A Discrete Dipole Approximation Forward Solver for Microwave Breast Imaging

SAMAR HOSSEINZADEGAN

Department of Electrical Engineering
Group of Biomedical Electromagnetics
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2019
A Discrete Dipole Approximation Forward Solver for Microwave Breast Imaging

SAMAR HOSSEINZADEGAN

© SAMAR HOSSEINZADEGAN, 2019.

Technical report number: R003/2019
ISSN 1403-266X

Department of Electrical Engineering
Group of Biomedical Electromagnetics
CHALMERS UNIVERSITY OF TECHNOLOGY
SE–412 96 Göteborg
Sweden
Telephone: +46 (0)31 – 772 1000
Email: samarh@chalmers.se

Front Cover: The figure on the front cover shows magnitude of an electric field distribution generated with the DDA code.

Typeset by the author using \LaTeX.
Chalmers Reproservice
Göteborg, Sweden 2019
To my family
Abstract

Breast cancer has the highest incidence rate of cancers in women worldwide. Early detection results in a higher survival rate. Drawbacks in conventional imaging modalities, including painful exams, have limited periodic breast screening. Microwave tomography has the potential to be a compelling alternative or complement to other imaging techniques. Advantages of microwave tomography is that it is harmless, comfortable, and cost effective. Microwave tomography has not yet fully been translated into the clinic, even if clinical trials are ongoing.

One important challenge is high computational demands of microwave tomography algorithms. 3D tomography algorithms require multiple hours and a large amount of hardware resources to produce images. 3D imaging algorithms are usually implemented and tested for simulations setup and barely used in clinical settings. 2D microwave tomography algorithms are computationally less expensive compared to 3D algorithms. Few imaging groups have been successful in integrating the acquired 3D data into the 2D tomography algorithms for clinical applications.

The microwave tomography algorithms include two main computation problems; forward and inverse. The forward problem has to be solved multiple times and the resulting computational cost is the time limiting step in microwave tomography algorithms, and this thesis is devoted to addressing it.

In this thesis, the two-dimensional forward problem is modelled and formulated. In particular, the two-dimensional discrete dipole approximation (DDA) is proposed as a new forward solver for microwave tomography. The accuracy of the 2D DDA with respect to sampling number, size, and contrast of target are investigated. Moreover, the 2D DDA time efficiency and computation time are studied.

The forward solver computation times for direct, iterative, and iterative combined with fast Fourier transformation (FFT) solvers are calculated. The observations imply that the 2D DDA is an accurate, reliable, and rapid forward solver, and using the Krylov subspace methods combined with the FFT accelerate the computation time significantly.

Keywords: breast cancer, microwave imaging, tomography, forward problem, computation time.
Acknowledgments

I thank all who in one way or another contributed to the completion of this thesis. First I would like to thank my main supervisor Professor Paul Meaney who always encourages and directs me. His patience towards younger researchers including me is endless. I also want to thank my co-supervisors Associate Professor Andreas Fhager and Professor Mikael Persson for their support and advice. Thanks for all the fruitful discussions and understanding me even in the hard situations.

I am thankful to my colleagues in the Biomedical Electromagnetics group. Thanks for all the nice discussions specially during our group meetings. I also would like to thank everyone at the department. You have made Chalmers an excellent workplace for me.

Without naming each single person, I would like to thank all my friends from Chalmers and elsewhere. I feel so lucky to meet you and have you in my life. Thanks for all the unforgettable moments.

Last but not the least, I would like to thank my parents Fariba and Hamid Reza, my sisters Dorreh and Samane, and my best friend Ramin for their constant love and support.

Samar, Göteborg, 2019
List of Publications

This thesis is based on the work contained in the following appended papers:

Paper A
Samar Hosseinzadegan, Andreas Flager, Mikael Persson, and Paul Meaney, “Application of Two-Dimensional Discrete Dipole Approximation in Simulating Electric Field of a Microwave Breast Imaging System,” Accepted for publication in IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology, Early access is available.

Paper B

Other related publications of the Author not included in this thesis:


*Other publications by the Author not related to the thesis:*

Contents

Abstract i

Acknowledgments iii

List of Publications v

Contents vii

I Introductory Chapters

1 Introduction 1
  1.1 Breast cancer overview ........................................ 2
    1.1.1 Global statistics ......................................... 2
    1.1.2 Risk factors ............................................... 3
    1.1.3 Breast screening .......................................... 5

2 Microwave Imaging 9
  2.1 Imaging methods ............................................. 10
    2.1.1 Radar-based imaging .................................... 10
    2.1.2 Holography ............................................... 10
    2.1.3 Microwave tomography .................................. 11
    2.1.4 Chalmers/Dartmouth breast imaging system .......... 13

3 Computational Aspects for Microwave Tomography 15
  3.1 Microwave tomography algorithms ......................... 16
    3.1.1 Overview of image reconstructions ...................... 16
    3.1.2 Important challenges .................................. 18
  3.2 The forward problem ....................................... 18
    3.2.1 Electromagnetics field theory ........................ 18
Contents

3.2.2 Scattering problem formulation ........................................... 19
3.3 Numerical methods ................................................................. 20
  3.3.1 Finite-Difference Time-Domain method .................................. 20
  3.3.2 Finite element method ....................................................... 21
  3.3.3 Other numerical methods ................................................... 22

4 The Discrete Dipole Approximation ........................................... 23
  4.1 Introduction ........................................................................... 23
  4.2 DDA and the 2D forward problem .......................................... 24
    4.2.1 The 2D DDA formulations ............................................... 24
    4.2.2 Direct and iterative solvers overview .................................. 28
    4.2.3 The 2D DDA as a system of equations ................................. 31
    4.2.4 Krylov subspace methods for the DDA ............................... 32
    4.2.5 Matrix-vector multiplication using the FFT ......................... 34

5 Summary of Papers .................................................................. 39
  5.1 Paper A ................................................................................. 39
  5.2 Paper B ................................................................................. 40

6 Conclusion and Future Work .................................................. 41

II Included Papers
Part I

Introductory Chapters
Introduction

According to the World Health Organization (WHO), breast cancer has the highest rate of incidence in women compared to the other cancer types, and a notable incidence growth has been observed in developing countries [1]. Changes in women’s lifestyle have been a primary factor for this increase [1]. To reduce the mortality rate, extensive cancer screening programs targeting early detection have been implemented in more developed regions for women after age 40. It would be desirable to also screen younger women but it is currently not recommended partially because the accumulated exposure of X-ray radiation over long periods increases the cancer risk. In developing countries, however, lack of financial resources constitutes a major obstacle for launching regular screening programs. New low-cost systems would pave the path for a wider deployment of such programs.

Conventional imaging technologies suffer from several drawbacks and improving state-of-the-art imaging techniques or introducing new screening technologies is needed to improve the accuracy in detection. X-ray mammography is a standard device for breast screening but has a relatively high rate of missed breast tumour detection, termed false negatives, but the false positive rate is also high. Moreover, the use of ionizing X-rays is not completely safe [2,3].

Generally, screening devices should be designed such that they would be cost-effective, sensitive to various tissue types, user-friendly, harmless and comfortable for patients [4]. Microwave technology has great potential to meet these requirements. Microwave tomography is in a developing stage, and has not yet been fully translated into the clinic. One of the known barriers is the substantial computational resources required to reconstruct images. The existing limitations in computational aspects of microwave tomography is the driving goal of this project. We are currently developing a new fast reconstruction algorithm for the systems at Chalmers University of Technology, and at Dartmouth College. The focus of this thesis is on the electromagnetic simulations, i.e. forward models, for these microwave imaging systems.
Chapter 1. Introduction

1.1 Breast cancer overview

Cancer is formed by mutations in genes which lead to an uncontrolled cell-growth of normal body cells. These mutations cause irregular and chaotic cell division and the formation of tumours in affected areas. Breast cancer tumours form mainly in the fibroglandular tissue, including the lobules and the ducts. In addition, the adipose (fatty tissue) within the breast can also be affected [5]. Also, men can get breast cancer but it is less common.

Breast cancer types are mainly categorized according to their growth patterns and location. They are categorized as either invasive or non-invasive (in situ). These two main categories are divided into subcategories based on the location within the breast as [5]:

- Ductal carcinoma in situ (DCIS)
- Lobular carcinoma in situ (LCIS)
- Invasive ductal carcinoma (IDC)
- Invasive lobular carcinoma (ILC)

Among these types, the invasive ductal carcinoma (IDC) is the most common type which initially forms in ducts and in later stages invades the surrounding tissues. Furthermore, breast cancer is staged based on the size of the tumour from 0 to 4, where 0 stage refers to non-invasive, in situ, i.e. tumours localized to its original location. In stage 4, the tumour has metastasized to nearby tissues as well as possibly to distant organs including lungs, brain, liver, or bones [6].

For the purpose of determining the stage of the tumour or monitoring treatment progress, imaging techniques such as X-ray mammography, magnetic resonance imaging (MRI), computerized tomography (CT), and position emission tomography (PET) are used. Brief summaries of these techniques are given in Section 1.1.3.

1.1.1 Global statistics

In this section, the 2018 global breast cancer statistics are summarized. These data are provided by the GLOBOCAN 2018 platform and are compiled by the International Agency for Research on Cancer (IARC) [7].

Figure 1.1a shows that lung cancer has the highest incidence rate with more than two million cases. This is an epidemic proportion of about 11%, i.e. the proportion of lung-cancer to all cancer types, for both genders. Breast cancer is estimated to be second, with only slightly lower epidemic proportions. Figure 1.1b illustrates the mortality rates due to different cancer types for both genders worldwide. Figures 1.2a and 1.2b show the estimated percentages for female breast cancer incidence and mortality. As is noted, breast cancer is the most common cancer type among
women with a proportion of about 24% of all types of cancer. The mortality rate due to this disease is the highest of all cancer types and estimated to be about 15% of all cancers, or 626,679 over the year. In North America, Australia, Northern and Southern Europe, higher incidence rates have been reported. Rates for South America and Eastern Europe and most parts of Africa and Asia are low, as are shown in Figure 1.3.

As can be seen in Figures 1.3 and 1.4, there is no direct relation between incidence and mortality rates. The incidence of breast cancer is shown with colors, darker colors correspond to higher incidence rate, and is seen that it is highest in the industrialized world. In Figure 1.4, the mortality rate is shown, with darker regions corresponding to higher rates, which are found in developing regions such as Africa. This data illustrates that mortality rates could be reduced with improved economic development and a corresponding expansion of welfare and health care sectors of these countries. One key factor to reduce the mortality rate is early detection. Moreover, other factors such as genetics can also impact the incidence and mortality rates.

1.1.2 Risk factors

In this section, risk factors are briefly discussed and are summarized based on the study performed by P. Maisonneuve [8]. The global incidence and mortality patterns are an indication of existing risk factors. According to this study, breast cancer risk factors can be divided into three main groups as:

- Demographic, lifestyle, and environmental factors
- Reproductive and hormonal factors
- Genetic factors

The two main factors in the first group are age and race. The incidence rate increases with age. Breast cancer is very common in women over age 50 and its incidence increases until age 70 and decreases afterwards [8]. It is also mentioned that breast cancer rates are higher among white women compared to black women. Moreover, there is a lower incidence rate among those of Asian origin and Pacific Islanders [8]. In addition to age and race, other demographic and lifestyle factors such as obesity, diet, and physical activity can also affect breast cancer risks which are categorized within the demographic and lifestyle group. The breast cancer risk is also related to the women reproductive and hormonal changes during their lives. Various reasons including the early age at menarche, first pregnancy age, breastfeeding and etc stimulate cell growth and consequently risk of breast cancer. The chance of developing cancer for women with a first-degree relative diagnosed with breast cancer is higher compared to others. It is predicted that the risk for these cases is doubled compared to others [8].
Figure 1.1: Estimated number of a) new cases and b) deaths worldwide for both sexes in 2018, *Source* GLOBOCAN 2018, IARC.

Figure 1.2: Estimated number of a) new cases and b) deaths worldwide for females in 2018, *Source* GLOBOCAN 2018, IARC.
1.1. Breast cancer overview

Early detection of breast cancer is the key to improving the chance of survival. The validity of this claim has been confirmed universally [9–11]. One of the most cited studies on the correlation between regular screening and mortality was performed by Shapiro et al. [11] in the 1970s. The objective was to evaluate periodic breast cancer screening which included mammography and palpation. Two randomized women groups of size 31,000, aged 40 to 64 years were selected. Study group women were
offered a screening examination, and the control group received usual practices and medical care. The results of this study showed that regular screening contributed to lower mortality compared to the control group. Similar observations from studies in various research groups over the last few decades have also been reported [12–15].

Breast imaging is usually performed for detection, diagnosis (after the clinician has seen an abnormality in a mammogram), and tumour progression monitoring during treatment. Conventional imaging techniques include X-ray based mammography screening, computed tomography (CT scan), magnetic resonance imaging (MRI), and positron emission tomography (PET scan). In addition to these, newer techniques such as digital breast tomosynthesis, contrast mammography, and contrast-enhanced MRI can be used in both screening and diagnosis (ch. 18 in [16]).

The primary method for routine breast cancer screening is X-ray mammography, Figure 1.5. A standard mammogram produces two-dimensional projections of the breast from two different angles, craniocaudal (CC: top-bottom view) and mediolateral oblique (MLO: side view at a specific angle). Additional CC views or spot compression can be performed in cases of the presence of an abnormality. The operating principle for mammography is based on imaging of the contrast between various tissue-types within the breast as a greyscale image. This technique requires firm and even compression to ensure enhanced resolution and higher contrast. Unfortunately, mammography is considered uncomfortable and painful by many patients.

One study on assessing diagnosis accuracy for different imaging techniques was performed by Berg et al. [17]. In that study the diagnostic performance of different imaging devices on 258 patients with proven lesions was tested. It was shown that mammography had an accuracy of about 70.2%, i.e. 181 patients out of 258 were correctly diagnosed. It was also limited to 67.8% in sensitivity (i.e. true positive) and 75% in specificity (i.e. true negative), respectively [17]. Distracting lesions and breast density are primary confounding features for missing a breast tumor. Moreover, due to the radiation exposure young women are sometimes not screened regularly because of the risk/benefit balance.
1.1. Breast cancer overview

Magnetic resonance imaging, MRI, which uses strong magnetic (non-ionizing) fields, is an alternative for breast imaging. MRI is mostly used as a complement to mammography for further evaluation of suspicious findings. One of the advantages with MRI over X-ray mammography is that it is insensitive to normal dense breast tissue [19]. It is also used with a gadolinium contrast agent which can be toxic to women with compromised kidneys [20]. Therefore, it has also been used more frequently for screening and diagnosing young women. MRI is, however, significantly more expensive than other imaging methods, and is, therefore, not suitable for use in screening. Another problem is a high rate of false-positives which leads to over diagnosis in screening [19,21].

Ultrasound screening can be used as a complementary imaging tool to both X-ray mammography and MRI. It sends high-frequency sound waves from the transducer into the breast, and the reflected echos are collected for further processing and image generation [22]. One caveat about ultrasound screening is that its precision is highly operator-dependant, which introduces considerable risk of errors in the diagnosis. It is also known to have a high number of false positives. Berg et al. [23] investigated the impact of combined screening, ultrasound and X-ray mammography, compared to X-ray mammography alone. In that study, they concluded that although ultrasound may reveal small and node-negative breast cancers not visible in mammography, the increased number of false positives is significant [23].

A CT scanner encircles the body and creates an image from measured X-ray projections at a large number of angles. The measured projections are processed to create cross-sectional images of the body. CT scanning is usually not used for screening, but sometimes for further investigations after detection [24]. As an example, it can be used to inspect a tumor (location, size, and shape) before and after treatment planning, or as guidance in needle biopsy [24].

PET scanning is a technology based on nuclear medicine. In PET scans, a radioactive sugar is injected into the patient. Its decay produces radioactivity that is measured outside the body with detectors encircling the body. Images are formed in a similar way as CT-scans. This imaging technique is mostly used during treatment planning since the metabolic activity frequently illustrates changes earlier under treatment than morphological changes such as tumor size (ch. 19 in [16]).

The aforementioned drawbacks in conventional imaging techniques illustrate the need and room for developing new image acquisition techniques to further improve diagnosis. Microwave technology appears to be a promising technique for this situation. The technology is harmless, cost-effective, comfortable, and sensitive enough to various tissue types. These characteristics make it an intriguing candidate to complement or replace other modalities.
Microwave Imaging

With microwave imaging, it is possible to retrieve characteristic information of the interior of an object. Such information could be position, shape, size, and dielectric properties and also information about smaller objects immersed within the large objects. The operating principle is based on transmitting electromagnetic waves through the object and then measuring and processing the forward and back-scattered waves.

Microwave imaging has been studied for a variety of applications over recent decades. These applications range from medical diagnosis, i.e. breast imaging [25], brain imaging [26–28], kidney imaging [29], cardiac imaging [30], and bone density measurements [31] to other fields including geophysical monitoring [32], civil engineering [33], and industrial engineering [34].

For cancer detection, microwave medical imaging relies on contrast in dielectric parameters between malignant tumour tissue and the surrounding normal breast-tissue. Although microwave imaging for medical applications has been extensively investigated since the 1970s, this imaging modality has not yet been widely transferred to the clinic.

One of the earliest medical microwave imaging systems was introduced by Larsen and Jacobi who used microwave signals to image canine kidneys [29,35]. From the 1970s until now, multiple imaging systems have been built by many microwave imaging groups, and their drawbacks and limitations have been investigated extensively.

In microwave imaging, the computational aspects are known as a barrier to the development and usability of the technology. It will be discussed further in Chapter 3. Aside from these known challenges, medical microwave imaging has been introduced as an alternative or complement to conventional imaging systems such as mammography, with the motives being recognized as low-cost, harmless, non-invasive, and convenience.
2.1 Imaging methods

The information to be recovered from scattered field measurements varies in practical applications based on different operating principles. The objectives can range from localization and detection of the presence of an object to full reconstructions of quantitative images of the object. For biomedical purposes the most common application is to completely image the electromagnetic parameter distributions (e.g. dielectric permittivity, electric conductivity) of the objects under test [36].

Microwave imaging methods could be divided into at least two categories, qualitative and quantitative methods, based on the characteristics and information content of the output image. In qualitative imaging methods, the reconstructed information is restricted to relative maps of scattering signal strengths revealing only location, size and shape of scatterers. One advantage of these methods is their computational efficiency and short reconstruction time. A disadvantage is that they require broad band measurements. This could lead to longer measurement times depending on the techniques used. Conversely, quantitative imaging methods aim at reconstructions of dielectric maps of the object under-test. These methods are known to be computationally demanding and in many implementations restricted to cases where a priori information about the object is available (chapter 5 in [36]).

Qualitative methods that have been researched include holography and radar-based imaging. Quantitative methods usually refer to tomography. These three different imaging methods will be further described here.

2.1.1 Radar-based imaging

Multiple radar-based systems have been developed and tested in both experiments and clinical settings [37–39], and their fundamental assumption is based on the differences between dielectric properties of healthy and malignant tissues over the microwave frequency range [37].

In active radar-based imaging, the target is illuminated by signals transmitted and received by an array of antennas. Data are also acquired in multiple planes and are simultaneously processed to produce the whole 3D image via a reconstruction algorithm [40]. The operating principle is based on transmission and reception of pulses. The short high energy pulses are transmitted and the returning echos are recorded. The commonly used imaging radar systems are Real Aperture Radar (RAR) and Synthetic Aperture Radar (SAR).

2.1.2 Holography

Another microwave-based imaging technique, microwave holography, has also been used to image dielectric targets [41]. This technique is relatively new in the field of medical imaging but has been employed for concealed weapons and contraband at
airports and other secure locations [42]. In modern microwave holography, coherent back-scattered signals are acquired on a surface and the reconstruction of object is performed through multiple direct and inverse Fourier transformations. Unlike the case in tomography, the reconstructions are not typically done in 2D slices. The measured data from the surface are simultaneously processed into a single reconstruction to produce a 3D distribution of the target [41,43].

2.1.3 Microwave tomography

Microwave tomography utilizes scattered electric field data from an object to reconstruct quantitative maps of the property distributions, usually permittivity and conductivity. For breast cancer detection, this technology aims at imaging the differences in properties between healthy and malignant tissues. Systems typically operate in the range of 500 MHz - 3 GHz.

Microwave tomography is in a developing stage, and has reached the stage of clinical trials in some cases [25,31]. Historically, the large property contrast between different biological tissues has been a major barrier in microwave tomography [44] and [45].

2.1.3.1 Microwave tomography systems

In microwave tomography, the target is illuminated with microwave signals from multiple directions, and the resulting scattered signals are measured at multiple locations. The illumination is made by consecutive transmissions from one antenna at a time in an array of antennas enclosing the imaging region. Antenna arrays are often configured in a circular pattern. For the Chalmers system, all antennas act as both transmitters and receivers, where the antennas that are not transmitting, act as receivers [46,47]. Measurements are repeated until all combinations of transmit and receive antennas have been measured. A schematic representation of such systems is illustrated in Figure 2.1.

The challenges for these systems are signal power loss, limited penetration depth, low resolution, and a need for enormous computational resources [40]. These challenges have hindered the development process for imaging systems and their algorithms. In microwave imaging, one important aspect is to optimize power penetration into the body and between antennas and the skin. To accomplish this, various coupling liquids such as water, glycerin and others are used. This coupling liquid also has an indirect impact on the image resolution. A better resolution (the smallest distinguishable pixel), requires higher frequency with a concomitant shorter wavelength; however, at the price of a shallower penetration depth. This in turn becomes a challenge for imaging, especially for deep-body organs. The remarkable signal attenuation in both biological tissues and the matching liquid lowers the penetration depth as well as the usable frequency range.
For these imaging systems, it is crucial that the transmitted waves also penetrate through the object under test. A biological tissue placed in air exhibits a significant contrast to the background and thus a high degree of scattering from the air/skin interface. With a lower coupling medium contrast, waves are better able to penetrate and thus allow for higher signal strength inside the body. Therefore, more accurate measurements can be acquired for signals reaching into the body. The lower scattered waves are capable of deeper penetration into the organs. For deeper penetration, two hypotheses have been suggested and employed [40]. The first approach is based on the same principle as was introduced by Larson and Jacobi [29]. In their experimental set-up, the antennas and the organ were immersed in water, which, then acted as a coupling liquid and lowered the contrast. Many different coupling liquids have been explored later. For example, Dartmouth College has used saline solutions in their early imaging systems. The coupling liquid used in the current imaging system is a mixture of glycerin and water. The disadvantages of using coupling liquids include the power loss and difficulty in maintenance of the imaging system [40]. The second approach is to place antennas in direct contact with the body to provide high coupling efficiency. In this case, wasted power is less of an issue.
2.1.4 Chalmers/Dartmouth breast imaging system

The first clinical microwave breast imaging system was developed at Dartmouth College [25]. In this system, a circular array consisting 32 monopole antennas was submerged in a saline bath, as the coupling fluid between the antennas and the breast. The operating frequency range for this system was from 300 to 900 MHz. The data from multiple imaging planes, from the patient chest wall to nipple position, were acquired, and this data processed to reconstruct images.

Several generations of systems have been developed at Dartmouth College [25,31]. The recently developed imaging system at Chalmers University of Technology is also built on a similar basis [46,47]. The microwave-based tomography system consists of a set of 16 monopole antennas. These antennas, located in the imaging tank, are equiangularly spaced on a circle of diameter 15.2 cm. The imaging tank is filled with a liquid mixture of glycerin and water. In contrast to other coupling liquids (i.e. saline or water), the mixture of glycerin and water is used to lower the contrast between background and breast tissue. With a reduced contrast, the reflections from the skin are correspondingly reduced. This coupling liquid also ensures that the surface waves and the undesirable multi-path signals are virtually eliminated by the high-conductivity nature of this liquid.

This conductive material is also beneficial in modeling the imaging domain and antennas, i.e. its contribution to computational aspects will be further discussed in Chapter 3.
In this section, state-of-the-art in medical microwave tomography algorithms is discussed. A number of methods and algorithms have been investigated by multiple research groups since the possibility of employing microwave signals for medical imaging was first studied [25, 29, 48, 49]. Notwithstanding the centralized focus on these algorithms, they have scarcely been translated into the real world. Principal among the causes for this are the low efficiency of different numerical methods and the presence of high-contrast among different tissue types. Additionally, the design of imaging systems and improving their measurement accuracy can contribute to future success or usefulness in microwave tomography. For medical imaging, high sensitivity and specificity are desirable, and these characteristics generally narrow the choices of algorithms.

The objective of microwave tomography is to create images of the dielectric properties, i.e., permittivity and conductivity. To construct an image, an inverse problem of Maxwell’s equations is typically solved; that is, to infer the source from observations of electromagnetic waves. One important aspect of inverse problems is whether it is well-posed or ill-posed. A well-posed problem satisfies the following postulates [50]:

- Existence: meaning that the solution to the problem exists,
- Uniqueness: meaning that the existing solution is unique,
- Continuity: meaning that the solution is continuous with respect to initial conditions and boundary conditions.

These requirements were formulated based on a concept introduced by Jacques Hadamard [50]. The problem is well-posed if all of the three postulates are satisfied. If at least one of these conditions is not fulfilled, the problem is referred to as an ill-posed problem.
The inverse scattering problems in microwave tomography are known to be ill-posed due to the existence of non-unique solutions, and also that the solutions do not depend continuously on the data. To obtain a unique solution, a large amount of measurement data is required, and this impacts the size of tomographic-system and the problem. Moreover, the solutions do not depend continuously on the data which is due to the irregular dependency of reconstructing dielectric property distributions to the changes in noisy measurement data [51]. To overcome the lack of continuity, reconstruction algorithms need to be regularized.

3.1 Microwave tomography algorithms

Here we present a brief introduction to image reconstruction algorithms for microwave tomography.

3.1.1 Overview of image reconstructions

Electromagnetic image reconstruction is considered as two problems:

- Direct or forward problem
- Inverse problem

The electromagnetic forward problem is designed to predict the electromagnetic propagation in a domain with known dielectric parameters and initial and boundary conditions. The inverse problem recovers the dielectric properties as a function of scattering during electromagnetic propagation through the domain. The inverse problem in microwave tomography is ill-posed due to the large property contrasts between tissues and is also non-linear. Therefore, more computationally expensive iterative optimization algorithms are required.

The main principle of these types of optimization algorithms is to minimize the difference between measured data and a computed solution (i.e. forward solution). In each iteration the dielectric properties within the domain are updated and the process is repeated until convergence criteria are satisfied. A general overview for reconstruction algorithms employed in microwave tomography [52] is given in Figure 3.1. An examination of reconstruction algorithms reveals that the forward solution is recalculated at each iteration such that a majority of the computational complexity in these algorithms is due to the forward problem. Accordingly, the forward problem formulation and its computational complexity constitute a notable degree of importance for actual applications from the medical imaging point of view. The forward problem will be further discussed in detail in Section 3.2.
3.1. Microwave tomography algorithms

Start reconstruction

Give initial property distribution

Map the properties to the numerical grid

Compute the forward problem solutions for all transmitters

\[ E_{\text{simulated}} - E_{\text{measured}} \]

Check Convergence Criteria

Update property distribution

Compute the property updates

Form the Hessian matrix (\( J^T J \)) and determine regularization parameter

Calculate the complex form Jacobian matrix and its log-transformed version

End construction

Yes

No

Figure 3.1: Microwave tomography algorithm overview.
3.1.2 Important challenges

A primary challenge is to accurately model the field distributions for the dielectric property distributions while maintaining reasonable computation times. The amount of computational resources required for a 3D reconstruction algorithm is exceptionally high in comparison to 2D algorithms. Consequently, some medical microwave tomography research groups have concentrated on 2D tomography, and 3D modeling has to a large extent been used only in simulations and barely in clinical applications.

Most of the study cases on biomedical microwave tomography indicate that there exists a limitation in incorporating acquired measurements data with reconstruction algorithms. One challenge is the use of 2D reconstruction algorithms integrated with actual 3D measured data. The clinical prototype developed at Dartmouth College has been used successfully in 2D imaging [25]. Apart from the Dartmouth research group, the imaging group at the Chalmers University of Technology has also been able to produce 2D images using the 3D acquired data from the prototype system [46,47].

3.2 The forward problem

To solve the forward problem effectively, we need to understand electromagnetics and its fundamentals.

3.2.1 Electromagnetics field theory

Maxwell’s equations characterize the physical interactions between fields, fluxes and material properties in an environment. The complex form of these equations in the frequency domain is written as

\[
\nabla \times \vec{E} = -i\omega \vec{B},
\]

\[
\nabla \times \vec{H} = \vec{J} + i\omega \vec{D},
\]

\[
\nabla \cdot \vec{D} = \rho,
\]

\[
\nabla \cdot \vec{B} = 0.
\]

where \(\vec{E}\) is the electric field, \(\vec{B}\) is the magnetic flux density, \(\vec{H}\) is the magnetic field, \(\vec{D}\) is the displacement field, \(\vec{J}\) is the free current density, \(\rho\) is the charge density, \(\omega\) is frequency, and \(i\) is the imaginary unit. The following constitutive relationships can be used to replace fields and fluxes:

\[
\vec{J} = \sigma \vec{E},
\]

\[
\vec{B} = \mu \vec{H},
\]

\[
\vec{D} = \epsilon \vec{E}.
\]
3.2. The forward problem

where \( \sigma \) is the electrical conductivity, \( \epsilon = \epsilon_0 \epsilon_r \) is the environment permittivity which can also have a complex value in form of \( \epsilon = \epsilon_0 \epsilon_r + \frac{\sigma}{\omega} \), and \( \mu = \mu_0 \mu_r \) is the environment permeability. \( \epsilon_0 \) and \( \mu_0 \) denote free-space permittivity and permeability, respectively. \( \epsilon_r \) and \( \mu_r \) denote relative permittivity and permeability of the material, respectively.

In absence of free charge or current, inserting 3.5 to 3.7 into 3.1 to 3.4 gives:

\[
\nabla \times \tilde{E} = -i\omega \mu \tilde{H}, \\
\nabla \times \tilde{H} = i\omega \epsilon \tilde{E}, \\
\n\nabla \cdot (\mu \tilde{H}) = 0, \\
\n\nabla \cdot (\epsilon \tilde{E}) = 0.
\]

(3.8)  (3.9)  (3.10)  (3.11)

Taking the curl of 3.8 and inserting 3.9 into it gives:

\[
\nabla \times (\nabla \times \tilde{E}) = \omega^2 \mu \epsilon \tilde{E}
\]

(3.13)

Rearrangement of this equation results in

\[
\nabla \cdot (\nabla \cdot \tilde{E}) - \nabla^2 \tilde{E} = k^2 \tilde{E}
\]

(3.14)

where \( k^2 = \omega^2 \mu \epsilon \) is the squared wavenumber. The wavenumber can be a real constant or a complex number for energy absorbing media. For nonmagnetic material (i.e. \( \mu_r = 1 \)) with a complex permittivity \( \epsilon \), squared wavenumber is in form of

\[
k^2 = \omega^2 \mu_0 \epsilon_0 \epsilon_r + i\omega \mu_0 \sigma
\]

(3.15)

In regions where \( \rho = 0 \), \( (\nabla \cdot \tilde{E}) \) goes to zero which leads to the Helmholtz equation

\[
\nabla^2 \tilde{E} - k^2 \tilde{E} = 0
\]

(3.16)

The solutions to this equation in inhomogeneous media are needed in order to solve the forward problem. Analytic solutions are typically known for simple geometries such as spheres and infinite cylinders. For more realistic cases, different numerical methods have been developed. Some of these numerical solvers will be discussed in Section 3.3.

3.2.2 Scattering problem formulation

One application for the Helmholtz equation is to separate its formulation into the incident and scattered fields. The problem is solved in a bounded region \( V \subset \mathbb{R}^d \), \( d = 2, 3 \), as shown in Figure 3.2. The problem is to determine the scattered electric field, \( \tilde{E}_{\text{scat}} \), due to the scattering object and the incident field such that

\[
\tilde{E}_{\text{tot}} = \tilde{E}_{\text{inc}} + \tilde{E}_{\text{scat}}
\]

(3.17)
Chapter 3. Computational Aspects for Microwave Tomography

The total electric field $\vec{E}_{\text{tot}}$ satisfies (3.16) in $\mathbb{R}^d \setminus \bar{D}$ as well as the boundary conditions on $\partial D$. With a given incident field $\vec{E}_{\text{inc}}$ and dielectric properties of the background, the problem is to solve for the unknown scattered field $\vec{E}_{\text{scat}}$.

3.3 Numerical methods

To solve the electromagnetic equations involved in microwave scattering problems, different numerical methods have been employed. The most popular schemes are the finite-difference (FD), finite-difference time-domain (FDTD), finite element methods (FEM), and method of moments (MoM). Among these numerical methods, the FDTD and FEM methods have been employed as the forward solver for the microwave imaging systems at Chalmers University of Technology and Dartmouth College. A brief overview of these methods is given in Sections 3.3.1 and 3.3.2.

3.3.1 Finite-Difference Time-Domain method

The FDTD method is one of the most commonly applied numerical solvers to microwave problems. Its popularity is due to being efficient and straightforward in implementations. In the FDTD method, the problem is solved on a structured Cartesian grid, and the equations are formulated in the time domain [53]. In this method, the entire solution domain is discretized, and the spacing discretization must be sufficiently small to resolve the object and to provide enough resolution over a wavelength. For example this is necessary for handling both the small electromagnetic wavelengths and the small tumors in the breast. In this format, a very dense com-
putational domain is obtained. The computation time scales up rapidly for larger problem sizes. Moreover, modeling some components such as antennas are difficult in FDTD because of their complex geometry requiring even denser meshes than usual which subsequently adds to the computational complexity.

A related drawback of this method in imaging is its inflexibility for handling geometries of arbitrary shape. Moreover, the boundary conditions are not explicitly considered in the original FDTD scheme. For imaging purposes, we are interested in an open space for the electric field calculations which require enlarging the computation domain. Consequently, the computational complexity is expected to grow in this case [54].

The FDTD scheme developed for the Chalmers microwave imaging system has been implemented and tested [54], and a complete description of the implementation of this method is given in [54].

### 3.3.2 Finite element method

The FE method has been widely used as a numerical solver for the forward problem. In this method, the imaging domain is subdivided into small domains of simple shapes such as tetrahedra (3D) or triangles (2D) [55]. The resulting sub-domains comprise a complete mesh in this scheme. The unstructured sub elements have made the FE method a powerful tool in dealing with complex geometries, especially those with an arbitrary form. For example, in a situation where a complex shape of object resides in a relatively uniform medium, the shape and the area immediately surrounding it can be finely discretized, while the more homogeneous zones can be more coarsely sampled without loss of accuracy. The electric field within the imaging domain is expressed with the help of a linear polynomial on each of these meshes. For an accurate approximation to the electric field distributions, a minimal size of elements is necessary, and the finer meshes result in a massive system of equations with a huge number of unknowns. Consequently, the computational complexity grows in this case and requires many hours to produce an image. However, these costs are largely instigated by the fact that the resulting matrices are banded for which there are numerous efficient computation methods. Another drawback of this method is that it also uses large computer resources. The FE method usually requires more CPU time and memory compared to FDTD for the 3D setting but are comparable in 2D problems for similar resolution.

The Dartmouth research group has used the FE scheme for calculating the electric field of the clinical imaging system. In their scheme, they implemented the dual mesh approach [56] to image tissue properties. First, a uniform, dense mesh is used to calculate the electric field distributions over the entire domain and then a less dense mesh is used for representing the dielectric property distributions in the imaging domain. In this way, they have been successful in obtaining images in near real time.
3.3.3 Other numerical methods

In addition to the finite difference and finite element methods, there exist other numerical solvers that are usually some derivative of integral equations. The methods which are meant for the integral equations are typically converted into a linear system of equations, which are solved to obtain the electric field in the domain. The method of moments (MoM) and the discrete dipole approximation (DDA) belong to this category. The DDA will be discussed further in the next chapter.
The Discrete Dipole Approximation

In this section, an overview of the discrete dipole approximation is given. We briefly discuss its fundamentals, and possible applications and implementations.

4.1 Introduction

The idea of the discrete dipole approximation was first introduced in 1964 to study the optical properties of monomer units, such as a molecular aggregates, molecular crystals, or polymers [57]. Later in 1973, this method was proposed by Purcell and Pennypacker [58] to calculate absorption and scattering for dielectric grains of arbitrary shape using a set of point dipoles, but they did not name this method as the discrete dipole approximation or the coupled dipole approximation as is used today. Later Draine and his colleagues further explored this method [59–61]. They implemented the DDA and made it publicly available as a computer code named DDSCAT [61,62]. The more theoretical aspects of the convergence and accuracy of the DDA and comparisons of available open source codes of the DDA was further studied in detail by Yurkin and his colleagues [63–65], who also developed another open source code, ADDA [66] with additional computational efficiency improvements. One good comprehensive review of the DDA and its development is given in [67]. To the best of our knowledge, a comprehensive review paper on the DDA after this publication does not exist.

The DDA, sometimes called as the coupled dipole method or approximation, is used to calculate light scattering properties of an object with an arbitrary geometry. The DDA principle is based on the idea of replacing the object with an array of dipoles (Figure 4.1).

The DDA can be formulated and developed in different ways [68]. Although other derivations deviate from the original implementation by Purcell and Pennypacker, and have typically not been called DDA, their equivalence has been proven [69,70], and Yurkin [68] has explained and compared these derivations briefly.
4.2 DDA and the 2D forward problem

In general, the forward model zone consists of a coupling medium and set of antennas enclosing the scatterer. The electric field due to this scattering is the sum of the incident and scattered electric fields in the form of Equation 3.17. We employ the DDA, instead of the FDTD or FE schemes, to calculate the electric field in the forward model zone. In the following sections, \( \vec{\mathbf{}} \) and \( \overline{\mathbf{}} \) are vector and matrix notations, respectively.

4.2.1 The 2D DDA formulations

In the context of the DDA, the forward model zone is replaced with arrays of dipoles equally spaced in \( N \) small cells such that

\[
D = \lim_{N \to \infty} \bigcup_{i=1}^{N} D_i
\]

where \( D \) and \( D_i \) represent the forward model zone and cells, respectively. The two-dimensional incident electric field due to a transmitting antenna is caused by a vertical electrical line source (ELS) and is described as

\[
\vec{E}_{\text{inc}}(\vec{r}_i) = \frac{I_0 \omega \mu_0}{4} H_0^2(k_b \vec{R}_i)
\]

where \( \vec{E}_{\text{inc}}(\vec{r}_i) \) denotes the incident electric field at any dipole location \( \vec{r}_i \). \( I_0 \), \( \omega \), and \( \mu_0 \) are the current amplitude, operating frequency and free-space permeability, respectively. \( H_0^2 \) is the zero-order Hankel function of the second kind, \( k_b \) is the background wavenumber, \( \vec{R}_i \) are the dipole distances from the transmitting antenna.
Moreover, the scattered electric field at a particular location can be considered as the joint responses from all other dipoles, and in this format reads as:

\[ E_{\text{scat}}(\vec{r}_i) = \sum_{j=1}^{N} \vec{G}_{ij} \vec{P}_j \] (4.3)

The operator \( \vec{G} \) contains the information about all dipole interactions (i.e. \( \vec{G}_{ij} \) are the elements of the interaction matrix for all \( i \) and \( j \)), and the term \( \vec{P}_j \) is the polarization at each dipole location. Its value represents the material properties which are assumed to be constant inside each of these cells. Using (3.17), (4.1), and (4.3) produces the total electric field at any dipole location \( r_i \) [61]:

\[ E_{\text{tot}}(\vec{r}_i) = E_{\text{inc}}(\vec{r}_i) + \sum_{j=1}^{N} \vec{G}_{ij} \vec{P}_j, \] (4.4)

The \( \vec{G}_{ij} \) term from the Green’s function for the 2D Helmholtz equation [71] describes the interactions between two dipoles located at \( r_i \) and \( r_j \) and is written as

\[ \vec{G}_{ij} = \frac{-i}{4} H_0^{(2)}(k b \vec{r}_{ij}) \] (4.5)

where \( \vec{r}_{ij} \) is the distance between nodes \( i \) and \( j \). Moreover, at the macroscopic level, the total electric field is proportional to the polarization \( P \) via

\[ P = \varepsilon_0 \chi E_{\text{tot}} \] (4.6)

where \( \chi \) is electric susceptibility. Additionally, the vector \( P \) and the total electric field in the microscopic analysis are also related to each other via

\[ P = N \varepsilon_0 \alpha E_{\text{loc}} \] (4.7)

where \( N \) is the number of molecules per unit volume and \( E_{\text{loc}} \) is the microscopic electric field sensed by each molecule. This is also the field that polarizes them to exhibit dipole moment \( p \). The term \( \alpha \) expresses the relationship between the susceptibility as a macroscopic parameter and the molecular polarizability as a microscopic parameter.

The polarizability term \( \alpha \) is usually modelled via the Clausius-Mossotti polarizability for the 3D case. The Clausius-Mossotti polarizability relates the relative permittivity \( \varepsilon_r \) of a dielectric to the polarizability \( \alpha \) of the molecules constituting the dielectric property. For a spherical inclusion located in free space, the dipole polarizability is [60,72]

\[ \alpha_{CM} = 3v \frac{\varepsilon_r - 1}{\varepsilon_r + 2} \] (4.8)
Figure 4.2: Photograph of the measurement set up including the tank, antennas and the phantom cylinders.

Here $\epsilon_r$ and $v$ are the relative permittivity and volume of the sphere. In [73], this equation is generalized to an arbitrary background as

$$\alpha_{3D} = 3v\frac{\epsilon_t - \epsilon_b}{\epsilon_t + 2\epsilon_b}$$

(4.9)

where $\epsilon_b$ and $\epsilon_t$ are the complex permittivities of the background and target, respectively. The applicability of this formulation to the 3D case has been previously tested in [73].

To increase the accuracy, we suggest to use a more suitable polarizability term for the microwave imaging systems. Since our imaging system, Figure 4.2, consists of a tank filled with a liquid mixture of glycerin and water with different polarization behaviour on the molecular level, we select a more general form of the Clausius-Mossotti relationship, the Maxwell-Garnett formula. The main advantage to the Clausius-Mossotti model is that it is also valid for composite media which is the normal situation for many applications including our microwave imaging system. The Maxwell-Garnett formula can be expressed as

$$\frac{\epsilon_{\text{mix}} - \epsilon_b}{\epsilon_{\text{mix}} + 2\epsilon_b} = \sum_{k=1}^{K} c_k \frac{\epsilon_k - \epsilon_b}{\epsilon_k + 2\epsilon_b}$$

(4.10)

In Equation 4.10, $\epsilon_{\text{mix}}$ denotes the composite medium consisting of $K$ media with permittivities, $\epsilon_k$. The coefficients, $c_k$, relate the volume of the inclusion to its concentration which is defined as $\frac{m}{M_rV}$ where $V$, $m$, $M_r$ are volume, mass, and the molecular weight of the inclusion, respectively. The general formula for the dimensionless $\alpha$ in m-dimensions is [74]:

$$\alpha = m\frac{\epsilon_t - \epsilon_b}{\epsilon_t + (m - 1)\epsilon_b}$$

(4.11)
4.2. DDA AND THE 2D FORWARD PROBLEM

Based on Equations (4.10) and (4.11), the two dimensional $\alpha$ in the form of the Maxwell-Garnett formula is expressed as:

$$\alpha_{2D} := \frac{\epsilon_{\text{mix}} - \epsilon_b}{\epsilon_{\text{mix}} + \epsilon_b} = \sum_{k=1}^{K} c_k \frac{\epsilon_k - \epsilon_b}{\epsilon_k + \epsilon_b}$$ (4.12)

The concentration coefficients $c_k$ are modified based on the area of the inclusion for the 2D case.

In this format, using (4.6) and (4.7), the equation (4.4) converts to:

$$\bar{E}_{inc}(\bar{r}_i) = \frac{\bar{P}_i}{\alpha_i} - \sum_{j=1}^{N} G_{ij} \bar{P}_j$$ (4.13)

The system of equations (4.13) is solved for vector $\bar{P}$, and then the equation (4.4) is used to calculate the total electric field in the forward model zone. The solution to the linear system of equations can be obtained via direct or inverse solvers. Such solvers are further discussed in section 4.2.2.

Samples of the results from the proposed 2D DDA are presented here. COMSOL multiphysics is also used as baseline results to benchmark the 2D DDA forward solutions. The 2D DDA solver is employed to calculate the electric field distributions in the forward model zone. To accomplish this, an array of dipoles is utilized to model the physical geometry of our imaging system. Figure 4.3 presents a 2D geometry of a uniform grid with different numbers of dipoles. In the forward model, an extended square modeling zone with size length of 40 cm is considered. Two circular inclusions are assumed within the imaging domain. An inclusion (target #1) of diameter 6 cm is at $x = 4$ and $y = -2$ and another inclusion (target #2) is of diameter 3 cm located at $x = -3$ and $y = 4$. The transmitting antenna is located at $x = -7.62$ cm. The 2D DDA forward solutions are investigated with respect to multiple parameters:

- Number of dipoles for representing the geometry
- Target size
- Property contrast

A comprehensive study of these parameters and their impact is provided in Paper A. Here some cases for electric field distributions are presented.

In these simulations, accuracy is tested for the number of dipoles to be 900 and 3600. For two circular inclusions, two cases with different relative permittivities are investigated:

1. Relative permittivities of targets #1 and #2 are set to 25 and 45, respectively,

2. Relative permittivities of targets #1 and #2 are set to 45 and 25, respectively.
Figures 4.4 and 4.5 show the magnitude and phase distributions for the 900 and 3600 number of dipole cases along with the COMSOL solution. Our observations show that the 900 dipole case does not capture the perturbation around the objects, while the 3600 case does. Additionally, when adequate numbers of dipoles within the forward model zone are used, the 2D DDA and COMSOL results are comparable.

4.2.2 Direct and iterative solvers overview

In general, systems of equations are represented in the form of

$$\bar{A} \bar{x} = \bar{b}$$

(4.14)

where $\bar{A}$ is the given coefficient matrix, $\bar{x}$ is the unknown vector, and $\bar{b}$ is also known. A linear system of equations can be solved via direct or iterative solvers. The direct solver refers to methods such that the solutions are obtained in one step. One way to obtain $\bar{x}$ is to invert the matrix $\bar{A}$ and multiply the right hand side, $\bar{b}$, by it. This approach is problematic in certain conditions including:

- The matrix $\bar{A}$ is not invertible,
- The matrix $\bar{A}$ is not square,
- The matrix $\bar{A}$ is singular or close to singular.

Even if the matrix $\bar{A}$ is invertible, calculation of its inversion matrix, $\bar{A}^{-1}$, is usually very expensive, especially for large numbers of unknowns. In such cases, other direct solvers are employed. Common methods to directly calculate the solution include Gaussian elimination, LU decomposition, Cholesky, and QR decomposition. The details of these algorithms are given in reference [75]. Direct solvers are not practical if the matrix size is large. For example, if the matrix is of size $n$, Gaussian elimination has complexity of $O(n^3)$. Consequently, for systems with large numbers of unknowns, iterative solutions are often favored instead of direct solutions. The iterative solutions are not obtained in one step, they start by assuming an approximate solution for the unknown $\bar{x}$. It requires multiple intermediate-steps to determine the "exact" solution to the system. The Krylov subspace based methods such as the conjugate gradient (CG) algorithms and its variants are popular for solving linear systems of equations due to their computational efficiency.

If the matrix $\bar{A}$ in (4.14) is positive definite, the system of equations can be solved with the CG method. For the cases where these properties are not satisfied, other CG-type methods are typically used. The body of iterative algorithms consists of three main parts: initialization, inside loop, and iteration check. The inside loop reveals that the matrix-vector multiplication is the most computationally expensive operation in an iterative method. Therefore, it requires substantial attention to restrain the computation time.
Figure 4.3: Diagrams of the 2D forward model zone utilizing (top) 900, and (bottom) 3600 dipoles, respectively.
Figure 4.4: Plots of magnitude (top) and phase (bottom) distributions for $f=1.0$ GHz, $\sigma = 1$, $\epsilon_{r,b} = 22$, $\epsilon_{r,t1} = 25$, $\epsilon_{r,t2} = 45$, $d_1 = 3$ (cm), $d_2 = 6$ (cm).

Figure 4.5: Plots of magnitude (top) and phase (bottom) distributions for $f=1.0$ GHz, $\sigma = 1$, $\epsilon_{r,b} = 22$, $\epsilon_{r,t1} = 45$, $\epsilon_{r,t2} = 25$, $d_1 = 3$ (cm), $d_2 = 6$ (cm).
The following system of equations is used to calculate the polarization $\tilde{P}$ [59]:

$$\tilde{E}_{\text{inc}}(\tilde{r}_i) = \frac{\tilde{P}_i}{\alpha_i} - \sum_{j=1 \atop j\neq i}^{N} \tilde{G}_{ij} \tilde{P}_j$$  \hspace{1cm} (4.15)

The system of equations (4.15) can be reduced to the following master relationship [73]:

$$\alpha_i \tilde{E}_{\text{inc}}(\tilde{r}_i) = \tilde{P}_i - \alpha_i \sum_{j \neq i} \tilde{G}_{ij} \tilde{P}_j$$  \hspace{1cm} (4.16)

This can be rearranged in matrix format to produce:

$$\begin{bmatrix}
1 & -\alpha_1 G_{12} & -\alpha_1 G_{13} & \ldots & -\alpha_1 G_{1N} \\
-\alpha_2 G_{21} & 1 & -\alpha_2 G_{23} & \ldots & -\alpha_2 G_{2N} \\
-\alpha_3 G_{31} & -\alpha_3 G_{32} & 1 & \ldots & -\alpha_3 G_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\alpha_N G_{N1} & -\alpha_N G_{N2} & -\alpha_N G_{N3} & \ldots & 1
\end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_N \end{bmatrix} = \begin{bmatrix} \alpha_1 E_{\text{inc}}(r_1) \\ \alpha_2 E_{\text{inc}}(r_2) \\ \alpha_3 E_{\text{inc}}(r_3) \\ \vdots \\ \alpha_N E_{\text{inc}}(r_N) \end{bmatrix}$$  \hspace{1cm} (4.17)

Here $r_i$ is the spatial location of each dipole. This is a general square matrix which can be solved with any number of direct or iterative solvers. In this form, the computation time is not substantially less than other numerical solvers such as FDTD or FE schemes. However, iterative solvers are known to be more time efficient compared to direct solvers. The set of equations in (4.15) can also be rearranged in matrix format as

$$\begin{bmatrix}
\frac{1}{\alpha_1} & -G_{12} & -G_{13} & \ldots & -G_{1N} \\
-G_{21} & \frac{1}{\alpha_2} & -G_{23} & \ldots & -G_{2N} \\
-G_{31} & -G_{32} & \frac{1}{\alpha_3} & \ldots & -G_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-G_{N1} & -G_{N2} & -G_{N3} & \ldots & \frac{1}{\alpha_N}
\end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_N \end{bmatrix} = \begin{bmatrix} E_{\text{inc}}(r_1) \\ E_{\text{inc}}(r_2) \\ E_{\text{inc}}(r_3) \\ \vdots \\ E_{\text{inc}}(r_N) \end{bmatrix}$$  \hspace{1cm} (4.18)

where $\tilde{A} = (\text{diag}(\frac{1}{\alpha}) - \tilde{G})$. As was mentioned, the matrix-vector multiplication is the most computationally expensive operation in inside loop calculations for iterative solvers such as conjugate gradient (CG) based approaches. This matrix-vector multiplication is typically operated over many iterations until a converged solution.
Chapter 4. The Discrete Dipole Approximation

is obtained. To increase the computational efficiency, the matrix $\bar{A}_{ij}$ is broken into two parts - the diagonal and off diagonal terms, respectively such that:

$$\bar{A}_{ij} \bar{P}_j = \begin{bmatrix}
\frac{1}{\alpha_1} & 0 & 0 & \ldots & 0 \\
0 & \frac{1}{\alpha_2} & 0 & \ldots & 0 \\
0 & 0 & \frac{1}{\alpha_3} & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & \frac{1}{\alpha_N}
\end{bmatrix}
\begin{bmatrix}
P_1 \\
P_2 \\
P_3 \\
\vdots \\
P_N
\end{bmatrix} +
\begin{bmatrix}
0 & -G_{12} & -G_{13} & \ldots & -G_{1N} \\
-G_{21} & 0 & -G_{23} & \ldots & -G_{2N} \\
-G_{31} & -G_{32} & 0 & \ldots & -G_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-G_{N1} & -G_{N2} & -G_{N3} & \ldots & 0
\end{bmatrix}
\begin{bmatrix}
P_1 \\
P_2 \\
P_3 \\
\vdots \\
P_N
\end{bmatrix} \quad (4.19)$$

The first matrix-vector operation on the right hand side can be performed pointwise as a vector-vector multiplication. The second matrix-vector operation on the right hand side is the dominant operation in computation time for the iterative CG-type algorithms. The key to handling this matrix-vector operation efficiently lies in the matrix properties of $\bar{G}$. The matrix $\bar{G}$ is complex and symmetric and more detail on these formulations and the implementations is given in next section.

4.2.4 Krylov subspace methods for the DDA

Krylov subspace algorithms are used to solve large scale systems of equations. These methods are usually named as conjugate gradient based algorithms. Moreover, they have been involved in solving linear system of equations raised in various applications including those from the electromagnetic field simulations [76–78]. These methods usually handle sparse matrices (meaning most elements are zero) often having narrow bandwidth. Particularly, discretization of partial differential equations (PDEs) with the FE method or FD methods leads to these conditions. The most commonly used method from Krylov subspace algorithms is the conjugate gradient (CG) algorithm. The iterative CG algorithm is employed for real, symmetric, and positive definite systems, meaning that $\bar{A}^T \bar{y} \bar{A} > 0$ for an arbitrary vector $\bar{y}$. Here $T$ indicates the transpose of a matrix. The idea behind iterative CG type algorithms is to form an (orthogonal) basis for the Krylov subspace.

We introduced the DDA to solve the forward problem in microwave tomography. For this application, we are ultimately required to solve a system of linear equations (4.18) considered as

$$\bar{A} \bar{x} = \bar{b}, \quad \bar{A} = (a_{ij}) \in \mathbb{C}^{n \times n}, \quad \bar{x}, \bar{b} \in \mathbb{C}^n \quad (4.20)$$

which is complex symmetric (non-Hermitian), meaning that $\bar{A} = \bar{A}^T \neq \bar{A}^H$. For such system of equations, different iterative solvers such as the quasi-minimal residual (QMR) [79], the BiCG-type methods such as (SCBiCG) [76], the conjugate orthogonal conjugate gradient (COCG) [77], and the conjugate orthogonal conjugate residual (COCR) [78] have been previously employed. Regarding computational efficiency, the COCG and COCR algorithms are preferred to other types since they require one
matrix-vector multiplication per iteration and thus impacts the overall computation time. The principles of these algorithms are similar to those of the CG algorithms. The main difference is in constructing a useful basis for the \( j \)-dimensional Krylov subspace \( K_j(\tilde{A}; \tilde{r}_0) \) for a linear system. \( \tilde{r}_0 = \tilde{b} - \tilde{A}\tilde{x}_0 \) is the initial residual for the initial guess to the solution, \( \tilde{x}_0 \). Details on the COCG and COCR methods are given in Algorithms 1 and 2 [77,78].

\[
x \leftarrow x_0; \\
r \leftarrow b - Ax_0; \\
\rho \leftarrow (r, r^i); \\
p \leftarrow r^i; \\
\text{while } ||r|| > TOL \text{ do} \\
\quad q \leftarrow Ap; \\
\quad \mu \leftarrow (p, q^t); \\
\quad \text{if } \mu = 0 \text{ then} \\
\quad \quad \text{Quit (failure)} \\
\quad \text{end} \\
\quad \alpha \leftarrow \frac{\rho}{\mu}; \\
\quad x \leftarrow x + \alpha p; \\
\quad r \leftarrow r - \alpha q; \\
\quad \text{if } x \text{ is accurate enough then} \\
\quad \quad \text{Quit (Convergence)}; \\
\quad \text{else} \\
\quad \quad \beta \leftarrow \rho; \\
\quad \quad \rho \leftarrow (r, r^i); \\
\quad \quad \beta \leftarrow \frac{\rho}{\beta}; \\
\quad \quad p = r^i + \beta p; \\
\quad \text{end} \\
\text{end} \\
\]

**Algorithm 1:** The COCG algorithm

\[
x \leftarrow x_0; \\
r \leftarrow b - Ax_0; \\
\rho \leftarrow (r, r^i); \\
p \leftarrow r^i; \\
q \leftarrow Ap; \\
\text{while } ||r|| > TOL \text{ do} \\
\quad \mu \leftarrow (q, q^t); \\
\quad \text{if } \mu = 0 \text{ then} \\
\quad \quad \text{Quit (failure)} \\
\quad \text{end} \\
\quad \alpha \leftarrow \frac{\rho}{\mu}; \\
\quad x \leftarrow x + \alpha p; \\
\quad r \leftarrow r - \alpha q; \\
\quad r^i \leftarrow r^i - \alpha q^t; \\
\quad t \leftarrow Ar^i; \\
\quad \text{if } x \text{ is accurate enough then} \\
\quad \quad \text{Quit (Convergence)}; \\
\quad \text{else} \\
\quad \quad \beta \leftarrow \rho; \\
\quad \quad \rho \leftarrow (r, r^i); \\
\quad \quad \beta \leftarrow \frac{\rho}{\beta}; \\
\quad \quad p = r^i + \beta p; \\
\quad \quad q = t + \beta q; \\
\quad \text{end} \\
\text{end} \\
\]

**Algorithm 2:** The COCR algorithm

In Algorithms 1 and 2, the most prominent operations regarding computational aspects are matrix-vector multiplications \( \tilde{A} \times \tilde{p} \) and \( \tilde{A} \times \tilde{r}^i \), respectively. These operations can be performed in the form of relationship (4.19). Generally, the operator \( \tilde{G} \times \tilde{P} \) (matrix-vector multiplication) has computational cost of order \( O(N^2) \). The matrix-vector multiplications can be done efficiently under specific circumstances. One approach is to employ the fast Fourier transformation (FFT). Details on this topic are given in following section.
4.2.5 Matrix-vector multiplication using the FFT

The matrix-vector multiplication can be done fast for some structured matrices such as Toeplitz and circulant matrices [80]. These matrices can be multiplied by vectors in $O(n \log n)$ computational cost, instead of $O(n^2)$. A Toeplitz matrix $\tilde{T}$ is unique such that it has regular repetition of the individual coefficients, but they are not necessarily symmetric:

$$
\tilde{T} = \begin{pmatrix}
a & b & c & d & e \\
f & a & b & c & d \\
g & f & a & b & c \\
h & g & f & a & b \\
i & h & g & f & a 
\end{pmatrix}
$$

(4.21)

A circulant matrix $\tilde{C}$ is a special Toeplitz matrix of the form

$$
\tilde{C} = \begin{pmatrix}
a & b & c & d & e \\
e & a & b & c & d \\
d & e & a & b & c \\
c & d & e & a & b \\
b & c & d & e & a 
\end{pmatrix}
$$

(4.22)

The unique property of a circulant matrix is that it can be represented by the values in a single row. With respect to computational efficiency, there is a substantial savings in terms of memory requirements. The matrix $\tilde{G}$ in (4.19) is symmetric block-Toeplitz, meaning that each of the $N$ portions of the matrix in size of $n \times n$ (where $n = \sqrt{N}$) are themselves Toeplitz matrices (Figure 4.6). The Toeplitz submatrices are symmetric and can be converted to circulant matrices. To accomplish this, columns and rows for the Toeplitz submatrices are padded appropriately to produce circulant matrices in the form of:
Figure 4.6: Block-Toeplitz matrix $\tilde{G}$. 
Chapter 4. The Discrete Dipole Approximation

The remaining circulant matrices $\tilde{C}_i$, where $2 \leq i \leq N$ are also calculated in a similar way. Consequently, the system of equations with coefficient matrix $\tilde{G}$ is converted to a large circulant linear system.

Multiplication of a vector solution $\tilde{x} = (x_0, x_1, \ldots, x_n)$ by circulant matrices $\tilde{C}$, is equivalent to an operation called a circular convolution. According to the convolution theorem, this convolution is performed using the $n$-point Fourier transformation, and it means that the Fourier transform of a convolution is the point-wise product of Fourier transforms. This indicates that we can multiply $\tilde{C} \times \tilde{x}$ for any circulant matrix $\tilde{C}$ in $O(n \log n)$ operations. This implies that to multiply $\tilde{C} \times \tilde{x}$, we just need to:

- Employ the fast Fourier transform (FFT) of the first column of the matrix and the vector,
- Multiply the two together pointwise,
- Take the inverse FFT to recover the result.

Subsequently, the matrix-vector multiplication with $O(N^2)$ computational complexity is reduced to an $O(N \log N)$ operation. With this formulation, the computation time is reduced dramatically and it significantly improves to overall computation time efficiency.

The computational aspects of the 2D DDA is studied in Paper B. Here we present some samples for the computation times. The computation times for the direct and iterative solvers are summarized in Table 4.1. For comparison, the computation times for both MATLAB, as an interpretive language, and C++, a classic compiler-based code, implementations are investigated. The results indicate that it is useful to compare performance both with MATLAB and C++ code. While the interpretive
4.2. DDA and the 2D forward problem

codes struggle computational time-wise with constructs such as loops, they contain highly optimized matrix operations which can often overcome such disadvantages. These examinations show that the computation time for the 2D DDA is significantly decreased in the iterative approach combined with FFT. The best performance is achieved in C++ using an open source C++ package for fast Fourier transform calculations denoted as FFTW packages [81].
Table 4.1: Computation time (s)

<table>
<thead>
<tr>
<th>#dipoles</th>
<th>Direct</th>
<th>COCG</th>
<th>COCR</th>
<th>COCG-FFT</th>
<th>COCR-FFT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MATLAB</td>
<td>C++ (Armadillo)</td>
<td>MATLAB</td>
<td>C++ (Armadillo)</td>
<td>MATLAB</td>
</tr>
<tr>
<td>900</td>
<td>0.0255</td>
<td>0.3780</td>
<td>0.01825</td>
<td>0.0317</td>
<td>0.0229</td>
</tr>
<tr>
<td>3600</td>
<td>0.9378</td>
<td>22.9960</td>
<td>1.1122</td>
<td>1.1027</td>
<td>1.3329</td>
</tr>
<tr>
<td>4225</td>
<td>1.5337</td>
<td>32.2780</td>
<td>1.4979</td>
<td>1.5455</td>
<td>1.6131</td>
</tr>
</tbody>
</table>
Summary of Papers

5.1 Paper A

The focus of this paper is on assessing the validity of the 2D DDA for computing the 2D electric field distributions in the microwave imaging systems. In this paper, the novel concept of using the 2D DDA as the forward solver of the imaging systems is investigated. The 2D electric field distribution for the microwave imaging system is numerically simulated for a simplified breast tumor model. The primary motivation is that 2D DDA has the potential to decrease the computation time, which has been intensively explored in Paper B.

Paper A starts with the modeling and formulations of the 2D DDA for electric field distributions. Next, the formula for adequate sampling size and number of dipoles are proposed based on the available sampling size formulae for the 3D case. Three different parameters and their contributions to the accuracy of the 2D DDA are studied. The three parameters include:

- Number of dipoles
- Inclusion size
- Inclusion property contrast

The aim is to explore the impact of these parameters on solution accuracy. To accomplish that, the COMSOL Multiphysics software is used for baseline results to benchmark the DDA simulations. Additionally, the simulation results are compared to measurement data. These comparisons show that the DDA, FE, and measurements show comparable results regarding both magnitude and phases. Our observations illustrate that 2D DDA is accurate for relatively large objects with high contrast permittivity properties when employing the adequate number of dipoles representing the geometry. In this way, the 2D DDA can be used as a forward solver in reconstruction.
algorithms for microwave imaging to replace FDTD or FE forward solvers previously implemented.

5.2 Paper B

In this paper, we focus on the computational aspects of the 2D DDA. We explore the 2D DDA formulations and the possibilities for integrating these formulations into the microwave imaging systems while retaining accuracy with a robust algorithm concerning the polarization terms.

Paper B starts with the 2D DDA formulations in detail. Next, the system of equations provided by the 2D DDA is investigated. In this case, the iterative conjugate orthogonal conjugate gradient (COCG) method is used for solving the forward problem where the time-limiting step is the matrix-vector multiplication per iteration. For the sake of reduced computation time, the matrix-vector multiplication step is formulated such that the Block-Toeplitz matrix is multiplied by the vector solutions. The reason to do that is to exploit the reduced computation costs for the block-Toeplitz matrices times a vector operation. The idea is to expand each Toeplitz matrix into a circulant matrix for which the convolution theorem is applied for fast computation utilizing the fast Fourier transform (FFT).

The results demonstrate that this formulation is accurate and efficient. Comparing the computation times for the direct solvers, the iterative solvers (COCG and COCR), and the iterative solver using the fast Fourier transformation (COCG-FFT and COCR-FFT), indicates that this approach significantly decreases the computation time. For the purpose of further algorithm optimization, we study the grid size and analysis of appropriate error tolerance levels for the conjugate gradient type iterative process. The results indicate that employing the iterative solver with appropriate error tolerance level combined with the FFT on the suitable grid size allow the forward solution time to decrease to the order of 0.1 second.
Conclusion and Future Work

In this thesis, validity of using the 2D DDA to calculate the electric field distributions for the microwave imaging systems is studied. Our investigations show that 2D DDA can be used as a viable and accurate forward solver. The accuracy of the proposed method is dependant on the number of dipoles. In Paper A, we have also proposed a formula for the minimum number of dipoles, which can be considered as a modified version of the 3D case applied to a microwave imaging system. The results indicate that this method can also handle a large contrast range within the forward model zone. The larger targets with high contrast contribute to a higher error compared to smaller targets with lower contrast.

The main motive to employ the 2D DDA is to reduce the computational complexity in the microwave tomography algorithms, and produce images in few minutes instead of multiple hours. To accomplish this, in Paper B, the linear system of equations from the 2D DDA is solved using Krylov-based iterative algorithms combined with the FFT. The proposed solver is time efficient for cases where the forward model zone is finely discretized and size of it is chosen appropriately. Additionally, choice of an interpretive language such as MATLAB or a classic compiler-based code such as C++ impacts time efficiency. Although the 2D DDA have been studied and investigated, there is a possibility to improve the accuracy, and it can be done by investigating the dipole polarizability for the microwave imaging systems.

We would like to employ the 2D DDA in a microwave tomography algorithm. Our plan is to build a new, 2D tomographic imaging algorithm utilizing the 2D DDA forward solver as a fundamental building block. Within that construct, we will examine how best to structure the forward solution grid and the parameter mesh (or grid) as a means of optimizing the computation time required to form the Jacobian matrix, a key to reducing computation time.

In addition, because the forward solutions only change marginally from one reconstruction iteration to the next, there may be ways to optimize the forward solution time further since the CG-based algorithms require a starting estimate for the
forward solution. In this context, instead of supplying the forward solution for a homogeneous medium, the integrated reconstruction algorithm may be able to supply better estimates of the forward solution. This may substantially reduce the number of CG iterations required for an accurate solution and significantly reduce the overall computation time further.

Besides, study of the convergence rate for the COCG and COCR can also contribute to time efficiency. The 2D DDA based reconstruction algorithm will be benchmarked with the FE and FDTD based algorithms developed at Dartmouth and Chalmers groups. To accomplish this, experimental data such as measured data from phantoms can be used. In addition, the dipole polarizability for the microwave imaging systems should be studied, as it substantially impacts the accuracy with respect to target’s characteristics.
Bibliography


Bibliography


diffraction tomography,” IEEE transactions on microwave theory and techniques,

for biomedical applications: Critical and prospective analysis of reconstruction

[46] T. Rydholm, A. Fhager, M. Persson, and P. M. Meaney, “A first evaluation of
the realistic supelec-breast phantom,” IEEE Journal of Electromagnetics, RF

plastic of the realistic geeps-l2s-breast phantom,” Diagnostics, vol. 8, no. 3, p. 61,
2018.

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, vol. 43, no. 9, pp. 869–877, 7 1996.

[49] E. Fear and M. Stuchly, “Microwave detection of breast cancer,” IEEE Trans-
actions on Microwave Theory and Techniques, vol. 48, no. 11, pp. 1854–1863,
2000.

[50] J. Hadamard, Lectures on Cauchy’s problem in linear partial differential equa-


totype development of a microwave tomography system for breast cancer detec-

[53] A. Taflove and S. C. Hagness, Computational electrodynamics: the finite-

[54] A. Fhager, Microwave tomography. PhD Dissertation, Chalmers University of


50