vec{H}(\vec{r}) &= j\omega\epsilon(\vec{r})\vec{E}(\vec{r}) + \sigma(\vec{r})\vec{E}(\vec{r}) \end{aligned} \tag{3.62}$$

where

$$\nabla_s = \frac{\partial}{\partial\tilde{x}}\hat{x} + \frac{\partial}{\partial\tilde{y}}\hat{y} + \frac{\partial}{\partial\tilde{z}}\hat{z} \tag{3.63}$$

and

$$\begin{aligned} \tilde{x} &= \int_0^x s_x(x')dx' \\ \tilde{y} &= \int_0^y s_y(y')dy' \\ \tilde{z} &= \int_0^z s_z(z')dz' \end{aligned} \tag{3.64}$$

are the stretched coordinates by function $s_x(x)$, $s_y(y)$ and $s_z(z)$ which are referred as the coordinate stretching coefficients. Substituting (3.64) into (3.63) we get

$$\nabla_s = \frac{1}{s_x}\frac{\partial}{\partial x}\hat{x} + \frac{1}{s_y}\frac{\partial}{\partial y}\hat{y} + \frac{1}{s_z}\frac{\partial}{\partial z}\hat{z} \tag{3.65}$$

from which the meanings of the coordinate stretching coefficient can be readily illus-

trated. Consequently, if $s_x = s_y = s_z = 1$, the electric and magnetic field distributions computed from (3.62) will be identical to the solution of the normal Maxwell's equations.

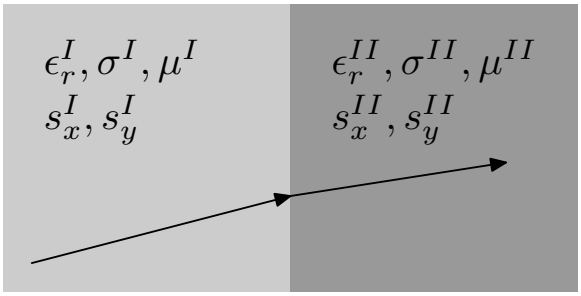


Figure 3.6: Matching condition at an interface perpendicular to x-axis

Considering a vertical interface perpendicular to x-axis between two lossy medium with each one characterized by a set of dielectric properties and coordinate stretching coefficients as noted in Figure 3.6, Fang and Wu showed in [49] that if the following conditions are met at the interface

$$\begin{aligned} \epsilon^I &= \epsilon^{II} \\ \mu^I &= \mu^{II} \\ s_y^I &= s_y^{II} \\ s_z^I &= s_z^{II} \end{aligned} \tag{3.66}$$

the plane waves incident at any direction with any polarization and frequency will be transmitted to the second medium with zero reflection. Applying the rotational rule to the subscripts in (3.66) (i.e. replacing x by y, y by z and z by x denoted as $x \rightarrow y \rightarrow z \rightarrow x$), condition (3.66) will match the interfaces perpendicular to y and z axis.

To transmit waves without reflections from medium 1 to medium 2 is only part of the rationale behind the PML absorbing boundary. The other is to attenuate the wave efficiently in medium 2 such that even if the medium 2 is terminated with a reflecting boundary (such as perfect electrical conductor - PEC), only minimal signals are

reflected back into medium 1. To achieve the second goal, the artificial anisotropic conductivity, σ_ξ, and magnetic conductivity, σ^*_ξ satisfying the matching condition as demonstrated by Berenger [8], i.e.

$$\frac{\sigma_\xi}{\epsilon} = \frac{\sigma^*_\xi}{\mu} \tag{3.67}$$

are introduced and embedded in the coordinate stretching coefficient s function. One possible form for s is

$$s_\xi(\xi) = s_{\xi 0}(\xi)\left(1 + \frac{\sigma_\xi(\xi)}{j\omega\epsilon}\right) \tag{3.68}$$

where $\xi = x, y, z$ and $s_{\xi 0}(\xi)$ and $\sigma_\xi(\xi)$ are functions that need to be determined. Fang and Wu proposed the following form

$$\begin{aligned} s_{\xi 0}(\xi) &= 1 + s_{max}\left(\frac{\xi}{\Delta\xi}\right)^{s_{exp}} \\ \sigma_{\xi 0}(\xi) &= \sigma_{max}\sin^2\left(\frac{\pi\xi}{2\Delta\xi}\right) \end{aligned} \tag{3.69}$$

for $s_{\xi 0}(\xi)$ and $\sigma_\xi(\xi)$ where ξ denotes the distance to the boundary (note that σ^*_ξ can be consequently determined from (3.67)); $\Delta\xi$ is the grid size along that direction; s_{max}, s_{exp} and σ_{max} are constants.

In the 2D Cartesian coordinate system, in order to match all boundaries, the PML media are typically arranged in the manner suggested in Figure 3.7. Eight PML slabs are used to surround the working volume among which four edge slabs are used to match the working volume boundaries and four corner slabs are used to match the neighboring edge slabs. The working volume contains the inhomogeneous target structure which is surrounded by the homogeneous background medium with properties ϵ_{bk}, μ_{bk} and σ_{bk}. As has been previously shown, the working volume can be treated as a special PML medium by simply setting $s_x = s_y = 1$ and $\sigma_\xi = \sigma^*_\xi = 0$. From (3.69), this

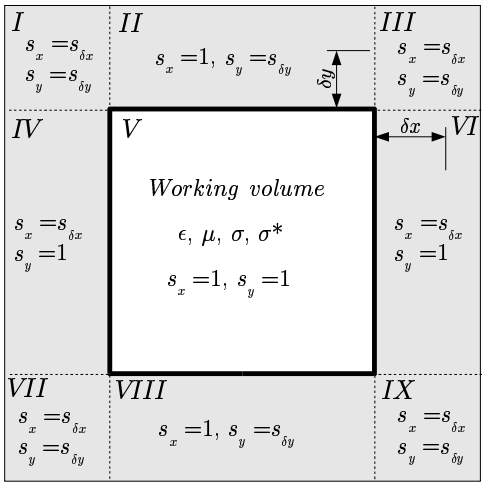


Figure 3.7: Coordinate stretching coefficients in various PML slabs.

implies that $s_{x0} = s_{y0} = 1$ and $\sigma_{x0} = \sigma_{y0} = 0$. Additionally, from the matching condition (3.66), the coefficient s_y in the PML media of regions IV and VI should match that of the working volume. Likewise, so should be s_x in regions II and VIII. In order to effectively attenuate the outgoing wave, the s function in the parallel direction to the interface inside the edge slabs should have gradually increasing values as depicted in equation (3.69). This gradually increasing function is denoted as s_δ where δ is the perpendicular distance to the interface. In the corner regions, the attenuated waves in the neighboring edge slabs enter through interior interfaces (denoted by dashed lines in Figure 3.7). In order to match the medium at these interfaces, the s functions in the corner regions should match the perpendicular species of their neighbors. i.e., $s_x = s_y = s_\delta$. Besides the settings of s functions, the ϵ, μ and σ value of all PML slabs are identical to the most exterior value of the working volume, i.e. ϵ_{bk}, μ_{bk} and σ_{bk}. To summarize the settings of the PML medium, Table 3.3 lists the dielectric values and coordinate stretching functions in all 8 PML slabs.

Since the EM waves are sufficiently attenuated inside the surrounding PML region, the most exterior boundaries (denoted by thin solid lines in Figure 3.7) can safely use a

Table 3.3: PML settings in the various domains

region	PML parameter settings
I	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = s_{\delta x}, s_y = s_{\delta y}$
II	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = 1, s_y = s_{\delta y}$
III	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = s_{\delta x}, s_y = s_{\delta y}$
IV	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = s_{\delta x}, s_y = 1$
V	inhomogeneities, $s_x = 1, s_y = 1$
VI	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = s_{\delta x}, s_y = 1$
VII	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = s_{\delta x}, s_y = s_{\delta y}$
VIII	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = 1, s_y = s_{\delta}$
IX	$\epsilon_{bk}, \mu_{bk}, \sigma_{bk}, s_x = s_{\delta x}, s_y = s_{\delta y}$

PEC boundary condition, i.e. setting the tangential E field to zeros.

A split-field Maxwell's equation was employed by Fang and Wu [49] by letting

$$E_z(t,\vec{r}) = E_{zx}(t,\vec{r}) + E_{zy}(t,\vec{r}) \tag{3.70}$$

and the frequency domain Maxwell's equations (3.62) in the 2D TM wave case were expanded to

$$\begin{aligned}
-\frac{\partial H_x(\vec{r})}{s_{y0}(y)\partial y} &= \left(j\omega\epsilon + \sigma + \sigma_y + \frac{\sigma\sigma_y}{j\omega\epsilon}\right) E_{zy}(\vec{r}) \\
\frac{\partial H_y(\vec{r})}{s_{x0}(x)\partial x} &= \left(j\omega\epsilon + \sigma + \sigma_x + \frac{\sigma\sigma_x}{j\omega\epsilon}\right) E_{zx}(\vec{r}) \\
\frac{\partial E_z(\vec{r})}{s_{x0}(x)\partial x} &= \left(j\omega\mu + \sigma^* + \sigma_x^* + \frac{\sigma^*\sigma_x^*}{j\omega\mu}\right) H_y(\vec{r}) \\
-\frac{\partial E_z(\vec{r})}{s_{y0}(y)\partial y} &= \left(j\omega\mu + \sigma^* + \sigma_y^* + \frac{\sigma^*\sigma_y^*}{j\omega\mu}\right) H_x(\vec{r})
\end{aligned} \tag{3.71}$$

Performing Fourier transforms to both sides of equation (3.71), the time-domain representation is obtained as follows

$$\begin{aligned}
-\frac{\partial H_x(t,\vec{r})}{s_{y0}(y)\partial y} &= \epsilon\frac{\partial E_{zy}(t,\vec{r})}{\partial t} + (\sigma + \sigma_y)E_{zy}(t,\vec{r}) + \frac{\sigma\sigma_y}{\epsilon}E^I_{zy}(t,\vec{r}) \\
\frac{\partial H_y(t,\vec{r})}{s_{x0}(x)\partial x} &= \epsilon\frac{\partial E_{zx}(t,\vec{r})}{\partial t} + (\sigma + \sigma_y)E_{zx}(t,\vec{r}) + \frac{\sigma\sigma_x}{\epsilon}E^I_{zx}(t,\vec{r}) \\
\frac{\partial E_z(t,\vec{r})}{s_{x0}(x)\partial x} &= \mu\frac{\partial H_y(t,\vec{r})}{\partial t} + (\sigma^* + \sigma_x^*)H_y(t,\vec{r}) + \frac{\sigma^*\sigma_x^*}{\mu}H^I_y(t,\vec{r}) \\
-\frac{\partial E_z(t,\vec{r})}{s_{y0}(y)\partial y} &= \mu\frac{\partial H_x(t,\vec{r})}{\partial t} + (\sigma^* + \sigma_y^*)H_x(t,\vec{r}) + \frac{\sigma^*\sigma_y^*}{\mu}H^I_x(t,\vec{r})
\end{aligned} \tag{3.72}$$

where

$$\begin{aligned} E^I_{zx}(t,\vec{r}) &= \int_{-\infty}^{t} E_{zx}(t',\vec{r})dt' \\ E^I_{zy}(t,\vec{r}) &= \int_{-\infty}^{t} E_{zy}(t',\vec{r})dt' \\ H^I_{x}(t,\vec{r}) &= \int_{-\infty}^{t} H_{x}(t',\vec{r})dt' \\ H^I_{y}(t,\vec{r}) &= \int_{-\infty}^{t} H_{y}(t',\vec{r})dt' \end{aligned} \tag{3.73}$$

Equation (3.72) can consequently be discretized using the E-grid scheme mentioned above. Notice that the split fields E_{zx} and E_{zy} are located at the same point as E_z. The discretized update equations for the 2D cases are listed below:

$$\begin{aligned} H_x^{n+\frac{1}{2}}(p_B) &= cAHx(p_B)H_x^{n-\frac{1}{2}}(p_B) + cBHx(p_B)\left(\frac{E_z^n(p_{BB})-E_z^n(p_{CC})}{\Delta y}\right) \\ H_y^{n+\frac{1}{2}}(p_L) &= cAHy(p_L)H_x^{n-\frac{1}{2}}(p_L) + cBHy(p_L)\left(\frac{E_z^n(p_{CC})-E_z^n(p_{LL})}{\Delta x}\right) \\ E_{zx}^{n+1}(p_{CC}) &= cAEzx(p_{CC})E_{zx}^n(p_{CC}) + cBEzx(p_{CC})\left(\frac{H_y^{n+\frac{1}{2}}(p_R)-H_y^{n+\frac{1}{2}}(p_L)}{\Delta x}\right) + cCEzx(p_{CC}){E_{zx}^I}^n(p_{CC}) \\ E_{zy}^{n+1}(p_{CC}) &= cAEzy(p_{CC})E_{zy}^n(p_{CC}) + cBEzy(p_{CC})\left(\frac{H_x^{n+\frac{1}{2}}(p_T)-H_x^{n+\frac{1}{2}}(p_B)}{\Delta y}\right) + cCEzy(p_{CC}){E_{zy}^I}^n(p_{CC}) \\ {E_{zx}^I}^{n+1}(p_{CC}) &= {E_{zx}^I}^n(p_{CC}) + \Delta t E_{zx}^{n+1}(p_{CC}) \\ {E_{zy}^I}^{n+1}(p_{CC}) &= {E_{zy}^I}^n(p_{CC}) + \Delta t E_{zy}^{n+1}(p_{CC}) \end{aligned} \tag{3.74}$$

where

$$\begin{aligned} cAHx(p) &= \frac{2\mu(p)-(\sigma^*(p)+\sigma_x^*(p))\Delta t}{2\mu(p)+(\sigma^*(p)+\sigma_x^*(p))\Delta t} \\ cBHx(p) &= \frac{2\Delta t}{2\mu(p)+(\sigma^*(p)+\sigma_x^*(p))\Delta t} \\ cAHy(p) &= \frac{2\mu(p)-(\sigma^*(p)+\sigma_x^*(p))\Delta t}{2\mu(p)+(\sigma^*(p)+\sigma_x^*(p))\Delta t} \\ cBHy(p) &= \frac{2\Delta t}{2\mu(p)+(\sigma^*(p)+\sigma_x^*(p))\Delta t} \\ cAEzx(p) &= \frac{2\epsilon(p)-\sigma(p)\Delta t}{2\epsilon(p)+\sigma(p)\Delta t} \\ cBEzx(p) &= \frac{2\Delta t}{2\epsilon(p)+\sigma(p)\Delta t} \\ cCEzx(p) &= \frac{2\Delta t\sigma(p)\sigma_x(p)}{2\epsilon(p)+\sigma(p)\Delta t} \\ cAEzy(p) &= \frac{2\epsilon(p)-\sigma(p)\Delta t}{2\epsilon(p)+\sigma(p)\Delta t} \\ cBEzy(p) &= \frac{2\Delta t}{2\epsilon(p)+\sigma(p)\Delta t} \\ cCEzy(p) &= \frac{2\Delta t\sigma(p)\sigma_y(p)}{2\epsilon(p)+\sigma(p)\Delta t} \end{aligned} \tag{3.75}$$

and P_B, P_L,P_{CC} are defined in Figure 3.5.

Equation (3.74) and (3.75) provide the update scheme for all interior field components. Among the fields located at the most exterior boundary, the E_z component is the

tangential electric field for 2D problems. Therefore, to apply the PEC BC on the outer boundary, these E vectors are simply initialized to zeros and left un-updated. The H field components on the exterior boundary are treated similarly since they are also zero.

A minor problem with the $N_x = N_y = 4$ and $N_{PML} = 3$ case is illustrated in Figure 3.8. From the figure, the array size for E_z (including E_{zx} and E_{zy}) is $(N_x + 2N_{PML} + 1) \times (N_y + 2N_{PML} + 1)$, that of H_x is $(N_x + 2N_{PML} + 1) \times (N_y + 2N_{PML})$, and that of H_y is $(N_x+2N_{PML})\times(N_y+2N_{PML}+1)$. Because the fields on the outer boundary do not update, the active array sizes are $(N_x+2N_{PML})\times(N_y+2N_{PML})$, $(N_x+2N_{PML}-1)\times(N_y+2N_{PML})$ and $(N_x + 2N_{PML}) \times (N_y + 2N_{PML} - 1)$ for E_z, H_x and H_y, respectively.

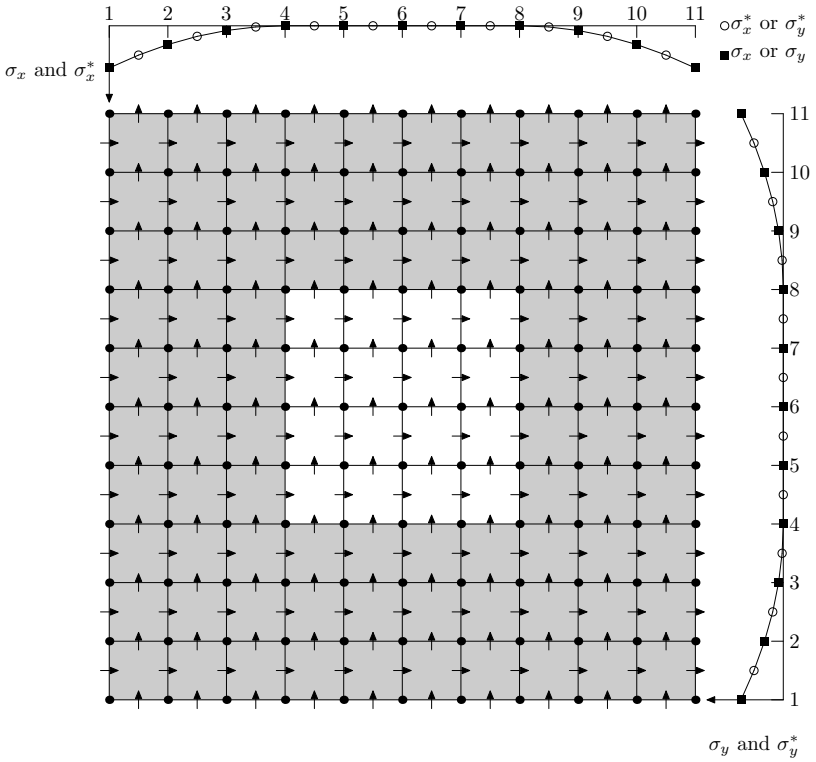


Figure 3.8: Illustration of the FDTD mesh with GPML boundary condition and values of the isotropic σ functions. The shaded cells are within the PML layers.

It should be noted that since $s_x = s_y = 1$ in the working volume, the differential equation with stretched coordinate is identical to the conventional curl equations;

therefore, the update scheme (3.74) can be replaced by (3.60) without altering the field values. This replacement can enhance the time savings because the total algebraic operation number is smaller with the conventional update scheme. However, to implement this hybrid update scheme, one needs to split the spatial update for each region individually. This lengthens the program and complicates both debugging and code maintenance. We applied the GPML update scheme throughout all subregions including the working volume in our programs for simplicity.

Source implementation and steady state extraction

Equation (3.60) and (3.74) provide a mechanism to model the time-marching EM fields over 2D grids. However, in order to initiate the wave propagation, sources or excitations need to be incorporated into the update equation at each time step by means of initial values, electrical currents, voltages or magnetic currents [102]. For the monopole antenna used in the microwave imaging system at Dartmouth, the source is modelled as a z-orientated time-harmonic point current source in form of

$$\vec{J}_z(t, \vec{r}) = J_0 \cos(\omega t)\delta(\vec{r} - \vec{r}_s) \tag{3.76}$$

where $J_0 = 1/(\omega\mu_0)$ is the amplitude and $\vec{r}_s$ is the spatial location of the point source. Notice that the phasor form of the source in (3.76) is $J_z(\vec{r}) = J_0 \exp(\omega t)\delta(\vec{r} - \vec{r}_s)$ whose Fourier transform gives the source similar to that in (3.45). When the working volume is filled with the homogeneous background medium, the implementation of this source produces the Green's function (3.47) in the frequency domain.

Incorporating the source term into (3.72), we get the discretized update equations

for the E field as follows

$$\begin{aligned} E_{zx}^{n+1}(p_{CC}) &= cAEzx(p_{CC})E_{zx}^{n}(p_{CC}) + cBEzx(p_{CC})\left(\frac{H_y^{n+\frac{1}{2}}(p_R)-H_y^{n+\frac{1}{2}}(p_L)}{\Delta x}\right) \\ &\quad +cCEzx(p_{CC})E_{zx}^{I\ n}(p_{CC}) + \frac{|J|}{2}\cos(\omega(n\Delta t)) \\ E_{zy}^{n+1}(p_{CC}) &= cAEzy(p_{CC})E_{zy}^{n}(p_{CC}) + cBEzy(p_{CC})\left(\frac{H_x^{n+\frac{1}{2}}(p_T)-H_x^{n+\frac{1}{2}}(p_B)}{\Delta y}\right) \\ &\quad +cCEzy(p_{CC})E_{zy}^{I\ n}(p_{CC}) + \frac{|J|}{2}\cos(\omega(n\Delta t)) \end{aligned} \tag{3.77}$$

where $|J| = \frac{2\Delta t}{(2\epsilon+\sigma\Delta t)(\omega\mu_0)\Delta x\Delta y}$ is the discretized source amplitude. (Here we assume $\vec{r}_s$ is located at P_{CC}. If $\vec{r}_s$ is not exactly located at a grid node, the source term in (3.77) is divided into weighted fragments and assigned to the four nodes closest to vector $\vec{r}_s$.)

In a time-harmonic imaging system, the fields required for the microwave imaging algorithm are typically the steady-state amplitude and phase distributions, i.e. the phasor or frequency domain solution. Therefore, an amplitude and phase extraction process needs to be performed to the time-varying fields provided by the update scheme (3.74). In order to obtain the steady-state field distribution, the time marching-on process needs to run sufficiently long to allow the reflections from various objects to propagate through the domain. Fortunately, due to the lossiness of the background medium used in our microwave imaging system, the time for the wave to propagate from one side to the other and back is typically sufficient to allow the field to reach steady-state.

At steady-state, all field components at all locations oscillate sinusoidally. One way to extract the amplitude and initial phase of the sine curve is to record the values for a period and perform fast Fourier transform (FFT). However, Oğuz and Gürel introduced a much faster and more convenient two-point extraction scheme [150]. Assuming the field values at two consecutive time steps, time steps n and $n+1$, are recorded and can be expressed as

$$\begin{aligned} f^{(n)} &= A\sin(n\Delta t+\phi_0) \\ f^{(n+1)} &= A\sin(n\Delta t+\Delta t+\phi_0) \end{aligned} \tag{3.78}$$

the amplitude A and the initial phase ϕ_0 can be analytically solved for by

$$\begin{aligned} A &= \csc(\Delta t)\sqrt{\left(\sin(\Delta t) f^{(n+1)}\right)^2 + \left(\cos(\Delta t) f^{(n+1)} - f^{(n)}\right)^2} \\ \phi_0 &= \tan^{-1}\left(\cot(\Delta t) - \frac{f^{(n)}}{f^{(n+1)}}\csc(\Delta t)\right) - n\Delta t \end{aligned} \tag{3.79}$$

where $\sin(\Delta t)$,$\cos(\Delta t)$,$\csc(\Delta t)$ and $\cot(\Delta t)$ can be pre-computed. This algorithm only requires a record of the field value from the previous time step and the computational expense arises primarily from the evaluation of the $\tan^{-1}$ and the square root which are not significant compared with the expenses for the time update.

Gürel and Oğuz also found that applying a low-pass filter [70] to the source (3.76) reduces the steady-state numerical noise. The filtered current source can be expressed as

$$\vec{J}_z(t, \vec{r}) = w(t) J_0 \cos(\omega t)\delta(\vec{r} - \vec{r}_s) \tag{3.80}$$

where

$$w(t) = \begin{cases} 0.5 - 0.5\cos(\omega t/\tau) & 0 \le x \le \tau \quad \text{(Hamming)} \\ 0.54 - 0.46\cos(\omega t/\tau) & 0 \le x \le \tau \quad \text{(Hanning)} \\ 0.42 - 0.5\cos(\omega t/\tau) + 0.08\cos(\omega t/\tau) & 0 \le x \le \tau \quad \text{(Blackman)} \\ 1 & x > \tau \end{cases} \tag{3.81}$$

is the filter function where τ is the length of the filter. The Hamming window filter is selected because of the improved performance demonstrated in [70].

Stability and dispersion error

The FDTD update equations in (3.60) or (3.74) use an explicit leap-frog time-stepping scheme. This scheme is conditionally stable when the spatial and temporal step sizes

satisfy the Courant-Friedrichs-Lewy (CFL) stability condition [188]:

$$\Delta t \leq \frac{1}{c_{max}\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}}} \tag{3.82}$$

where c_{max} is the maximum wave speed throughout the computational domain. The CFL number (CFLN) is defined by

$$CFLN = c_{max}\Delta t\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}} \tag{3.83}$$

which is a number smaller than 1. On uniform grids where $\Delta x = \Delta y = \Delta$, (3.82) requires $\Delta t \leq \frac{\Delta}{\sqrt{2}c_{max}}$. If inhomogeneities are present in the computational domain, c_{max} is determined utilizing the minimum permittivity and conductivity values by

$$c_{max} = \frac{\omega}{\Im\mathrm{m}(k)_{min}} \tag{3.84}$$

where k is defined in (3.8).

The alternative-direction-implicit (ADI) FDTD method [209, 140] is unconditionally stable and is used in both 2D and 3D reconstructions where it can operate beyond the CFL limit. This ADI approach is introduced in Section 5.1.7 to accommodate additional optimizations in the forward field computation.

The numerical dispersion in the FDTD method refers to the phenomena that the wave propagates with different wave speeds along different directions on the FDTD grid. In [188], Taflove demonstrated that the numerical dispersion at different angle, θ, is related to the spatial and temporal step sizes by

$$\left(\frac{1}{c\Delta t}\sin\frac{\omega\Delta t}{2}\right)^2 = \left(\frac{1}{\Delta x}\sin\frac{|k_\theta|\cos\theta\Delta x}{2}\right)^2 + \left(\frac{1}{\Delta y}\sin\frac{|k_\theta|\sin\theta\Delta y}{2}\right)^2 \tag{3.85}$$

where $|k_\theta|$ is the absolute value of the wave number along the θ direction. The wave speed along that direction is

$$v_p = c\frac{2\pi}{|k_\theta|} \tag{3.86}$$

The denser the mesh, the smaller the dispersion error. Low-dispersion FDTD methods have attracted considerable interests in the last decade. Oğuz and Gürel developed a simple compensation scheme to reduce the dispersion error [150]. High order FDTD methods [47, 193] also demonstrate effectiveness in reducing the dispersion error. In Section 5.1.6, we implement a 4-th order FDTD method in the construction of 3D FDTD method.

Accuracy of the 2D FDTD solver

In this subsection, we investigate the accuracy of the 2D FDTD method proposed in previous subsections. The optimal values of the PML parameters, the impact of the mesh density to dispersion error and other related issues are discussed in the context of two benchmark problems: B1) a homogeneous lossy background medium, and B2) a 2D cylindrical lossy object within a homogeneous lossy background medium.

As demonstrated in Section 3.6.1, the GPML medium is characterized by four parameters, the number of layers N_{PML}, and the values of s_{max}, s_{exp} and σ_{max} in (3.69). For simplicity, a GPML medium is denoted by a quadruplet in the form of (N_{PML}, s_{exp}, s_{max}, σ_{max}). The selection of these parameters affects the accuracy and the computational expense of this algorithm. Theoretical analysis as well as simulations were performed for lossless cases in [62]. We have repeated the numerical analysis for the lossy cases, i.e. investigating the absorption efficiency of the GPML characterized at different parameter settings.

Problem B1 is utilized to investigate the absorption performance of the GPML

ABC. The working volume is a uniform rectangular grid with $N_x = N_y = 100$ and $\Delta x = \Delta y = \Delta$. The background medium has $\epsilon_{bk} = 25$ and $\sigma_{bk} = 1.0$ S/m to simulate an 83% glycerin solution. A point current source in the form of (3.76) is located at the center of the grid to excite cylindrical waves at 900 MHz. The spatial step size Δ is determined by $\Delta = \frac{\lambda_{bk}}{R}$ where R is the mesh resolution and λ_{bk} is the wavelength in the background medium, in this case $R = 20$ and $\Delta = 0.3$ cm. The time step is subsequently determined by letting CFLN=0.90 from (3.83). The forward solution at $t = 100\Delta t$ is computed from the update scheme (3.74) and (3.77) under a given GPML setting. As an experimental control, a reference field distribution is computed within a larger mesh centered at the source with $N_x = N_y = 400$ and the same PML configuration. Since the wave front is still far away from the domain boundary in the larger mesh case, the reference solution can be considered to be approaching the unbounded numerical solution. The differences between the E_z components computed from the small mesh solution $E_z(i, j)$ and large mesh solution $E_z^r(i, j)$ are measured by the l^1-norm of the error defined by

$$e_{L1}(N_{PML}, s_{exp}, s_{max}, \sigma_{max}) = \sum_{i=1}^{100} \sum_{j=1}^{100} \left| |E_z(i, j)| - \left|E_z^r(i, j)\right| \right| \tag{3.87}$$

Equation (3.87) provides a 5D data array with axes along N_{PML}, s_{exp}, s_{max}, σ_{max} and e_{L1}. For simplicity, we computed the error norm distribution at $N_{PML} = 5, 8, 10, 12, 15$ while varying the remaining three parameters over a rectangular grid ($0.5 \leq s_{exp} \leq 7$,$0.5 \leq s_{max} \leq 10$,$0.5 \leq \sigma_{max} \leq 10$ in step sizes of 0.5). The contour plots of these error distributions are shown in Figures 3.9 (a)-(e). From all five contour plots, the diminishing error norm with respect to increasing N_{PML} is evident. Additionally, there exist optimal values for each of the other three parameters for a given N_{PML} (the contour slices cut through these optimal values in Figure 3.9). The approximate optimal

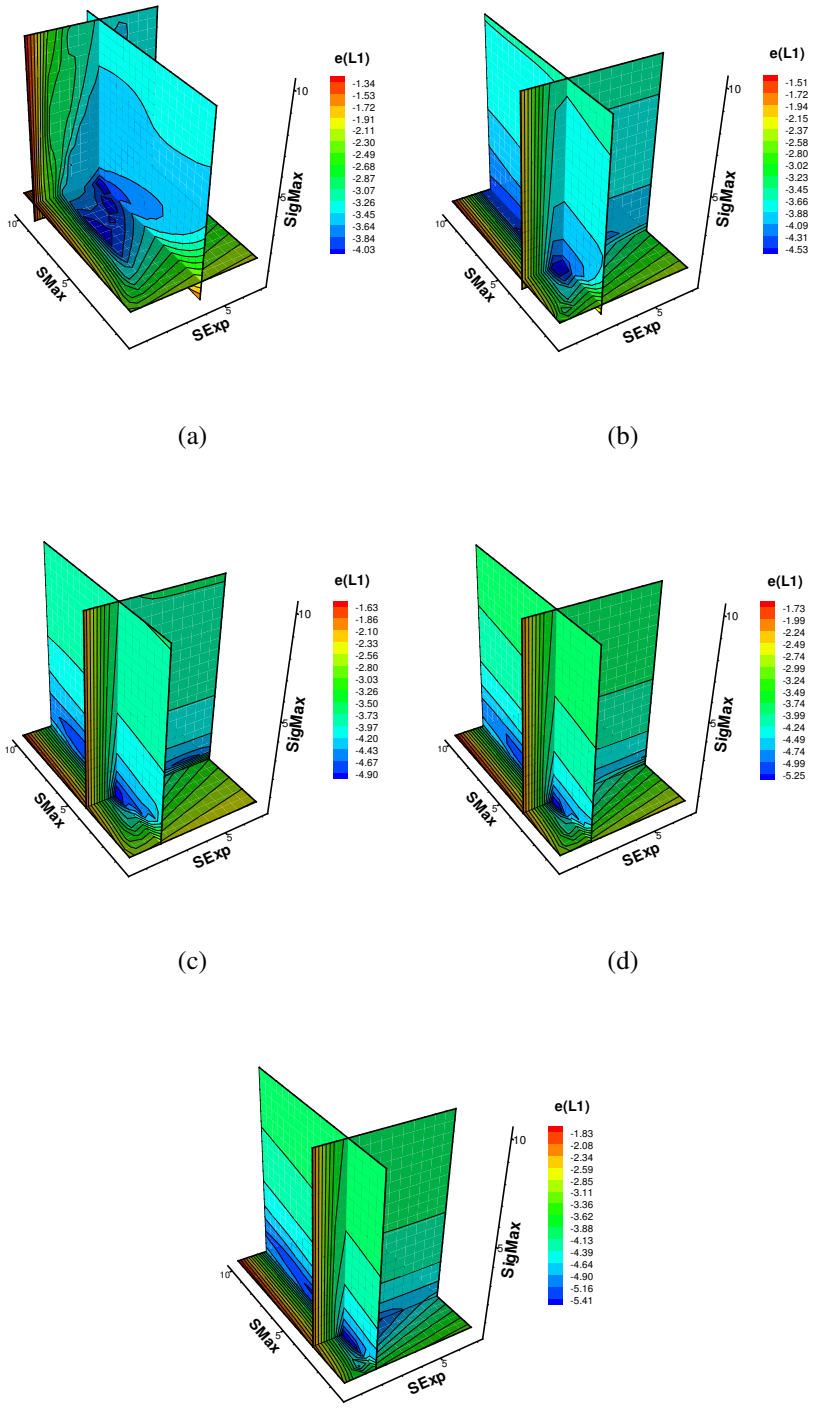


Figure 3.9: L1 reflection error for various PML layers (a) $N_{PML} = 5$, (b) $N_{PML} = 8$, (c) $N_{PML} = 10$, (d) $N_{PML} = 12$, (e) $N_{PML} = 15$. Contours are plotted along the planes cutting through the point with minimum error.

Table 3.4: Optimal PML parameters at various thicknesses

N_{PML}	s_{exp}	s_{max}	σ_{max}	e_{L1}
5	4.0	9.0	1.5	0.60103233E-04
8	2.5	3.5	1.0	0.18133276E-04
10	2.0	4.0	0.5	0.73576640E-05
12	2.0	3.5	0.5	0.31877565E-05
15	2.0	3.0	0.5	0.21373473E-05

stretching coefficients at different PML layer numbers ($R = 20$) are listed in Table 3.4.

A second study investigated the dispersion error of the 2D FDTD scheme. The frequency-domain solution or steady-state phasor solution for the benchmark problem B1 can be analytically expressed using equation (3.47). In this case, we use the uniform rectangular grid centered at the source to compute the steady-state forward field whose amplitude and phase are extracted by the two-point scheme (3.79). The background medium consists of the 83% glycerin solution and 12 layers of GPML cells are applied while the stretching coordinate parameters were chosen using the optimal values in Table 3.4. The numerical frequency domain responses $E_z(i, j)$ are obtained over various mesh resolutions while fixing the domain physical size. The corresponding grid sizes at mesh resolution $R = 10, 20, 30, 40$ are $N_x = N_y = 50$, $N_x = N_y = 100$, $N_x = N_y = 150$ and $N_x = N_y = 200$, respectively.

The differences between the analytical and numerical solutions are illustrated in Figures 3.10 and 3.11. From these plots, it is quite obvious that the dispersion error is significantly reduced when the mesh resolution is increased. The root-mean-square (RMS) amplitude (dB) and the phase (radians) differences are computed by

$$\begin{aligned} e_{amp} &= 20\sqrt{\frac{1}{N_xN_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left(\log_{10}|E_z(i,j)| - \log_{10}|E_z^a(i,j)|\right)^2} \\ e_{pha} &= \sqrt{\frac{1}{N_xN_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left(\angle E_z(i,j) - \angle E_z^a(i,j)\right)^2} \end{aligned} \tag{3.88}$$

at the four mesh resolutions and are plotted in Figure 3.12 (a) and (b). From the RMS

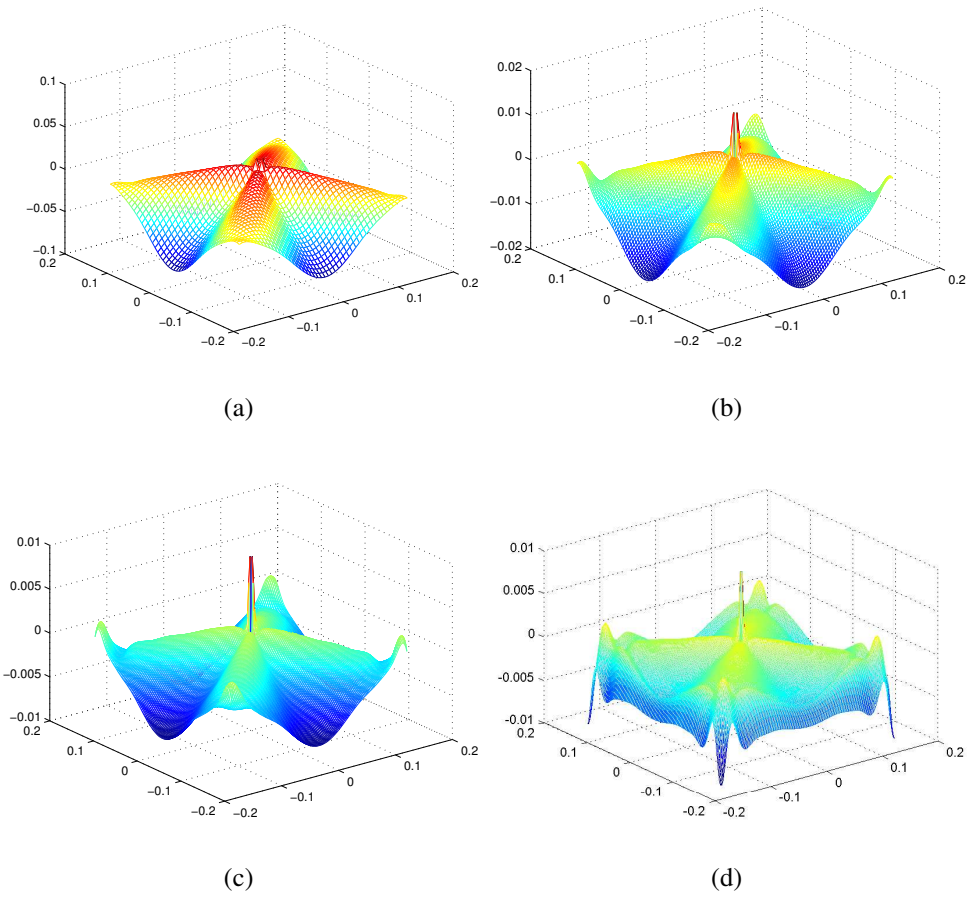


Figure 3.10: Dispersion error (dB amplitude) at various mesh densities for the benchmark problem B1. (a) R=10, (b) R=20, (c) R=30, (d) R=40.

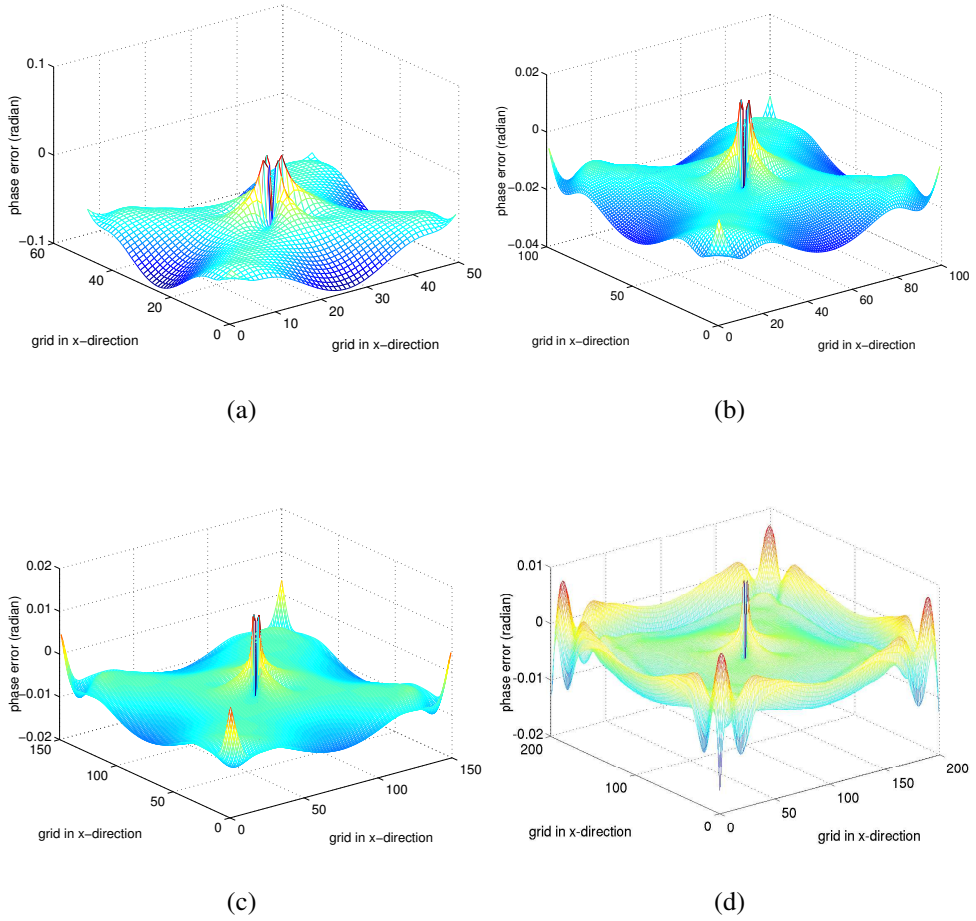
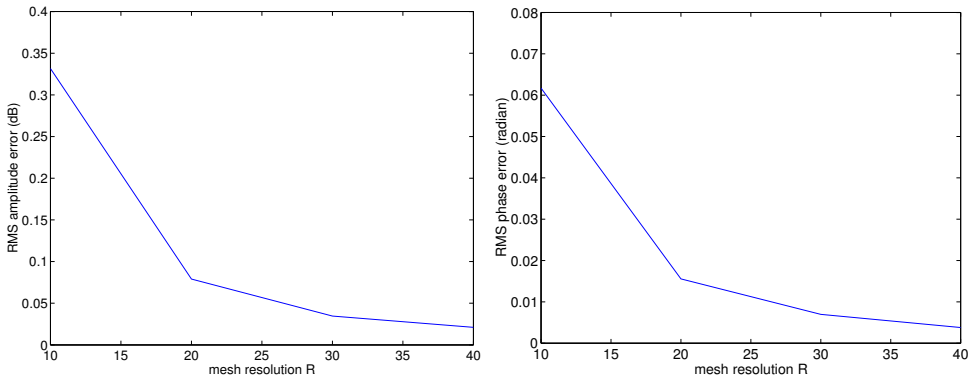


Figure 3.11: Dispersion error (phase in radians) at various mesh densities for the benchmark problem B1. (a) R=10, (b) R=20, (c) R=30, (d) R=40.

error curves, the averaged amplitude difference at $R = 20$ is roughly 0.08 dB while that for the phase is 0.9 degrees.



(a) (b)

Figure 3.12: RMS dispersion error at various mesh densities for the benchmark problem B1 (a) amplitude, (b) phase.

The third study tested the number of time steps required to reach the steady-state. As mentioned briefly above, if the background medium is very lossy, the time-step number to reach stead-state is much less than in lossless cases. To verify this, the benchmark B1 configuration with the identical settings as the $R = 20$ case in the dispersion error study is investigated except with the central source being replaced by a circular antenna array as shown in Figure 3.3. The amplitudes and phases are extracted at time-steps $t = 150\Delta t$ to $t = 600\Delta t$ in $50\Delta t$ increments. The L1 amplitude errors at these time steps are normalized by the L1 amplitude error at $t = 1500\Delta t$. This normalized error curve is plotted vs. time-step Figure 3.13. The amplitudes at the receiver directly opposite the transmitter are computed for all time-step (Figure 3.14 (a)). From both plots, it is reasonable to assert that 250 to 350 time steps is sufficient to reach steady-state. Recognizing that the round-trip time-step number, N_r, for this mesh is 200, we can basically use N_r plus a fixed number of extra steps to estimate the steady-state time step where

the extra steps are used to account for the effect of the source low-pass filter in (3.80).

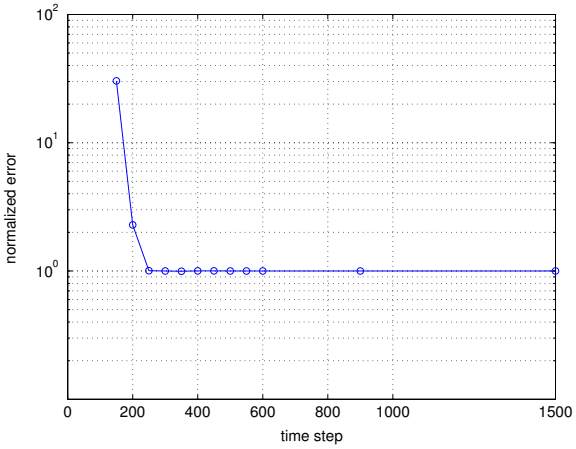


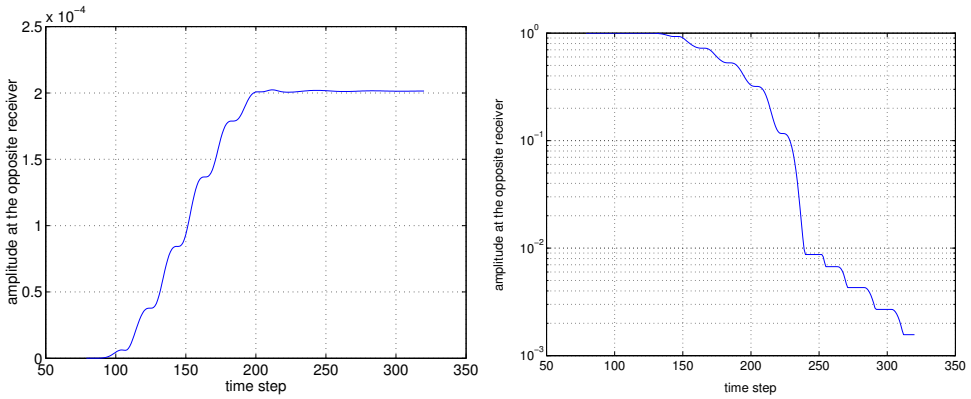
Figure 3.13: Normalized L1 amplitude error at various time-steps.

Due to the oscillatory nature of the curve in Figure 3.14 (a), we devised a monotonic decreasing metric for determining the steady-state time-step from the amplitudes at the opposing receiver. This metric is defined as

$$\gamma = \frac{|E|_{max} - |E|_{min}}{|E|_{max}} \tag{3.89}$$

where $|E|_{max}$ and $|E|_{min}$ are the maximum and minimum values within a sliding window in the amplitude sequence. The length of the window must be at least one full period of the EM wave to avoid oscillations. The corresponding plot of γ for the amplitude sequence in Figure 3.14 (a) is plotted in Figure 3.14 (b). In practice, we typically use $\gamma < 10^{-2}$ as the FDTD time-stepping stopping criterion.

For inhomogeneous problems, we used the B2 benchmark to repeat the above analysis where similar conclusions can be drawn. The plots are omitted here.



(a) (b)

Figure 3.14: Study of the time step number required for achieving steady-state (a) the amplitude extracted at the opposite receiver for all time-steps, (b) values of γ for all time steps.

Efficiency of 2D FDTD method compared with 2D FE/BE method

In this subsection, we discuss the computational efficiency of the FDTD approach compared with the scalar method using the FE/BE hybrid in evaluating the frequency-domain EM field solutions for tomographic imaging settings.

In order to make these comparisons, we used a uniform rectangular grid of size $N_x = N_y = N + 2N_{PML}$ (N_{PML} is the thickness of the PML absorbing boundary layer) as the mesh for the 2D FDTD method while the corresponding circular FE mesh is located at the center of the grid within a circular area of radius $r = N/2$. The meshes for both methods are illustrated in Figure 3.15. The comparison is performed by counting the total floating-point operation (flop) numbers [67] for both methods in producing a single steady-state solution. For simplicity, the operations for constructing the FE matrix (A in (3.14)), building and solving the BE matrix and those for computing the FDTD update coefficients in (3.75) are neglected as are in the amplitude/phase extraction calculations.

For the 2D FE method, the total unknown number in the FE domain is roughly $\frac{\pi}{4}N^2$.

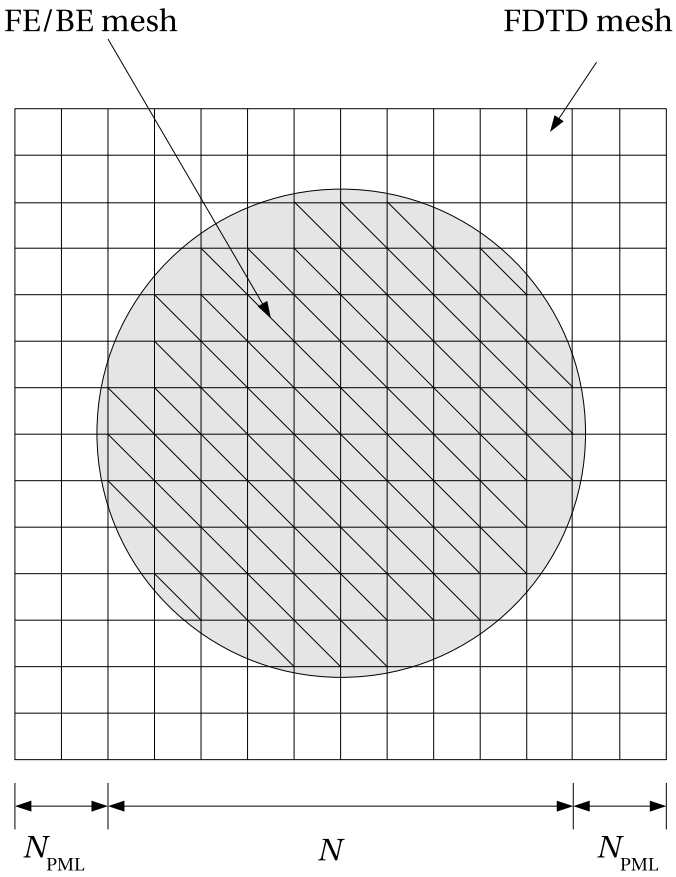


Figure 3.15: Meshes used for the efficiency comparison.

Therefore, the size of the LHS matrix is $(\frac{\pi}{4}N^2)\times(\frac{\pi}{4}N^2)$ and is sparse. The half-bandwidth of the matrix is related to the node numbering scheme. As discussed in Section 3.4, the incorporation of the BE equation effects the bandwidth of the FE matrix which is at least the number of boundary nodes. In this case, the number of boundary nodes is roughly πN which approximately equals the half-bandwidth. Utilizing the symmetry feature of the FE matrix, a banded Cholesky factorization algorithm is used for its high efficiency and produces a total floating-point operations count of roughly $np^2 + 7np + 2n$ [67] (here we ignore the n square root operations needed for the factorization given that the value of p is sufficiently large) where n is the dimension of the matrix and p is the half-bandwidth. For the FE/BE method, $n = \frac{\pi}{4}N^2$ and $p = \pi N$. This produces a total flop count of $\frac{\pi^3}{4}N^4 + \frac{7\pi^2}{4}N^3 + \frac{\pi}{2}N^2$.

For the 2D FDTD method, on the other hand, the total flop number can be represented by a two-term expression as

$$F = F_{steady}F_{iter} \tag{3.90}$$

where F_{steady} is the number of time steps to reach steady state, and F_{iter} is the number of operations within a single time step.

F_{iter} can be easily computed by counting the algebraic operations in (3.74) which is roughly $28(N+2N_{PML})^2$. Assuming the mesh resolution is given by R, the CFL number is given by $CFLN$, the wave speed in the background medium is c_{bk} and the maximum wave speed among all inhomogeneities is c_{max}, the number of time steps required to

reach steady state can be estimated by

$$\begin{aligned} F_{steady} &= \frac{t_{steady}}{\Delta t} \\ &= \frac{\frac{2N\times\Delta}{c_{bk}}}{\frac{CFLN\times\Delta}{(\sqrt{2}c_{max})}} \\ &= \frac{2\sqrt{2}N\times c_{max}}{CFLN\times c_{bk}} \end{aligned} \tag{3.91}$$

and the total flop number for the FDTD method can be written as

$$F_{FDTD} = 56\sqrt{2}N(N+2N_{PML})^2\frac{c_{max}}{CFLN\times c_{bk}} \tag{3.92}$$

If $CFLN = 0.999 \approx 1$, equation (3.92) becomes

$$F_{FDTD} = 56\sqrt{2}N(N+2N_{PML})^2\frac{c_{max}}{c_{bk}} \tag{3.93}$$

where $\frac{c_{max}}{c_{bk}}$ is related to the contrast of the object to the background: if the object has higher dielectric values than the background then $\frac{c_{max}}{c_{bk}} = 1$; if the object has lower dielectric values than the background then $\frac{c_{max}}{c_{bk}} \approx \sqrt{\frac{\epsilon_{bk}}{\epsilon_{min}}}$ is the square root of the permittivity contrast.

Assuming the permittivity contrast in a sample benchmark problem is $\frac{\epsilon_{bk}}{\epsilon_{min}} = 10$ (which is reasonably high for typical microwave imaging cases) and $N_{PML} = 8$, the total flop number curve for the FDTD method is plotted as a function of mesh size (N) in Figure 3.16 along with that for the FE/BE method. From the plot, the advantage of using FDTD method is significant because the highest order term in F_{FDTD} is N^3 while that for $F_{FE/BE}$ is N^4. For most cases, the FDTD method (with GPML ABC) outperforms the hybrid scalar technique especially when mesh grows bigger. It should be noted that for smaller problems, the advantage is reduced. For the current FE/BE algorithm in the breast imaging reconstructions, the number of nodes on the circumference is 216 which

would make $N = 72$ for this analysis. For $N = 72$, $N_{PML} = 8$ and $\frac{\epsilon_{bk}}{\epsilon_{min}} = 10$, the flop count for the FE/BE method is only 54% higher than that for the FDTD technique.

Note that this comparison does not fully reflect the performance differences between the FE/BE and FDTD forward solvers in the actual image reconstructions because the FE/BE method only needs to decompose the LHS matrix once for multiple-source forward field calculations followed by multiple back-substitutions. The cost for the back-substitutions is relatively small. On the other hand, the flop counts for the FDTD method also vary because we used PML update equation (3.74) throughout the domain for simplicity. If we write separate loops for PML slabs and working volume, the total flop count can be reduced roughly by half. In addition, the FDTD for the cylindrical coordinate system is also helpful in enhancing its performance. The actual computation times for both methods can be found in the result section of this chapter.

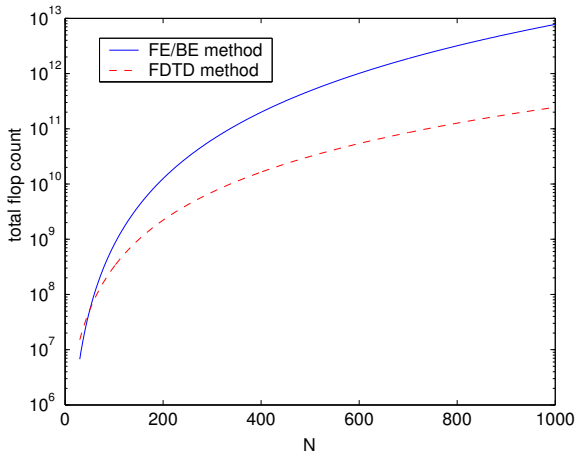


Figure 3.16: Comparison of the total floating-point operation numbers between FE/BE and FDTD methods for different mesh sizes.

Summary for the 2D FDTD forward method

We have investigated the implementation of a 2D FDTD method with GPML absorbing boundary condition for modelling the forward fields in a lossy medium. The up-

date scheme, source implementation, stability and numerical dispersion were discussed. With simple benchmark problems, the accuracy and efficiency of the proposed method were analyzed and compared to the FE method to demonstrate the advantages.

3.6.2 2D FDTD forward method coupled with 2D reconstruction

Similar to the FE/BE hybrid method, the 2D FDTD method provides the frequency domain forward field response for each transmitter and the measurement data are extracted at receivers to facilitate the reconstruction of the dielectric properties. Considering the tomographic configuration of the measurement system, the dual-mesh settings for the 2D FDTD field solution driving 2D reconstruction algorithm (2Ds$^{\text{FDTD}}$/2D) is shown in Figure 3.17 from which the rectangular uniform FDTD grid, the circular antenna array and the concentric reconstruction mesh can be identified. Due to the superior performance of the GPML medium, the source antennas can be placed very close to the boundary (as close as 5 cells in general).

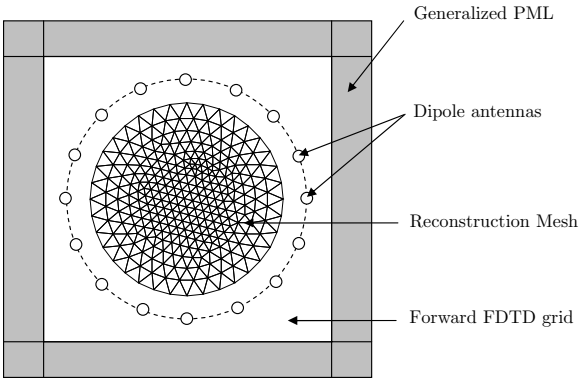


Figure 3.17: Dual-mesh configuration for 2D FDTD (GPML) forward solver and Gauss-Newton reconstruction.

The overall structure of the 2Ds$^{\text{FDTD}}$/2D reconstruction is simply an extension of the general diagram Figure 3.1 by replacing the forward solver by the 2D FDTD method

formulated above where the steps for forward field evaluation are illustrated in Figure 3.18. Besides the alteration in the forward field evaluation, the method for building the Jacobian matrix is also improved significantly. The new algorithm is referred as the nodal adjoint method which will be discussed in details in Section 5.1.2 in the context of the 3D vector reconstruction.

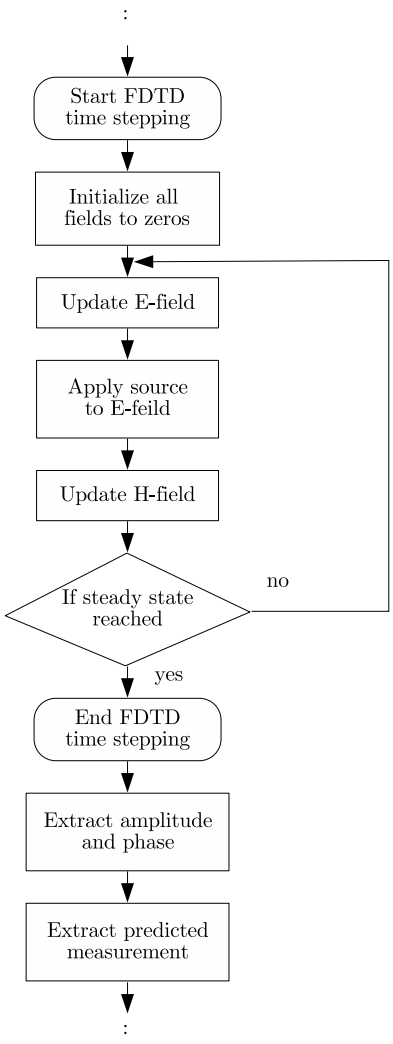


Figure 3.18: Flow chart of the forward field evaluation in 2Ds^{FDTD}/2D reconstructions.

3.7 Results

In this section, we present several simulation, phantom and *in vivo* animal and patient data reconstructions to demonstrate the effectiveness of the 2D algorithms. These reconstructions are primarily computed by the $2Ds^{FDTD}/2D$ while solutions from the 2Ds/2D reconstructions are provided for comparison. The dual-mesh for the $2Ds^{FDTD}/2D$ reconstruction is shown in Figure 3.17. The diameter of the antenna array is 15.1 cm which contains 16 monopole antennas modelled as time-harmonic point sources. The forward FDTD mesh is larger than the antenna array and typically 12 layers of GPML medium are incorporated to surround the forward domain to absorb outgoing waves. The stretching coordinate coefficients of the GPML use values provided in Table 3.4. The background medium consisted of an 83% glycerin solution ($\epsilon_r = 25.7, \sigma = 0.87$ S/m). The reconstruction mesh was placed concentrically within the antenna array with diameter $d = 15.2$ cm. The circular reconstruction mesh ($r = 6.5$ cm) is comprised of 473 nodes with 872 triangular elements. Linear Lagrangian basis function were used to model the distribution of the inhomogeneity inside the mesh. All reconstructions were initialized from the homogeneous background medium and were allowed to run for 20 iterations. The Tikhonov regularization with the empirical method for selecting regularization parameter, λ, was used in conjunction with the spatial filter technique (Section 3.2.3) (the smoothing coefficient α is set to 0.1 for all simulation reconstructions and 0.3 for phantom, animal and patient data reconstructions).

In all simulations, the measurement data was obtained by using a high-resolution FDTD grid of $R = 40$ with an operating frequency of 900 MHz. Two types of errors were computed to assess the convergence of the reconstruction. The relative error

(normalized misfit error) at each iteration i is defined by

$$e(i) = \frac{\|\mathbf{E}_z^{(i)} - \mathbf{E}_z^{\text{meas}}\|_2^2}{\|\mathbf{E}_z^{(1)} - \mathbf{E}_z^{\text{meas}}\|_2^2} \tag{3.94}$$

This relative error measures the appropriateness of the solution without requiring additional information. Since the true values of the dielectric property distribution were known in the simulations, the normalized RMS error between the reconstructed properties and their true values can be defined by

$$e_{RMS}(i) = \frac{\|\boldsymbol{\theta}^{(i)} - \boldsymbol{\theta}_{\text{true}}\|_2^2}{\|\boldsymbol{\theta}_{bk} - \boldsymbol{\theta}_{\text{true}}\|_2^2} \tag{3.95}$$

where $\boldsymbol{\theta}$ represents either the vector of the unknown permittivity or conductivity. $\boldsymbol{\theta}_{\text{true}}$ is the true value of the selected property and $\boldsymbol{\theta}_{bk}$ is the vector of the background properties which is the initial guess for all the reconstructions.

3.7.1 Simulations

The first simulation experiment utilized a centered cylindrical target with an offset inclusion. The cylinder had permittivity $\epsilon_r = 9$ and conductivity $\sigma = 0.3$ S/m with a diameter $d = 8$ cm to simulate a fatty breast. The 1.8 cm diameter cylindrical inclusion had $\epsilon_r = 35$ and $\sigma = 1.2$ S/m to simulate fibroglandular tissue and was located at ($x = -2$ cm,$y = -2$ cm). The 83% glycerin solution was used as the background medium. The simulated measurement data was utilized by the dual-mesh 2D reconstruction algorithm along with a homogeneous background initial guess. The reconstructed permittivity and conductivity images for the 20-th iteration is shown in Figure 3.19. A plot of the normalized misfit error vs. the iteration number is shown in Figure 3.20. After 6 iterations, the boundaries and properties of the object and inclusion are

readily apparent and the normalized misfit error decreased to 0.07.

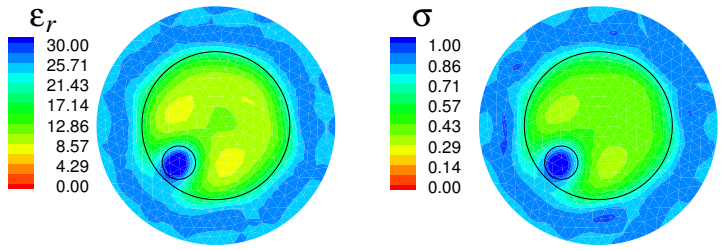
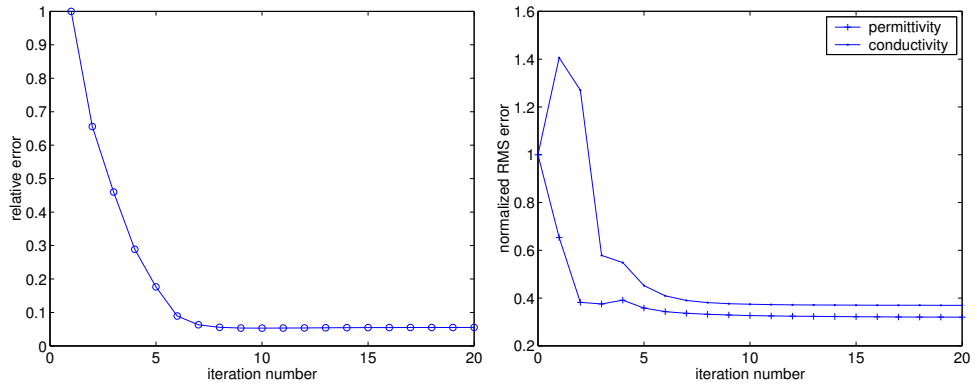


Figure 3.19: Reconstructed dielectric profiles of the breast-like object after 20 iterations (a) relative permittivity, (b) conductivity.



(a) (b)

Figure 3.20: (a) Relative error and (b) normalized RMS error plots for the breast-like object reconstruction.

The second simulation was designed to investigate the image resolution of the algorithm. Six objects with various sizes were placed inside the imaging zone. Their dimensions and dielectric values are noted in Table 3.5. Once again, the reconstruction started from the homogeneous background values and the images at the 20-th iteration are shown in Figure 3.21. The associated relative field and RMS errors curves are plotted in Figure 3.22.

Table 3.5: Object properties for simulation with varied sizes.

	shape	center (cm)	size * (cm)	ϵ_r	σ (S/m)
object 1	square	(-2,2)	4	10	0.3
object 2	square	(2,-2)	1	10	0.3
object 3	circle	(-3,-3)	2	10	0.3
object 4	circle	(-1,-1)	1.6	10	0.3
object 5	circle	(1,1)	1.2	10	0.3
object 6	circle	(3,3)	0.8	10	0.3

* edge lengths for square objects and diameters for circular objects

In the reconstructed images, all six objects were successfully recovered with the correct positions and sizes. Artifacts are noticeable within the background medium. However, the variations due to these artifacts are still less than the recovered properties of the smallest object, which is 0.8 cm in diameter. This demonstrates that the reconstruction algorithm is able to recover objects smaller than the half wavelength resolution limit (3.3 cm in this case) in a complex setting with the iterative algorithm. It should also be pointed out that due to the smoothing effects of regularization and spatial filter technique, the recovered object values appear to be higher than their true values. This is more noticeable for the smaller objects. It is interesting to note that the relative field error decreases monotonically and seems to have achieved steady-state by 7-th iteration. However, it appears that the permittivity values reach their minimum error very quickly with the conductivity error catching up quite slowly indicating a possible bias towards the permittivity recovery.

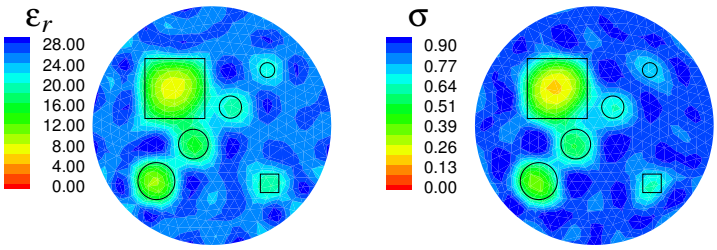
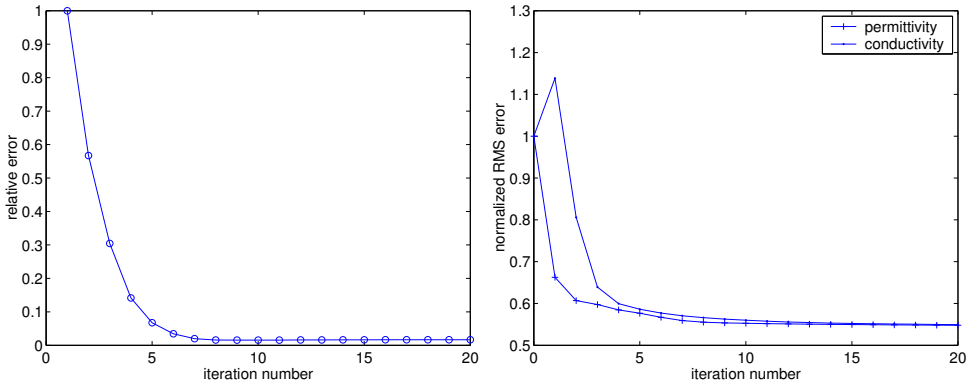


Figure 3.21: Reconstructed dielectric profiles of the size simulation.

Table 3.6: Object properties for simulation with varied contrasts.

	center (cm)	ϵ_r	σ (S/m)
object 1	(-3,-3)	10	0.1
object 2	(-1,-1)	15	0.2
object 3	(1,1)	20	0.4
object 4	(3,3)	25	0.6
object 5	(-3,3)	30	1.0
object 6	(-1,1)	40	1.2
object 7	(1,-1)	50	1.4
object 8	(3,-3)	60	1.6



(a) (b)

Figure 3.22: (a) Relative error and (b) normalized RMS error histograms for the reconstruction of objects with varied sizes.

The third simulation investigated object contrast. Eight circular objects with identical diameter $r = 1.6$ cm filled with different materials were studied with their values summarized in Table 3.6. Using the same strategy as applied in the previous reconstructions, the recovered dielectric profiles are shown in Figure 3.23 after 20 iterations. The relative and RMS error curves for this simulation are shown in Figure 3.24. From the reconstructed images, it is difficult to distinguish the top two objects from the background because they have the smallest contrast. The remainder are all readily detected. Again, it is interesting that the permittivity RMS error curve converges faster than its conductivity counterpart further suggesting some bias.

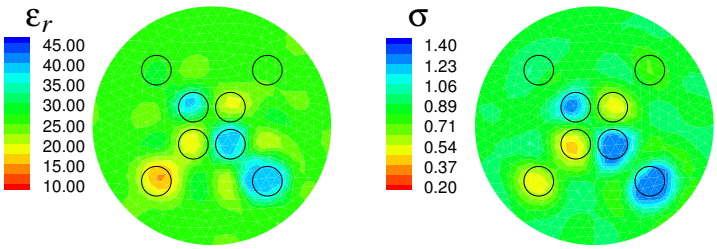
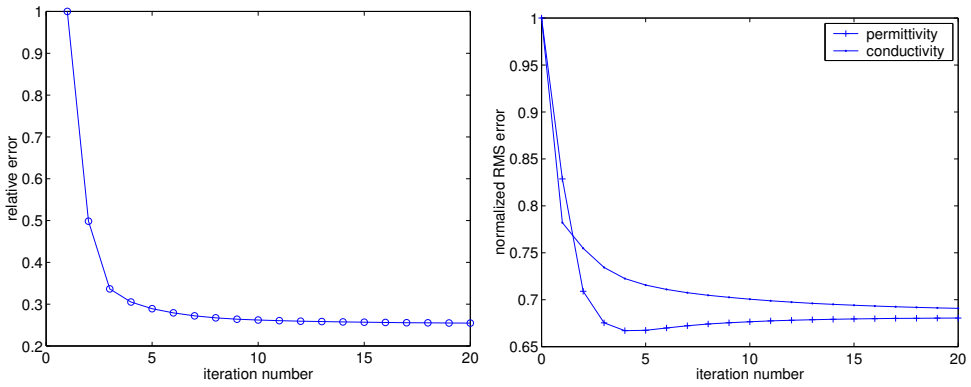


Figure 3.23: Reconstructed dielectric profiles for the contrast simulation.



(a) (b)

Figure 3.24: (a) Relative error and (b) normalized RMS error histograms for the reconstruction of objects with varied contrasts.

The fourth simulation was designed to assess the cross-talk between the permittivity and the conductivity during the reconstruction process. Following the settings from the second simulation, six objects were divided into two groups with the first group appearing in the permittivity and the second group appearing in the conductivity image. After the 20-th iteration, the permittivity and conductivity images are successfully reconstructed with negligible crosstalk between images as shown in Figure 3.25. The associated relative and RMS error curves are plotted in Figure 3.26. Comparing the RMS error curve with the one in Figure 3.22, the error for the conductivity was reduced significantly. This indicates that the convergence of the algorithm might also be effected by the complexities of the true distribution.

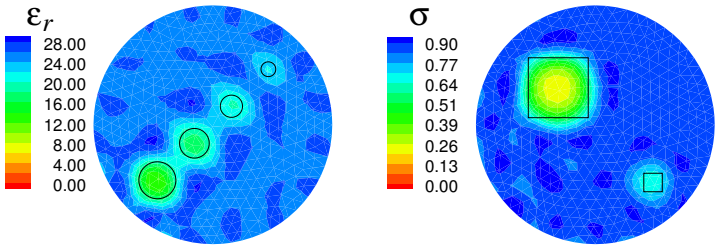
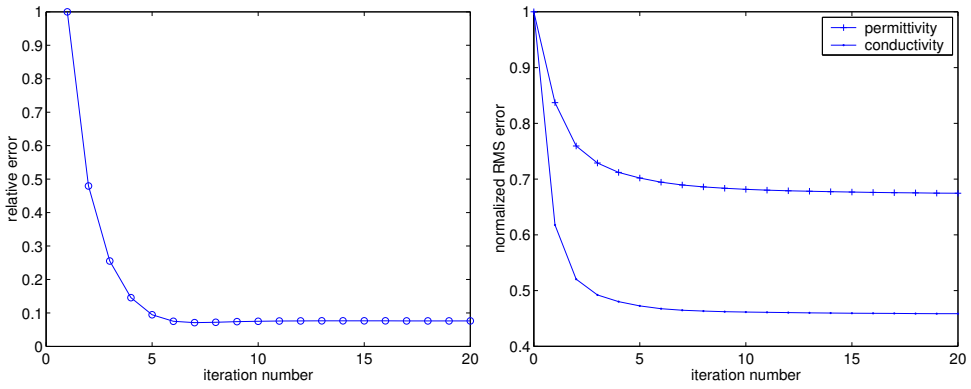


Figure 3.25: Reconstructed dielectric profiles for assessing the cross-talk between permittivity and conductivity.

In all simulations, the 2D FDTD forward/2D reconstruction algorithm correctly recovered the locations and property values of the targets except for minor distortions and smoothing due to the nature of the algorithm and limited amount of measurement data. The reconstruction time for a single iteration consisted of roughly 4 seconds (R = 20) for forward solution modelling and 5 seconds for computing the update. A parallel version of this reconstruction code was also implemented which reduced the forward solution modelling time by a factor of 3.5 when 4 CPU's were used simultaneously.



(a) (b)

Figure 3.26: (a) Relative error and (b) normalized RMS error histograms for the cross-talk simulation.

3.7.2 Phantom experiments

Two groups of phantoms were studied with the 2Ds^{FDTD}/2D reconstruction algorithm to demonstrate the viability of this method in realistic imaging settings. The first study involved two solid cylindrical phantoms which both had low dielectric properties (ϵ_r = 5 and σ = 0.3 S/m) to simulate bone/fat tissue while the background medium was an 83% glycerin solution (ϵ_r = 22.8, σ = 0.89 S/m). Measurements were acquired at a number of frequencies which only the 1100 MHz measurement data was used for reconstructing the dielectric profiles. The dielectric images after 20 iterations were shown in Figure 3.27. The actual positions and sizes of the targets are drawn over the recovered object in the reconstructed images. Figure 3.28 shows the corresponding relative error curve.

The locations, sizes and shapes of the two cylinder phantoms were correctly reconstructed in the permittivity image. The location of the small cylinders on the conductivity image is slightly shifted toward the boundary. The distortion and artifacts on both images can probably be partially explained by the mismatch between the numerical for-

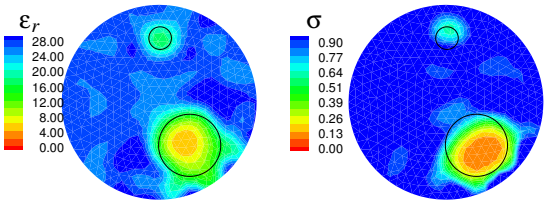


Figure 3.27: Reconstructed bone/fat phantom dielectric images.

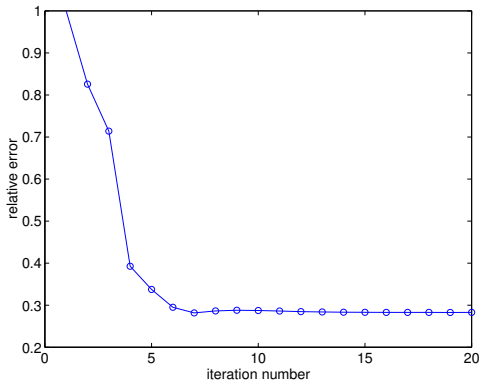


Figure 3.28: Relative error histogram for the two-cylinder phantom reconstruction.

ward model and actual illumination configuration (In terms of mismatch, this includes the fact that the actual fields scatter into 3D space while those in the model only scatter into 2D space).

In the second experiment, we utilized a 10 cm liquid cylinder to simulate a breast and two cylindrical inclusions to mimic glandular and tumor tissue, respectively. A photograph of this experiment is shown in Figure 3.29. The inclusion for mimicking glandular tissue was 1.8 cm in diameter with dielectric properties of $\epsilon_r = 32.3$ and $\sigma = 1.3$ S/m while those for the 2.1 cm diameter tumor phantom had $\epsilon_r = 53.5/\sigma = 1.12$ S/m. The large phantom was filled with different glycerin solutions such that the properties were close to those for a range of breast density classifications: extremely dense ($\epsilon_r = 22.2, \sigma = 1.06$ S/m), heterogeneously dense ($\epsilon_r = 18.3, \sigma = 0.95$ S/m), scattered ($\epsilon_r = 14.4, \sigma = 0.77$ S/m) and fatty ($\epsilon_r = 9.5, \sigma = 0.44$ S/m), respectively.

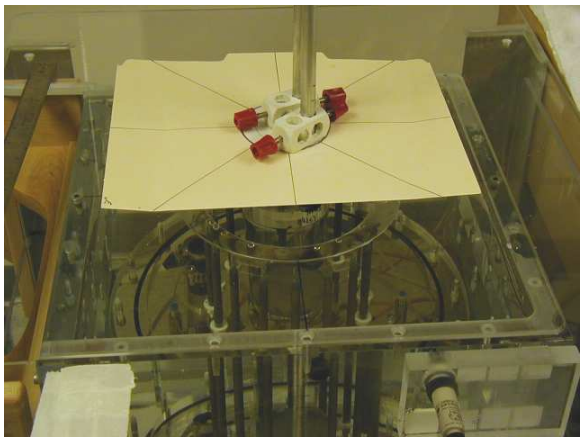


Figure 3.29: Photograph of the phantom experiment with various contrasts. The illumination tank, antenna array, large cylindrical object and two tubes for inclusions are shown.

For all four cases, images of the phantoms were reconstructed at 900 MHz by the 2D algorithm and are shown in Figure 3.30 (a)-(d). For comparison, we also reconstructed the images (Figure 3.31) using the 2D scalar technique described in Section 3.4.

The reconstructed images accurately captured the locations and values of the object and inclusions. The recovery of the tumor becomes more accurate with a decrease in contrast between the breast and background. For the extremely dense case, the outline of the breast is not visible because the contrast between the background and phantom are negligible in this case. These images are generally consistant with the results reconstructed from the 2Ds/2D technique. Certain amount of artifacts can be observed around the image boundaries which may be related to signal noise and forward model mismatch issues.

3.7.3 *In vivo* animal measurement reconstructions

A series of *in vivo* animal experiments utilizing piglets were conducted to demonstrate the thermal monitoring capability of our microwave imaging system. These experi-

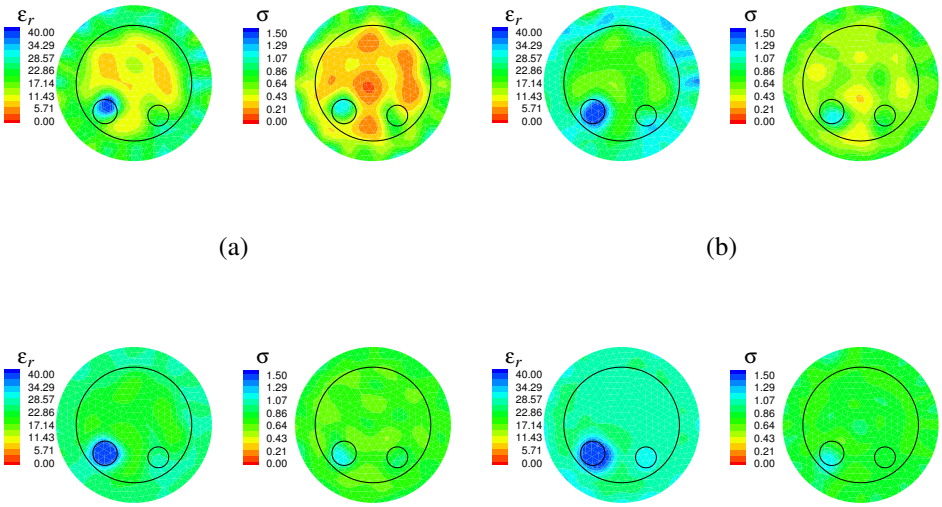


Figure 3.30: Reconstructed dielectric profiles for four 10 cm diameter phantoms with a range of contrasts mimicking (a) fatty, (b) scattered, (c) heterogeneously dense and (d) extremely dense breasts, respectively. In each there is a 2.1cm diameter inclusion to the lower left simulating a tumor and a second 1.8 cm diameter inclusion to the lower right simulating glandular tissue.

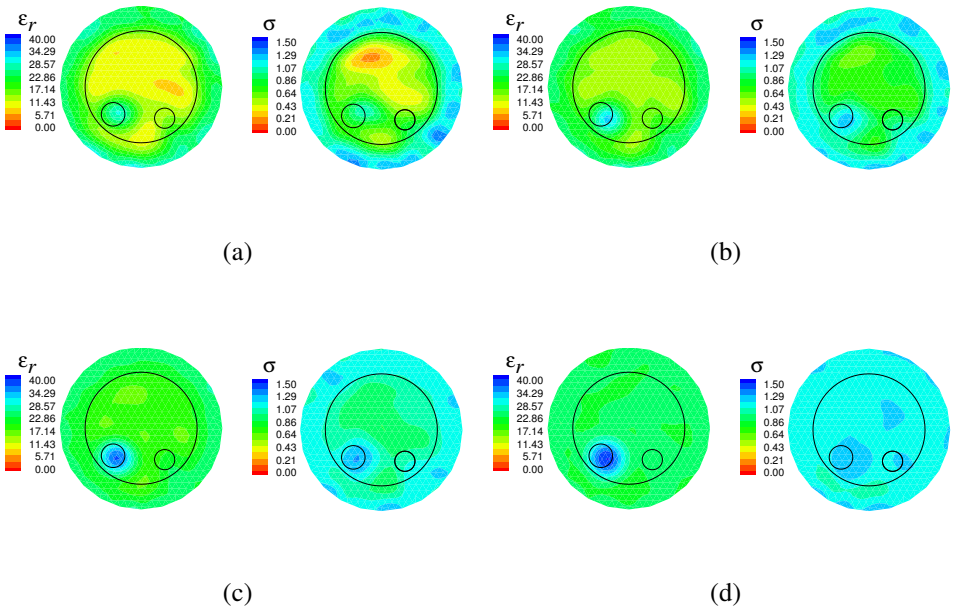


Figure 3.31: Reconstructed dielectric profiles for four 10 cm diameter phantoms with a range of contrasts mimicking (a) fatty, (b) scattered, (c) heterogeneously dense and (d) extremely dense breasts, respectively, with the FE/BE forward field technique.

ments were performed with the first generation system (Chapter 1) during years 2001 and 2002. In the experiment, a plastic tube was surgically inserted through the abdomen of a living piglet. The piglet as well as the tubings was submerged into the illumination tank and placed within the center of the circular antenna array. During the experiment, the tube was supplied with saline at different temperatures and the measurement of the scattered microwave field was collected accordingly. CT scans were performed before and after the data acquisition to assess the position shift during the experiment. A photograph of the experimental set up is shown in Figure 3.32. A sample CT image is shown in Figure 3.33 to demonstrate the structure of the target and its relative position to the antenna array.

The temperature of the saline inside the tube was raised from 33°C to 45°C in 3°C increments. Afterwards, the temperature was decreased back to 33°C with the same

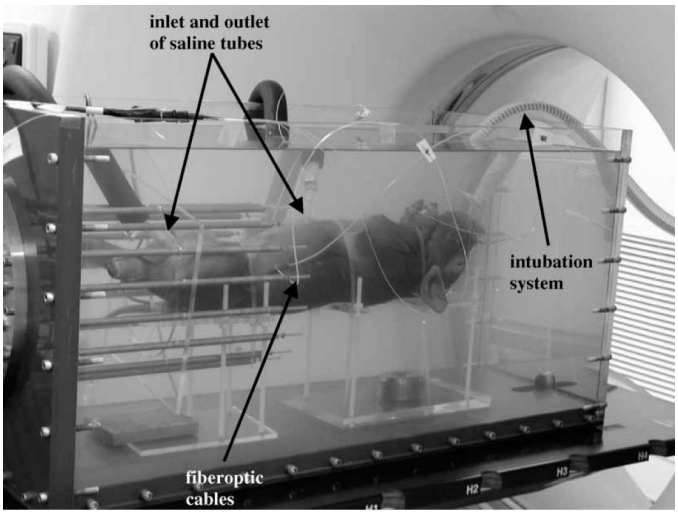


Figure 3.32: Photograph of the system settings in the piglet experiment [123].

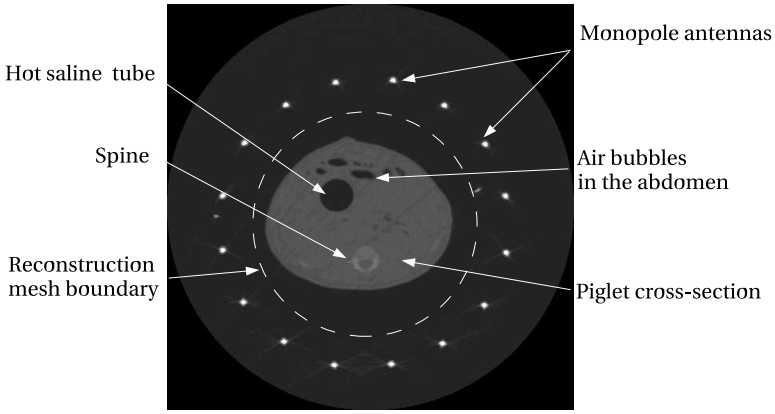


Figure 3.33: Axial view CT image of the piglet abdomen as well as the microwave antenna array.

step size. Dielectric property profiles at different temperature points were reconstructed with our 2Ds^{FDTD}/2D algorithm and are shown in Figure 3.34 (note that the temperature points during the rising phase have a "u" suffix, while those in the decreasing phase have a "d" suffix). The images were also reconstructed for the case with an empty tube in the room temperature (denoted by "air" in the figure). To assess the correlation between the temperature and the dielectric properties, we used the images for the case "33u" as the reference which was subtracted from the rest of the images. The difference images in the conductivity part of the dielectric properties are shown in Figure 3.35.

A few observations can be made from the absolute images (Figure 3.34): 1) the contours of the pig abdomen in these images agrees well with those in the CT images (Figure 3.33), 2) the saline tube can be observed in both the permittivity and conductivity images and agrees with the location in the CT image, 3) low dielectric value zones can be seen at the top and the center bottom of the pig abdomen in the permittivity images where the former agrees with the locations of the fatty tissue and air bubbles inside the abdomen and the latter is consistant with the location of the pig spine, 3) the location of the tube is highlighted in the "air" case. From the conductivity difference images, the linear relationship between the conductivity and the temperature can be readily seen (the conductivity difference in the tube area reaches the maximum when the temperature difference is maximum). This experiment not only demonstrates the capability of imaging the anatomy of a living animal, but also provides rationale for utilizing microwave imaging as a means of thermal monitoring. More details about this experiment can be found in [123].

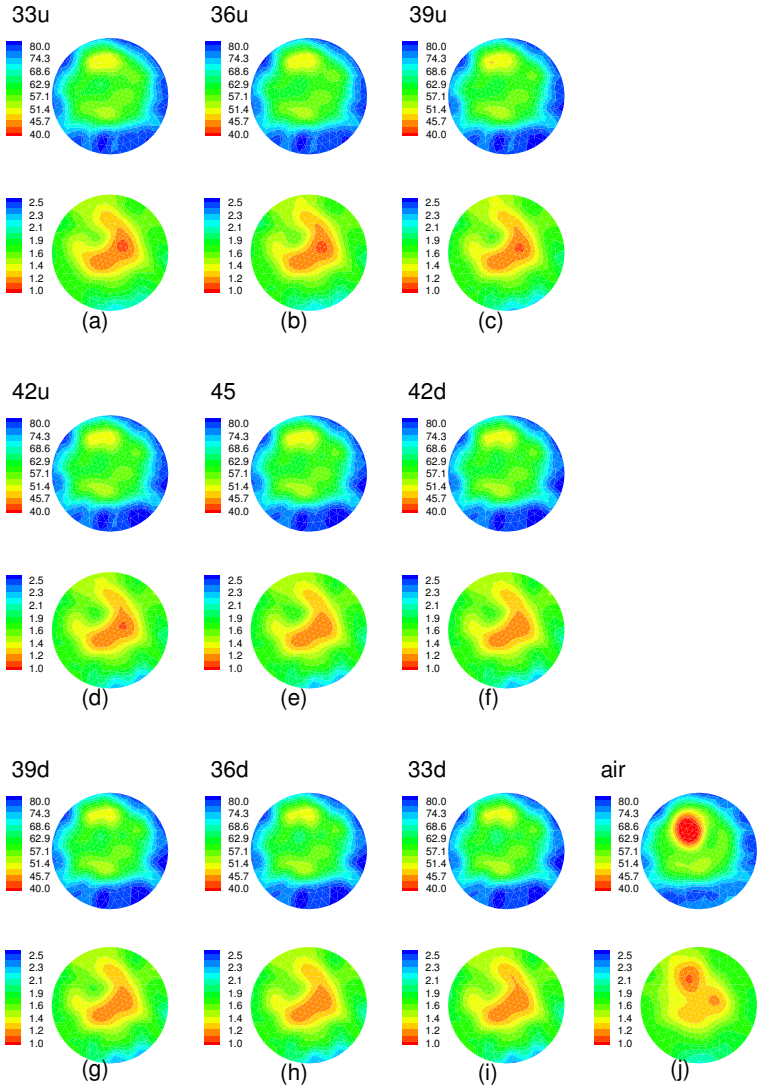


Figure 3.34: Reconstructed images at different tube saline temperatures during the rising phase: (a) 33°C, (b) 36°C, (c) 39°C, (d) 42°C, (e) 45°C, and the decreasing phase: (f) 42°C, (g) 39°C, (h) 36°C, (i) 33°C, (j) room temperature (tube was filled with air).

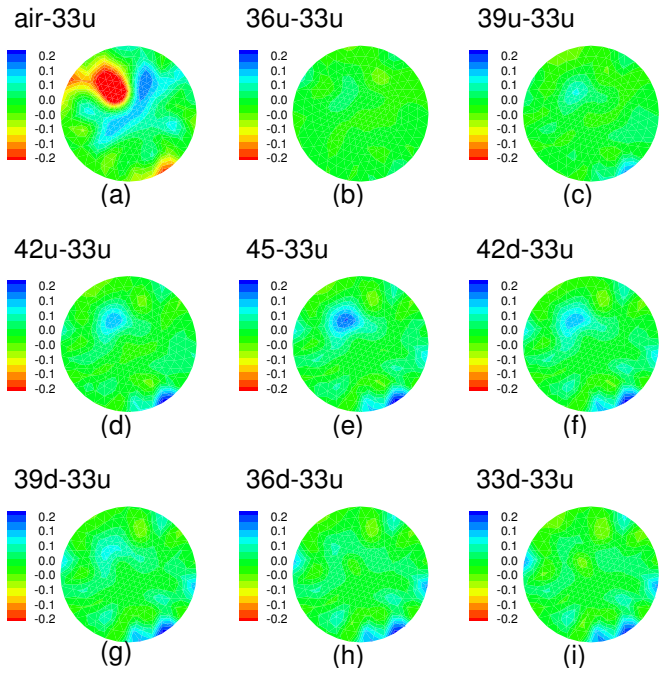


Figure 3.35: Difference images of the conductivity between the "33u" case and (a) air, (b) 36u, (c) 39u, (d) 42u, (e) 45, (f) 42d, (g) 39d, (h) 36d, (i) 33d.

3.7.4 Patient data reconstructions

With our clinical imaging system, we have performed microwave breast exams for over 200 women. In these patient exams, the planar antenna array collected measurement data at 7 vertical positions with either a 0.5 cm or 1 cm vertical separation where position 1 corresponded to the plane closest to the patient chestwall and position 7 refers to the plane closest to the nipple. The imaging plane is parallel to the anatomical coronal plane of the patient. This process is performed for both breasts. The whole exam typically takes roughly 16 minutes. We selected a patient with a known tumor to demonstrate the performance of the 2D algorithm.

The patient (patient ID: 1082) had two tumors in her left breast: the first was close to the chestwall, and the second was towards the nipple. The locations and sizes of these two tumors are indicated in Figure 3.36. The reconstructed images from the data

measured at 1300 MHz are shown in Figure 3.37.

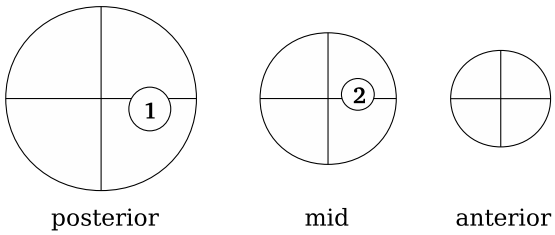
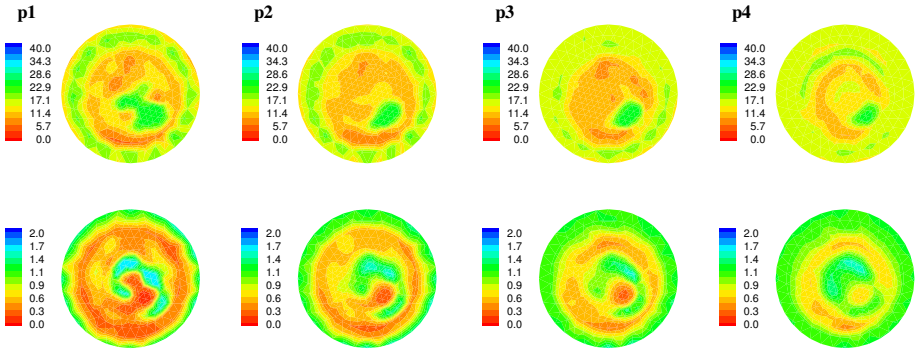


Figure 3.36: Locations of the tumors on the left breast of patient 1082. The largest circle corresponds to the posterior region (closest to chestwall), the smallest corresponds to the anterior zone and the center one refers to the mid breast zone.

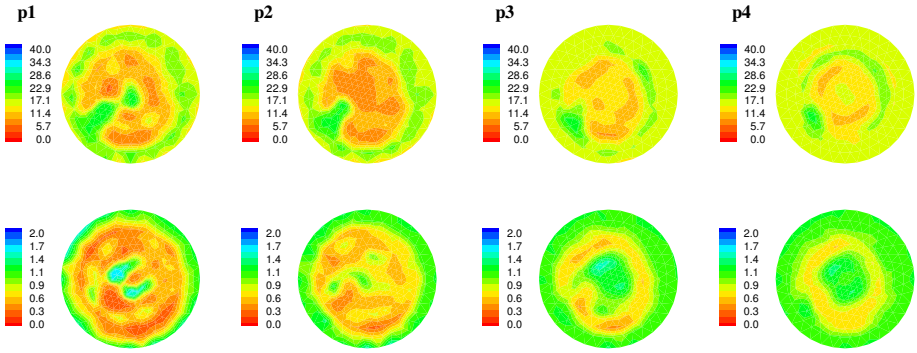
From these images, the contours of the breast at different slices are readily distinguishable. The glandular tissue on both sides of the breast can also be observed predominantly in the permittivity images of both breasts. On the "p1" slice of the left breast conductivity images, the enlarged glandular area with elevated properties compared with the contralateral breast indicates the existence of the tumor. The circular indentation (i.e. elevated property value) at 3 o'clock on the conductivity image (p1) agrees well with the clinical information.

Another tumor patient example is shown in Chapter 10 in the discussion of the image reconstructions of strong scattering objects such as a large tumor case utilizing unwrapped phase information.

From the recovered images, the proposed 2D algorithm demonstrated good correlation with the results of the previous algorithm in terms of accurate recovery of the phantom and inclusion shapes, locations and properties, respectively. The computational time for the FDTD-based reconstruction was roughly half as long as the FE/BE reconstruction while the FDTD grid size had roughly twice as many nodes as the FE/BE approach. The advantage of using the FDTD method as the forward modelling technique is significant for large N from these experiments in terms of computational effi-



(a)



(b)

Figure 3.37: Reconstructed image slices for patient 1082: (a) left breast, (b) right breast.

ciency.

3.8 Conclusions

A number of important issues have been discussed in this chapter. We first outlined the overall structure of an iterative image reconstruction scheme for tomographic microwave imaging. The concept of the dual-mesh was mathematically defined with respect to the image reconstruction. Two 2D reconstruction algorithms were proposed including the FE/BE scalar forward method driving the 2D reconstruction algorithm (2Ds/2D) and the 2D FDTD forward method driving the 2D reconstruction ($2Ds^{FDTD}/2D$). The forward field modelling in both 2D algorithms were discussed including topics such as their formulations, radiation or absorbing boundary conditions, accuracy and efficiency. From the comparison of these two algorithms, the 2D FDTD method demonstrated promising advantages in modelling the steady-state field in a lossy media compared with the FE/BE counterpart. Numerical simulations, phantom and *in vivo* experiment cases were presented to assess the accuracy and efficiency of these algorithms for image reconstruction in real-world settings. For all cases, the targets were successfully reconstructed with the $2Ds^{FDTD}/2D$ algorithm and the convergence was assessed in terms of the relative field error and the RMS error (for simulations).

In conclusion, we have demonstrated significant algorithmic flexibilities utilizing the dual-mesh based algorithm in terms of accommodating various forms of the forward field modelling method and the parameter representations in a single reconstruction. These flexibilities play an important role in the 3D image reconstruction algorithms discussed in the following two chapters. Due to the straightforward formulation,

highly efficient absorbing boundary condition and overall computational advantages, the 2D FDTD method is quite attractive for the forward field modelling role. The high efficiency of the 2D FDTD method in conjunction with state-of-the-art computer power make it possible for quasi-real-time image reconstructions in medical applications. In Chapter 5, this algorithm is expanded to 3D with various optimizations to further improve efficiency. These 3D optimizations can also be applied to the 2D algorithm.

Chapter 4

3D scalar field driving 2D reconstruction algorithm

In this chapter, an efficient Gauss-Newton iterative image reconstruction technique utilizing a three-dimensional field solution coupled to a two-dimensional parameter estimation scheme (denoted as 3Ds/2D) is presented. As an intermediate step towards 3D microwave imaging and a direct extension of the 2D algorithms discussed in Chapter 3, the 3Ds/2D algorithm employs a 3D scalar model for the forward field computation under reasonable approximations. This simplified model together with the supplementary iterative block linear equation solver demonstrates remarkable efficiency for modelling 3D wave propagation and limited reduction in accuracy compared with the full 3D vector field solutions. In addition, the resultant image recovery has been restricted to a 2D plane and the interconnections between the forward and reconstruction problem are once again organized with respect to the dual-mesh scheme discussed in Chapter 3.3. As demonstrated in the phantom reconstructions in Section 3.7.1, using 2D methods to reconstruct a slice of a 3D object can cause distortions in the image due to the mismatch in the forward model. With the implementation of the 3Ds/2D algorithm,

these 3D artifacts are perceivably reduced when compared with results from entirely two-dimensional inversions (2D/2D). Important advances in terms of improving algorithmic efficiency include the use of a block solver for computing the field solutions and application of the dual-mesh adjoint approach for assembling the Jacobian matrix. Results obtained from synthetic measurement data show that the new 3Ds/2D algorithm consistently outperforms its 2D/2D counterpart in terms of reducing the effective imaging slice thickness in both permittivity and conductivity images over a range of inclusion sizes and background medium contrasts.

The next section provides a summary of the underlying methodology. This is followed by the results (Section 4.2), which reports and analyzes images recovered with the algorithm for various phantom experiments that are representative of cases previously studied with the 2D algorithms. Metrics previously developed in the 2D studies are applied to the output of these 3D reconstructions to assess potential improvements. The results and innovations of the 3Ds/2D algorithm are summarized in the concluding section of the chapter (Section 4.3).

4.1 Theory and techniques

For biological tissues, the governing equation of the electric field propagation in the frequency domain is generally expressed as a vector equation, i.e. (3.13). To correctly formulate the forward equation, we need to reconsider the three basic assumptions used in the 2D algorithms, i.e. 2D dielectric property distribution, the line source and TM wave propagation (Section 3.4). First, the assumption of 2D dielectric profiles is still valid since the reconstruction in 3Ds/2D algorithm is still modelled by 2D spatial bases. Second, in 3D space, it is no longer appropriate to require the source to have infinite

length, instead, a finite length source or point source is desired. The associated electric field distribution of a finite-length source is a 3D vector field where the E_x and E_y components are not necessarily zero. This jeopardizes the validity of the second and the third assumptions above and creates difficulties in obtaining the forward solution since (3.13) cannot be decoupled into scalar equations along each spatial axis. Fortunately, the comparison between the vector form forward solution with the scalar solution indicates that the electric field in the far field of the antenna is approximately a TM wave where the $\vec{E}$ field in the $x-y$ plane is negligible compared with its z component. As a result, approximations are introduced by imposing the TM assumption in the 2D forward model. The impact of this approximation on the forward model and image reconstruction accuracy is discussed in the results section of this chapter.

As a result of above approximations, the frequency domain equation involving only the E_z component is once again applied similarly to the 2D case except that this equation now needs to be solved in the 3D space. A dual-mesh pair is constructed where the forward mesh is a cylindrical mesh in the 3D space and the reconstruction mesh is a planar circular mesh as depicted in Figure 4.1. The forward mesh extends radially beyond the circular array of driving antennas (oriented vertically). The top and bottom of the cylindrical mesh are well away from the antenna cross-sectional plane to minimize interactions with finite boundaries resulting from mesh truncation (Figure 4.1). Radiation boundary conditions (RBC's) are applied to the exterior surfaces of the forward mesh to accurately represent unbounded field propagation while facilitating truncation of the problem to an acceptable size (Section 4.1.1).

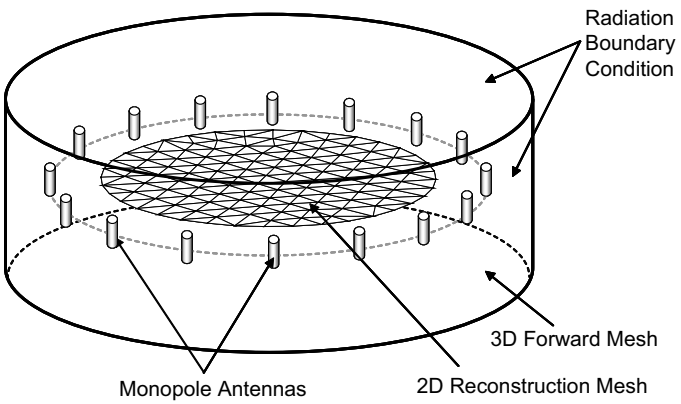


Figure 4.1: Schematic of the 3Ds/2D imaging problem – the 2D reconstruction area is centered within an array of 16 monopole antennas with the 3D cylindrical volume extending radially beyond the antennas and a substantial distance above and below the 2D imaging plane.

4.1.1 Forward problem

3D scalar model

The three-dimensional finite element technique was chosen to solve the frequency domain equation (3.13) due to its appealing features such as the sparseness of the associated matrix system and the capability of modelling curved boundaries. In the 3Ds/2D algorithm, the forward domain Ω is a 3D volume and the basis/weighting functions are defined over 3D elements instead of planar triangles. Utilizing linear basis functions, the weak form system of equations can be constructed with the implementation of Galerkin method, similarly to the 2D scalar cases, to produce

$$\begin{aligned} \langle \nabla E_z(\vec{r}), \nabla \phi_\ell \rangle - \left\langle k^2 E_z(\vec{r}), \phi_\ell \right\rangle - \oint_{\partial\Omega} \nabla \phi_\ell E_z(\vec{r}) \cdot \hat{n} ds \\ = -jw\mu_0 \langle J_z(\vec{r}), \phi_\ell \rangle \end{aligned} \tag{4.1}$$

where $\langle \cdot, \cdot \rangle$ is the integration of the product of the two terms over the entire forward volume, $\partial\Omega$ is the surface of that volume, ϕ_ℓ is the weighting function, $\hat{n}$ is the unit

normal vector to the volume surface. By satisfying (4.1) for the weighting functions associated with all nodes in the 3D FE mesh, a system of N equations with N unknowns (the electric field values at all nodes) can be constructed and organized in matrix form

$$A\mathbf{E} = \mathbf{b} \tag{4.2}$$

where the (i,ℓ)-th element of A is

$$a_{i,\ell} = \langle \nabla\phi_i, \nabla\phi_\ell \rangle - \left\langle k^2\phi_i, \phi_\ell \right\rangle \tag{4.3}$$

and the ℓ-th element of $\mathbf{b}$ is

$$b_\ell = -j\omega\mu_0 \left\langle J_z\left(\vec{r}\right), \phi_\ell \right\rangle \tag{4.4}$$

Note that the contribution from the surface integral in (4.1) is discussed in the following section. It is also important to note that A contains all of the information pertaining to the electrical property distribution within the modelled zone while $\mathbf{b}$ contains all of the source antenna data.

Radiation boundary conditions

The modelled domain must be truncated at a finite distance to limit the problem size. One possible choice was to use the 3D extension of the boundary element method similarly to the 2D algorithm in the last chapter. A surface integration would need to be evaluated and the BE matrix equation constructed similar to (3.49). Fortunately, due to the highly lossy nature of the coupling medium, an approximate radiation boundary condition, the first order Bayliss-Turkel RBC [6, 5, 162, 92], is implemented introduc-

ing only moderate errors while significantly simplifying the evaluation. This RBC is derived from the Sommerfeld radiation condition [182] described by

$$\lim_{r\to\infty} r\left(\frac{\partial E_z(r)}{\partial r} - jkE_z(r)\right) = 0 \tag{4.5}$$

evaluated on an infinitely large 3D spherical surface S_r with radius r where the sources and scatterers are assumed to be at the center of the sphere. The approximated asymptotic expansion of the scalar function E_z in spherical coordinates on S_r is given by [6, 92]

$$E_z(r) = \frac{\exp(jkr)}{\sqrt{k\rho}} \sum_{i=0}^{\infty} f_i(\theta)\,(kr)^{-i} \tag{4.6}$$

where $f_i(\theta)$ is the angular projection. The gradient of E_z on the surface S_r can be approximated by the i-th order Bayliss-Turkel RBC [6, 5, 92] expressed as

$$\nabla E_z(r) = \left(\alpha_i(r)E_z(r) + \beta_i(r)\frac{\partial E_z(r)}{\partial^2\theta}\right)\hat{r} \tag{4.7}$$

where $\alpha_i(r)$ and $\beta(r)$ are coefficients defined by

$$\begin{aligned} \alpha_1(r) &= jk - \tfrac{1}{2r} \\ \beta_1(r) &= 0 \end{aligned} \tag{4.8}$$

for the first order approximation ($i = 1$) and by

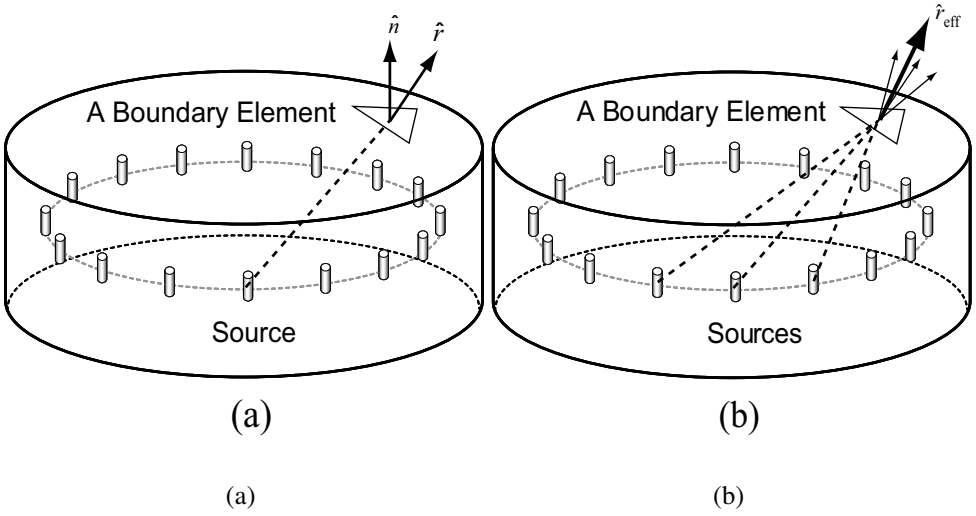
$$\begin{aligned} \alpha_2(r) &= \frac{jk - \frac{3}{2r} + \frac{3j}{8kr^2}}{1+\frac{j}{kr}} \\ \beta_2(r) &= \frac{\frac{j}{2kr^2}}{1+\frac{j}{kr}} \end{aligned} \tag{4.9}$$

for the second order approximation ($i = 2$).

Choosing the first order Bayliss-Turkel RBC and substituting (4.8) into (4.1), we eliminate the unknown gradient on the boundary and produce a continuous equation of

the unknown field which can be easily discretized and solved with linear matrix solvers. Since the forward mesh boundary $\partial\Omega$ is not a sphere, the surface integration involves the vector product of radial vector $\hat{r}$ and the normal vector of the surface patch $\hat{n}$ as

$$\langle \nabla E_z(\vec{r}), \nabla \phi_\ell \rangle - \left\langle k^2 E_z(\vec{r}), \phi_\ell \right\rangle - \oint_{\partial\Omega} \left(jkE_z(\vec{r}) - \frac{E_z(\vec{r})}{2r} \right) \phi_\ell \hat{r} \cdot \hat{n} ds = -jw\mu_0 \langle J_z(\vec{r}), \phi_\ell \rangle \tag{4.10}$$



(a) (b)

Figure 4.2: Schematic of a) the vector from a portion of a single line source to a boundary element on the cylindrical volume, and b) the vectors from multiple antennas simultaneously projected to produce an effective $\hat{r}$ vector at the boundary element.

Because the antennas are finite length, the direction of $\hat{r}$ at $\partial\Omega$ changes as a function of the section of the antenna that is referenced. To account for this nonlinear variation in the direction of $\hat{r}$, its direction is integrated along the antenna length to produce an effective $\hat{r}$.

In our geometric configuration, the sources are not at the center of the volume. This produces an immediate impact on the surface integral term in (4.10). If $\hat{r}$ is taken as the

unit vector from each individual source corresponding to $J_z(\vec{r})$ on the right-hand-side of (4.10), matrix A will vary for each transmitter. This has significant computational consequences in that (4.2) will then have to be solved independently for each source. However, an alternative approach is to construct the $\hat{r} \cdot d\hat{s}$ product utilizing the weighted sum of $\hat{r}$'s from all source antennas in the array (even though only one is active at a time) (Figure 4.2 (b)). In this way, the contribution from the surface integral in (4.10) becomes independent of the active antenna, making A identical for all sources. Table 4.1 shows a summary of the computed forward electric field magnitude and phase differences at 500 and 900 MHz for a single transmitter averaged over the associated 15 receiver antennas for the two different BC configurations along with the analytical solution. The row for case A of the table contains the differences in a homogeneous saline solution between the numerical case with $\hat{r}$ referenced just to the transmitter (ANT1) and the analytical solution The row for case B lists the differences between the solutions computed by the average of the 16 antennas (ANT16) and the analytical solution. The rows for case C list the differences between the numerical models utilizing ANT1 and ANT16 BC's for the homogeneous and inhomogeneous cases. In general, the differences between the ANT1 and ANT16 cases were less than the differences between the ANT1 and the exact analytical solutions except for the phase at 500 MHz. For both RBC's, the differences in both magnitude and phase compared with the analytical solution are generally acceptable considering the mesh used in this comparison is relatively coarse (about 8 nodes per wavelength at 900 MHz). Essentially this shows that any errors introduced by moving from the exact analytical solution to that of the ANT1 BC are more significant than that introduced by replacing the ANT1 BC's with the ANT16 BC's. To verify this in a more complicated situation, we computed the same ANT1 and ANT16 differences for the case where there was a high contrast, breast-like object

Table 4.1: Average 500 and 900 MHz forward solution magnitude (dB) and phase (degrees) differences for signals computed at the 15 associated receivers for a single transmitter.

A. ANT1-Analytical B. ANT16-Analytical C. ANT1-ANT16		**500M**		**900M**	
		Mag (dB)	Phase (deg)	Mag (dB)	Phase (deg)
Homogeneous	A.	0.56	2.4	1.1	25.7
	B.	0.098	3.7	1.3	21.5
	C.	0.52	5.5	0.28	4.2
Breast-Like Target	C.	0.46	5.4	0.28	4.9

1. Homogeneous saline solution:
   A. 1 antenna BC's versus the analytical solution, and
   B. the averaged 16 antenna BC's versus the analytical solution, and
   C. 1 antenna BC's versus the averaged 16 antenna BC's;
2. Breast target in saline solution:
   C. 1 antenna BC's versus the averaged 16 antenna BC's.

with a tumor inclusion (see Section 4.2.2) suspended in the saline solution. In general, the differences were roughly unchanged for the homogeneous case. Implementation of this concept of an effective $\hat{r} \cdot \hat{n}$ contribution is essential for facilitating the use of the multi-right-hand-side implementation of the matrix system solver described in the next section.

Iterative solver with multiple right hand sides

In general, matrix A in (4.2) is too large to compute by direct methods (Choleskey or LU decomposition) but is well suited to iterative solvers [18], of which the bi-conjugate gradient (BCG) and the quasi-minimum residual (QMR) methods are two of the most common. We have focused on the QMR method because it has been demonstrated to be superior to the BCG approach for applications involving large sparse matrices [57].

A number of strategies can be applied to precondition A in order to cluster its eigenvalues which has the intended consequence of accelerating solution convergence [67]. Possible options include the incomplete Cholesky and the incomplete LU (ILU) preconditioners of which the latter was chosen because of its superior performance [18]. The QMR routine can be implemented in block form utilizing the block Lanczos algorithm to process multiple right-hand-side vectors of (4.2)), simultaneously. We performed several studies to determine whether the grouping of specific right-hand-side vectors along with the block size had any effect on the solution convergence rate. In general, grouping right-hand-side vectors corresponding to adjacent antennas provided faster convergence than when antenna vectors were arranged in random order. In addition, the size of the block had a significant impact on the convergence rate. Figure 4.3 shows a plot of the forward solution time per right-hand-side vector versus block size for the configuration in Figure 4.2 (b) with 10,571 nodes and 54,720 tetrahedral elements. In this case, 16 antennas on a 15 cm diameter circle in a single plane were used to generate the source fields. The convergence time per right-hand-side is minimum at the block size of seven. We generally used a size of eight as a convenient denominator for organizing the complete set of antennas with only nominal degradation in solution time.

4.1.2 Image reconstruction and the dual-mesh adjoint method

We previously utilizes the regularized Gauss-Newton approach described in Section 3.2.3 for reconstructing 2D permittivity and conductivity maps of a desired imaging zone. In contrast to the direct differentiation approach used in the 2D method (Section 3.5), we have derived a fast algorithm to compute the elements of the Jacobian matrix by exploiting the principle of reciprocity (a signal transmitted from antenna A and

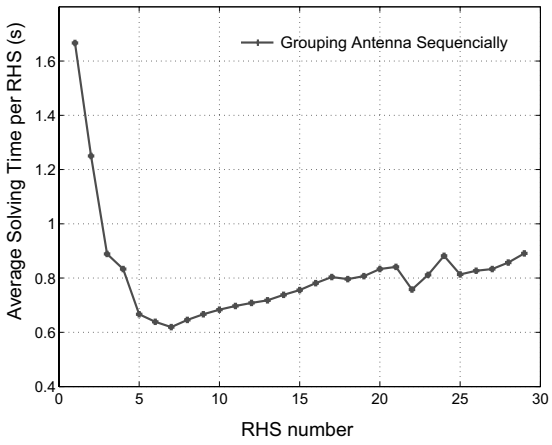


Figure 4.3: Plot of the forward problem computation time per source antenna as a function of block size (number of right hand sides computed simultaneously) when the block QMR solver is used.

received at antenna B equals the signal transmitted from B and received at A).

Revisiting the forward solution equation (3.13), we can rewrite it in an operator equation form as

$$\mathcal{L}E_z(\vec{r}) = b(\vec{r}) \tag{4.11}$$

where the forward operator $\mathcal{L}$ is given by

$$\mathcal{L} = \nabla^2 + k^2(\vec{r}) \tag{4.12}$$

and b is the source term.

We shall demonstrate that the operator $\mathcal{L}$ is a linear self-adjoint operator. $\mathcal{L}$ is essentially a differential operator which is linear because the two conditions in (2.5) are satisfied in the functional space. The property of self-adjointness is proven below: for two arbitrary functions, $f(\vec{r})$ and $g(\vec{r})$, defined over domain Ω, the adjoint operators $\mathcal{L}^a$ of $\mathcal{L}$ satisfy

$$\langle g(\vec{r}), \mathcal{L}^a f(\vec{r})\rangle = \langle f(\vec{r}), \mathcal{L}g(\vec{r})\rangle \tag{4.13}$$

Utilizing (4.12), the RHS of (4.13) can be expanded as

$$\begin{aligned} \langle f(\vec{r}), \mathcal{L}g(\vec{r})\rangle &= \int_\Omega f(\vec{r})\left(\nabla^2 g(\vec{r}) + k^2(\vec{r})g(\vec{r})\right)d\vec{r} \\ &= \int_\Omega f(\vec{r})\nabla^2 g(\vec{r})d\vec{r} + \int_\Omega f(\vec{r})k^2(\vec{r})g(\vec{r})d\vec{r} \end{aligned} \tag{4.14}$$

If $f(\vec{r})$ and $g(\vec{r})$ are orthogonal at the boundary of the domain, i.e $\int_{\partial\Omega} f(\vec{r})g(\vec{r})d\vec{r} = 0$, and applying the Green's second identity (3.48), (4.14) can be written as

$$\begin{aligned} \langle f(\vec{r}), \mathcal{L}g(\vec{r})\rangle &= \int_\Omega g(\vec{r})\nabla^2 f(\vec{r})d\vec{r} + \int_\Omega g(\vec{r})k^2(\vec{r})f(\vec{r})d\vec{r} \\ &= \langle g(\vec{r}), \mathcal{L}f(\vec{r})\rangle \end{aligned} \tag{4.15}$$

Thus, from (4.15) and (4.13), we have $\mathcal{L}^a = \mathcal{L}$.

For a point source located at $\vec{r}_s$, (4.11) can be rewritten as

$$\mathcal{L}E_s(\vec{r}) = b_s(\vec{r}) \tag{4.16}$$

where $b_s(\vec{r}) = -J_{s0}\delta(\vec{r} - \vec{r}_s)$ and $E_s(\vec{r})$ is the radiation field of this source. Taking the derivative of (4.16) with respect to the τ-th parameter, k_τ^2, yields

$$\mathcal{L}S_a(\vec{r}) = b_a(\vec{r}) \tag{4.17}$$

where $b_a(\vec{r}) = -\frac{\partial \mathcal{L}}{\partial k_\tau^2}E_s(\vec{r})$ and $S_a = \frac{\partial E_s(\vec{r})}{\partial k_\tau^2}$. For $\mathcal{L}^a = \mathcal{L}$, equation (4.17) is referred as the adjoint equation of (4.16).

Additionally, assuming a normalized point source (defined in (3.45)) is applied at the receiver location $\vec{r}_r$, the resultant field distribution $E_r(\vec{r})$ can be computed by

$$\mathcal{L}E_r(\vec{r}) = b_r(\vec{r}) \tag{4.18}$$

where $b_r(\vec{r}) = -\delta(\vec{r} - \vec{r}_r)$.

Reciprocity of the electromagnetic field ensures that

$$\langle b_r(\vec{r}), S_a(\vec{r})\rangle = \langle b_a(\vec{r}), E_r(\vec{r})\rangle \tag{4.19}$$

Substituting the definitions of b_r, b_a, E_r and S_a into (4.19) produces

$$\int_\Omega \delta(\vec{r}-\vec{r}_r)\frac{\partial E_s(\vec{r})}{\partial k_\tau^2}d\vec{r} = \int_\Omega -\left(\frac{\partial \mathcal{L}}{\partial k_\tau^2}E_s(\vec{r})\right)E_r(\vec{r})d\vec{r} \tag{4.20}$$

Considering the sampling property of the Dirac delta function, the LHS of (4.20) becomes $\frac{\partial E_s(\vec{r}_r)}{\partial k_\tau^2}$ which represents the sensitivity of the measured field at receiver $\vec{r}_r$ with respect to a perturbation of the τ-th parameter, k_τ^2. This is precisely the definition of a single element of the Jacobian matrix. As a result, we can rewrite (4.20) as

$$J((s,r),\tau) = \int_\Omega -\left(\frac{\partial \mathcal{L}}{\partial k_\tau^2}E_s(\vec{r})\right)E_r(\vec{r})d\vec{r} \tag{4.21}$$

where s, r and τ are the indices of the source, receiver and parameter, respectively.

If the forward operator $\mathcal{L}$ is discretized using the Galerkin scheme, (4.21) can be further simplified by incorporating the dual-mesh basis functions in (3.33) and the weak-form equation (3.42). A single element of the Jacobian matrix in the discretized form can be expressed as

$$J((s,r),\tau) = (D_\tau \mathbf{E}_s)^T \mathbf{E}_r \tag{4.22}$$

where the (i,ℓ)-th element of the matrix D_τ is defined as

$$d_{i,\ell} = \int_\Omega \phi_i(\vec{r})\phi_\ell(\vec{r})\varphi_\tau(\vec{r})d\vec{r} \tag{4.23}$$

It is interesting that the matrix D_τ in (4.23) is only related to the forward and pa-

rameter basis functions. Once the dual-mesh pair is determined, for all τ, D_τ can be constructed ahead of time and stored for use in the reconstruction process. Moreover, the basis function $\varphi_\tau(\vec{r})$ is only non-zero at a small subzone of Ω, i.e. the neighboring parameter elements of the τ-th parameter mesh node. We denote the domain where $\varphi_\tau(\vec{r}) \neq 0$ as Ω_τ. Thus, for most of the (i, ℓ) pairs where ϕ_i and ϕ_ℓ are defined outside Ω_τ, the integration in (4.23) yields zero and D_τ becomes a very sparse matrix. The evaluation of the integration becomes complicated when the boundaries of the forward element mesh are not precisely conformal to the boundaries of the parameter elements. In this chapter, we only consider the conformal dual-mesh pair (an example is shown in Figure 4.4).

In all cases, evaluating the Jacobian matrix J with (4.22) involves only an inner product of field distributions (which are already computed at each iteration) multiplied by weighting coefficients near each parameter node which is an $O(N)$ operation compared with $O\left(N^2\right)$ solving the sensitivity equation (3.54) for all permutations of sources, s, and reconstruction parameters, τ. As N increases, the computational savings become considerable. Note that equation (4.22) is general and applicable to both 2D and 3D reconstruction problems (for both dual-mesh or single-mesh schemes).

4.2 Results

We have organized several experiments utilizing simulated measurement data to demonstrate the viability of the 3Ds/2D algorithm. The scope of the investigations reported here centers on computational efficiency, suitability of convergence behavior and whether there are improvements over the 2D/2D approach with respect to metrics devised to quantify 3D wave propagation effects. Specifically, in Sections 4.2.1 and 4.2.2, we

have purposely eliminated the issue of data-model mismatch caused by using the scalar field approximation through the incorporation of simulated measurements which were generated by the scalar model. This synthetic data set also assumed no measurement noise. The intent of these studies is to highlight the ideal performance of the inversion algorithm prior to provoking any image quality degradations resulting from modelling error. In the spirit of establishing the ideal performance, we computed the measurement data on the same mesh employed for image recovery in these simulations. In Section 4.2.3, on the other hand, full vector field solutions were produced on a different high resolution mesh in order to construct the synthetic measurement data for image reconstruction which includes the effects of scalar model approximation. We also quantify the differences that can be expected between the 3D vector and scalar models under representative conditions at the start of this section.

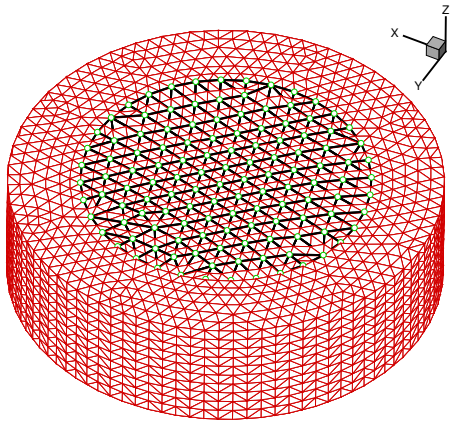


Figure 4.4: 3D forward mesh overlapped with 2D reconstruction mesh. (Recognizing that the boundaries of parameter elements are conformal to the forward mesh).

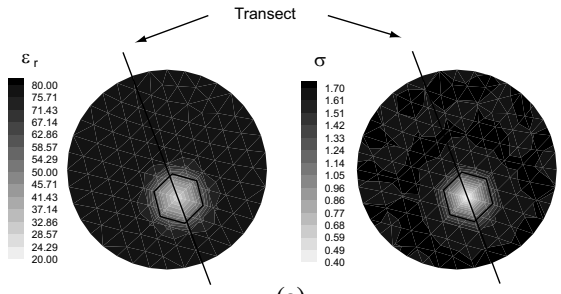
For all of the imaging experiments described here, the 3D FE region is an 18 cm diameter cylinder with a height of 5 cm comprised of 11 uniform layers (as shown in Figure 4.4) with a total of 10,571 nodes and 54,720 tetrahedral elements. The circular

array (7.1 cm radius) of 16 antennas (1 cm in length) is concentric within the vertical axis of the cylindrical FE mesh. In each imaging experiment, 16 transmitters were used while signals were received at 9 opposing antennas for a total of 144 measurements. The 2D image reconstruction mesh was a 12 cm diameter circle concentric with the cylindrical volume comprised of 126 nodes and 214 triangular elements. For the reconstruction process, the Tikhonov regularization with the empirical regularization parameter as discussed in Section 3.2.3 was used in conjunction with a spatial filter applied at each iteration to remove high frequency variations through an averaging factor of 0.1 (Section 3.2.3). In all cases the images converged to a stable solution within roughly 6 iterations which required approximately 1 minute to execute on a Compaq Alpha 833 MHz ES40 workstation. For the images presented in Sections A and B, the background medium was 0.9% saline ($\epsilon_r = 77$ and $\sigma = 1.7$ S/m) with an operating frequency of 900 MHz. The experiments in Section 4.2.3 utilized a range of backgrounds to illustrate the influence of contrast on 3D effects. All of these reconstructions started from an initial estimate consisting of the values for the homogeneous background.

4.2.1 Simple cylindrical phantom

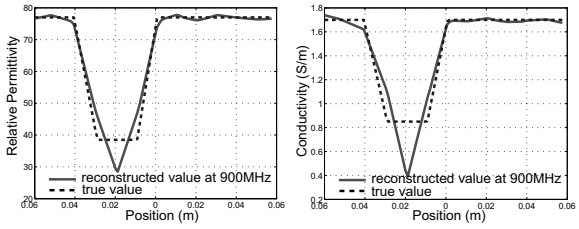
Figure 4.5 (a) shows the 900 MHz reconstructed images of a 2.9 cm diameter cylinder (the geometry of the object is superimposed on the images) with $\epsilon_r = 38.5$ and $\sigma = 0.85$ S/m for a contrast of 1:2 with the background properties. Both the permittivity and conductivity distributions are recovered quite well with only minor artifacts appearing in the conductivity image background. Transects through both images plotted in Figure 4.5 (b) illustrate the uniformity of the background recovery along with the position and size of the inclusion with respect to the exact distribution (also shown). It is interesting to note that the recovered properties underestimate those of the actual object in the cen-

ter. This may be a consequence of the spatial filtering which constrains the algorithm from exactly recovering a property step-distribution at the object background interface, forcing the algorithm to compensate for this limitation by exaggerating the properties in the center of the object.



(a)

(a)



(b)

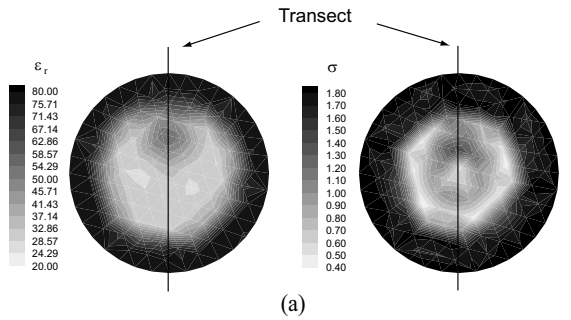
(b)

Figure 4.5: a) 900 MHz reconstructed permittivity and conductivity images for a 2.9 cm diameter cylinder within a homogeneous saline background, and b) the associated property transects through the imaging domain including the recovered object compared with the actual distributions.

4.2.2 Breast-like cylindrical phantom

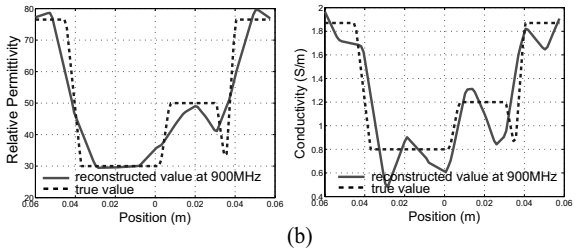
Figure 4.6 (a) shows the 900 MHz permittivity and conductivity images recovered for a centered breast-like region with an inclusion. The properties of the large 8 cm (roughly 2.2λ) diameter cylinder were $\epsilon_r = 30.0$ and $\sigma = 0.8$ S/m while those for the offset, 3

cm diameter inclusion were $\epsilon_r = 50.0$ and $\sigma = 1.2$ S/m. In contrast to the previous example, our log-magnitude/unwrapped phase minimization (Section 3.2.3) was used for this case because the standard complex form diverged as a result of the high proportion of measured data phase wrapping [151]. The excessive phase wrapping of the scattered fields is directly related to the target size, contrast and operating frequency. It is important that concepts successfully developed in the 2D/2D approach extend to the 3Ds/2D implementation.



(a)

(a)



(b)

(b)

Figure 4.6: a) 900 MHz reconstructed permittivity and conductivity images for an 8 cm diameter breast-like phantom with a 3 cm diameter inclusion within a homogeneous saline background, and b) the associated property transects through the imaging domain including the recovered breast and inclusion compared with the actual distributions.

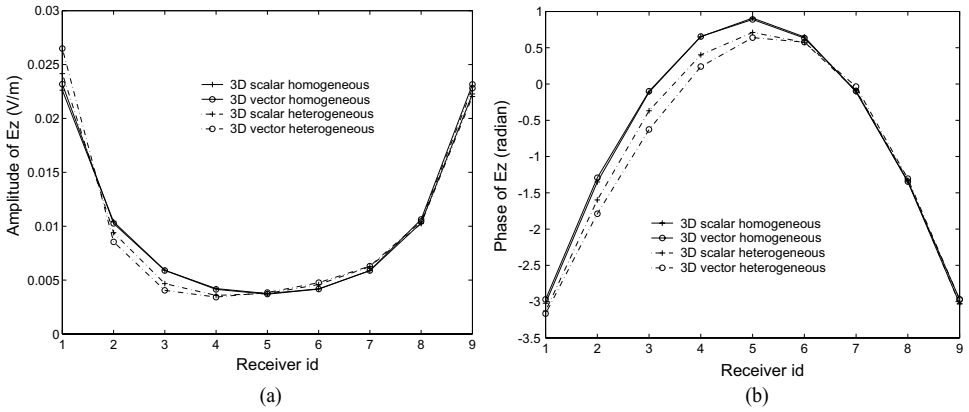
From the image pair in Figure 4.6, it is clear that the permittivity component is recovered more accurately than its conductivity counterpart. For instance, there is a

considerably higher level of artifacts in both the background properties and internal breast composition for the conductivity image and the reconstructed breast geometry appears smaller for the conductivity relative to the permittivity map. These observations are generally consistent with previously reported findings [124]. For the plots of the recovered properties along the vertical transects through the phantom (Figure 4.6 (b)), previous observations are also confirmed in that both images recover accurate property profiles, although the permittivity component generally has fewer artifacts. Additionally, it also appears that the location of the recovered inclusion is correct for the permittivity image while it is skewed noticeably downwards in the conductivity case. This is again consistent with previously reported images obtained from the 2D/2D configuration [110], and is clearly exacerbated by the high contrast background.

4.2.3 Reduction in 3D propagation effects

Prior to investigating image reconstruction differences in a 3D problem between our existing 2D/2D algorithm and the 3Ds/2D approach developed here, we begin this section by quantifying the differences in the underlying 3D scalar model of field propagation with its more appropriate 3D vector version under two representative conditions – a homogeneous imaging volume and an heterogeneous volume. Figure 4.7 shows field values computed for the 3D scalar and full vector models at the imaging array antenna sites under 500 MHz illumination of a homogeneous and heterogeneous volume. In the homogeneous case a background medium with $\epsilon_r = 70$ and $\sigma = 1.7$ S/m was used while in the heterogeneous problem this same background included an off-centered sphere ($r = 1.78$ cm, 2.54 cm offset from the center of the antenna array) with $\epsilon_r = 20$ and $\sigma = 0.18$ S/m. The solutions illustrated in Figure 4.9 result in mean amplitude differences of roughly 0.4% and 1.2% in the homogeneous and heterogeneous prob-

lems, respectively, and a mean difference of 1.5° and 5.4° in phase in the two cases (in both cases, the solutions were computed with a mesh of 56,636 nodes and 312,453 elements).



(a) (b)

Figure 4.7: Comparison of the a) magnitude and b) phase of the fields at antenna array measurement sites for the 3D scalar and vector propagation models in a homogeneous background and a background containing a spherical object within the array.

A previous study based on the 2D/2D algorithm presented results utilizing a metric to estimate the imaging slice thickness [126] in order to quantify imaging artifacts due to 3D field propagation effects. We present data here which directly compares the 3Ds/2D and 2D/2D algorithms in terms of this measure. The simulated scattered data was computed using a full finite element 3D vector formulation [154, 153] at 900 MHz. The study involved raising low permittivity spheres of different diameters through the imaging plane defined by the array of monopole antennas. Permittivity and conductivity images were recovered for each sphere (large: 4.6 cm diameter, 1.3λ; medium: 3.6 cm diameter 1λ; small: 2.5 cm diameter, 0.7λ) at each vertical position separated in 1.27 cm increments. The electrical properties of the spheres were $\epsilon_r = 20.0$ and $\sigma = 0.18$ S/m while the relative permittivity of the background varied from 30 to

70 with the conductivity fixed at 1.78 S/m. As an indication of the averaging effect along the z-axis due to 3D microwave propagation, imaging slice thicknesses derived from the recovered sphere half width and peak values of its estimated properties (see Meaney *et al.* [126] for a complete definition) were computed for both the permittivity and conductivity, for all three spheres and over the complete range of background permittivities, respectively. Figure 4.8 compares plots of the imaging slice thickness as a function of background permittivity for both the 2D/2D and 3Ds/2D algorithms. For almost all cases, the slice thickness for the 3Ds/2D algorithm is smaller. In general, the slice thickness is greater for larger sized spheres and the permittivity slice thickness is smaller than in the corresponding conductivity images. The conductivity thickness decreases consistently for both algorithms and appears to converge to similar values for the lowest background permittivity. The permittivity thicknesses for the 2D/2D algorithm are relatively flat as a function of background permittivity – decreasing slightly for the large sphere and increasing slightly for the two smaller spheres. In contrast, the permittivity thicknesses for the 3Ds/2D cases consistently decrease with respect to a lower background permittivity. In general, both algorithms demonstrate improvement in reducing the imaging slice thicknesses with reduced background contrast, where the 3Ds/2D algorithm produces consistently smaller values and is generally better at handling larger targets. As a visual example of the 3Ds/2D algorithm enhancement in a representative case, Figure 4.9 shows a pair of conductivity images obtained from the reconstructions of the large sphere with the antenna array positioned in the azimuthal plane by the two methods. The improvement is evident in terms of the sharper and more accurate contrast with the background that is achieved with the 3Ds/2D method.

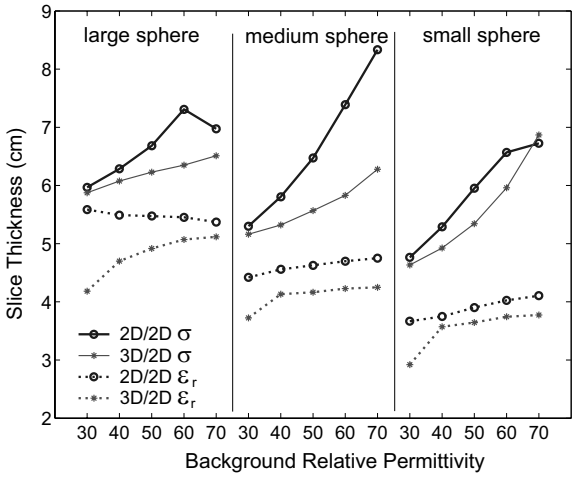
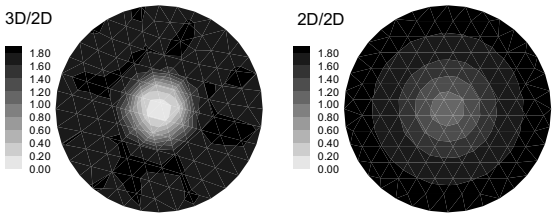


Figure 4.8: Plots of the slice thickness computed at 900 MHz for the recovered permittivity and conductivity images using 4.6, 3.6, and 2.5 cm diameter spheres (ϵ_r = 20.0, and σ = 0.18 S/m) as a function of background permittivity (σ = 1.78 S/m). Plots are compared with corresponding results using the 2D/2D algorithm.

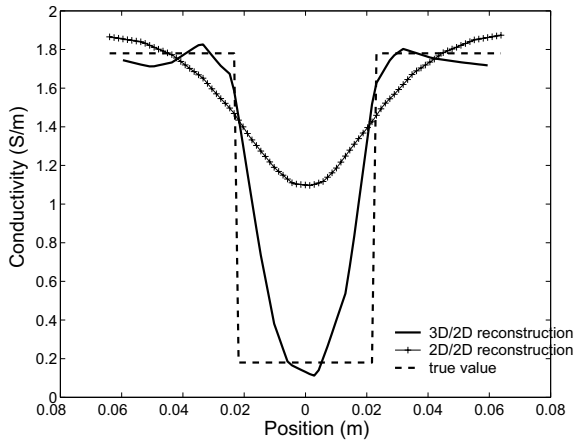
4.3 Conclusions

We have implemented a 3D scalar field solution/2D Gauss-Newton iterative parameter inversion algorithm (3Ds/2D) for microwave imaging. Various strategies including exploitation of a 3D scalar formulation and a truncated mesh with radiation boundary conditions were deployed to limit the computational overhead of the problem. Additionally, as an important initial step, we developed a 2D reconstruction procedure that is integrated with the 3D field solution. This is significant in a practical sense because measurement data is a precious commodity and restricting the parameter reconstruction to 2D allows this new algorithm to be applied to the limited microwave signal channels now in place [122]. However, it is also important because the 2D inversion portion of the algorithm has been developed to readily generalize to a full 3D reconstruction when the requisite amount of measurement data can be acquired. Implementation of the adjoint method to dramatically accelerate computation of the Jacobian matrix has made 3D approaches much more attainable.



(a)

(a)



(b)

(b)

Figure 4.9: a) Plots of the reconstructed conductivity for the large sphere (4.6cm diameter) in background $\epsilon_r = 60$ at 900 MHz by 3Ds/2D and 2D/2D methods, b) transects of the reconstructed conductivity profiles together with the true value of the distribution.

The range of results presented here demonstrates the capabilities of the 3Ds/2D algorithm in a variety of settings – specifically in a simple 2D cylinder and a more complex, large cylindrical geometry. In addition, a full set of experiments were performed to illustrate that the 3Ds/2D algorithm is an overall improvement over the 2D/2D algorithm in terms of reducing previously observed 3D artifacts. In general, these results are encouraging and set the stage for development of more advanced 3D/3D methods in the next chapter.

Chapter 5

Three dimensional microwave imaging

Three-dimensional microwave imaging utilizing inverse scattering techniques has received increased attentions over the past decade. With the dramatic increase in problem size of both the forward field modelling and parameter reconstruction associated with the transition from 2D to 3D problems, enhancement of the computational efficiency is critically important in making microwave imaging viable. In this chapter, we present two 3D image reconstruction methods utilizing 3D scalar and vector forward computational models, respectively. The proposed methods, together with the previously proposed 2D and 3D/2D hybrid methods, outline a spectrum of dual-mesh based algorithms which enable us to investigate the fine balance between the model accuracy and efficiency. Several strategies developed for the preceding approaches were improved upon and incorporated into the new 3D reconstruction algorithms including the iterative block solver and the adjoint method for constructing the Jacobian matrix. For the 3D vector forward method, an optimized FDTD method with a uniaxial perfectly matched layer (UPML) technique was developed to obtain more complete field information within acceptable computational time and is considerably faster than the 3D reconstructions previously reported. These algorithms were tested with both simu-

lated and experimental data where the measurement was acquired utilizing our new 3D data acquisition system for microwave breast imaging. The new reconstructions were evaluated in terms of the computational efficiency and compared with other previously developed algorithms.

This chapter is organized in the following manner: the theory and method section (Section 5.1) focuses on explaining the method details and innovations of the proposed 3D reconstruction algorithms, i.e. the 3D scalar forward/3D reconstruction (3Ds/3D) and the 3D vector forward/3D reconstruction (3Dv/3D). 3D dual-meshes, the nodal adjoint method and optimizations of 3D vector field solver are discussed sequentially. A brief introduction to our 3D data acquisition system is provided in Section 5.2. The results section contains simulation and measurement data reconstructions to test and evaluate the performance of the proposed algorithms by making comparisons to the existing 2D methods. In particular, the computation complexities for the range of dual-mesh based algorithms are compared. We conclude this chapter with a summary discussion in Section 5.4.

5.1 Theory and method

The 3D reconstruction algorithms in this chapter are direct extensions of the 2D/2D and the 3Ds/2D algorithms discussed in Chapter 3 and 4. The iterative regularized Gauss-Newton method and the dual-mesh scheme are employed to build the algorithmic infrastructure in the 3D space. One of the most significant changes compared with the preceding algorithms is that the reconstruction meshes are 3D regions instead of 2D slices. Consequently, these 3D algorithms should no longer be called as "tomographic" imaging methods in a strict sense since tomography refers to cross-sectional imaging.

Another difference is the dramatic increase in the sizes of forward field and parameter modelling problems. In addition, since more unknowns must be reconstructed, this requires more measurement data.

The introduction of the adjoint method in the last chapter dramatically shortened the computation time for constructing the Jacobian matrix, which, consequently, transferred the burden of the reconstruction time-limiting step to the forward field calculation. For 3D imaging, improvement of the forward modelling efficiency is essential for the viability of the algorithm. The 3D finite element method with iterative block solver devised in the previous chapter has demonstrated promise in solving the scalar forward problem simplified form the vector form with the 2D property distribution and TM wave approximations. For 3D reconstructions, the dielectric profiles are also 3D distributions and the 2D assumption becomes yet another approximation. Alternatively, the computational advantages of the FDTD method in 2D problems has been demonstrated in Section 3.6.1 which motivates the use of the 3D FDTD method to model the full 3D vector equation. The marching-on-time (MOT) feature of the FDTD method also provides additional opportunities to further reduce the computational time. Finally, an approximated adjoint method, the nodal adjoint method, is also derived in this chapter which further enhances the reconstruction efficiency while introducing only nominal accuracy degradation.

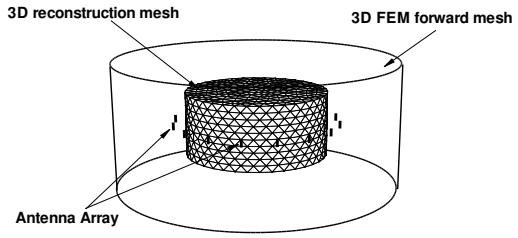
In this section, the dual-mesh configurations for the 3D scalar forward/3D reconstruction and the 3D vector forward/3D reconstruction are first demonstrated. Then the nodal adjoint method is derived with a general form which can be used in a range of algorithms, including the 3D algorithms discussed in this chapter. Formulation of the 3D FDTD algorithm is subsequently discussed together with the implementation details especially with respect to the absorbing boundary condition and the computational

efficiency analysis. Finally, enhanced FDTD methods implementing high-order FDTD versions, an ADI update scheme and various additional accelerations via initial field computations are discussed in conjunction with their overall efficiencies are analyzed and compared with the traditional 3D FDTD forward solver.

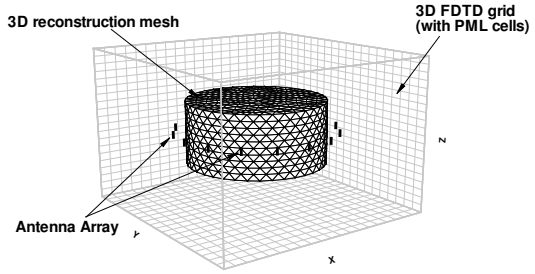
5.1.1 3D dual-meshes

The dual-mesh scheme discussed in Section 3.3 was developed in a general manner such that it could be utilized in modelling the 3D distributions of both the forward fields and parameters. With respect to the imaging system configuration at Dartmouth College, the dual-meshes of the 3Ds/3D and the 3Dv/3D methods are illustrated in Figure 5.1 (a) and (b).

The forward mesh for the 3Ds/3D method is similar to the one used in the 3Ds/2D method (Figure 4.1), i.e. a cylindrical mesh concentrically aligned with the circular antenna array while extending beyond the antennas in the radial direction. The forward mesh for the 3D vector reconstruction is a cubic shaped 3D Yee-grid which will be discussed in the following subsection. The reconstruction meshes for both methods are identical 3D cylindrical meshes which are concentric to both the antenna array and the forward meshes to model the inhomogeneities within a 3D domain. Facilitated by the dual-mesh representation, the forward and reconstruction meshes can possess different mesh densities. Moreover, the nodal adjoint method, which will be derived soon, significantly simplifies the computations of the Jacobian matrix in cases where the the forward mesh and the reconstruction mesh are not precisely conformal at their boundaries (such as the case in Figure 5.2). This simplifies the preprocessing of the reconstruction and greatly enhances the flexibility.



(a)



(b)

Figure 5.1: Forward and reconstruction mesh orientations for (a) the 3Ds/3D and (b) the 3Dv/3D methods

5.1.2 Nodal adjoint method

The adjoint formula of the Jacobian matrix (4.22) and (4.23) can be rewritten in terms of a summation over forward elements as

$$J((s,r),\tau) = \sum_{e\in\Omega_\tau} (D_\tau^e \mathbf{E}_s^e)^T \mathbf{E}_r^e \tag{5.1}$$

where Ω_τ denotes the region within which $\varphi_\tau \neq 0$ and $\sum_{e\in\Omega_\tau}$ denotes the summation over the forward elements which are located within Ω_τ. Matrix D_τ^e is a square matrix with each element defined by

$$d^\tau_{i_e,\ell_e} = \int_{\Omega_e} \phi_{i_e}(\vec{r})\phi_{\ell_e}(\vec{r})\varphi_\tau(\vec{r})d\vec{r} \tag{5.2}$$

where $i_e = 1,2,\cdots,M$ and $\ell_e = 1,2,\cdots,M$ are the local node indices and M is the total node number for a single forward element ($M = 3$ in 2D and 4 in 3D). ϕ and φ represent the basis functions over the forward and reconstruction meshes, respectively. Ω_e is the spatial domain occupied by the e-th forward element. $\mathbf{E}_s^e = \{E_s(\vec{p}_\kappa^e)\}_{\kappa=1}^M$ and $\mathbf{E}_r^e = \{E_r(\vec{p}_\kappa^e)\}_{\kappa=1}^M$ are the fields at the vertices $\{\vec{p}_\kappa^e\}_{\kappa=1}^M$ of the element due to source antennas at s and r, respectively. Equation 5.1 is referred as the element-based form of the adjoint formula.

For cases where the boundaries of the forward elements do not precisely match those of the parameter elements (such as the 2D case shown in Figure 5.2), the evaluation of the integration in (5.2) becomes more difficult since integrations often have to be evaluated over partial elements of the forward mesh. A nodal adjoint method is introduced to simplify the integration for a given dual-mesh pair under the assumption that the averaged size of the forward elements is significantly smaller than that of the parameter elements (a discussion on the accuracy of the nodal adjoint method can be

found at the end of this subsection). We shall derive the expression for this method.

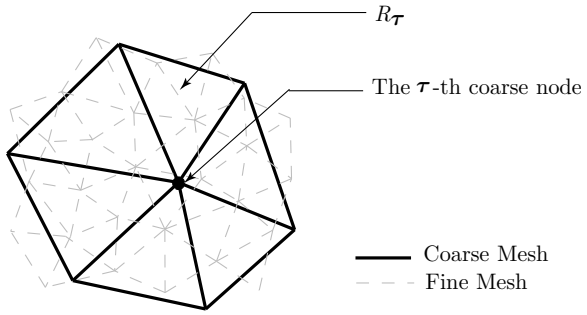


Figure 5.2: Challenging dual-mesh configuration for the element based adjoint method.

Within domain Ω_e where $e \in \Omega_\tau$, the parameter basis function φ_τ can be expanded as a linear combination of the forward basis functions:

$$\varphi_\tau(\vec{r}) = \sum_{\kappa=1}^{M} \varphi_\tau(\vec{p}_\kappa)\phi_\kappa(\vec{r}) \tag{5.3}$$

Inserting (5.3) into (5.1), we get:

$$J((s,r),\tau) = \sum_{e\in\Omega_\tau} \sum_{\kappa=1}^{M} \varphi_\kappa(\vec{p}_\kappa)(D^e_{\tau\kappa}\mathbf{E}^e_s)^T\mathbf{E}^e_r \tag{5.4}$$

where $D^e_{\tau\kappa}$ is an $M \times M$ matrix defined as

$$D^e_{\tau\kappa} = \begin{pmatrix} \langle\phi_1\phi_1\phi_\kappa\rangle & \langle\phi_1\phi_2\phi_\kappa\rangle & \cdots & \langle\phi_1\phi_M\phi_\kappa\rangle \\ \langle\phi_2\phi_1\phi_\kappa\rangle & \langle\phi_2\phi_2\phi_\kappa\rangle & \cdots & \langle\phi_2\phi_M\phi_\kappa\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi_M\phi_1\phi_\kappa\rangle & \langle\phi_M\phi_2\phi_\kappa\rangle & \cdots & \langle\phi_M\phi_M\phi_\kappa\rangle \end{pmatrix} \tag{5.5}$$

where $\langle\cdot\rangle$ denotes the volume integration over Ω_e. Notice that the nonzero off-diagonal elements in $D^e_{\tau\kappa}$ result in cross-multiplication terms of the fields at different nodes in (5.4). To simplify the analysis, we approximate the weighting matrix $D^e_{\tau\kappa}$ by summing each column (or row) and adding the off-diagonal elements to the diagonal and simul-

taneously zeroing out all off-diagonal terms

$$\tilde{D}^{e}_{\tau\kappa} = \begin{pmatrix} \sum_{i=1}^{M}\langle\phi_i\phi_1\phi_\kappa\rangle & 0 & \cdots & 0 \\ 0 & \sum_{i=1}^{M}\langle\phi_i\phi_2\phi_\kappa\rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=1}^{M}\langle\phi_i\phi_M\phi_\kappa\rangle \end{pmatrix} \tag{5.6}$$

It is not difficult to prove that (see Appendix E)

$$\sum_{\kappa=1}^{M} \tilde{D}^{e}_{\tau\kappa} = \frac{V_e}{M} I \tag{5.7}$$

where V_e is the volume of the e-th forward element (in 2D, V_e is the area of the element) and I is an $M \times M$ identity matrix. By substituting (5.7) back into (5.4) and expanding the vector multiplications, the reorganized equation can be written as

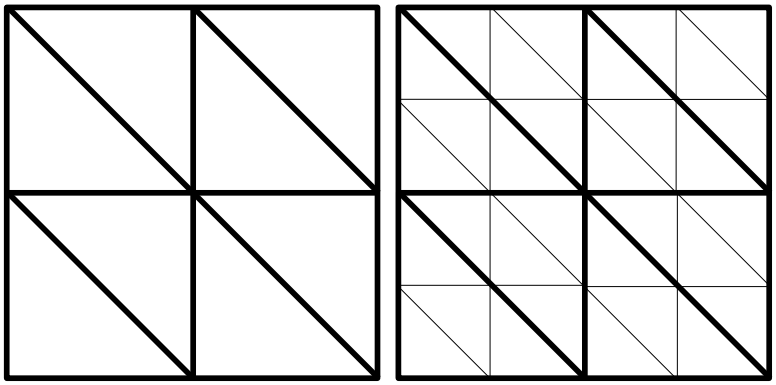
$$J((s,r),\tau) = \sum_{n\in\Omega_\tau} \left(\frac{\sum_{e\in\Omega_n} V_e}{M} \right) \varphi_\kappa(\vec{p}_n) E_s(\vec{p}_n) E_r(\vec{p}_n) \tag{5.8}$$

where $\sum_{n\in\Omega_\tau}$ refers to the summation over the forward nodes which fall inside Ω_τ and $\sum_{e\in\Omega_n}$ refers to the summation over the forward elements that share the n-th forward node. The term $\left(\frac{\sum_{e\in\Omega_n} V_e}{M}\right)$ is a scalar term associated with the n-th forward node which can be simplified as V_n.

The nodal form adjoint formula (5.8) allows us to compute the Jacobian matrix for both conformal and nonconformal dual-meshes quite easily: $E_s(\vec{p}_n)$ and $E_r(\vec{p}_n)$ are the nodal electrical field values computed directly from the forward problem; V_n and $\varphi_\tau(\vec{p}_n)$ require only simple algebraic operations and can be built on-the-fly. This is important for forward techniques which might dynamically generate their meshes, such as FDTD and some adaptive methods. Note that the reconfiguration of the weighting matrix $D^{e}_{\tau\kappa}$ is only valid when the forward element is substantially small with respect to the

parameter mesh elements such that the field values at its vertices are approximately equal.

To validate our derivations, we computed the Jacobian matrices using the nodal adjoint formula over a series of refined dual-meshes (40×40 nodes grid). Plots of a segment of the original single-mesh and the refined forward mesh (level 2) are shown in Figure 5.3. The maximum relative error between the nodal adjoint and the adjoint Jacobian are plotted vs. the ratios between the averaged parameter and forward element sizes (Figure 5.4). From this plot, it is reasonable to assume that when the forward element is small compared with the parameter element, the nodal adjoint Jacobian is a good approximation to the accurate Jacobian matrix.



(a) (b)

Figure 5.3: A fraction of the dual-meshes with different parameter/forward element area ratios: (a) 1:1 and (b) 4:1. The forward and parameter meshes are denoted by thin and thick lines, respectively. Note that in both diagrams, part of the forward mesh is overlapped by the parameter mesh.

5.1.3 3D vector forward solution coupled with 3D reconstruction

All of the dual-mesh based algorithms which have previously been developed rely on certain assumptions or approximations in order to simplify the forward model and re-

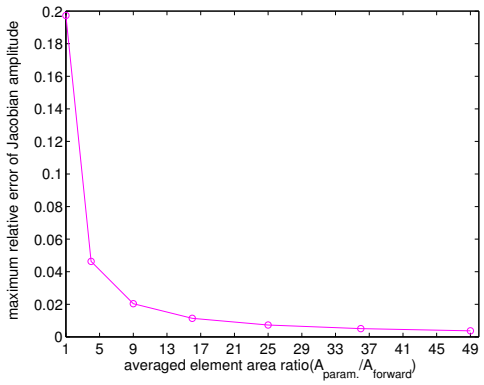


Figure 5.4: Plot of the maximum relative error of the nodal adjoint Jacobian at various parameter/forward element area ratios.

duce the reconstruction time. The 3D vector forward solution provides the most complete forward field solution and is clearly the most challenging case. A 3D FDTD method is chosen in this case for its simplicity and efficiency coupled with a uniaxial perfectly matched layer (UPML) technique as the absorbing boundary condition similarly to the GPML technique in the 2D case.

In the following subsections, we discuss a variety of issues associated with the 3Dv/3D method, including the formulation of the UPML update scheme in lossy media, the nodal adjoint expression for constructing the Jacobian matrix, and the possibility of reducing the computation time by utilizing high order spatial difference and initial field distributions. An alternative-directional-implicit (ADI) update scheme for lossy media using UPML is developed to implement the acceleration via setting the initial field values.

Forward model and dual-mesh

The derivation of the standard FDTD method in 3D space is very similar to that of the 2D case, except that all 6 components, i.e. E_ξ and H_ξ ($\xi = x, y, z$), are nonzero and there are 6 equations in the expanded curl relations instead of 3 as in 2D FDTD method

under the TM wave assumption (equation (3.1)). For conciseness, we focus only on the 3D Yee-grid and absorbing boundary condition while omitting the derivation of the standard 3D FDTD update equations.

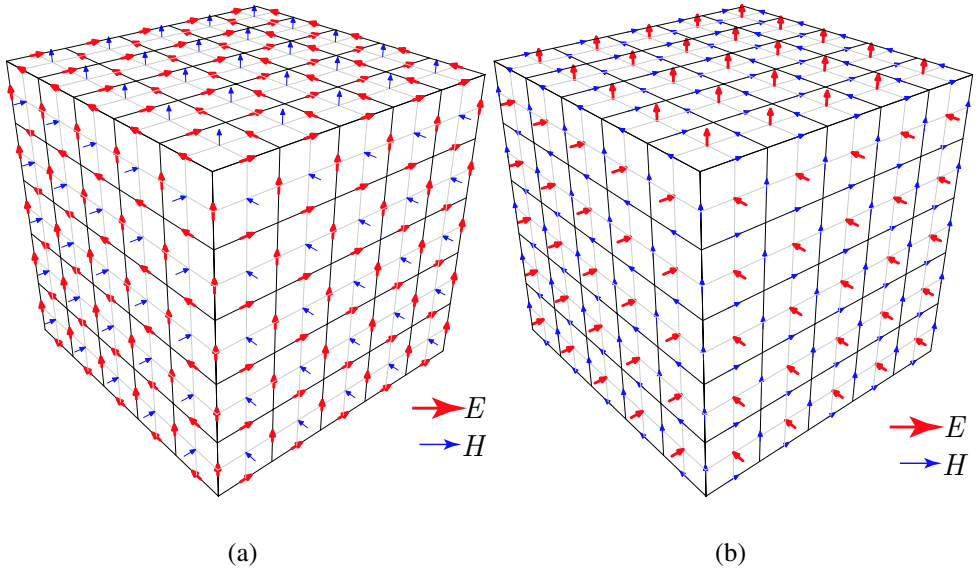


Figure 5.5: FDTD grid in 3D space: (a) E-grid, (b) H-grid.

The FDTD grids in the 3D space also have two variants, the E-grid and the H-grid shown in Figure 5.5. Similar to the 2D cases, these two grids represent the identical spatial relationships except that the origin of each grid is located at (1/2,1/2,1/2) of the other. The uniform E-grid ($\Delta x = \Delta y = \Delta z = \Delta$) is used in all of our 3D FDTD computations.

Gedney [61] developed a PML absorbing boundary condition with non-split-field representation, i.e. the uniaxial PML, and extended it for lossy medium. This technique is implemented in our 3D vector reconstruction algorithm and its formulation is presented briefly below.

Similar to the implementation in 2D cases, the PML in 3D space requires 26 PML slabs to terminate the cubic FDTD grid in all directions. With the stretching coordinate

notations, the central cube is the working volume with coordinate stretching coefficients $s_x = s_y = s_z = 1$. In order to match the media at all interfaces, the stretching coordinate coefficients at the non-perpendicular direction in all surface PML slabs must be identical to those of the working volume, i.e. 1; the coefficients in the perpendicular direction are an increasing function when stepping away from the interface. The stretching coordinate coefficients [35, 62] in the edge and corner slabs have similar characteristics. A sample setting of the PML slabs is depicted in Figure 5.6.

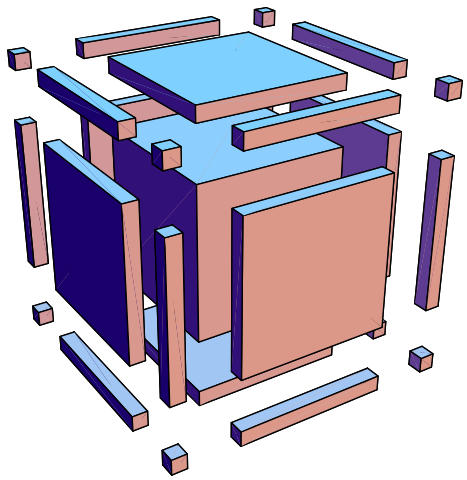


Figure 5.6: Configuration of the 3D UPML absorbing boundary condition (the surface, edge and corner slabs were positioned slightly away from the working volume, i.e. the center cube, to illustrate their spatial positions).

The stretching coordinate expression of Maxwell's equations (3.3) for all subzones in Figure 5.6 can be rewritten in a concise form as [188]:

$$\begin{aligned} \nabla \times \vec{E}(\vec{r}) &= -j\omega\mu\bar{\bar{s}}\vec{H}(\vec{r}) \\ \nabla \times \vec{H}(\vec{r}) &= j\omega\epsilon\bar{\bar{s}}\vec{E}(\vec{r}) \end{aligned} \tag{5.9}$$

where $\bar{\bar{s}}$ is the stretching coefficient tensor defined by

$$\bar{\bar{s}} = \begin{pmatrix} \frac{s_y s_z}{s_x} & 0 & 0 \\ 0 & \frac{s_z s_x}{s_y} & 0 \\ 0 & 0 & \frac{s_x s_y}{s_z} \end{pmatrix} \tag{5.10}$$

s_x,s_y and s_z can be chosen from a variety of forms including (3.68). The following expression is used by Gedney:

$$s_\xi(\xi) = \kappa_\xi(\xi) + \frac{\sigma_\xi(\xi)}{j\omega\epsilon} \tag{5.11}$$

where $\xi = x, y, z$ and $\kappa_{\xi 0}(\xi)$ and $\sigma_\xi(\xi)$ are defined by

$$\begin{aligned} \kappa_{\xi 0}(\xi) &= 1 + \kappa_{max}\left(\frac{\xi}{\Delta_\xi}\right)^m \\ \sigma_\xi(\xi) &= \sigma_{max}\left(\frac{\xi}{\Delta_\xi}\right)^m \end{aligned} \tag{5.12}$$

with κ_{max},σ_{max} and m being the parameters.

For the lossy case, the first equation in (5.9) is transformed into time domain and expanded as [62]

$$\begin{pmatrix} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \end{pmatrix} = \frac{\partial}{\partial t}\begin{pmatrix} \kappa_y & 0 & 0 \\ 0 & \kappa_z & 0 \\ 0 & 0 & \kappa_x \end{pmatrix}\begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} + \frac{1}{\mu_0}\begin{pmatrix} \sigma_y & 0 & 0 \\ 0 & \sigma_z & 0 \\ 0 & 0 & \sigma_x \end{pmatrix}\begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} \tag{5.13}$$

with

$$\begin{aligned} \frac{\partial}{\partial t}(\kappa_x B_x) + \frac{\sigma_x}{\epsilon_0}B_x &= \frac{\partial}{\partial t}(\kappa_z H_x) + \frac{\sigma_z}{\epsilon_0}H_x \\ \frac{\partial}{\partial t}(\kappa_y B_y) + \frac{\sigma_y}{\epsilon_0}B_y &= \frac{\partial}{\partial t}(\kappa_x H_y) + \frac{\sigma_x}{\epsilon_0}H_y \\ \frac{\partial}{\partial t}(\kappa_z B_z) + \frac{\sigma_z}{\epsilon_0}B_z &= \frac{\partial}{\partial t}(\kappa_y H_z) + \frac{\sigma_y}{\epsilon_0}H_z \end{aligned} \tag{5.14}$$

where

$$\begin{aligned} B_x &= \mu\frac{s_z}{s_x}H_x \\ B_y &= \mu\frac{s_x}{s_y}H_y \\ B_z &= \mu\frac{s_y}{s_z}H_z \end{aligned} \tag{5.15}$$

are the magnetic flux densities in the stretched space. Expansions of the the second equation in (5.9) yields

$$\begin{pmatrix} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \end{pmatrix} = \frac{\partial}{\partial t}\epsilon_0\epsilon_r \begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} + \sigma \begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} \tag{5.16}$$

$$\frac{\partial}{\partial t}\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \frac{\partial}{\partial t}\begin{pmatrix} \kappa_y & 0 & 0 \\ 0 & \kappa_z & 0 \\ 0 & 0 & \kappa_x \end{pmatrix}\begin{pmatrix} Q_x \\ Q_y \\ Q_z \end{pmatrix} + \frac{1}{\epsilon_0}\begin{pmatrix} \sigma_y & 0 & 0 \\ 0 & \sigma_z & 0 \\ 0 & 0 & \sigma_x \end{pmatrix}\begin{pmatrix} Q_x \\ Q_y \\ Q_z \end{pmatrix} \tag{5.17}$$

$$\begin{aligned} \frac{\partial}{\partial t}(\kappa_x Q_x) + \frac{\sigma_x}{\epsilon_0}Q_x &= \frac{\partial}{\partial t}(\kappa_z E_x) + \frac{\sigma_z}{\epsilon_0}E_x \\ \frac{\partial}{\partial t}(\kappa_y Q_y) + \frac{\sigma_y}{\epsilon_0}Q_y &= \frac{\partial}{\partial t}(\kappa_x E_y) + \frac{\sigma_x}{\epsilon_0}E_y \\ \frac{\partial}{\partial t}(\kappa_z Q_z) + \frac{\sigma_z}{\epsilon_0}Q_z &= \frac{\partial}{\partial t}(\kappa_y E_z) + \frac{\sigma_y}{\epsilon_0}E_z \end{aligned} \tag{5.18}$$

where

$$\begin{aligned} P_x &= s_y E_x \\ P_y &= s_z E_y \\ P_z &= s_x E_z \\ Q_x &= \frac{s_z}{s_x}E_x \\ Q_y &= \frac{s_x}{s_y}E_y \\ Q_z &= \frac{s_y}{s_z}E_z \end{aligned} \tag{5.19}$$

are auxiliary variables.

Applying central differences in both time and space, the discretized update scheme for H_x is a two-step process as

$$\begin{aligned} B_x^{n+1/2}(p_C) &= cABy(p_C)B_x^{n-1/2}(p_C) \\ &- cBBy(p_C)\left(\frac{E_z^n(p_R) - E_z^n(p_L)}{\Delta y} - \frac{E_y^n(p_T) - E_y^n(p_B)}{\Delta z}\right) \\ H_x^{n+1/2}(p_C) &= cAHz(p_C)H_x^{n-1/2}(p_C) \\ &- cBHz(p_C)\left(cCHx(p_C)B_x^{n+1/2}(p_C) - cDHx(p_C)B_x^{n+1/2}(p_C)\right) \end{aligned} \tag{5.20}$$

where

$$\begin{aligned} cCH\xi(p) &= \kappa_\xi(p) - \frac{\sigma^*_\xi(p)\Delta t}{2\epsilon_0} \\ cDH\xi(p) &= \kappa_\xi(p) + \frac{\sigma^*_\xi(p)\Delta t}{2\epsilon_0} \\ cAB\xi(p) &= \frac{cCH\xi(p)}{cDH\xi(p)} \\ cBB\xi(p) &= \frac{1}{cDH\xi(p)} \\ cAH\xi(p) &= cAB\xi(p) \\ cBH\xi(p) &= \frac{1}{\mu(p)\Delta t} cBB\xi(p) \end{aligned} \tag{5.21}$$

are coefficients where $\xi = x, y, z$ and spatial points p_C, p_L, etc. are marked on Figure 5.7. Similarly, we can derive the update equations for E_x from (5.16), (5.17) and (5.18) as

$$\begin{aligned} P_x^{n+1/2}(p_{CC}) &= cAP(p_{CC})P_x^{n-1/2}(p_{CC}) \\ &- cBP(p_{CC})\left(\frac{H_z^{n+1/2}(p_N) - H_z^{n+1/2}(p_S)}{\Delta y} - \frac{H_y^{n+1/2}(p_W) - H_y^{n+1/2}(p_E)}{\Delta z}\right) \\ Q_x^{n+1/2}(p_{CC}) &= cAEy(p_{CC})Q_x^{n-1/2}(p_{CC}) \\ &- cBEy(p_{CC})\left(P_x^{n+1/2}(p_{CC}) - P_x^{n+1/2}(p_{CC})\right) \\ E_x^{n+1/2}(p_{CC}) &= cAEz(p_{CC})E_x^{n-1/2}(p_{CC}) \\ &- cBEz(p_{CC})\left(cCEx(p_{CC})Q_x^{n+1/2}(p_{CC}) - cDEx(p_{CC})Q_x^{n+1/2}(p_{CC})\right) \end{aligned} \tag{5.22}$$

where

$$\begin{aligned} cAP(p) &= \frac{\epsilon(p)/\Delta t - \sigma/2}{\epsilon(p)/\Delta t + \sigma/2} \\ cBP(p) &= \frac{1}{\epsilon(p)/\Delta t + \sigma/2} \\ cCE\xi(p) &= \kappa_\xi(p) - \frac{\sigma_\xi(p)\Delta t}{2\epsilon_0} \\ cDE\xi(p) &= \kappa_\xi(p) + \frac{\sigma_\xi(p)\Delta t}{2\epsilon_0} \\ cAE\xi(p) &= \frac{cCEx(p)}{cDE\xi(p)} \\ cBE\xi(p) &= \frac{1}{cDE\xi(p)} \end{aligned} \tag{5.23}$$

are update coefficients. Analogously to the derivations of the equation for H_x and E_x in (5.20) and (5.22), respectively, the corresponding y and z components of the fields can be easily obtained by rotating the subscripts, i.e. $x \to y \to z \to x$, along with the relative positions of the points on Figure 5.7. To apply the PEC/PMC boundary condition at the exterior of the grid, one simply needs to leave these fields un-updated after initializing them with zeros before the time stepping. Note that most of the update

coefficients in (5.21) and (5.23) can be stored in 1D arrays which provides significant memory savings.

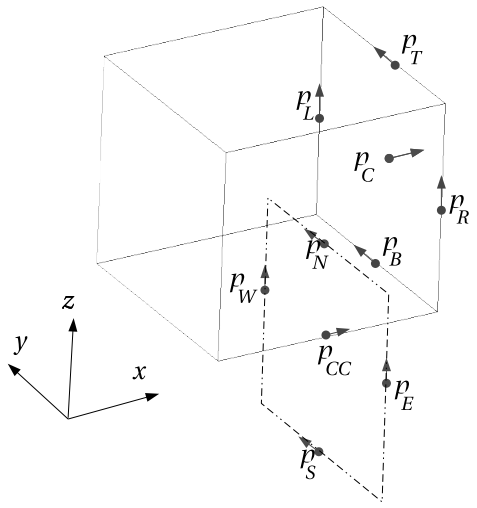


Figure 5.7: EM field vector positions for deriving the update relationships in the 3D FDTD method.

The implementation of the current source, the amplitude/phase extraction and the source low-pass filtering are also applied in similar manners as those in the 2D cases (refer to Chapter 3 for details). The CFL stability condition for the 3D FDTD method (coupled with the UPML ABC) can be written as

$$\Delta t \leq \frac{1}{c_{max}\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}+\frac{1}{\Delta z^2}}} \tag{5.24}$$

and the CFL number (CFLN) is defined by

$$CFLN = c_{max}\Delta t\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}+\frac{1}{\Delta z^2}} \tag{5.25}$$

Details of the numerical dispersion analysis for the 3D FDTD method can be found

in [189].

5.1.4 Accuracy of the 3D FDTD solver for lossy media

We computed the steady-state frequency-domain solution of the radiation field of an infinitely small z-oriented dipole antenna with the 3D FDTD approach is outlined above. The background medium has electrical properties $\epsilon_r = 22.85$ and $\sigma = 1.02$ to simulate an 83% glycerin solution. A circular receiver antenna array included 15 receivers and one transmitter equally spaced on an 15.2 cm diameter circle (as shown in Figure 5.1 (b)) with an operating frequency of $f = 1100$ MHz.

The analytical solution of the radiation fields for this configuration is derived in [87]. The electric and magnetic field components in spherical coordinates are expressed as

$$\begin{aligned} E_r &= \frac{I_0 L_0}{j\omega\epsilon}\frac{\exp(-jkr)}{4\pi}\left(\frac{j2k}{r^2}+\frac{2}{r^3}\right)\cos(\theta) \\ E_\theta &= \frac{I_0 L_0}{j\omega\epsilon}\frac{\exp(-jkr)}{4\pi}\left(-\frac{k^2}{r}+\frac{jk}{r^2}+\frac{1}{r^3}\right)\sin(\theta) \\ E_\phi &= 0 \\ H_r &= H_\theta = 0 \\ H_\phi &= I_0 L_0 \frac{\exp(-jkr)}{4\pi}\left(\frac{jk}{r^2}+\frac{1}{r^3}\right)\sin(\theta) \end{aligned} \tag{5.26}$$

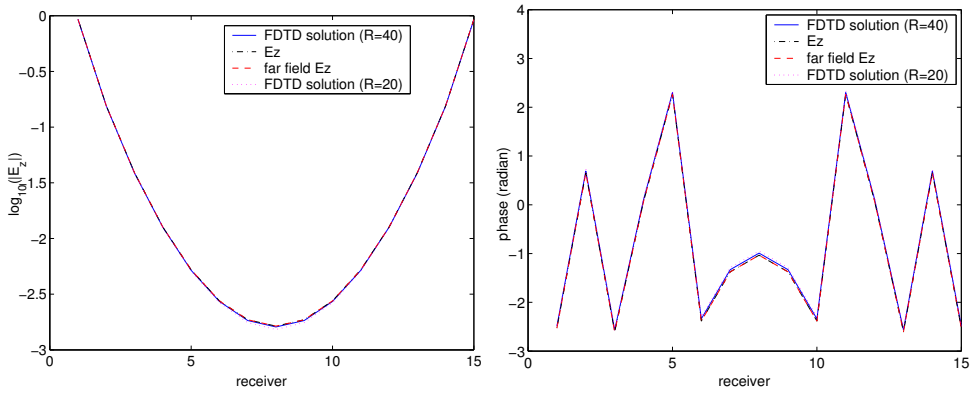
where I_0 is the current and L_0 is the length of the dipole, r, θ, ϕ are the spherical coordinates, k is the complex wave number and ω is angular frequency.

In the far field zone, the non-zero field components are given by

$$\begin{aligned} E_\theta &= jI_0 L_0 \frac{\exp(-jkr)}{4\pi r}\omega\mu\sin(\theta) \\ H_\phi &= jI_0 L_0 \frac{\exp(-jkr)}{4\pi r}k\sin(\theta) \end{aligned} \tag{5.27}$$

The amplitudes and phases at the receivers computed from the 3D FDTD method (with the source in form of equation (3.45)) utilizing two mesh densities, i.e. $R = 20$ and $R = 40$ where R is the number of nodes per wavelength, are compared with that from the previous analytical solutions (5.26) and (5.27) in Figure 5.8. From the curves in

Figure 5.8, the numerical and the analytical models match quite well in both amplitude and phase, especially for the high-density mesh case. In fact, the analytical solution in (5.26) and its far field approximation are almost indistinguishable in the plot.



(a) (b)

Figure 5.8: Comparison between FDTD solutions with analytical solutions: (a) amplitude and (b) phase.

In Appendix C, we model the radiation field distribution inside the illumination tank using this 3D FDTD forward technique under more realistic settings. These settings include the plastic tank walls, the air gap between the coupling medium surface and top wall and the air outside the tank. We demonstrate the negligible effect of the presence of the walls in such a lossy environment and validate our approach of treating the forward modelling as an unbounded problem for our imaging setting.

5.1.5 Computational complexity comparison to 3D FE/BE method

Similar to the analysis in Section 3.6.1, we compare the computational complexity of the 3D FDTD method and 3D FE method for the forward solution in this subsection.

A uniform grid with $N_x = N_y = N_z = N$ is used for both methods while in the FE mesh, each cube is split into 3 tetrahedral elements. The total node number for both

meshes is N^3. In assembling the FE matrix for the simplified scalar model (Equation (3.13)), the size of the matrix is $N^3 \times N^3$. The minimum half-bandwidth for finite element approach is roughly N^2 when numbering the nodes sequentially in each layer. If a BE matrix is incorporated to account for the boundary condition, the half-bandwidth increases to roughly $6N^2$ which is essentially the number of the boundary nodes. Thus, to solve this matrix equation with the Cholesky factorization algorithm, the total flop count for FE/BE hybrid approach is $36N^7 + 42N^5 + 2N^3$ while that for FE method with absorbing boundary conditions is $N^7 + 7N^5 + 2N^3$.

The total flop count for obtaining a 3D FDTD steady-state solution has a similar expression to that in (3.90). Summing the float operations in (5.20) and (5.22) and multiplying by 3 to account for the components in all directions, the flop number for a single iteration is $F_{iter} = 84N^3$. The expression for F_{steady} is once again approximated by the number of time-steps required for round-trip time-step of the radiation wave which leads to the identical result as that in the 2D case, i.e. (3.91). Consequently, the total flop count for 3D FDTD method (with UPML for lossy medium) is given by

$$F_{FDTD} = 168\sqrt{3}N^4 \frac{c_{max}}{CFLN \times c_{bk}} \tag{5.28}$$

A plot of the total flop number at various N values is shown in Figure 5.9 where $\frac{c_{max}}{c_{bk}} = \sqrt{10}$ and $CFLN \approx 1$ are used in the calculation. From the plot, the computational advantage of 3D FDTD method compared to the FE method is even more significant than for the 2D cases (see Figure 3.16).

Similar to the comparison in the 2D case (Section 3.6.1), the actual computational efficiency of the FE/BE method with our reconstruction settings is not as bad as in this example. Additionally, the implementation of iterative solvers in solving FE/BE equation can also significantly reduce the computational expense of this approach. In

Section 5.3, we list the forward field computation time for a range of methods.

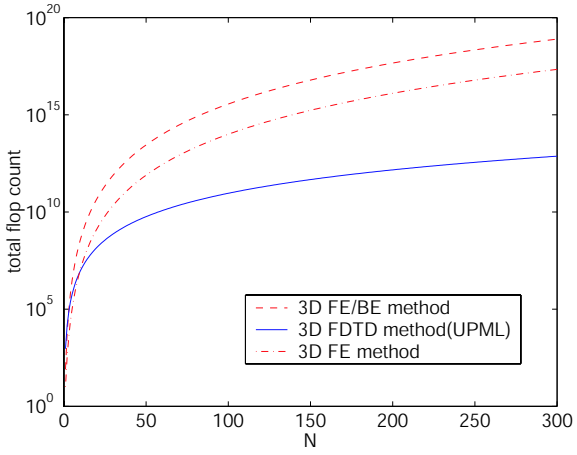


Figure 5.9: Comparison of the total floating-point operation counts between the 3D FE/BE and 3D FDTD methods for different mesh sizes.

Nodal-adjoint approach for the 3D FDTD method

Given the derivations in Section 5.1.2, the nodal adjoint formulation for 3Dv/3D method is straightforward to construct. In the 3D FDTD grid, the effective volume V_n for all interior forward nodes are identical, which is the volume of a single voxel, i.e $V_n = \Delta x \Delta y \Delta z$. The nodal adjoint formula in this case is correspondingly written as

$$J((s,r),\tau) = -\sum_{n\in\Omega_\tau} (\Delta x \times \Delta y \times \Delta z) \times \varphi_\tau(\vec{p}_n) \times E_s(\vec{p}_n) \times E_r(\vec{p}_n) \tag{5.29}$$

As a simple extension from (5.29), for the 2D dual-mesh reconstructions using FDTD as forward solver, the adjoint formula is easily computed by

$$J((s,r),\tau) = -\sum_{n\in\Omega_\tau} (\Delta x \times \Delta y) \times \varphi_\tau(\vec{p}_n) \times E_s(\vec{p}_n) \times E_r(\vec{p}_n) \tag{5.30}$$

where $\Delta x \times \Delta y$ is the area of a 2D FDTD cell. Equation (5.30) is the actual method used for all $2Ds^{FDTD}$/2D reconstructions in Chapter 3.

5.1.6 Enhancement of the 3D vector forward solver

High order spatial difference scheme

In the derivations of the FDTD update equations, a central difference scheme was used in both time and space which provided 2nd-order accuracy. Higher order difference schemes were investigated by Turkel [194, 193], Fang [47] and others in order to improve accuracy or equivalently reduce the problem size. The studies into high-order FDTD methods is an active topic in FDTD research.

A Ty(2,4) method proposed by Turkel *et al.* [194](i.e. implicit 2nd-order in time 4th-order in space) is implemented in our reconstructions along with the UPML FDTD algorithm. The implementation of this method is quite simple. The basic idea is to replace the spatial derivative terms in the curl operator (equation (5.13) and (5.16)) by 4th order implicit difference representations. Denoting the 4th order accuracy difference operator as D_4 representing the discretization form of $\frac{\partial}{\partial \xi}$, we have the following implicit relationship for the discretized difference at the neighboring nodes [193]:

$$\frac{D_4 u_{i+1}^n - D_4 u_{i-1}^n}{24} + \frac{11}{12} D_4 u_i^n = \frac{u_{i+1/2}^n - u_{i-1/2}^n}{\Delta \xi} \tag{5.31}$$

where $\xi = x, y, z$; n is the time step and u can be any one of the E or H components. Applying this implicit relationship to every spatial derivative term in the LHS of (5.13) and (5.16), we get 12 matrix equations with the following form

$$A \begin{pmatrix} D_4 u_{1/2} \\ D_4 u_{3/2} \\ \vdots \\ D_4 u_{(2N_\xi - 1)/2} \end{pmatrix} = \frac{1}{\Delta \xi} \left(\begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ u_{N_\xi + 1} \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N_\xi} \end{pmatrix} \right) \tag{5.32}$$

where

$$A = \frac{1}{24}\begin{pmatrix} 26 & -5 & 4 & -1 & \cdots & 0 \\ 1 & 22 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 22 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdots & 0 & 1 & 22 & 1 \\ 0 & \cdots & -1 & 4 & -5 & 26 \end{pmatrix} \tag{5.33}$$

is a matrix of size $N_\xi \times N_\xi$ (notice that fourth order backward and forward differences are used at the two end points). To solve for the 4th order differences from (5.32), the LU decomposition of A matrix is analytically obtained and can be pre-computed and stored before the time stepping. Twelve back-substitution processes need to be performed to construct all spatial derivatives in (5.13) and (5.16) during each iteration. Consequently, the spatial central difference terms in update equations (5.20) and (5.22) are substituted by the solutions to (5.32) for all field components.

The implementation of this algorithm successfully reduces the forward mesh to 1/8 of its original size while yielding similar accuracy. However, proportional time savings are not achieved due to the operations for the back-substitutions at each time step. The total flop number per time-step in the Ty(2,4) scheme is roughly 4 times more than that of the 2nd-order method. In general, the high-order methods are useful for improving the forward problem accuracy, but the improvement in terms of computational efficiency is not significant.

Computation time improvements by supplying initial fields

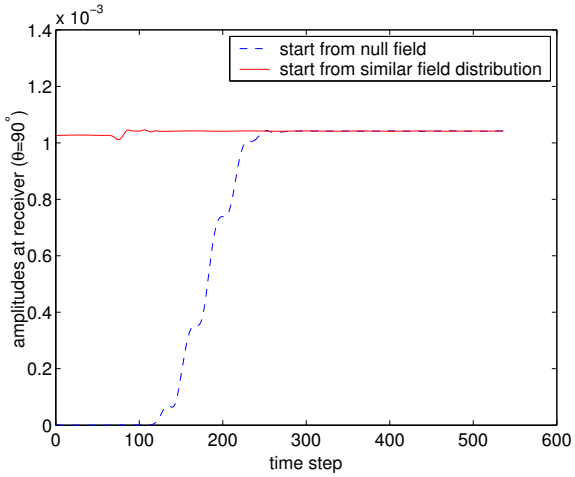
In the floating-point operation analysis of the FDTD method, the total flop number for the FDTD is proportional to the time steps for reaching steady-state. We have found that the steady-state time steps F_{steady} is related to the initial value of the field. If the FDTD time-stepping starts from a null field distribution (i.e. all components are zeros),

it takes longer to reach steady state than from a field distribution that resembles the final solution.

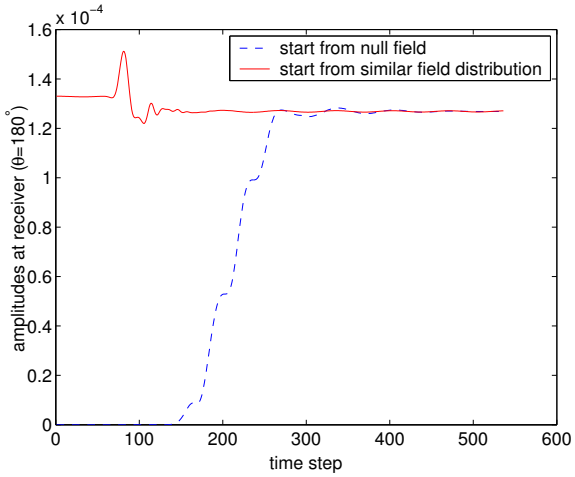
A simple 2D forward problem is computed to illustrate this finding. A 2.5 cm ×2.5 cm square dielectric object is located at the center of the antenna array whose properties are $\epsilon_r = 10$ and $\sigma = 0.5$ S/m and those of the background are 25 and 1.0 S/m, respectively. Utilizing polar coordinates, and the transmitter operating at $f = 900$ MHz and located at ($r = 7.6$ cm, $\theta = 0°$), the amplitudes of the receivers at $\theta = 90°$ and $\theta = 180°$ are recorded and plotted vs. time step in Figure 5.10 in comparison to the responses computed from the initial values of a similar field distribution, i.e. the fields due to the presence of a similar sized object that has $\epsilon_r = 12, \sigma = 0.7$ S/m. In both computations, the time step Δt is set to 1.64e-11s to ensure stability. From the plot, it is obvious that the second approach leads to significantly fewer time steps to achieve steady-state. It should be noted that the sharp oscillations in the solid lines are referred as spurious modes induced by sudden changes in the dielectric properties.

Moreover, in an iterative reconstruction process, the update of the parameters between successive iterations becomes smaller, resulting in increasingly similar field distributions as the iterative process advances. Therefore, the field distributions at the final time step of the previous iteration are good starting points for the subsequent iteration. Utilizing this finding, we derived an iterative FDTD approach in conjunction with the iterative reconstruction process to reduce the forward modelling time.

The implementation of this scheme is quite simple. Extra memory is required to store all field components and the accumulated elapsed time at the end of each iteration for each source. At the subsequent iteration, the fields are initialized by the stored fields from the previous iteration of the corresponding source. Then the fields start updating as part of the FDTD process until achieving steady-state, at which point the



(a)



(b)

Figure 5.10: Amplitudes at different time-steps for receivers located at (a) $\theta = 90°$ and (b) $\theta = 180°$.

above process is repeated. We shall demonstrate in Section 5.1.6 that it is possible to reduce the steady-state time step number to 1/2 to 1/3 of the original round-trip time step by supplying initial field estimates while not compromising convergence and image quality.

From a wave point of view, when a source is located close to the boundaries of the domain, the time to reach steady state takes approximately twice as long as when the excitation is located at the center of the domain because of the increased averaged distance between the source and the receivers. For our microwave imaging configuration, the unknown object is typically located at the center of the imaging zone. Therefore, the perturbed EM wave induced by the parameter updates during the iterations propagates to the receiver in less time than from the application of excitations near the borders of the reconstruction domain. This may partially explain the time step number reduction from supplying an initial field estimate during the iterative FDTD approach.

One significant impact of utilizing this scheme is on computing Δt. In our previous 2D and 3D algorithms, the value of Δt is determined dynamically by the minimum values of the permittivity and conductivity by (3.84) or (5.24) at each iteration. In reconstructions where the object has a lower permittivity value than the background, the value of Δt typically drops at each iteration in coordination with the recovery of the object. When applying the iterative FDTD approach in cases with a dynamic Δt, the spurious waves observed in Figure 5.10 become more severe and degrade the quality of the forward field. To avoid these spurious waves, utilization of a constant time-step throughout all iterations is important and the iterative FDTD process must be tailored accordingly. The minimum permittivity ϵ_{min} should be estimated before the reconstruction process so that Δt for all iterations can be determined from (5.24). At each iteration, we need to compare the updated permittivity values with ϵ_{min} and set ϵ_{min} as the lower

bound to ensure stability. Using this approach, the number of time steps for the first several iterations is greater than that for the original method; however, overall, the acceleration by providing an initial field distribution makes the iterative FDTD approach faster. Examples are demonstrated in Section 5.3.

5.1.7 ADI FDTD with lossy UPML absorbing boundary condition

As was been demonstrated in the previous subsection, utilizing an initial field distribution estimate can reduce the steady-state time step number. However, to meet the stability criteria, a constant time-step needs to be used which results in the computational redundancy in the first few iterations. In this subsection, we derive an unconditionally stable FDTD scheme, the alternative-directional-implicit (ADI) FDTD method, for forward field modelling in order to avoid the time step redundancy.

The ADI FDTD technique was initially proposed by Zheng *et al.* in 1999 and independently by Namiki in 2000. The general description and formulation of this technique can be found in [209, 140]. It has been extended to the lossy case by Lazzi *et al.* [106], to the PML absorbing boundary condition case by Liu *et al.* [114] and Wang *et al.* [200] and the UPML case by Zhao[208, 207]. However the ADI formulation for both lossy medium and UPML ABC has not previously been discussed in the literature to the best of our knowledge.

In this situation, for each time step of the ADI FDTD method, field updates are computed utilizing two sub-steps: 1) compute all field components at time step $n + 1/2$ from the field distributions of time step n where the second spatial difference terms in the discretized curl equations, i.e. (5.20) and (5.22), use the fields at $n + 1/2$; step 2), compute the fields at $n + 1$ from time step $n + 1/2$ where the first spatial difference terms use the field values at time step $n + 1$. In both sub-steps, the target time step

fields appear on both sides of the update equation; thus, this method yields an implicit difference update scheme. Based on this principle, the ADI form of the UPML update equation for lossy media is not difficult to derive. We used a symbolic software package, Mathematica™, to perform the derivations. The Mathematica codes for both sub-steps can be found in Appendix A.

With this ADI technique, the time step size Δt is not constrained by the CFL stability condition (5.24); instead, the dispersion error becomes the major factor that limits Δt. A detailed study on the impact on the dispersion error due to various Δt's in the ADI FDTD is given by Zhao [206]. The unconditional stability of the ADI FDTD allows for simultaneous use with the iterative FDTD approach introduced in the previous subsection. A few reconstruction examples are presented in the results section.

To estimate the computational efficiency, the total flop number needs to be calculated for this method. Assuming the 3D grid size is $N_x = N_y = N_z = N$, the floating-point operations per iteration for the ADI approach can be written as

$$F_{iter}^{ADI} \approx 2N^3(177 + 5 + 66) \tag{5.34}$$

where the number "2" is due to the two sub-steps, "177" is derived from the operations to assemble the RHS for the tri-diagonal equations for $\vec{P}$, the "5" comes from the number of back substitutions required to solve for the tri-diagonal equations and the "66" relates to the contributions of the remaining update equations. The total number of time steps to reach steady-state for the ADI FDTD method can now be written as

$$F_{steady}^{ADI} = \frac{CFLN}{CFLN_{ADI}} F_{steady} \tag{5.35}$$

where F_{steady} and $CFLN$ are the steady time step and CFL number defined in (3.91)

and (5.25), respectively. Combining (5.34) and (5.35), we get the total flop count for the ADI FDTD with lossy UPML ABC as

$$F_{ADI} = F_{steady}^{ADI} F_{iter}^{ADI} = 992\sqrt{3}N^4 \frac{c_{max}}{CFLN_{ADI} \times c_{bk}} \tag{5.36}$$

From (5.36) and (5.28), in order to achieve faster computations, the CFL number for the ADI FDTD should be at least 6 times that in the traditional FDTD.

5.2 3D microwave imaging system prototype

The illumination array for the new 3D data acquisition system consists of a 25 cm diameter by 19.3 cm height Plexiglas cylinder and is shown in Figure 5.11. The 16 monopole antenna array is positioned on a 15 cm diameter concentric circle with the antennas capable of traveling vertically a total of 11.4 cm. Note that we still utilize the technique of allowing the rigid coaxes feeding the monopole antennas to pass through hydraulic seals in the base of the tank facilitating partitioning of all the motors, array plates and coaxial cables outside of the tank. The active part of the antenna consists of a 3.8 cm length of exposed coaxial cable (both center conductor and surrounding dielectric). The 16 antennas are segregated into two interleaved arrays of 8 antennas with each sub-array being able to move independently from the other. The mounting plates for the antenna sets are attached to separate pairs of opposing, computer controlled motors positioned underneath the tank. In this way, a single antenna can still transmit a signal to all of the remaining 15 antennas within individual horizontal planes. However, this new arrangement now allows for a single antenna to transmit a signal that can be received by 8 of the 15 receiver channels at different vertical positions. While this does not facilitate acquisition of all cross plane transmit/receive pair permutations, the

amount of new data is considerable and the trade-offs in terms of increased array positioning complexity versus the merits of acquiring the new 3D data and possibly being able to reconstruct real 3D images were reasonable.

An additional innovation compared to the previous system involved incorporating two NI4472 24 bit, 8 channel data acquisition boards. This allowed for the parallel detection of all signals over the 140 dB dynamic range which is necessary in this system because of the broad operating frequency range (500 - 3000 MHz) and the wide range of propagation path distances between the antennas. The previous system utilized two 16 bit, 8 channel A/D boards which required a variable gain amplifier in front to achieve the desired dynamic range. In addition, the new boards actually sample all 8 channels simultaneously instead of multiplexing the channels into a single sampler circuit. These new innovations have enabled us to maintain our excellent signal dynamic range while greatly increasing the data acquisition speed. This is essential for the 3D system since the increase in the amount of data that will be collected is substantial and we will want to limit the breast exam times to as short as possible for patient comfort and to minimize the possibility of patient motion artifacts.

5.3 Results

In this section, we present reconstructions from simulated data to assess the performance of the proposed methods under ideal conditions. The computational efficiency of the enhanced algorithm, such as incorporating initial field estimates and the ADI FDTD were studied with the numerical simulations. More realistic reconstructions involving simple 3D phantoms were performed with measurement data collected from our 3D data acquisition system. Finally, we studied all five dual-mesh based algo-

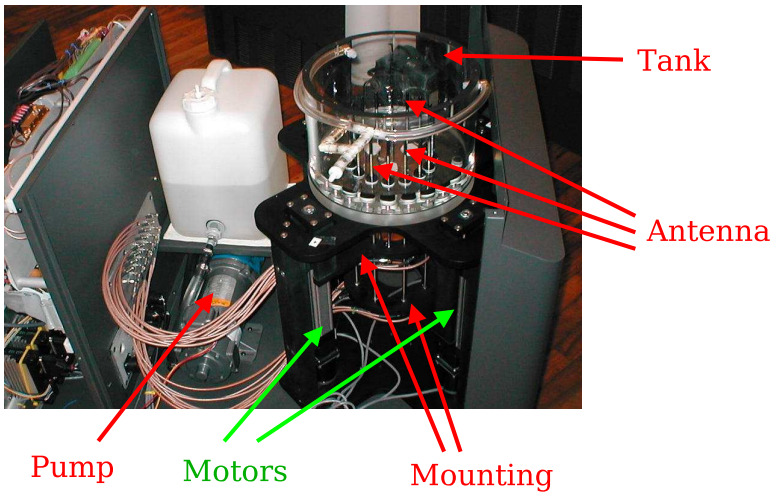


Figure 5.11: Photograph of the new illumination tank indicating the interleaved antenna sub-arrays with the mounting plates and linear actuator motors.

rithms, as listed in Table 3.2 including the 2D and 3D algorithms, with benchmark reconstructions to profile their computational complexity with respect to the incremental improvements in forward modelling accuracy.

In order to make reasonable comparisons across these approaches, we used a group of common settings for all experiments unless otherwise noted. In the simulation cases, the background medium used a 0.9% saline solution having $\epsilon_r = 77$ and $\sigma = 1.7$ S/m. The cylindrical reconstruction meshes for the two 3D methods are identical, comprised of 1660 nodes and 7808 tetrahedron elements. In this case, the Cartesian coordinate system origin is located at the center of the reconstruction mesh with the z-axis aligned along cylinder axis. A circular antenna array located on a radius $r = 7.62$ cm comprised of 16 equally spaced monopole antennas operating at 900 MHz is placed on the $x - y$ plane centered at the origin. Each individual antenna is modelled by an infinitely small z-oriented dipole. For cases where multiple layered antenna arrays are used, diagrams are provided to illustrate the positions of the antennas. For each iteration of the Gauss-Newton reconstruction, a Tikhonov regularization is imposed with the regular-

ization parameter computed by the empirical method discussed in Section 3.2.3. All reconstructions started from an initial estimate equal to the homogeneous background medium.

5.3.1 Simulated data reconstructions

The imaging target is an off-centered sphere ($\epsilon_r = 20$, $\sigma = 0.5$ S/m) with center location ($x = 0.0$ cm, $y = -2.5$ cm, $z = 0.0$ cm) and radius $r = 2$ cm. For the 3Ds/3D reconstructions, the forward mesh is a cylinder consisting of 56,636 nodes and 312,453 tetrahedral elements. The mesh has radius $r = 12$ cm and extends vertically from $z = -5$ cm to $z = 5$ cm. For the 3Dv/3D reconstructions, the interior grid has size $70 \times 70 \times 35$ nodes and is surrounded by 5 layers of a UPML (the final node size of the data array is $80 \times 80 \times 45$). The FDTD cells are cubes with uniform node spacing of $\Delta x = \Delta y = \Delta z = 2.47$ mm.

The simulated measurement data was generated using an FDTD 3D vector solution over a much finer forward mesh (40 nodes per wavelength of the background medium compared with 15 nodes per wavelength in the reconstruction problem) with the E_z components extracted at the receiver sites.

Four antenna array configurations (as illustrated in Figure 5.12) were investigated. The reconstructed 3D dielectric profiles from both scalar (for scheme A only) and vector methods are shown in Figures 5.13 to 5.17.

Several observations can be made from these images:

1. The permittivity images generally have less artifacts than their conductivity counterparts, similar to that observed in Chapter 4. This demonstrates that the imaging mechanism we are exploiting is more sensitive to permittivity variations.

2. The images reconstructed utilizing the scheme A antenna configuration from the

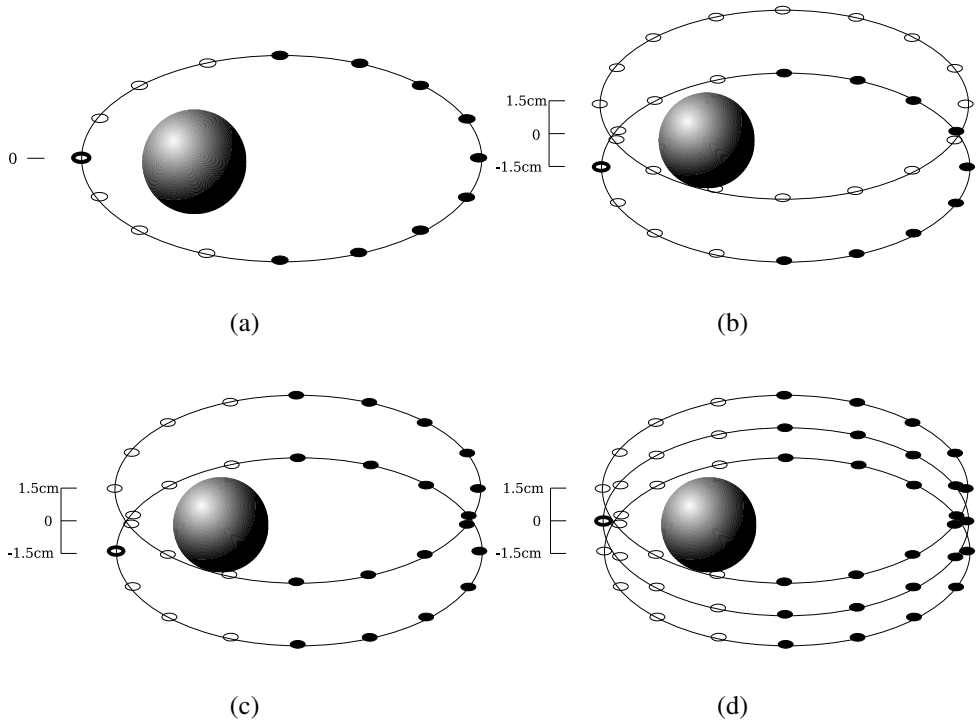


Figure 5.12: Source configurations for 3D simulation reconstructions: (a) scheme A, (b) scheme B, (c) scheme C, (d) scheme D. In each diagram, the bold circle represents a transmitter and the solid circles represent the corresponding receivers for that specific transmitter. In scheme D, only the antennas on the central plane were used as transmitters, while in the other schemes, all antennas operated as transmitters sequentially. Additionally, scheme B and C are distinguished from each other by the fact that in scheme B the receivers are only those antennas in the same plane as the transmitter while the receivers in scheme C can be in either plane with respect to the transmitter plane.

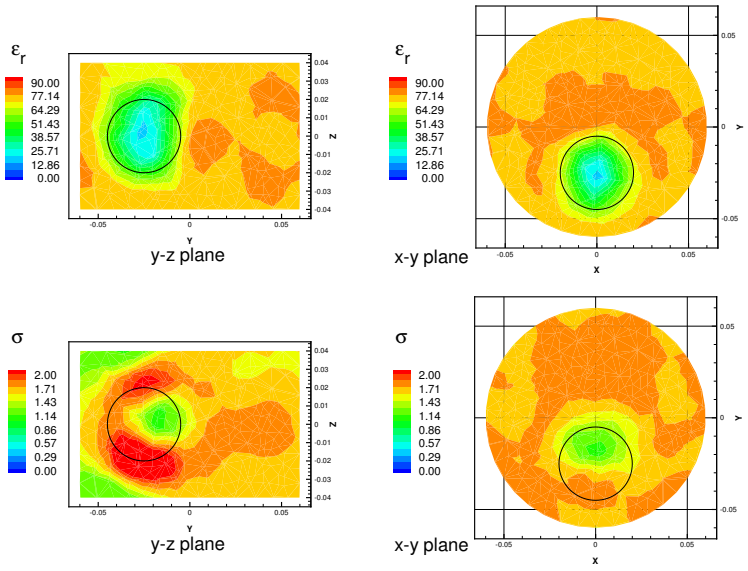


Figure 5.13: Cross-sectional images of the reconstructed dielectric profiles using the scheme A antenna configuration (3Ds/3D algorithm).

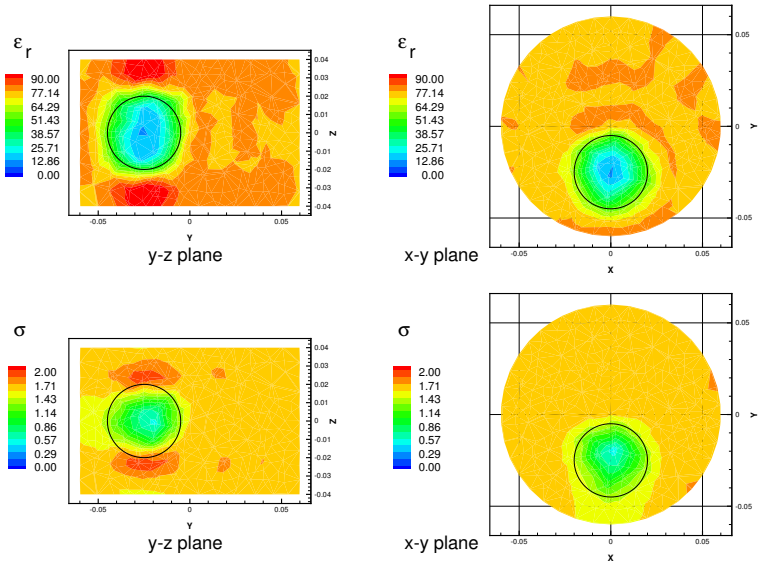


Figure 5.14: Cross-sectional images of the reconstructed dielectric profiles using the scheme A antenna configuration (3Dv/3D algorithm).

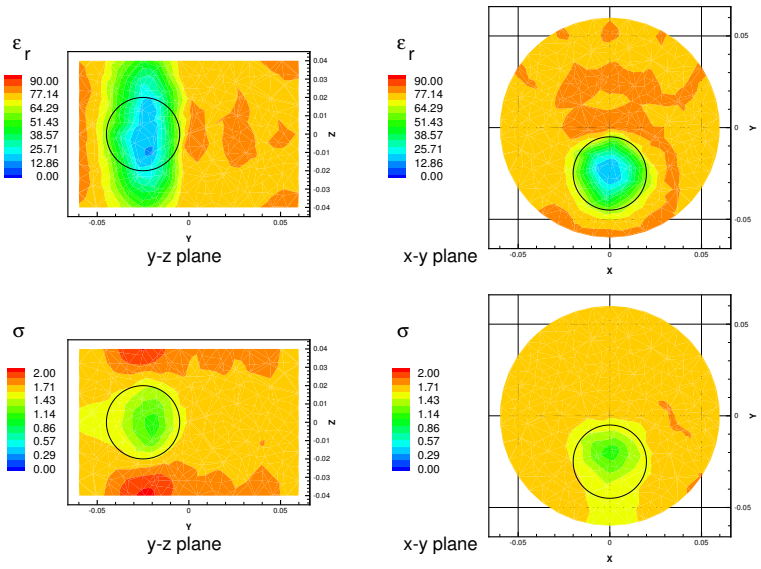


Figure 5.15: Cross-sectional images of the reconstructed dielectric profiles using the scheme B antenna configuration (3Dv/3D algorithm).

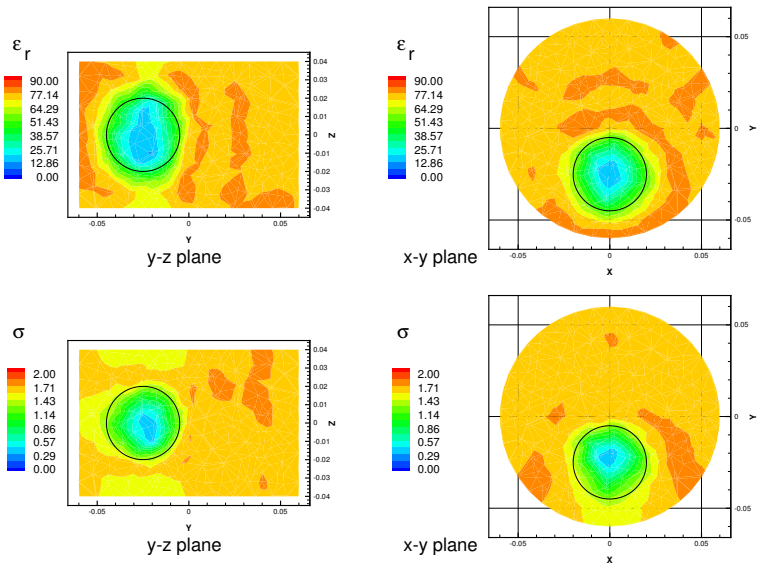


Figure 5.16: Cross-sectional images of the reconstructed dielectric profiles using the scheme C antenna configuration (3Dv/3D algorithm).

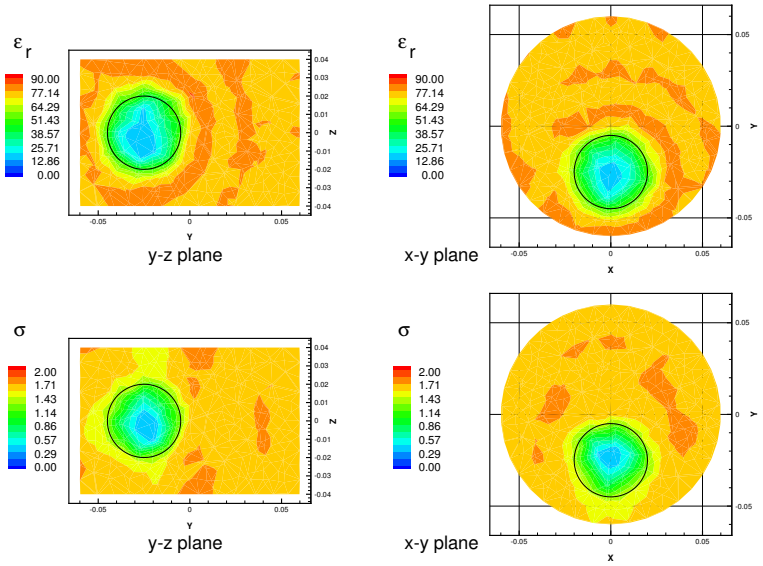


Figure 5.17: Cross-sectional images of the reconstructed dielectric profiles using the scheme D antenna configuration (3Dv/3D algorithm).

scalar method (3Ds/3D) have more artifacts above and below the object than the corresponding images produced with 3Dv/3D method. In particular, the conductivity component is overwhelmed by these artifacts and is unable to provide useful information concerning the target.

3. The permittivity contours for the single-layer antenna array are relative accurate in the plane where the antenna array is located. However, artifacts appear above and below the object. The artifacts in the conductivity image are more noticeable and distort the contour of the object in the z-direction.

4. From the images of the two-layer (32 sources, 32 transmitters and 18 receivers per transmitter) and three-layer (48 sources, 16 transmitters and 27 receivers per transmitter) antenna configurations, the conductivity image artifacts in the z-direction above and below the target are significantly reduced compared to the single plane images making the object generally appear more uniform. This

demonstrates that more measurement data especially out-of-plane data is helpful in improving the quality of the 3D reconstructions.

5. The images computed from scheme B have significant distortions in permittivity component along z-axis while the conductivity profile is not fully reconstructed. Even while the amount of measurement is doubled in this case compared with just the single-layer case, the image quality is not as good as the latter.

For a further analysis of the artifacts observed in the results using scheme B, we performed three more reconstructions. In the first reconstruction, we reduced the spacing of the two antenna arrays from 3 cm to 2 cm. The reconstructed images for this case are shown in Figure 5.18. In the second reconstruction, three antenna arrays similar to scheme D were used while only the receivers in the transmitter plane acquired measurement. The vertical space between the antenna arrays was 1 cm. The results for this experiment are shown in Figure 5.19. The third experiment utilized a five-layer antenna array with 1 cm spacing in z direction and the corresponding results are shown in Figure 5.20. Similar to the previous two examples, only planar measurement were collected. With each increase in data, the distortions in the z-axis of the permittivity image (especially above and below the recovered object) are reduced compared with the images in Figure 5.15. Significant artifacts can be observed above and below the object in the conductivity images but they too are reduced with increased data. This demonstrates that by utilizing measurements from multiple planar antenna arrays, we can recover improved 3D images. Given that there is a significant amount of multi-planar measurement data already acquired using our current clinical system (Section 3.7.4), the above finding indicates that we may be able to directly use these data set for 3D image reconstruction. Comparing the results from these cases and schemes C and D, it is evident that the cross-layer measurement is more efficient in helping the recon-

struction algorithm to recover the three-dimensional shape and location of the object, but that utilizing multiple single-plane data sets should not be discounted.

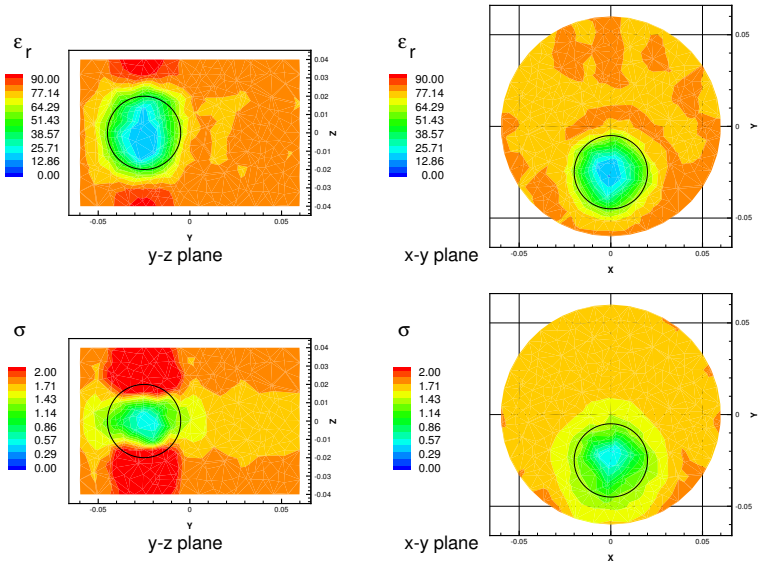


Figure 5.18: Cross-sectional images of the reconstructed dielectric profiles using two antenna arrays with 2 cm spacing.

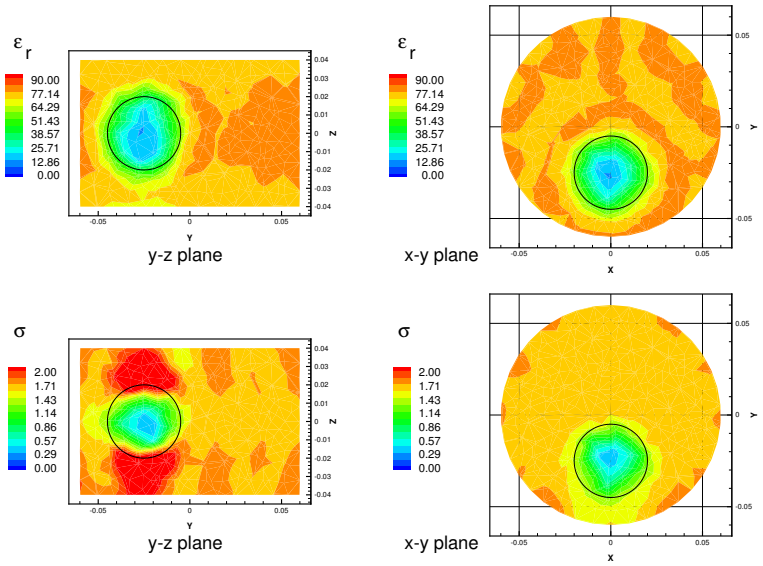


Figure 5.19: Cross-sectional images of the reconstructed dielectric profiles using three antenna arrays with 1 cm spacing.

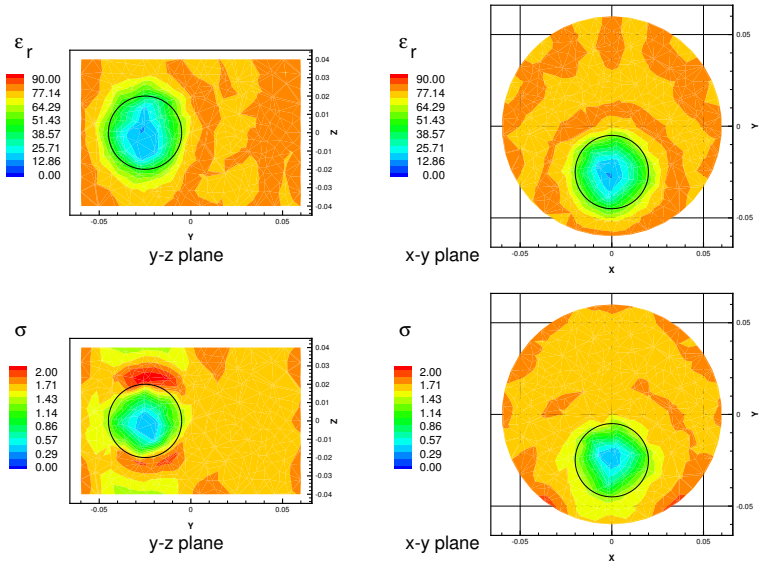


Figure 5.20: Cross-sectional images of the reconstructed dielectric profiles using five antenna arrays with 1 cm spacing.

With the simulated data for scheme A, we also investigated the performance of various enhancements discussed in Section 5.1.6. First, as predicted, the Ty(2,4) FDTD method did run slightly faster than the traditional 3D FDTD method due to the reduction in mesh size; however, the enhancement in efficiency was not significant. Second, the computation time decrease from supplying an initial field distribution was significant. For the reconstructions using this technique, we first estimated the maximum dielectric property at 1:5 with respect to the background and used $CLFN$=0.86 to compute the time step value for all iterations. The values of all field vectors and the accumulated time-steps were recorded starting from the second iteration and subsequently supplied to all iterations. Meanwhile, we deliberately reduced the steady-state time step number estimated from (3.91) by a factor of 2 or 3. Under these circumstances, the reconstructed images show no obvious degradation. The relative errors of these reduced computation time reconstructions are plotted with those from the unenhanced version in Figure 5.21 which confirms the benefits of this technique. Finally, we used

the ADI FDTD technique to evaluate the forward field computation for this example. Unfortunately, due to the small density of the 3D mesh, the flexibility for increasing the $CLFN_{ADI}$ is not large. We found that when $CLFN_{ADI} > 4$, the dispersion error in the solution significantly impacts the forward accuracy; consequently, the overall reconstruction time with the ADI FDTD method did not improve.

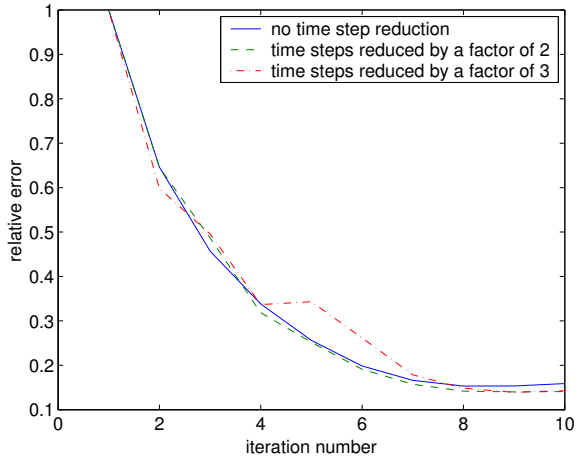


Figure 5.21: Relative error plot of the reconstructions with and without the initial field estimates.

5.3.2 Measured data reconstructions

A phantom experiment involving two spherical objects was performed using our new 3D system described in Section 5.2. The background medium is 83% glycerin and the spheres are made of solid bone/fat-like material with $\epsilon_r = 5$ and $\sigma = 0.2$ S/m. A picture of the experimental configuration is shown in Figure 5.22. In this experiment, two circular antenna arrays, Group A and Group B, were translated independently along 16 positions along the z-axis in 0.5 cm increments, resulting in $16 \times 16 = 256$ total array combinations. For each combination, each antenna transmitted a signal sequentially while the remainder acted as receivers. This experiment was repeated at frequencies from 500 MHz to 1100 MHz in 200 MHz increments. The total time for acquiring

the full set of data took roughly two and a half hours. Although the measurement data set is a large array, only a few positions worth of data at 900 MHz are used for these reconstructions. These positions are indicated in Figure 5.23.

Figure 5.22: Experimental setup for the sphere phantom measurement.

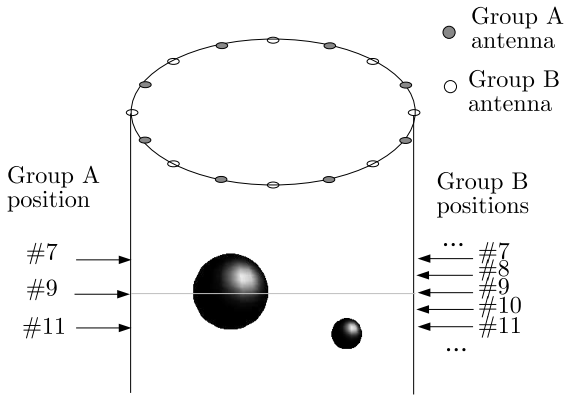


Figure 5.23: Antenna sub-group positions for the 3D phantom experiments.

Five sets of data were used for the reconstructions including

1. Scheme 1: {9,9}

2. Scheme 2: {{9,7},{9,9},{9,11}}

3. Scheme 3: {{9,7},{9,9},{9,11},{7,9},{11,9}}

4. Scheme 4: {{9,7},{9,8},{9,9},{9,10},{9,11}}

5. Scheme 5: {{9,5},{9,6},{9,7},{9,8},{9,9},{9,10},{9,11},{9,12},{9,13}}

where each pair of numbers represent the vertical position numbers for antenna group A and B, respectively. The images recovered for the 3Ds/3D and 3Dv/3D methods are shown in Figures 5.24 to 5.26 for the selected schemes.

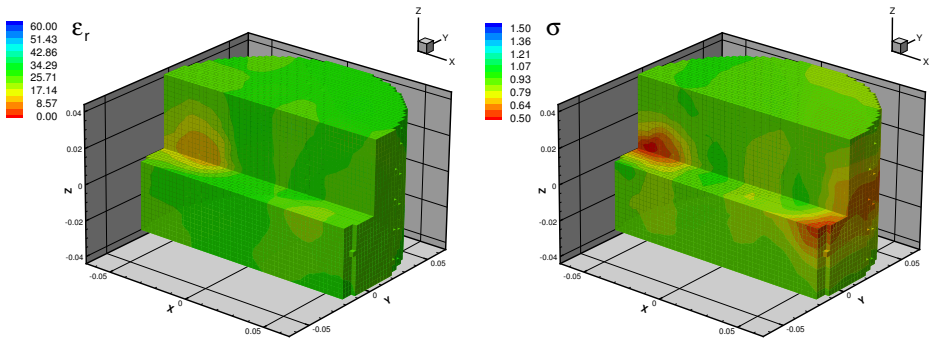


Figure 5.24: Contour slice images extracted from the results of 3D phantom experiment reconstructions utilizing antenna scheme 1.

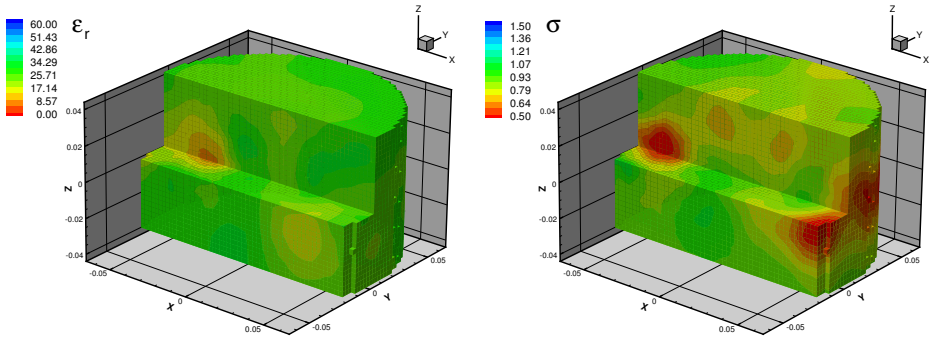


Figure 5.25: Contour slice images extracted from the results of 3D phantom experiment reconstructions utilizing antenna scheme 2.

From the reconstructed images, the recovered permittivity part of the object is more accurate than the conductivity images with the objects being shifted slightly toward the

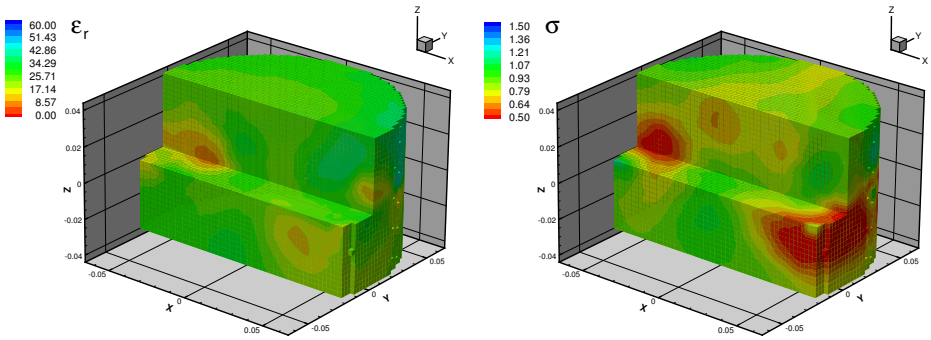


Figure 5.26: Contour slice images extracted from the results of 3D phantom experiment reconstructions utilizing antenna scheme 3.

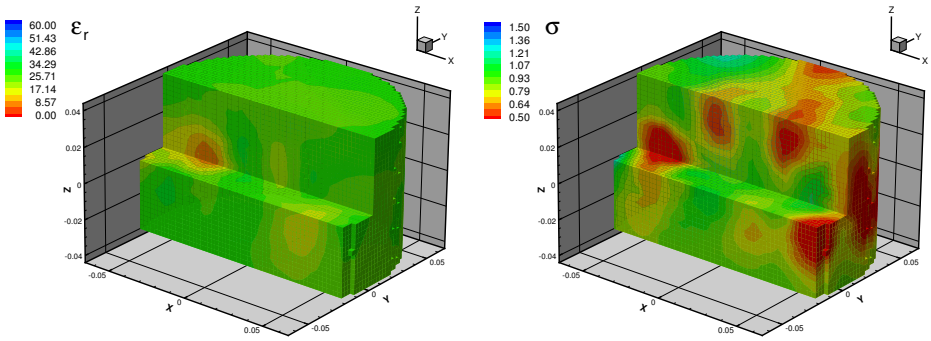


Figure 5.27: Contour slice images extracted from the results of 3D phantom experiment reconstructions utilizing antenna scheme 4.

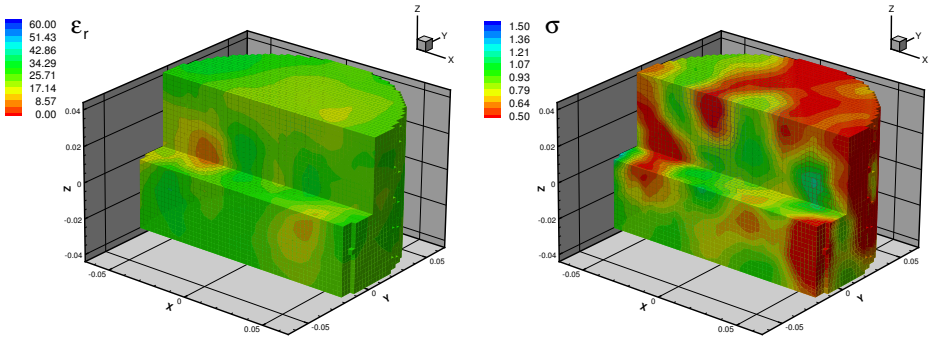


Figure 5.28: Contour slice images extracted from the results of 3D phantom experiment reconstructions utilizing antenna scheme 5.

boundary. For scheme 1, which contains only measurement data from a single slice, the large sphere was correctly reconstructed for both permittivity and conductivity images while the small one is not obvious (This is reassuring since the small sphere is not in this imaging plane). When incorporating some of the multi-slice measurement data (schemes 2 through 5), both spheres were successfully recovered on the 3D permittivity profiles. This reinforces the importance of utilizing cross-plane measurement data. Notice that the smaller sphere in the permittivity images appears as a smoothed low contrast object due to the filtering effect of the regularization. For the conductivity images, artifacts appear to become more significant with more measurement data. More analysis is required to assess why the conductivity images do not recover the object very well.

5.3.3 Comparisons of all dual-mesh based algorithms

Using the simulation reconstructions (Section 5.3.1) as a benchmark, we tested all five dual-mesh based methods and summarized the problem size and computational times in Table 5.3.3.

From Table 5.3.3, we can clearly see trends when increasing the problem size along with the transition from 2D to 3D reconstructions. In conjunction with this, due to the implementations of various techniques proposed in this and the previous chapters, such as iterative block solver, the FDTD technique and computation time enhancements associated with setting the initial fields, the forward computation time for all of these methods is within an acceptable range even for the full 3D vector approach. Additionally, when comparing the data from the first two columns, the use of the adjoint method is essential for making these reconstruction algorithms viable.

In general, the 2D algorithms in Table 5.3.3 demonstrate advantages in terms of

Table 5.1: Comparisons between dual-mesh based reconstructions

	2Ds/2D^{+}	2Ds/2D	2Ds^{FDTD}/2D	3Ds/2D
forward method	FE/BE	FE/BE	2D FDTD	3D FE
forward mesh size	3,903	3,903	12,100 17,955*	56,636
recon. mesh size	556	556	473	126
forward time/source	1.8s	1.8s	1s(0.3s**)	5.5s(13s***)
solving for update	1h	16s	2s	0.5s
	3Ds/3D	3Dv/3D	3Dv/3D^{++}	
forward method	3D FE	3D FDTD	3D FDTD	
forward mesh size	56,636	108,000 196,000*	108,000 196,000*	
recon. mesh size	1660	1660	1660	
forward time/source	5.5s(13s)	9s(1-3s**)	5s	
solving for update	4s	4s	4s	

$^{+}$ utilized sensitivity equation method to construct the Jacobian matrix
$^{++}$ used the initial field acceleration technique(factor=2)
* total unknown size including the nodes in PML slabs
** used 4 CPU's in parallel
*** without multiple-RHS option
Note: the forward field computation for 3Dv/3D algorithm uses single precision storage which is twice faster than using double precision computation.

speed. For instance, the reconstruction time for 2Ds^{FDTD}/2D method is close that of the actual data acquisition time and is promising in providing quasi-real-time image reconstructions by further incorporating the ADI and initial field estimation techniques. However, the image quality will need to be improved for the 3D cases utilizing actual measured data (especially for the conductivity component) before these improvements are fully realized.

For the 3D reconstructions, the scalar technique based on the FE method together with the iterative block solver provide an efficient approach for modelling the 3D field distribution with the understanding that the underlying scalar model imposes multiple approximations. The 3D FDTD algorithm used in 3Dv/3D method is very promising because of 1) the accuracy in field modelling, 2) the simplicity in programming and 3) the flexibility in accommodating various optimizations as discussed in Section 5.1.6. From Table 5.3.3, utilizing a mesh that is three times larger for the forward problem

(considering the vector nature, the actual number of unknowns is much larger than that for the scalar problem), the 3D FDTD method can still compute the full vector solution within 9 seconds which is less than twice that of the scalar technique. With the rapid increase in computing power, the use of vector techniques such as the FDTD method becomes increasingly important for producing accurate field representations and consequently improves the overall image quality.

5.4 Discussions and conclusions

We have developed two 3D image reconstruction algorithms in this chapter including the 3D scalar forward field/3D reconstruction method based on the FE technique and the 3D vector forward field/3D reconstruction method based on the FDTD technique. The forward models and assumptions of both techniques were discussed with emphasis on the 3D FDTD method. The adjoint method devised in the previous chapter was further extended to a nodal-based approximate formula which significantly simplifies the pre-processing computation stage of the reconstruction along with an associated reduction in computation time. Several enhancements of the 3D FDTD method applied to the image reconstruction problem were investigated including the use of a high-order difference scheme, initial field estimates and the ADI FDTD method.

Reconstructions utilizing both simulated and measured data were performed to validate the proposed algorithms. For most cases, the target objects were successfully reconstructed in both location and dielectric property values. Generally, the permittivity images have fewer artifacts than for the conductivity images. The high level of conductivity artifacts in the phantom data reconstructions may indicate a model-data mismatch or diminished system signal-to-noise which may not be fully appreciated at

this time.

Finally, we compared the series of algorithms we have developed utilizing the dual-mesh and iterative reconstruction framework. The 2D algorithms are superior in speed and the 3D algorithms are generally superior in accuracy. Among the 2D algorithms, the 2D FDTD based technique is quite promising and might facilitate quasi-real-time imaging. Sufficient measurement data, especially the cross-plane measurement data, has been shown to be essential for the 3D reconstructions. Consequently, the computational expenses in 3D are significantly greater than for the 2D cases and the computation speed enhancements in the forward modelling are necessary for practical use of these algorithms.

The investigations into the 3D image reconstruction algorithm and data acquisition system are still quite preliminary and there is significant work to be done in order to make them useful. Further studies of 3D microwave imaging include the further reductions in the forward and reconstruction computation time, improvements in the DAQ system performance and improved match between the numerical model and measured data from data acquisition system.

Chapter 6

Multiple-frequency dispersion reconstruction algorithm

A multiple frequency dispersion reconstruction (MFDR) algorithm utilizing a Gauss-Newton iterative strategy is presented for microwave imaging in this chapter. This algorithm facilitates the simultaneous use of multiple frequency measurement data in a single image reconstruction. Using the stabilizing effects of the low frequency measurement data, higher frequency data can be included to reconstruct images with improved resolution. The parameters reconstructed in this implementation are now frequency independent dispersion coefficients instead of the actual properties and may provide new diagnostic information. In this chapter, large high-contrast objects are successfully constructed utilizing assumed simple dispersion models for both simulation and phantom cases for which the traditional single frequency algorithm previously failed. Consistent improvement in image quality can be observed by involving more frequencies in the reconstruction; however, there appears to be a limit to how closely spaced the frequencies can be chosen while still providing independent new information which will be explored in the next chapter. Possibilities for fine-tuning the image reconstruction

performance in this context include 1) variations of the assumed dispersion model, and 2) Jacobian matrix column and row weighting schemes. Techniques for further reducing the forward solution computation time using time-domain solvers are also briefly discussed. The proposed dispersion reconstruction technique is quite general and can also be utilized in conjunction with other Gauss-Newton based algorithms including the log-magnitude phase-form (LMPF) algorithm.

6.1 Introduction

As introduced in the preceding chapters, our data acquisition and image reconstruction strategy involves collection of data at receiver sites about the imaging zone associated with multiple electromagnetic illuminations similarly to other tomographic microwave imaging techniques. In most cases, the reconstructed images generally improve with increased amounts of measured data [67]. The existing reconstruction algorithms have usually only been applied utilizing single operating frequencies with the lower frequency reconstructed images appearing smoother and with less detail but also exhibiting more stable convergence behavior to a viable solution compared with the less stable higher frequency cases (a further discussion on the impact of operating frequency to image reconstruction can be found in Section 7.4). However, increasing the amount of data through reconstructions utilizing data from multiple frequency (MF) illuminations could prove to be a powerful way to improve the image quality.

Unlike the frequency-hopping approach of Chew and Lin [33] and the multi-frequency work of Haddadin *et al.* [72], in which the spectral data were applied sequentially, we have developed a multi-frequency approach where the spectral data simultaneously contribute to a single image reconstruction. The following sections discuss the im-

plementation of this approach along with possible characteristic relationships for the permittivity and conductivity frequency dispersions. A parameter scaling approach is also discussed since scaling of the recovered dispersion coefficients is a considerably different problem than just scaling the electrical properties for a single frequency problem [129]. The results section illustrates the strength of this algorithm in three challenging cases: two simulations and an analogous phantom experiment. The large high-contrast object imaging cases were chosen because the standard, single frequency algorithm converged to non-useful images for the higher frequency cases and produced only very smoothed images for the lower frequency reconstructions. Only by using the combination of data from both the lower and higher frequencies was the algorithm able to recover well-resolved images of the targets.

6.2 Theory

6.2.1 Multiple frequency dispersion reconstruction algorithm

Assuming time dependence of $\exp(j\omega t)$, the complex wave number squared, k^2, for non-magnetic isotropic media can be written as

$$\begin{aligned} k^2 &= \omega^2\mu_0\epsilon(\omega) \\ &= \omega^2\mu_0\left(\epsilon_r(\omega)\epsilon_0 - j\tfrac{\sigma(\omega)}{\omega}\right) \\ &= k_R^2 - jk_I^2 \end{aligned} \tag{6.1}$$

where ω is the angular frequency, and $k_R^2 = \omega^2\mu_0\epsilon_0\epsilon_r(\omega)$ and $k_I^2 = \omega\mu_0\sigma(\omega)$ are the real and imaginary constituents of k^2.

Multiple dispersion models exist with varying degrees of complexity and appropriateness [36, 59, 43, 87]. For the microwave frequencies we are most interested, the frequency range is generally well within the range between the dipolar and atomic re-

laxation frequencies, such that the property variations are smooth and well-behaviored. Without loss of generality, we can express the dispersion relationships in terms of non-dispersive coefficients as

$$\begin{aligned} \epsilon_r(\omega) &= \epsilon_r\left(\omega, \lambda_1, \lambda_2, \cdots, \lambda_M\right) \\ \sigma(\omega) &= \sigma\left(\omega, \gamma_1, \gamma_2, \cdots, \gamma_N\right) \end{aligned} \tag{6.2}$$

where $\lambda_i \quad (i = 1, 2, \cdots, M)$ and $\gamma_i \quad (i = 1, 2, \cdots, N)$ are the frequency independent dispersion coefficients for the M and N term relationships, $\epsilon_r(\omega)$ and $\sigma(\omega)$, respectively.

The Gauss-Newton's method assumes (from a truncated Taylor series with respect to k_R^2 and k_I^2) [96]

$$\begin{aligned} \mathbf{\Delta E}_R &= \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_R^2}\mathbf{\Delta k}_R^2 + \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_I^2}\mathbf{\Delta k}_I^2 \\ \mathbf{\Delta E}_I &= \frac{\partial \mathbf{E}_I}{\partial \mathbf{k}_R^2}\mathbf{\Delta k}_R^2 + \frac{\partial \mathbf{E}_I}{\partial \mathbf{k}_I^2}\mathbf{\Delta k}_I^2 \end{aligned} \tag{6.3}$$

where vectors $\mathbf{\Delta E}_R$ and $\mathbf{\Delta E}_I$ are the real and imaginary part of the difference between measured and calculated fields, respectively. The lengths of vectors $\mathbf{\Delta E}_R$ and $\mathbf{\Delta E}_I$ are equal to the total measurement data $TR = T \times R$, where T denotes the number of transmitters and R denotes the number of receivers per transmitter. Vectors $\mathbf{k}_R^2$ and $\mathbf{k}_I^2$ are length P, which is the number of unknown property parameters. The derivative terms in (6.3) are all matrices of size $TR \times P$. Combining equations (6.1), (6.2) and (6.3), applying the chain rule and assuming single frequency operation initially yields

$$\begin{aligned} \mathbf{\Delta E}_R &= \sum_{i=1}^{M} \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_R^2}\frac{\partial \mathbf{k}_R^2}{\partial \lambda_i}\mathbf{\Delta}\lambda_i + \sum_{i=1}^{N} \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_I^2}\frac{\partial \mathbf{k}_I^2}{\partial \gamma_i}\mathbf{\Delta}\gamma_i \\ \mathbf{\Delta E}_I &= \sum_{i=1}^{M} \frac{\partial \mathbf{E}_I}{\partial \mathbf{k}_R^2}\frac{\partial \mathbf{k}_R^2}{\partial \lambda_i}\mathbf{\Delta}\lambda_i + \sum_{i=1}^{N} \frac{\partial \mathbf{E}_I}{\partial \mathbf{k}_I^2}\frac{\partial \mathbf{k}_I^2}{\partial \gamma_i}\mathbf{\Delta}\gamma_i \end{aligned} \tag{6.4}$$

which can subsequently be written in matrix form

$$\begin{pmatrix} J_R^R & J_I^R \\ J_R^I & J_I^I \end{pmatrix} \begin{pmatrix} \mathbf{\Delta l} \\ \mathbf{\Delta g} \end{pmatrix} = \begin{pmatrix} \mathbf{\Delta E}_R \\ \mathbf{\Delta E}_I \end{pmatrix} \tag{6.5}$$

The components of the Jacobian matrix J are

$$\begin{aligned} J_R^R &= \left(\frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_R^2} \frac{\partial \mathbf{k}_R^2}{\partial \lambda_1} \middle| \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_R^2} \frac{\partial \mathbf{k}_R^2}{\partial \lambda_2} \middle| \cdots \middle| \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_R^2} \frac{\partial \mathbf{k}_R^2}{\partial \lambda_M} \right) \\ J_I^R &= \left(\frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_I^2} \frac{\partial \mathbf{k}_I^2}{\partial \gamma_1} \middle| \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_I^2} \frac{\partial \mathbf{k}_I^2}{\partial \gamma_2} \middle| \cdots \middle| \frac{\partial \mathbf{E}_R}{\partial \mathbf{k}_I^2} \frac{\partial \mathbf{k}_I^2}{\partial \gamma_N} \right) \end{aligned}$$

with J_R^I and J_I^I having corresponding definitions. J_R^R and J_R^I are submatrices with dimensions $(TR)\times(P \times M)$ whereas J_I^R and J_I^I are $(TR)\times(P \times N)$. $\mathbf{\Delta l} = (\Delta\lambda_1, \Delta\lambda_2, \cdots, \Delta\lambda_M)^T$, and $\mathbf{\Delta g} = (\Delta\gamma_1, \Delta\gamma_2, \cdots, \Delta\gamma_N)^T$ are the frequency independent property updates solved for at each iteration. By solving equation (6.5) at each iteration, the dispersion coefficient lists, i.e. $(\lambda_1, \lambda_2, \cdots, \lambda_M)$ and $(\gamma_1, \gamma_2, \cdots, \gamma_N)$, can be updated by

$$\begin{aligned} (\lambda_1, \lambda_2, \cdots, \lambda_M)_{s+1} &= (\lambda_1, \lambda_2, \cdots, \lambda_M)_s + \mathbf{\Delta l}_s^T \\ (\gamma_1, \gamma_2, \cdots, \gamma_N)_{s+1} &= (\gamma_1, \gamma_2, \cdots, \gamma_N)_s + \mathbf{\Delta g}_s^T \end{aligned} \tag{6.6}$$

where s is the iteration index. Essentially, the images are comprised of the dispersion coefficient distributions. As before, the dielectric profiles at any specified frequency in the investigating band can be readily calculated from equation (6.2). Additionally, the reconstructed dispersion coefficients themselves might provide new diagnostic information by capturing the dispersion signature of the tissues over a range of frequencies. For a given dispersion relationship, the terms $\frac{\partial \mathbf{k}^2}{\partial \lambda_i}$ and $\frac{\partial \mathbf{k}^2}{\partial \gamma_i}$ in (6.4) can be computed analytically. The details for deriving $\frac{\partial \mathbf{E}}{\partial \mathbf{k}^2}$ can be found in Section 4.1.2.

Since $\mathbf{\Delta l}$ and $\mathbf{\Delta g}$ are frequency independent, equation (6.5) can be generalized to F

frequencies by expanding the Jacobian matrix on the left and electric field difference vector on the right

$$\begin{pmatrix} J_R^R(\omega_1) & J_I^R(\omega_1) \\ J_R^I(\omega_1) & J_I^I(\omega_1) \\ J_R^R(\omega_2) & J_I^R(\omega_2) \\ J_R^I(\omega_2) & J_I^I(\omega_2) \\ \cdots & \cdots \\ J_R^R(\omega_F) & J_I^R(\omega_F) \\ J_R^I(\omega_F) & J_I^I(\omega_F) \end{pmatrix} \begin{pmatrix} \mathbf{\Delta l} \\ \mathbf{\Delta g} \end{pmatrix} = \begin{pmatrix} \mathbf{\Delta} E_R(\omega_1) \\ \mathbf{\Delta} E_I(\omega_1) \\ \mathbf{\Delta} E_R(\omega_2) \\ \mathbf{\Delta} E_I(\omega_2) \\ \cdots \\ \mathbf{\Delta} E_R(\omega_F) \\ \mathbf{\Delta} E_I(\omega_F) \end{pmatrix} \tag{6.7}$$

Note that the Jacobian matrix and $\mathbf{\Delta} E$ terms are now functions of frequency. Equation (6.7) is the generic form for MFDR and is valid for both 2-D and 3-D cases since the dispersion characteristics for an isotropic medium are dimensionless. It is also valid for vector or scalar forward models with dispersive or non-dispersive medium. Additionally, the MFDR technique can be combined, without loss of generality, with the log-magnitude/phase form (LMPF) approach which was discussed in Section 3.2.3. The MFDR expression for the LMPF algorithm is

$$\begin{pmatrix} J_R^\Gamma(\omega_1) & J_I^\Gamma(\omega_1) \\ J_R^\Phi(\omega_1) & J_I^\Phi(\omega_1) \\ J_R^\Gamma(\omega_2) & J_I^\Gamma(\omega_2) \\ J_R^\Phi(\omega_2) & J_I^\Phi(\omega_2) \\ \cdots & \cdots \\ J_R^\Gamma(\omega_F) & J_I^\Gamma(\omega_F) \\ J_R^\Phi(\omega_F) & J_I^\Phi(\omega_F) \end{pmatrix} \begin{pmatrix} \mathbf{\Delta l} \\ \mathbf{\Delta g} \end{pmatrix} = \begin{pmatrix} \mathbf{\Delta}\Gamma(E(\omega_1)) \\ \mathbf{\Delta}\Phi(E(\omega_1)) \\ \mathbf{\Delta}\Gamma(E(\omega_2)) \\ \mathbf{\Delta}\Phi(E(\omega_2)) \\ \cdots \\ \mathbf{\Delta}\Gamma(E(\omega_F)) \\ \mathbf{\Delta}\Phi(E(\omega_F)) \end{pmatrix} \tag{6.8}$$

where Γ and Φ symbolize the log-magnitude and unwrapped phase of the electric fields,

respectively. In this situation, the modified Jacobian terms can be expressed as

$$\begin{aligned} J_R^{\Gamma} &= \frac{E_R J_R^R + E_I J_R^I}{E_R^2 + E_I^2} \\ J_I^{\Gamma} &= \frac{E_R J_I^R + E_I J_I^I}{E_R^2 + E_I^2} \\ J_R^{\Phi} &= \frac{E_R J_R^I - E_I J_R^R}{E_R^2 + E_I^2} \\ J_I^{\Phi} &= \frac{E_R J_I^I - E_I J_I^R}{E_R^2 + E_I^2} \end{aligned} \tag{6.9}$$

J^{Γ} and J^{Φ} are Jacobian submatrices. $\mathbf{\Delta\Gamma}\left(E(\omega)\right)$, where $\mathbf{\Delta\Gamma}\left(E(\omega)\right) = \ln\left(E^{\text{meas}}(\omega)\right) - \ln\left(E^{\text{calc}}(\omega)\right)$ are the differences in log-amplitude between measured and calculated field values at the receivers, and $\mathbf{\Delta\Phi}\left(E(\omega)\right)$, where $\mathbf{\Delta\Phi}\left(E(\omega)\right) = \arg\left(E^{\text{meas}}(\omega)\right) - \arg\left(E^{\text{calc}}(\omega)\right)$, are the differences in unwrapped phases [151]. In practice, the Gauss-Newton algorithms described in equations (6.7) and (6.8) are ill-posed and can only be successfully used by applying appropriate regularization techniques [191, 23, 126].

6.2.2 Dispersion model

The electrical property dispersion relationships can vary significantly from one material to another. Accurate characterization over a large frequency spectrum such as 10 MHz 100 GHz is quite difficult due to multiple relaxations mechanisms [54]. Fortunately, within a narrower microwave frequency band used in medical microwave imaging, most biological tissue and coupling media investigated to date [171, 128] follow a smooth characteristic function enabling us to utilize simple functional representations. The linear model is the most straightforward case where an individual electrical property can be represented in a two term expression as

$$\vartheta(\omega) = \alpha\, w + \beta \tag{6.10}$$

where ϑ can be either ϵ_r (or σ) or $\ln(\epsilon_r)$ (or $\ln(\sigma)$); w represents either ω or $\ln(\omega)$ and

α and β are linear coefficients. The use of the logarithm function allows us to assume linear-linear, log-linear, linear-log or log-log relationships, respectively. For example, $\epsilon_r(\omega) = \alpha \ \ln(\omega)+\beta$ is referred as the log-linear model. For sufficiently narrow frequency intervals, the linear-linear model is quite often a good approximation. The traditional single frequency reconstruction can be regarded as a limiting case of the linear-linear MFDR algorithm where α is simply set to zero.

Important factors to consider when choosing an appropriate frequency dispersion model for the image reconstruction algorithm include:

1. the dispersion model will be applied identically to all materials in the imaging zone. For this type of imaging, *a priori* knowledge concerning the dispersion characteristics of the target and medium may be useful, and

2. the MFDR algorithm is general enough to accommodate more complicated dispersion models than those suggested above. Such a model could be utilized for a variety of complex relationships over a large frequency range; however, the consequences would include reconstructing more unknowns which could increase the possibility of convergence instability.

6.2.3 Row and column weighting

From a statistical perspective, the optimal strategy to scale the linear equations (6.7) and (6.8) requires one to know the statistical properties of both the measurement and the unknowns to be reconstructed. In other words, the covariance matrices W and U in (2.53) should be known prior to the reconstruction. In general, W can be determined from the measurement data. Since the unknown parameters are treated as nonstochastic quantities, U can be replaced by the identity matrix.

However, choosing non-identity weighting matrices in the reconstruction was found to have numerical significance in the reconstruction which could provide more balanced images and minimize numerical noise. For example, we have previously explored a parameter scaling approach to balance the influence of the permittivity parameters to their associated conductivity values [129]. This scaling approach falls under the general heading of matrix row and column weighting [67, 105] in solving least-square problems. As we progress to reconstructing the dispersion coefficients themselves instead of the actual properties, the parameter scaling clearly becomes more complex given the fact that the reconstruction itself is ill-posed and nonlinear in nature. In general, the weighted system of equations can be written in the form

$$(D_R A D_C)\left(D_C^{-1}\mathbf{x}\right) = (D_R \mathbf{b}) \tag{6.11}$$

where D_R and D_C are the row and column diagonal weighting matrices, and A, x and b are the conventional left-hand-side (LHS) matrix, unknown and right-hand-side (RHS) vectors, respectively (where the equation $A\mathbf{x} = \mathbf{b}$ is formed from either equations (6.7) or (6.8). In this case, A is the Jacobian matrix and x is the same unknown vector as in (6.7) and (6.8)). For the problems described in equations (6.7) and (6.8), the dimensions of $A, \mathbf{x}$ and $\mathbf{b}$ are $(2TR \times F) \times (P \times M + P \times N)$, $P \times (M + N)$, and $(2TR \times F)$, respectively. In equation (6.11), the problem is initially solved for the least square solution of $\left(D_C^{-1}\mathbf{x}\right)$ instead of simply $\mathbf{x}$ from which $\mathbf{x}$ can then eventually be computed through multiplication by D_C

$$\mathbf{x} = D_C \left(D_C^{-1}\mathbf{x}\right)_{LS} \tag{6.12}$$

For linear least square problems, Sluis [196] showed that when the diagonal matrix D_C

is given by

$$D_C = \text{diag}\{1/\|\mathbf{a}_i\|_2, 1/\|\mathbf{a}_2\|_2, \cdots, 1/\|\mathbf{a}_N\|_2\} \tag{6.13}$$

where $\mathbf{a}_i$ are the column vectors of A. the condition number of the scaled LHS matrix is maximum.

We can also deliberately tweak the reconstruction equation by these scaling mechanisms. For example, by setting D_R, we can form a weighted least-square problem through which the weights of the measurement data at different frequencies or weights between log-magnitude and phase can be explicitly set depending on the problem needs.

6.2.4 Time-domain forward computation

As was demonstrated in the preceding chapters, the most significant computational time expense for a Gauss-Newton iterative approach (utilizing the adjoint technique) is the calculation of the forward electric field solutions at each iteration. Utilizing the previously developed frequency domain technique [151, 127], implementation of the MFDR approach would increase the computation time linearly based on the number of frequencies used. However, implementation of a time-domain electric field forward solution can offer significant benefits because the required multiple frequency solutions can be extracted from a single time-domain solution. In practice, the FDTD approach utilizing a differential Gaussian pulse could be used to generate the time domain response. The pulses would be applied individually at each transmitting antenna followed by a fast Fourier transform (FFT) of the signal responses at each receiver to recover the associated single frequency response. While the computation time for performing the Fourier transformations is not insignificant, it still constitutes a considerable time savings compared to solving multiple frequency domain forward solutions at each frequency (i.e. computing numerous matrix back substitutions). For a typical reconstruction problem

size, a factor of roughly 2 in computation time reduction is achieved by using the time-domain/FFT approach when 5 frequencies are applied. For this analysis, the main assumption is that the dielectric properties are constant with frequency. For more realistic dispersive of the dielectric property relationships, one may choose more sophisticated FDTD algorithms as discussed in [101].

6.3 Results

In this section, we present three examples to illustrate image reconstruction improvements utilizing this technique. These examples focus on 2-D reconstructions utilizing transverse-magnetic (TM) microwave illuminations. Sixteen monopole antennas are positioned equally about the perimeter on a 15 cm diameter circle. The data sets consist of electric field measurements at all 15 receivers for a given transmitting antenna with the target region being illuminated sequentially by each of the 16 antennas individually.

For all forward calculations including the generation of simulated measurements in example 1, the 2D FDTD approach introduced in Chapter 3 is used with a GPML boundary condition for truncating the mesh. The grid size for the forward domain is 110×110 for a total size of 18.8 cm×18.8 cm surrounded by 12 layers of the GPML. For each excitation, a monochromatic wave is applied at the location of each antenna. The dual-mesh settings of these reconstructions are described in Section 3.6. The circular parameter mesh consists of 281 nodes with 524 associated linear triangular elements concentrically placed within the antenna array.

The reconstruction algorithm utilizes the Tikhonov regularized Gauss-Newton method discussed in Chapter 3. The algorithm was allowed to proceed 30 iterations in example

1 and 20 iterations for example 2. Recognizing that the output of the MFDR algorithm is a set of dispersion coefficient distributions which by themselves do not actually have physical meaning, we have interpolated all results to the dielectric profiles at 900 MHz based on the selected dispersion model for all cases, unless otherwise noted, to simplify the comparisons.

The first example is a simulation of a large/high-contrast, two-region object consisting of frequency varying materials to mimic a breast with a large inclusion. This example is intended to demonstrate the performance of MFDR under ideal conditions. In addition to the first example, we present a second simulation which exploits the notion of visualizing the dispersion coefficients directly to enhance the low-contrast object recovery. In the third example, we reconstruct images of a cylindrical molasses phantom with a saline inclusion from actual measured data. The measurement data is acquired using the prototype system described in [122] with a background medium comprised of a 50:50 glycerin:water mixture [128]. The relative permittivities and conductivities of the background, object and inclusion were measured using an HP85070B dielectric probe kit in conjunction with an HP8753C Network Analyzer.

A log-log dispersion model was chosen for both the simulation and the reconstruction of the molasses phantom from actual measurement data. All reconstructions were initialized as a homogeneous domain with the actual background dispersion coefficients. Before starting the process, a least squared regression process was used to establish the dispersion coefficients of the background permittivity and conductivity from actual probe measurements. All computations were performed on a Compaq AlphaServer 833 MHz ES40 workstation. The computation time for each iteration included roughly 3 seconds per frequency to calculate the forward field solutions for 16 transmitting antennas using 4 CPU's in parallel and roughly 1 second to solve for the

dispersion coefficient update vector on a single CPU.

6.3.1 Simulation experiments

A 10.2 cm diameter cylindrical object with a 3.0 cm diameter inclusion located in the lower left quadrant is used in the simulation. Properties equivalent to that of 0.9% saline are used as the background coupling medium to generate a high contrast imaging problem that would normally be difficult to reconstruct in a single frequency scheme. The properties for the object and inclusion roughly mimic that for breast fat and glandular tissue, respectively [171]. The property dispersion curves for the background, object and inclusion used in this simulation are plotted in Figure 6.1.

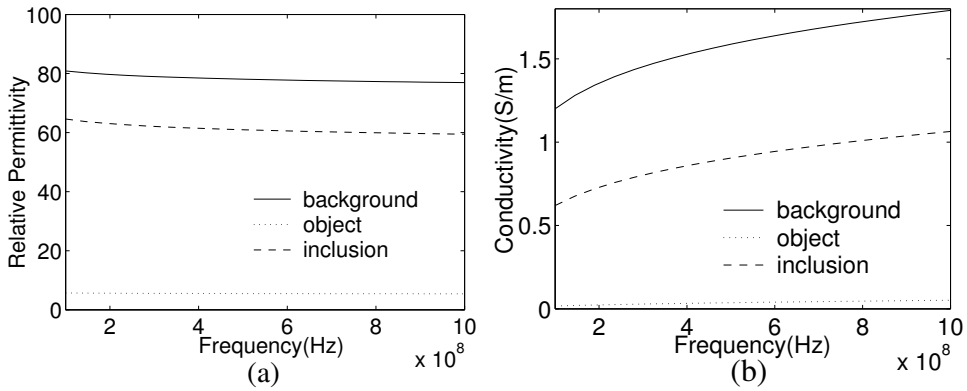


Figure 6.1: Simulated dispersion curves for the materials used in the simulation (a) relative permittivity, (b) conductivity.

Figure 6.2 shows the recovered relative permittivity and conductivity images for several single and multiple frequency reconstructions utilizing the log-log dispersion model. While the 300 MHz case converges to a stable image, the properties are quite smoothed over the domain (the inclusion appears only as an indentation in the object perimeter) as would be expected because of the reduced resolution associated with the lower frequencies. For both higher, single frequency cases (600 and 900 MHz), the

images have clearly converged to non-interesting solutions suggesting that the measurement data in these two cases individually do not contain sufficient information to recover stable images. The two frequency case utilizing 300 and 600 MHz recovers an accurate representation of the phantom with the inclusion more accurately defined than for the 300 MHz case alone while a similar two frequency case using 600 and 900 MHz cannot recover a useful image. It is clear that the algorithm benefits from both the stabilizing effects of the lower frequency data along with the higher resolution capabilities of the higher frequency reconstruction. The images for the three (300, 600 and 900 MHz) and four (300, 500, 700 and 900 MHz) frequency cases converge to similar solutions to that of the combined 300 and 600 MHz case suggesting again that the 300 MHz data is vital for convergence stability but that the increased higher frequency data has diminished impact.

Figure 6.3 illustrates the RMS error (e_{RMS}) between the true and recovered electrical property values as a function of iteration number for the seven cases discussed above. e_{RMS} is defined as

$$e_{RMS} = \sqrt{\frac{\sum_{i=1}^{P}\left(\vartheta_i^{\text{true}} - \vartheta_i^{\text{recon}}\right)^2}{P}} \tag{6.14}$$

where ϑ stands for either ϵ_r or σ and P is the number of reconstruction parameters. These are also plotted for the dispersion relationship defined 900 MHz values. Similar to the qualitative results in Figure 6.2, e_{RMS} does not decrease significantly for either the ϵ_r or σ cases with iteration for either the 600, 900 or 600/900 MHz cases which would be expected since all of the images in these cases converge to non-interesting solutions. Of the remaining cases, the 300 MHz error plots converge to the highest values for both ϵ_r and σ which would also be expected since these are the least spatially resolved. The remaining three converge to nearly the same σ error value; however, the three and four frequency reconstructions converge to a slightly improved ϵ_r error compared with

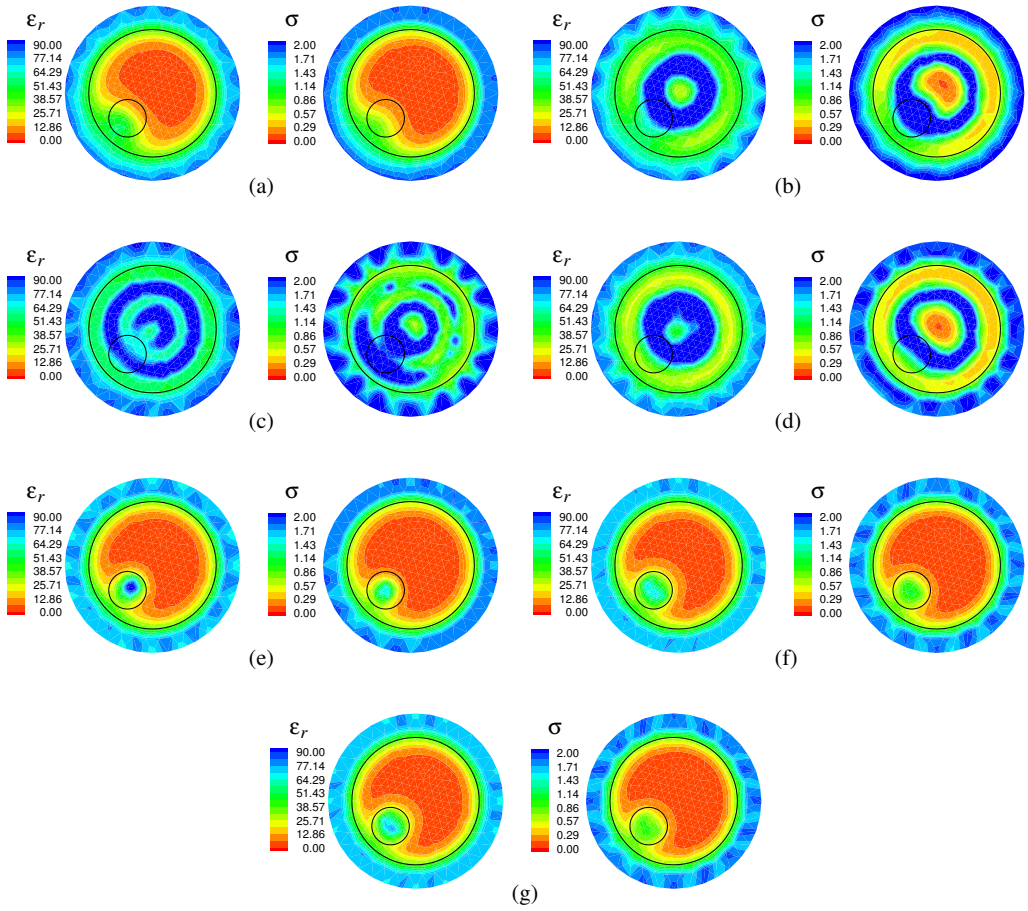
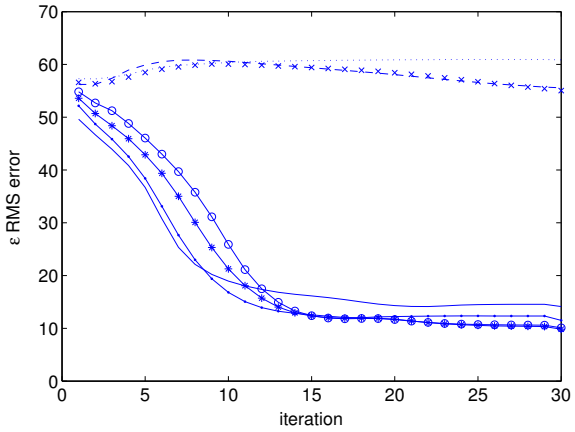
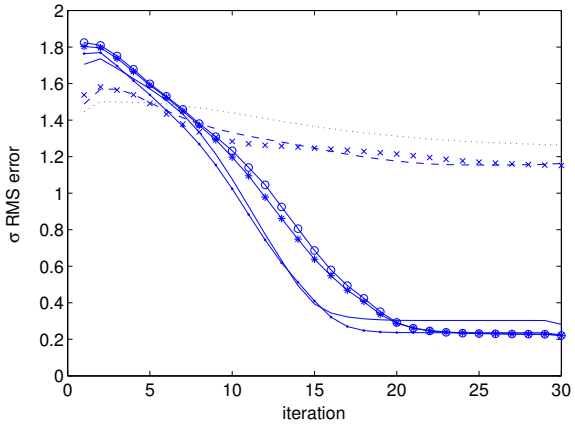


Figure 6.2: Reconstructed permittivity and conductivity images of a 10.2 cm diameter breast-like object with a 3.0 cm diameter tumor-like inclusion at (a) 300 MHz, (b) 600 MHz, (c) 900 MHz, (d) 600/900 MHz, (e) 300/600 MHz, (f) 300/600/900 MHz, (g) 300/500/700/900 MHz using simulated data.



(a)



(b)

Figure 6.3: Plots of the (a) ϵ_r and (b) σ RMS errors between the actual and recovered properties as a function of iteration for all seven imaging cases shown in Figure 6.2.

the 300/600 MHz case. This suggests that the addition of more frequency data does improve the images somewhat but that increasing the amount of data beyond the three frequency sets in this situation has minimal impact.

The previous example shows the advantages of MFDR in a high-contrast image reconstruction situation; however, the approach also works well in lower contrast cases. In this particular low contrast case, the object is difficult to distinguish from dielectric images at individual frequencies while its dispersion characteristics might provide significant contrast from that of the background which can be exploited by MFDR. For this example, we removed the 10.2cm diameter object from the previous case and retain the inclusion. The inclusion and background media were characterized by a linear-log dispersion model: permittivity (a) background, $\alpha_{\epsilon_r} = -6.98 \times 10^{-11}, \beta_{\epsilon_r} = 3.70$, (b) inclusion, $\alpha_{\epsilon_r} = 9.67E \times 10^{-11}, \beta_{\epsilon_r} = 2.85$ (Note that the signs of the dispersion slopes are opposite for the background and inclusion). For the conductivity component, the inclusion and background are identical having $\alpha_\sigma = 8.41 \times 10^{-11}, \beta_\sigma = -0.662$. The permittivity dispersion curves of the inclusion and background are plotted in Figure 6.4 (a). Using the LMPF-MFDR reconstruction with simulated measurement data at 600/900 MHz, $\alpha_{\epsilon_r}, \beta_{\epsilon_r}$ and conductivity dispersion coefficients were successfully reconstructed. The recovered α_{ϵ_r} image (Figure 6.4 (b)) clearly shows the distinct dispersion characteristics of the inclusion. Based on the dispersion model, equation (6.10), the permittivity distributions were also computed at 600 MHz and 900 MHz and shown in Figure 6.4 (c) and (d). The inverted contrast of the object relative to the background can be observed as the result of the dispersion reconstruction.

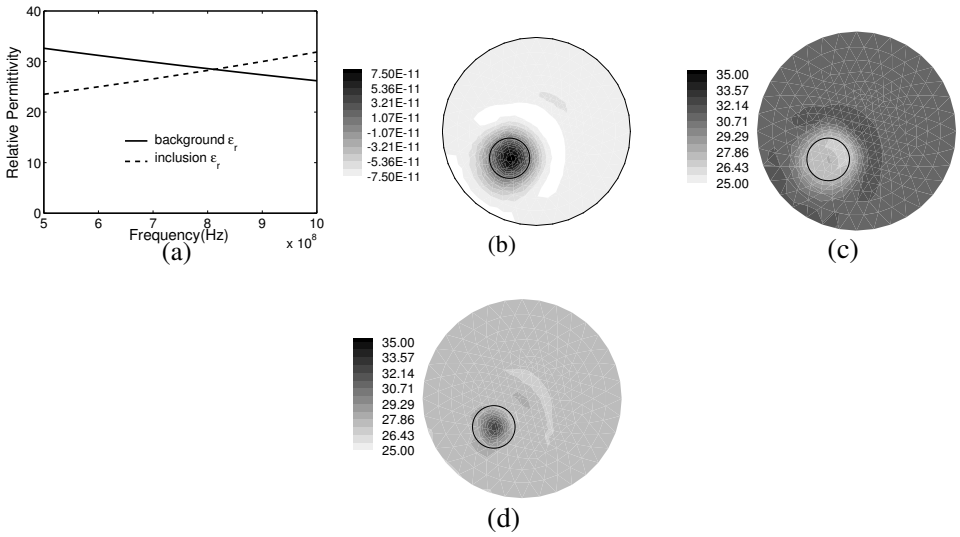


Figure 6.4: Direct utilization of dispersion coefficients: (a) relative permittivity dispersion curves, (b) reconstructed α_{ϵ_r}, (c) computed ϵ_r at 600 MHz and (d) 900 MHz.

6.3.2 Phantom experiments

For this experiment, the imaging target was a 10.1 cm diameter cylinder of molasses with a 3.1 cm diameter 0.9% saline inclusion offset upwards within the molasses. The entire molasses cylinder was positioned 0.6 cm upwards from the center of the array and is surrounded by a background medium comprised of 50% Glycerin and 50% water. The electrical properties of the three liquids are plotted versus frequency in Figure 6.5.

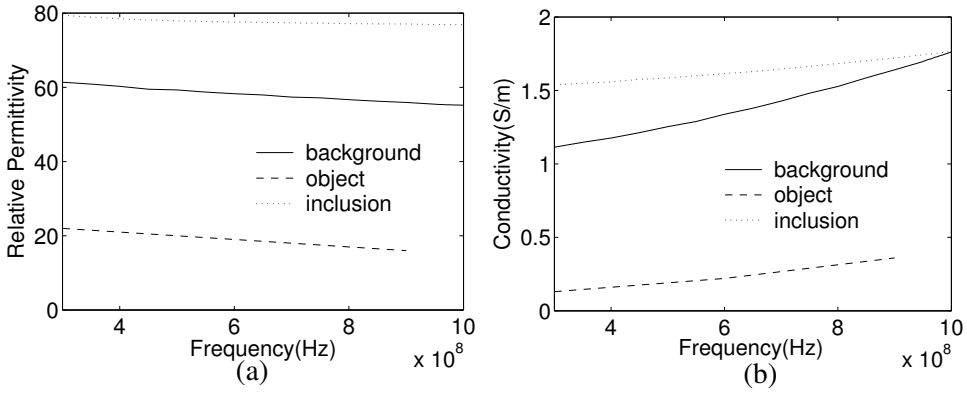


Figure 6.5: Measured electrical properties for the materials used in the phantom experiment: (a) relative permittivity, (b) conductivity.

Figure 6.6 shows the reconstructed images for the phantom utilizing various combinations of single and multiple frequency data sets. Similar to the first simulation case, the ϵ_r image for the 500 MHz case shows the rough outline of the cylinder with an indentation near its top surface corresponding to the saline inclusion. The recovered object in the associated conductivity image is smaller in size than its permittivity counterpart (similar to the observations in Chapter 4 and 5) with no apparent indication of any inclusion. The property values are nominally correct and the least squared electric field error (LSE) plot does not suggest that this solution has diverged (Figure 6.7). Similar to the simulation cases in Section 6.3.1, the higher frequency case (900 MHz) has converged to a non-interesting image. The two multi-frequency cases have converged to significantly better resolved images compared with the 500 MHz case. In both cases the outline of the molasses phantom is clearly defined in both permittivity and conductivity images with the location of the inclusion consistently more accurately recovered in the permittivity component. Additionally, the property distribution of the molasses appears to be more uniform and the recovered values of the inclusion are more accurate for the three frequency case. We also utilized the linear-linear dispersion model to reconstruct this phantom, the results are shown in 6.8. These images are similar to those in Figure 6.6 but the quality of the images is slightly worse, especially for the conductivity images.

6.4 Discussions and conclusions

We have developed a dispersion characteristic reconstruction technique which facilitates the synergy of multiple frequency measurements into a single image reconstruction process. Utilization of lower frequency data alone can often produce low resolution

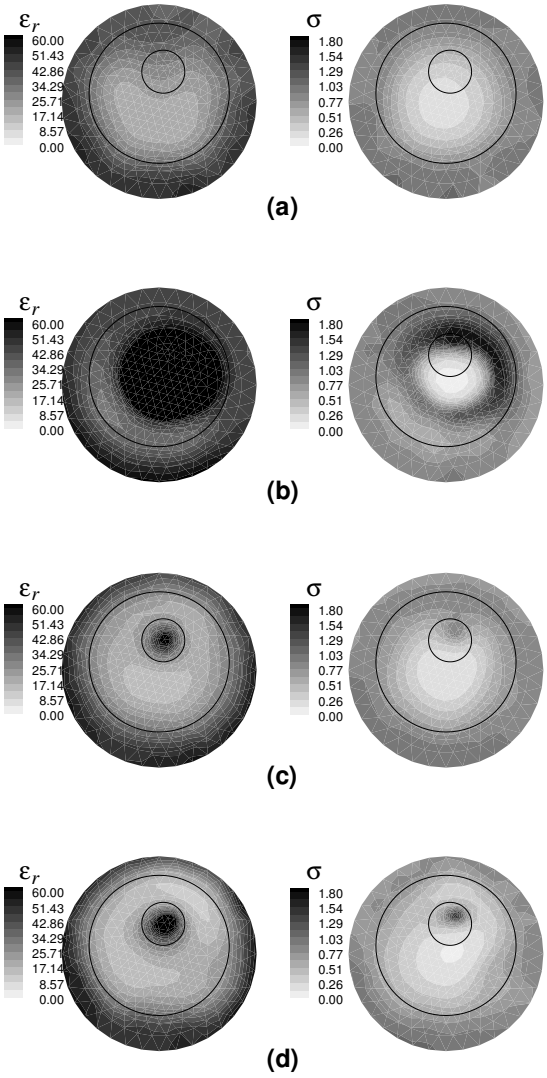


Figure 6.6: Reconstructed permittivity and conductivity images of a 10.1 cm diameter cylinder of molasses with a 3.1 cm diameter saline inclusion at (a) 500 MHz, (b) 900 MHz, (c) 500/900 MHz, (d) 300/500/900 MHz using measurement data (assuming log-log dispersion relationship).

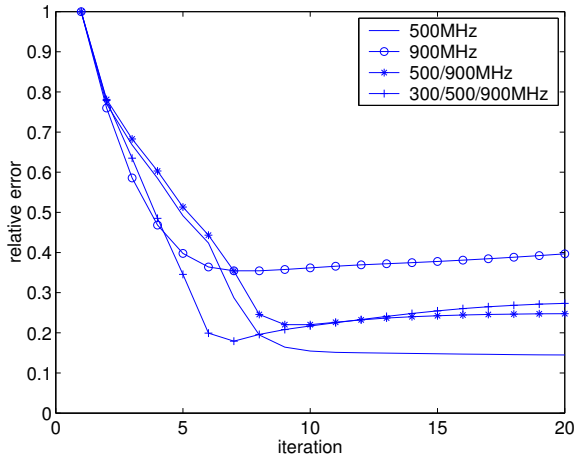


Figure 6.7: Relative error curves for the phantom reconstructions at various frequencies.

images in a stable manner while reconstructions using higher frequency data alone (especially when imaging large, high contrast objects such as the breast) often result in non-meaningful results. For the algorithms presented here, we simultaneously utilize measurement data over a broad frequency range and recover frequency independent coefficients associated with assumed underlying property dispersion relationships. While images at discrete frequencies can be extracted by applying the dispersion relationships after the reconstruction is processed, the dispersion coefficients themselves may provide additional diagnostic information.

We specifically chose imaging problems for large, high-contrast objects to demonstrate the capabilities of this approach; that is the single frequency algorithm was known to diverge for the higher frequency cases without the assistance of *a priori* information. In both simulations and phantom experiments, it is clear that we can only utilize the higher frequency data when combined with that for a lower frequency. In addition, as the amount of higher frequency data is increased, there is a slight improvement in the image quality. While we would naturally expect resolution improvement whenever any new data is added, the level of independence of the new data from the existing data may

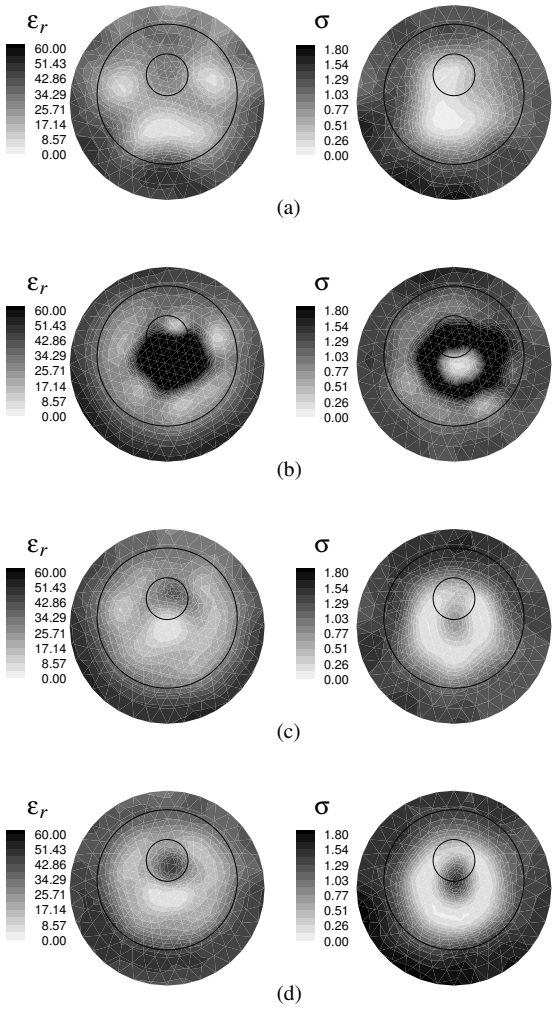


Figure 6.8: Reconstructed permittivity and conductivity images of a 10.1 cm diameter cylinder of molasses with a 3.1 cm diameter saline inclusion at (a) 500 MHz, (b) 900 MHz, (c) 500/900 MHz, (d) 300/500/900 MHz using measurement data (assuming linear-linear dispersion relationship).

be compromised when the selected frequencies are closely packed. This will be an important point of further investigation as we work towards utilizing the higher frequency (up to 2.5 GHz) data available with our new data acquisition system [111].

Additionally, there remain several techniques by which the performance of this approach can be fine-tuned. As we work towards the development of a lower contrast coupling medium for our breast imaging system, the single frequency algorithm can often recover a stable image, even at higher frequencies because of the lower contrast. It will be important to study the effects of the dispersion model choice, along with the amount and span of the additional frequency information in conjunction with existing reconstruction enhancement capabilities such as row and column weighting and our 2-step regularization approach [126] to optimize the system resolution.

Chapter 7

Singular value analysis of the Jacobian matrix

7.1 Introduction

The update equation (3.17) plays a central role in the reconstruction process of the dielectric properties. The significance of this equation lies in the fact that the final image of the reconstruction is directly comprised of the solutions yielded from this equation at a sequence of iterations. In the cases where the forward field is linearly dependent on the parameters, this equation leads to the final solution instantly. The measurement data, parameter update and the sensitivity map in terms of the Jacobian matrix are all components of this equation which along with any *a priori* information complete the ingredients for the image reconstruction. Moreover, as a matrix equation, (3.17) is comparatively easy to solve and analyze. Therefore, a thorough analysis of this equation, especially the Jacobian matrix, is of great importance in explaining the behavior of the reconstructor and estimating the performance of the imaging system. More importantly, these understandings could be useful in assisting the optimization of

the measurement system to maximize overall efficiency.

The principal questions we are interested in exploring with this analysis include

1. Is there a metric that reasonably characterizes the performance of a imaging scheme?

2. What is the resolution limit of the microwave imaging system and what factors impact the resolution?

3. How does the measurement noise effect the image reconstruction?

4. Given the measurement scheme, what are the optimal system parameters (such as operating frequency, source/detector number and distribution, etc.) that yield the best image quality?

From a linear equation perspective, the singular value decomposition on the LHS matrix in (3.17) is promising as a way to assist in answering the first question because of the rich information exposed by the SVD. In the remainder of this chapter, we focus on constructing the singular value decomposition (SVD) of the Jacobian matrix from which the analysis of the image resolution and the impact of the measurement noise are performed. For this purpose, we develop a metric in terms of the singular value spectrum and use it to explore the optimal imaging system configurations.

Temporarily ignoring the regularization term and various statistical assumptions, (3.17) is mathematically equivalent to the following equation

$$J(\mathbf{k}^2 - \mathbf{k}_0^2) = \mathbf{E} - \mathbf{E}^{\text{meas}} \tag{7.1}$$

where J is the Jacobian matrix defined in (2.21). Although the reconstruction of a high-contrast object in microwave imaging can not be completed by solving this equation in

a single step, the single step solution of (7.1) provides certain measures related to the recoverability of the algorithm in a linear functional sense. Therefore, investigating the single step solution, or the solution under the Born approximation, is valuable for an in-depth understanding of the reconstruction algorithm.

The single step reconstruction using a homogeneous background initial guess and far field assumption (Section 2.2) was studied by Brander and DeFacio [19]. In their study, the SVD analysis of the discretized forward operator referred as to the Born matrix is performed but with a quite general description. For our derivations, we approach the analysis of the near field, single-step reconstruction from the perspective of the nodal adjoint method. In this assessment, we compute the analytical SVD for the nodal adjoint form of the Jacobian matrix similarly to the analytical work performed by Nelson and Kahana [141] for the acoustic scattering problem. Various issues concerning the image resolution and noise are then subsequently investigated based on this decomposition.

7.2 Analytical SVD of the Jacobian matrix

The nodal adjoint formula (5.8) provides a simple but meaningful representation for the Jacobian matrix. With a uniform single-mesh, i.e. the forward and parameter meshes are identical, where the nodes inside the mesh are evenly distributed, (for each node, their effective volumes V_n $(n = 1, \cdots, P)$ are approximately identical, denoted as V), the Jacobian matrix can be written in the following form (defined in Section 5.1.2)

$$J = VB \tag{7.2}$$

where the $((s,r),\tau)$-th element of matrix B is given by

$$b_{(s,r),\tau} = E(\vec{r}_\tau, \vec{r}_s)\hat{E}(\vec{r}_\tau, \vec{r}_r) \tag{7.3}$$

where (s,r) represents a transmitter/receiver pair with s being the index of the source, r being the index of the receiver, τ is the index of the parameter and $E(\vec{r}_\tau, \vec{r}_s)$ and $\hat{E}(\vec{r}_\tau, \vec{r}_r)$ are the electric fields at node τ due to sources at either s or r, respectively. For simplicity, we use a single index ν to access all possible $(s,\ r)$ pairs and a one-to-one mapping is established denoted as $\nu : \{1, 2, \cdots, Q\} \rightarrow (s,r)$ where Q is the total number of the combinations. Thus, (7.3) can be re-written as

$$b_{\nu,\tau} = E(\vec{r}_\tau, \vec{r}_{\nu:s})\hat{E}(\vec{r}_\tau, \vec{r}_{\nu:r}) \tag{7.4}$$

where $\nu : s$ and $\nu : r$ are the corresponding source and receiver indices in the ν-th pair, respectively.

For the microwave imaging case, the receivers and sources are both point sources, therefore, the field E_s and $\hat{E}_r$ can be expressed in terms of Green's functions, i.e.

$$\begin{aligned} E(\vec{r}, \vec{r}_{\nu:s}) &= \alpha \times g(\vec{r}, \vec{r}_{\nu:s}) \\ \hat{E}(\vec{r}, \vec{r}_{\nu:r}) &= g(\vec{r}, \vec{r}_{\nu:r}) \end{aligned} \tag{7.5}$$

where α is a constant denoting the source strength. Substituting (7.5) into (7.3) and subsequently into (7.2), the Jacobian can be expressed as

$$J = \alpha V \begin{pmatrix} g(\vec{r}_1, \vec{r}_{1:s})g(\vec{r}_1, \vec{r}_{1:r}) & g(\vec{r}_2, \vec{r}_{1:s})g(\vec{r}_2, \vec{r}_{1:r}) & \cdots & g(\vec{r}_P, \vec{r}_{1:s})g(\vec{r}_P, \vec{r}_{1:r}) \\ g(\vec{r}_1, \vec{r}_{2:s})g(\vec{r}_1, \vec{r}_{2:r}) & g(\vec{r}_2, \vec{r}_{2:s})g(\vec{r}_2, \vec{r}_{2:r}) & \cdots & g(\vec{r}_P, \vec{r}_{2:s})g(\vec{r}_P, \vec{r}_{2:r}) \\ \cdots & \cdots & \cdots & \cdots \\ g(\vec{r}_1, \vec{r}_{Q:s})g(\vec{r}_1, \vec{r}_{Q:r}) & g(\vec{r}_2, \vec{r}_{Q:s})g(\vec{r}_2, \vec{r}_{Q:r}) & \cdots & g(\vec{r}_P, \vec{r}_{Q:s})g(\vec{r}_P, \vec{r}_{Q:r}) \end{pmatrix} \tag{7.6}$$

It is interesting to note, from (7.6), the curves satisfying

$$g(\vec{r}, \vec{r}_s)g(\vec{r}, \vec{r}_r) = c \tag{7.7}$$

can be generated for different constants c which depict isocurves (or isosurfaces) with respect to the sensitivity to the parameter perturbation. For certain special cases, we have plotted the iso-sensitivity curves and surfaces in the Appendix.

In order to decompose the matrix in (7.6), the approach by Nelson and Kahana is applied [141]. First, we assume that there exist two sets of orthogonal basis functions, one for the spatial domain where the unknown parameters are located, i.e. the parameter space, and one for the space where the source/receiver antennas are located, i.e. the excitation space. We denote the orthonormal basis for the parameter space as $\{\varphi_i(\vec{r}_\tau)\}$ and the orthonormal basis for excitation space as $\{\phi_i(\vec{r}_s, \vec{r}_r)\}$. Thus, we can expand each element of the matrix in (7.6) into the following form

$$g(\vec{r}_\tau, \vec{r}_{\nu:s})g(\vec{r}_\tau, \vec{r}_{\nu:r}) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \lambda_m^n \varphi_m(\vec{r}_\tau)\phi_n(\vec{r}_{\nu:s}, \vec{r}_{\nu:r}) \tag{7.8}$$

where λ_m^n are the coefficients.

Once again, all (m, n) pairs are denoted by a single index κ such that $\text{abs}(\lambda_\kappa)$ are sorted in non-decreasing order. Assuming the summation over the first K terms of the expansion is sufficiently accurate to represent the original function, the truncated expansion is then written as

$$g(\vec{r}_\tau, \vec{r}_{\nu:s})g(\vec{r}_\tau, \vec{r}_{\nu:r}) \approx \sum_{\kappa=1}^{K} \lambda_\kappa \varphi_{\kappa:m}(\vec{r}_\tau)\phi_{\kappa:n}(\vec{r}_{\nu:s}, \vec{r}_{\nu:r}) \tag{7.9}$$

Nelson and Kahana [141, 142] demonstrated that once the expansion (7.9) is con-

structed, matrix J can be decomposed into the following form

$$J = F\Lambda H^T \tag{7.10}$$

where

$$F = \begin{pmatrix} \phi_{1:n}(\vec{r}_{1:s}, \vec{r}_{1:r}) & \phi_{2:n}(\vec{r}_{1:s}, \vec{r}_{1:r}) & \cdots & \phi_{K:n}(\vec{r}_{1:s}, \vec{r}_{1:r}) \\ \phi_{1:n}(\vec{r}_{2:s}, \vec{r}_{2:r}) & \phi_{2:n}(\vec{r}_{2:s}, \vec{r}_{2:r}) & \cdots & \phi_{K:n}(\vec{r}_{2:s}, \vec{r}_{2:r}) \\ \cdots & \cdots & \cdots & \cdots \\ \phi_{1:n}(\vec{r}_{Q:s}, \vec{r}_{Q:r}) & \phi_{2:n}(\vec{r}_{Q:s}, \vec{r}_{Q:r}) & \cdots & \phi_{K:n}(\vec{r}_{Q:s}, \vec{r}_{Q:r}) \end{pmatrix} \tag{7.11}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_K \end{pmatrix} \tag{7.12}$$

$$H = \begin{pmatrix} \varphi_{1:m}(\vec{r}_1) & \varphi_{2:m}(\vec{r}_1) & \cdots & \varphi_{K:m}(\vec{r}_1) \\ \varphi_{1:m}(\vec{r}_2) & \varphi_{2:m}(\vec{r}_2) & \cdots & \varphi_{K:m}(\vec{r}_2) \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{1:m}(\vec{r}_P) & \varphi_{2:m}(\vec{r}_P) & \cdots & \varphi_{K:m}(\vec{r}_P) \end{pmatrix} \tag{7.13}$$

The dimensions of these matrices are: $K \times K$ for Λ, $Q \times K$ for F and $P \times K$ for H where K is the truncation level, Q is the combination number for all source/receiver pairs and P is the total number of the unknown parameters. Moreover, the matrices F and H are both column orthogonal, i.e. $F^T F = H^T H = I$ which is readily proved from the orthonomalities of the associated basis functions.

Nelson and Kahana also demonstrated in [141] that there exists an orthogonal matrix S such that

$$J = (FS)(S^T \Lambda S)(HS)^T \tag{7.14}$$

where $S^T \Lambda S$ is a non-negative-valued diagonal matrix. By letting

$$\begin{aligned} U &= FS \\ \Sigma &= S^T \Lambda S \\ V &= HS \end{aligned} \tag{7.15}$$

equation (7.14) becomes the SVD of the Jacobian matrix. From (7.15) we can see that the left and right singular vectors are linear combinations of the basis functions in the excitation space and parameter space, respectively.

From the above derivation, it is clear that determining the orthonormal bases of the parameter and excitation spaces is essential for the decomposition of the Jacobian matrix. Unfortunately, it is quite difficult to construct such basis functions for arbitrary spatial domains. In the following section, we focus on a special 2D case: circular parameter domain with an equispaced circular source/receiver array.

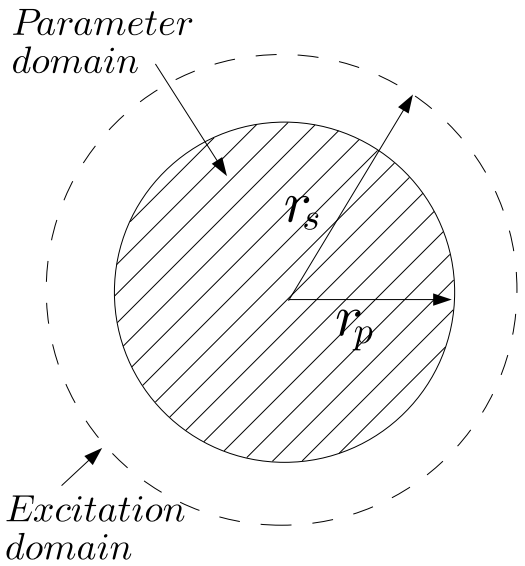


Figure 7.1: Circular parameter domain with equally spaced circular antenna array.

7.3 Jacobian SVD over a circular parameter domain

The problem analyzed in this section is illustrated in Figure 7.1. The parameter domain is a circular region with radius r_p. Sources and receivers are equally distributed along a concentric circle whose radius is $r_s \geq r_p$. The orthonormal basis function for the circular parameter domain is analytically known, called the Zernike polynomial [205, 17] and is defined by

$$Z_m^n(\rho, \theta) = R_m^n(\rho / r_p) \exp(jn\theta) \tag{7.16}$$

where the radial component $R_m^n(\rho)$ is defined by

$$R_m^n(\rho) = \sum_{l=0}^{(m-n)/2} \frac{(-1)^l (m-l)!}{l!((m+n)/2-l)!((m-n)/2-l)!} \rho^{m-2l} \tag{7.17}$$

Thus, H can be written as

$$H = \begin{pmatrix} \beta_1 Z_{1:m}(\vec{r}_1) & \beta_2 Z_{2:m}(\vec{r}_1) & \cdots & \beta_K Z_{K:m}(\vec{r}_1) \\ \beta_1 Z_{1:m}(\vec{r}_2) & \beta_2 Z_{2:m}(\vec{r}_2) & \cdots & \beta_K Z_{K:m}(\vec{r}_2) \\ \cdots & \cdots & \cdots & \cdots \\ \beta_1 Z_{1:m}(\vec{r}_P) & \beta_2 Z_{2:m}(\vec{r}_P) & \cdots & \beta_K Z_{K:m}(\vec{r}_P) \end{pmatrix} \tag{7.18}$$

where β_κ is defined by

$$\beta_\kappa = \frac{1}{\sqrt{|Z_\kappa|}} \tag{7.19}$$

and $|Z_\kappa| = \sqrt{\int Z_\kappa^2(\vec{r}) d\vec{r}}$ is the l^2-norm of the κ-th Zernike polynomial.

Given the orthogonal expansion expression (7.9), we have the following relationship

$$\lambda_{\kappa:n} \phi_{\kappa:n}(\theta_{v:s}, \theta_{v:r}) = \frac{1}{|Z_{\kappa:n}|} \int g(\vec{r}, \vec{r}_{v:s}) g(\vec{r}, \vec{r}_{v:r}) Z_{\kappa:n}(\vec{r}) d\vec{r} \tag{7.20}$$

If we denote ΛH^T in (7.10) by W, then (7.20) represents the (v, κ)-th element of matrix

W. By considering $H^T H = I$, the singular values λ can finally be computed by

$$\lambda_{\kappa:n} = \sqrt{||w_i||_2} \tag{7.21}$$

where w_i is the i-th row of matrix W.

The Green's function in the 2D homogeneous case is given by (2.6). By inserting it into (7.20), the integration on the RHS contains the multiplication term of three special functions

$$\lambda_{\kappa:n}\phi_{\kappa:n}(\theta_{\nu:s}, \theta_{\nu:r}) = \frac{-1}{16|Z_{\kappa:n}|}\int H_0^{(1)}(k_{bk}|\vec{r} - \vec{r}_{\nu:s}|)H_0^{(1)}(k_{bk}|\vec{r} - \vec{r}_{\nu:r}|)Z_{\kappa:n}(\vec{r})d\vec{r} \tag{7.22}$$

The possibility of analytically evaluating this integral is still under investigation. However, for the remainder of this section, numerical solution of the Jacobian matrix will be computed for this domain configuration to verify our analysis.

Figure 7.2: Circular parameter mesh.

In the previous derivation, we assumed that the mesh has uniform density within the parameter domain. A suitable mesh for this problem is shown in Figure 7.2. The mesh has a radius of 6 cm and is surrounded by a 7.6 cm radius antenna array consisting of 32 equally spaced monopole antennas. Each antenna transmits a signal at 900 MHz while

the remaining 31 antennas act as receivers. The Jacobian matrix under this scheme is computed using our nodal adjoint formula (5.8) and its SVD is numerically computed. Several right singular vectors were selected as examples whose values are plotted as distributions over the parameter mesh (Figure 7.3). From these plots, the Zernike polynomial Z_4^4, Z_4^{-4}, Z_4^0, Z_3^1, Z_5^{-3}, Z_5^3, Z_6^0, Z_6^2 and Z_6^4 can be identified as the major component of each pattern.

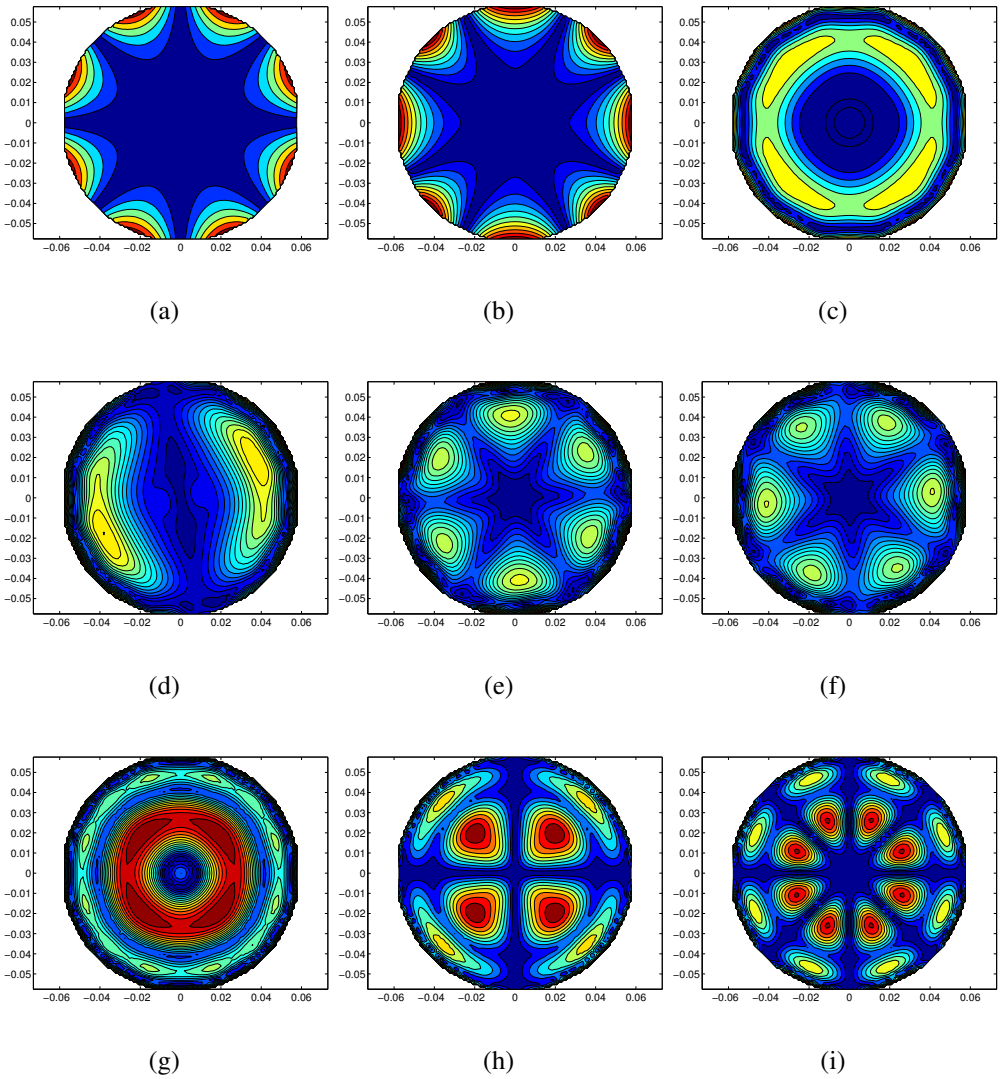


Figure 7.3: Right singular vector patterns: (a) $|v_8|$, (b) $|v_9|$, (c) $|v_{28}|$, (d) $|v_{29}|$, (e) $|v_{35}|$, (f) $|v_{36}|$, (g) $|v_{46}|$, (h) $|v_{54}|$, (i) $|v_{65}|$.

7.4 Numerical SVD and the degree of ill-posedness

Analytical decomposition of the Jacobian matrix is only available for a few special cases and evaluation of the integration in (7.20) is quite difficult in most cases. Alternatively, the numerical evaluation of the Jacobian matrix for the homogeneous background medium (i.e. using the Born approximation) can be easily computed even for irregular-shaped parameter domains and source distributions given the nodal adjoint representation. In this section, we study the impact of various imaging parameters, including operating frequency, source/receiver number, background properties and parameter mesh density, on the singular spectrum of the Jacobian matrix, and, consequently, the potential quality of the reconstructed image.

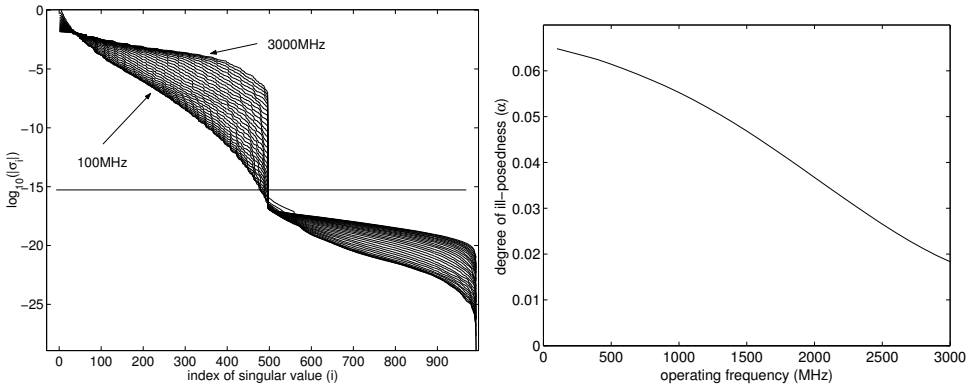
The problem configuration used in this section is identical to that of the previous section except where otherwise noted. The numerical singular spectrum of the Jacobian matrices were computed using different system parameters. In order to make comparisons between the different spectrum's, we modified the concept of the "degree of ill-posedness" discussed in [77] using the following definition

Definition 7.4.1 (The degree of ill-posedness). *if there exists a positive real number α, for a singular spectrum $\{\sigma_i\}_i^N$, and if $\frac{\sigma_i}{\sigma_1} = O(\exp(-\alpha i))$, then, α is called the degree of ill-posedness of the spectrum.*

A linear regression process was performed on the series $\{\log(\sigma_i) - \log(\sigma_1)\}_i$ for estimating α.

We first computed the Jacobian singular spectra at various frequencies within the working frequency range of our system, i.e. 100 MHz to 3000 MHz in 100 MHz increments. The plot of these spectrum curves is shown in Figure 7.4 (a). Notice that portion of the curves underneath roughly -14 (log scale) are essentially zeros with

respect to the computer resolution and are neglected. From the curve in Figure 7.4 (b), we can see that the degree-of-illposedness at 3000 MHz is roughly 1/3 of that in the 100 MHz case indicating that utilizing higher frequencies is advantageous in providing less redundant information about the target. However, one should recognize that with the increase in frequency, the nonlinearity between the measurement data and dielectric properties becomes more severe requiring more accurate initial estimate. In cases where no *a priori* information is provided, the higher frequency measurement data might lead to solutions trapped at local minima. Alternatively, image reconstructions at lower-frequencies may require additional regularization which essentially smooths the image. This may partially explain the distinct behaviors of the low frequency and high frequency measurements described in Chapter 6.

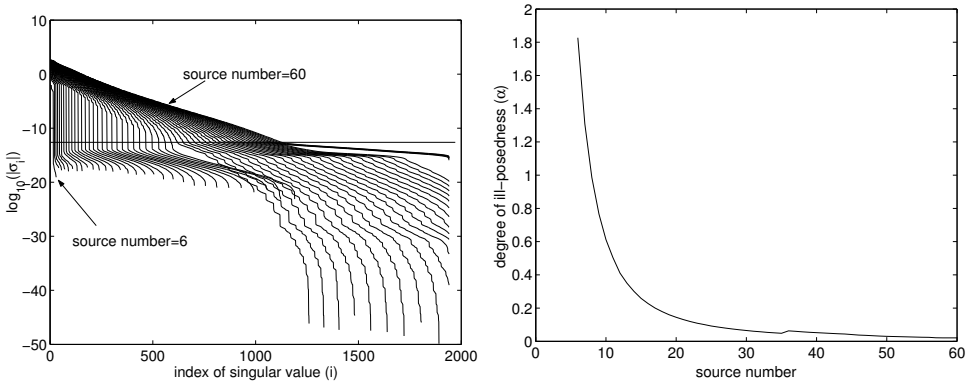


(a) (b)

Figure 7.4: (a) Singular spectra and (b) degree-of-illposedness for a range of frequencies.

In a second experiment, we varied the number of sources/receivers surrounding the parameter domain. The antennas were evenly distributed along the array with one transmitting and the remaining antennas acting as receivers. At 1100 MHz, their spectra are plotted in Figure 7.5 (a). The corresponding degree-of-illposedness plot is shown

in Figure 7.5 (b). These plots confirm our experience that more sources can provide improved imaging performance. The degree-of-illposedness is significantly reduced for increased number of antennas. At this frequency, above an array antenna count of 35, the rate of reduction in ill-posedness slows considerably.

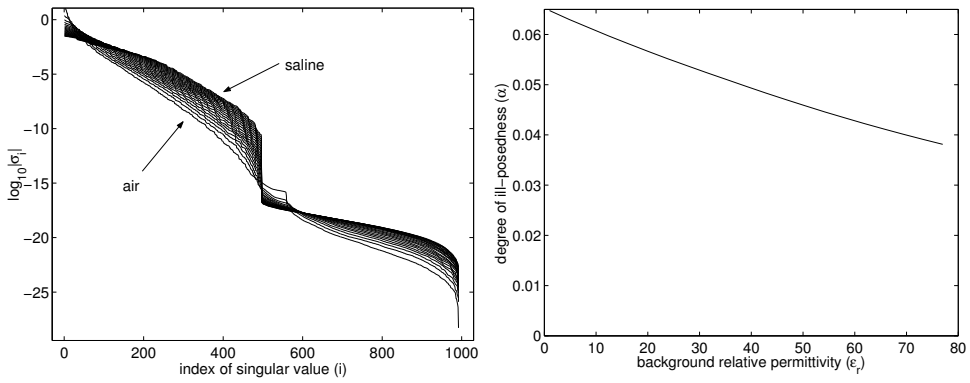


(a) (b)

Figure 7.5: (a) Singular spectra and (b) degree-of-illposedness for various source/receiver numbers.

The third experiment focused on the effects of the background medium properties. Utilizing 32 antennas operating at 1100 MHz, we varied the background dielectric properties linearly from air to 0.9% saline ($\epsilon_r = 77$, $\sigma = 1.7$ S/m). The singular spectrum curves are plotted in Figure 7.6 (a) while the corresponding degree-of-illposedness curve is shown in Figure 7.6 (b). These plots demonstrate that the more lossy the background medium (in most coupling media we used in our system, the lossier medium often has higher permittivity as well) produces a lower degree-of-illposedness. However, for very lossy media, the SNR of the measurement drops significantly which has a negative effect on the reconstruction. Therefore, when choosing the background medium, both factors should be considered.

The fourth experiment considered the density of the reconstruction mesh, or number

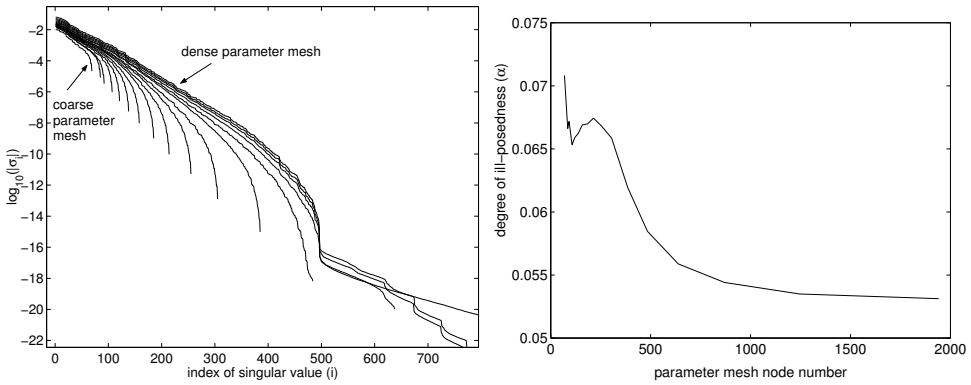


(a) (b)

Figure 7.6: (a) Singular spectra and (b) degree-of-illposedness for various background media.

of unknowns. Here, we varied the cell sizes of the reconstruction mesh and computed the corresponding Jacobian singular spectrum. The corresponding singular spectrum and degree-of-illposedness are shown in Figures 7.7 (a) and (b), respectively. From these curves, we can make three basic observations: 1) if the number of unknowns is smaller than the amount of independent measurement data (equal to the number of sources multiplied by the number of receivers per source and divided by two due to reciprocity), the singular spectrum drops precipitously at the number of parameter nodes; 2) when the number of the unknowns is greater than the number of independent measurements (i.e. underdetermined cases), the point where the spectrum drops is determined by the amount of measurement data; and 3) increasing the reconstruction mesh density is helpful in reducing the degree-of-illposedness of the problem; however, the significance of this reduction is diminished when the number of unknowns is greater than roughly twice that of the amount of independent measurement data (in this case, the amount of independent measurement data is $\frac{32\times31}{2} = 496$).

In the fifth experiment, we investigated the impact of using multi-frequency data in



(a) (b)

Figure 7.7: (a) Singular spectra and (b) degree-of-illposedness for various parameter densities.

the reconstruction. To accomplish this, we repeated the second experiment while incorporating two (700/1100 MHz) and three (300/700/1100 MHz) frequency data sets. The degree-of-illposedness curves were plotted in comparison with that of the single frequency case (1100 MHz) (Figure 7.8). From the plot, the reduction in α for low numbers of antennas is quite significant. This indicates that the amount of independent measurement increases when incorporating multi-frequency data into the reconstruction. However, the curve for triple-frequency case only drops slightly from the curve of the dual-frequency case indicating diminished improvement when adding data from more frequencies. It is also interesting to note that as the number of antennas increases, the differences between single and multiple frequencies is negligible.

One final experiment compared the log-magnitude phase form (LMPF) and traditional complex code reconstructions. As demonstrated in [151], the LMPF uses a real-valued Jacobian matrix computed from the real form Jacobian matrix. The latter is equivalent to the complex form Jacobian except that the unknowns are real variables. Thus, we compared the degree-of-illposednesses of the LMPF Jacobian and the real

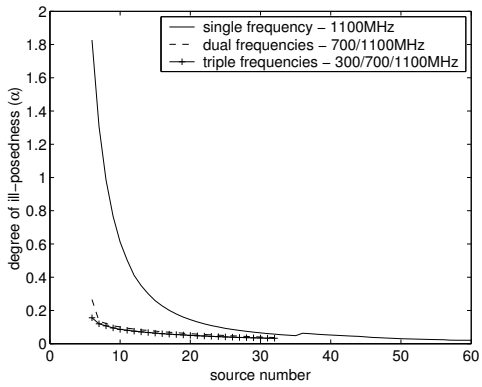


Figure 7.8: Degree-of-illposedness for source/receiver numbers computed at different frequency combinations.

form Jacobian at different source/receiver numbers. Interestingly, it turns out that the singular spectra of these two Jacobians are exactly the same. This demonstrates that the log-magnitude/unwrapped phase transform does not alter the singular spectrum of the Jacobian matrix.

Notice that the previous discussion does not take into account measurement noise which is a non-negligible factor for real applications. The presence of measurement noise essentially sets a truncation level above the matrix numerical rank in the singular spectrum and the maximum extent of the effective spectrum is consequently constrained by this noise level. In Hansen's work [77], a metric called the "effective resolution limit" is devised to denote the turning point along the singular spectrum where the measurement noise starts to dominate the solution. From our analysis above, the maximum extent of the effective spectrum is directly related to the image resolution of the final image and is evident from the orthogonal decomposition expression (7.8). The more singular spectrum values that are included in the reconstruction, the more Zernike polynomials will be used to construct the solution with correspondingly improved image resolution. This conclusion holds not only for a Jacobian matrix utilizing the Born approximation (single iteration cases), but also for iterative reconstructions. Assuming

smooth convergence in an iterative reconstruction, the final image resolution is related to the maximum angular and radial modes of the Zernike polynomials that correspond to the full effective spectrum for all iterations and these high order angular and radial modes are directly constrained by the noise level in the measurements.

7.5 Discussions and conclusions

We analyzed the singular value decomposition of the Jacobian matrix from the starting point of the nodal adjoint expression derived in Chapter 5. This analysis provides increased insight into the structure of the Jacobian matrix and the significance of its components in the process of image reconstruction. We focused on 2D cases with a homogeneous background implying the assumption of the Born approximation. However, it can be easily extended to 3D space and other simple parameter domain geometries. For iterative reconstruction schemes, we indicated that the measurement noise level plays an important role in determining the overall image quality. The concepts from the research field of ill-posed problems, such as the degree-of-illposedness and the effective resolution limit, were introduced into our application to assess the effects of various system parameters.

Among the efforts of further understanding the Jacobian matrix and its impact to the reconstruction, investigations to the analytical derivation of the integration in (7.20) for simple geometries is needed as well as characterizing the measurement noise level for our the current system in order to apply our analysis to optimizing the system performance.

Part III

Phase unwrapping and phase singularities

Chapter 8

A mathematical framework of phase unwrapping

Spatial unwrapping of the phase component of time varying electromagnetic fields has important implications in a range of disciplines including MRI, optical confocal microscopy and microwave tomography. This chapter presents a fundamental framework based on the phase unwrapping integral, especially in the complex case where phase singularities are enclosed within the closed path integral. With respect to the phase unwrapping required when utilized in Gauss-Newton iterative microwave image reconstruction, the concept of dynamic phase unwrapping is introduced where the singularity location varies as a function of the iteratively modified property distributions. Strategies for dynamic phase unwrapping in the microwave problem were developed and successfully tested in simulations and phantom experiments utilizing large, high contrast targets to validate the approach which can be found in the following chapters.

8.1 Introduction

The complex (or real and imaginary) representation is used in many physical problems to simplify the associated mathematical derivations into succinct and meaningful expressions, especially when analyzing wave phenomenon in the frequency domain. In such problems, it is quite natural to transform the complex forms of such equations into their magnitude and phase representations. These representations are also convenient since the magnitude and phase features of many physical quantities can be measured and processed directly and serve as quantitative descriptions of the associated phenomenon. Consequently, understanding the magnitude and phase in both the mathematical and physical context is important. In this paper we focus primarily on the phase which is often the more difficult quantity to process.

In many applications, the phase encodes rich temporal and spatial information [147]. Decoding it is an important aspect of signal processing especially in image processing. However, an important and often difficult aspect of the phase signal is the wrapping due to its periodical nature. The result of wrapping introduces mathematical discontinuities in the phase signal which often cause major challenges in various applications. Phase unwrapping techniques have been developed to restore the "continuous" phase from the wrapped version and has been widely studied in interferometric synthetic aperture radar (InSAR) [204, 63] and MRI image processing [38] along with other implementations. In InSAR phase data processing, investigations into the phase unwrapping imposed by the presence of phase singularities has received considerable attention over the last decade. Robust phase unwrapping strategies were developed to account for the effect of these phase singularities in static 2-D phase maps [65, 158].

Another area of investigation involving phase unwrapping and phase singularities stems from the study of nulls in wave scattering problems. Scattering nulls re-

fer to the spatial locations where the amplitudes of the field goes to zero, rendering the associated phase unspecified. Nye and Berry [146] performed the first systematic study of this phenomenon in optical scattering problems. It has also been discussed under different labels such as wave dislocations [146], phase singularities [172] and optical vortices [118, 99, 143] among various publications where its curious nature related to phase unwrapping has received the majority of attention. The equivalence of a discrete scattering null and phase singularity was demonstrated by Fried and Vaughn [58] while the evolution and structure of these nulls were investigated by several researchers [172, 42, 145].

Our work on phase unwrapping and scattering nulls is motivated by the utilization of the unwrapped phase in our iterative reconstruction algorithm for microwave medical imaging. A log-magnitude phase form (LMPF) image reconstruction algorithm was proposed by Meaney *et al.* [151] and demonstrated improved performance in both convergence and resulting image qualities over the more traditional complex form algorithm (Section 3.2.3). However, as we applied this algorithm in various clinical cases, we observed that LMPF algorithm failed under certain circumstances and produce noisy or diverged results. We later observed that the failure of the LMPF algorithm in these situations was related to the presence of scattering nulls within the imaging domain. In some cases, the nulls were only present at intermediate solutions of the Gauss-Newton iterative reconstruction process. To be able to explore the advantages of the LMPF reconstruction in these situations, robust phase unwrapping strategies are required to correctly process the scattering nulls in the iterative image reconstruction scenario.

This part of the thesis is comprised of three closely related topics. First, we will develop a mathematical framework for the phase unwrapping analysis. Conclusions from both complex analysis and the field of topology are utilized to formalize the phase

unwrapping properties in mathematical terms. The concept of static and dynamic phase unwrapping problems are introduced to facilitate the application of the phase analysis theory. In contrast to the first topic, we investigated the phenomenon of scattering nulls in the microwave scattering problem from a physical perspective as a natural extension of its optical counterpart in Chapter 9. The relationship of these phase unwrapping problems in 2-D and 3-D spaces are also briefly discussed. The last topic (Chapter 10) focuses on the application of the two preceding analyses to microwave image reconstruction utilizing the LMPF algorithm. Simple and robust unwrapping strategies are proposed to solve the challenges associated with the current algorithm. Image reconstructions utilizing simulated and clinical measured data are performed to test the validity and efficiency of these analyses and strategies.

8.2 A mathematical framework of phase unwrapping

8.2.1 Phase function and the single-valued interval

In these derivations, functions with domain X and range in Y are denoted by $W : X \rightarrow Y$. More specifically, if W is a complex-valued function in an n-dimensional Euclidean space, where $X = \mathbb{R}^n$ and $Y = \mathbb{C}$, one can write W in the form of $W(r) = u(r) + jv(r) = (u(r), v(r))$ where $r \in \mathbb{R}^n$. If real functions $u(r)$ and $v(r)$ are both continuously differentiable functions in $\mathbb{R}^n$, W is said to be continuously real-differentiable [168], and consequently continuous.

Applying Euler's formula, W can also be written in exponential form:

$$W(r) = \rho(W(r)) \exp(j\Phi(W(r))) \tag{8.1}$$

where $\rho : \mathbb{C} \to \mathbb{R}^+ \cup \{0\}$ is the amplitude function and $\Phi : \mathbb{C} \to \mathbb{R}$ is the phase function. Particularly, for a complex number $w = (u, v) \in \mathbb{C}$, phase function $\Phi(w)$ can be written as

$$\Phi(w) = \text{atan}_2(v, u) + 2n\pi \tag{8.2}$$

where $n \in \mathbb{Z}$ is an arbitrary integer due to the periodic nature of the complex exponential function. Function atan_2 is defined as an extension of arctan function as:

$$\text{atan}_2(v, u) = \begin{cases} \arctan(v/u) & u > 0 \\ \pi \times \text{sign}(v) + \arctan(v/u) & u < 0 \\ \pi/2 \times \text{sign}(v) & u = 0, v \neq 0 \\ \text{arbitrary} & u = 0, v = 0 \end{cases} \tag{8.3}$$

where sign(v) is the sign of v, defined as sign(v) = 1 if $v > 0$, -1 if $v \leq 0$.

From (8.2) we see that $\Phi(W(r))$ is a multi-valued real function. To simplify the analysis, we define the single-valued interval S for any given real number $\varphi_0 \in \mathbb{R}$ as $S = [\varphi_0, \varphi_0 + 2\pi)$. For any $r \in \mathbb{R}^n$ if $(u(r), v(r)) \in \mathbb{C}\backslash\{0\}$ ($\mathbb{C}\backslash\{0\}$ is also denoted as $\mathbb{C}^\times$), there exists one and only one integer $n_S \in \mathbb{Z}$ that satisfies (as a direct consequence of the Pigeonhole principle [28])

$$\text{atan}_2(v(r), u(r)) + 2n_S\pi \in S$$

As a result, we define a single-valued function

$$\phi_S(W(r)) = \text{atan}_2(v(r), u(r)) + 2n_S\pi \in S \tag{8.4}$$

where $\phi_S(W(r))$ is called a single-valued branch of $\Phi(W(r))$ over interval S. For convenience, we denote interval $[-\pi, \pi)$ as S_*. For any $W(r) \in \mathbb{C}^\times$, the output of function

$\text{atan}_2(v, u)$ always resides within S_*; therefore, (8.4) can be written as

$$\phi_S(W(r)) = \phi_{S_*}(W(r)) + 2n_S\pi \in S \tag{8.5}$$

8.2.2 Path and Phase unwrapping integral

Before defining the phase unwrapping integral, we must examine another important concept: the path. In complex analysis, a path is defined as a continuous map between the real interval $I = [0, 1]$ to a metric space Y [168]. For example, a curve in $\mathbb{C}$ is denoted as $\gamma : I \to \mathbb{C}$. $\gamma(0)$ and $\gamma(1)$ designate the initial point and the terminal point of curve γ. If $\gamma(0) = \gamma(1)$, we say γ is a closed path. Paths are sometimes called curves. A path-sum is defined as $\gamma = \sum_{i=1}^{N} \gamma_i$ with overlapped initial points and terminal points among the γ_i, $i = 1, 2, \cdots, N$. A path γ is said to be piecewise continuously differentiable if it can be written as a finite path-sum of continuously differentiable paths [168](p.172). Let Γ be a piecewise continuously differentiable path, $W : \mathbb{R}^n \to \mathbb{C}$ be a real-differentiable function, for any $r \in \Gamma$, if $W(r) \neq 0$, we call Γ a piecewise unwrappable path. Since the definition of the phase unwrapping integral will involve the gradient of the phase function, we will only consider piecewise unwrappable paths unless otherwise noted.

Given a single-valued interval $S = [\varphi_0, \varphi_0 + 2\pi)$ and a real-differentiable function $W : \mathbb{R}^n \to \mathbb{C}$, we call path $\Gamma : I \to \mathbb{R}^n$ an unwrappable path of S under W if for all $r \in \Gamma$

$$\phi_S(W(r)) \in S \setminus \{\phi_0\} \tag{8.6}$$

We denote $\Gamma \in P(S, W)$, where $P(S, W)$ is the set of all unwrappable paths of complex function W over the single-valued interval S. $\phi_S(W(r))$ is a composite map such that

$\phi_S \circ W \circ \Gamma : I \to \mathbb{R}$ which is readily shown to be piecewise continuously differentiable over I if $\Gamma \in P(S, W)$. For the identity map $Id : \mathbb{C} \to \mathbb{C}$, any path in the $\mathbb{C}$ plane that does not cross phase branch cut $\phi_S(z) = \phi_0$ and the origin is an unwrappable path. The previous explanation of the basic concepts sets the stage for us to define the phase unwrapping integral as:

Definition 8.2.1 (Phase Unwrapping Integral). *Let $W : \mathbb{R}^n \to \mathbb{C}$ be a continuously real-differentiable function, and $\Gamma : I \to \mathrm{dom}(W)$ be a path in the domain of W. The phase unwrapping integral of W over path Γ is then defined as*

$$\int_{\Gamma} \nabla\phi(W(r)) \cdot \mathbf{dl} = \sum_{i=1}^{N} \int_{\Gamma_i} \nabla\phi_{S_i}(W(r)) \cdot \mathbf{dl} \tag{8.7}$$

where $r \in \mathbb{R}^n$, $\Gamma = \sum_{i=1}^{N} \Gamma_i$ and $\Gamma_i \in P(S_i, W)$. $\nabla\phi_{S_i}(W(r))$ is the gradient of the single-valued phase function with single-valued interval S_i.

Γ_i is called a segment of Γ, and set $\{\Gamma_i\}_{i=1}^{N}$ is a segmentation of path Γ. All (Γ_i, S_i) pairs comprise a set of $\{(\Gamma_i, S_i)\}_{i=1}^{N}$ referred to as a partition of Γ. We use $Q(\Gamma)$ to designate the set of all partitions of Γ. The left-hand-side of (8.7) can be reduced as $\mathcal{U}(W(r), \Gamma)$ or $\mathcal{U}(W(r), \{\Gamma_i, S_i\})$.

Here we provide a constructive proof on the existence of a partition for any piecewise unwrappable path $\Gamma : I \to \mathbb{R}^n$. Interval I is a topology space, for $\forall t \in I$, we construct the quotient space $I/\sim$ based on an equivalence map [107](p.54) with $\sim = \{t | t \in I \text{ and } \phi_{S_*}(W(\Gamma(t))) \in [-\pi, 0]\}$. Since Γ and W are continuous maps, the subspaces of quotient space $\Gamma/\sim$ are continuous subintervals of I. The subintervals satisfying map $\sim$ are closed intervals whose images under Γ are unwrappable paths with $S = [-\pi/2, 3\pi/2)$. The remainder are open subintervals where $\phi_{S_*}(W(\Gamma(t))) \in (0, \pi)$. One can add boundary points to transform them into closed intervals whose images

under Γ are the unwrappable paths of $S = [\pi/2, 5\pi/2)$.

8.2.3 Properties of the phase unwrapping integral

Definition 8.2.1 illustrates the phase unwrapping integral for any valid partition of a unwrappable path. However, for a given path, the number of possible valid partitions is infinite. In this subsection, we shall first demonstrate the independence of the integral with respect to partitions of the path, followed by a illustration of the properties of the closed-path phase unwrapping integral.

Lemma 8.2.2. *Let $W : \mathbb{R}^n \to \mathbb{C}$ be a continuously real-differentiable function, and S be a single-valued interval. For any unwrappable path $\Gamma : I \to \mathrm{dom}(W)$ of S with $\Gamma(0)$ and $\Gamma(1)$ being the initial and terminal points, respectively, $\mathcal{U}(W(r), \Gamma)$ can be expressed as*

$$\mathcal{U}(W(r), \Gamma) = \int_{\Gamma} \nabla\phi_S(W(r)) \cdot \mathbf{dl} = \phi_S(W(\Gamma(1))) - \phi_S(W(\Gamma(0))) \tag{8.8}$$

Lemma 8.2.2 can be easily verified by using the fundamental theorem of a line integral [45](p.976) given the piecewise continuous differentiability of ϕ_S over I.

Lemma 8.2.3. *Let $W : \mathbb{R}^n \to \mathbb{C}$ be a continuously real-differentiable function, S_1 and S_2 be two single-valued intervals, and $\Gamma : I \to \mathrm{dom}(W)$ be a path. If $\Gamma \in P(S_1, W)$ and $\Gamma \in P(S_2, W)$, then*

$$\int_{\Gamma} \nabla\phi_{S_1}(W(r)) \cdot \mathbf{dl} = \int_{\Gamma} \nabla\phi_{S_2}(W(r)) \cdot \mathbf{dl} \tag{8.9}$$

Proof. From lemma 8.2.2, equation (8.9) can be re-written as

$$\begin{aligned}\int_\Gamma \nabla\phi_{S_1}(W(r))\cdot\mathbf{dl} &= \phi_{S_1}(W(\Gamma(1))) - \phi_{S_1}(W(\Gamma(0)))\\ \int_\Gamma \nabla\phi_{S_2}(W(r))\cdot\mathbf{dl} &= \phi_{S_2}(W(\Gamma(1))) - \phi_{S_2}(W(\Gamma(0)))\end{aligned} \tag{8.10}$$

Since path Γ lies in a single-valued branch of the phase function, the change of the single-valued interval will result in simultaneous addition or subtraction of $2n\pi$ for all points in Γ. Therefore, the phase differences between $W(\Gamma(0))$ and $W(\Gamma(1))$ remain constant, and from (8.10), lemma 8.2.3 is proven. □

Lemma 8.2.4. *Let $W:\mathbb{R}^n\to\mathbb{C}$ be a continuously real-differentiable function. For any path $\Gamma: I\to \mathrm{dom}(W)$, if $\{(\Gamma_i, S_i)\}_{i=1}^N \in Q(\Gamma)$, then we have*

$$\mathcal{U}(W(r),\Gamma) = \sum_{i=1}^{N}\left(\phi_{S_i}(W(\Gamma_i(1))) - \phi_{S_i}(W(\Gamma_i(0)))\right) \tag{8.11}$$

This is obvious from definition 8.2.1 and Lemma 8.2.2.

Lemma 8.2.5. *Let $W:\mathbb{R}^n\to\mathbb{C}$ be a continuously real-differentiable function, $\Gamma: I\to \mathrm{dom}(W)$ be a path, and $\{(\Gamma_i, S_i)\}_{i=1}^N \in Q(\Gamma)$ be a partition of Γ. If $\Gamma_i = \sum_{j=1}^{N_i}\Gamma_{i,j}$, then $\bigcup_{i,j}\{(\Gamma_{i,j}, S_i)\} \in Q(\Gamma)$ and*

$$\mathcal{U}\left(W(r),\{(\Gamma_i,S_i)\}\right) = \mathcal{U}\left(W(r),\bigcup_{i,j}\{(\Gamma_{i,j},S_i)\}\right) \tag{8.12}$$

Proof. If $\Gamma_i \in P(S_i, W)$, the continuous subspaces $\Gamma_{i,j}$ are also unwrappable paths of S_i, i.e. $\Gamma_{i,j}\in P(S_i,W)$ for all j. Therefore, $\bigcup_{i,j}\{(\Gamma_{i,j},S_i)\}$ is a valid partition of path Γ, referred to as a refinement of partition $\{(\Gamma_i,S_i)\}$. From Lemma 8.2.4,

$$\mathcal{U}(W(r),\Gamma_i) = \sum_{j=1}^{N_i}\left(\phi_{S_i}(W(\Gamma_{i,j}(1))) - \phi_{S_i}(W(\Gamma_{i,j}(0)))\right) \tag{8.13}$$

From the path-sum definition, segments $\Gamma_{i,j}$ overlap at initial and terminal points. There-

fore, the phases at intermediate points cancel and the summation in (8.13) leaves only the phase difference between the initial and terminal points of Γ_i. □

Theorem 8.2.6. *Let $W : \mathbb{R}^n \to \mathbb{C}$ be a continuously real-differentiable function, $\Gamma : I \to \mathrm{dom}(W)$ be a path, and $\{(\Gamma_i^a, S_i^a)\}_{i=1}^{N_a}$ and $\{(\Gamma_i^b, S_i^b)\}_{i=1}^{N_b}$ be two partitions of Γ. Then*

$$\mathcal{U}(W(r), \{(\Gamma_i^a, S_i^a)\}) = \mathcal{U}(W(r), \{(\Gamma_i^b, S_i^b)\}) \tag{8.14}$$

Proof. Construct a refinement $\{(\Gamma_i^c, S_i^c)\}_{i=1}^{N_c}$ of $\{(\Gamma_i^a, S_i^a)\}_{i=1}^{N_a}$, which satisfies: for $\forall \Gamma_i^c$, there exist two positive integers $M, N \in \mathbb{N}$ with $\Gamma_i^c \subset \left(\Gamma_M^a \cap \Gamma_N^b\right)$. Letting $S_i^c = S_M^a$, from Lemma 8.2.5, we can write

$$\mathcal{U}(W(r), \{(\Gamma_i^c, S_i^c)\}) = \mathcal{U}(W(r), \{(\Gamma_i^a, S_i^a)\}) \tag{8.15}$$

Furthermore, letting let $S_i^c = S_N^b = \tilde{S}_i^c$ produces

$$\mathcal{U}(W(r), \{(\Gamma_i^c, \tilde{S}_i^c)\}) = \mathcal{U}(W(r), \{(\Gamma_i^b, S_i^b)\}) \tag{8.16}$$

From Lemma 8.2.3 , the left-hand-sides of (8.15) and (8.16) are equal. Therefore, the associated right-hand-sides are also equivalent. □

Since the complex function W is a continuous map, the image of a piecewise unwrappable path in the range of W is also a piecewise unwrappable path. As a result, we can transform the integral variable to the $\mathbb{C}$ plane:

Lemma 8.2.7. *Let $W : \mathbb{R}^n \to \mathbb{C}$ be a continuously real-differentiable function, and $\Gamma : I \to \mathrm{dom}(W)$ be a path. Let $Id : \mathbb{C} \to \mathbb{C}$ be an identity map over $\mathbb{C}$, and Γ' be the*

image of the Γ *under map* W*, i.e.* $W : \Gamma \to \Gamma'$*. Then*

$$\mathcal{U}(W(r), \Gamma) = \mathcal{U}(Id(z), \Gamma') \tag{8.17}$$

Proof. Utilizing a partition of Γ, $\{(\Gamma_i, S_i)\}$, from Lemma 8.2.4 , we can write

$$\mathcal{U}(W(r), \{(\Gamma_i, S_i)\}) = \sum_i \left(\phi_{S_i}(W(\Gamma_i(1))) - \phi_{S_i}(W(\Gamma_i(0)))\right) \tag{8.18}$$

For any segment Γ_i, the image in $\mathbb{C}$ under W is Γ'_i and $\Gamma' = \sum_i \Gamma'_i$. Since $\Gamma \in P(S_i, W)$, $\Gamma'_i \in P(S_i, Id)$. Consequently, $\{(\Gamma'_i, S_i)\}$ is a partition of path Γ'. From Lemma 8.2.4

$$\mathcal{U}(Id(r), \{(\Gamma'_i, S_i)\}) = \sum_i \left(\phi_{S_i}(Id(\Gamma'_i(1))) - \phi_{S_i}(Id(\Gamma'_i(0)))\right) \tag{8.19}$$

Considering $\Gamma'_i(0) = W(\Gamma_i(0))$, $\Gamma'_i(1) = W(\Gamma_i(1))$ and $Id(z) = z$, Lemma 8.2.7 is proven.

□

The phase unwrapping integral over a closed path possesses similarities to that of the complex integral. The following two theorems are quite useful in real applications.

Theorem 8.2.8. *Let* $W : \mathbb{R}^n \to \mathbb{C}$ *be a continuously real-differentiable function,* $\Gamma : I \to \mathrm{dom}(W)$ *be a closed path, and* $\Gamma' \subset \mathbb{C}$ *be the image of* Γ*. If* Γ' *does not enclose* $z = 0$ *in the* $\mathbb{C}$ *space, we can write*

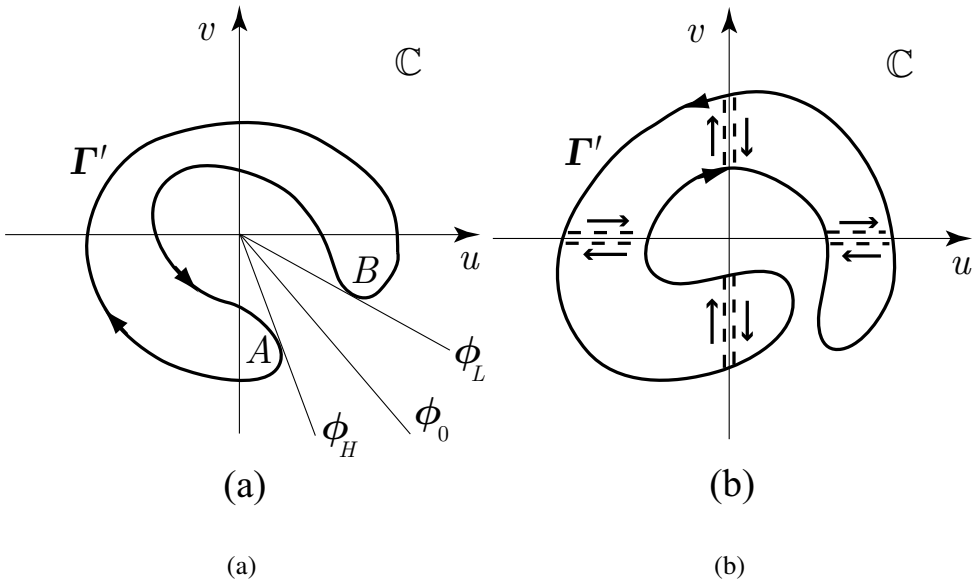
$$\mathcal{U}(W(r), \Gamma) = 0 \tag{8.20}$$

Proof. From Lemma 8.2.7, we can produce $\mathcal{U}(W(r), \Gamma) = \mathcal{U}(Id(r), \Gamma')$. The closed path integral $\mathcal{U}(Id(r), \Gamma')$ in $\mathbb{C}$ space is subsequently discussed in two cases:

1. If there exists a real number φ_0 making $\Gamma' \in P(S, Id)$, where $S = [\varphi_0, \varphi_0 + 2\pi)$, and since phase map ϕ_S is a continuous function over Γ', from the extreme value theo-

rem [107](p.76), there must exist two points $z_A, z_B \in \Gamma'$, where $\phi(z_A) = \sup(\phi_S(\Gamma')) = \phi_H$ and $\phi(z_B) = \inf(\phi_S(\Gamma')) = \phi_L$ as shown in Figure 8.1 (a). z_A and z_B divide Γ' into two segments Γ'_{AB} and Γ'_{BA} both of which are unwrappable paths of S. From Lemma 8.2.2, the summation of the phase unwrapping integral over the two paths is zero.

2. if such φ_0 does not exist (Figure 8.1(b) for example), we add pairs of paths along the u and v axes between the intersections of Γ' with the axes. This results in a collection of closed curves, each covering a single quadrant. Therefore, we can apply the conclusion from case 1 and produce zeros for all of these paths.



(a) (b)

Figure 8.1: Proof of Theorem 8.2.8

□

Theorem 8.2.8 can be extended to a more general case if we incorporate the winding number concept [168, 144, 83] from topology to get:

Theorem 8.2.9. *Let* $W : \mathbb{R}^n \to \mathbb{C}$ *be a continuously real-differentiable function,* $\Gamma : I \to \mathrm{dom}(W)$ *be a closed path,* Γ' *be the image of* Γ *in* $\mathbb{C}$*, and* $\mathrm{Ind}_{\Gamma'}(0)$ *be the winding*

number (or index) of close path Γ' *with respect to the origin in* $\mathbb{C}$*. Then*

$$\mathcal{U}(W(r),\Gamma) = 2\pi \cdot \mathrm{Ind}_{\Gamma'}(0) = \frac{1}{j}\oint_{\Gamma'}\frac{dz}{z} \tag{8.21}$$

Proof. 1. if $\mathrm{Ind}_{\Gamma'}(0) = 0$ (i.e. closed path Γ' does not enclose the origin in $\mathbb{C}$), theorem 8.2.9 is proven by theorem 8.2.8.

2. if $\mathrm{Ind}_{\Gamma'}(0) = 1$, we create a rectangular mesh within Γ' as shown in Figure 8.2. The integral over Γ' is converted into a summation of closed path integrals along the boundary of each subdivided zone. Because the directions of integration along any one line segment are opposite each other for adjacent subzones sharing that line segment, all internal line integrations within Γ' cancel. Only one rectangle Γ'_R that encloses the origin while the rest of them yield zeros from theorem 8.2.8. Denoting the two intersections of Γ'_R with u-axis by A and B, path Γ'_R is broken into Γ'_{AB} and Γ'_{BA}. Choosing $S_1 = [-\pi/2, 3\pi/2)$ and $S_2 = [\pi/2, 5\pi/2)$ where $\Gamma'_{AB} \in P(S_1, Id)$ and $\Gamma'_{BA} \in P(S_2, Id)$ produces

$$\begin{aligned}\mathcal{U}(Id(z),\Gamma'_R) &= \mathcal{U}(Id(z),\Gamma'_{AB}) + \mathcal{U}(Id(z),\Gamma'_{BA}) \\ &= (\pi - 0) + (2\pi - \pi) = 2\pi\end{aligned} \tag{8.22}$$

For paths with $\mathrm{Ind}_{\Gamma'}(0) = -1$, the analysis is similar.

3. if $|\mathrm{Ind}_{\Gamma'}(0)|$ is larger than 1, such as in Figure 8.3, we can always add pairs of paths along u-axis between its intersections with path Γ', resulting in a collection of enclosed paths Γ'_i with $\mathrm{Ind}_{\Gamma'_i}(0) = \pm 1$. From the discussion above and the addition property of winding number [168](p.288), theorem 8.2.9 is readily proven.

□

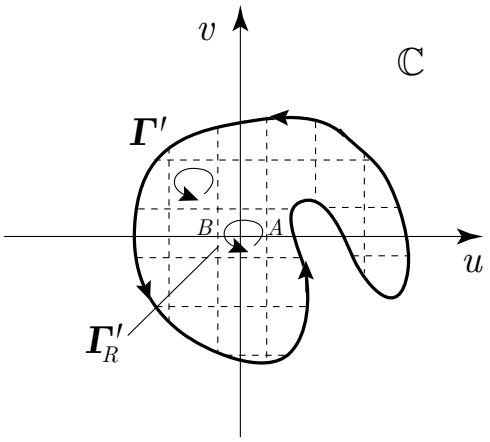


Figure 8.2: Closed path integral over Γ' with $\mathrm{Ind}_{\Gamma'}(0) = 1$.

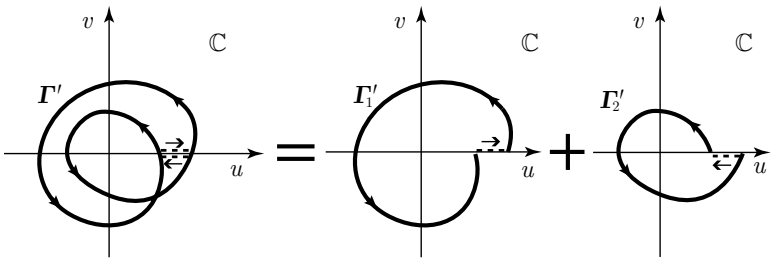


Figure 8.3: Decomposition of a multi-wound closed path into simple closed paths ($\mathrm{Ind}_{\Gamma'_1}(0) = \mathrm{Ind}_{\Gamma'_2}(0) = \pm 1$)

8.2.4 Closed path phase unwrapping integral in $\mathbb{R}^n$ space

The conclusions in previous section are general for any continuously real-differentiable complex function over $\mathbb{R}^n$. Notice that theorems 8.2.8 and 8.2.9 specifically involve the image of the unwrapping path in $\mathbb{C}$ space, while, in many applications, the unwrapping paths are chosen directly in $\mathbb{R}^n$ space (more specifically, in $\mathbb{R}^2$ or $\mathbb{R}^3$). An extension of theorem 8.2.9 from $\mathbb{C}$ to $\mathbb{R}^n$ space would prove useful.

In $\mathbb{R}^n$ space, the complete inverse image of $z = 0 \in \mathbb{C}$ is defined as the point set $\{r | r \in \mathbb{R}^n, W(r) = 0\}$, denoted by $W^{-1}(0)$. Equivalently, $W^{-1}(0)$ can be defined as the solution of

$$\begin{cases} u(r) = 0 \\ v(r) = 0 \end{cases} \qquad r \in \mathbb{R}^n \tag{8.23}$$

where $u(r)$ and $v(r)$ are the real and imaginary part of W, respectively.

Assuming map $W : \mathbb{R}^n \to \mathbb{C}$ has full rank over every point in $W^{-1}(0)$, from Theorem 5.8 and Corollary 5.9 of [16], we can be assured that $W^{-1}(0)$ is a closed regular submanifold with dimension $n - 2$. With this assumption, instead of looking for the topological relation between Γ' and z =0 in $\mathbb{C}$ as we did in Theorem 8.2.9, we can investigate the reciprocal graph pair, i.e. Γ and $W^{-1}(r)$, in $\mathbb{R}^n$ space since both are closed regular manifolds. In higher dimension space, the corresponding concept to the winding number is the linking number [97, 144](p.8). In such cases, it is not difficult to prove from the definition that the following is true

$$\mathrm{Ind}_{\Gamma'}(0) = \mathrm{Lk}(\Gamma, W^{-1}(0)) \tag{8.24}$$

where $\mathrm{Lk}(\Gamma, W^{-1}(0))$ is the linking number between the unwrapping path Γ and the complete inverse image of $z = 0$.

Consequently, we have the following conclusion:

Theorem 8.2.10. *Let $W : \mathbb{R}^n \to \mathbb{C}$ be a continuously real-differentiable function with $n \geq 2$, and $\Gamma : I \to \mathrm{dom}(W)$ be a closed paths. If W has full rank at every point in $W^{-1}(0) = \{r | r \in \mathbb{R}^n, W(r) = 0\}$, then*

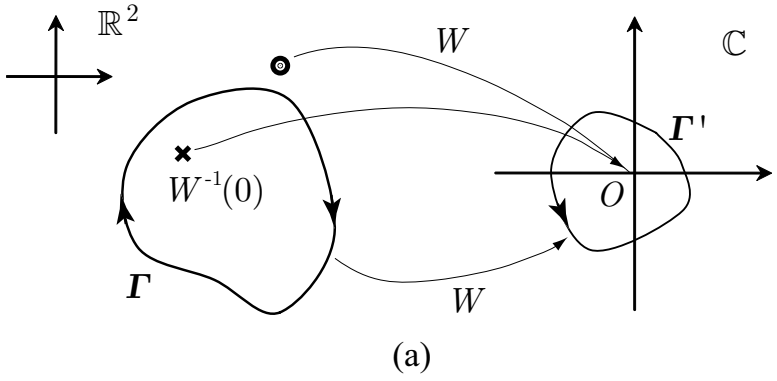
$$\mathcal{U}(W(r), \Gamma) = 2\pi \cdot \mathrm{Lk}(\Gamma, W^{-1}(0)) \tag{8.25}$$

Theorem 8.2.10 reveals more realistic pictures about phase unwrapping over closed path, which turns out to be 2π multiplies of the linking number between the chosen path and $W^{-1}(0)$, which is sometimes referred as a phase singularity. In two dimensional space, $W^{-1}(0)$ is a set of oriented point pairs, such as shown in Figure 8.4 (a) where we used crosses and circles to denote the orientation of the points (note that the crosses and circles are always in pairs since $W^{-1}(0)$ is a closed manifold); whereas in $\mathbb{R}^3$, it manifests itself as an oriented closed curve or curves, as in Figure 8.4 (b), or even more complicated geometries [11].

8.2.5 Static and dynamic phase unwrapping problems

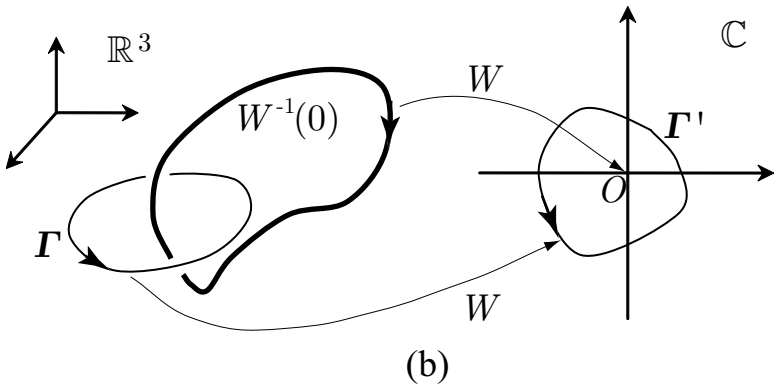
In general, there are two types of applications of the phase unwrapping integral. The first type typically involves only one static distribution of the complex field (in 2-D or 3-D) requiring the evaluation of the phase unwrapping integral at given field points with respect to specified reference points. We refer to this as the static phase unwrapping problem. From Theorem 8.2.6 the phase unwrapping integral is unambiguously defined for any valid unwrapping path. Consequently, the static phase unwrapping problem can be simplified to:

Definition 8.2.11 (Static Phase Unwrapping). *Let $W(r) = u(r) + jv(r)$, and $r \in D \subset \mathbb{R}^n$ be a complex field distribution over domain D. The observation set consists of a finite*



(a)

(a)



(b)

(b)

Figure 8.4: Mapping relationships between $\mathbb{R}^n$ and $\mathbb{C}$. The cross and circle in $\mathbb{R}^2$ and bold line in $\mathbb{R}^3$ are the pre-images of the origin in $\mathbb{C}$.

number of field points $\{r_i \in D\}_{i=1}^{N}$. *For each observation point* r_i, *a reference field point* $r_i^0 \in D$ *and a piecewise smooth path* $\Gamma_i : I \to D$ *from* r_i^0 *to* r_i *are selected. The phase unwrapping integral for each path yields the associated unwrapped phases* $\{\mathcal{U}(W(r), \Gamma_i)\}_{i=1}^{N}$.

Quite often, a single reference field point $r^0 \in D$ is used for all observation points. In that case, all unwrapping paths Γ_i share a common initial point r^0. From Theorem 8.2.6 and 8.2.9, if there is no phase singularity in the complex field distribution $W(r)$ over domain D, the results of the phase unwrapping integrals are independent of the unwrapping paths and unique solutions are produced for all observation points in domain D. However, when phase singularities appear in domain D, we can no longer recover unique unwrapped phases at observation points. In order to generate meaningful phase information, unwrapping path selection criteria must be imposed.

The second type of phase unwrapping problems involves a sequence of complex field distributions governed by a set of parameters p, denoted by $\{W(r, p(t))|r \in D\}_{t=1}^{N_t}$, where parameter p is either a single or vector valued function of time, t. In such cases, it is often essential to unwrap the phases at observation points for the complete sequence of the complex fields with additional constrains such as continuity requirements. This is called dynamic phase unwrapping and is defined below:

Definition 8.2.12 (Dynamic Phase Unwrapping). *Let* $W : \mathbb{R}^n \times T \to \mathbb{C}$ *be a differentiable function over* $\mathbb{R}^n$ *and parameter space* T, *and* $\{W(r, p(t)) = u(r, p(t)) + jv(r, p(t))\}_{t=1}^{N_t}$ *be a sequence of complex field distributions over* $D \subset \mathbb{R}^n$ *with respect to a finite parameter sequence* $p : \{1, 2, ..., N_t\} \to T$ *[107]. For each complex field distribution (subsequently defined as a frame) in the sequence, static phase unwrapping is performed. The reference point is generally pre-determined and the unwrapping paths*

are constrained by the continuity condition of the sequence usually in the form of

$$\lim_{p(t)\to p(t-1)} |\mathcal{U}(W(r, p(t)), \Gamma_i) - \mathcal{U}(W(r, p(t-1)), \Gamma_i)| = 0 \tag{8.26}$$

which results in a sequence of unwrapped phases $\{\mathcal{U}(W(r, p(t)), \Gamma_i)\}_{i=1,t=1}^{N,N_t}$ *at the observation points.*

Similar to static phase unwrapping, in the case where none of the frames contain phase singularities, for a given set of reference points, the unwrapped phases of each frame of the sequence are uniquely determined. The continuity condition is automatically satisfied since W is a continuous function of parameter p.

In the case where phase singularities do exist in some frames of the sequence, condition (8.26) will play an important role in determining the unwrapped phases in those field distributions. Since $W(r, p)$ is a continuous function of both location r and parameter p, the location of a single phase singularity in $\mathbb{R}^n$ space will continuously depend on parameter p. For $\mathbb{R}^2$ space, we expect the singularities to follow a continuous curve $\Lambda : T \to D$ connecting the locations of each phase singularity from one frame to the next if present. We call these curves the phase singularity trajectories and define the trajectory set $\{\Lambda_i\}_{i=1}^{N_s}$ as the collection of all the trajectories within the sequence, with N_s denoting the total number of the trajectories. Analogously, in $\mathbb{R}^3$ space, the trajectories of the phase singularities are surfaces. Note that, the trajectory Λ is not simply a linear connection of the phase singularity locations in discrete successive frames, but rather a continuous map from parameter space T to $\mathbb{R}^n$ space. Once the trajectory set of a complex field distribution sequence is identified, the evaluation of dynamic phase unwrapping problem satisfying condition (8.26) is more obvious. Assuming that identical reference points are used for all frames in the sequence, one useful criterion for

selecting the unwrapping paths is:

The phase unwrapping path Γ *for any observation point* r *and associated reference for a given frame shall not cross the singularity trajectory set* $\{\Lambda_i\}_{i=1}^{N_s}$ *for the complete sequence of the frames.*

The rationale for the above criterion is straightforward: given an observation point r and its associated reference, the unwrapping paths at frames with p and $p+\Delta p$ are both within the complimentary space of trajectory set in domain D, i.e. within $D\backslash\{\Lambda_i\}$. Since neither of the paths intersect the phase singularity trajectories, the region enclosed by the two paths will not contain any phase singularity. As $\Delta p \rightarrow 0$, the two paths will yield the same unwrapped phases and equation (8.26) is satisfied. However, if one path crosses the trajectory, there exists a frame at which the phase singularity falls inside the region enclosed by the two paths. Then as $\Delta p \rightarrow 0$, the unwrapped phase difference between the two paths will approach 2π and (8.26) is violated.

In Chapter 9, we will use the notion of dynamic phase unwrapping and the path selection criteria to investigate the iterative image reconstruction process in microwave imaging.

Chapter 9

Phase singularities in microwave scattering problems

Scalar or vector Helmholtz equations with associated boundary conditions are the governing equations describing the scattering phenomenon of time-harmonic electromagnetic waves [31]. The Helmholtz equation requires the existence of the second order derivatives meaning that the electric and magnetic field components solved by the equation are all continuously real-differentiable complex functions up to the continuity at internal boundaries.

Moreover, the scattering field may contain scattering nulls, where the amplitude in the complex representation of the scattering field is zero. At these nulls, the Helmholtz equation is reduced to the Laplace equation. There is a very small subset of solutions which are rank-deficient maps and satisfy

$$\begin{cases} \partial A_r/\partial x &= \alpha \partial A_i/\partial x \\ \partial A_r/\partial y &= \alpha \partial A_i/\partial y \\ \partial A_r/\partial z &= \alpha \partial A_i/\partial z \end{cases} \tag{9.1}$$

where A_r and A_i are the real and imaginary parts of the field components, respectively, and α is a nonzero constant. Spatially linear-varying static fields are examples of these rank-deficient cases. However, in most cases, the solutions of scattering problems are full-ranked maps. As a result, the conclusions from Section 8.2.3 and 8.2.4 can be applied here.

In this chapter, we will examine the phase of the scattering field of an infinitely long, lossy cylinder with an incident TM cylindrical wave to demonstrate the scattering null phenomenon and the related 2-D phase unwrapping problem. For the 3-D phase unwrapping problem, we examine the scattering field of a lossy sphere illuminated by a dipole antenna to illustrate differences in features of scattering nulls between 2-D and 3-D spaces.

9.1 Scattering nulls in 2-D problems

Consider an infinite lossy cylinder with radius ρ_a and its axis oriented along the z-axis with a line source placed in parallel to the z-axis at polar location (ρ_s, ϕ_s) (ϕ_s is zero in the configuration in Figure 9.1) with current density $\vec{J} = \hat{z}\exp(j\omega t)$ (essentially a 2-D problem). The background medium has relative permittivity ϵ_1 and conductivity σ_1 with those of the cylinder being ϵ_2 and σ_2, respectively. Assuming time dependence $\exp(j\omega t)$, the complex wave number of the background and the cylinder can be written in form of $k_i = \sqrt{\omega^2\mu_0\epsilon_0\epsilon_i - \omega\mu_0\sigma_i}$, $i = 1, 2$ where ϵ_0 and μ_0 are the free space permittivity and permeability, respectively. By separating variables and matching the boundary conditions, the series solution of the E_z component can be obtained in the similar way as in [80]. The incident or primary (p), scattered (s) and total (t) E_z field distributions in region I can be written as

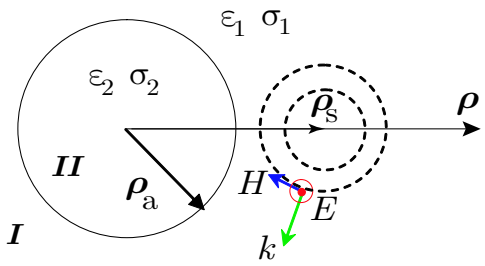


Figure 9.1: Scattering of a cylindrical TM wave by an infinite cylinder

$$\begin{aligned} E_p^I &= E_0 \sum_{n=-\infty}^{+\infty} H_n^{(2)}(k_1\rho_>) \cdot J_n(k_1\rho_<) e^{jn(\phi-\phi_s)} \\ E_s^I &= E_0 \sum_{n=-\infty}^{+\infty} c_n \cdot H_n^{(2)}(k_1\rho_>) \cdot H_n^{(2)}(k_1\rho_<) e^{jn(\phi-\phi_s)} \\ E_t^I &= E_p^I + E_s^I \end{aligned} \tag{9.2}$$

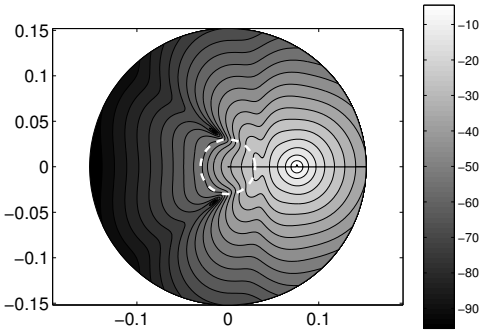
while in region II, the total (t) E_z field distribution can be written as

$$E_t^{II} = E_0 \sum_{n=-\infty}^{+\infty} d_n \cdot H_n^{(2)}(k_1\rho_s) \cdot J_n(k_2\rho) e^{jn(\phi-\phi_s)} \tag{9.3}$$

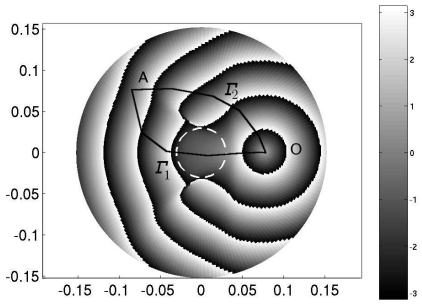
where E_0 is the amplitude, J_n stands for the n-th Bessel function of the first kind, and $H_n^{(2)}$ for the n-th Hankel function of the second kind. $\rho_> = \max(\rho, \rho_s)$ and $\rho_< = \min(\rho, \rho_s)$. In Figure 9.1, ϕ_s is simply 0. c_n and d_n are parameters defined by

$$\begin{aligned} c_n &= -\frac{k_1 J_n(k_2\rho_a) J_n'(k_2\rho_a) - k_2 J_n(k_1\rho_a) J_n'(k_2\rho_a)}{k_1 J_n(k_2\rho_a) H_n'^{(2)}(k_1\rho_a) - k_2 H_n^{(2)}(k_1\rho_a) J_n'(k_2\rho_a)} \\ d_n &= \frac{k_1 J_n(k_2\rho_a) H_n'^{(2)}(k_1\rho_a) - k_1 H_n^{(2)}(k_1\rho_a) J_n'(k_1\rho_a)}{k_1 J_n(k_2\rho_a) H_n'^{(2)}(k_1\rho_a) - k_2 H_n^{(2)}(k_1\rho_a) J_n'(k_2\rho_a)} \end{aligned}$$

When the permittivity and conductivity of the cylinder is much larger or lower than that of the background medium (i.e. high-contrast), scattering nulls will emerge in the total field. For example, in the case where the operating frequency f=800 MHz, ρ_a = 3 cm, ρ_s = 7.6 cm, where the 0.9% saline background medium has $\epsilon_1 = 76$ and $\sigma_1 = 1.7$ S/m, and the scattering cylinder (breast fat tissue [156]) has $\epsilon_2 = 5$ and $\sigma_2 = 0.1$ S/m, the contour plot of the total field amplitude (in dB) in both regions I and II is shown in



(a)



(b)

Figure 9.2: Amplitude(dB) and phase(radians) plot of the total field in regions I and II at f =800 MHz.

Figure 9.2 (a), and that of the wrapped phase, $\phi(E_t)$, is shown in Figure 9.2 (b). (The dashed white circles indicate the location and size of the scattering cylinder.) From Figure 9.2, two scattering nulls (or phase singularities) can be identified where either the amplitude abruptly drops to zero and correspondingly where the phase changes abruptly. Similarly, at a higher frequency, more nulls appear both inside and outside the cylinder (Figure 9.3). In this case, the electrical properties of the background and cylinder have been the previous case and the operating frequency is 2 GHz. In Figure 9.3, 12 phase singularities are visible.

From a wave perspective, scattering nulls can be explained as the destructive inter-

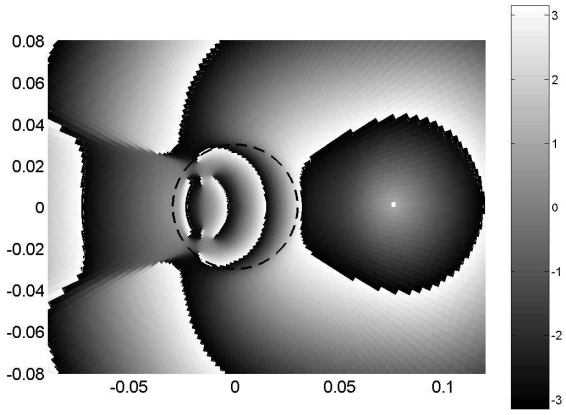


Figure 9.3: Phase plot of the total field in regions I and II at f =2 GHz.

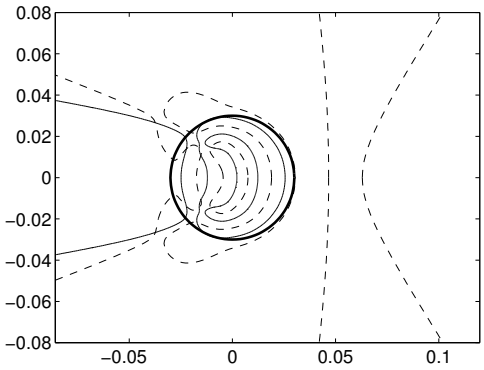


Figure 9.4: Out-of-phase curves (dash lines) and equal-amplitude curves (thin solid lines) at f =2 GHz. Their intersections illustrate the scattering null locations

ference points caused by the interaction of the incident and scattered waves. The nulls appear where the incident and scattered waves have equal amplitudes but $(2n + 1)\pi$ phase differences. For the cylinder scattering problem shown in Figure 9.3, the equal amplitude curves (Figure 9.4), where $|E_p| = |E_s|$, and the out-of-phase curves, where $\phi(E_p) = \phi(E_s)+(2n+1)\pi$, are drawn and their intersections clearly indicate the locations of the nulls.

For this problem, it is quite difficult to derive a closed solution for the complex null locations. For more complex scattering problems, numerical techniques must be applied to determine the locations.

9.2 Phase unwrapping in 2-D scattering fields

The static phase unwrapping problem in the previous example is relatively straightforward. Typically the source point is chosen as the zero-phase reference point. As discussed earlier, if the scattered field is null free, for any given observation point, arbitrary unwrapping paths will lead to identical solutions. However, if there are scattering nulls, the unwrapped phase will be path dependent. For example, in Figure 9.2 (b), path Γ_1 and path Γ_2 yield different unwrapped phase values with a 2π phase difference.

If we take either the frequency, or the dielectric properties of either the background or cylinder, ρ_s or ρ_a of the scattered field, we can form a group of dynamic phase unwrapping problems with respect to the selected parameter. For example, if the frequency is varied from 590 to 1000 MHz in 10 MHz increments in the above problem, a sequence of electrical field distributions can be obtained. For each distribution, the locations of the nulls vary. Given the continuous nature of the scattering field, a simple linear connection between the positions of the nulls in the two successive field distributions is a relatively good approximation to the continuous trajectory of the nulls for this sequence with respect to frequency change. The approximate trajectory for the frequency sequence discussed above is shown in Figure 9.5.

9.3 Phase unwrapping in 3-D scattering fields

In the 3-D scattering problem, the finite difference-time domain (FDTD) method was used to compute the scattering field of a lossy sphere ($\epsilon_r = 10$, $\sigma = 0.4$ S/m) for the same saline background with a dipole source illumination. The sphere was centered at the origin with radius $r_a = 3$ cm and the z-oriented dipole antenna was positioned at $(r_s, 0, \phi_s)$ with $r_s = 7.6$ cm and $\phi_s = \pi/2$ in spherical coordinates. A ring-like scattering

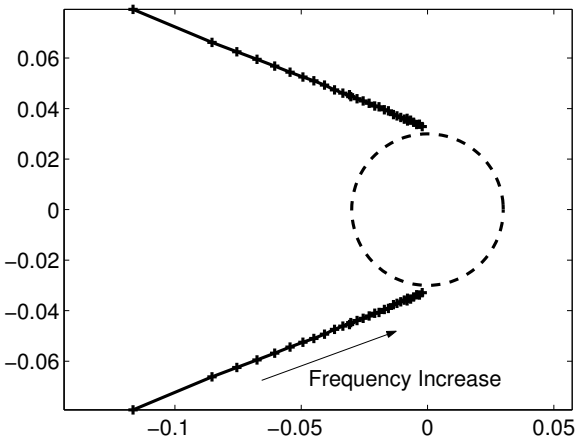


Figure 9.5: The trajectories of scattering nulls for the frequency varying from 590 to 900 MHz.

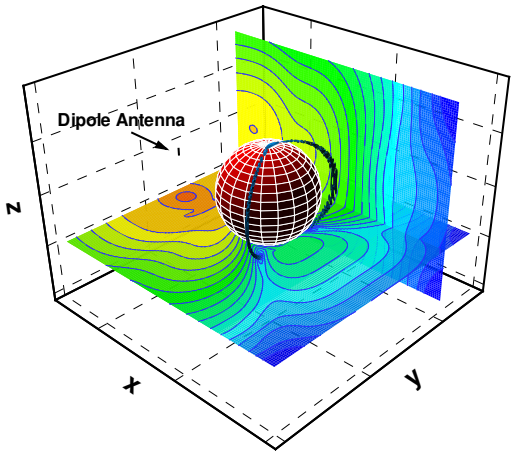


Figure 9.6: 3-D scattering null in the field scattered by a lossy sphere at f =900 MHz. The ring-like null curve is on the opposite side of the sphere with respect to the short dipole antenna location.

null is extracted from the 3-D amplitude plot of the total field at 900 MHz which is shown in Figure 9.6 (contour plots of the field magnitudes are also shown in the figure for two orthogonal planes).

For the dynamic phase unwrapping problem in 3-D, the phase singularity trajectories due to gradual changes in the selected parameter will form a surface, referred to

as a trajectory surface. A set of unwrapping paths that does not intersect the trajectory surface will yield unambiguous unwrapped phase satisfying the continuity condition.

Chapter 10

Applications of phase unwrapping theory in microwave image reconstruction

10.1 Method

As described in Chapters 3 and 6, in our log-magnitude phase form (LMPF) reconstruction algorithm, both the complex measurement and computed field data are transformed into their amplitude and phase components. A phase unwrapping process is applied to the phase portion with respect to the transmitter reference to make it continuous and differentiable with respect to dielectric properties. Over a relatively large contrast range and large imaging targets, this algorithm exhibits faster convergence and yields superior images. However, when the target is large and the contrast is high, scattering nulls can appear in the domain. Without properly choosing the unwrapping paths, the algorithm may diverge. Even when the measured field data does not exhibit any complex null be-

havior, nulls can occur in the computed distributions at intermediate iteration steps. If not accounted for properly, these phase singularities can cause the algorithm to diverge to an unwanted solution.

Unwrapping the computed field phases in the Gauss-Newton iterative reconstruction approach can be regarded as a dynamic phase unwrapping problem, where the complex dielectric property $k(r)$ varies from one iteration to the next as the algorithm converges to a solution as discussed in Definition 8.2.12. The continuity condition of the unwrapped phase imposed by the Gauss-Newton method requires the existence of the first order derivative:

$$\lim_{\Delta k \to 0} \frac{\|\mathcal{U}(E_z(k+\Delta k),\Gamma) - \mathcal{U}(E_z(k),\Gamma)\|}{\|\Delta k\|} < \infty \tag{10.1}$$

where $\|\cdot\|$ denotes the l^2-norm. As a result, the conclusions from Section 8.2.5 can be applied directly to this situation. For convenience, we define the problem configuration such that the source point is always the reference point and the unwrapping paths are fixed for all measurement sites over the full set of iterations (or frames). If the unwrapping paths do not have intersections with the trajectory set associated with parameter k, then the unwrapped phases will be defined unambiguously and the continuity condition (8.26) is satisfied.

This strategy is relatively straightforward, however, the key to its success is developing an algorithm for effectively detecting the cross over between unwrapping path null trajectories. For 2-D or 3-D image reconstruction cases, we have devised a two-path unwrapping strategy to cope with scattering nulls associated with high-contrast scatterers. A diagram of this method in a tomographic imaging context is shown in Figure 10.2. In this algorithm, we assume that the reconstruction process is initialized at a low contrast state where no nulls are present in the domain. Therefore, at the first

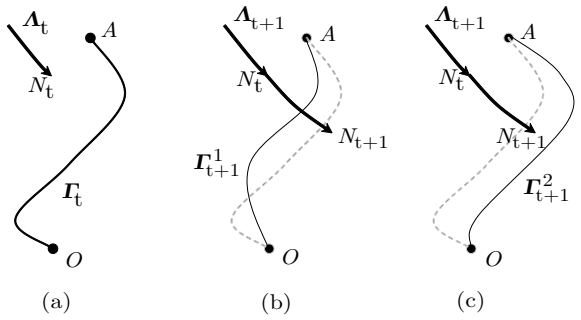


Figure 10.1: Selection of an unwrapping path during image reconstruction. (a) unwrapping path at the t-th iteration, (b) invalid unwrapping path, (c) valid unwrapping path.

iteration, the unwrapped phases computed from two separate paths will be identical. In all subsequent iterations, we compute the unwrapped phases $\Phi^t_{\Gamma_A}$ and $\Phi^t_{\Gamma_B}$ along path Γ_A and Γ_B. Comparing absolute differences $|\Phi^t_{\Gamma_A} - \Phi^{t-1}|$ and $|\Phi^t_{\Gamma_B} - \Phi^{t-1}|$, we choose the path corresponding to the smaller difference as the valid path and use its unwrapped phase for iteration t.

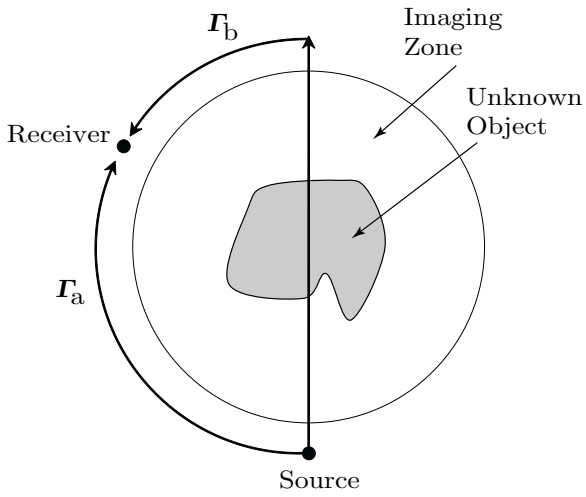


Figure 10.2: Schematic plot of the two-path unwrapping strategy used in microwave tomographical imaging reconstruction.

Below, we present results utilizing both simulated and experimental (i.e. an *in vivo* breast imaging example) data. In the former, we illustrate the effectiveness when inter-

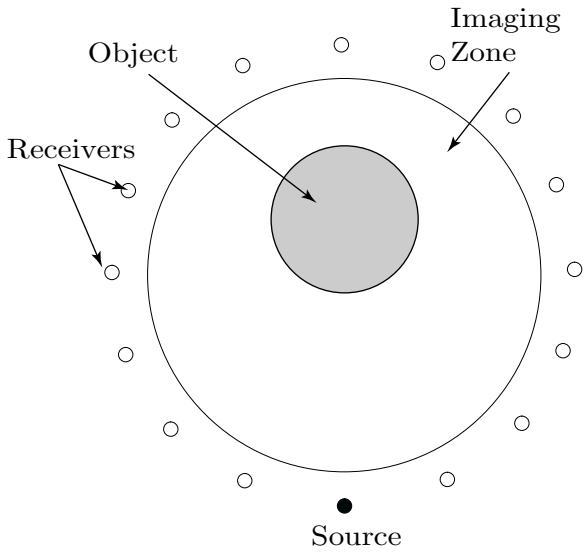


Figure 10.3: Schematic diagram of the imaging configuration for the simulation.

mediate complex nulls appear in the computed forward solution. The latter illustrates a case where the algorithm has been used to reconstruct an image of a breast which was not possible without the use of the new unwrapping strategy.

10.2 Results

10.2.1 Reconstruction with the presence of scattering nulls

Here we present an imaging example to illustrate the robustness of the algorithm consisting of a lossy cylinder with radius $\rho_a = 3cm$ and electrical properties $\epsilon_r = 10$ and $\sigma = 0.4$ S/m submerged in a background medium of 0.9% saline with $\epsilon_r = 77.1$ and $\sigma = 1.76$ S/m. The cylinder is surrounded by an antenna array on a radius of 7.6 cm consisting of 16 line sources parallel to the axis of the cylinder (Figure 10.3). A 2D dual-mesh pair is used with the circular reconstruction parameter mesh (radius $\rho = 6cm$) conformal with the imaging zone. We utilized the LMPF-MFDR reconstruction algorithm as described in Section 3.2.3 and Section 6.2 operating at 900 MHz. The

synthesized measurement data was computed by the analytical solution expressed in terms of equation (9.2). We used three unwrapping strategies for each measurement site as shown in Figure 10.4: (a) unwrapping through the domain and subsequently traversing the shortest arc to the receiver, (b) starting from the source and unwrapping along the shortest arc to the receiver and (c) the two-path strategy as described above.

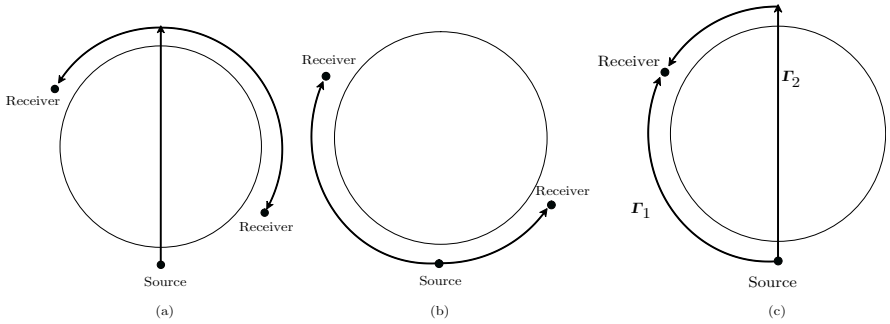


Figure 10.4: Unwrapping strategies (a) strategy A, (b) strategy B, (c) strategy C.

The contrast of the object is sufficient to excite nulls in the imaging zone for all sources. The reconstructed images are dramatically different for all three unwrapping strategies shown in Figure 10.5 (a)-(c). The conductivity images for both (a) and (b) are very noisy along with the permittivity image for (b). The permittivity image for (a) appears to recover the proper value for the object but the background property is incorrect. The recovered permittivity and conductivity images for case (c) match the actual distribution quite well. In addition, Figure 10.5 (d) shows the reconstructed images using the complex Gauss-Newton algorithm instead of the LMPF algorithm and illustrates an example of the reconstruction converging to a local minima. A plot of the relative error between the calculated and measured field values as a function of iteration (Figure 10.6.) clearly indicates that the first two strategies diverged due to inadequate processing of the scattering nulls. In both cases, the nulls appeared within the imaging zone during the 4th iteration.

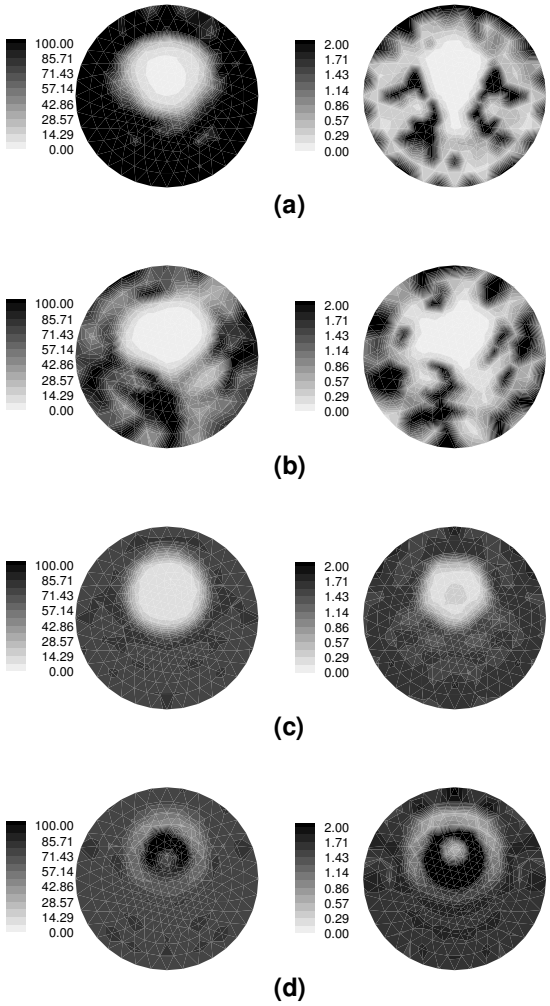


Figure 10.5: Reconstructed permittivity and conductivity images using the different unwrapping strategies. (a) unwrapping through imaging zone and subsequently with the shortest arc to receiver (b) shortest arc to receiver, (c) 2-path strategy and (d) the complex Gauss-Newton reconstruction.

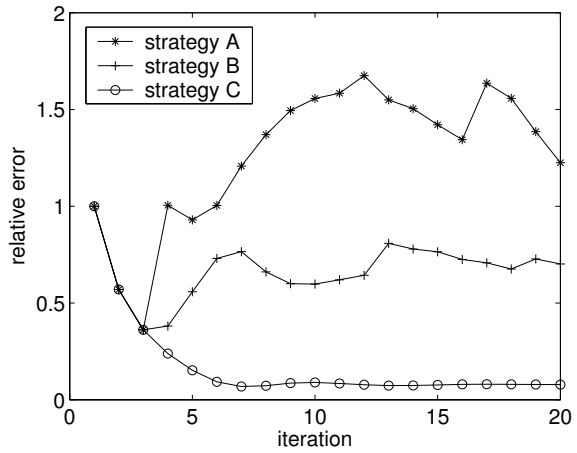


Figure 10.6: Relative errors of the three unwrapping strategies with respect to iteration number

10.2.2 Reconstruction with intermediate nulls

The notion of intermediate nulls refers to cases where the phase singularities do not appear in the true scattering field; however, they are created and propagate into the imaging zone during intermediate steps of the iterative reconstruction process and eventually exit the zone by the time the algorithm has converged (if it does converge to an adequate solution). During the Gauss-Newton iterative process, the reconstruction parameters do not necessarily converge monotonically. In fact, the values often overshoot the final solution and very often oscillate about the desired values until the oscillations are almost completely damped at convergence. If the unwrapped phase continuity condition is violated at intermediate steps, it could significantly alter the algorithm behavior and cause it to diverge. These types of nulls are readily processed by the two-path unwrapping strategy.

The 2-D simulation example we presented here is for a "panda face" pattern shown in Figure 10.7. A 15.2 cm diameter circular antenna array consisting of 16 dipole antennas encircled the object. The imaging zone is a 14 cm diameter concentric circular region. The "panda face" is a 9 cm diameter circle with the diameters for the eyes and

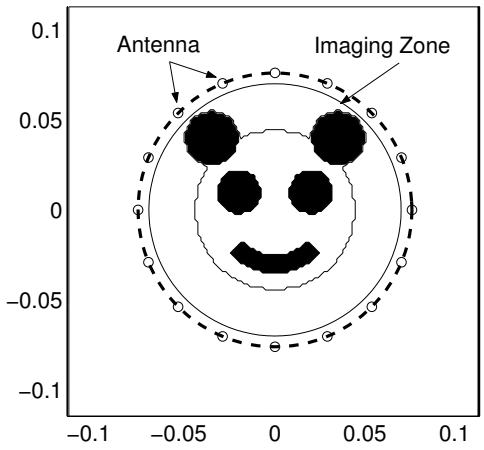


Figure 10.7: Schematic diagram of the object and imaging configuration (dimensions in meters).

ears being 2.4 and 3 cm, respectively. The "panda mouth" is a quarter of a concentric annulus with an inner radius of 2.5 cm and outer radius of 3.5 cm between the angle of $-3\pi/4$ and $-\pi/4$. The electrical properties of the background and different zones of the "panda face" are listed in Table 1.

Regions	ϵ_r	$\sigma(S/m)$
background	76.9	1.8
panda face	55.0	1.2
panda eyes	15.0	0.3
panda ears	30.5	0.6
panda mouth	15.0	0.3

Table 10.1: The exact relative permittivity and conductivity values at 1000 MHz for all zones in the "panda face" simulation.

The measurement data was computed at 1 GHz using a 2-D finite difference-time domain (FDTD) method with a generalized perfectly matched layer(G-PML) as the absorbing boundary condition [51]. The forward solution domain was a 110×110 grid surrounded by 12 layers of G-PML. For all 16 sources, the electric fields for the exact property distribution do not contain any scattering nulls. We added noise (maximum amplitude of -100 dB and 1° phase) to the amplitude and phase data respectively, which is representative of our current hardware system [122]. The reconstruction utilized

the LMPF algorithm. The reconstruction mesh conformed to the imaging zone and was comprised of 281 parameter nodes with 524 triangular elements. The Levenberg-Marquardt regularization parameter, λ, was fixed at 0.05 together with our spatial filtering scheme (Section 3.2.3) with the averaging factor set to 0.1 for stabilizing the convergence. The algorithm was initialized with a homogeneous distribution equal to that of the background.

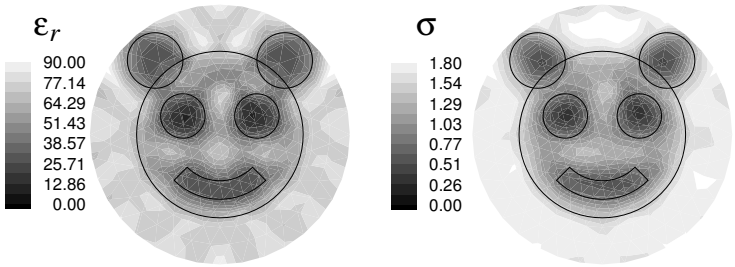


Figure 10.8: Recovered dielectric profiles after 10 iterations using the two-path unwrapping strategy: (a) relative permittivity, (b) conductivity.

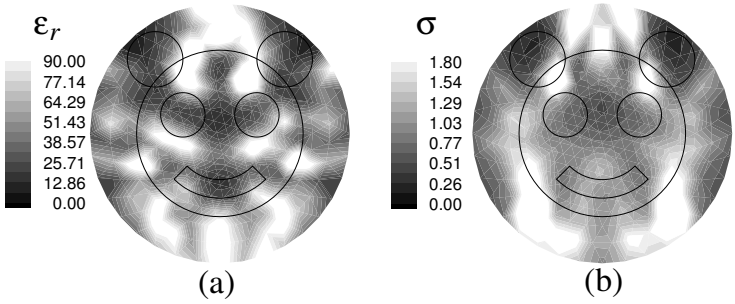


Figure 10.9: Recovered dielectric profiles after 10 iterations without considering the scattering nulls: (a) relative permittivity, (b) conductivity.

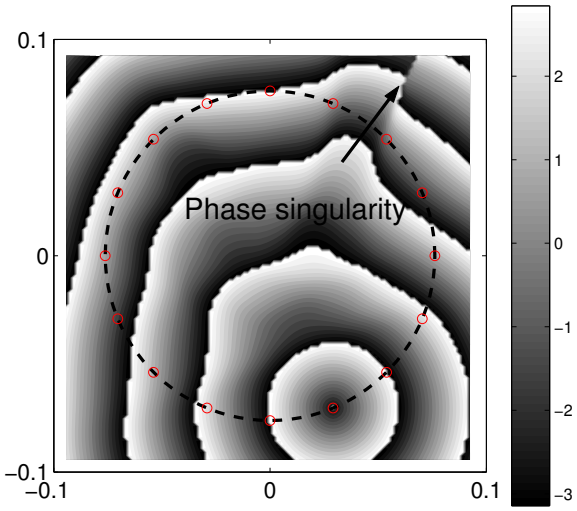
Using the two-path unwrapping strategy, the object was successfully reconstructed after 10 iterations with a relative electric field least squared error of 5% as shown in Figure 10.8 (a) and (b). The locations, shapes and values of the features in the pattern are quite close to their true values. It is interesting to note that at the third iteration,

intermediate nulls appeared in the computed field solutions for several sources and disappeared after the fourth iteration. Wrapped phase plots for the true object scattering field distribution and the forward field computation at the third iteration for a single antenna are compared in Figure 10.10. Note that the scattering null migrated inside the antenna array for the third iteration (Figure 10.10 b), but retreated immediately after that. Given that the unwrapping paths are usually either along the arc of the antenna or along a path through the imaging zone, phase singularities within this zone are the ones that impact the unwrapping. Without the correct phase unwrapping strategy, the reconstruction diverged quickly after the third iteration (the solution after the 10th iteration is shown in Figure 10.9).

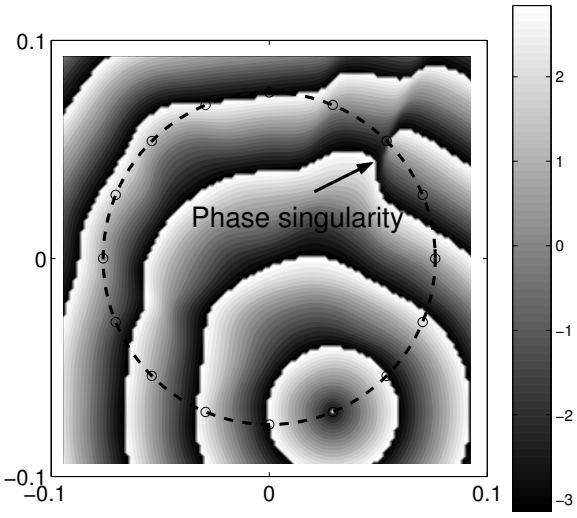
10.2.3 Reconstruction of patient measurement

The scattering nulls together with intermediate nulls are frequently encountered in the processing of breast cancer patient measurement data especially where there are high contrast inclusions such as large tumors or cysts. Even for normal breasts, the scattered field from the glandular tissue may also induce scattering nulls. A sample MRI image is shown in Figure 10.11 to demonstrate inhomogeneities in a normal breast due to its internal structures. In these cases, in order to use the LMPF algorithm, we must incorporate the two-path unwrapping strategy into the reconstruction algorithm to obtain valid unwrapped phases.

In this example we reconstructed an image slice of a patient's breast where the woman was being treated with chemotherapy for a large tumor. The measurement data was obtained with the tomographic microwave imaging system as in [122]. The patient had a large tumor located at her upper half breast close to chest wall. The high contrast of the tumor to the fatty tissue background in the breast caused multiple phase



(a)



(b)

Figure 10.10: Wrapped phase plots for (a) the true scattering field, (b) the forward field computation at the 3rd iteration for a single transmitter (singularity present).

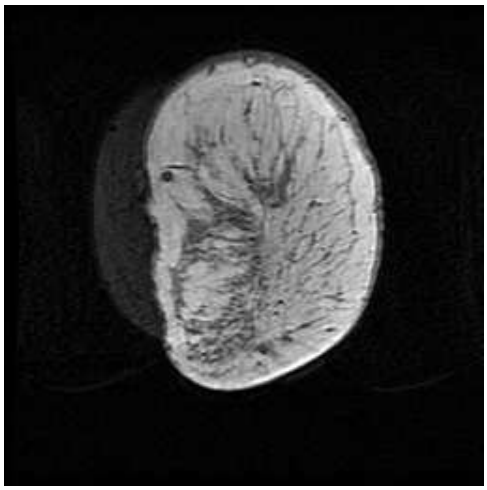


Figure 10.11: MRI scan of a normal breast. The dark regions are fibroglandular tissues which may have significantly high dielectric property values compared with the fatty tissues in the background.

singularities during the reconstruction. We utilized the LMPF reconstruction with and without the two-path unwrapping strategy, of which only the former one yielded reasonable images (Figure 10.12) while the latter diverged. The tumor is clearly visible in the reconstructed image which is in the correct location and has reasonably appropriate values for the electrical properties of a typical tumor.

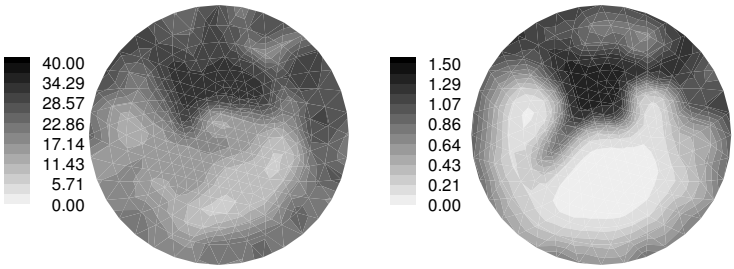


Figure 10.12: Reconstructed single plane dielectric profiles of a patient breast that has a large tumor, left: relative permittivity, right: conductivity.

10.3 Conclusions

In summary, we have established a general mathematical framework for explaining phase unwrapping including definitions and illustrations of particular properties related to the uniqueness and closed-path phase unwrapping. The concept of dynamic versus static phase unwrapping problems was introduced with special attention to applications in microwave imaging. These included the phenomenon of scattering nulls in the high contrast and high operating frequency cases and their behavior (i.e. paths of their trajectories) as these parameters varied.

The path selection criteria for the dynamic phase unwrapping problem was implemented in several microwave tomographic image reconstruction examples. The challenges of utilizing the LMPF algorithm were discussed from a dynamic phase unwrapping perspective along with efficient unwrapping strategies. The success of these reconstructions demonstrates the importance and efficiency of the our theory and analysis.

Appendix A

Mathematica code for the ADI FDTD method update equations

The Mathematica code for the ADI update equation of the first substep, i.e. $n \rightarrow n+1/2$, is written as

```
(*********************************************************************)
(********         ADI FDTD step 1: n->n+1/2                 ********)
(*********************************************************************)
(*=========================Update Ex================================*)
Px[i_,j_,k_]:=cAP[i,j-1,k-1]*Px_o[i,j,k]+
              cBP[i,j-1,k-1]*((Hz[i,j,k]-Hz[i,j-1,k])-(Hy_o[i,j,k]-Hy_o[i,j,k-1]));
Qx[i_,j_,k_]:=cAEy[j-1]*Qx_o[i,j,k]+cBEy[j-1]*(Px[i,j,k]-Px_o[i,j,k]);
Ex[i_,j_,k_]:=cAEz[k-1]*Ex_o[i,j,k]+cBEz[k-1]*(cCHx[i]*Qx[i,j,k]-cDHx[i]*Qx_o[i,j,k]);

(*=========================Update Ey================================*)
Py[i_,j_,k_]:=cAP[i-1,j,k-1]*Py_o[i,j,k]+
              cBP[i-1,j,k-1]*((Hx[i,j,k]-Hx[i,j,k-1])-(Hz_o[i,j,k]-Hz_o[i-1,j,k]));
Qy[i_,j_,k_]:=cAEz[k-1]*Qy_o[i,j,k]+cBEz[k-1]*(Py[i,j,k]-Py_o[i,j,k]);
Ey[i_,j_,k_]:=cAEx[i-1]*Ey_o[i,j,k]+cBEx[i-1]*(cCHy[j]*Qy[i,j,k]-cDHy[j]*Qy_o[i,j,k]);
```

```
(*=========================Update Ez=============================*)
Pz[i_,j_,k_]:=cAP[i-1,j-1,k]*Pz_o[i,j,k]+
              cBP[i-1,j-1,k]*((Hy[i,j,k]-Hy[i-1,j,k])-(Hx_o[i,j,k]-Hx_o[i,j-1,k]));
Qz[i_,j_,k_]:=cAEx[i-1]*Qz_o[i,j,k]+cBEx[i-1]*(Pz[i,j,k]-Pz_o[i,j,k]);
Ez[i_,j_,k_]:=cAEy[j-1]*Ez_o[i,j,k]+cBEy[j-1]*(cCHz[k]*Qz[i,j,k]-cCHz[k]*Qz_o[i,j,k]);

(*=========================Update Hx=============================*)
Bx[i_,j_,k_]:=cAHy[j]*Bx_o[i,j,k]+
              cBBy[j]*((Ey[i,j,k+1]-Ey[i,j,k])-(Ez_o[i,j+1,k]-Ez_o[i,j,k]));
Hx[i_,j_,k_]:=cAHz[k]*Hx_o[i,j,k]+cBHz[k]*(cCEx[i-1]*Bx[i,j,k]-cDEx[i-1]*Bx_o[i,j,k]);

(*=========================Update Hy=============================*)
By[i_,j_,k_]:=cAHz[k]*By_o[i,j,k]+
              cBBz[k]*((Ez[i+1,j,k]-Ez[i,j,k])-(Ex_o[i,j,k+1]-Ex_o[i,j,k]));
Hy[i_,j_,k_]:=cAHx[i]*Hy_o[i,j,k]+cBHx[i]*(cCEy[j-1]*By[i,j,k]-cDEy[j-1]*By_o[i,j,k]);

(*=========================Update Hz=============================*)
Bz[i_,j_,k_]:=cAHx[i]*Bz_o[i,j,k]+
              cBBx[i]*((Ex[i,j+1,k]-Ex[i,j,k])-(Ey_o[i+1,j,k]-Ey_o[i,j,k]));
Hz[i_,j_,k_]:=cAHy[j]*Hz_o[i,j,k]+cBHy[j]*(cCEz[k-1]*Bz[i,j,k]-cDEz[k-1]*Bz_o[i,j,k]);
```

where (i, j, k) is the index of the 3D field arrays and index (1,1,1) for each array is located at the closest vector near the origin in Figure 5.5. The field components with their name ended with "_o" denote the field at n-th time step, while those without this suffix represent the field at the $n + 1/2$ time step. Similarly, the relationships for the second substep, i.e. $n + 1/2 \rightarrow n + 1$ is written as

```
(*****************************************************************)
(********        ADI FDTD step 2: n+1/2->n+1              ********)
(*****************************************************************)
(*=========================Update Ex=============================*)
```

```
Px[i_,j_,k_]:=cAP[i,j-1,k-1]*Px_o[i,j,k]+
              cBP[i,j-1,k-1]*((Hz_o[i,j,k]-Hz_o[i,j-1,k])-(Hy[i,j,k]-Hy[i,j,k-1]));
Qx[i_,j_,k_]:=cAEy[j-1]*Qx_o[i,j,k]+cBEy[j-1]*(Px[i,j,k]-Px_o[i,j,k]);
Ex[i_,j_,k_]:=cAEz[k-1]*Ex_o[i,j,k]+cBEz[k-1]*(cCHx[i]*Qx[i,j,k]-cDHx[i]*Qx_o[i,j,k]);

(*=========================Update Ey==============================*)
Py[i_,j_,k_]:=cAP[i-1,j,k-1]*Py_o[i,j,k]+
              cBP[i-1,j,k-1]*((Hx_o[i,j,k]-Hx_o[i,j,k-1])-(Hz[i,j,k]-Hz[i-1,j,k]));
Qy[i_,j_,k_]:=cAEz[k-1]*Qy_o[i,j,k]+cBEz[k-1]*(Py[i,j,k]-Py_o[i,j,k]);
Ey[i_,j_,k_]:=cAEx[i-1]*Ey_o[i,j,k]+cBEx[i-1]*(cCHy[j]*Qy[i,j,k]-cDHy[j]*Qy_o[i,j,k]);

(*=========================Update Ez==============================*)
Pz[i_,j_,k_]:=cAP[i-1,j-1,k]*Pz_o[i,j,k]+
              cBP[i-1,j-1,k]*((Hy_o[i,j,k]-Hy_o[i-1,j,k])-(Hx[i,j,k]-Hx[i,j-1,k]));
Qz[i_,j_,k_]:=cAEx[i-1]*Qz_o[i,j,k]+cBEx[i-1]*(Pz[i,j,k]-Pz_o[i,j,k]);
Ez[i_,j_,k_]:=cAEy[j-1]*Ez_o[i,j,k]+cBEy[j-1]*(cCHz[k]*Qz[i,j,k]-cCHz[k]*Qz_o[i,j,k]);

(*=========================Update Hx==============================*)
Bx[i_,j_,k_]:=cAHy[j]*Bx_o[i,j,k]+
              cBBy[j]*((Ey_o[i,j,k+1]-Ey_o[i,j,k])-(Ez[i,j+1,k]-Ez[i,j,k]));
Hx[i_,j_,k_]:=cAHz[k]*Hx_o[i,j,k]+cBHz[k]*(cCEx[i-1]*Bx[i,j,k]-cDEx[i-1]*Bx_o[i,j,k]);

(*=========================Update Hy==============================*)
By[i_,j_,k_]:=cAHz[k]*By_o[i,j,k]+
              cBBz[k]*((Ez_o[i+1,j,k]-Ez_o[i,j,k])-(Ex[i,j,k+1]-Ex[i,j,k]));
Hy[i_,j_,k_]:=cAHx[i]*Hy_o[i,j,k]+cBHx[i]*(cCEy[j-1]*By[i,j,k]-cDEy[j-1]*By_o[i,j,k]);

(*=========================Update Hz==============================*)
Bz[i_,j_,k_]:=cAHx[i]*Bz_o[i,j,k]+
              cBBx[i]*((Ex_o[i,j+1,k]-Ex_o[i,j,k])-(Ey[i+1,j,k]-Ey[i,j,k]));
```

```
Hz[i_,j_,k_]:=cAHy[j]*Hz_o[i,j,k]+cBHy[j]*(cCEz[k-1]*Bz[i,j,k]-cDEz[k-1]*Bz_o[i,j,k]);
```

where the symbols with suffix "_o" represent the fields at time step $n + 1/2$ and those without the suffix are at time step $n + 1$. To explicitly solve for fields at the newer time step from the above relationships, the following code is submitted in Mathematica. For example, to compute the update equation for D_x for the first sub-step, we need to execute the following code

```
(*****************************************************************)
(********        ADI FDTD step 1: n->n+1/2                ********)
(*****************************************************************)
Px[i_,j_,k_]=.;
RHS =cAP[i,j-1,k-1]*Px_o[i,j,k]+
     cBP[i,j-1,k-1]*((Hz[i,j,k]-Hz[i,j-1,k])-(Hy_o[i,j,k]-Hy_o[i,j,k-1]));
a1=-Coefficient[RHS-Px[i,j,k],Px[i,j-1,k]]//Simplify;
a2=-Coefficient[RHS-Px[i,j,k],Px[i,j,k]]//Simplify;
a3=-Coefficient[RHS-Px[i,j,k],Px[i,j+1,k]]//Simplify;
a4=RHS-Px[i,j,k]-a1*Px[i,j-1,k]-a2*Px[i,j,k]-a3*Px[i,j+1,k]//Simplify;
{a1, a2, a3, a4} // TableForm
```

which gives the implicit relationship in form of

$$a_1 P_x[i, j-1, k] + a_2 P_x[i, j, k] + a_3 P_x[i, j+1, k] = a_4 \tag{A.1}$$

where

```
a1=-cBBx[i]cBEy[j-2]cBEz[k-1]cBHy[j-1]cBP[i,j-1,k-1]cCEz[k-1]cCHx[i]
a2=1+cBBx[i]cBEy[j-1]cBEz[k-1](cBHy[j-1]+cBHy[j])cBP[i,j-1,k-1]cCEz[k-1]cCHx[i]
a3=-cBBx[i]cBEy[j]cBEz[k-1]cBHy[j]cBP[i,j-1,k-1]cCEz[k-1]cCHx[i]
a4=cBBx[i]cBEy[j-2]cBEz[k-1]cBHy[j-1]cBP[i,j-1,k-1]cCEz[k-1]cCHx[i]Px[i,j-1,k]-
   Px[i,j,k]-(1+cBBx[i]cBEy[j-1]cBEz[k-1](cBHy[j-1]+
   cBHy[j])cBP[i,j-1,k-1]cCEz[k-1]cCHx[i])Px[i,j,k]+
   cBBx[i]cBEy[j]cBEz[k-1]cBHy[j]cBP[i,j-1,k-1]cCEz[k-1]cCHx[i]Px[i,j+1,k]+
```

```
cAP[i,j-1,k-1]Px_o[i,j,k]+cBP[i,j-1,k-1](Hy_o[i,j,k-1]-Hy_o[i,j,k]-
cAHy[j-1]Hz_o[i,j-1,k]+cAHy[j]Hz_o[i,j,k]-cBHy[j-1](-cDEz[k-1]Bz_o[i,j-1,k]+
cCEz[k-1](cAHx[i]Bz_o[i,j-1,k]+cBBx[i](-cAEz[k-1]Ex_o[i,j-1,k]+
cAEz[k-1]Ex_o[i,j,k]+Ey_o[i,j-1,k]-Ey_o[1+i,j-1,k]-
cBEz[k-1](-cDHx[i]Qx_o[i,j-1,k]+cCHx[i](cBEy[j-2](Px[i,j-1,k]-
Px_o[i,j-1,k])+cAEy[j-2]Qx_o[i,j-1,k]))+cBEz[k-1](-cDHx[i]Qx_o[i,j,k]+
cCHx[i](cBEy[j-1](Px[i,j,k]-Px_o[i,j,k])+cAEy[j-1]Qx_o[i,j,k])))))+
cBHy[j](-cDEz[k-1]Bz_o[i,j,k]+cCEz[k-1](cAHx[i]Bz_o[i,j,k]+
cBBx[i](-cAEz[k-1]Ex_o[i,j,k]+cAEz[k-1]Ex_o[i,j+1,k]+Ey_o[i,j,k]-
Ey_o[1+i,j,k]-cBEz[k-1](-cDHx[i]Qx_o[i,j,k]+cCHx[i](cBEy[j-1](Px[i,j,k]-
Px_o[i,j,k])+cAEy[j-1]Qx_o[i,j,k]))+cBEz[k-1](-cDHx[i]Qx_o[i,j+1,k]+
cCHx[i](cBEy[j](Px[i,j+1,k]-Px_o[i,j+1,k])+cAEy[j]Qx_o[i,j+1,k])))))
```

A tridiagonal matrix equation is formed by cascading (A.1) for P_x at various j indices, which can be efficiently solved by traditional matrix solvers. The implicit equation of P_y and P_z can be derived in a similar fashion. Notice that the coefficients of the tridiagonal matrix, i.e. `a1,a2` and `a3` do not contain any field qualities, so that the LU decompositions of these tridiagonal matrices can be performed in advance of the time stepping and only back-substitutions are required in each time-step.

Once the values of $\vec{P}$ are computed, the remaining field vectors, i.e. $\vec{Q}$, $\vec{E}$, $\vec{B}$ and $\vec{H}$, can be updated explicitly with the traditional UPML scheme (as the expressions in the Mathematica code) since their RHS's are already computed. Similarly, the relationships for the second sub-step can be derived utilizing the identical process.

Appendix B

Statistical analysis of the reconstruction algorithm with measurement data

As was demonstrated in Chapter 2, the statistical properties of the measurement noise is critically important in selecting appropriate parameter estimation strategies. In all reconstruction approaches used in this thesis, we chose the OLS (ordinary least-square) estimator for reconstructing the dielectric properties which assumes that the measurement noise is additive and iid (identical independently distributed) satisfying the normal distribution with zero mean and constant variance (Section 2.5). However, we did not characterize the actual measurement noise properties obtained from our imaging system to justify the above assumptions. This is the central task of this appendix.

A series of experiments were performed to facilitate the investigation to the measurement noise. The scattered fields (in terms of dB amplitude and phase) were repeatedly measured 18 times with a small cylindrical object inside the imaging zone (the

smaller cylinder used in Section 3.7.2) and is referred to as the raw measurement. In the raw measurement, the amplitude and phase data are the differences between the field scattered by the inhomogeneous structure and by the homogeneous background medium, i.e.

$$\begin{aligned} \boldsymbol{\Delta\gamma}_{dB} &= |\mathbf{E}^{\text{inhomo}}|_{dB} - |\mathbf{E}^{\text{homo}}|_{dB} \\ \boldsymbol{\Delta\phi} &= \Phi(\mathbf{E})^{\text{inhomo}} - \Phi(\mathbf{E})^{\text{homo}} \end{aligned} \tag{B.1}$$

where the superscript "inhomo" refers to the case with the target inside the imaging zone while "homo" refers to the homogeneous background medium case; $|\cdot|_{dB}$ denotes the dB amplitude and $\Phi(\cdot)$ denotes the phase. The quantities on the LHS are the raw measurement which can be subsequently converted into complex form as

$$\begin{aligned} \boldsymbol{\Delta}\mathbf{E}_R &= 10^{\frac{\boldsymbol{\Delta\gamma}_{dB}}{20}} * \cos(\boldsymbol{\Delta\phi}) \\ \boldsymbol{\Delta}\mathbf{E}_I &= 10^{\frac{\boldsymbol{\Delta\gamma}_{dB}}{20}} * \sin(\boldsymbol{\Delta\phi}) \end{aligned} \tag{B.2}$$

The subtraction of $\mathbf{E}^{\text{homo}}$, referred to as the calibration data, from $\mathbf{E}^{\text{inhomo}}$ in (B.1) can significantly reduce the systematic error of the imaging system including the cancellation of the the gain imbalance between different channels, phase shift due to varied cable lengths and so on. However, from the reconstruction algorithm perspective, only the raw measurement data on the LHS, i.e. $\boldsymbol{\Delta\gamma}_{dB}$ and $\boldsymbol{\Delta\phi}$, are the input quantities. Prior to the actual iterative reconstruction, a forward field solution with the homogeneous background is computed, denoted as E_c^{homo}, and the restored inhomogeneous field is written as

$$\mathbf{E}_r^{\text{inhomo}} = \mathbf{E}_c^{\text{homo}} \times (\boldsymbol{\Delta}\mathbf{E}_R + j\boldsymbol{\Delta}\mathbf{E}_I) \tag{B.3}$$

and from which the dielectric properties are reconstructed. For a fixed forward method and a given background medium, $\mathbf{E}_c^{\text{homo}}$ is a constant distribution, therefore, the statistical properties of $(\boldsymbol{\Delta}\mathbf{E}_R + j\boldsymbol{\Delta}\mathbf{E}_I)$ should be analyzed.

To evaluate the appropriateness of the parameter estimation method, we first assume

that the forward model is accurate. In other words, if the imaging system is noise free, the predicted measurement computed from the forward model should be identical to the actual measurement. In this case, we need to follow the traditional parameter estimation theory (Section 2.5) and investigate the noise in the raw measurement data. However, for reconstructions using real measured data, the forward model only has limited accuracy. The justification of the estimation model becomes more difficult. A practically useful strategy is to look at the statistical properties of the residual error $\mathbf{\Delta E}_{\text{res}}$ produced by the iterative reconstruction algorithm [41]. The residual error can be expressed as

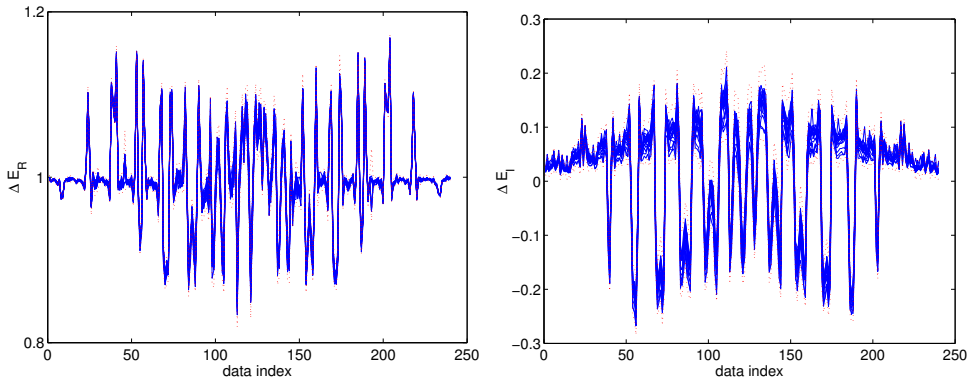
$$\mathbf{\Delta E}_{\text{res}} = \mathbf{E}_r^{\text{inhomo}} - F(\mathbf{k}_{\text{recon}}^2) \tag{B.4}$$

where $\mathbf{E}_r^{\text{inhomo}}$ is the LHS of (B.3), F denotes the forward model and $\mathbf{k}_{\text{recon}}^2$ is the reconstructed dielectric property vector. The residual error reflects both the influences from the measurement data and the forward model accuracy. The investigation on the residual error can be found in Section B.2.

B.1 Analysis of the raw measurement

With the repeated measurement data at 1100 MHz (which is a typical frequency used in our reconstructions), the variances and the means of the data were computed at each data point, from which, the error bound diagram of the data is plotted in Figure B.1. In the figure, the dotted lines above and below all solid lines are located at $\mathbf{m} \pm 3\mathbf{s}$ where $\mathbf{m}$ is the vector of mean values and $\mathbf{s}$ is the vector of standard deviations. Note that the x-axis in the plot is the index for all the transmitter/receiver pairs. In this case, there are 16 transmitters and 15 receivers per transmitter making the total length of 240.

In order to characterize the distribution of the measurement noise, the data is first



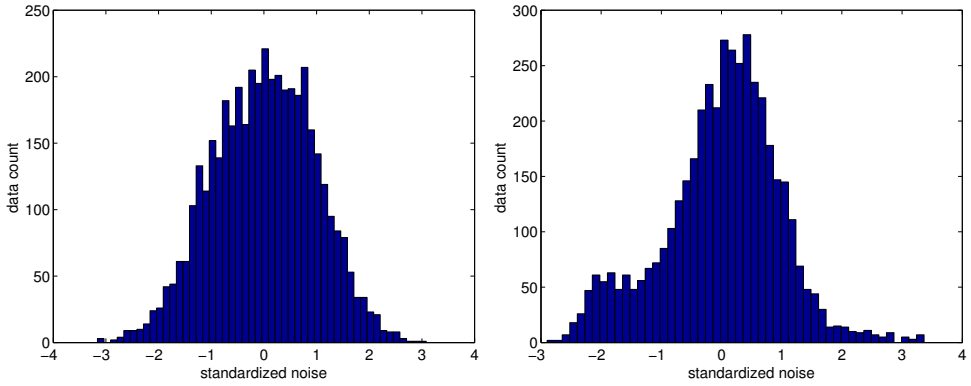
(a) (b)

Figure B.1: Error bound plots of the (a) real and (b) imaginary parts of the raw measurement.

standardized by the mean and standard deviation by

$$\tilde{\mathbf{x}} = \frac{\mathbf{x} - \mathbf{m}}{\mathbf{s}} \tag{B.5}$$

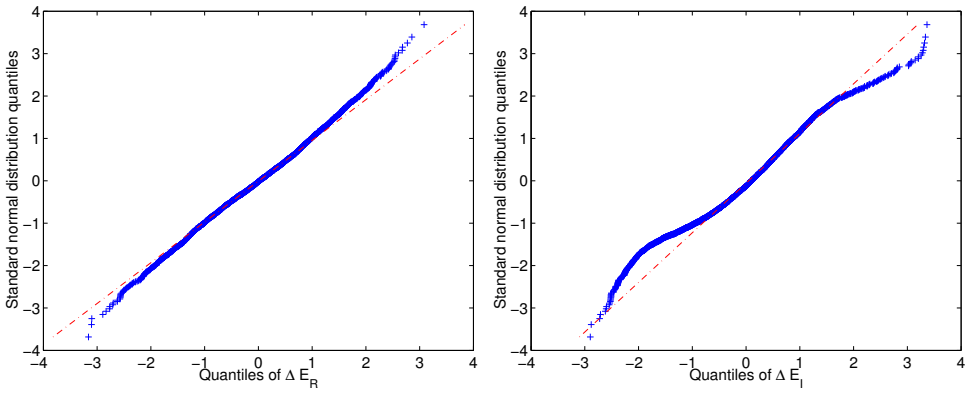
then, we plotted the histogram plots of the standardized measurement noise which are shown in Figure B.2 from which a symmetric distribution feature can be observed. The



(a) (b)

Figure B.2: Histogram plots of the (a) real and (b) imaginary parts of the standardized measurement noise.

standardized data is subsequently analyzed by the quantile-quantile plot (QQ-plot) with respect to different symmetric probability density functions (PDF) using a MATLAB software package. The output of the analysis for the real and imaginary data against the normal distribution is shown in Figure B.3.

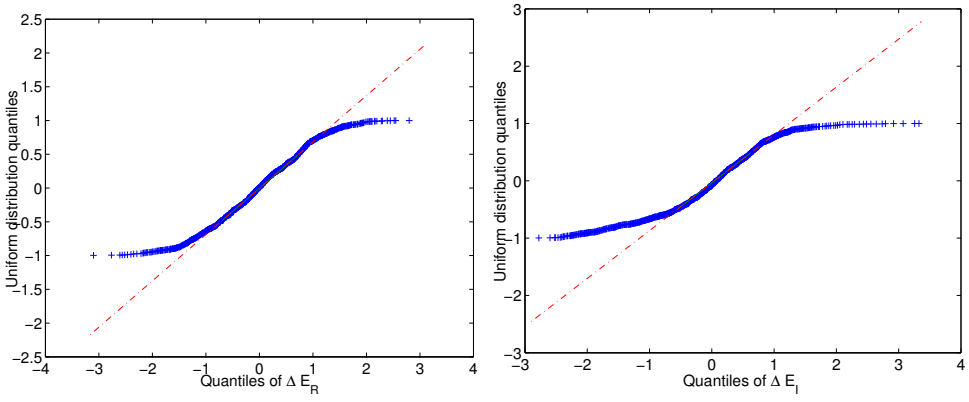


(a) (b)

Figure B.3: Quantile-quantile plots of the (a) real and (b) imaginary parts of the standardized measurement noise against normal distribution.

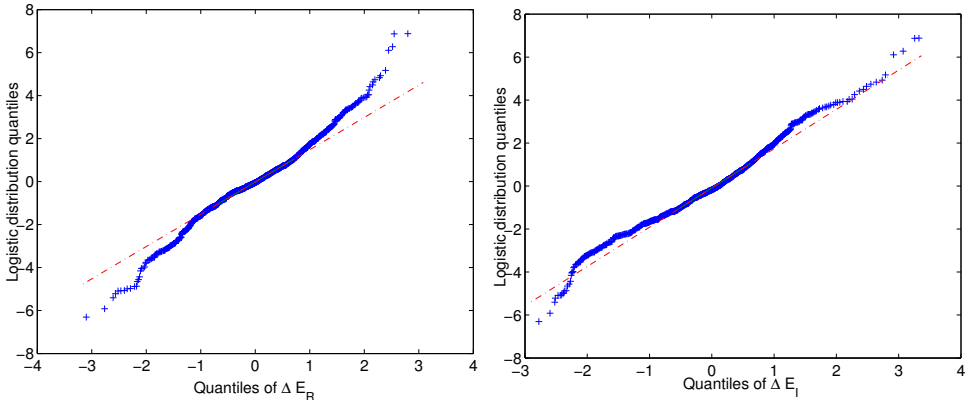
We also tested the measurement noise with respect to a uniform distribution (Figure B.4), a logistic distribution (Figure B.5) and a Laplace distribution (Figure B.6) [201]. From all of these QQ-plots, the normal distribution seems to be the most appropriate model to describe the raw measurement noise. This is an expected conclusion for most measurement systems because within these systems, a large number of independent factors effect the data. From central limit theorem [135], the summation of these random effects is approximately a normal distribution.

In our log-magnitude phase form (LMPF) reconstruction (Section 3.2.3), the log-amplitude and phase measurements are directly used in the estimation process based on minimizing sum-of-square functions. Therefore, we need to investigate the noise in the dB amplitude and phase data as well. We performed a similar analysis as that for



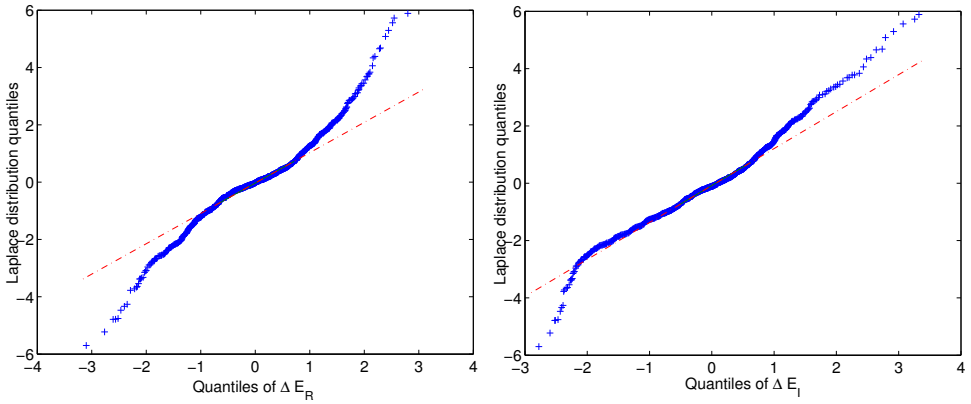
(a) (b)

Figure B.4: Quantile-quantile plot of the (a) real and (b) imaginary parts of the standardized measurement noise against a uniform distribution.



(a) (b)

Figure B.5: Quantile-quantile plot of the (a) real and (b) imaginary parts of the standardized measurement noise against a logistic distribution.

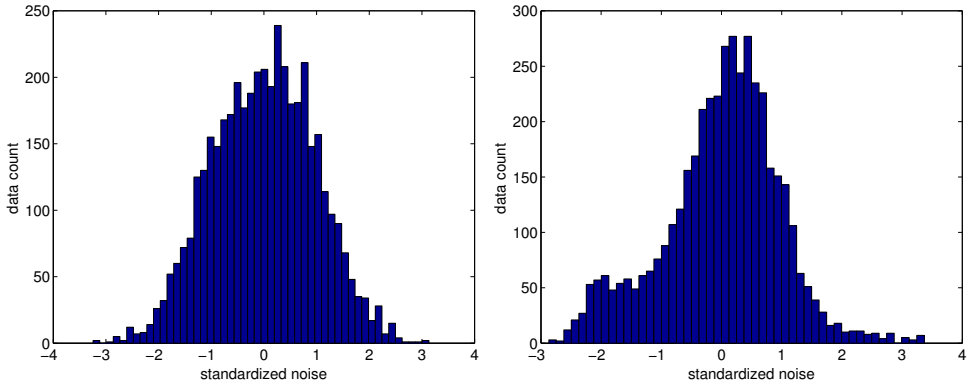


(a) (b)

Figure B.6: Quantile-quantile plot of the (a) real and (b) imaginary parts of the standardized measurement noise against a Laplace distribution.

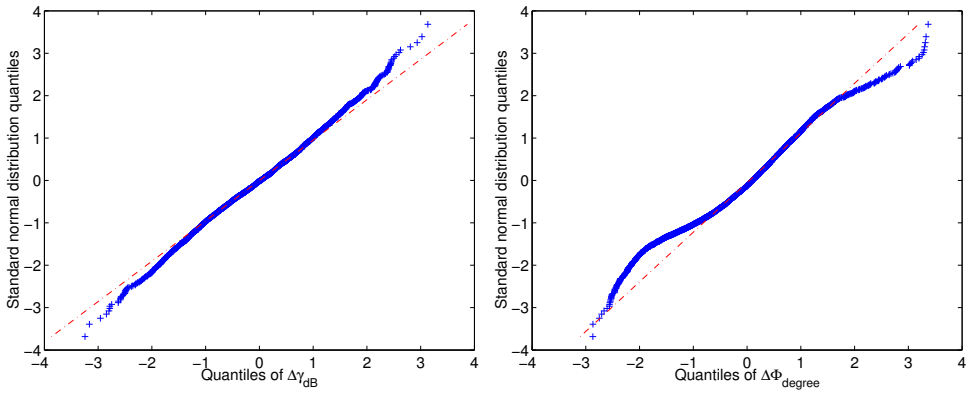
the real/imaginary data, and the histogram plot as well as the QQ-plot against a normal distribution are shown in Figures B.7 and B.8, respectively. From the QQ-plot, one may notice that the amplitude distribution is very close to that of the real part in Figure B.3, and the distribution of the phase data is similar to that of the imaginary part of the raw measurement. This is reasonable because when the size or contrast of the scatterer is small, the entries in the phase measurement $\mathbf{\Delta}\phi$ are small which results in $\cos(\mathbf{\Delta}\phi) \to \mathbf{1}$ and $\sin(\mathbf{\Delta}\phi) \to \mathbf{\Delta}\phi$. The entries in $\mathbf{\Delta}\gamma_{dB}$ are also close to zero. From B.2, the statistical properties of $\mathbf{\Delta}\gamma_{dB}$ are directly transferred to $\mathbf{\Delta E}_R$ and similarly from $\mathbf{\Delta}\phi$ to $\mathbf{\Delta E}_I$. This demonstrates that the dB amplitude/phase measurements can also be approximated as normal distributions when the scattered field is weak. Consequently, in the cases where the forward model is accurate, both the traditional complex version or LMPF version of Gauss-Newton reconstructions are both practically appropriate methods in estimating the dielectric properties for these cases.

One may also notices from Figure B.1 that even the distributions of the measurement can be viewed as a normal distribution. The variances of the measurement vary



(a) (b)

Figure B.7: Histogram plot of the (a) dB amplitude and (b) phase of the raw measurement.



(a) (b)

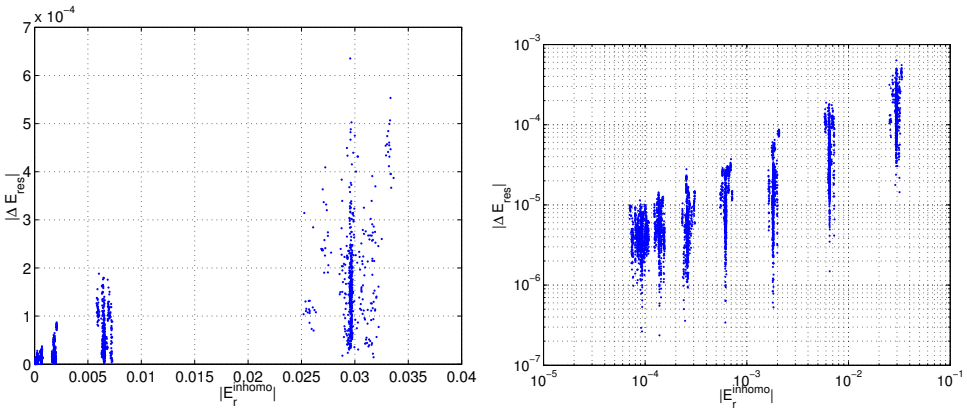
Figure B.8: Quantile-quantile plot of the (a) dB amplitude and (b) phase of the raw measurement noise against a normal distribution.

with respect to the relative positions of the receiver to the transmitter as well as with respect to the operating frequency (not shown). In this case, the weighted least-square (WLS) estimator is the more appropriate choice than the OLS estimator. A covariance matrix C of the raw measurement noise should be estimated at the selected reconstruction frequency prior to the reconstruction and a weighting matrix W needs to be constructed from equation (B.3) and used in the iterative least-square update equation (2.53).

B.2 Analysis on residual error

In this section, the statistical properties of the residual error in an iterative scheme are investigated. Intuitively, these properties are influenced by both the measurement noise and a complex image reconstruction process (including forward model accuracy, regularization, smoothing and so on). Utilizing the repetition measurement data set used in the previous section, we performed reconstructions with our traditional Gauss-Newton iterative algorithm. The configuration of the reconstruction was identical to the phantom experiments in Chapter 3. After 20 iterations, the residual error vector was recorded for each input data set. Consequently, 18 groups of residual error vectors were obtained for all measurement data sets.

A scatter plot of the residual error with respect to the amplitude of the input measurement data $\mathbf{E}_r^{\text{inhomo}}$ is shown in Figure B.9 (a). From this figure, it is quite obvious that the variation of the residual error amplitude becomes larger when $|\mathbf{E}_r^{\text{inhomo}}|$ increases. Transforming the scatter plot to a log scale in both axes (Figure B.9 b), the data points are approximately bounded by two parallel lines which indicates that the variance of the residual error in the log scale is approximately equal across various



(a) (b)

Figure B.9: Scatter plots between the amplitude of the residual error and the amplitude of the measurement data in (a) linear-linear scale, and (b) log-log scale.

measurement amplitude scales. In the variance analysis in parameter estimation, this is referred to as "homoscedasticity". For the results presented in Figure B.9 (a), the error is heteroscedasticity. It has been shown that heteroscedasticity in the data will result in low efficiency in the parameter estimation [139, 7]. This indicates that the logarithm transform used in the LMPF algorithm may substantially enhance the efficiency of the estimator and partially explains the improved image quality generally observed using this algorithm.

Appendix C

3D FDTD modelling of the illumination tank

In all of our forward field computations, we solved an unbounded radiation problem outside the imaging zone where we assume the space is filled by the background medium. However, in actual experimental settings, the lossy background medium only extends to the boundary of our illumination tank and the exterior of the tank is filled by air. In order to test the validity of our approximation, we compute the forward field with and without the presence of the tank. The tank has dimension $40 \times 40 \times 40$ cm^3 centered at the origin. A circular antenna array with radius 7.62 cm is placed at $z = 0$ plane and is also centered at the origin. The tank is filled with 83% glycerin solution ($\epsilon_r = 25, \sigma = 1.0$ S/m) while the rest of the space is filled by air ($\epsilon_r = 1, \sigma = 0$ S/m). The forward fields computed at 900 MHz for the two cases are shown in Figure C.1 and C.2 where the perimeter of the antenna array and the tank are marked by the circle and the square, respectively. From these two plots, the field distributions within the antenna array (the circular area) are very close to each other which validates the effectiveness of our exterior radiation assumptions. Consequently, the modelling space

can be compressed close to the antenna array providing a considerable computational savings compared with modelling the whole tank.

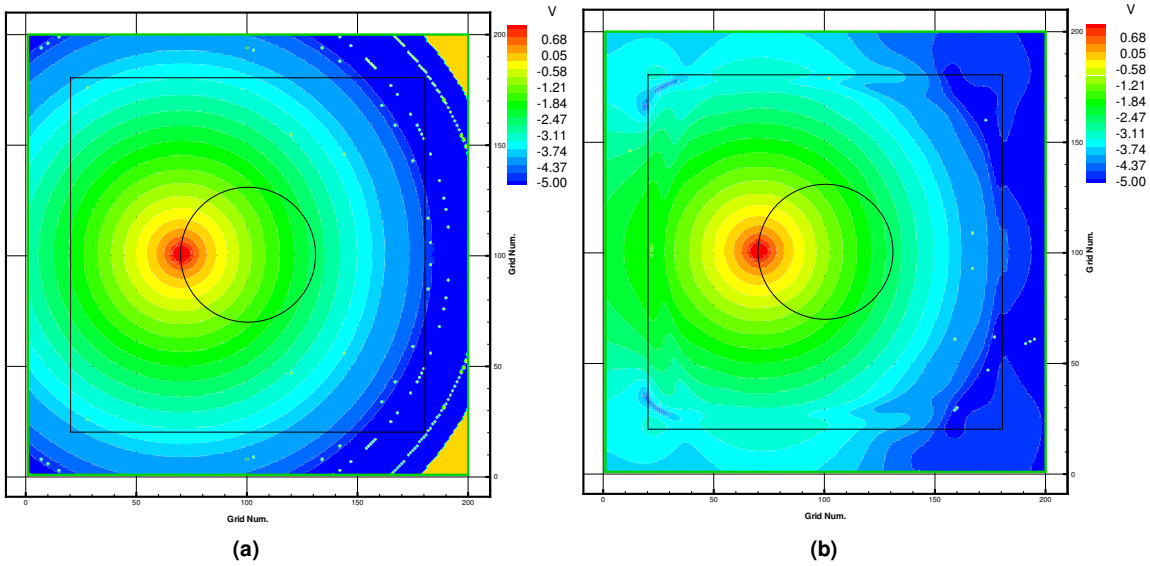


Figure C.1: The computed field amplitude ($\log_{10}(|E_z|)$) along $z = 0$ plane (a) without and (b) with the presence of the tank.

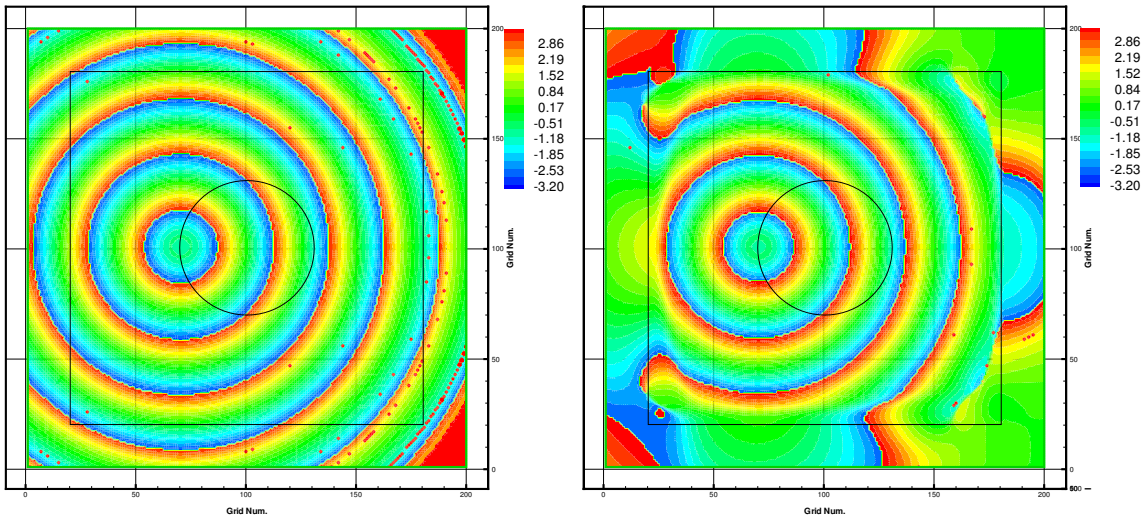


Figure C.2: The computed field phases (radian) along $z = 0$ plane (a) without and (b) with the presence of the tank.

Appendix D

Iso-sensitivity ovals and surfaces

A single row of the Jacobian matrix represents the sensitivity map across the imaging domain for a given transmit/receive antenna pair with respect to perturbations of the parameters at different locations. In our discussion in Chapter 7, this sensitivity map can be expressed by the multiplication of two Green's functions of the Helmholtz equation, i.e.

$$J((\vec{r}_s, \vec{r}_r), \vec{r}) = g(\vec{r}, \vec{r}_s)g(\vec{r}, \vec{r}_r) \tag{D.1}$$

and the iso-sensitivity curve (in 2D) or surface (in 3D) is defined by

$$g(\vec{r}, \vec{r}_s)g(\vec{r}, \vec{r}_r) = c \tag{D.2}$$

where c is a constant. For different c values, the curves defined by (D.2) comprise a contour plot which illustrates the measurement sensitivity over space. We present two examples here. The first example is the homogeneous medium in the 2D space. In this case, the Green's function is written as

$$g(\vec{r}, \vec{r}_s) = \frac{j}{4} H_0^{(1)}(k_0|\vec{r} - \vec{r}_s|) \tag{D.3}$$

where k is the complex wave number, and the iso-sensitivity curve is defined by

$$H_0^{(1)}(k_0|\vec{r}-\vec{r}_s|)H_0^{(1)}(k_0|\vec{r}-\vec{r}_r|) = c \tag{D.4}$$

Let $\vec{r}_s = (7,0)$ cm and $\vec{r}_r = (-7,0)$ cm, let $\epsilon_r = 77, \sigma = 1.7$ S/m be the dielectric properties of the background, the sensitivity contour plot is shown in Figure D.1.

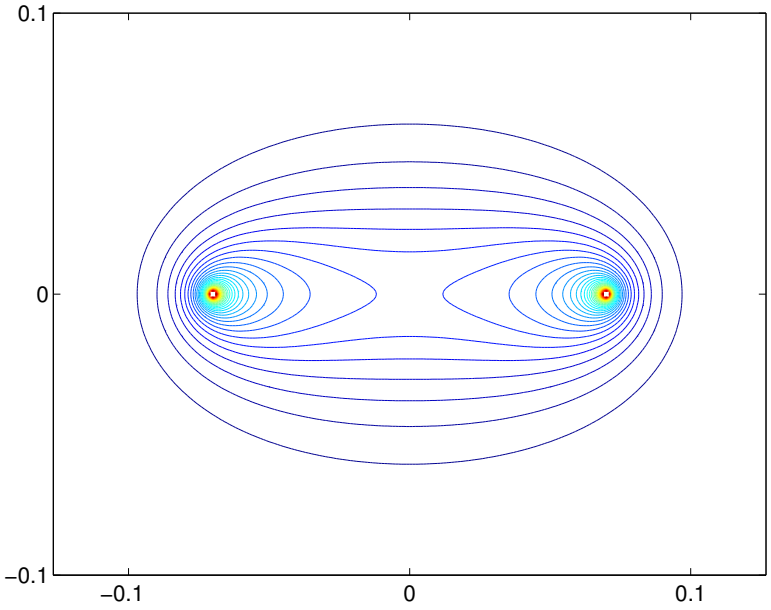


Figure D.1: Iso-sensitivity curves for infinitely large 2D homogeneous background medium.

Similarly, for infinitely large homogeneous medium in 3D space, the corresponding Green's function is written as

$$g(\vec{r},\vec{r}_s) = \frac{\exp jk|\vec{r}-\vec{r}_s|}{4\pi|\vec{r}-\vec{r}_s|} \tag{D.5}$$

We computed the iso-sensitivity surfaces for the case where $\vec{r}_s = (7,0,0)$ cm and $\vec{r}_r = (-7,0,0)$ cm with saline background, which are shown in Figure D.2.

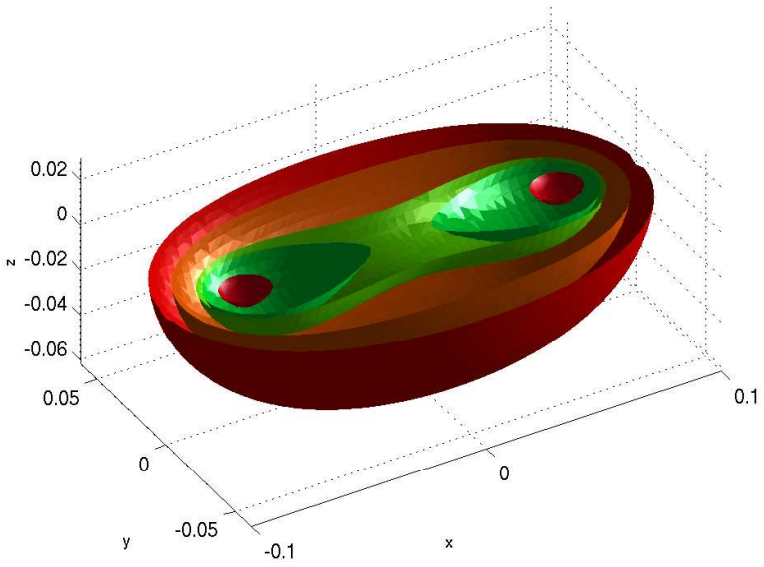


Figure D.2: Iso-sensitivity surfaces for infinitely large 3D homogeneous background medium (cut from $z = 0$ plane).

Appendix E

Proof of the nodal adjoint matrix reconditioning

For linear Lagrange elements, the integration of the multiplications of the basis functions can be analytically computed by

$$< \phi_1^m \phi_2^n \phi_3^l >= 2A \frac{m!n!l!}{(m+n+l+2)!} \tag{E.1}$$

for 2D elements and

$$< \phi_1^m \phi_2^n \phi_3^l \phi_4^k >= 6V \frac{m!n!l!k!}{(m+n+l+k+3)!} \tag{E.2}$$

for 3D elements where A is the area of a 2D triangle and V is the volume of a 3D tetrahedron. From (E.1), we can derive

$$\begin{aligned}\sum_{i=1}^{M}\sum_{j=1}^{M} <\phi_i\phi_j\phi_1> &= <\phi_1\phi_1\phi_1> + <\phi_1\phi_2\phi_1> + <\phi_1\phi_3\phi_1> \\ &\quad + <\phi_2\phi_1\phi_1> + <\phi_2\phi_2\phi_1> + <\phi_2\phi_3\phi_1> \\ &\quad + <\phi_3\phi_1\phi_1> + <\phi_3\phi_2\phi_1> + <\phi_3\phi_3\phi_1> \\ &= \frac{A}{10}+\frac{A}{30}+\frac{A}{30}+\frac{A}{30}+\frac{A}{30}+\frac{A}{60}+\frac{A}{30}+\frac{A}{60}+\frac{A}{30} \\ &= \frac{A}{M}\end{aligned} \tag{E.3}$$

where $M = 3$ is the node number per element. Similarly, we can prove $\sum_{i=1}^{M}\sum_{j=1}^{M} <\phi_i\phi_j\phi_2> = \sum_{i=1}^{M}\sum_{j=1}^{M} <\phi_i\phi_j\phi_3> = \frac{A}{M}$. For 3D elements, similar conclusions can be made, i.e. $\sum_{i=1}^{M}\sum_{j=1}^{M} <\phi_i\phi_j\phi_k> = \frac{V}{M}$ for $k = 1, \cdots, M$ where $M = 4$. With these conclusions, the reconditioning of the D matrices in Section 5.1.2 can be easily processed.

Appendix F

Common methods in computational electromagnetics

The following diagram summarizes the common computational methods for electromagnetic field calculations.

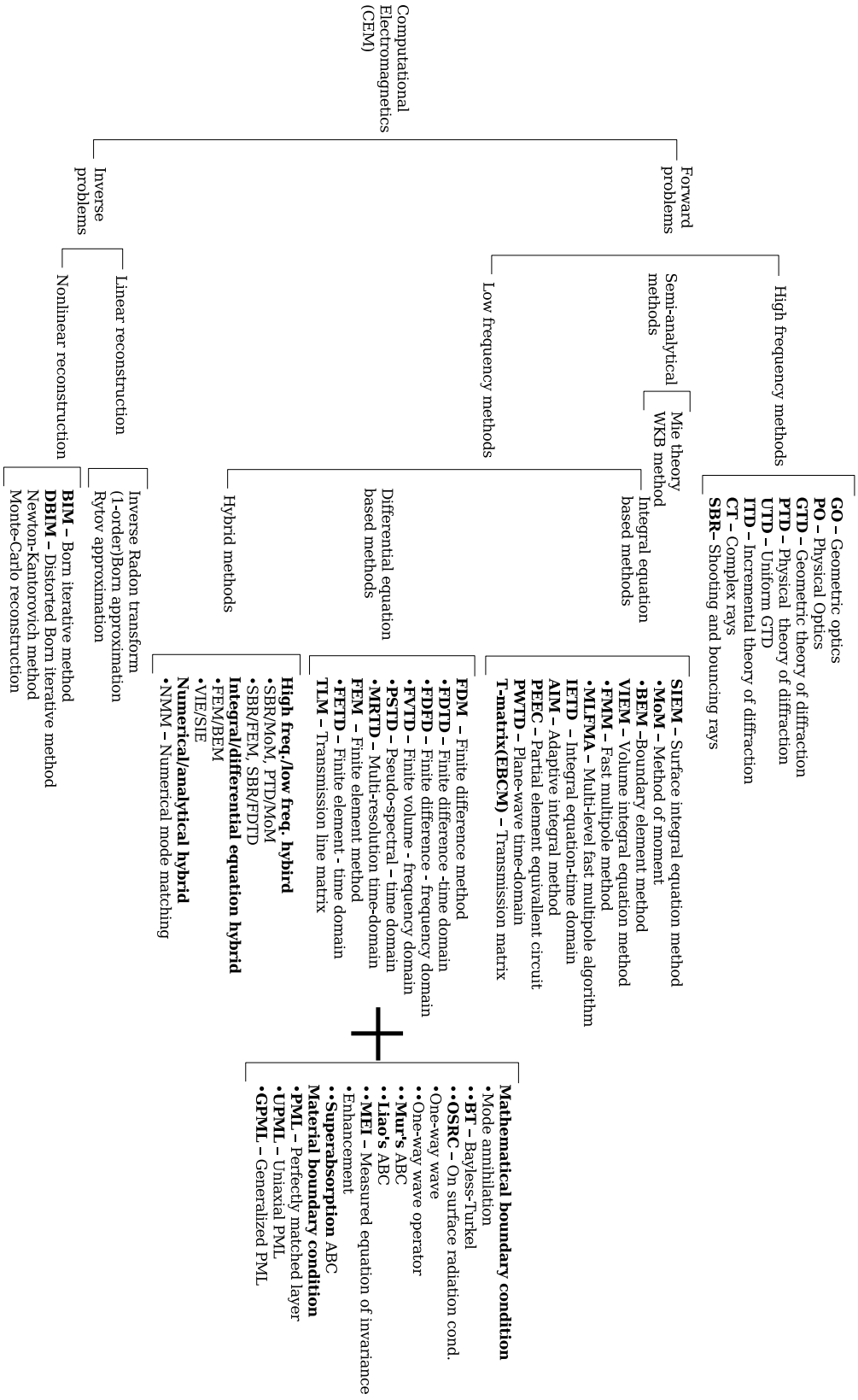


Figure F.1: Computational methods for EM modelling

Bibliography

[1] M. F. Adams and A.P. Anderson. Synthetic aperture tomographic (sat) imaging for microwave diagnostics. *IEE Proc.*, 129:83–88, 1982.

[2] R. Azaro, S. Caorsi, and M. Pastorino. A 3-GHz microwave imaging system based on a modulated scattering technique and on a modified Born approximation. *International Journal of Imaging Systems and Technology*, 9(5):395–403, Dec 1998.

[3] Y. Bard. *Nonlinear Parameter Estimation*. Academic Press, New York, 1974.

[4] F. Bardati, V. J. Brown, M. P. Ross, and P. Tognolatti. Microwave radiometry for medical thermal imaging: theory and experiment. *Microwave Symposium Digest,IEEE MTT-S International*, 3:1287–1290, Jun 1992.

[5] A. Bayliss, M. Gunzburger, and E. Turkel. Boundary Conditions for the Numerical Solution of Elliptic Equations in Exterior Regions. *SIAM J. Appl. Math.*, 42(2):430–451, 1982.

[6] A. Bayliss and E. Turkel. Radiation Boundary Conditions for Wave-Like Equations. *Comm. Pure Appl. Math.*, 33:707–725, 1980.

[7] J. V. Beck and K. J. Arnold. *Parameter Estimation in Engineering and Science*. John Wiley & Sons, 1977.

[8] J.-P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200, 1994.

[9] J.-P. Berenger. Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 127:363–379, 1996.

[10] E. Bermani, S. Caorsi, and M. Raffetto. A microwave object recognition approach based on neural networks. *Instrumentation and Measurement Technology Conference, 1999. IMTC/99. Proceedings of the 16th IEEE*, 3(24-26):1582–1585, May 1999.

[11] M. V. Berry and M. R. Dennis. Knotted and linked phase singularities in monochromatic waves. *Proc. R. Soc. A*, 457 2251C63, 2001.

[12] Å. Bjorck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.

[13] J. C. Bolomey, J. Izadnegahdar, L. Jofre, C. H. Pichot, G. Peronnet, and M. Solaimani. Microwave diffraction tomography for biomedical applications. *IEEE Trans. Microwave Theory Tech.*, 30:1990–2000, 1982.

[14] J. C. Bolomey, L. Jofre, G. Peronnet, A. Izadnegahdar, N. Berkane, M. Solaimani, G. Collignon, Y. Michel, J. Saint, C. Szeles, M. Gautherie, and J. L. Guerquin-Kern. A quasi real time microwave inverse scattering technique applicable to biomedical tomography. *Proc. URSI Symp. Nat. Radio Sci. Meet.*, page 93, 1982.

[15] E. J. Bond, X. Li, S. C. Hagness, and B. D. van Veen. Microwave imaging via space-time beamforming for early detection of breast cancer. *IEEE Trans. on Antennas Propagat.*, 51:1690–1705, 2003.

[16] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian Geometry, 2nd edition*. Academic Press, 1986.

[17] M. Born and E. Wolf. *Principles of Optics: Electromagnetic Theory of Propagation Interference, and Diffraction of Light,6th ed.* New York: Pergamon Press, 1989.

[18] W. E. Boyse and A. A. Seidl. A block QMRmethod for computing multiple simultaneous solutions to complex symmetric systems. *SIAM J. Sci. Comput.*, 17:263–274, 1996.

[19] O. Brander and B. DeFacio. The role of filters and the singular-value decomposition for the inverse Born approximation. *Inverse Problem*, 2:375–393, 1986.

[20] A. Broquetas, J. Romeu, J. M. Rius, A. R. Elias-Fuste, A. Cardama, and L. Jofre. Cylindrical geometry: a further step in active microwave tomography. *IEEE Transactions on Microwave Theory and Techniques*, 39(5), May 1991.

[21] J. T. Bushberg, J. Anthony Seibert, E. M. Leidholdt Jr., and J. M. Boone. *The Essential Physics of Medical Imaging (2nd Edition)*. Lippincott Williams and Wilkins, 2002.

[22] A. M. Campbell and D. V. Land. Dielectric properties of female human breast tissue measured in vitro at 3.2 GHz. *Phys Med Biol.*, 37(1):193–210, Jan 1992.

[23] S. Caorsi, S. Ciaramella, G. L. Gragnani, and M. Pastorino. On the use of regularization techniques in numerical inverse scattering solutions for microwave imaging applications. *IEEE Trans. Microwave Theory Tech.*, 43:632–640, 1995.

[24] S. Caorsi, G. L. Gragnani, S. Medicina, M. Pastorino, and G. Zunino. Microwave imaging method using a simulated annealing approach. *IEEE Microwave and*

Guided Wave Letters [see also IEEE Microwave and Wireless Components Letters], 1(11):331–333, Nov 1991.

[25] S. Caorsi and M. Pastorino. Two-dimensional microwave imaging approach based on a genetic algorithm. *IEEE Transactions on Antennas and Propagation*, 48(3):370–373, Mar 2000.

[26] A. Caponnetto and M. Berteron. Tomography with a finite set of projections: Singular value decomposition and resolution. *Inverse Problems*, 13:1191–1205, 1997.

[27] J. T. Chang, K. F. Paulsen, P. M. Meaney, and M. Fanning. Noninvasive thermal assessment of tissue phantoms using an active near field microwave imaging technique. *Int J Hyperthermia*, 14, 1998.

[28] G. Chartrand. *Introductory Graph Theory*. New York: Dover, 1985.

[29] S. S. Chaudhary, R. K. Mishra, A. Swarup A, and J. M. Thomas. Dielectric properties of normal & malignant human breast tissues at radiowave & microwave frequencies. *Indian J Biochem Biophys*, 21:76–79, 1984.

[30] C. Chavent and P. Lemonnier. Identification de la Non-Linéarité D'Une Équation Parabolique Quasilinéare. *Applied Mathematics and Optimization*, 1:121–162, 1974.

[31] W. C. Chew. *Waves and Fields in Inhomogeneous Media*. New York: Van Nostrand-Reinhold, 1990.

[32] W. C. Chew. *Computational electrodynamics: the finite-difference time-domain method*. Artech House, Norwood, London, 1998.

[33] W. C. Chew and J. H. Lin. A frequency-hopping approach for microwave imaging of large inhomogeneous bodies. *IEEE Microwave and Guided Wave Letters*, 5:439–441, 1995.

[34] W. C. Chew and Y. M. Wang. Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method. *IEEE Transactios on Medical Imaging*, 9(2):218–225, Jun 1990.

[35] W. C. Chew and W. H. Weedon. A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates. *Microwave and Optical Technology Letters*, 7(13):599–604, 1994.

[36] K. S. Cole and R. H. Cole. Dispersion and absorption in dielectrics: I. Alternating current characteristics. *J. of Chem. Phys.*, pages 341–351, 1941.

[37] D. Colton and P. Monk. A new approach to detecting leukemia: using computational electromagnetics. *IEEE Trans. on Comp. Sci. and Eng.*, 2:46–52, 1995.

[38] R. Cusack and N. Papadakis. New Robust 3-D Phase Unwrapping, Algorithms: Application to Magnetic Field Mapping and Undistorting Echoplanar Images. *NeuroImage*, 16:754–764, Jul 2002.

[39] R. Kress D. Colton. *Inverse Acoustic and Electromagnetic Scattering Theory, 2nd ed.* Springer-Verlag,New York, 1998.

[40] M. E. Davison. A singular value decomposition for the Radon transform in n-dimensional Euclidean space. *Num. Funct. Anal. Optimiz.*, 51(3):321–340, 1981.

[41] Eugene Demidenko. personal communication with the author. 2004.

[42] M. R. Dennis. Local phase structure of wave dislocation lines: twist and twirl. *Journal of Optics A: Pure and Applied Optics, special issue on singular optics*, Apr 2004.

[43] F. A. Duck. *Physical Properties of Tissue: A Comprehensive Reference Book*.

[44] C. H. Durney, H. Massoudi, and M. F. Iskander. *Radiofrequency radiation dosimetry handbook*. USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC), Brooks Air Force Base, TX, 1986.

[45] C. Henry Edwards and D. E. Penney. *Multivariable calculus with analytic geometry*. Prentice-Hall, Inc, 1998.

[46] H. Ermert, G. Fulle, and D. Hiller. Microwave computerized tomography. *Proc. XIth European Microwave Conf.*, 1981.

[47] J. Fang. *Time-Domain Finite Difference Computations for Maxwell's Equations*. Ph.D. dissertation, EECS Dept. University of California, Berkeley, CA, 1989.

[48] J. Fang and Z. Wu. Generalized perfectly matched layer—An extension of Berenger's perfectly matched layer boundary condition. *IEEE Microwave Guided Wave Letters*, 5(12):451–453, Dec 1995.

[49] J. Fang and Z. Wu. Generalized perfectly matched layer for the absorption of propagating and evanescent waves in lossless and lossy media. *IEEE Trans. Microwave Theory Tech.*, 44:2216–2222, 1996.

[50] Q. Fang, P. M. Meaney, S. D. Geimer, and and K. D. Paulsen A. V. Streltsov. Microwave image reconstruction from 3D fields coupled to 2D parameter estimation. *IEEE Trans. Med. Imag.*, 23:475–484, Apr 2004.

[51] Q. Fang, P. M. Meaney, and K. D. Paulsen. Microwave Image Reconstruction of Tissue Property Dispersion Characteristics Utilizing Multiple Frequency Information. *IEEE Trans. Microwave Techniques and Technology*, 52:1866–1875, Aug 2004.

[52] E. C. Fear, X. Li, S. C. Hagness, and M. A. Stuchly. Confocal microwave imaging for breast cancer detection: localization of tumors in three dimensions. *IEEE Trans. Biomed. Eng.*, 49:812–822, 2002.

[53] E. C. Fear, J. Sill, and M. A. Stuchly. Experimental feasibility study of confocal microwave imaging for breast tumor detection. *IEEE Trans. Microwave Theory Tech.*, 51:887–892, 2003.

[54] K. R. Foster and J. L. Schepp. Dielectric properties of tumor and normal tissues at radio through micro-wave frequencies. *Journal of Microwave Power*, 16:107–119, 1981.

[55] K. R. Foster and H. P. Schwan. Dielectric properties of tissues and biological materials: A critical review. *Critical Reviews in Biomedical Engineering*, 17(1):25–104, 1989.

[56] A. Franchois and C. Pichot. Microwave imaging-complex permittivity reconstruction with a Levenberg-Marquardt method. *IEEE Trans. Antennas and Propagat*, 45(2):203–215, Feb 1997.

[57] R. W. Freund. Conjugate gradient type methods for linear systems with complex symmetric coefficient matrices. *SIAM J. Sci. Statist. Comput.*, 13:425–448, 1992.

[58] D. L. Fried and J. L. Vaughn. Branch cuts in the phase function. *Applied Optics*, 31(15):2866–2882, 1992.

[59] C. Gabriel, S. Gabriel, and E. Corthout. The dielectric properties of biological tissues: I. Literature survey. *Phys. Med. Biol.*, 41:2231–2249, 1996.

[60] S. Gabriel, R. W. Lau, and C. Gabriel. The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. *Phys. Med. Biol.*, 41:2271–2293, 1996.

[61] S. D. Gedney. An anisotropic PML absorbing media for the FDTD simulation of fields in lossy and dispersive media. *Electromagnetics*, 16(4):399–415, 1996.

[62] S. D. Gedney. *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Norwood, London, 1998.

[63] D. C. Ghiglia and M. D. Pritt. *Two-dimensional phase unwrapping : theory, algorithms, and software*. New York: J. Wiley, 1998.

[64] P. F. Goldsmith, C.-T. Hsieh, G. R. Huguenin, J. Kapitzky, and E.L. Moore. Focal plane imaging systems for millimeter wavelengths. *IEEE Transactions on Microwave Theory and Techniques*, 41(10):1664–1675, Oct 1993.

[65] R. M. Goldstein, H. A. Zebker, and C. L. Werner. Satellite Radar Interferometry: Two-Dimensional Phase Unwrapping. *Radio Science*, 23:713–720, 1988.

[66] G. H. Golub and C. H. van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17(4):883–893, 1980.

[67] G. H. Golub and C. H. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, 1991.

[68] P. Gould. The rise and rise of medical imaging. *Physics World*, Aug 2003.

[69] L. Greengard and V. Rokhlin. A fast algorithm for particle simulation. *Journal of Computational Physics*, 73:325–348, 1987.

[70] L. Gürel and U. Oğuz. Signal-processing techniques to reduce the sinusoidal steady-state error in the FDTD method. *IEEE Antennas and Propagat. Soc. Int. Symp., Montreal, Canada*, 48(4):585–593, Apr 2000.

[71] J. Hadamard. *Lectures on Cauchy's Problems in Linear Partial Differential Equations*. New Haven, CT: Yale University Press, 1923.

[72] O. S. Haddadin and E. S. Ebbini. Imaging strongly scattering media using a multiple frequency distorted Born iterative method. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, 45:1485–1496, 1998.

[73] S. C. Hagness, A. Taflove, and J. E. Bridges. Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Fixed-focus and antenna-array sensors. *IEEE Trans. Biomedical Engineering*, 45:1470–1479, Dec 1998.

[74] S. C. Hagness, A. Taflove, and J. E. Bridges. Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Design of an antenna-array element. *IEEE Trans. on Antenna and Propagation*, 47(5):783–779, May 1999.

[75] M. Hanke and T. Raus. A general heuristic for choosing the regularization parameter in ill-posed problems. *SIAM Journal on Scientific Computing*, 17(4):956–972, 1996.

[76] P. C. Hansen. *Regularization Tools - A matlab package for analysis and solutoin of discrete ill-posed problems*. version 3.1 for matlab 6.0, 1992.

[77] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, 1997.

[78] R. C. Hansen, editor. *Geometric Theory of Diffraction*. IEEE Press, New York, 1981.

[79] G. W. Hanson and A. B. Yakovlev. *Operator Theory for Electromagnetics, An Introduction*. Springer Verlag, New York, 2002.

[80] R. F. Harrington. *Time-Harmonic Electromagnetic Fields*. New York: McGrawHill Book Company, 1961.

[81] R. F. Harrington. *Field Computation by Moment Methods*. IEEE Press, 1968.

[82] J. R. Harris, M. E. Lippman, U. Veronesi, and W. Willett. Medical progress: Breast cancer. *New Eng. J. Med.*, 327(5):319–328, 1992.

[83] M. Henle. *A Combinatorial Introduction to Topology*. Dover Publication Inc., New York, 1994.

[84] J. H. Holland. *Adaptation in natural and artificial system*. The University of Michigan Press, Ann Arbor, 1975.

[85] R. P. Hsia, S. Cheng, W. R. Geck, and C. W. Domier N. C. Luhmann Jr. Millimeter-wave Schottky diode imaging array development. *Microwave Conference Proceedings, 1993. APMC '93., 1993 Asia-Pacific*, 1:1–9–1–12, 1993.

[86] J. P. Hugonin, N. Joachimowicz, and C. Pichot. *Inverse Methods in Action*. Springer-Verlag, Berlin, 1990.

[87] A. Ishimaru. *Electromagnetic wave propagation radiation and scattering*. Prentice Hall, New Jersey, 1991.

[88] Institute for Applied Physics Italian National Research Council. *Dielectric Properties of Body Tissues: Home page*. http://niremf.ifac.cnr.it/tissprop/, 1997-2002.

[89] J.H. Jacobi, L. E. Larsen, and C. T. Hast. Water-Immersed Microwave Antennas and Their Application to Microwave Interrogation of Biological Targets. *IEEE Transactions on Microwave Theory and Techniques*, 27(1), Jan 1979.

[90] S. Jacobsen and P. R. Stauffer. Multifrequency radiometric determination of temperature profiles in a lossy homogeneous phantom using a dual-mode antenna with integral water bolus. *IEEE Trans. Microwave Theory Tech.*, 50:1737–1746, 2002.

[91] S. L. S. Jacoby, J. S. Kowalik, and J. T. Pizzo. *Iterative methods for nonlinear optimization problems*. Prentice-hall, Inc, Englewood Cliffs, NJ, 1972.

[92] J. Jin. *The Finite Element Method in Electromagnetics, 2nd Edition*. Wiley-IEEE Press, 2002.

[93] N. Joachimowicz, C. Pichot, and J. P. Hugonin. Inverse scattering: an iterative numerical method for electromagnetic imaging. *IEEE Transactions on Antennas and Propagation*, 39:1742–1752, 1991.

[94] W. T. Joines, Y. Zhang, C. Li, and R. L. Jirtle. The measured electrical properties of normal and malignant human tissues from 50 to 900 MHz. *Med. Phys.*, 21:547–550, 1994.

[95] A. C. Kak and M. Slaney, editors. *Principles of Computerized Tomographic Imaging*. Society of Industrial and Applied Mathematics, 2001.

[96] B. Kaltenbacher. Newton-type methods for ill-posed problems. *Inverse Problems*, 13:729–753, 1997.

[97] L. Kauffman. *Knots and Physics*. Teaneck, NJ: World Scientific, 1991.

[98] J. B. Keller. Geometric Theory of Diffraction. *J. Opt. Soc. Am.*, 52:116–130, 1962.

[99] Yu. S. Kivshar, J. Christou, V. Tikhonenko, B. Luther-Davies, and L. M. Pismen. Dynamics of optical vortex solitons. *Opt. Commun.*, 152:198–206, 1998.

[100] J. A. Kong. *Electromagnetic Wave Theory*. Wiley-Interscience, New York, 1990.

[101] P. Kosmas, C. M. Rappaport, and E. Bishop. Modeling with the FDTD method for microwave breast cancer detection. *IEEE Transactions on Microwave Theory and Techniques*, 52(8):1890– 1897, 2004.

[102] K. S. Kunz and R. J. Luebbers. *The Finite Difference Time Domain Method for Electromagnetics*. CRC Press, Boca Raton, FL, 1993.

[103] L. E. Larsen and J. H. Jacobi. Microwave scattering parameter imagery of an isolated canine kidney. *Med Phys*, 6:394–403, 1979.

[104] M. M. Lavrent'ev, V. G. Romanov, and S. P. ShishatSkiĭ. *Ill-posed problems of mathematical physics and analysis*. AMS, 1986.

[105] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. Englewood Cliffs, NJ: PrenticeHall, 1974.

[106] G. Lazzi. Unconditionally stable D-H FDTD formulation with anisotropic PML boundary conditions. *Microwave and Wireless Components Letters, IEEE [see also IEEE Microwave and Guided Wave Letters]*, 11(4):149–151, Apr 2001.

[107] J. M. Lee. *Introduction to Topological Manifolds*. Springer-Verlag New York, 2000.

[108] K. Levenberg. A Method for the Solution of Certain Problems in Least Squares. *Quart. Appl. Math.*, 2:164–168, 1944.

[109] B. Lewin. *Genes VIII*. Pearson Prentice Hall, Upper Saddle River, NJ, 2004.

[110] D. Li, P. M. Meaney, and K. D. Paulsen. Conformal microwave imaging for breast cancer detection. *IEEE Transactions on Microwave Theory and Techniques*, 51(4):1179–1186, Apr 2003.

[111] D. Li, P. M. Meaney, T. Raynolds, S. A. Pendergrass, M. W. Fanning, and K. D. Paulsen. A parallel-detection microwave spectroscopy system for breast imaging. *Review of Scientific Instruments*, 75(7):2305–2313, 2004.

[112] C.-T. Liauh, R. G. Hills, and R. B. Roemer. Comparison of the Adjoint and Influence Coefficient Methods for Solving the Inverse Hyoperthermia Problem. *J. of Biomechanical Eng.*, 115:63–71, 1993.

[113] H. Ling, R.-C. Chou, and S.-W. Lee. Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity. *IEEE Transactions on Antennas and Propagation*, 37(2):194–205, Feb 1989.

[114] G. Liu and S. D. Gedney. Perfectly matched layer media for an unconditionally stable three-dimensional ADI-FDTD method. *Microwave and Guided Wave Letters, IEEE [see also IEEE Microwave and Wireless Components Letters]*, 10(7):61–263, Jul 2000.

[115] Q. H. Liu, Z. Q. Zhang, T. T. Wang, J. A. Bryan, G. A. Ybarra, L. W. Nolte, and W. T. Joines. Active microwave imaging. I. 2-D forward and inverse scattering methods. *IEEE Trans. Microwave Theory Tech.*, 50:123–133, 2002.

[116] Michael A. Lombardi, Lisa M. Nelson, Andrew N. Novick, and Victor S. Zhang. Time and frequency measurements using the global positioning system. *Technical report, National instutute of standards and technology, time and frequency division*, Aug 2001.

[117] C. C. Lu and W. C. Chew. A multilevel algorithm for solving boundary integral equations of wave scattering. *Microwave Opt. Technol. Lett.*, 7(10):466–470, Jul 1994.

[118] B. LutherDavies, J. Christou, and Yu. S. Kivshar V. Tikhonenko. Optical vortex solitons: experiment versus theory. *J. Opt. Soc. Am. B*, 14:3045–3053, 1997.

[119] R. Maini, M. F. Iskander, and C. H. Durney. On electromagnetic imaging using linear reconstruction techniques. *Proceedings of the IEEE*, 68:1550–1552, Dec 1980.

[120] J.J. Mallorqui. Active microwave tomography. *IEE Colloquium on Functional Imaging*, 5:6/1 – 6/3, Jan 1994.

[121] D. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM J. Appl. Math*, 11:431–441, 1963.

[122] P. M. Meaney, M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen. A clinical prototype for active microwave imaging of the breast. *IEEE Trans. Microwave Theory Tech.*, 48:1841–1853, 2000.

[123] P. M. Meaney, M. W. Fanning, K. D. Paulsen, D. Li, S. A. Pendergrass, Q. Fang, and K. L. Moodie. Microwave thermal imaging: Initial *in vivo* experience with a single heating zone. *International Journal of Hyperthermia*, 19(6):617–641, Nov 2003.

[124] P. M. Meaney, K. D. Paulsen, and J. T. Chang. Near-field microwave imaging of biologically-based materials using a monopole transceiver system. *IEEE Trans. on Microwave Theory and Tech.*, 46(1):31–45, 1998.

[125] P. M. Meaney, K. D. Paulsen, J. T. Chang, and M. Fanning. Compensation for nonactive array element effects in a microwave imaging system: part II – imaging results. *IEEE Trans. on Med. Imag.*, 18:508–518, 1999.

[126] P. M. Meaney, K. D. Paulsen, S. Geimer, S. Haider, and M. W. Fanning. Quantification of 3D field effects during 2D microwave imaging. *IEEE Trans. on Biomed. Eng.*, 49:708–720, 2002.

[127] P. M. Meaney, K. D. Paulsen, and T. P. Ryan. Two-dimensional hybrid element image reconstruction for TM illumination. *IEEE Trans. Antennas Propagat.*, 43:239–247, 1995.

[128] P. M. Meaney, S. A. Pendergrass, M. W. Fanning, D. Li, and K. D. Paulsen. Importance of using a reduced contrast coupling medium in 2D microwave breast imaging. *Journal of Electromagnetic Waves and Applications*, 17:333–355, 2003.

[129] P. M. Meaney, N. K. Yagnamurthy, and K. D. Paulsen. Pre-scaling of reconstruction parameter components to reduce imbalance in image recovery process. *Physics in Medicine and Biology*, 47:1101–1119, 2002.

[130] R. Mittra and M. Kuzuoglu. A review of some recent advances in perfectly-matched absorbers for mesh truncation in FEM. *Antennas and Propagation Society International Symposium, 1997. IEEE., 1997 Digest*, 2:1302 – 1305, Jul 1997.

[131] M. Miyakawa. Tomographic measurement of temperature change in phantoms of the human body by chirp radar-type microwave computed tomography. *Med. Biol. Eng. Comput.*, 31:S31–S36, 1993.

[132] M. Miyakawa, K. Orikasa, M. Bertero, F. Conte P. Boccacci, and M. Piana. Experimental validation of a linear model for data reduction in chirp-pulse microwave CT. *IEEE Trans. Med. Imag.*, 21:385–395, 2002.

[133] M. Moghaddam and W. C. Chew. onlinear two-dimensional velocity profile inversion using time domain data. *IEEE Transactions on Geoscience and Remote Sensing*, 30(1):147 – 156, Jan 1992.

[134] M. Moghaddam, W. C. Chew, and M. Oristaglio. Comparison of the Born iterative method and Tarantola's method for an electromagnetic time-domain inverse problem. *Int. J. Imaging Syst. Tech.*, 3:318–333, 1991.

[135] D. C. Montgomery and G. C. Runger. *Applied Statistics and Probability for Engineers, 2nd edition*. John Wiley & Sons, New York, 1999.

[136] S. K. Moore. Better breast cancer detection. *IEEE Spectrum*, 38(5):50–54, 2001.

[137] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet Mathematics Doklady*, 7:414–417, 1966.

[138] V. A. Morozov. *Methods for solving incorrectly posed problems*. Springer Verlag, New York, 1984.

[139] H. G. Muller and U. Stadtmuller. Estimation of heteroscedasticity in regression analysis. *Ann. Statist.*, 15:610–625, 1987.

[140] T. Namiki. Unconditionally stable FDTD algorithm for solving three-dimensional Maxwell's equations. *Microwave Symposium Digest., 2000 IEEE MTT-S International*, 1:11–16, Jun 2000.

[141] P. A. Nelson and Y. Kahana. Spherical harmonics, singular value decomposition and the head related transfer function. *Special edition of the Journal of Sound and Vibration*, 239(4):607–637, 2001.

[142] P. A. Nelson and S. H. Yoon. Estimation of acoustic source strength by inverse methods: Part I, Conditioning of the inverse problem; Part II, experimental investigation of methods for choosing regularisation parameters. *Journal of Sound and Vibration*, 233(4):639–701, 2000.

[143] D. Neshev and Yu. S. Kivshar A. Nepomnyashchy. Nonlinear Aharonov-Bohm scattering by optical vortices. *Phys. Rev. Lett.*, 87, 043901-4, 2001.

[144] S. P. Novikov. *Topology I: General Survey*. Springer-Verlag Berlin Heidelberg, 1996.

[145] J. F. Nye. Unfolding of higher-order wave dislocations. *J. Opt. Soc. Am. A*, 15(5), May 1998.

[146] J. F. Nye and M. V. Berry. Dislocations in wave trains. *Proc. Roy. Soc. Lond. A*, 336:165–190, 1974.

[147] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *In Proceedings of the IEEE*, 69:529–541, May 1981.

[148] A. V. Oppenheim, A. S. Willsky, and N. S. Hamid. *Signals and Systems (2nd Edition)*. Prentice Hall Engineering, Science & Math, 1996.

[149] G. P. Otto and W. C. Chew. Microwave inverse scattering - local shape function imaging for improved resolution of strong scatterers. *IEEE Trans. Microwave Theory and Techniques*, 42(1):137–141, Jan 1994.

[150] U. Oğuz and L. Gürel. Interpolation techniques to improve the accuracy of the plane wave excitations in the finite-difference time-domain method. *Radio Sci.*, 32:2189–2199, 1997.

[151] K. D. Paulsen P. M. Meaney, B. W. Pogue, and M. I. Miga. Microwave image reconstruction utilizing log-magnitude and unwrapped phase to improve high-contrast object recovery. *IEEE Transactions on Medical Imaging*, 20:104–116, 2001.

[152] M. Pastorino, A. Massa, and S. Caorsi. A microwave inverse scattering technique for image reconstruction based on a genetic algorithm. *IEEE Transactions on Instrumentation and Measurement*, 49(3):573–578, Jun 2000.

[153] K. D. Paulsen, S. Geimer, J. Tang, and W. E. Boyse. Optimization of pelvic heating rate distributions with electromagnetic phased arrays. *Int. J. Hyperthermia*, 15:157–166, 1999.

[154] K. D. Paulsen, X. Jia, and Jr. J. M. Sullivan. Finite element computations of specific absorption rates in anatomically conforming full-body models for hyperthermia treatment analysis. *IEEE Trans. on Biomed. Eng.*, 40:933–945, 1993.

[155] K. D. Paulsen, P. M. Meaney, M. J. Moskowitz, and J. M. Sullivan Jr. A dual mesh scheme for finite element based reconstruction algorithms. *IEEE Trans. Med. Imag.*, 14:504–514, 1995.

[156] R. Pethig. Dielectric properties of biological materials: biophysical and medical appliications. *IEEE Trans. on Elec. Insulation*, 19:453–474, Oct 1984.

[157] C. Pichot, L. Jofre, Peronnet G., and J. Bolomey. Active microwave imaging of inhomogeneous bodies. *IEEE Transactions on Antennas and Propagation*, 33(4), Apr 1985.

[158] C. Prati, M. Giani, and N. Leuratti. SAR Interferometry: a 2-D Phase Unwrapping Technique Based on Phase and Absolute Value Informations. *IGARSS '90*, pages 2043–2046, 1990.

[159] A. W. Preece, J. L. Green, N. Pothecary, and R. H. Johnson. Microwave imaging for tumour detection. *IEE Colloquium on Radar and Microwave Imaging*, 16:9/1 – 9/4, Nov 1994.

[160] A. W. Preece, R. H. Johnson, J. L. Green, and M. P. Robinson. Dielectric imaging for localisation and detection of breast tumours. *Microwave Symposium Digest, 1993., IEEE MTT-S International*, 2:1145–1146, Jun 1993.

[161] H. R. Raemer. *Radar Systems Principles*. CRC Press Inc., FL, 1997.

[162] O. Ramahi and R. Mittra in M. A. Morgan. *Progress in Electromagnetics Research (PIER2)*. Elsevier Science Publishing Company Inc., 1990.

[163] P. S. Rao, K. Santosh, and E. C. Gregg. Computer tomography with microwaves. *Radiology*, 135, 1980.

[164] S. S. Rao. *Optimization Theory and Applications, 2nd edition*. John Wiley & Sons, New Delhi, 1984.

[165] C. M. Rappaport. Perfectly matched absorbing boundary conditions based on anisotropic lossy mapping of space. *Microwave and Guided Wave Letters, IEEE*, 5:90–92, 1995.

[166] J. R. Reitz and F. J. Milford. *Funcations of electromagnetic theory*. Addison Wesley Publishing Company, 1967.

[167] R. F. Remis and P. M. van den Berg. On the equivalence of the Newton-Kantorovich and distorted Born methods. *Inverse Problems*, 16(1):L1–L4, Feb 2000.

[168] R. Remmert. *Theory of Complex Functions*. Springer-Verlag, New York, 1991.

[169] A. Roger. A Newton-Kantorovich algorithm applied to an electromagnetic inverse problem. *IEEE Trans. Antennas Propag.*, 29:232–238, 1981.

[170] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *Journal of Computational Physics*, 60(2):187–207, 1985.

[171] R. W. Lau S. Gabriel and C. Gabriel. The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Phys. Med. Biol.*, 41:2251–2269, 1996.

[172] H. F. Schouten, T. D. Visser, G. Gbur, D. Lenstra, and H. Blok. Creation and annihilation of phase singularities near a sub-wavelength slit. *Opt. Express*, 11:371–380, 2003.

[173] H. P. Schwan. *Electrical properties measured with alternating currents; body tissues*. Handbook of Biological Data. Philadelphia: W B Saunders Co, 1956.

[174] S. Y. Semenov, A. E. Bulyshev, R. H. Svenson A. E. Souvorov, Y. E. Sizov, V. Y. Borisov, I. M. Kozlov V. G. Posukh, A. G. Nazarov, and G. P. Tatsis. Microwave tomography: Theoretical and experimental investigation of the iteration reconstruction algorithm. *IEEE Trans. Microwave Theory Tech.*, 46:133–141, 1998.

[175] S. Y. Semenov, R. H. Svenson, A. E. Boulyshev, A. E. Souvorov, V. Y. Borisov, Y. Sizov, A. N. Starostin, K. R. Dezern, G. P. Tatsis, and V. Y. Baranov. Microwave tomography: two-dimensional system for biological imaging. *IEEE Transactions on Biomedical Engineering*, 43(9):869–877, Sep 1996.

[176] S. Y. Semenov, R. H. Svenson, A. E. Bulyshev, A. E. Souvorov, A. G. Nazarov, Y. E. Sizov, A. V. Pavlovsky, V. Y. Borisov, B. A. Voinov, G. I. Simonova, A. N. Starostin, V. G. Posukh, G. P. Tatsis, and V. Y. Baranov. Three-dimensional microwave tomography: experimental prototype of the system and vector Born reconstruction method. *IEEE Transactions on Biomedical Engineering*, 46(8):937–946, Aug 1999.

[177] S. Y. Semenov, R. H. Svenson, A. E. Bulyshev, A. E. Souvorov, A. G. Nazarov, Y. E. Sizov, V. G. Posukh, and A. Pavlovsky. Three-dimensional microwave tomography: Initial experimental imaging of animals. *IEEE Trans. Biomed. Eng.*, 49:55–63, 2002.

[178] S. Y. Semenov, R. H. Svenson, and G. P. Tatsis. Microwave spectroscopy of myocardial ischemia and infarction: 1. experimental study. *Ann. Biomed. Eng.*, 28:48–54, 2000.

[179] S. Silver, editor. *Microwave Antenna Theory and Design*. McGraw-Hill, New York, 1949.

[180] American Cancer Society. Cancer Facts & Figures 2004. Technical report, 2004.

[181] American Cancer Society. Cancer Prevention & Early Detection Facts & Figures 2004. Technical report, 2004.

[182] A. Sommerfeld, editor. *Partial differential equations in physics*. Pure and applied mathematics. Academic Press Inc., New York, NY, 1949.

[183] J. M. Song and W. C. Chew. Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering. *Microwave Opt. Technol. Lett.*, 10(1):14–19, Sep 1995.

[184] J. M. Song, C.-C. Lu, , and W. C. Chew. Multilevel fast multiple algorithm for electromagnetic scattering by large complex objects. *IEEE Trans. Ant. Propag.*, 45(10):1488–1493, Oct 1997.

[185] G. Strang and G. J. Fix, editors. *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs, 1973.

[186] M. A. Stuchly and S. S. Stuchly. Dielectric properties of biological substances - tabulated. *Journal of Microwave Power*, 15(1):19–26, 1980.

[187] A. J. Surowiec, S. S. Stuckly, J. R. Barr, and A. Swarup. Dielectric properties of breast carcinoma and the surrounding tissues. *IEEE Trans Biomed Eng*, 35:257–263, 1988.

[188] A. Taflove, editor. *Computational electrodynamics: the finite-difference time-domain method*. Artech House, Boston, London, 1995.

[189] A. Taflove, editor. *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Norwood, London, 1998.

[190] A. Taflove and S. C. Hagness, editors. *Computational electrodynamics: the finite-difference time-domain method*. Artech House, Boston, MA, 2000.

[191] A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. John Wiley & Sons, New York, 1977.

[192] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola. *Numerical Methods for the Solution of Ill-posed Problems*. Dordrecht: Kluwer, 1995.

[193] E. Turkel. *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Boston, MA, 1998.

[194] E. Turkel and A. Yefet. Fourth order method for Maxwell's equations on a staggered mesh. *IEEE Antennas and Propagat. Soc. Int. Symp., Montreal, Canada*, 4:2156–2159, Jul 1997.

[195] P. Ya. Ufimtsev. *Method of edge waves in the Physical Theory of Diffraction*. Air Force System Command, Foreign Tech. Div. Document ID No. FTD-HC-23-259-71, 1971.

[196] A. van der Sluis. Condition numbers and equilibration of matrices. *Numerische Mathematik*, 14:14–23, 1969.

[197] C. R. Vogel. *Computational Methods for Inverse Problems*. Society for Industrial & Applied Mathematics, 2002.

[198] J. L. Volakis. A uniform geometrical theory of diffraction for an imperfectly conducting half-plane. *IEEE Trans. Antennas and Propagat*, 34:172–180, 1986.

[199] G. Wahba. *Spline models for observational data*. SIAM, Philadelphia, 1990.

[200] S. Wang and F. L. Teixeira. An efficient PML implementation for the ADI-FDTD method. *Microwave and Wireless Components Letters, IEEE [see also IEEE Microwave and Guided Wave Letters]*, 13(2):72–74, Feb 2003.

[201] S. Wolfram. *The Mathematica book, 5th edition*. Wolfram Media, Inc., 2003.

[202] J. W. Woods and H. Stark. *Probability and Random Processes with Applications to Signal Processing (3rd Edition)*. Prentice-Hall, New Jersey, 2002.

[203] K. S. Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, Mar 1966.

[204] H. A. Zebker and R. M. Goldstein. Topographic Mapping from Interferometric Synthetic Aperture Radar Observations. *Journal of Geophysical Research*, 91:4993–4999, 1986.

[205] F. Zernike. Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. *Physica*, 1:689–704, 1934.

[206] A. P. Zhao. Analysis of the numerical dispersion of the 2D alternating-direction implicit FDTD method. *IEEE Transactions on Microwave Theory and Techniques*, 50(4):1156–1164, Apr 2002.

[207] A. P. Zhao. Two special notes on the implementation of the unconditionally stable ADI-FDTD method. *Microwave and Optical Technology Letters*, 33(4):273–277, May 2002.

[208] A. P. Zhao. Uniaxial perfectly matched layer media for an unconditionally stable 3-D ADI-FD-TD method. *Microwave and Wireless Components Letters, IEEE*

[see also IEEE Microwave and Guided Wave Letters], 12(12):1531–1309, Dec 2002.

[209] F. Zheng, Z. Chen, and J. Zhang. A finite-difference time-domain method without the Courant stability conditions. *Microwave and Guided Wave Letters, IEEE [see also IEEE Microwave and Wireless Components Letters]*, 9(11):441–443, Nov 1999.