Efficient Domain Decomposition Methods for the Electromagnetic Analysis of Dielectric Objects

By

Bijilash Babu

A dissertation submitted in fulfillment of the
requirements for the award of
Doctor of Philosophy
to the

Dublin City University

Faculty of Engineering and Computing

School of Electronic Engineering

Supervisor: Dr. Marissa Condon

September 2010

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Bijilash Babu

ID No: 55156657

Dated: 22 September 2010

To My Parents

Table of Contents

Abstract

Even with the advent of cutting-edge computing facilities on desktop computers, electromagnetic scattering simulations of moderately sized problems still faces major challenges. This is primarily due to the finer resolution requirement for lower wavelengths in the scatterer. Domain Decomposition (DD) methods are an apparent candidate to alleviate this difficulty. In this dissertation, the primary focus will be on the use of DD methods for the solution of Electric Field Integral Equations for dielectric objects.

Novel variants of two classes of DD methods are investigated in this thesis; namely, the Buffered Block Forward Backward Methods (BBFB) and the Characteristic Basis Function Methods (CBFM). BBFB methods solve the discretised equations for finding the unknown field in a semi-iterative block-wise manner and march towards a complete solution. The BBFB method and its variants investigated in this work are confirmed not to be suitable for dielectric objects with higher permittivity and complex shapes. Some of the difficulties faced by the BBFB methods are due to the propagation of spurious edge effects appearing near the artificial boundaries that the BBFB method creates to decompose the object.

The original CBFM is a Macro-domain Basis Function (MBF) based method which is free from some of the shortcomings of the FBM. Detailed investigation of the method shows that the CBFM can be tuned easily for the efficient analysis of electromagnetic scattering problems of dielectric objects. A novel modification of the CBFM is investigated and successfully implemented. This latest version of the CBFM is free from all the shortcomings that were reported for the other DD methods we came across. In principle, any large EM scattering problem can be solved on a personal computer using this novel method.

Acknowledgements

I would like to express my sincere thanks to my supervisor Dr. Marissa Condon for her advise and support throughout the course of this work. My thanks to Dr. Conor Brennan for his advise on several occasions. My sincere thanks to Prof. Raj Mittra, Dr. Kai Du and Mr. Kadappan Panayappan of the Electromagnetic Communication Laboratory, Pennsylvania State University, for their support.

My heartfelt thanks to Achan, Amma and Bakku for their unconditional support and affection. So many people have been of constant help throughout my term at DCU. Without their support this work would not have been so successful. My heartfelt thanks to Aldo, Anoop, Daniel, Dileep, Gerry, Joe, Lolo, Nijil, Rejish, Sabine, Selin, Shihab, Tao and Yasar for their support. Thanks to Prof. P. Vivekanandan, Prof. P. Pradeep and Prof. P. N. Bhasi for introducing me to the subject. I would like to thank Prof. N. D. Hari Dass, Prof. P. G. Vaidya, Prof. V. S. Murty, Prof. S. V. Sudheer and Prof. G. Rangarajan for their encouragement to pursue a research career.

I extend my thanks to the technical support team at the school, staff at the Registry and Invent for their support on several occasions.

Notations

Mathematical Statement	Descriptive Statement
I	Identity matrix, $i_{mm} = 1$ and $i_{mn} = 0$
$*$	Indicates complex conjugation
$[A]^T = [A]$	Symmetric
$[A] = [A^*]^T$	Hermitian (self-adjoint)
$[A] = -[A^*]^T$	Skew Hermitian
$[A]^{-1} = [A^*]^T$	Unitary
$[A^T][A] = [I]$	Orthogonal
$[A^*]^T[A] = [A][A^*]^T$	Normal
$A_{mn} > 0$	Positive-definite
$A_{mn} \geq 0$	Nonnegative or positive semidefinite
$A_{mn} = 0$ for $m \neq n$	Diagonal
$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of complex numbers
$\mathbb{R}^{m \times n}$	Vector space of real matrices
$\mathbb{C}^{m \times n}$	Vector space of complex matrices
$\mathcal{O}(n)$	Complexity (n) of an algorithm
$\texttt{span}\{a_1, a_2, \ldots, a_n\}$	Subspace spanned by a_j
$\mathcal{K}_n = \texttt{span}\{b, Ab, \ldots, A^{n-1}b\}$	n-dimensional Krylov subspace spanned by A

Operator type	Eigenvalue (λ) Properties
Hermitian	Real eigenvalues
Unitary	Eigenvalues on unit circle
Skew Hermitian	Eigenvalues on imaginary axis
Positive semidefinite	Eigenvalues ≥ 0
Positive-definite	Eigenvalues > 0

Symbols

Symbol	Descriptive Statement	Unit
ϵ_0	free space permittivity	8.854×10^{-12} farads/meter (F/m)
μ_0	free space permeability	$4\pi \times 10^{-7}$ henrys/meter (H/m)
σ	electric current conductivity	mhos/m $(\frac{1}{\Omega}/m)$
σ_m	magnetic current conductivity	ohms/m (Ω/m)
ρ_e	electric charge density	Coulombs per cubic meter (C/m^3)
ρ_m	magnetic charge density	Webers per cubic meter (Wb/m^3)
ϵ	permittivity	Farad per meter (F/m)
μ	permeability	Henries per meter (H/m)
λ	wavelength	meter (m)
k	wavenumber	reciprocal meter (m^{-1})
f	frequency	Hertz (Hz)
ω	angular frequency	$2\pi f$, radians per second (rads/sec)
λ_0	free space wavelength	reciprocal meter (m^{-1})
$\mathbf{E}$	electric field intensity	Volts per meter $(V/m$ $)$
$\mathbf{H}$	magnetic field intensity	Amperes per meter $(A/m$ $)$
$\mathbf{B}$	magnetic flux density	Webers per square meter (Wb/m^2)
$\mathbf{D}$	electric flux density	Coulombs per square meter (C/m^2)
$\mathbf{K}$	source magnetic current density	Volts per square meter (V/m^2)
$\mathbf{J}$	source electric current density	Amperes per square meter (A/m^2)
ϵ_r	medium's relative permittivity constant	
μ_r	medium's relative permeability constant	

Chapter 1

Introduction

Electromagnetics (EM) is the study of the propagation and interaction of electromagnetic fields in matter. Whether it is in home appliances, medical imaging equipment, communication devices, military applications like radar, or computers, EM find applications everywhere. Knowledge of the EM phenomena is necessary during the design stage of the aforementioned applications. Electromagnetic Theory (EMT) is the basic underlying theory that guides a designer when computationally modelling prototypes of these appliances. Electromagnetic Compatibility (EMC) is the branch of EM that studies the unintentional generation, propagation and reception of electromagnetic energy producing unwanted effects. During the design of such appliances, computation of the EM fields in and around the device is necessary.

The basic theory and equations governing the propagation and interaction of electromagnetic waves have been developed a century ago [35, 23, 71]. The number of cases for which classical solutions exist is limited. Many practical problems that need electromagnetic analysis are analytically intractable using available analytic solution methods [16, 41, 29]. Often, the only option left is numerical solutions with the aid of computers [63, 24, 44]. The branch of computational science where such numerical modelling is done is known as Computational Electromagnetics (CEM). The advent

of ever-increasing computing power has accelerated the CEM research. Recent advances in computer hardware and numerical algorithms have significantly improved the possibility of numerical investigations of several problems, where analytical methods are not applicable. However, as the application domain becomes computationally expensive or memory-wise impractical, the computational approach gets into difficulty. CEM techniques can be broadly classified into differential equation (DE) and integral equation (IE) methods of various forms [44, 66, 63]. Methods in the former category consist of the Finite Element Methods (FEM) and the Finite Difference Time Domain (FDTD) methods [24, 72, 68, 63, 37]. In this dissertation, we devote our attention only to the IE based approach.

Our primary focus in this work is to develop efficient IE based CEM methods for the EM analysis of dielectric objects. The problem of electromagnetic scattering from inhomogeneous dielectric objects can be formulated in terms of an Electric Field Integral Equation (EFIE) over the domain of the object. The first method for solving the EFIE over a dielectric object of arbitrary cross-section was developed by Richmond [49, 50]. Before Richmonds's work, only approximate solutions were available for dielectric objects with arbitrary cross section. In this article, Richmond uses the Method of Moments (MoM) with pulse expansion functions and point matching. The dielectric object is discretised into small cells. The total electric field intensity within each cell is initially considered to be an unknown quantity. The fields inside each cell are considered uniform. A system of linear equations is obtained by enforcing at the center of each cell the condition that the total field must be equal to the sum of the incident and scattered fields. This system of equations is solved numerically to compute the electric field intensity in each cells.

The requirement of a large number of basis functions in the Method of Moments (MoM) solution method has limited the use of the IE-MoM modelling approach to

smaller objects in terms of wavelength [49, 44]. Typical basis functions, such as the Rao-Wilton-Glisson (RWG) [48] or rooftops [3], do not incorporate any phase variation, so they are not able to model the behavior of the induced surface currents on large domains [31]. In addition, these basis functions are defined over flat surfaces, and are not suitable for modeling currents on large arbitrary shaped surfaces, except by using a subdomain approach [31]. As the object size increases in terms of wavelength, the size of the associated MoM matrix becomes prohibitively large. Entire domain basis functions [3] have not been successful and they tend to make the resulting matrices highly ill-conditioned. One class of techniques to alleviate this problem is to use domain decomposition (DD) method and to define Macro Basis Functions (MBF) [62] on decomposed domains. The strategy is to divide-and-conquer the large scale problem into a number of small sub-problems that are computationally manageable and have a physical interpretation.

A number of techniques have been developed recently to enhance the MoM analysis of large electromagnetic problems. To handle large problems, one common strategy is to use asymptotic methods like the Geometrical Theory of Diffraction (GTD) [28] or the Physical Optics (PO) [67] solutions. These methods are geometry dependent and not valid for all objects with fine structures. A recently prposed method for the MoM analysis of large objects is known as the Fast Multipole Method (FMM) [10, 55]. The FMM's extension, known as the Multilevel Fast Multipole Algorithm (MLFMA), saves memory by storing only the near-field interaction part of the large matrix. These methods are mostly suitable for far-field computations. Another recently proposed method known as the IES^3 [26] is very efficient for certain types of problems, especially for far-field computations. Computation of near-field quantities requires better methods. Most of these methods still require iterative methods for their solution. It would be desirable to avoid the use of iterative methods as they would sometimes

diverge or take too many iterations to produce reasonably accurate results. One idea is to compress the MoM matrix some how so that direct matrix solution methods [14] can be brought in for their solution. The Macro Basis Function (MBF) and a recursive technique called the Subdomain Multilevel Approach (SMA) were recently developed for large antenna arrays [61, 22]. These methods analyse portions of the object separately and compute macro-basis for each domain. The above approach does not directly account for the mutual coupling of different portions. Usage of this sort of method must be investigated for electromagnetic scattering problems taking care of the mutual coupling of various portions of the object. Though these methods have been vastly implemented for various problems [61, 46, 45, 8, 11, 38, 13], EM scattering problems for dielectric objects were not thoroughly investigated.

1.1 Domain decomposition methods

Domain decomposition has different meanings in different contexts [54].

- In parallel computing, DD often denotes the process of distributing data belonging to different regions of the problem to different processors in the computing machinery.
- In computational sciences, DD means the separation of physical domain into regions that can be modelled with different equations (in the most trivial case, only matrix indices differs) with interfaces between the domain handled by various conditions.

Throughout this thesis, DD methods with various names will denote either or both of these terminologies. The attractive feature of DD methods are [54]:

- They can handle problems of any size.

- They are easy to parallelise.
- Simplification of complicated geometry.
- Superior convergence properties.

The widespread availability of parallel computers with distributed and shared memory offers endless potential for the numerical solution of large problems in computational sciences. In the CEM community, the migration from serial to parallel algorithms is comparatively slower than in other branches of computational sciences. Domain decomposition methods are intrinsically suitable for parallel computation. However, they are not frequently implemented for CEM problems. In this thesis, an effort is made to investigate the computational efficiency of some DD methods with an intention to implement them for shared memory parallel processing. Shared memory systems with several processors on board are becoming a very common member in computational science laboratories. This facilitates efficient implementation of DD based methods for CEM simulations.

1.2 DD methods used in this dissertation

A class of *physically inspired iterative solutions*, known as the *current marching* method is a DD method that was successfully implemented for perfect electrical conductors. Examples include the *method of ordered interactions* [25] and the *Forward Backward Method* (FBM) [20, 70, 69]. These methods solve scattering problems in portions and use direct matrix inversion methods for the solution of MoM submatrices. The Buffered Block Forward Backward (BBFB) Method was implemented to investigate the EM wave scattering by infinite dielectric cylinders [7]. In the BBFB method, basis functions that belong sub-regions are grouped together, and buffers represent a set of adjacent basis functions that belong to neighbouring sub-blocks. The

introduction of the buffer regions allow the suppression of spurious edge effects at the boundaries of decomposition. These FB methods work well when applied to problems involving objects that do not have the possibility of multiple scattering. These methods are not efficient for complex geometrical objects. The FB methods converge at a much faster rate compared to iterative techniques when applied to problems of scattering from simple structures. Our goal in this work is to develop a DD method that would efficiently handle dielectric objects irrespective of its shape and electrical properties. After a thorough investigation, the FBM was found not to be suitable for implementing a standalone software for the EM analysis of dielectric objects.

In contrast to the conventional IE-MoM technique and the FEM which uses subdomain basis functions, the Characteristic Basis Function Method (CBFM) [46, 21, 42, 64, 60, 31, 12, 33, 43] defines a new kind of macro-domain basis function. These MBFs are constructed from traditional low-level basis functions that depend on the geometry of the problem. Consequently, they are automatically adapted to the shape of the body and are also capable of modeling the phase variations in large domains [31]. Once these Characteristic Basis Functions (CBFs) have been generated, a reduced matrix is computed using these CBFs. The reduced matrix system can be computed by evaluating the matrix-vector products involving the sub-blocks of the conventional MoM matrix and their corresponding CBFs. The reduced matrix is so called because it is a compressed version of the conventional MoM matrix. It is usually much smaller in size and, hence, is amenable to direct solvers. This avoids the need for iterative methods. Even though the problem is split into isolated pieces, the algorithm takes into account mutual coupling at a global level.

1.3 Novel elements in the dissertation

- Investigation of an improved FBM for dielectric objects.
- Implementation of the CBFM for dielectric objects.
- A novel version of the CBFM tailored for dielectric objects.
- Efficient implementation of the CBFM on desktop computers.

1.4 Outline of the dissertation

This dissertation is organised as follows. In Chapter 2, fundamental EM field equations are presented and the mathematical framework is reviewed and summarised in the context of this thesis. In Chapter 3, the electric field integral equation is formulated and subsequently discretised using the method of moments. Numerical results are validated with available analytical results. In Chapter 4, linear algebraic methods that are used in various contexts in the dissertation are summarised. In Chapter 5, some novel modifications of the forward backward method are implemented. The CBFM is implemented for the EM analysis of dielectric objects in Chapter 6. In Chapter 7, a novel modification of the CBFM especially tailored for dielectric objects is developed and implemented for electrically large objects.

1.5 Notation

Unless otherwise specified, vectors are distinguished from scalars using bold typeface. For example, $\mathbf{r} \in \mathbb{R}^3$ denotes position in three-dimensional space. The three components are given by $\mathbf{r} = (x, y, z)^T$, here T denotes transpose. The electric field vector is denoted $\mathbf{E}$ while the scalar magnetic potential is denoted A.

Chapter 2

Basic Equations for CEM

In general, all scattering problems are formulated using fundamental classical electrodynamics equations. In Section (2.1) of this chapter, the equations used for modelling the propagation of electromagnetic waves will be reviewed. Section (2.4) summarises the advantages of Integral Equation (IE) methods over Differential Equation methods. Throughout this thesis, the focus will be on the Integral Equation methods. A formulation leading to the integral equation approach of the scattering theory is illustrated in Section (2.4.2).

2.1 Governing Equations

This section is a brief review of the classical mathematical models for the propagation of electromagnetic waves. The most important set of equations in electromagnetics were formulated by James Clerk Maxwell in 1864 [35], known as *Maxwell's Equations*. These equations consist of two pairs of coupled partial differential equations relating six fields. These equations are not sufficient to uniquely determine the electromagnetic field. Additional conditions are required to model the interaction of an EM wave with a target.

2.1.1 Maxwell's Equations

In Maxwell's formulation, the electromagnetic field in a scattering domain is described by five vectors, which are functions of position $\mathbf{r} \in \mathbb{R}^3$. Maxwell's equations relate these five vectors in a compact form:

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{2.1.1}$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \tag{2.1.2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.1.3}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{2.1.4}$$

In Equations (2.1.1)- (2.1.4), the vector fields $\mathbf{E}, \mathbf{H}, \mathbf{D}, \mathbf{B}$ and $\mathbf{J}$ are

$\mathbf{E}$, the electric field intensity (volts/meter)

$\mathbf{H}$, magnetic field intensity (amps/meter)

$\mathbf{D}$, Electric flux density (coulombs/meter^2)

$\mathbf{B}$, Magnetic flux density (webers/meter^2)

$\mathbf{J}$, Electric current density (amps/meter^2)

$\mathbf{M}$, Magnetic current density (volts/meter^2)

ρ, Electric charge density (coulombs/meter^3)

$\rho_{\mathbf{m}}$, Magnetic charge density (webers/meter^3)

Note that:

coloumb = amp.second

weber = volt.second

ohm = volt/amp

The scalar quantity ρ is the electric charge density.

The EM field is created by a distribution of static electric charges and a directed flow of electric charge. Equation (2.1.1) is termed *Faraday's law*. This gives the effect of the changing magnetic field on the electric field. Equation (2.1.2) is called *Ampère's*

law which was modified by Maxwell to this form. Equation (2.1.3) shows why free magnetic poles do not exist. Equation (2.1.4) is called *Gauss's law*. This gives the effect of charge density on electric displacement. In an isotropic media, where the electrical properties are independent of the direction of propagation of the EM wave:

$$\mathbf{D} = \epsilon\mathbf{E} \tag{2.1.5}$$

$$\mathbf{B} = \mu\mathbf{H} \tag{2.1.6}$$

$$\mathbf{J} = \sigma\mathbf{E} \tag{2.1.7}$$

The scalar quantities ϵ, μ and σ are the electrical permittivity, magnetic permeability and electrical conductivity respectively. Note that, $\epsilon = \epsilon_0\epsilon_r$. Substitution of Equations (2.1.5)-(2.1.7) into (2.1.1)-(2.1.4) results in:

$$\nabla \times \mathbf{B} - \mu_0\epsilon_0\frac{\partial \mathbf{E}}{\partial t} = \mu_0\mathbf{J} \tag{2.1.8}$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{2.1.9}$$

Also, implicit in the Maxwell equations is the continuity equation for charge density and current density:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \tag{2.1.10}$$

This follows from combining the divergence of Equation (2.1.2) and time derivative of Equation (2.1.4). The speed of light in vacuum, given by $c = \frac{1}{\sqrt{(\mu_0\epsilon_0)}}$, is an essential quantity in electromagnetics. Another important equation worth mentioning here is the *Lorentz force equation* arising from the motion of charged particles:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{2.1.11}$$

This gives the force acting on a point charge q in the presence of electromagnetic fields. The electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, in Maxwell's equations and elsewhere were actually introduced by means of the Lorentz force equation. In Coulomb's

experiments [71], forces acting between localised charge distributions were observed. Therefore, it was useful to introduce $\mathbf{E}$ as *force per unit charge*. Similarly, in Ampère's experiments [71], the mutual forces on current carrying loops were studied [23]. The current I in a loop of cross sectional area A with N charge carriers per unit volume moving at a velocity $\mathbf{v}$ is given by $I = NAq\mathbf{v}$. $\mathbf{B}$ in Maxwell's Equations is defined in magnitude as the force per unit current.

Consider the case where the source currents vary sinusoidally in time. If the EM radiation has a temporal frequency $\omega > 0$, the field is said to be time-harmonic. The time dependent Equations (2.1.1)-(2.1.4) can be rewritten as time-harmonic Maxwell's Equations. Time-harmonic electromagnetic waves of the form

$$\mathbf{E}(z,t) = E_0 e^{j(kz-wt)} \tag{2.1.12}$$

$$\mathbf{H}(z,t) = H_0 e^{j(kz-wt)} \tag{2.1.13}$$

satisfy the time-harmonic Maxwell equations.

$$\nabla \times \mathbf{E} + jk\mathbf{H} = 0 \tag{2.1.14}$$

$$\nabla \times \mathbf{H} - jk\mathbf{E} = 0 \tag{2.1.15}$$

$$\nabla \cdot (\mu\mathbf{H}) = 0 \tag{2.1.16}$$

$$\nabla \cdot (\epsilon\mathbf{E}) = \rho \tag{2.1.17}$$

In Equations (2.1.12) and (2.1.13), E_0 and B_0 are the complex amplitudes, ω is the angular frequency. The physical fields are the real parts of $\mathbf{E}$ and $\mathbf{B}$. k is the wave number given by:

$$k^2 = (\epsilon + \frac{j\sigma}{\omega})\mu\omega^2 \tag{2.1.18}$$

In Equations (2.1.14)-(2.1.17), $\mathbf{E}$ and $\mathbf{H}$ are complex-valued phasors, representing the vector amplitude and phase angle of electromagnetic waves.

2.1.2 Boundary Conditions

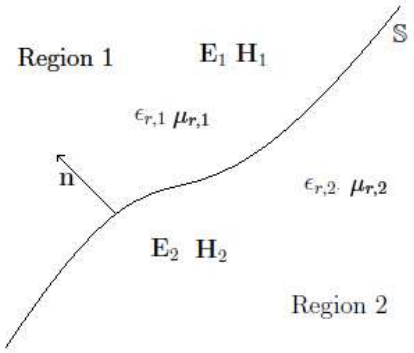

Figure 2.1: Two electrically different media separated by surface $\mathbb{S}$

In general, electromagnetic fields are discontinuous at a boundary between two different media, or at a surface that carries charge density ρ or surface current density $\mathbf{K}$ [3, 16, 44]. The Maxwell equations must be supplemented with suitable boundary conditions. Consider the interface between two different media denoted as Region 1 and Region 2 as in Figure 2.1. $\mathbf{n}$ is unit vector normal to the surface of separation denoted as $\mathbb{S}$. The tangential component of the electric field must be continuous across $\mathbb{S}$, i.e., $\mathbf{n} \times \mathbf{E}$ is continuous across $\mathbb{S}$ [41]. If $\mathbf{E}_1$ and $\mathbf{E}_2$ denote the limiting value of the electric fields in Region 1 and 2 as $\mathbb{S}$ is approached, we must have:

$$\mathbf{n} \times (\mathbf{E}_1 - \mathbf{E}_2) = 0 \text{ on } \mathbb{S} \tag{2.1.19}$$

At the same time, the normal components of $\mu_r \mathbf{H}$ are continuous across $\mathbb{S}$ so that:

$$\mathbf{n} \cdot (\mu_{r,1}\mathbf{H}_1 - \mu_{r,2}\mathbf{H}_2) = 0 \text{ on } \mathbb{S} \tag{2.1.20}$$

In general,

$$\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_{\mathbb{S}} \tag{2.1.21}$$

where $\mathbf{J}_{\mathbb{S}}$ is the surface current density on $\mathbb{S}$. In most cases, the magnetic field has continuous tangential components, (i.e, $J_{\mathbb{S}} = 0$). An exception to this is when $\mathbb{S}$ is a thin conductive layer. This gives rise to the *conductive boundary condition*. But when $\mathbf{J}_{\mathbb{S}} = 0$, usually Equation (2.1.21) is generalised as:

$$\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) = 0 \text{ on } \mathbb{S} \tag{2.1.22}$$

The presence of singularities in the charge density ρ may cause jumps in the normal component of $\epsilon_r \mathbf{E}$.

$$\mathbf{n} \cdot (\epsilon_{r,1}\mathbf{E}_1 - \epsilon_{r,2}\mathbf{E}_2) = \rho_{\mathbb{S}} \text{ on } \mathbb{S} \tag{2.1.23}$$

where, $\rho_{\mathbb{S}}$ is the surface charge density. Thus it is clear that the electric and magnetic fields are not always continuous if ϵ_r or μ_r are discontinuous across a boundary. Numerical discretisation schemes for Maxwell's equations in the presence of a discontinuity must take into account that although, the tangential components of the fields are continuous, the normal components have a jump at the boundary.

If the second medium is a perfect conductor, the tangential component of $\mathbf{E}^t$ and tangential component of $\mathbf{H}^t$ vanish at the interface. This gives rise to the *perfect conductor boundary condition* on the interface. According to *Ohm's law* [41], if conductivity $\sigma \to \infty$ and the current density $\mathbf{J}$ is to remain bounded, then $\mathbf{E} \to 0$. This suggests that inside a perfect conductor the electric field vanishes. If Region 1 is a perfect conductor, then $\mathbf{E}_1 = 0$ and the *perfect conductor boundary condition* for $\mathbf{E}_1$ is:

$$\mathbf{n} \times \mathbf{E}_1 = 0 \text{ on } \mathbb{S} \tag{2.1.24}$$

When the second interface is not a perfect electric conductor (PEC), but allows electromagnetic waves to penetrate to some extent, then the *impedance boundary condition* is given by:

$$\mathbf{n} \times \mathbf{H}_1 - \kappa(\mathbf{n} \times \mathbf{E}_1) \times \mathbf{n} = 0 \tag{2.1.25}$$

where, impedance κ is a suitably chosen positive function of position on the surface of the material.

2.2 Various CEM Methods

To solve Maxwell's equations numerically, the usual procedure is to discretise the scattering domain into finite regions. There are several methodologies employed in CEM modelling. Selection of a particular method for any specific problem depends on various facts such as the available computational resources, the required accuracy, the scattering domain's electrical parameters, size and so on. The following sub-sections give an overview of the most widely employed methodologies.

2.2.1 Differential Equation (DE) methods

DE methods evolve from the direct discretisation of Maxwell's equations. Derivatives are approximated using finite differences and the final system to be solved is a linear equation. Various DE methods are the finite difference method, finite element method and its various forms. They are local in nature and the computational domain includes additional regions of space outside the scatterer [66]. This adds to the computational burden and memory requirements. However, the favourable properties of differential formulations are the ease of implementation and resultant sparse matrices [24, 66]. Development of parallel software for DE methods is relatively easy.

2.2.2 Variational Methods

Variational methods are based on the variational formulation of Maxwell's Equations [24]. Examples include method of moments (MoM), finite element methods, finite volume methods, etc. They offer many algorithmic advantages like ease of

implementation, parallelisation, involvement of sparse matrices etc. The main disadvantage is that they require a large computational domain. This results in large memory requirements and high computational costs.

2.2.3 Integral Equation methods

The Integral Equation (IE) formulation represents the entire physics of the problem in a very compact form [66]. This is because the integral equation describes the physical global features, including the boundary conditions. In IE formulations, unlike DE methods, only active regions are modelled. That is, the IE formulation is exact, even though the discretisation schemes used for the computational solution of IEs introduce a numerical nature to the overall formulation. The power of IEs lie in their ease of use for arbitrary geometries and excitations. Allthough IEs can be applied to any frequency range, there are two main practical limitations [27] in the high frequency region. The interaction matrix that needs to be inverted becomes very large. This is a major issue and in this thesis we will be focussing on methods to circumvent these problems. Secondly, the interactions between different regions decrease as the frequency increases and eventually becomes very much local.

2.2.4 Asymptotic Methods

Unlike the other methods in Section 2.2, asymptotic models do not solve the full Maxwell model but an approximation to the Maxwell model. These methods are very efficient since they do not solve for the whole field. However, they are asymptotic in nature and thus not error-controllable and give rise to significant error for large frequencies.

2.2.5 Some final comments on the above methods

All of the aforementioned schemes are successfully employed in a variety of CEM problems. One or more of these methods may be suitable for certain specific problems. The suitability of any method is decided after analysing the solution requirements. However, in this thesis, the focus shall be on Integral Equation Methods.

2.3 Integral Equations

An integral equation is an equation involving an integral of an unknown function $\phi(\mathbf{r})$. Integral equations are classified into two major groups depending on the integration domain. If the integration domain depends on a variable, the equation is termed a *Volterra* integral equation. If the integration domain is fixed, it is termed a *Fredholm* integral equation [66]. Integral equations can be further classified. If the unknown function $\phi(\mathbf{r})$ appears only in-side the integral, it is a Fredholm integral equation of the *first kind*. If the unknown function appears both inside and outside the integral, it is a Fredholm integral equation of the *second kind*. Throughout our study, we deal with integral equations where the integration domain is fixed. Thus, we will only be concerned with Fredholm integral equations. Fredholm integral equations of the first and second kind are defined by:

$$\int_{\Omega} \mathbf{K}(\mathbf{r}, \mathbf{r}')\phi(\mathbf{r}')d\Omega = f(\mathbf{r}),\ \mathbf{r}' \in \Omega \tag{2.3.1}$$

and

$$\phi(\mathbf{r}) - \lambda \int_{\Omega} \mathbf{K}(\mathbf{r}, \mathbf{r}')\phi(\mathbf{r}')d\Omega = f(\mathbf{r}),\ \mathbf{r}' \in \Omega \tag{2.3.2}$$

respectively. Here, $\mathbf{K}$ is a function of two variables, known as the kernel. $f(\mathbf{r})$ is a known function, often termed as data function. Note that in the second kind, the unknown function $\phi(\mathbf{r})$ appears both inside and outside the integral.

2.4 Why Integral Equation Methods?

DE Methods are easy to implement, and are extensively used in EM modelling [24]. However, they are constrained to use structured grids even for complex structured bodies. Therefore, one has to discretise regions outside the object, and this adds to the memory burden [24].

When using DE methods in scattering simulations, one has to supply boundary conditions. This is possible with *Dirichlet to Neumann* maps [19]. However, the resulting boundary conditions are not local. These boundary conditions are, in general, known as artificial boundary conditions. A type of artificial boundary, that requires the scattered field to vanish at the artificial boundary is often used [19]. However, this causes significant reflections from the artificial boundary and this affects the results, unless the artificial boundary is placed away from the scatterer. A solution to this problem is the introduction of an absorbing medium known as *perfectly matched layer* [5].

Integral equations counter many of the DE difficulties. The boundary conditions are explicitly enforced in IEs. In IE formulations, unlike DE methods, the discretised domains contain only the active regions. However, IE methods result in highly dense matrices. Inversion of a dense matrix of order N by a direct solver is an $\mathcal{O}(N^3)$ operation [14]. While an iterative method can reduce this to $\mathcal{O}(N^2)$ [51], this requirement again puts a constraint on the problem size that can be computationally solved. In the following chapters, we address some recently developed methods which use domain decomposition techniques. These methods avoid the requirement of storing the whole matrix. In this thesis, we report some novel improvements in the above mentioned domain decomposition methods. In particular, it will be shown that these improved methods can be applied successfully to dielectric objects modelled using the Electric Field Integral Equation (EFIE).

2.4.1 Volume Equivalence Principle

The *Volume Equivalence Principle* [44] helps to simplify the formulation of integral equations. The dielectric or magnetic material is replaced by equivalent induced polarisation current and charges. The equations in Section 2.1 can be used for formulating integro-differential equations for modelling the interaction of an EM wave with penetrable dielectric scatterers. Maxwell equations (2.1.14)-(2.1.17)can be rewritten as:

$$\nabla \times \mathbf{E} = -\jmath\omega\mu_0\mathbf{H} - \mathbf{K} \tag{2.4.1}$$

$$\nabla \times \mathbf{H} = \jmath\omega\epsilon_0\mathbf{E} + \mathbf{J} \tag{2.4.2}$$

$$\nabla \cdot (\epsilon_0\mathbf{E}) = \rho_e \tag{2.4.3}$$

$$\nabla \cdot (\mu_o\mathbf{H}) = \rho_m \tag{2.4.4}$$

where,

$$\mathbf{K} = \jmath\omega\mu_0(\mu_r - 1)\mathbf{H} \tag{2.4.5}$$

$$\mathbf{J} = \jmath\omega\epsilon_0(\epsilon_r - 1)\mathbf{E} \tag{2.4.6}$$

$$\rho_e = \epsilon_0\epsilon_r\mathbf{E} \cdot \nabla(\frac{1}{\epsilon_r}) \tag{2.4.7}$$

$$\rho_m = \mu_0\mu_r\mathbf{H} \cdot \nabla(\frac{1}{\mu_r}) \tag{2.4.8}$$

One can think of the source terms $\mathbf{J}$ and $\mathbf{K}$ as replacements for $\mathbf{E}$ and $\mathbf{H}$ in Equations (2.1.14)-(2.1.17). Equations (2.4.1)-(2.4.4) are equivalent to Equations (2.1.14)-(2.1.17). The procedure of replacing the dielectric or magnetic material by induced sources is known as the *volume equivalence principle*. Sources in Equations (2.4.5)-(2.4.8) radiate in free space. The scattering problem can thus be modified with these sources as unknowns. If the scatterer is homogeneous with ϵ_r and μ_r constant throughout, Equations (2.4.7) and (2.4.8) show that there is no induced charge density in those regions. If ϵ_r and μ_r vary continuously and are differentiable, Equation (2.4.7) and (2.4.8) produce an induced volume charge density.

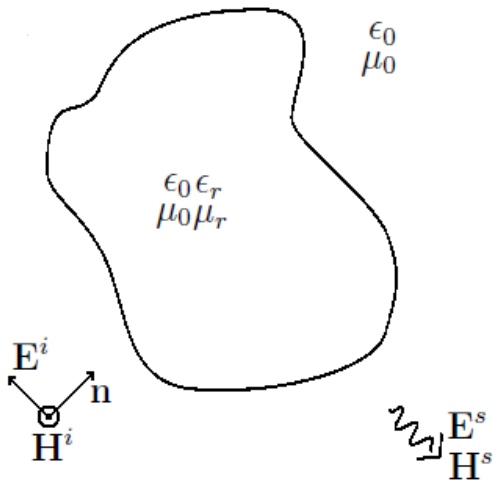

Figure 2.2: An inhomogeneous scatterer illuminated by an incident electromagnetic field

2.4.2 Electric Field Integral Equation (EFIE)

Consider a source free region containing an inhomogeneity characterised by a relative permittivity ϵ_r and a permeability μ_r which are functions of space. When this region is illuminated by an electromagnetic wave, the electromagnetic field in the region must satisfy Maxwell equations (2.1.14)-(2.1.17). The central problem in computational electromagnetics is the calculation of the electromagnetic field inside and near the scatterer. We wish to determine the EM field inside the inhomogeneity and the scattered field around it. The scattered field is determined by the currents induced in the object. These currents are related to the polarisation of the incident fields, the direction of the incident wave and how various parts of the object interact. The later is dependent on the electrical properties of the body

Suppose the scatterer in Figure (2.2) is illuminated by a field produced by a source located outside the scatterer. Then the inhomogeneous scatterer can be replaced by equivalent induced sources radiating in free space [44]. Now we can split the total field into two parts, one associated with the primary source and the other associated

with the induced sources. The fields produced by the primary source are denoted by $\mathbf{E}^i$, the incident electric field, and $\mathbf{H}^i$, the incident magnetic field. The fields due to induced sources are the scattered electric field, $\mathbf{E}^s$, and the scattered magnetic field $\mathbf{H}^s$. The superposition of incident and scattered fields constitute the original fields in the presence of scatterer which can be denoted as below:

$$\mathbf{E}^t = \mathbf{E}^i + \mathbf{E}^s \tag{2.4.9}$$

$$\mathbf{H}^t = \mathbf{H}^i + \mathbf{H}^s \tag{2.4.10}$$

Here, E^t and H^t are the total electric and magnetic fields. The incident field near the scatterer (i.e. at reasonable distance from the source) satisfies the Helmholtz equations.

$$\nabla^2\mathbf{E}^i + k^2\mathbf{E}^i = 0 \tag{2.4.11}$$

$$\nabla^2\mathbf{H}^i + k^2\mathbf{H}^i = 0 \tag{2.4.12}$$

The scattered fields are solutions to the equations:

$$\nabla^2\mathbf{E}^s + k^2\mathbf{E}^s = \jmath\omega\mu_0\mathbf{J} - \frac{\nabla\nabla\cdot\mathbf{J}}{\jmath\omega\epsilon_0} + \nabla\times\mathbf{K} \tag{2.4.13}$$

$$\nabla^2\mathbf{H}^s + k^2\mathbf{H}^s = -\nabla\times\mathbf{J} + \jmath\omega\epsilon_0\mathbf{K} - \frac{\nabla\nabla\cdot\mathbf{K}}{\jmath\omega\mu_0} \tag{2.4.14}$$

Here, $\mathbf{J}$ and $\mathbf{K}$ are the equivalent sources in Equations (2.4.5) and (2.4.6). In order to find the solution of Equation (2.4.13) and (2.4.14), the classical approach is to express the fields in terms of magnetic and electric vector potentials denoted by $\mathbf{A}$ and $\mathbf{F}$, respectively.

$$\mathbf{E}^s = \frac{\nabla\nabla\cdot\mathbf{A} + k^2\mathbf{A}}{\jmath\omega\epsilon_0} - \nabla\times\mathbf{F} \tag{2.4.15}$$

$$\mathbf{H}^s = \frac{\nabla\nabla\cdot\mathbf{F} + k^2\mathbf{F}}{\jmath\omega\mu_0} + \nabla\times\mathbf{A} \tag{2.4.16}$$

On substituting Equations (2.4.15) and (2.4.16) into Maxwell's equations, the vector potentials satisfy:

$$\nabla^2 \mathbf{A} + k^2 \mathbf{A} = -\mathbf{J} \tag{2.4.17}$$

$$\nabla^2 \mathbf{F} + k^2 \mathbf{F} = -\mathbf{K} \tag{2.4.18}$$

A solution to these equations can be written as:

$$\mathbf{A} = \mathbf{J} * G \tag{2.4.19}$$

$$\mathbf{F} = \mathbf{K} * G \tag{2.4.20}$$

where, G is the Green's function given by:

$$G = \frac{e^{-\jmath k|\mathbf{r}|}}{4\pi|\mathbf{r}|} \tag{2.4.21}$$

In equation (2.4.19) and (2.4.20), $*$ denotes 3-dimensional convolution of the form:

$$\mathbf{A}(\mathbf{r}) = \int\int\int \mathbf{J}(\mathbf{r}') \frac{e^{-\jmath k|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \tag{2.4.22}$$

In 2-dimensional problems, the integration over the third dimension involves only the Green's function and this can be performed analytically. Once $\mathbf{J}$ and $\mathbf{K}$ are calculated, $\mathbf{E}^s$ and $\mathbf{H}^s$ are determined using Equations (2.4.15) and (2.4.16). Note that this procedure involves *integration followed by differentiation*. Often, integrals arising from source field relations can not be evaluated in closed form. Numerical quadrature algorithms are used to evaluate these integrals which introduces some error. This error can be further magnified if the derivative of this integral is evaluated using a finite-difference scheme. Such an order of operations is not suitable for numerical implementation [44]. Green's functions can be easily differentiated analytically and then brought into the integral. Thus a change of order gives *differentiation followed by integration*. This reduces the error in numerical implementation of the EFIE.

An alternative to the pure vector potential source-field relations is the mixed potential source-field relation [44].

$$\mathbf{E}^s = -\jmath\omega\mu_0\mathbf{A} - \nabla\Phi_e - \nabla\times\mathbf{F} \tag{2.4.23}$$

$$\mathbf{H}^s = \nabla\times\mathbf{A} - \jmath\omega\epsilon_0\mathbf{F} - \nabla\Phi_m \tag{2.4.24}$$

where, Φ_e and Φ_m are scalar potential functions. which are given by convolutions similar to Equation (2.4.19) and (2.4.20).

In Chapter 3, methods leading to the discretisation of the EFIE is presented. Results obtained are verified using analytically obtained solutions.

Chapter 3

Discretisation of Integral Equations

Continuous integral equations can be discretised and solved for the unknown fields using numerical methods. Several numerical techniques have been developed for the solution of computational electromagnetic problems [16]. The advantage of these computer oriented numerical techniques over the eigenfunction expansion method generated huge interest a few decades ago [16, 52]. The *Method of Moments* (MoM) is one among these numerical methods that is still widely used with great success. The mathematical foundation of the MoM will be reviewed in this chapter and an EFIE will be discretised using it. All discretisation procedures place a limit on the accuracy of the numerical results. Before using them, questions like whether or not a numerical method will converge to exact solution must be addressed. Also, the suitability of selected *basis functions* for a particular problem must be examined. This chapter is intended to investigate, characterise and finally implement an Integral Equation-Method of Moments (IE-MoM) discretisation scheme.

3.1 Mathematical foundation

Tools of *functional analysis*, a branch of mathematics, throw insight into fundamentals and theory behind numerical approximations [41, 15]. These tools characterise and

help us to choose approximation methods that shall be used for the problem at hand.

$$\mathfrak{L}f = g \tag{3.1.1}$$

In Equation (3.1.1), the linear operator $\mathfrak{L}$ maps functions in its *domain* to functions in its *range*. The domain $\mathcal{D}_{\mathfrak{L}}$, and the range $\mathcal{R}_{\mathfrak{L}}$ are different linear spaces. Two functions a and b in the inner product space are said to be *orthogonal* if:

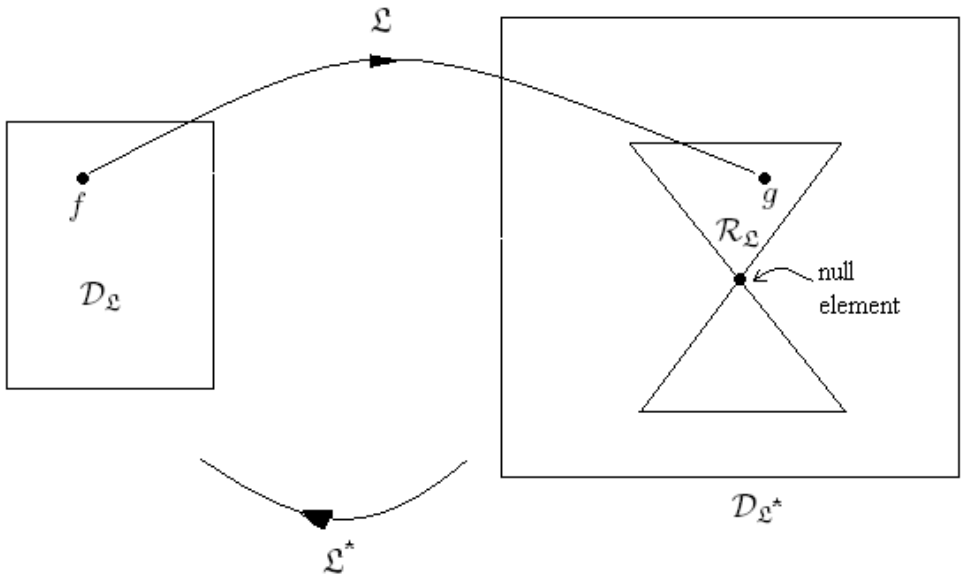

Figure 3.1: Domain and range space of $\mathfrak{L}$ and $\mathfrak{L}^*$

$$\langle a, b \rangle = 0 \tag{3.1.2}$$

Similarly, functions $\{B_n\}$ in an inner product space form an orthogonal set if:

$$\langle B_m, B_n \rangle = 0 \qquad m \neq n \tag{3.1.3}$$

The set $\{B_n\}$ is said to be *complete* if the zero function is the only function in the inner product space orthogonal to each member of the set. A set of functions that is

both orthogonal and complete is said to be a *basis*. A basis can be used to represent any function f in the inner product space such that:

$$\left\| f - \sum_n \alpha_n B_n \right\| = 0 \tag{3.1.4}$$

Here, α_n are scalar coefficients given by:

$$\alpha_n = \frac{\langle B_n, f \rangle}{\langle B_n, B_n \rangle} \tag{3.1.5}$$

The functions of interest are actually projected to a finite-dimensional subspace of the original infinite dimensional inner product space. In the subspace, the basis is truncated to the form $\{B_1, B_2, \ldots, B_N\}$ and the function is represented by:

$$f \cong f^N = \sum_{n=1}^{N} \alpha_n B_n \tag{3.1.6}$$

The scalar coefficients $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ are selected to minimise the distance between f and f^N. The error

$$d\left(f, f^N\right) = \left\| f - f^N \right\| \tag{3.1.7}$$

is minimised when the coefficients are chosen to make the error orthogonal to $\{B_n\}$, ie.,

$$\langle B_n, f - f^N \rangle = 0 \qquad n = 1, 2, \ldots, N \tag{3.1.8}$$

This technique is known as *orthogonal projection* [57]. Due to the orthogonality of the basis functions, the coefficients are the same in the subspace as in the original inner product space. The best approximation for f is obtained when coefficients from Equation (3.1.5) are used. Since f is unknown, α_n cannot be determined directly, a quantity similar to α defined using the range g might be useful. If the set $\{T_n\}$ forms a basis for the range space $\mathcal{R}_{\mathfrak{L}}$, any function in the range may be represented in the N-dimensional subspace spanned by $\{T_1, T_2, \ldots, T_N\}$:

$$g \cong g^N = \sum_{m=1}^{N} \beta_m T_m \tag{3.1.9}$$

The projection that minimises the error $d(g, g^N)$ employs coefficients:

$$\beta_m = \frac{\langle T_m, g \rangle}{\langle T_m, T_m \rangle} \tag{3.1.10}$$

Likewise, the function $\mathfrak{L}B_n$ can be represented by:

$$\mathfrak{L}B_n \cong \sum_{m=1}^{N} l_{mn} T_m \tag{3.1.11}$$

The coefficients that minimise the error $\|\mathfrak{L}B_n - l_{mn}T_m\|$ are given by:

$$l_{mn} = \frac{\langle T_m, \mathfrak{L}B_n \rangle}{\langle T_m, T_m \rangle} \tag{3.1.12}$$

Coefficients in (3.1.12) achieve an orthogonal projection in $\mathcal{R}_{\mathfrak{L}}$ and therefore, give the best approximation. In order to get the approximate solution of $\mathfrak{L}f = g$, the unknown f can be represented in the form (3.1.6). This produces a function on $\mathcal{R}_{\mathfrak{L}}$ of the form:

$$\mathfrak{L}f^N = \sum_{n=1}^{N} \alpha_n \mathfrak{L}B_n \tag{3.1.13}$$

Projecting this function onto $\{T_1, T_2, \ldots, T_N\}$ gives:

$$\mathfrak{L}f^N \cong \sum_{m=1}^{N} \sum_{n=1}^{N} l_{mn} \alpha_n T_m \tag{3.1.14}$$

Coefficients l_{mn} are calculated from (3.1.12). Equating (3.1.14) with (3.1.9) produces a system of equations:

$$\sum_{n=1}^{N} l_{mn} \alpha_n = \beta_m \qquad m = 1, 2, \ldots, N \tag{3.1.15}$$

The $N \times N$ linear equation in (3.1.15) can be solved for the unknowns $\{\alpha_n\}$. The coefficients $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ determined from (3.1.15) are not always the same as that

given by (3.1.5). The fact that projections in $\mathcal{R}_{\mathfrak{L}}$ are orthogonal does not always ensure that the projection in $\mathcal{D}_{\mathfrak{L}}$ is orthogonal and will usually not produce the approximation as measured by the metric $d(f, f^N)$.

3.2 The Method of Moments (MoM)

Following on from the development in Section (3.1), an approximate solution of the linear equation $\mathfrak{L}f = g$ can be written in the form:

$$f \cong \sum_{n=1}^{N} \alpha_n B_n \tag{3.2.1}$$

$\{B_n\}$ are known as the basis functions, defined on the domain $\mathcal{D}_{\mathfrak{L}}$ and the scalars $\{\alpha_n\}$ are unknown coefficients to be determined. Substituting Equation (3.2.1) to $\mathfrak{L}f = g$ and by forcing the residual

$$\mathfrak{L}\left(\sum_{n=1}^{N} \alpha_n B_n\right) - g = \sum_{n=1}^{N} \alpha_n \mathfrak{L} B_n - g \tag{3.2.2}$$

to be orthogonal to a set of testing functions $\{T_1, T_2, \ldots T_N\}$ produces a matrix equation:

$$\mathfrak{L}\alpha = \beta \tag{3.2.3}$$

with entries:

$$l_{mn} = \langle T_m, \mathfrak{L}B_n \rangle \qquad \text{and} \qquad \beta_m = \langle T_m, g \rangle \tag{3.2.4}$$

The matrix equation (3.2.3) can be solved for α using matrix solution algorithms which will be described in Chapter (4). The system of equations (3.2.3) is obtained by forcing the residuals to be orthogonal to the testing functions and hence the method is called the *weighted-residual method*. In electromagnetics, this is popularly known as the Method of Moments. The discretisation of a continuous integral equation by the

MoM involves the projection of the continuous linear operator onto finite dimensional subspaces defined by the basis and testing functions.

In reality, the aforementioned process can be applied to produce an approximate solution even if $\{B_n\}$ and $\{T_n\}$ are not complete and orthogonal [44]. In practice, the basis and testing functions used are often not orthogonal sets [44]. However, even if the basis functions are orthogonal, the projection of the domain space onto the basis functions is not guaranteed to be orthogonal. Thus in the limiting case $N \to \infty$ convergence of the numerical solution to the exact solution cannot be guaranteed and the result obtained will always be approximate (In any case, N will always be finite). The choice of the basis and testing functions is the main issue arising in the implementation of the MoM. Basis functions will be selected depending on the required accuracy, computational resources and on the upper limit on the matrix size.

3.3 Various basis functions

A crucial choice in the numerical computation of the electromagnetic problems using the method of moments is that of the basis functions. A basis set that has the ability to accurately represent the anticipated function must be chosen as the basis functions. However, at the same time, the computation involved in employing this basis set to represent the unknown function must also be considered. A basis function with smoother properties than the unknown being represented must not be selected as the basis function [53]. Even though there are numerous basis sets, only a very limited number of them are used in practice. Basis functions fall into two main classes *sub-domain basis functions* and *entire domain basis functions*

3.3.1 Sub-domain functions

Sub-domain functions are nonzero only over a part of the domain of f. They can be used without the prior knowledge of the nature of the function they must represent. Some of the most commonly used sub-domain basis functions are outlined below.

Pulse basis function

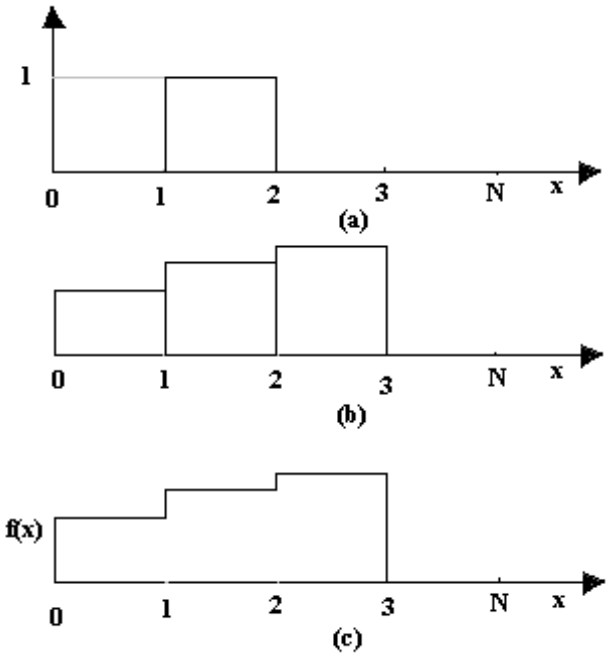

Figure 3.2: Piecewise constant or pulse basis functions: (a) single, (b) multiple, (c) function representation.

The piecewise constant 1-dimensional scalar *pulse basis* function in Figure (3.2) is defined as:

$$f_n(x) = \begin{cases} 1 & x_{n-1} \leq x \leq x_n \\ 0 & \text{elsewhere} \end{cases} \tag{3.3.1}$$

When the associated coefficients are determined, the pulse basis functions will produce a stair case representation of the unknown function.

Triangle function

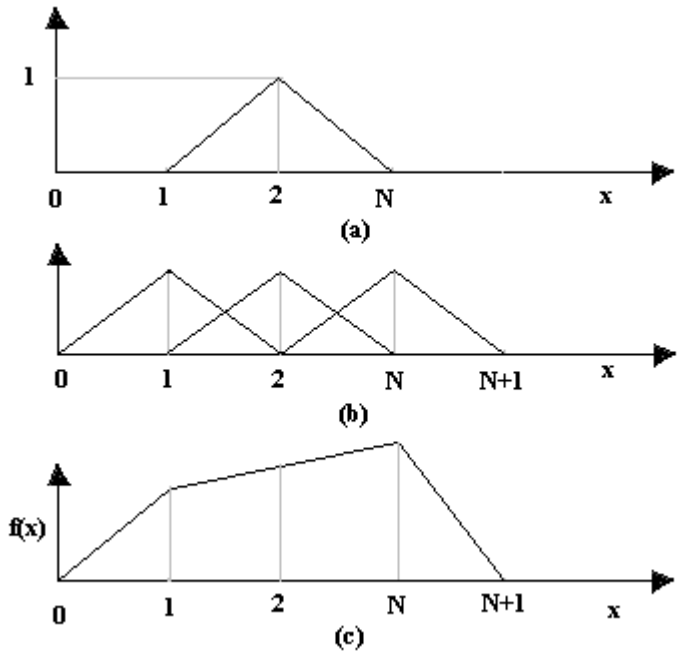

Figure 3.3: Piecewise linear or triangle basis functions: (a) single, (b) multiple, (c) function representation.

The piecewise linear 1-dimensional scalar *triangle basis* function in Figure (3.3) is defined as:

$$f_n(x) = \begin{cases} \dfrac{x - x_{n-1}}{x_n - x_{n-1}} & x_{n-1} \leq x \leq x_n \\ \dfrac{x_{n+1} - x}{x_{n+1} - x_n} & x_n \leq x \leq x_{n+1} \\ 0 & \text{elsewhere} \end{cases} \tag{3.3.2}$$

Triangle functions cover two cells in the discretised domain and adjacent functions

overlap. They provide a smoother representation than that of the pulse basis, also it can go to *zero*. Their usage results in increased computational complexity compared with that of pulse basis [4].

Other basis functions

Increasing the sophistication beyond the level of the triangle function may not guarantee or even help with the improvement of accuracy. In some cases, more specialised functions shall be used as they might help when the integrals involved are multiplied by a $sin(kx)$ or $cos(kx)$ function [3] where x is the variable of integration. Since numerical integration can be avoided, numerical round off errors can be avoided. Significant advantages in error minimisation and computation time can be gained by using basis functions like the *piecewise sinusoid* or *truncated cosines*.

Piecewise sinusoid:

$$f_n(x) = \begin{cases} \dfrac{\sin[k(x - x_{n-1})]}{\sin[k(x_n - x_{n-1})]} & x_{n-1} \leq x \leq x_n \\ \dfrac{\sin[k(x_{n+1} - x)]}{\sin[k(x_{n+1} - x_n)]} & x_n \leq x \leq x_{n+1} \\ 0 & \text{elsewhere} \end{cases} \tag{3.3.3}$$

Truncated cosine:

$$f_n(x) = \begin{cases} cos\left[k\left(x - \dfrac{x_n - x_{n-1}}{2}\right)\right] & x_{n-1} \leq x \leq x_n \\ 0 & \text{elsewhere} \end{cases} \tag{3.3.4}$$

3.3.2 Entire domain functions

Entire domain functions exist over the entire domain of the unknown function and no segmentation is involved in their use. A common entire domain basis set is the set

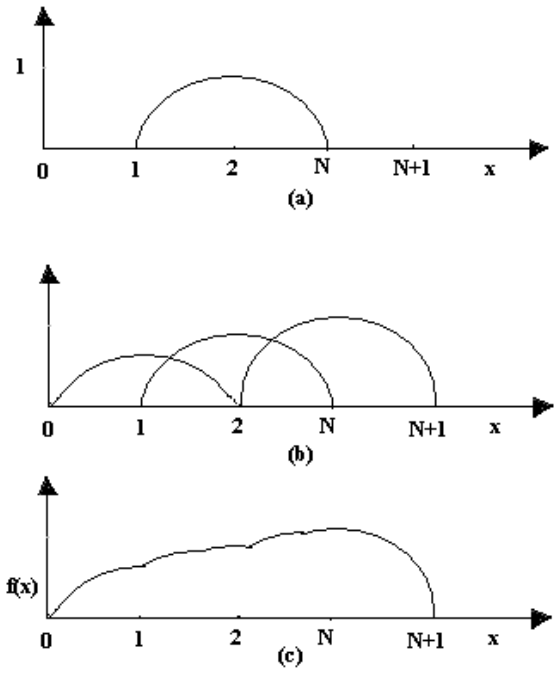

Figure 3.4: Piecewise sinusoids subdomain functions: (a) single, (b) multiple, (c) function representation.

of sinusoidal functions:

$$f_n(x') = cos\left[\frac{(2n-a)\pi x'}{l}\right] \qquad -\frac{l}{2} \leq x' \leq \frac{l}{2} \tag{3.3.5}$$

This basis set is useful for modelling the current distribution on a wire dipole which is known to have a sinusoidal distribution [3]. Entire domain basis functions are used mainly when the unknown quantity to be modelled has a known pattern. Thus one can use fewer number of points to render the assumed approximate pattern of the unknown.

3.4 Galerkin's method

When an identical set of basis functions is used for the basis and testing for the MoM discretisation, the procedure is termed *Galerkin's* method [44]. For Galerkin's

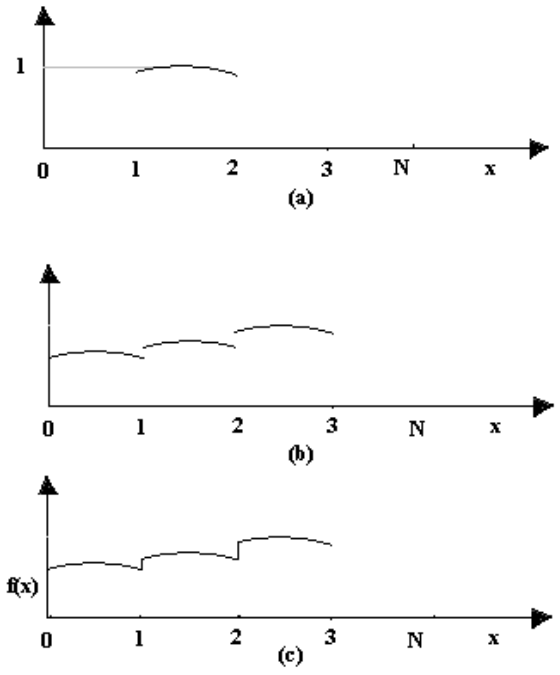

Figure 3.5: Truncated cosines subdomain functions: (a) single, (b) multiple, (c) function representation.

method, the weighting functions are chosen as piecewise continuous functions defined over a subinterval. The assumption is that one can get arbitrarily good accuracy by making the subintervals smaller [52]. For the Galerkin method to be applied, the expansion function must span both the domain and the range of the operator. The procedure is advantageous when the operator $\mathfrak{L}$ is self-adjoint with respect to the inner product. The idea behind the method is as follows. If the operator equation $\mathfrak{L}$ is self adjoint with respect to the inner product, such that:

$$\langle \mathfrak{L}a, b\rangle = \langle a, \mathfrak{L}b\rangle \tag{3.4.1}$$

and the operator is positive definite:

$$\langle a, \mathfrak{L}a\rangle > 0 \quad \text{for all non zero} \quad a \tag{3.4.2}$$

Another inner product can be defined:

$$\langle a, b\rangle_2 = \langle a, \mathfrak{L}b\rangle \tag{3.4.3}$$

where the subscript 2 denotes a 2nd inner product. In this case,

$$\left\langle B_m, \mathfrak{L}(f^N - f)\right\rangle = 0 \tag{3.4.4}$$

In the original inner product space, this can be expressed as:

$$\left\langle B_m, f^N - f\right\rangle_2 = 0 \tag{3.4.5}$$

This means that the projection of the error in the approximation f^N is orthogonal to the basis with this new inner product. This means that the method of moments projection of the unknown function f onto the basis functions is an orthogonal projection and the important consequence of this is that $f^N \to f$ as $N \to \infty$.

Using Galerkin's method for the discretisation of EFIE gives matrices with diagonal symmetry [44]. For a problem discretised using N discretisation points, this requires the generation and storing of only $\frac{N^2}{2}$ elements of the MoM matrix.

3.5 Selection of basis and testing functions

Mathematical aspects defining the proper selection of the basis and testing functions were outlined in Section (3.1). Computational aspects must also be considered when selecting the basis and testing functions. Evaluation of matrix elements may be computationally demanding or analytically impossible for higher-order basis functions and due to the non-availability of closed form solutions. Various basis functions were used in the course of this work. The pulse functions were found to be the most suitable one for the domain decomposition methods used to model dielectric objects analysed in this thesis. Recently, after their successful use in signal-processing, *wavelets* are used as basis functions for representing current in integral equations [44]. Because of

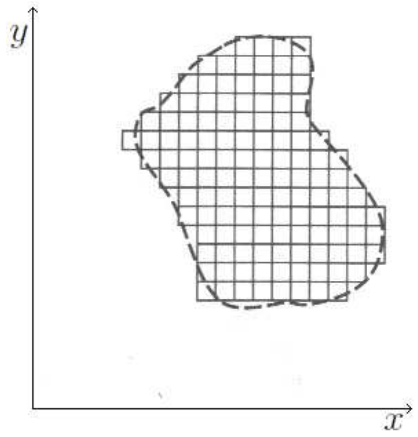

Figure 3.6: The cross-section of a dielectric cylinder divided into N square cells of area Δ_n.

their oscillatory nature, wavelets produce sparse matrices that can be exploited for computational storage and speed up. However, we will not use them.

3.6 Discretisation of the EFIE

Consider a cross section perpendicular to the axis of an infinite dielectric cylinder as shown in Figure (3.6). The relative permittivity ϵ_r is assumed to vary from point to point within the body. A source is radiating a TM plane wave propagating along the $+x$ direction with polarisation parallel to cylinder's axis. This wave produces an incident field which interacts with the infinite cylinder of arbitrary cross-section to produce a scattered field E^s. The field components excited by the TM wave are E_z, H_x and H_y. The total electric field at any point is the sum of incident and scattered fields. Note that, the time dependence $e^{(\jmath\omega t)}$ will be omitted in the formulation.

$$E_z^t(\mathbf{r}) = E_z^i(\mathbf{r}) + E_z^s(\mathbf{r}) \tag{3.6.1}$$

From now on, we drop the z subscript as it is assumed that all electric fields are

polarised in this direction. For a penetrable dielectric, Equation (3.6.1) can be written as [44]:

$$E^i = E^t(\mathbf{r}) + \jmath k\eta A + \nabla\Phi_e + \nabla \times F \tag{3.6.2}$$

where, $\eta = \sqrt{\mu_o/\epsilon_0}$. Equation (3.6.2) is known as the EFIE and holds everywhere throughout the scatterer. Throughout this thesis, we will deal with scatterers composed entirely of dielectric material, so the term involving magnetic current drops out.

$$E^i = E^t(\mathbf{r}) + \jmath k\eta A + \nabla\Phi_e \tag{3.6.3}$$

Using the volume equivalence principle, the dielectric material can be replaced by equivalent polarisation currents:

$$\mathbf{J}(x, y) = \hat{z}\jmath\omega\epsilon_0[\epsilon_r(x, y) - 1]E_z(x, y) \tag{3.6.4}$$

For TM polarisation, the EFIE can be written as:

$$E^i_z(x, y) = \frac{J_z}{\jmath\omega\epsilon_0(\epsilon_r - 1)} + \jmath\omega\mu_0 A_z \tag{3.6.5}$$

where

$$A_z(x, y) = \int\int J_z(x', y')\frac{1}{4j}H^{(2)}_0(kR)dx'dy' \tag{3.6.6}$$

here

$$R = \sqrt{(x - x')^2 + (y - y')^2} \tag{3.6.7}$$

In this formulation, J_z is the primary unknown. The scatterer can be divided into N square cells denoted as Δ_n as in Figure (3.6). The unknown polarisation current density J_z can be approximated by the superposition of pulse basis functions defined on the discretised cells as:

$$p_n(x, y) = \begin{cases} 1 & \text{if } (x, y) \in \Delta_n; \\ 0 & \text{otherwise.} \end{cases}$$

Now, the current density can be written as:

$$J_z(x,y) \cong \sum_{n=1}^{N} j_n p_n(x,y) \tag{3.6.8}$$

Using Equation (3.6.8), Equation (3.6.5) can be written as:

$$E_z^i(x,y) \cong \sum_{n=1}^{N} j_n \left(\frac{\eta p_n(x,y)}{\jmath k[\epsilon_r(x,y)-1]} + \jmath k\eta \int\int_{\Delta_n} \frac{1}{4\jmath} H_0^2(kR) dx'dy' \right) \tag{3.6.9}$$

where, k is the wave-number at any cell, $\eta = \sqrt{\mu_0/\epsilon_r}$ and H_0^2 is the zeroth-order Hankel function of the second kind.

Equation (3.6.9) can be imposed at the cell centres in Figure (3.6) to produce a $N \times N$ matrix equation:

$$\begin{pmatrix} E^i(x_1,y_1) \\ E^i(x_2,y_2) \\ \vdots \\ E^i(x_N,y_N) \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} & \dots & Z_{1N} \\ Z_{21} & Z_{22} & \dots & Z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \dots & Z_{NN} \end{pmatrix} \begin{pmatrix} j_1 \\ j_2 \\ \vdots \\ j_N \end{pmatrix} \tag{3.6.10}$$

Off-diagonal entries in the $N \times N$ matrix are given by:

$$Z_{mn} = \frac{k\eta}{4} \int\int_{\Delta_n} H_0^2(kR_m) dx'dy', \qquad m \neq n \tag{3.6.11}$$

Diagonal entries are given by:

$$Z_{mm} = \frac{\eta}{\jmath k(\epsilon_{r,m}-1)} + \frac{k\eta}{4} \int\int_{\Delta_n} H_0^2(kR_m) dx'dy', m = n \tag{3.6.12}$$

where

$$R_m = \sqrt{(x_m - x')^2 + (y_m - y')^2} \tag{3.6.13}$$

Integrals in Equations (3.6.11) and (3.6.12) can be evaluated analytically [49], assuming that the cell shape is approximately circular but of the same area as the square cells.

$$\int_{\phi'=0}^{2\pi}\int_{\rho'=0}^{0} H_0^2(kR)\rho' d\rho' d\phi' = \begin{cases} \frac{2\pi a}{k} J_0(k\rho)H_1^2(ka) - \frac{\jmath 4}{k^2} & \text{if } \rho < a; \\ \frac{2\pi a}{k} J_1(ka)H_0^2(k\rho) & \text{if } \rho > a; \end{cases}$$

Here, ρ and ϕ are the cylindrical coordinates, a is the radius of the equivalent circle. J_0 and J_1 are the Bessel functions of order zero and one. H_0^2 and H_1^2 are the second kind Hankel functions of order zero and one, respectively. Using the circular cell approximation [49], Equations (3.6.11) and (3.6.12) can be evaluated as:

$$Z_{mn} = \frac{\eta\pi a_n}{2} J_1(ka_n)H_0^2(kR_{mn}), m \neq n \tag{3.6.14}$$

$$Z_{mm} = \frac{\eta\pi a_m}{2} H_1^2(ka_m) - \frac{\jmath\eta\epsilon_{rm}}{k(\epsilon_{rm}-1)} \tag{3.6.15}$$

Equation (3.6.10) can be solved to compute the coefficients j_n. These can be used for calculating the total field E_z^t or other useful quantities like the bi-static scattering cross section given by:

$$\sigma_{TM}(\phi) \cong \frac{k\eta^2}{4}\left|\sum_{n=1}^{N} j_n \frac{2\pi a_n}{k} J_1(ka_n) e^{\jmath k(x_n cos(\phi)+y_n sin(\phi))}\right|^2 \tag{3.6.16}$$

For practical problems, often the size of this matrix equation will be so large that direct inversion is not practical. Iterative solutions are widely used for the solution of such matrices. In the following chapters, we will discuss various Krylov subspace iterative methods and some recently proposed novel methods based on domain decomposition techniques.

3.7 Validation of the discretisation scheme

Before closing this chapter, validation of the results obtained using the discretisation scheme in Section (3.6) is done with an available analytical solution. For this purpose, a homogeneous infinite dielectric cylinder of circular cross-section is analysed.

The radius of the cylinder is $r = 0.5\lambda$ and the relative permittivity $\epsilon_r = 2 - \jmath 0.8$. The incident field is produced by a TM plane wave of frequency 800 MHz propagating along the $+x$ direction with its polarisation parallel to the cylinder's axis. The scattered field E_z^s produced by this field will be computed using an analytical solution method [32]. It is compared to the results from the IE-MoM approach in Section (3.6).

3.7.1 Analytical formulation of dielectric cylinder

A standard analytical method for calculating the scattered field is the classical eigenfunction solution (Mie series) [16, 59]. The Mie series is applicable only to geometries which permits formulation in separable coordinate systems. In this section we review a method [32] for calculating the scattered field from a dielectric cylinder using the eigenfunction expansion of the total field. Consider an incident TM plane wave propagating along the $+x$-direction with its polarisation parallel to the cylinder's axis. It is expressed as:

$$E_z^i = e^{-\jmath k_0 \rho cos\phi} = \sum_{n=-\infty}^{\infty} j^{-n} J_n\left(k_0\rho\right) e^{\jmath n\phi} \tag{3.7.1}$$

Here, k_0 is the wavenumber in free space. The scattered field external to the cylinder (E_{ze}^s) is given by [32]:

$$E_{zd}^s = \sum_{n=-\infty}^{\infty} j^{-n} A_n^{TM} H_n^{(2)}\left(k_0\rho\right) e^{\jmath n\phi} \tag{3.7.2}$$

where,

$$A_n^{TM} = \frac{\frac{\eta_0}{\eta_d} J_n\left(k_0 a\right) J_n^{'}\left(k_d a\right) - J_n^{'}\left(k_0 a\right) J_n\left(k_d a\right)}{J_n\left(k_d a\right) H_n^{(2)'}\left(k_0 a\right) - \frac{\eta_0}{\eta_d} J_n^{'}\left(k_d a\right) H_n^{(2)}\left(k_0 a\right)} \tag{3.7.3}$$

Here, η_o and η_d are the wave impedences in free space and within the dielectric object, respectively. k_d is the wavenumber in the dielectric. The scattered field inside the dielectric is:

$$E_{z_d}^s = \sum_{n=-\infty}^{\infty} j^{-n} B_n^{TM} J_n\left(k_d\rho\right) e^{\jmath n\phi} \tag{3.7.4}$$

where,

$$B_n^{TM} = \frac{-2\jmath/\pi k_0 a}{J_n\left(k_d a\right) H_n^{(2)'}\left(k_0 a\right) - \frac{\eta_0}{\eta_d} J_n'\left(k_d a\right) H_n^{(2)}\left(k_0 a\right)} \tag{3.7.5}$$

3.7.2 IE-MoM formulation of dielectric cylinder

A TM planewave propagating in the $+x$-direction impinges an infinite dielectric object with a circular cross-section. The incident field of frequency 800 MHz is given by:

$$E_z^i = e^{-\jmath k_0 \rho \ cos\phi} \tag{3.7.6}$$

From now on, the z subscript will be avoided as the electric field is always polarised in the z direction. This incident field interacts with the object and produces a scattered field E^s. The scattered field can be expressed as a volume integral convolving the total electric field at each point with the two-dimensional free space Green's function [6]. The resultant integral equation is:

$$E^i\left(\mathbf{r}\right) = E^t\left(\mathbf{r}\right) + E^s\left(\mathbf{r}\right) \tag{3.7.7}$$

where

$$E^s\left(\mathbf{r}\right) = \frac{\jmath}{4}\int_s \chi\left(\mathbf{r}'\right) E_z^t\left(\mathbf{r}\right) H_0^{(2)}\left(k_0\left|\mathbf{r}-\mathbf{r}'\right|\right) ds' \tag{3.7.8}$$

In Equation (3.7.8), $\chi(\mathbf{r})$ is the contrast at $\mathbf{r}$ given by:

$$\chi\left(\mathbf{r}\right) = k^2\left(\mathbf{r}\right) - k_0^2 \tag{3.7.9}$$

k_0 is the background wavenumber and $k(\mathbf{r})$ is the wavenumber at the point $\mathbf{r}$. $H_0^{(2)}$ is the zeroth order Hankel function of the second kind. The electric fields throughout the volume of the scatterer are solved by introducing N pulse basis and testing functions. The pulse basis functions are defined on identical square cells Δ_n and point matching at the cell centres, that is:

$$E^t(\mathbf{r}) = \sum_{n=1}^{N} J_n p_n(\mathbf{r}) \tag{3.7.10}$$

where,

$$p_n(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in \Delta_n \\ 0 & \text{elsewhere} \end{cases} \tag{3.7.11}$$

The Galerkin's procedure results in a dense matrix equation of order N.

$$ZE^t = E^i. \tag{3.7.12}$$

The matrix elements Z_{mn} are given by:

$$Z_{mn} = C_1 J_1(k_0 a) J_1(k_0 R_{mn}) - \jmath H_0(k_0 R_{mn}), \qquad m \neq n \tag{3.7.13}$$

$$Z_{mm} = 1 + \frac{C_1}{k_0} J_1(k_0 a) - \jmath H_1(k_0 a) - \frac{\jmath 4}{k_0^2}, \qquad m = n \tag{3.7.14}$$

$$\text{where,} \quad C_1 = \frac{\jmath 2\pi a \gamma}{4}$$

$$\text{and} \quad \gamma = k^2(n) - k_0^2$$

Here, R_{mn} is the distance between cells m and n, η_0 and k_0 are the impedance and wave-number of the background medium respectively. a is the radius of a circle of equal area ds.

3.7.3 Comparison of the analytical and IE-MoM results

A surface plot of the magnitude of E_z^t on the cross-section at $z = 0$ computed using the analytical solution is shown in Figure (3.7). The MoM matrix in (3.7.12) is ivnverted and the solution is plotted in Figure (3.8). In Figure (3.9), a comparison is made between the above two solutions at $y = 0$. In Figure (3.10) a comparison is made between the above two solutions at $x = 0$. These results confirm that the IE-MoM method will converge to exact results.

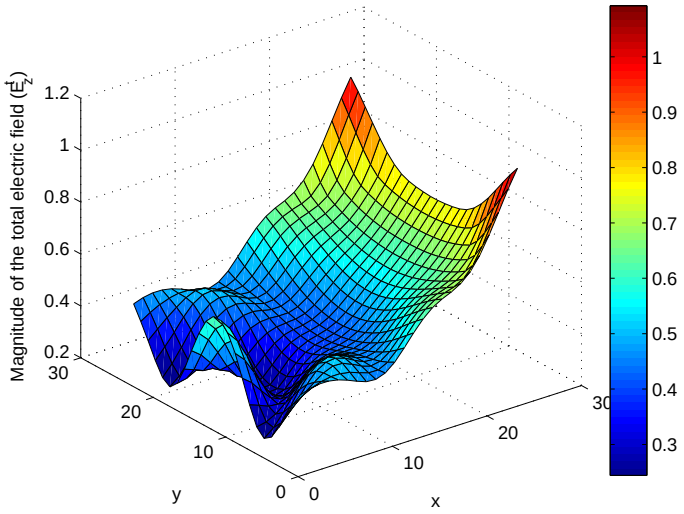

Figure 3.7: Magnitude of the total electric field (E_z^t) computed using the analytical method.

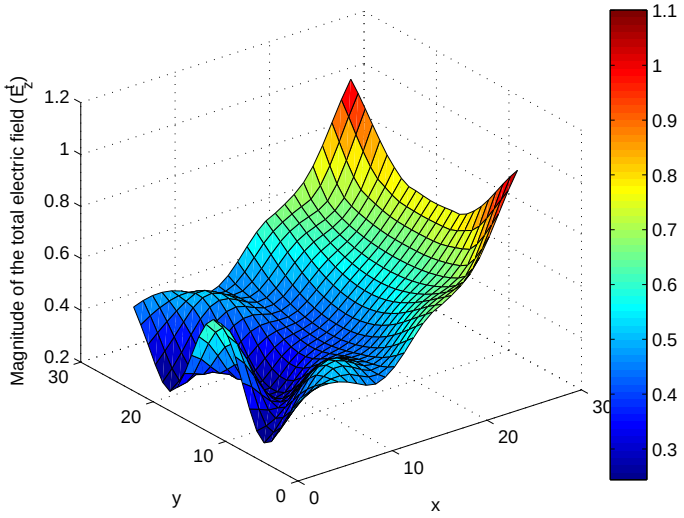

Figure 3.8: Magnitude of the total electric field (E_z^t) computed with the EFIE-MoM method.

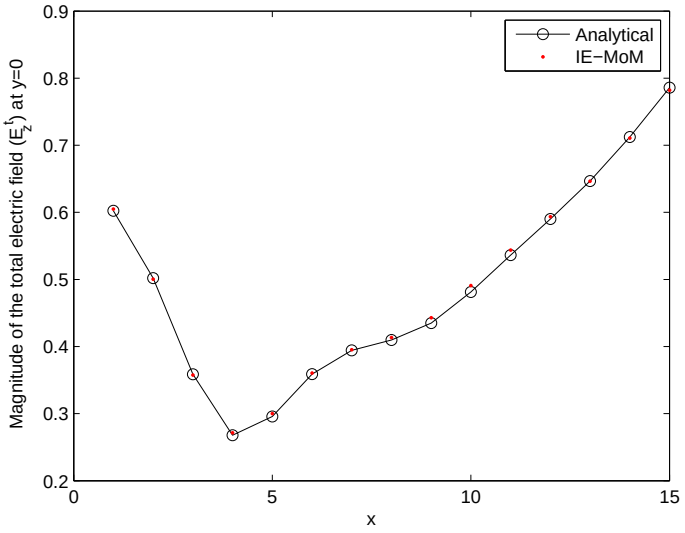

Figure 3.9: Comparison of the magnitude of the total electric field (E_z^t) at $y = 0$ in the object.

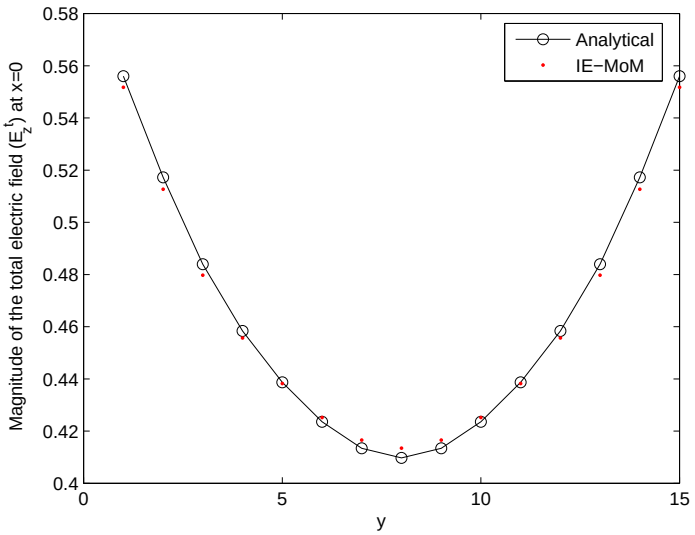

Figure 3.10: Comparison of the magnitude of the total electric field (E_z^t) at $x = 0$ in the object.

3.8 Discretisation requirement

For successful implementation of the IE-MoM approach, the size of the resultant linear equation to be solved depends on the electrical size of the object. To take care of the material-dependent wavelength inside the dielectric, one must be careful about the discretisation size used. For 2-dimensional surface areas modelled using the EFIE, an estimation of the minimum number of discretisation cells needed to discretise the object is given by [44]:

$$N = \frac{A}{\lambda_d^2} D_\lambda^2 \tag{3.8.1}$$

A is the area of the object (m^2), D_λ is the number of discretisation cells used per λ_d where λ_d is the wavelength inside the object, given by:

$$\lambda_d = \frac{\lambda_0}{\sqrt{|\epsilon_r|}} \tag{3.8.2}$$

Here, λ_0 is the free-space wavelength and ϵ_r is the relative permittivity of the dielectric object. Typically, D_λ must be greater than 10 [44]. According to Equation (3.8.1), when the object size or ϵ_r increases, N grows rapidly. For electrically large objects operating at high frequencies, N becomes too large for computations on desktop computers. Domain decomposition methods can be used to circumvent this limitation and to implement EM solvers for small scale computers.

Chapter 4

Numerical methods for CEM

The discretisation of the integral equations governing the scattering problem results in linear equations of the form:

$$Ax = b \tag{4.0.1}$$

of size $N \times N$. From a purely mathematical point of view, determination of x is trivial but, when N is large, the task of computing x may become computationally expensive or even impossible. Often, the main constraint in computing the solution x is the available computational resources. Even with the availability of high-end number-crunching machines and huge parallel and grid computing resources, inversion of large matrices is still a computationally demanding job. The traditional way to solve (4.0.1) is to employ *Gaussian elimination* [14]. This process can be implemented using $\mathcal{O}(N^3)$ floating point operations. The memory requirement is $\mathcal{O}(N^2)$. These methods are not suitable if A is large and sparse. A variant of the Gaussian elimination is known as *Gauss Jordan Method*, this requires $\mathcal{O}(N^3/2)$. For large system of equations this difference will be significant. Anyway all the direct methods require operations of the order $\mathcal{O}(N^3)$.

Another approach to solve (4.0.1) is based on the solution of a nearby linear system, with a matrix that is computationally easy to manipulate both in terms of

storage and number of operations needed. These methods generate a sequence of approximate solutions $x^{(k)}$. Well known examples for methods of this type are Gauss-Seidel, Jacobi method, Successive Over Relaxation(SOR), etc. All iterative methods require the computation of a matrix-vector product which is an $\mathcal{O}(N^2)$ operation. This leads to operations of order $\mathcal{O}(nN^2)$ for n iterations. However, this may be a bottleneck. Therefore, in Section (4.5) a survey of more efficient iterative methods for solving such large matrices will be reviewed. This highly successful class of methods, known as *Krylov subspace* iterative methods, are widely used in CEM.

4.1 Direct methods for solving $Ax = b$

Most text books in numerical linear algebra start with the introduction of these methods as they play an essential role in understanding how the numerical solutions might encounter instabilities if the algorithm has draw backs. A common trend in scientific computing is to break complicated objects into more elementary components. Factoring big matrices into matrices that are easy to deal with helps in solving them easily in parts. Most of the direct methods utilise this idea and create equivalent triangular systems that are easy to solve. The most basic and commonly used direct method is the Gaussian elimination, Gaussian elimination with partial pivoting is the most fundamental algorithm used for the solution of linear algebraic equations [36]. It solves the linear equation in two parts. *1) Forward elimination:* this transforms the the system of equations to a triangular or echelon form [36]. *2) Back substitution:* This part finds the solution of the triangular system. The only method we use from this class of direct methods is the LU decomposition.

4.2 Gaussian elimination

If $Ax = b$ is a nonsingular system, then the Gaussian elimination reduces A to an upper-triangular matrix using elementary row operations. In this process, the algorithm systematically transforms one system into another simpler, but equivalent, system. The elimination process relies on simple operations such as row interchange, addition, subtraction and multiplication. To illustrate the method, let us consider a 3×3 linear equation:

$$\begin{pmatrix} 2 & 1 & 1 \\ 6 & 2 & 1 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 7 \end{pmatrix} \tag{4.2.1}$$

Forward elimination

For convenience, we will adapt a compact notation for the coefficients in $Ax = b$:

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 6 & 2 & 1 & -1 \\ -2 & 2 & 1 & 7 \end{array}\right) \tag{4.2.2}$$

At each step, the strategy is to select one position, called the *pivot position*, and to eliminate all terms below this position in the same column using the aforementioned elementary operations. The coefficient in the pivot position is called a *pivot.* The equation in which the pivot lies is called the pivotal equation. Only nonzero numbers can be taken as pivots. If a coefficient in a pivot position is zero, then the pivotal equation is interchanged with an equation below it to get a nonzero pivot. To start with, if not zero, a_{11} is taken as the first pivot. In (4.2.2), to eliminate all terms below the first pivot 2, we subtract three times the first equation from the second to produce an equivalent system:

$$R_2 - 3R_1 \rightarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ & -1 & -2 & -4 \\ -2 & 2 & 1 & 7 \end{array}\right) \tag{4.2.3}$$

In (4.2.3), add the first equation to the third equation to produce the next equivalent system:

$$R_1 + R_3 \rightarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ & -1 & -2 & -4 \\ & 3 & 2 & 8 \end{array}\right) \tag{4.2.4}$$

Select the next pivot by moving down and to the right (a_{22}). If this coefficient is not 0, then take it as the next pivot. Otherwise, interchange with an equation below this position to bring a nonzero number into a_{22}. In (4.2.4) we can select -1 as the next pivot. Add three times the second equation to the third equation to produce the equivalent system in (4.2.5).

$$R_3 + 3R_2 \rightarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ & -1 & -2 & -4 \\ & & -4 & -4 \end{array}\right) \tag{4.2.5}$$

(4.2.5) is a diagonal system as required. If no zero pivots are encountered, then row interchanges are not necessary, and the reduction can be accomplished by using only elementary row operations. A triangular system can be easily solved using a procedure known as *back substitution*.

Back substitution

Solving the system in (4.2.5) gives:

$$z = \frac{-4}{-4} = 1 \tag{4.2.6}$$

Substituting $z = 1$ back into the second equation in (4.2.5):

$$y = 4 - 2z = 4 - 2(1) = 2 \tag{4.2.7}$$

Substituting $z = 1$ and $y = 2$ back into the first equation in (4.2.5):

$$x = \frac{1}{2}(1 - y - z) = \frac{1}{2}(1 - 2 - 1) = 1 \tag{4.2.8}$$

Now the complete solution is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \tag{4.2.9}$$

If no zero pivots are encountered, then row interchanges are not necessary, and the reduction can be accomplished by using only elementary row operations. Gaussian elimination with back substitution applied to an $N \times N$ system requires:

$$\frac{N^3}{3} + N^2 - \frac{N}{3} \qquad \text{multiplications or divisions} \tag{4.2.10}$$

and

$$\frac{N^3}{3} + \frac{N^2}{2} - \frac{5N}{6} \qquad \text{additions or subtractions} \tag{4.2.11}$$

As N grows, the $\frac{N^3}{3}$ term dominates in both expressions. Gaussian elimination with back substitution on an $N \times N$ system requires operations of the order $\mathcal{O}(N^3)$.

4.3 LU decomposition

If A in (4.0.1) is nonsingular, then $A = LU$ is a product of a lower-triangular matrix L and an upper-triangular matrix U. The process is called an LU factorization of A. Notice that U is the end product of the Gaussian elimination, it has pivots on its diagonal. L has 1's on its diagonal. In L the lower diagonal entries l_{ij} are the multipliers used in the elimination. Once the LU factors of the matrix A have been obtained, it is relatively easy to solve the linear system (4.0.1) by rewriting it as $L(Ux) = b$ and setting $y = Ux$ The two triangular systems $Ly = b$ and $Ux = y$ is

equivalent to (4.0.1). To explain the method, let us consider the linear system in (4.2.2). Operations on equivalent systems (4.2.3) - (4.2.3) can be represented by a series of matrix multiplications.

$$R_2 - 3R_1 \rightarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ & -1 & -2 & -4 \\ -2 & 2 & 1 & 7 \end{array}\right) \tag{4.3.1}$$

$$R_1 + R_3 \rightarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ & -1 & -2 & -4 \\ & 3 & 2 & 8 \end{array}\right) \tag{4.3.2}$$

$$R_3 + 3R_2 \rightarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ & -1 & -2 & -4 \\ & & -4 & -4 \end{array}\right) \tag{4.3.3}$$

Operations on the LHS of these equivalent systems can be executed using left-hand multiplication with elementary matrices G_i

$$G_3G_2G_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -7 & 3 & 1 \end{pmatrix} \tag{4.3.4}$$

That is $G_3G_2G_1A = U$, so that $A = G_1^{-1}G_2^{-1}G_3^{-1}U = LU$, L is the lower-triangular matrix.

$$L = G_1^{-1}G_2^{-1}G_3^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.3.5}$$

Clearly A is a product of L and U. Note that U has pivots in its diagonal while L has $1'$s in its diagonal.

Once L and U are determined $Ax = b$ can be rewritten as:

$$L(Ux) = b \tag{4.3.6}$$

$$\text{if} \quad y = Ux \tag{4.3.7}$$

$$Ly = b \tag{4.3.8}$$

The lower-triangular system $Ly = b$ can be solved by forward substitution:

$$y_1 = b_1, \quad y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k \quad \text{for} \;\; i = 2, 3, \ldots, n. \tag{4.3.9}$$

Once y is known, the upper-triangular system $Ux = y$ can be solved using back substitution:

$$x_n = \frac{y_n}{u_{nn}} \tag{4.3.10}$$

$$x_i = \frac{1}{u_{ii}} \left(y_i - \sum_{k=i+1}^{n} u_{ik} x_k \right) \quad \text{for} \;\; i = n-1, n-2, \ldots, 1. \tag{4.3.11}$$

If the LU factors of A were computed and saved when the original system was solved, then they need not be recomputed for a new RHS. That is, the operation counts for each subsequent system are on the order of N^2. We will use the LU decomposition for inverting matrices when N is of the order of a few hundred. Beyond this limit we will use iterative methods outlined in Section (4.5).

4.4 Classical iterative methods

All classical iterative methods for solving $Ax = b$ start with an initial guess x_0 (zero, if nothing better is available). At the m^{th} iteration step, the updated estimate for x is :

$$x_m = x_{m-1} + h_m \tag{4.4.1}$$

Here, h_m is a chosen correction step. The best possible correction step would be the current error vector:

$$f_{m-1} = x_* - x_{m-1} \tag{4.4.2}$$

But as the unique solution x_* is unknown, f_{m-1} is also unknown and cannot be used as a correction. The error f_{m-1} would solve the *residual equations*:

$$Af_{m-1} = r_{m-1} \tag{4.4.3}$$

where, $r_{m-1} = b - Ax_{m-1}$. The idea behind classical iterative methods is to solve a related problem [58]:

$$Mh_m = r_{m-1} \tag{4.4.4}$$

where the matrix M is close to the original matrix A and is easy to manipulate in the iteration. The m^{th} iteration step of a basic iterative method can be summarised as:

$$\texttt{Compute residual vector} \;:\; r_{m-1} = b - Ax_{m-1} \tag{4.4.5}$$

$$\texttt{Solve the linear system} \;:\; Mh_m = r_{m-1} \tag{4.4.6}$$

$$\texttt{Update the previous iterate} \;:\; x_m = x_{m-1} + h_m \tag{4.4.7}$$

The principal operations in this iteration are the matrix-vector multiplications in (4.4.5) and the solution of the linear system with M in (4.4.6). If M is independent of the iteration step m, then the resulting iteration schemes are called *Stationary iterative methods*. Classical methods such as the Richardson, Jacobi, Gauss-Seidel, Successive Over Relaxation(SOR) and Symmetric Successive Over Relaxation (SSOR) methods, etc. are stationary iterative methods [14]. If A is decomposed as:

$$A = D - L - U \tag{4.4.8}$$

where, D is diagonal, $-L$ and $-U$ are strictly lower and upper triangular respectively, then the stationary methods are classified as in Table (4.1).

M	Stationary method
$\frac{1}{\omega}I, (\omega \neq 0)$	Richardson
D	Jacobi
$D - L$	Gauss Seidel
$\frac{1}{\omega}D - L, (\omega \neq 0)$	SOR
$\frac{\omega}{2-\omega}\left(\frac{1}{\omega}D - L\right)D^{-1}\left(\frac{1}{\omega}D - U\right), (\omega \neq \{0, 2\})$	SSOR

Table 4.1: Classification of stationary iterative methods based on the structure of iteration matrix

There are block versions of these methods [14]. Stationary methods are implemented by splitting the coefficient matrix $A = M - (M - A)$. Then $Ax = b$ is transformed into a fixed point form:

$$Mx = (M - A)x + b \tag{4.4.9}$$

$$x = M^{-1}(M - A)x + M^{-1}b \tag{4.4.10}$$

$$x = Tx + c \tag{4.4.11}$$

where, $T := I - M^{-1}A$ is the iteration matrix and $c := M^{-1}b$. The m^{th} iteration step can be written as:

$$x_m = Tx_{m-1} + c \tag{4.4.12}$$

The suitability of implementing an iteration of the form (4.4.12) lies in the ease of solving a linear system with T as the matrix. Convergence of (4.4.12) depends on the eigenvalues of T. These methods will converge only if the spectral radius $\rho(T) := max|\lambda| : \lambda \in \Lambda(T)$ of T is strictly less than *one* [14].

4.5 Why Krylov methods?

The origin of this class of methods is deeply rooted in approximation theory and, in particular, in orthogonal polynomials. Aleksei Nikolaevich Krylov [36] showed how a sequence of the form $b, Ab, A^2b, \ldots$ can be used to construct the characteristic polynomial for a matrix. The minimal polynomial [36] of A can be used to express A^{-1} in terms of powers of A. This shows the solution to (4.0.1) $x = A^{-1}b$ lies in a *Krylov space* whose dimension is the degree of the minimal polynomial of A. If the minimal polynomial A has a lower degree, then the dimension of the space in which an algorithm will have to search for a solution will be smaller than N. Thus, the Krylov space, when used as a search space for x, might achieve faster convergence.

Another advantage of these methods is that they do not require the matrix A to be stored, but merely needs an implicit subroutine which when given a vector x returns Ax. In electromagnetics, one often comes across such linear equations where, A is an implicit linear operator acting on some vector. *Krylov methods* solve $Ax = b$ by repeatedly performing matrix-vector multiplications involving A. Krylov methods can be again subdivided according to the constituent procedures used in the algorithm [65, 36]. All Krylov methods start with an initial guess x_0 and correct it towards a more accurate solution in each iteration. At the kth iteration, Krylov methods produce an approximate solution x_k from a Krylov space spanned by k vectors. A good candidate for this vector is b, because b is problem dependent.

$$\mathcal{K}(A, b) = \mathtt{span}\left\{b, Ab, A^2b, \ldots, A^{k-1}b\right\} \tag{4.5.1}$$

It was in the early 50s, that a number of Krylov subspace methods appeared mainly due to Lanczos [30], Arnoldi [1] and Hestenes and Stiefel [18], who proposed different versions of the well known *Conjugate Gradient* method. Krylov methods are

projection methods, where N-dimensional problems will be projected into a lower-dimensional (say k) Krylov subspace. So these methods are sometimes termed as *Krylov projection methods*. The associated Krylov sequence for Equation (4.0.1) is the set of vectors $b, Ab, A^2b, A^3b, \ldots, A^{k-1}b$, which can be computed using a subroutine $A(\ )$ to give $A(b), A(Ab), A(A(Ab)), \ldots$. The corresponding Krylov subspaces are the spaces spanned by larger groups of these vectors. $K = (b, Ab, \ldots, A^{k-1}b)$ is termed as the Krylov matrix.

It is easily observable that the Krylov sequence tends to become nearly linearly dependent. As k increases, vectors in Krylov sequence almost become multiples of each other. The directions of the vectors A^kb tend to converge to the direction of the largest eigenvalue of A. This means that K is ill conditioned. This difficulty can be overcome by replacing K by its QR factorization [65]:

$$K = Q_{n\times k}R_{k\times k} \tag{4.5.2}$$

If A is hermitian, the reduced matrix will be tridiagonal, otherwise it will have Hessenberg form [65]. A complete reduction of a large A into Hessenberg form by an orthogonal similarity transform can be written as $A = QHQ^*$ or $AQ = QH$. If A is large, computation of the full reduction is not practical. Instead, the first n columns of $AQ = QH$ is often considered. There are numerous methods based on the Krylov approach, all of them are built upon a few fundamental ideas. In the next section, the Arnoldi and Lanczos methods are reviewed, which are the foundations behind the methods used in this thesis. The orthonormal basis for the Krylov subspace can reduce a matrix to upper Hessenberg form. Matrices with orthonormal columns are perfectly conditioned. The eigenvalues of H are called the *Ritz values* for A. They are often a good approximation to the extreme eigenvalues of A, especially when A is Hermitian. Now the task is the computation of orthonormal bases for the *Krylov subspaces*. As mentioned above, the connection between the tridiagonalisation and

QR factorisation of $\mathcal{K}(A, b)$ will be used. Forthcoming subsections explore this idea.

4.6 QR Factorisation

The QR factorisation is a sequential construction of orthonormal vectors $q_1, q_2, \ldots$ that span a column space of $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$, assuming that A has full rank n. The sequence should have the property:

$$\texttt{span}\{q_1, q_2, \ldots, q_j\} = \texttt{span}\{a_1, a_2, \ldots, a_j\} \qquad j = 1, 2, \ldots, n \tag{4.6.1}$$

Consider an equation of the form [65]:

$$\left(\begin{array}{c|c|c} & & \\ a_1 & \ldots & a_n \\ & & \end{array}\right) = \left(\begin{array}{c|c|c} & & \\ q_1 & \ldots & q_{n+1} \\ & & \end{array}\right) \begin{pmatrix} r_{11} & \ldots & r_{1n} \\ & \ddots & \vdots \\ & & r_{n,n} \end{pmatrix} \tag{4.6.2}$$

Here, the diagonal entries r_{jj} are nonzero. Expanding (4.6.2):

$$\begin{aligned} a_1 &= r_{11}q_1 \\ a_2 &= r_{12}q_1 + r_{22}q_2 \\ a_3 &= r_{13}q_1 + r_{23}q_2 + r_{33}q_3 \\ a_n &= r_{1n}q_1 + r_{2n}q_2 + \ldots + r_{nn}q_n \end{aligned} \tag{4.6.3}$$

Equation (4.6.3) can be written as a compact matrix equation.

$$A = QR \tag{4.6.4}$$

Equation (4.6.4) is known as the QR factorisation of A. Equation (4.6.3) can be implemented as an algorithm to construct Q and R from A. The *Gram-Schmidt orthogonalisation* algorithm does this. At the j^{th} step, the process finds a unit vector $q_j \in \texttt{span}\ \{a_1, a_2, \ldots, a_j\}$ that is orthogonal to $q_1, q_2, \ldots, q_{j-1}$.

$$
\begin{aligned}
q_1 &= \frac{a_1}{r_{11}} \\
q_2 &= \frac{a_2 - r_{12}q_1}{r_{22}} \\
q_3 &= \frac{a_3 - r_{13}q_1 + r_{23}q_2}{r_{33}} \\
&\vdots \\
a_n &= \frac{a_n - \sum_{i=1}^{n-1} r_{in}q_i}{r_{nn}}
\end{aligned}
\tag{4.6.5}
$$

An algorithm that implements (4.6.5) is known as the *classical Gram-Schmidt orthogonalisation*.

Algorithm *Classical Gram-Schmidt*

1. **for** $j = 1, 2, \ldots, n$
2. $\quad v_j \leftarrow a_j$
3. $\quad$ **for** $i = 1, 2, \ldots, j-1$
4. $\quad\quad r_{ij} \leftarrow q_i^* a_j$
5. $\quad\quad v_j \leftarrow v_j - r_{ij}q_i$
6. $\quad r_{jj} \leftarrow \|v_j\|$
7. $\quad q_j \leftarrow v_j / r_{jj}$
8. **return**

The Gram-Schmidt algorithm is a process of triangular-orthogonalisation, making the columns q_j of an orthonormal matrix Q via sequential multiplication of vectors r_i of an upper-triangular matrix R. Numerically this turns out to be unstable due to rounding errors on a computer [36]. The *modified Gram-Schmidt orthogonalisation* that is free from this draw back is given in Section (4.7). Once the QR factorisation is done, to solve $Ax = b$ we can write $QRx = b$ or $Rx = Q^*b$. If Q and R are computed, owing to the triangular nature of the LHS one can easily solve the new system.

4.7 Modified Gram-Schmidt

In the classical Gram-Schmidt algorithm, a single orthogonal projection of rank $m-(j-1)$ is computed for every j. In the modified Gram-Schmidt the same projection is computed by a sequence of $j-1$ projections of rank $m-1$. The projector notation can be developed as follows [65], consider a sequence:

$$q_1 = \frac{P_1 a_1}{\|P_1 a_1\|}, \quad q_2 = \frac{P_2 a_2}{\|P_2 a_2\|}, \ldots \quad q_n = \frac{P_n a_n}{\|P_n a_n\|} \tag{4.7.1}$$

In Equations (4.7.1), P_j denotes an orthogonal projector which is a $m \times m$ matrix of rank $m-(j-1)$ that projects $\mathbb{C}^m$ onto $\texttt{span}\{q_1, \ldots, q_n\}$. Note that $P_1 = I$ and $P_j = I - Q_{j-1}Q_{j-1}^*$ The rank-one orthogonal projector that isolates the components in the direction of q can be denoted as:

$$P_q = qq^* \tag{4.7.2}$$

For each j the classical Gram-Schmidt algorithm computes the orthogonal projection of rank $(m-j-1)$ given by :

$$v_j = P_j a_j \tag{4.7.3}$$

$$P_j = P_{\perp_{q_{j-1}}}, \ldots, P_{\perp_{q_2}} P_{\perp_{q_1}} \tag{4.7.4}$$

$\perp$ denotes the orthogonality. In the modified Gram-Schmidt algorithm, (4.7.3) is replaced with (4.7.5):

$$v_j = P_{\perp_{q_{j-1}}}, \ldots, P_{\perp_{q_2}} P_{\perp_{q_1}} a_j \tag{4.7.5}$$

Mathematically (4.7.3) and (4.7.5) are equivalent, but the sequence of arithmetic operations in them are different. In finite precision computer arithmetic the new sequence introduces smaller errors [65]. The modified Gram-Schmidt algorithm is given below.

Algorithm *Modified Gram-Schmidt*

1. **for** $i = 1, 2, \ldots, n$

2. $v_i \leftarrow a_i$

3. **for** $i = 1, 2, \ldots, n$

4. $r_{ii} \leftarrow \|v_i\|$

5. $q_i \leftarrow v_i / r_{ii}$

6. **for** $j = i + 1, \ldots, n$

7. $r_{ij} \leftarrow q_i^* v_j$

8. $v_j \leftarrow v_j - r_{ij} q_i$

9. **return**

4.8 Lanczos iteration

When A is Hermitian, $A^* = A$ implies $H^* = (Q^* A Q)^* = H$ and symmetric Hessenberg matrices are tridiagonal. Denote H as T to symbolise the tridiagonal nature:

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ & & & \beta_{n-1} & \alpha_n \end{pmatrix} \tag{4.8.1}$$

Here, $Q = (q_1, q_2, \ldots, q_n)$ and consider $\beta_0 = 0$ and $q_{n+1} = 0$. Equating jth column in $AQ = QT$ gives a three-term recurrence relation:

$$A q_j = \beta_{j-1} q_{j-1} + \alpha_j q_j + \beta_j q_{j+1} \tag{4.8.2}$$

or,

$$\beta_j q_{j+1} = v_j, \quad \text{where, } v_j = A q_j - \alpha_j q_j - \beta_{j-1} q_{j-1} \quad \text{for } j = 1, 2, \ldots, n \tag{4.8.3}$$

The orthogonality of q_j implies that:

$$\alpha_j = q_j^T A q_j \quad \text{and} \quad \beta_j = \|v_j\| \tag{4.8.4}$$

The Lanczos iteration is presented below [65]:

Algorithm ***Lanczos Iteration***

1. $b \neq 0$, $\beta_0 = 0$, $q_0 = 0$ $q_1 = b/\left\|b\right\|$
2. **for** $j = 1, 2, \ldots, n$
3. $\quad v \leftarrow Aq_j$ or $Aq_n - \beta_{n-1}q_{n-1}$ for stability
4. $\quad \alpha_j \leftarrow q_j^T v$
5. $\quad v \leftarrow v - \alpha_j q_j - \beta_{j-1}q_{j-1}$
6. $\quad \beta_j \leftarrow \|v\|$
7. $\quad$ **if** $\beta_j = 0$ *then quit*
8. $\quad q_{j+1} \leftarrow v/\beta_j$
9. **return**

After the j^{th} step we have an $n \times (j+1)$ matrix $Q_{j+1} = (q_1, q_2, \ldots, q_{j+1})$ of orthonormal columns. The q_js are called *Lanczos vectors*. If the iteration terminates because $\beta = 0$ for $j < n$ then it can be restarted with a new vector b which is orthogonal to $q_1, q_2, \ldots, q_j$ and a complete orthonormal set q_n can be computed. Alternatively, to compute a reduced tridiagonal decomposition, one can stop the Lanczos iteration before completion

4.9 Arnoldi iteration

One way to extend the Lanczos process to asymmetric matrices is the *Arnoldi method* [14]. Reduction of $A_{m\times m}$ to Hessenberg form by an orthogonal similarity transform can be written as $A = QHQ^*$ or $AQ = QH$. The Arnoldi iteration involves a column-by-column generation of the orthogonal matrix Q. If m is large full reduction is tedious, instead we can consider the first n columns of $AQ = QH$. Let Q_n be the $m \times n$ matrix whose columns are the first n columns of Q:

$$Q_n = \left(\begin{array}{c|c|c} & & \\ q_1 & \dots & q_n \\ & & \end{array} \right) \tag{4.9.1}$$

It is straightforward to notice that here, $Q^*AQ = H$, where H is a Hessenberg matrix. Let H_n be the $(n+1) \times n$ upper left section of H

$$H_n = \begin{pmatrix} h_{11} & & \dots & h_{1n} \\ h_{21} & h_{22} & & \vdots \\ & \ddots & \ddots & \\ & & h_{n,n-1} & h_{n,n} \\ & & & h_{n+1,n} \end{pmatrix} \tag{4.9.2}$$

Now, one can write:

$$AQ_n = Q_{n+1}H_n \tag{4.9.3}$$

that is:

$$\begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} \left(\begin{array}{c|c|c} & & \\ q_1 & \dots & q_n \\ & & \end{array} \right) = \left(\begin{array}{c|c|c} & & \\ q_1 & \dots & q_{n+1} \\ & & \end{array} \right) \begin{pmatrix} h_{11} & \dots & h_{1n} \\ & \ddots & \vdots \\ & & h_{n+1,n} \end{pmatrix} \tag{4.9.4}$$

The nth column of this equation can be written as:

$$Aq_n = h_{1n}q_1 + \ldots + h_{nn}q_n + h_{n+1,n}q_{n+1} \tag{4.9.5}$$

Equation (4.9.5) shows that q_{n+1} satisfies an $(n+1)$-term recurrence involving itself and the previous Krylov vectors. The Arnoldi Iteration [65] is the Gram-Schmidt iteration that implements Equation (4.9.5).

Algorithm *: Arnoldi Iteration*

1. $b =$ arbitrary, $q_1 = b/\left\|b\right\|$

2. **for** $n = 1, 2, 3, \ldots$

3. $\quad v \leftarrow Aq_n$

4. $\quad$ **for** $j = 1$ **to** n

5. $\quad\quad h_{jn} \leftarrow q_j^* v$

6. $\quad\quad v \leftarrow v - h_{jn} q_j$

7. $\quad h_{n+1,n} \leftarrow \|v\|$

8. $\quad q_{n+1} \leftarrow v / h_{n+1,n}$

9. **return**

After the k^{th} step, we have an $n \times (k+1)$ matrix $Q_{k+1} = (q_1, q_2, \ldots, q_{k+1})$ of orthonormal columns. The Arnoldi algorithm will not be as efficient as the Lanczos because of the Hessenberg matrix instead of a tridiagonal one. It can be easily noticed that the Arnoldi algorithm is just the modified *Gram-Schmidt* process [14].

It is evident from (4.9.5) that q_j form bases of the successive Krylov subspaces generated by A and b:

$$\mathcal{K}_n = \mathtt{span}\{b, Ab, \ldots, A^{n-1}b\} = \mathtt{span}\{q_1, q_2, \ldots, q_n\} \subseteq \mathbb{C}^m \tag{4.9.6}$$

Note that vectors q_j are orthonormal, so these are orthonormal bases. The Arnoldi process is therefore a method for construction of orthonormal bases for successive Krylov subspaces. q_j are called *Arnoldi vectors*.

4.10 Various Krylov methods

In our study, we will deal with inhomogeneous objects and the MoM matrices involved in computations will be asymmetric. A classification of suitable Krylov subspace iterative methods based on matrix properties is given in Table (4.2).
In Table (4.2), CG is the Conjugate Gradient, and GMRES is the Generalised Minimal RESidual method. The CG algorithm requires that the matrix A is positive-definite

	Ax=b	$Ax = \lambda x$
A=A*	CG	Lanczos
$A \neq A^*$	GMRES	Arnoldi

Table 4.2: Comparison of matrix properties and Krylov methods

and hermitian. There are other variants of the Krylov family where A need not be symmetric. In all of these methods projection into the Krylov subspaces reduces the original problem to a sequence of matrix problems of dimensions $n = 1, 2, 3, \ldots$.

4.10.1 Conjugate Gradient Method

The Conjugate Gradient Method (CGM), famously known as the CG-method is the *original* Krylov subspace iteration. Discovered independently by Hestenes and Stiefel [18] in 1952, the CGM is the most common iterative method. It solves symmetric positive definite systems of equations quickly if eigenvalues are well distributed [65] If $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite, consider a function defined by:

$$\phi(x) = \frac{1}{2} x^T A x = x^T b \tag{4.10.1}$$

where $b \in \mathbb{R}^n$, Its gradient is given by:

$$\Delta\phi(x) = Ax - b \tag{4.10.2}$$

A unique minimizer for ϕ is the the solution $x = A^{-1}b$. Thus any attempt to minimize ϕ can be used to find the solution x. Since x is somewhere in $\mathcal{K}$ one can try to minimize ϕ over $\mathcal{K}$. One approach is the *method of steepest descent*. In this, the current approximation x_j is updated by adding a correction term directed along the negative gradient:

$$-\nabla\phi(x_j) = b - Ax_j = r_j \tag{4.10.3}$$

$$x_{j+1} = x_j + \alpha_j r_j \tag{4.10.4}$$

where, $\alpha_j = \dfrac{r_j^T r_j}{r_j^T a r_j}$, because this α minimizes $\phi(x_{j+1})$. The rate of convergence can be slow if the ratio of eigenvalues $\lambda_{max}(A)/\lambda_{min}(A)$ is large. Geometrically, this means that the surface ϕ can be highly distorted and in the steepest decent, the corrections can move back and forth without going to the lowest point quickly.

A workaround for this is to replace search directions r_j by vectors $q_1, q_2, \ldots$ that are conjugate to each other such that $q_i^T A q_j = 0$ for $i \neq j$. The idea is to start the iteration with x_0,

$$x_1 = \alpha_1 q_1, \quad \text{where, } q_1 = r_0 = b \quad \text{and} \quad \alpha_1 = \frac{r_0^T r_0}{r_0^T A r_0} \tag{4.10.5}$$

In the second step, use the direction vector:

$$q_2 = r_1 + \beta_1 q_1, \quad \text{where, } \beta_1 \text{ is chosen to force } q_2^T A q_1 = 0 \tag{4.10.6}$$

where, $\beta_1 = \dfrac{r_1^T r_1}{r_0^T r_0}$, then set $x_2 = x_1 + \alpha_2 q_2$.

Algorithm *CG Iteration*

1. $x_0 = b, r_0 = b, q_1 = b$
2. **for** $j = 1, 2, 3, \ldots$
3. $\alpha_j \leftarrow r_{j-1}^T r_{j-1} / q_j^T A q_j$ (step size)
4. $x_j \leftarrow x_{j-1} + \alpha_j q_j$ (approximate solution)
5. $r_j \leftarrow r_{j-1} - \alpha_j A q_j$ (residual)
6. **if** $\|r_j\| = 0$, or is satisfactorily small
7. **then** set $x = x_j$ and **stop**
8. **else**
9. $\beta_j \leftarrow r_j^T r_j / r_{j-1}^T r_{j-1}$ (conjugation factor)

10. $q_{j+1} \leftarrow r_j + \beta_j q_j$ (search direction)

Vectors produced by the *CG algorithm* after j steps are such that:

$$\mathtt{span}\{x_1, x_2, \ldots, x_j\} = \mathtt{span}\{q_1, q_2, \ldots, q_j\} = \mathtt{span}\{r_0, r_1, \ldots, r_{j-1}\} = \mathcal{K}_j \quad (4.10.7)$$

The Algorithm will find the solution in $k \leq n$ steps. The CG iteration involves many vector manipulations and one matrix-vector product. If A is dense the matrix-vector product dominates the operation count, which is $\sim 2n^2$ for each iteration.

4.10.2 GMRES

After j Arnoldi iterations, one obtains Q_j and Q_{j+1} with orthonormal columns that span $\mathcal{K}_j$ and $\mathcal{K}_{j+1}$, respectively. Also, a $j \times j$ upper-Hessenberg matrix H_j:

$$AQ_j = Q_{j+1}\tilde{H}_j, \quad \text{here} \quad H_j = \begin{pmatrix} \tilde{H}_j \\ h_{j+1,j}e_j^T \end{pmatrix} \quad (4.10.8)$$

Equation (4.10.8) suggests that least squares solution $AQ_j z = b$ is the same as $Q_{j+1}\tilde{H}z = b$. Also, $\tilde{H}_j z = Q_{j+1}^T b$. But $Q_{j+1}^T b = \|b\|_2 e_1$, because the first column in Q_{j+1} is $b/\|b\|_2$

Algorithm ***GMRES Iteration***

1. $b =$ arbitrary, $q_1 = b/\|b\|_2$
2. **for** $n = 1, 2, 3, \ldots$
3. $\quad v \leftarrow Aq_n$
4. $\quad$ **for** $j = 1$ **to** n
5. $\quad\quad h_{jn} \leftarrow q_j^* v$
6. $\quad\quad v \leftarrow v - h_{jn}q_j$
7. $\quad h_{n+1,n} \leftarrow \|v\|$
8. $\quad q_{n+1} \leftarrow v/h_{n+1,n}$

9. Usr QR method to compute the least squares solution of $\tilde{H}_j z = \|b\|_2 e_1$

10. **if** $\|b - AQ_j z\|_2 <$ tolerance, set $x = Q_j z$

11. **else** go to line 3

12. **return**

The upper Hessenberg least square problem can be solved using rotations. The main problem with GMRES is that the j^{th} iteration involves $\mathcal{O}(kn)$ operations. Thus a restart strategy is employed to minimize iterations and to limit memory usage. If m steps are tolerable, the algorithm can be restarted with x_m as initial guess for another GMRES iteration.

4.11 Singular Value Decomposition

Many problems in Linear Algebra can be better dealt with by using the Singular value Decomposition. The Singular Value Decomposition (SVD) is often seen as a significant topic in linear algebra due to its potential application to a wide range of data analysis. SVD has many practical and theoretical values. One special feature of SVD is that it can be performed on any $m \times n$ matrix. It factors A into three matrices U, S, V such that, $A = USV^T$, where U and V are orthogonal matrices and S is a diagonal matrix.

If A is a real $m \times n$ matrix, then there exists orthogonal matrices:

$$U = [u_1, u_2, \ldots, u_n] \in \mathbb{R}^{m \times m} \quad \text{and} \quad V = [v_1, v_2, \ldots, v_n] \in \mathbb{R}^{n \times n}$$

such that,

$$U^T AV = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{m \times n} \quad p = \min(m, n) \tag{4.11.1}$$

where,

$$\sigma_1 \geq \sigma_2 \geq \sigma_p \geq 0$$

If A is a complex $m \times n$ matrix, then

$$A = U\Sigma V^* \tag{4.11.2}$$

In this case, there exists unitary matrices:

$$U = [u_1, u_2, \ldots, u_n] \in \mathbb{C}^{m\times m} \quad \text{and} \quad V = [v_1, v_2, \ldots, v_n] \in \mathbb{C}^{n\times n}$$

The σ_i are called the singular values, the first p columns of V the right singular vectors and the first p columns of U the left singular vectors. Singular values are the eigenvalues of A^TA, and the singular vectors are specialised sets of eigenvectors for A^TA and AA^T. Constructive method used to derive the SVD can be used as an algorithm but in practice sophisticated techniques exist, and all good matrix computation packages contain numerically stable SVD implementations. The details of a practical SVD algorithm are too complicated to be discussed in this thesis. A detailed description of such algorithms can be found in [65]. The practical algorithm for computing the SVD is an implementation of the QR iteration applied to A^TA without explicitly computing A^TA.

Chapter 5

The Forward Backward Method

Recently, there has been much research into what may be termed *physically inspired iterative solutions*, or informally known as *current marching* methods. Examples include the *method of ordered interactions* [25] and the *Forward Backward (FB) method* [20]. This chapter investigates an adaptive strip-size current-marching algorithm which is applicable to dielectrics modelled using the Electric Field Integral Equation (EFIE). The algorithm presented in this chapter constitutes an extension of the ideas presented in [7] for conducting objects. These ordinary FB methods [7] are seen to be not suitable for dielectric objects modelled with the EFIE. The new algorithm introduced here is termed the Adaptive-Strip Forward-Backward (ASFB) method. The performance of the ASFB when it is applied to dielectric scattering problems will be studied in this chapter.

5.1 Numerical Procedure

Consider an infinite cylindrical dielectric scatterer as in Figure (5.1), illuminated by a normally incident TM plane-wave. E^i is the incident electric field that would exist in the absence of the scatterer. The total field anywhere in space is then given by the sum of the incident and scattered electric fields where the scattered field E^s is given

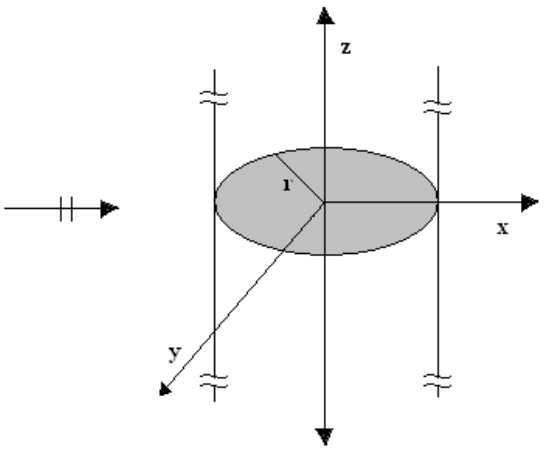

Figure 5.1: *An infinite homogeneous cylinder whose electrical properties are invariant in the z-direction. The shaded region is the cross-section of the cylinder at $z = 0$. The radius of the cylinder is denoted as r. A plane wave in the $+x$-direction with frequency f illuminates the object normal to the surface of the cylinder.*

by the following integral expression. For TM polarisation, the EFIE can be written as [44]:

$$E_z^i(x,y) = \frac{J_z}{\jmath\omega\epsilon_0(\epsilon_r - 1)} + \jmath\omega\mu_0 A_z \tag{5.1.1}$$

The object is impinged by a TM plane wave of frequency $f = 300$ MHz. The radius of the object is $r = \lambda$ where, the vector potential A_z is given by:

$$A_z(x,y) = \int\int J_z(x',y')\frac{1}{4j}H_0^{(2)}(kR)dx'dy' \tag{5.1.2}$$

where

$$R = \sqrt{(x-x')^2 + (y-y')^2} \tag{5.1.3}$$

In this formulation, the z component of current density J_z is the primary unknown. Following the discretisation process in Section (3.7.2), Equation (5.1.1) can be written

as an $N \times N$ matrix equation (5.1.4). N is the number of basis functions used to discretise the cross section of the object.

$$ZJ = E^i \tag{5.1.4}$$

The matrix elements Z_{mn} are given by:

$$Z_{mn} = \frac{\eta\pi a_n}{2} J_1(ka_n) H_0^{(2)}(kR_{mn}) \qquad m \neq n \tag{5.1.5}$$

$$Z_{mm} = \frac{\eta\pi a_m}{2} H_1^{(2)}(ka_m) - \frac{j\eta\epsilon_{rm}}{k(\epsilon_{rm} - 1)} \qquad m = n \tag{5.1.6}$$

In Equations (5.1.5) and (5.1.6) R_{mn} is the distance between cells m and n, η and k are the impedance and wave-number of the dielectric medium respectively. a is the radius of a circle of equal area $dxdy$. In the discretisation process, according to Equation (3.8.1) N will be large. As the electrical size of the object increases, it becomes impossible to store or invert the impedance matrix Z. This leads to the need for domain decomposition methods where it is possible to solve bigger problems in steps, without having to deal with bigger matrices.

5.2 The Ordinary FB method

The FB method is implemented by breaking down the scatterer into smaller blocks of size $n_{ofb} \times n_{ofb}$. Then each sub-block is solved sequentially for the unknown basis function amplitudes J_z in that block (henceforth denoted J for convenience) in a manner that attempts to mimic the physical processes that create the current. This can often yield fast convergence in a reduced number of iterations [7]. Mathematically, the algorithm involves decomposing the Z matrix into blocks, the $\tilde{Z}_{ij}$ block containing the interactions between basis functions residing in the i^{th} and j^{th} sub-regions on the scatterer. Each iteration of a forward-backward algorithm involves solving two equations. The first equation (5.2.1) is solved for the sub-regions $i = 1 \ldots m$ in turn and is termed the forward sweep.

$$\tilde{Z}_{ii}\tilde{J}_i^k = \tilde{E}_i - \sum_{j=1}^{i-1} \tilde{Z}_{ij}\tilde{J}_j^k - \sum_{j=i+1}^{m} \tilde{Z}_{ij}\tilde{J}_j^{k-1} \tag{5.2.1}$$

where, $\tilde{E}_i$ and $\tilde{J}_i$ are the ith sub-block vectors of E^i and J, respectively. Equation (5.2.1) is a matrix equation for the k^{th} estimate of the currents on sub-region i. Note that the right-hand side incident fields have been modified to include the effects of the most up to date current estimates available for the other sub-regions. As it involves a matrix of relatively low order, Equation (5.2.1) can be efficiently solved using a conjugate gradient solver.

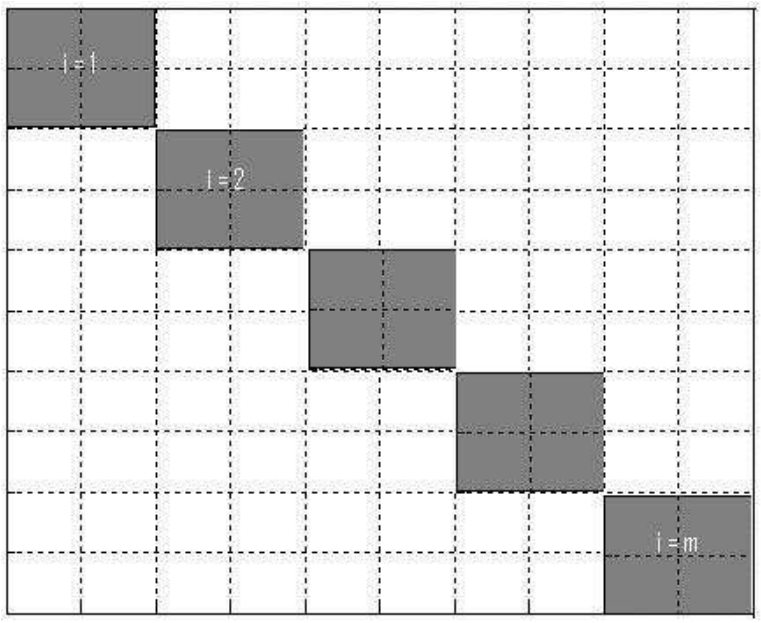

Figure 5.2: *Interaction matrix Z subdivided into m blocks ($\tilde{Z}$) of equal size $n_{ofb} \times n_{ofb}$ in the ordinary FB method.*

$$\tilde{Z}_{ii}\tilde{J}_i^{k+1} = \tilde{E}_i - \sum_{j=1}^{i-1} \tilde{Z}_{ij}\tilde{J}_j^k - \sum_{j=i+1}^{m} \tilde{Z}_{ij}\tilde{J}_j^{k+1} \tag{5.2.2}$$

Equation (5.2.1) is solved for $i = m \ldots 1$ in turn and corresponds to a backward sweep. The right-hand side incident fields have been modified to include the effects

of the most up to date current estimates available for the other sub-regions. Being of low order, Equation (5.2.2) can be efficiently solved using a conjugate gradient solver. For the sake of clarity in comparing the FB method's performance with other iterative methods, a full FB iteration is considered to be one complete forward sweep followed by a complete backward sweep. The convergence or divergence of the iteration process is defined by computing how well the governing equation (5.1.4) is satisfied at each iteration. The error δ_{ofb}^{n} in satisfying (5.1.4) after completion of the n^{th} iteration is given by:

$$\delta_{ofb}^{n} = \frac{\|ZJ_z - E_z^i\|_2}{\|E_z^i\|_2} \tag{5.2.3}$$

5.2.1 OFBM applied to homogeneous 2D cylinder

The OFB method is applied to the dielectric cylinder shown in Figure (5.1). The interaction matrix Z is sub-divided into m blocks $\tilde{Z}$ of equal size $n_{ofb} \times n_{ofb}$ as shown in Figure (5.2). The absolute error in the OFB iteration process given by Equation (5.2.3) is plotted in Figure (5.3) The experiment is repeated for several block sizes. Each fails to achieve convergence.

Figure (5.3) shows that the FB method is inappropriate even for a simple object. This inability of the ordinary FB method is due to an improper decomposition of the interaction matrix. Another reason for the OFB method's failure is due to spurious edge effects arising due to the improper decomposition. These edge effects will propagate and consequently, distort the computation. In the next section, a new method to overcome these problems of the OFB method is introduced.

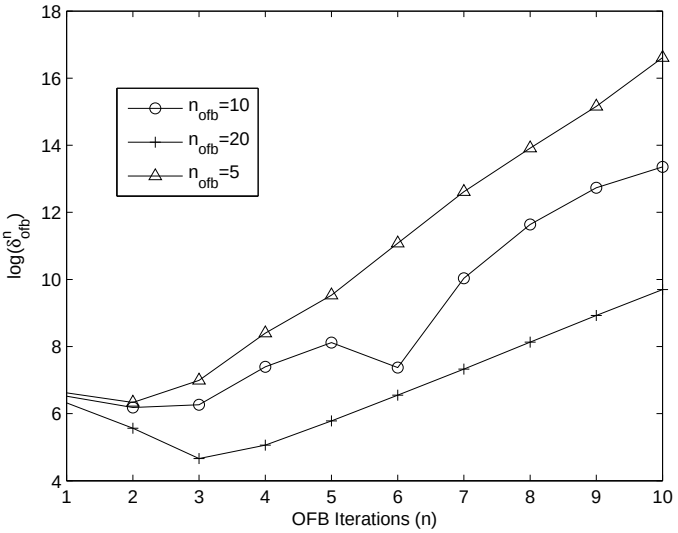

Figure 5.3: *The OFB iteration applied to a 2D dielectric cylinder, with* $\epsilon = 3$ *and* $\sigma = 1.2$*. The object is impinged by a TM plane wave at frequency* $f = 300$ MHz*. The radius of the object is* $r = \lambda$*.* n_{ofb} *is the number of basis functions in each sub-block.*

5.3 Adaptive-Strip FB method

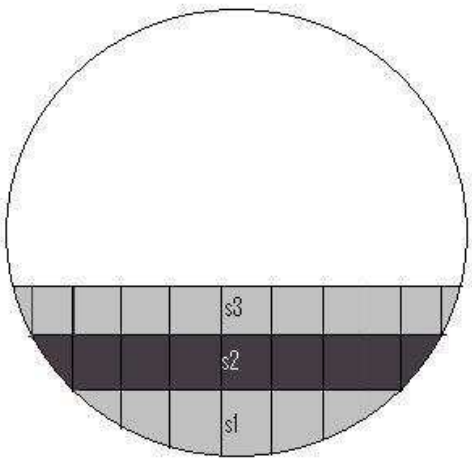

Figure 5.4: *2D dielectric cylinder, with* $\epsilon = 3$ *and* $\sigma = 1.2$ *radius* $r = \lambda$*. The number of strips is given by* $N_s = 2r/dx$*,* dx *is the discretisation size.*

Consider the corss-section of the dielectric cylinder impinged by a TM plane wave of frequency $f = 300$ MHz. The radius of the object is $R = \lambda$. The object can be broken into N_s strips of different sizes as shown in Figure (5.4). n_{asfb} is the number of basis functions in each strip. The complex permittivity of the dielectric is set to be $\epsilon = 3$ and $\sigma = 1.2$. The interaction matrix Z is subdivided into N_s sub-blocks according to the geometry of the object as shown in Figure (5.5). The ASFB algorithm determines the strip width from the information obtained from the geometry of the object. The added storage requirement due to this is just an N_s dimensional vector. Using this information, the ASFB algorithm suitably subdivides the interaction matrix Z. The interactions between basis function in each strip will be grouped together in a single sub-matrix $\tilde{Z}$ and solved using the ASFB method. Equation (5.2.1) and (5.2.2) are modified for the ASFB method. Equation (5.3.1) corresponds to the forward sweep

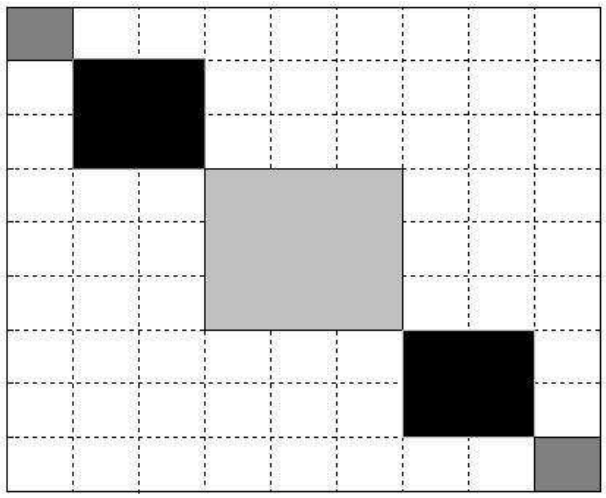

Figure 5.5: *Interaction matrix subdivided according to strip width. Note the adaptive step size, this gives a superior performance compared to ordinary FB methods (For convenient representation, the number of basis functions in each strip is set to be* $1, 2, 3, 2$ *and* 1*)*

and Equation (5.3.2) corresponds to the backward sweep.

$$\tilde{Z}_{s(i)}\tilde{J}^{k}_{s(i)} = \tilde{E}_{s(i)} - \sum_{strip(1)}^{strip(i-1)} \tilde{Z}_{s(j)}\tilde{J}^{k}_{s(j)} - \sum_{strip(i+2)}^{strip(m)} \tilde{Z}_{s(j)}\tilde{J}^{k-1}_{s(j)} \tag{5.3.1}$$

$$\tilde{Z}_{s(i))}\tilde{J}^{k+1}_{s(i))} = \tilde{E}_{s(i))} - \sum_{strip(1)}^{strip(i-2)} \tilde{Z}_{s(j)}\tilde{J}^{k}_{s(j)} - \sum_{strip(i+1)}^{strip(m)} \tilde{Z}_{s(j)}\tilde{J}^{k+1}_{s(j)} \tag{5.3.2}$$

In Equations (5.3.1) and (5.3.2) $\tilde{Z}_{s(i)}$ is the sub-matrix of Z containing basis functions from the i^{th} strip. A similar notation is used for sub-vectors of E^i and J. In an adaptive FB method, the user can select the number of strips n_s that can be joined together for forming a single block $\tilde{Z}$.

The ASFB method makes room for a new approach for solving highly inhomogeneous and complex shaped objects. Since the sub-blocks in the interaction matrix will be of different sizes $n_{asfb} \times n_{asfb}$, one can use direct inversion methods for relatively

smaller sub-blocks and various other Krylov methods for larger blocks. For highly inhomogeneous objects divided into smaller strips, various Krylov methods will show different performances depending on the electrical parameters in various strips. Thus the algorithm can itself select suitable solvers for each sub-block from the information provided in a database describing which solver is most suitable for different dielectric properties.

5.3.1 ASFB method applied to homogeneous 2D Cylinder

The ASFB method is applied for a homogeneous 2D cylindrical dielectric impinged by a TM plane wave of frequency $f = 300$ MHz. Dielectric properties are varied in the range: $\epsilon = \{1, 7\}$ and $\sigma = \{0, 4\}$. The object can be broken into N_s strips of different sizes as shown in Figure (5.4). The ASFB algorithm determines the strip width from information about the geometry of the object. n_{asfb} is the number of basis functions in each strip. Also, a number of strips n_s may be clubbed together to form a single sub-block $\tilde{Z}$ in the ASFB method. Figure (5.6) shows ASFB method's convergence compared to that of the OFB method. It is seen that as the electrical size of the object increases, ASFB method's performance deteriorates. In order to over come this inability of the ASFB, a number of buffer strips can be added to a number of strips n_s. The next section investigates this extension of the ASFB.

5.4 Buffered ASFB Method (BASFB)

As a means to improve the convergence properties of the ASFB, the Buffered ASFB is investigated for dielectric objects. The ASFB method treats each strip as a physically isolated scatterer. This results in spurious edge effects that arise from this isolation of the strips. These edge effects will propagate and consequently, distort the

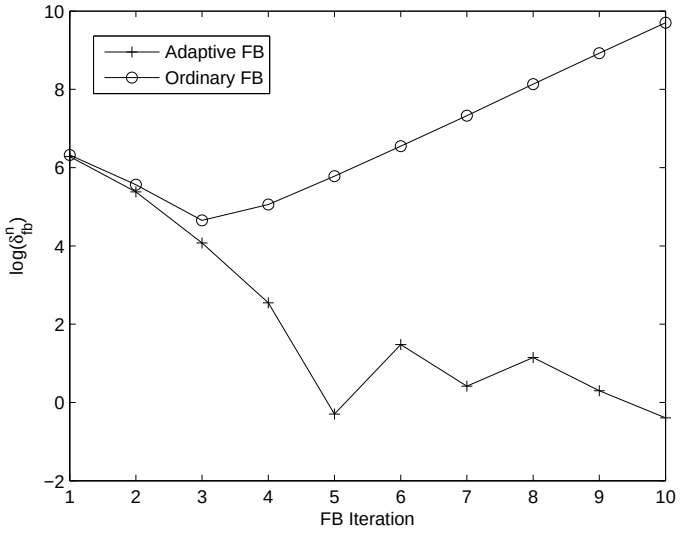

Figure 5.6: *The ASFB method applied to 2D cylinder ($\epsilon = 3.0$ and $\sigma = 1.2$). ASFB shows a superior performance compared to ordinary FB methods with 20 basis functions in each sub-block.*

computation if not suppressed. To eliminate this problem, one or more buffer strips b_i can be added to each strip. The inclusion of these buffer regions suppresses the spurious edge effects that occur within the isolated strips of a dielectric object. The buffering scheme in [7] can be applied with the ASFB method to improve it. Equations (5.3.1) and (5.3.2) can be modified to include buffer strips in ASFB method. Equation (5.4.1) corresponds to the forward sweep and Equation (5.4.2) corresponds to the backward sweep with buffers.

$$\tilde{Z}_{s(i\cup(i+1))}\tilde{J}^{k}_{s(i\cup(i+1))} = \tilde{E}_{s(i\cup(i+1))} - \sum_{strip(1)}^{strip(i-1)} \tilde{Z}_{s(j)}\tilde{J}^{k}_{s(j)} - \sum_{strip(i+2)}^{strip(m)} \tilde{Z}_{s(j)}\tilde{J}^{k-1}_{s(j)} \qquad (5.4.1)$$

$$\tilde{Z}_{s(i\cup(i-1))}\tilde{J}^{k+1}_{s(i\cup(i-1))} = \tilde{E}_{s(i\cup(i-1))} - \sum_{strip(1)}^{strip(i-2)} \tilde{Z}_{s(j)}\tilde{J}^{k}_{s(j)} - \sum_{strip(i+1)}^{strip(m)} \tilde{Z}_{s(j)}\tilde{J}^{k+1}_{s(j)} \qquad (5.4.2)$$

Equation (5.4.1) is solved for $i = 1 \ldots m$ and Equation (5.4.2) for $i = m \ldots 1$. In Equation (5.4.1), $\tilde{Z}_{s(i\cup(i+1))}$ denotes the sub-matrix $\tilde{Z}$ that contains interactions from

strip(i) and *strip*$(i+1)$, likewise for $\tilde{J}$ and $\tilde{E}$. At any update i only the $\tilde{J}_{s(i)}$ in that strip will be updated. In Figure (5.7), the interaction matrix Z is subdivided into

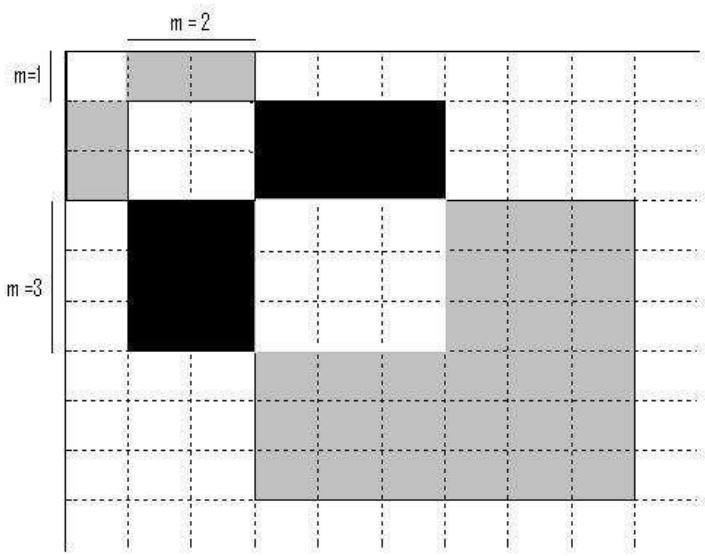

Figure 5.7: *The interaction matrix subdivided according to the strip width, during forward sweep from* 1 *to* i*. For* i^{th} *strip the very next strip* $(i+1)$ *is used as a buffer. During the backward sweep from* i *to* 1*, the* $(i-1)^{\text{th}}$ *strip will be taken as a buffer.*

blocks of different sizes $\tilde{Z}_i$ and a buffer is added to each block. For clarity of the discussion, consider the strip sizes for $i = 1, \ldots, N_s$ to be $1, 2, 3, etc.$ In Figure (5.7) for the ith strip the very next strip $(i+1)$ is used as a buffer. That is, during the forward sweep, when $i = 1$, $\tilde{Z}_1$ will be a 3×3 matrix which consists of interactions from *strip* 1 and *strip* 2. Before moving to $i = 2$, only $\tilde{J}_1$ will be updated. When $i = 2$ in the BASFB iteration $\tilde{Z}_2$ will be a 5×5 matrix which consists interactions from *strip* 2 and *strip* 3, after this only $\tilde{J}_2$ will be updated. This is repeated and a forward sweep will be completed updating the whole J. When $i = N_s$, there will be no buffer. During the backward sweep from $i = m \ldots 1$, the $(i-1)$th strips will be taken as a buffer for the ith strip.

5.4.1 BASFB method applied to homogeneous 2D cylinder

The BASFB method is applied to a homogeneous 2D cylindrical dielectric impinged by a TM plane wave of frequency $f = 300$ MHz. Dielectric properties are varied in the range: $\epsilon = \{0, 7\}$ and $\sigma = \{0, 4\}$. The object can be broken into N_s strips of different sizes as shown in Figure (5.4). A number of strips n_s may be clubbed together to form a single sub-block $\tilde{Z}$ in the ASFB method. b is the number of strips added as buffer strips with n_s strips. Figure (5.4.1) plots the error in Equation (5.2.3) against the iteration number.

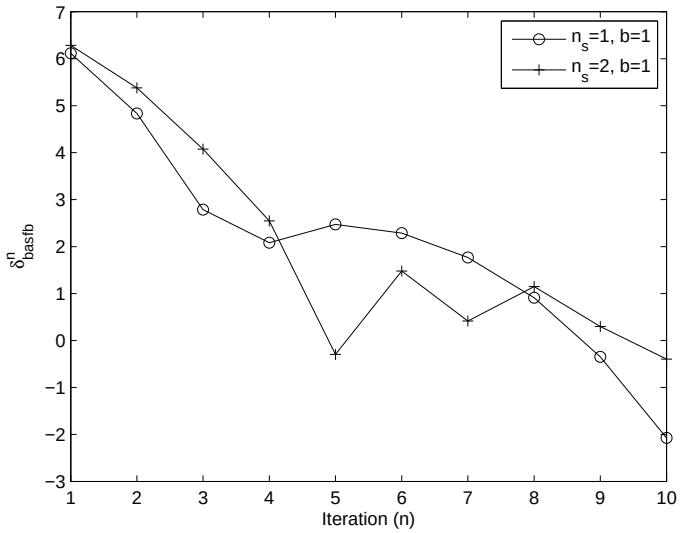

Figure 5.8: *Absolute error given by Equation (5.2.3) for the object in Figure (5.1) with $\epsilon = 3.0$ and $\sigma = 1.2$. The BASFB shows a superior performance compared to the ASFB method. n_s is the number of strips and b is the number of buffers.*

Figure (5.4.1) shows that the BASFB method is comparatively better than the ASFB method. However, this marginally better performance comes form a relatively expensive buffering scheme. It is seen that the buffer width required is often of the same size of the width of the strip or even wider that it. The following figures show

surface plots of the magnitude of J_z and the distribution of δ.

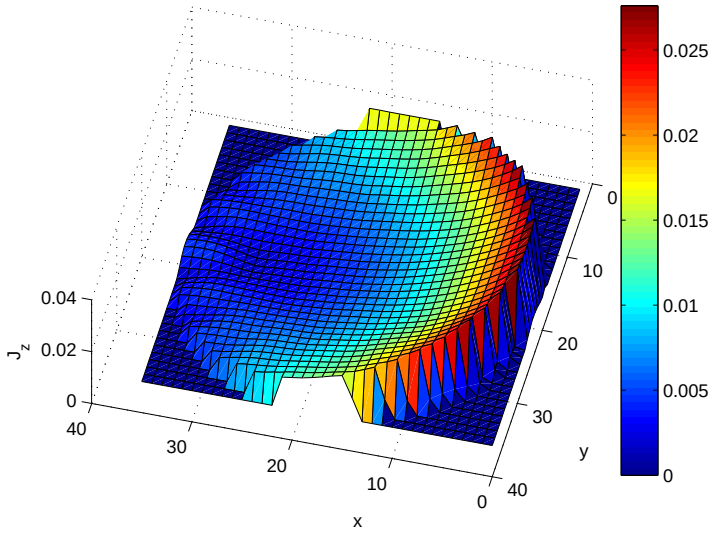

Figure 5.9: *Analytical solution for the cylindrical object with* $\epsilon = 3.0$ *and* $\sigma = 1.2$

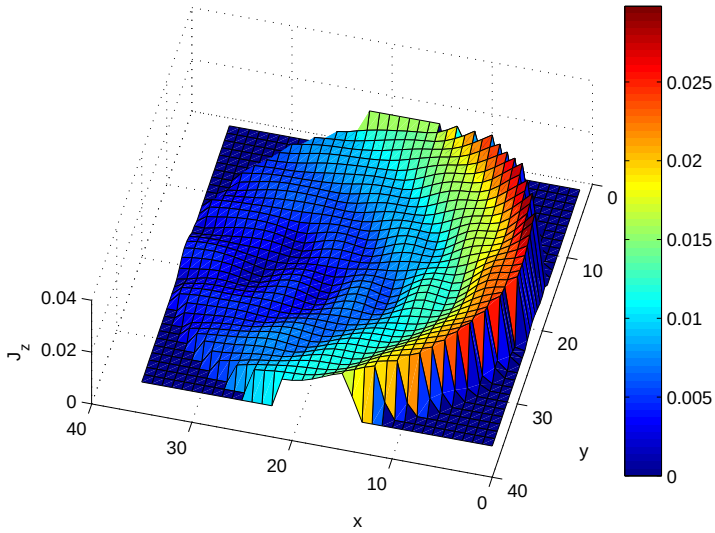

Figure 5.11: *Polarisation current density* J_z *on the 2D cylindrical dielectric object discretised into 1020 cells.* $\epsilon = 3.0$ *and* $\sigma = 1.2$. *To solve for* J_Z, *with the BASFB method with* $n_s = 1$ *and* $b = 1$ *is used.*

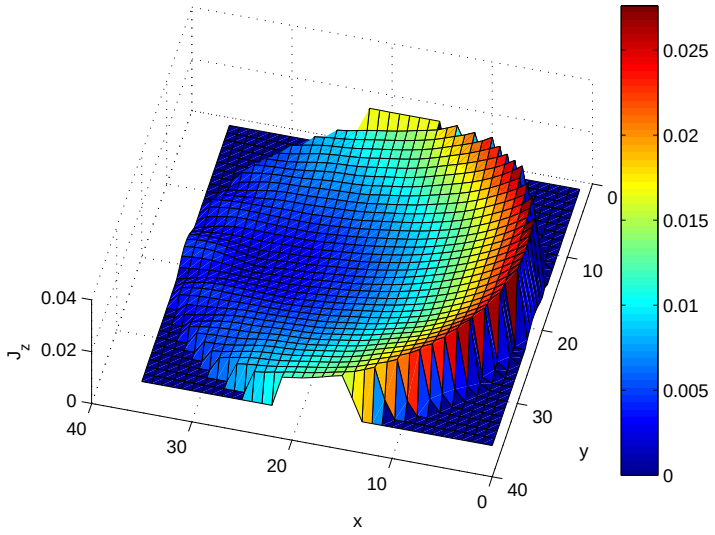

Figure 5.10: *Polarisation current density J_z on the 2D cylindrical dielectric object discretised into 1020 cells. $\epsilon = 3.0$ and $\sigma = 1.2$. To solve for J_Z, GMRES(10) was used, which converged at outer iteration 9 (inner iteration 7), with relative residual* $0.94e-4$

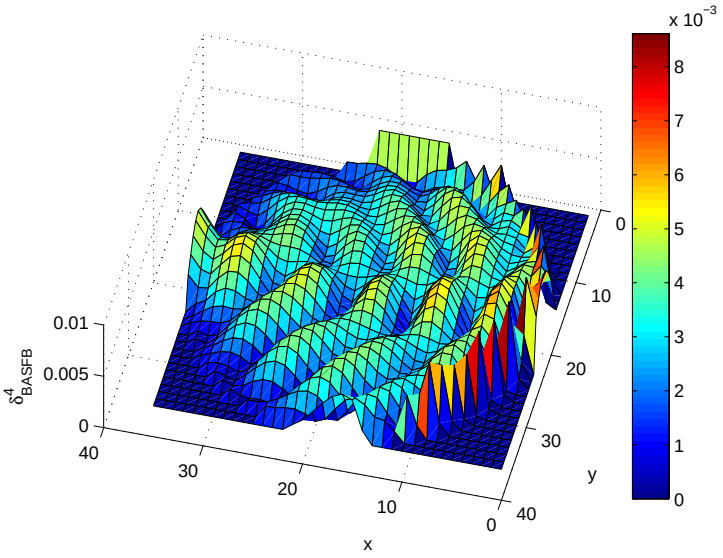

Figure 5.12: *Surface distribution of the error δ^4_{basfb} in polarisation current density J_z obtained from the BASFB method compared to the GMRES solution.*

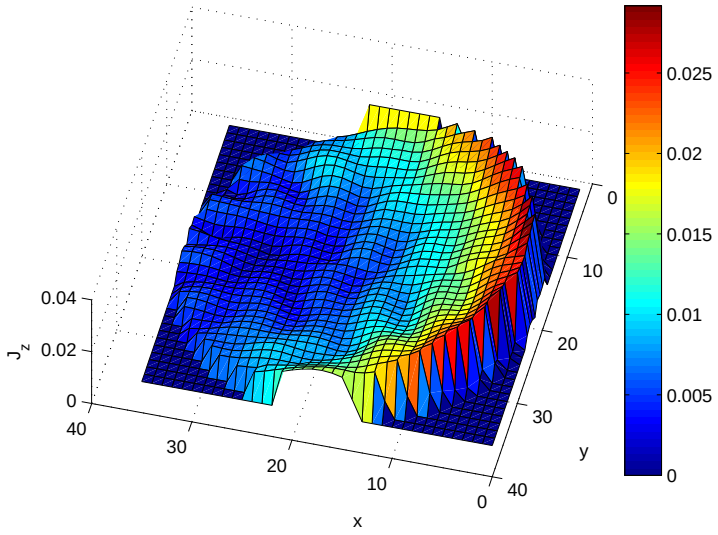

Figure 5.13: *Polarisation current density J_z on the 2D cylindrical dielectric object discretised into 1020 cells. $\epsilon = 3.0$ and $\sigma = 1.2$. To solve for J_Z, BASFB method with $n_s = 2$ and $b = 1$ is used.*

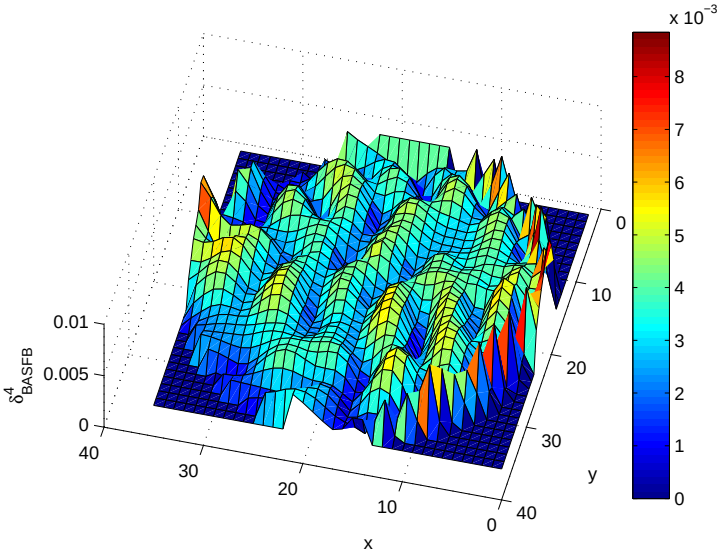

Figure 5.14: *Surface distribution of the error δ^4_{basfb} in polarisation current density J_z obtained from the BASFB method compared to the GMRES solution.*

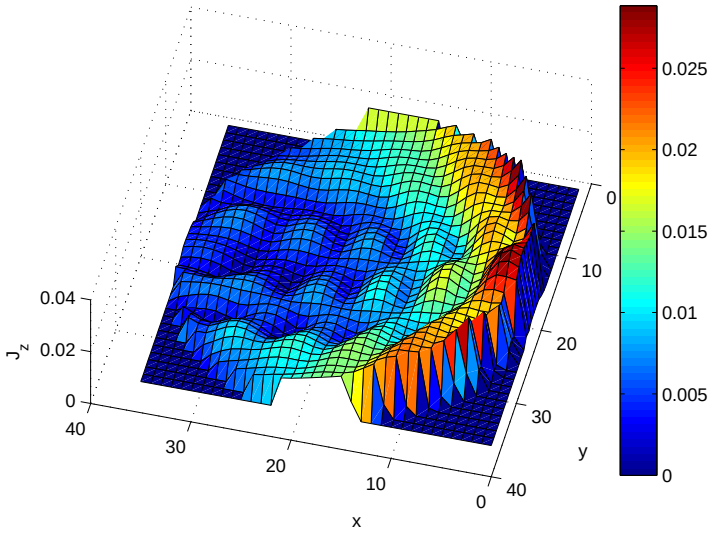

Figure 5.15: Polarisation current density J_z on the 2D cylindrical dielectric object discretised into 1020 cells. $\epsilon = 3.0$ and $\sigma = 1.2$. To solve for J_Z, ASFB method with $n_s = 1$ is used.

5.5 Final Comments on the FB Methods

As the electrical size of the object is increased, all the versions of the forward backward methods presented in this chapter fail to converge. When the object size is increased the BASFB method requires buffer regions wider than the width of the strip. This is not satisfactory for a domain-decomposition method. We conclude that the forward-backward algorithm based methods are not suitable for the electromagnetic analysis of dielectric objects.

The FBM is semi-iterative in nature and lacks a global nature in its algorithm. That is, in each iteration individual blocks are solved at different steps to find the currents in that block. This is one reason behind the unpredictable behaviour of solutions obtained via FBMs. In the following chapters of this thesis we verify the

suitability of another domain decomposition method known as the Characteristic Basis Function Method (CBFM) and develop it further.

Chapter 6

The Characteristic Basis Function Method (CBFM)

A recently developed domain decomposition method, termed the Characteristic Basis Function Method (CBFM) [39, 46], is designed for solving large-scale electromagnetic problems [34, 17] with limited computing resources. The CBFM uses a kind of *macro basis function* [62] termed the *characteristic basis function*. The CBFM algorithm permits the user to set an upper limit on the size of the matrix equation that must be inverted while modelling a variety of electromagnetic problems. Even if the object is electrically large, one can suitably decompose the object geometry, use the CBF method and thus constrain the size of the matrix equation such that a direct method in Chapter 4 can be used for its inversion. To date, the CBFM has not been extensively used for the analysis of dielectrics objects [40]. A thorough study of the CBFM's efficiency for solving scattering problems involving dielectric objects is carried out in this chapter.

6.1 Motivation and Advantages

Perfectly conducting objects (PEC) are quite successfully modelled using domain decomposition (DD) methods like the FB Method and its various versions [7]. In

Chapter 5 we saw that, as the object size and dielectric constant increases, the FB algorithm based methods are shown to be incapable of solving scattering problems involving dielectric objects. This is mainly due to the greater wave-effects inside the object, which in turn brings about greater edge effects due to the imaginary edges constructed for partitioning the object [2].

Rather than solving the whole object in a semi-iterative and block-wise way as in the FB methods, the CBFM creates a single *reduced or compressed* matrix that can be inverted using a stationary method. Theoretically, any large object can be solved on a personal computer using the CBFM and a non-iterative solver. Depending on the required accuracy and available computing resources, one can suitably choose the CBFM parameters for a complete solution of objects of any size. The efficiency of the CBF method is independent of the type of basis and testing functions used to discretise the object [46]. The aforementioned advantages over contemporary methods which often need iterative methods makes the CBFM a suitable candidate for scattering problems involving large objects.

Using the CBFM, the linear equation resulting from the MoM with N unknowns can be compressed to another linear equation with M^2 unknowns, where, M is the number of imaginary patches that the CBFM creates to partition the object. Thus, in the CBFM, the size of the resultant linear equation to be solved does not depend on the dimension of the object. It depends only on the number of imaginary sub-domains that the CBFM creates in the object. Even if the object to be modelled is electrically large, the number of patches M that the CBFM creates on the object can be kept at a manageable level. This property enables the CBFM to solve electrically large objects using limited computing resources. In reality, the bigger the scatterer, the higher will be the number of discretisation points used. This, in turn, requires more patches or needs bigger patches covering larger portions of the object. In any case, the user can

fix a tolerance for the accuracy required and select a suitable number of patches of appropriate size so that the reduced matrix can be solved using a stationary method.

This chapter investigates the CBFM's suitability for computing EM fields inside dielectric objects modelled with the EFIE. Results calculated using the CBFM are compared with the analytic solution obtained using the *Mie series* [32], where the later exists. The CBF method is easily parallelisable and most suitable for shared memory parallel computation. This will facilitate the use of massively parallel processors emerging in the market to solve large EM problems quickly on desktop computers.

6.2 The CBF Formulation

The matrix equation resulting from the Method of Moments (MoM) solution method of the EFIE [49] is denoted as:

$$Z\mathbf{J} = \mathbf{R} \tag{6.2.1}$$

Z is the $N \times N$ matrix containing the coefficients of the interaction between different cells in the object [44]. $\mathbf{R}$ is the excitation field vector of size N. $\mathbf{J}$ is the unknown solution vector of size N that contains information related to the total electric field on the object. According to Equation (3.8.1), for electrically large objects, N will be large.

The CBFM begins by dividing the scatterer geometry into several patches as shown in Figure (6.1) The MoM matrix elements that belong to each patch can be easily computed at once and stored separately on the main memory. In Figure (6.1), the object is partitioned into $M = 16$ distinct patches. Now, basis functions that are *characteristic* of each domain are constructed. The *characteristic basis functions* constructed for each patch are of two categories:

- **Primary basis:** These arise from the self-interactions within the same patch, one each for every patch, M primary basis functions in total.

- **Secondary basis:** These account for the mutual coupling of patch i with all other patches. There will be $M-1$ secondary basis functions for each patch, $M^2 - M$ in total.

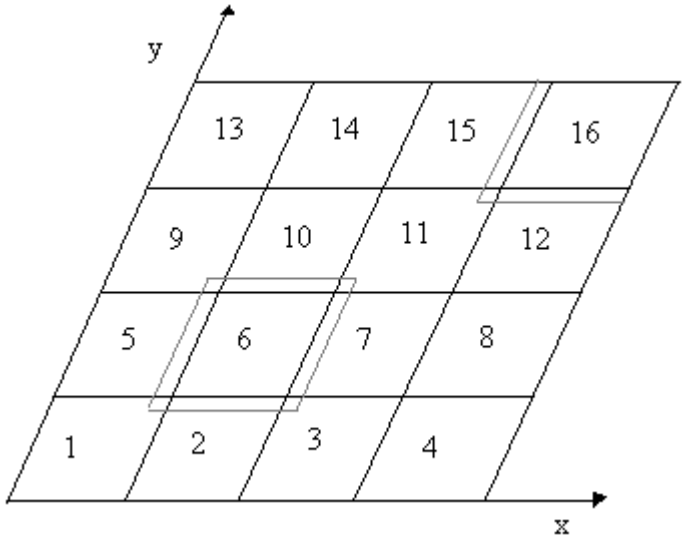

Figure 6.1: *A 2-dimensional rectangular dielectric object divided into* $M = 16$ *patches for constructing the CBFs, note the extended blocks* $i = 6$ *and* 16.

6.2.1 Computation of primary basis functions

Let N_i be the number of unknowns in patch i. Each patch will be extended on all sides to include buffer regions of width Δ_b as shown in Figure (6.1) for patches 6 and 16. For the sake of clarity, consider an object divided into $M = 3$ patches, in Figure (6.2). The coefficient matrix Z is subdivided into 9 blocks. Let N_i^e be the number of unknowns in these extended block-matrices. Elements of Z that belong to the i^{th} domain will be extracted from Z and a new matrix $\tilde{Z}_e^{(i)}$ of dimension $N_i^e \times N_i^e$ shall

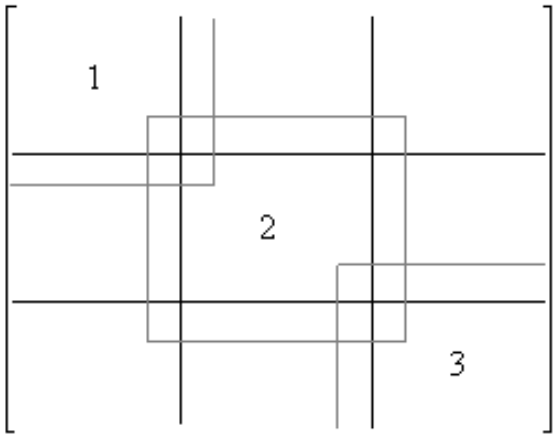

Figure 6.2: *The coefficient matrix Z divided into* 9 *blocks. Note the extension of each block to include a buffer region to form the extended block* $\tilde{Z}_e$.

be stored to the main memory. The primary basis J_i for the i^{th} domain is computed by solving Equation (6.2.2).

$$Z_e^{(i)} J_i^{(i)} = R^{(i)} \tag{6.2.2}$$

$R^{(i)}$ is the $N_i^e \times 1$ dimensional vector taken out from $\mathbf{R}$ that corresponds to block i. Equation (6.2.2) will be solved for $i = 1, \ldots, M$ to find the primary basis for each sub-block.

6.2.2 Computation of secondary basis functions

To construct the secondary basis functions arising from the mutual interaction of block i with other blocks, Equation (6.2.3) will be solved $M - 1$ times.

$$Z_e^{(i)} J_k^{(i)} = R_k^{(i)} \tag{6.2.3}$$

where,

$$R_k^{(i)} = -Z^{(i,k)} J_k^{(k)} \tag{6.2.4}$$

Here, $J_k^{(i)}$ is the k^{th} secondary basis for the i^{th} block. $R_k^{(i)}$ is the excitation vector resulting from the mutual coupling between block i and block k. If there is no overlap between block i and block k, the excitation sub-vector $R_k^{(i)}$ is given by Equation (6.2.4). $Z^{(i,k)}$ is an $N_i^e \times N_k$ matrix formed from the MoM matrix Z. When the i^{th} block shares some unknowns with block k, let $N_{i,k}^c$ be the number of common elements. These elements will be eliminated from $Z^{(i,k)}$ and $J_k^{(k)}$ making them $N_i^e \times (N_k - N_{i,k}^c)$ and $(N_k - N_{i,k}^c) \times 1$ respectively. The excitation vector $R_k^{(i)}$ is still determined using (6.2.4).

6.2.3 Generation of the reduced matrix

M characteristic basis functions constructed for each patch will be ortho-normalised using the modified Gram-Schmidt [14] procedure as described in Section (4.7). The solution $\mathbf{J}$ can be expressed as a linear combination of the ortho-normalised CBFs as given in Equation (6.2.5). Here, $\alpha_k^{(i)}$ are unknown coefficients to be determined. $J_k^{(i)}$ is the k^{th} CBF of block i. Substituting (6.2.5) in (6.2.1) results in an over-determined system of equations given by Equation (6.2.6):

$$\sum_{k=1}^{M} \alpha_k^{(1)} \vartheta_k^{(1)} + \sum_{k=1}^{M} \alpha_k^{(2)} \vartheta_k^{(2)} + \ldots + \sum_{k=1}^{M} \alpha_k^{(M)} \vartheta_k^{(M)} = (\mathbf{R}) \tag{6.2.6}$$

where,

$$\vartheta_k^{(i)} = \left((A_{1,i}) \left(J_k^{(i)} \right) (A_{2,i}) \left(J_k^{(i)} \right) \ldots (A_{M,i}) \left(J_k^{(i)} \right) \right)^T \tag{6.2.7}$$

The over-determined system of equations (6.2.6) with α as unknowns can be compressed to a linear equation (6.2.8) of size $M^2 \times M^2$. This is done by taking the inner

$$\begin{pmatrix} \mathbf{J} \end{pmatrix} = \sum_{k=1}^{M} \alpha_k^{(1)} \begin{pmatrix} (J_k^{(1)}) \\ . \\ . \\ (0) \end{pmatrix} + \sum_{k=1}^{M} \alpha_k^{(2)} \begin{pmatrix} (0) \\ \left(J_k^{(2)}\right) \\ . \\ (0) \end{pmatrix} + \ldots + \sum_{k=1}^{M} \alpha_k^{(M)} \begin{pmatrix} (0) \\ (0) \\ . \\ \left(J_k^{(M)}\right) \end{pmatrix} \tag{6.2.5}$$

product of (6.2.6) on both sides with $\vartheta_q^{(j)\dagger}$, where $\{q, j\} = 1, 2 \ldots, M$ and $\dagger$ denotes the conjugate transpose. Note that M^2 will be termed N_c in further discussions.

$$Z_c \alpha = R_c \tag{6.2.8}$$

N_c, the size of the compressed equation, can be kept low so that it can be solved using a stationary method like LU decomposition.

6.3 Numerical Results

Scattering problems involving infinite cylindrical dielectric objects of various sizes and cross-sectional shapes will be formulated and numerically implemented in the following sections. The field components that are excited by a normally incident TM plane wave are E_z, H_x and H_y. Only E_z, which is invariant in the z direction is computed in this work. A cross-section of the dielectric object can be discretised into N square cells, each with a side of $width = \Delta$. The EFIE formulation in Chapter 3

results in a matrix equation:

$$Z\mathbf{E}^t = \mathbf{E}^i \tag{6.3.1}$$

$\mathbf{E}^i$ is an incident plane-wave [32] and $\mathbf{E}^t$ is the unknown total electric field. The matrix elements in Z are given by:

$$Z_{mn} = C_1 J_1(k_0 a) J_1(k_0 R_{mn}) - \jmath H_0(k_0 R_{mn}) \qquad m \neq n \tag{6.3.2}$$

$$Z_{mm} = 1 + \frac{C_1}{k_0} J_1(k_0 a) - \jmath H_1(k_0 a) - \frac{\jmath 4}{k_0^2} \qquad m = n \tag{6.3.3}$$

$$\text{where,} \quad C_1 = \frac{\jmath 2\pi a \gamma}{4}$$

$$\text{and} \quad \gamma = k^2(n) - k_0^2$$

In Equations (6.3.2) and (6.3.3), R_{mn} is the distance between cells m and n, $\eta = \sqrt{\mu_0/\epsilon_r}$, H_0^2 is the zeroth order Hankel function of the second kind, $k(n)$ is the wave-number at cell n and a is the radius of a circle with equivalent area Δ^2. The EFIE solution will be compared with the Mie-series solution [32]. A varying buffer size of width Δ_b will be used in all test cases. $\mathbf{E}^i$ is a TM plane wave with frequency $f = 300 - 800$ MHz in the $+x$ direction with a polarisation parallel to the cylinder's axis [32]. D_λ is the number of discretisation cells used per wavelength.

6.3.1 Validation of the CBFM

In order to validate the method, a homogeneous dielectric cylinder with circular cross-section and $\epsilon_r = 2$ and $\sigma = 0.8$ is considered. The object is impinged by a TM plane wave of frequency $f = 300$ MHz. The radius of the object is $r = 0.5\lambda_0$ and $D_\lambda = 10$. The interaction matrix Z is created using Equations (6.3.2) and (6.3.3). The MoM matrix (6.3.1) is directly solved using the GMRES method [14] with a tolerance $\delta_g \leq 1e-05$.

The CBFM with $M = 16$ is applied to compute the total field $\mathbf{E}^t$ produced on the object. A buffer size of $\Delta_b = \Delta$ is used. Surface plots of the magnitude of E_z^t computed using the Mie series, EFIE-MoM and CBFM are shown in Figures (6.3), (6.4) and (6.5) respectively. A comparison of all the above solutions at $z = y = 0$ is plotted in Figure (6.6). A comparison of these solutions at $z = x = 0$ is plotted in Figure (6.7). Figures (6.6) and (6.7) show excellent agreement between the analytical solution and the CBFM solutions.

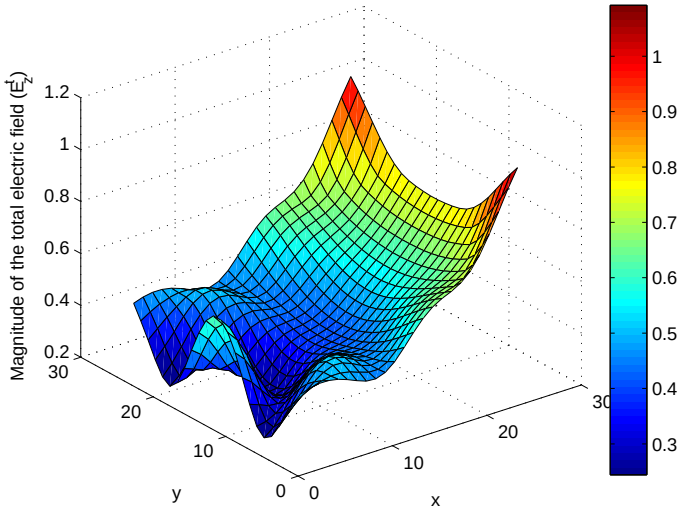

Figure 6.3: *A surface plot of the magnitude of E_z^t on the object computed using the Mie series, $D_\lambda = 10$.*

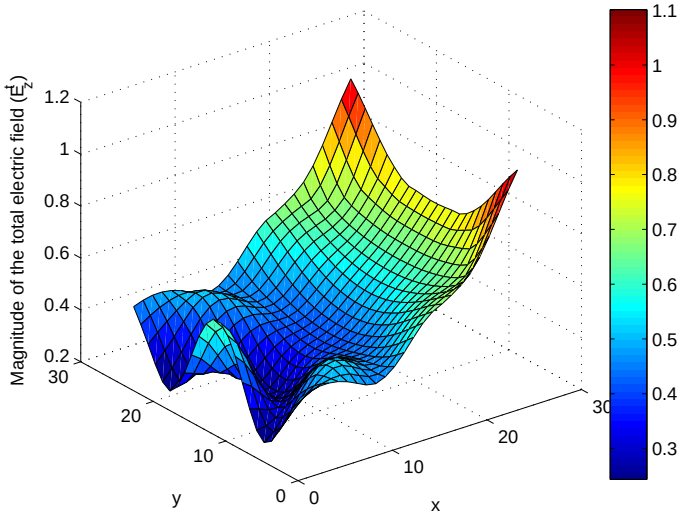

Figure 6.4: *A surface plot of the magnitude of E_z^t on the object computed using the EFIE, $D_\lambda = 10$.*

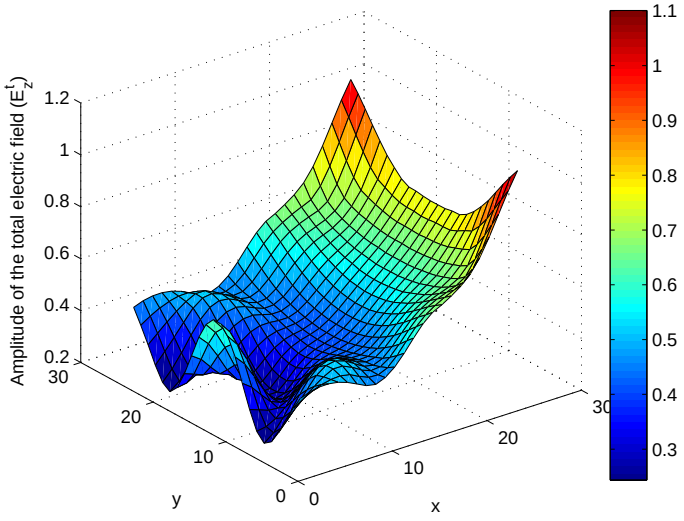

Figure 6.5: *A surface plot of the magnitude of E_z^t on the object computed using the CBFM with $M = 16$ patches, $D_\lambda = 10$.*

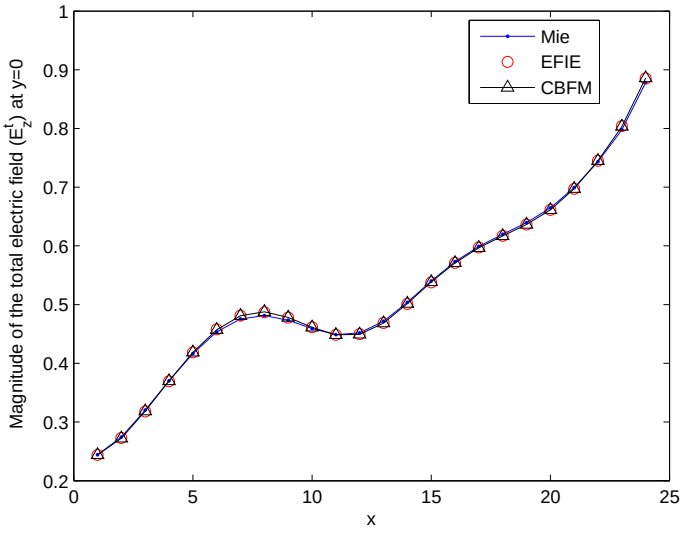

Figure 6.6: *A comparison of the magnitude of the E_z^t on the object at $y = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

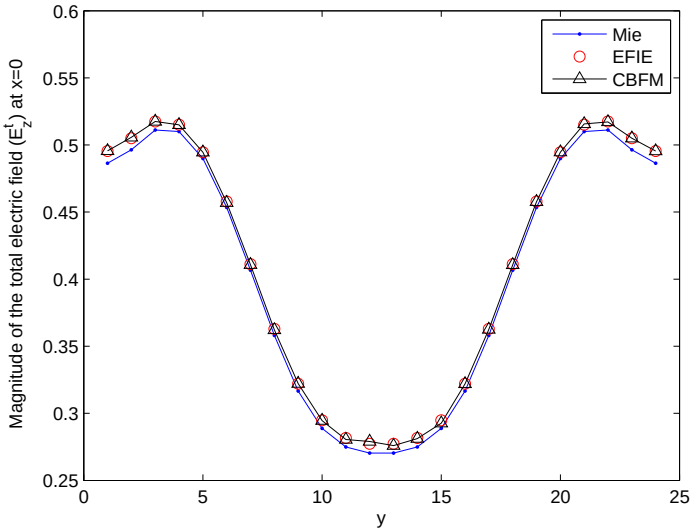

Figure 6.7: *A comparison of the magnitude of the E_z^t on the object at $x = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

Figures (6.3) - (6.7) show that the CBFM can be used for the analysis of dielectric

cylinders with circular cross-section. In next sections, we consider several test cases with different cross-sectional shapes and sizes.

6.3.2 Homogeneous Infinite Circular cylinder

A number of infinite homogeneous dielectric cylinders with different electrical properties and sizes are examined. In all cases, TM plane waves of frequency in the range $f = 300 - 800$ MHz are used as incident field. The radius of the object is varied from $r = 0.5\lambda$ to 2λ. The linear equation (6.3.1) is directly solved using the GMRES method [14] with an error $\delta_g \leq 1e-05$. The CBFM with $M = 16$ patches is applied to find the total field $\mathbf{E}^t$ produced on the object. A buffer size of Δ_b is used in all cases. Table (6.1) summarises results from all of the test cases.

Test Case	f(MHz)	R	ϵ	σ	D_λ	N	Δ_b	κ	δ
1	300	λ	2	0.8	15	576	2Δ	5.114	0.0084
2	300	λ	5	0.8	10	576	2Δ	44.635	0.0052
3	300	2λ	2	0.8	10	1024	1Δ	10.185	0.0141
4	300	2λ	2	0.8	10	1024	2Δ	10.148	0.0100
5	300	2λ	2	0.8	10	1024	3Δ	10.135	0.0084
6	300	2λ	2	0.8	10	1024	5Δ	10.135	0.0064

Table 6.1: *Summary of results from test cases involving scattering from homogeneous circular infinite cylinders with various parameters. R is the radius of the object, f is the frequency of the incident field, D_λ is the number of discretisation cells used per wavelength, Δ_b is the buffer width used, N_c is the size of the reduced matrix Z_c, κ is the condition number of Z_c and δ is the normalised error in satisfying Equation (6.3.1).*

In Figures (6.8)-(6.20), magnitude of E_z^t computed using the CBFM is compared with the Mie series and EFIE solutions. It can be observed that as the buffer size Δ_b is increased, CBFM solutions get closer to the EFIE solution. This confirms the fact that the CBFM can be tuned to achieve accurate solutions for EM computations involving dielectric objects. The main computational constraint in the use of CBFM for large objects is the size of the reduced equation that needs to be inverted. This in turn restricts the number of patches that has to be created to decompose the scattering domain. If the object gets larger, one can keep M constant considering the available RAM size on the computer and increase the buffer size appropriately. Note that this does not add any significant burden to the overall computation time.

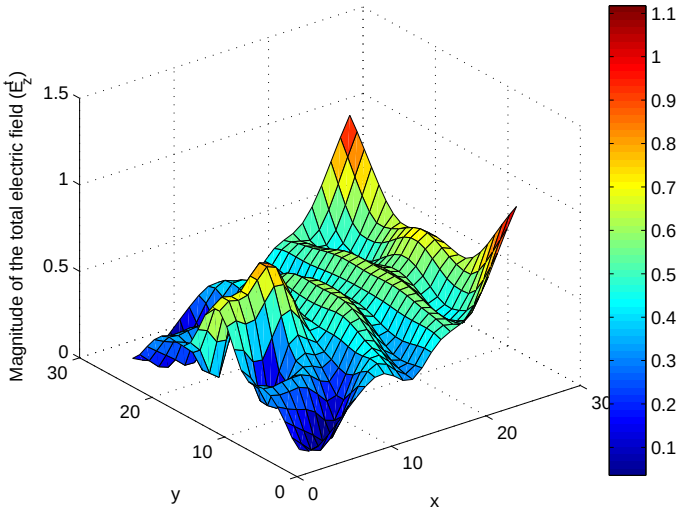

Figure 6.8: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 2 at $z = 0$ computed using the Mie series.*

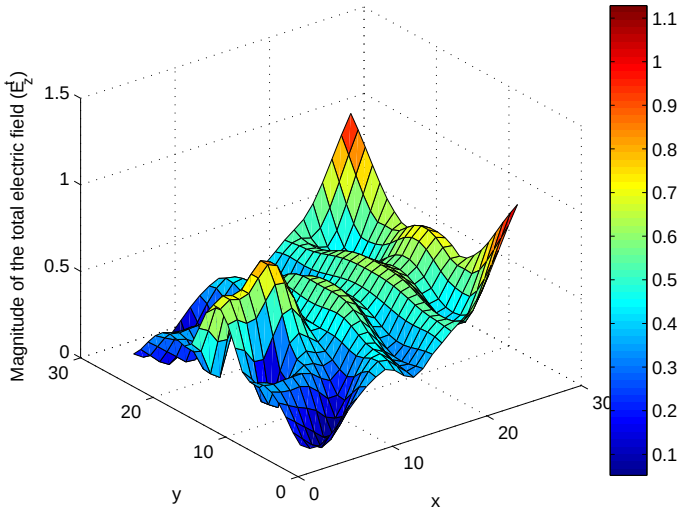

Figure 6.9: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 2 at $z = 0$ computed using the EFIE-MoM approach.*

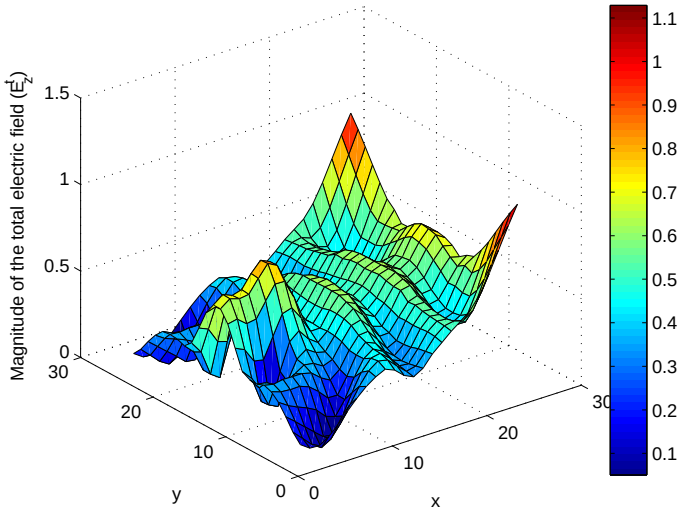

Figure 6.10: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 2 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

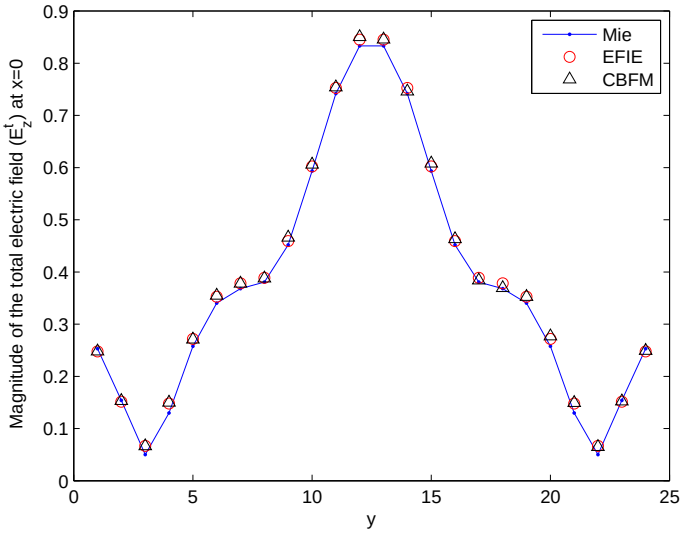

Figure 6.11: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 2 at $z = x = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

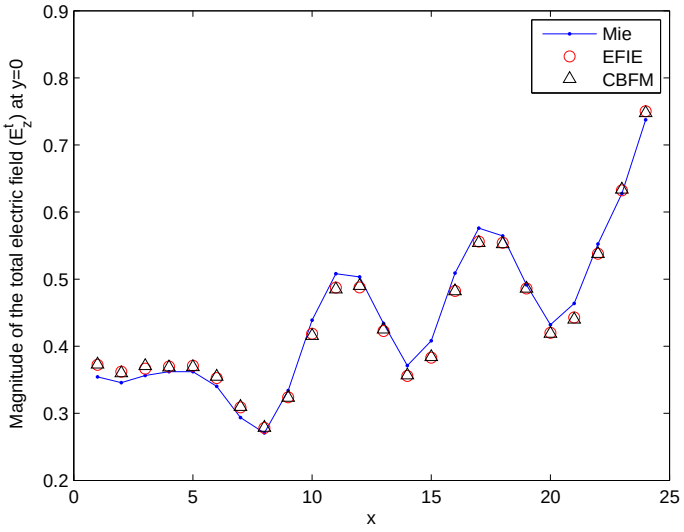

Figure 6.12: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 2 at $z = y = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

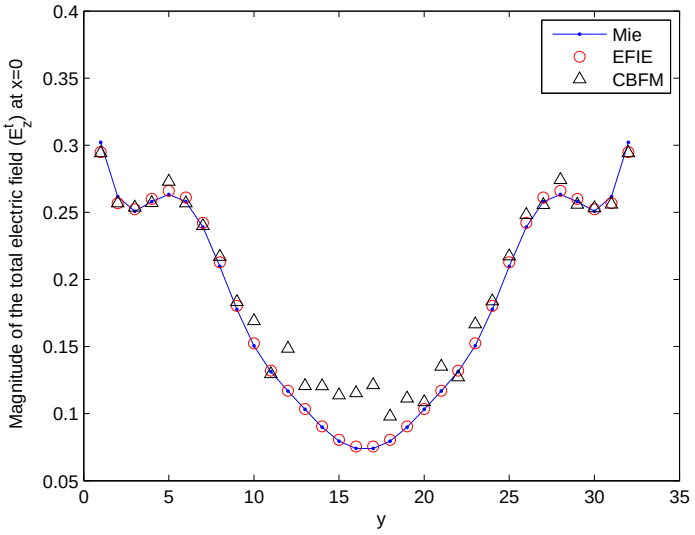

Figure 6.13: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = x = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

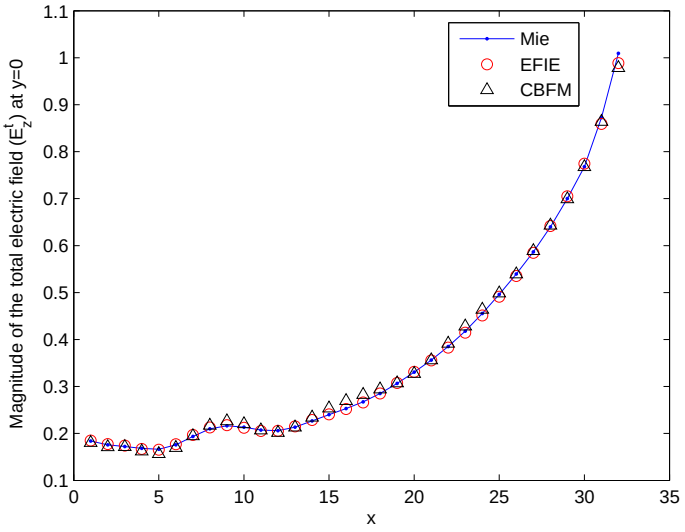

Figure 6.14: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = y = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

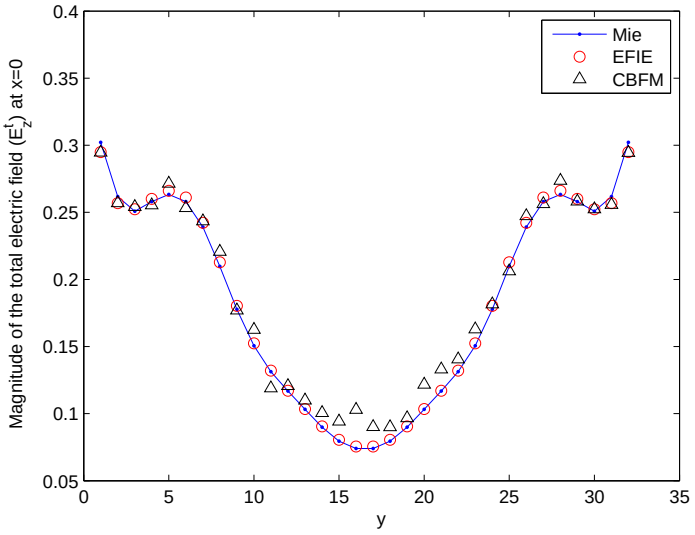

Figure 6.15: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 4 at $z = x = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

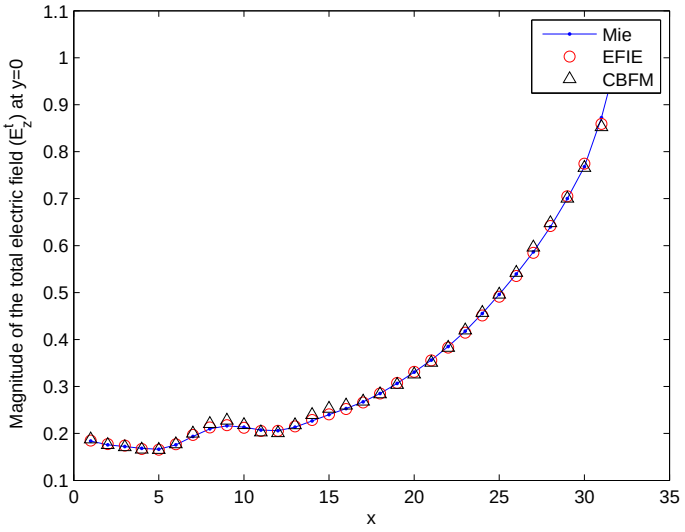

Figure 6.16: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 4 at $z = y = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

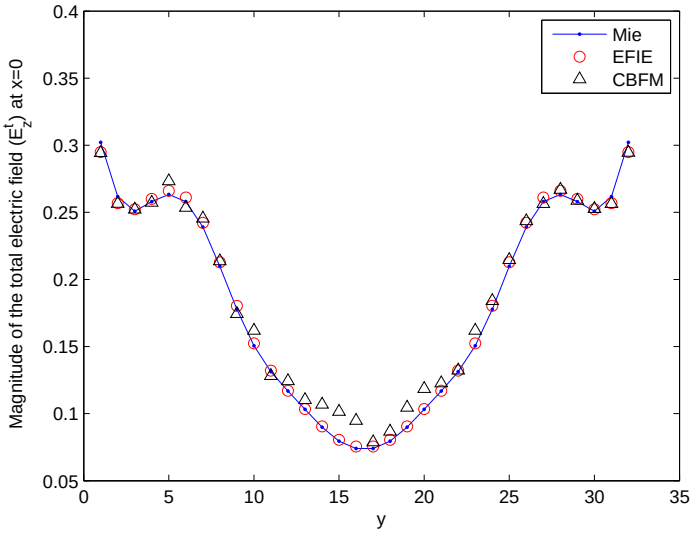

Figure 6.17: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 5 at $z = x = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

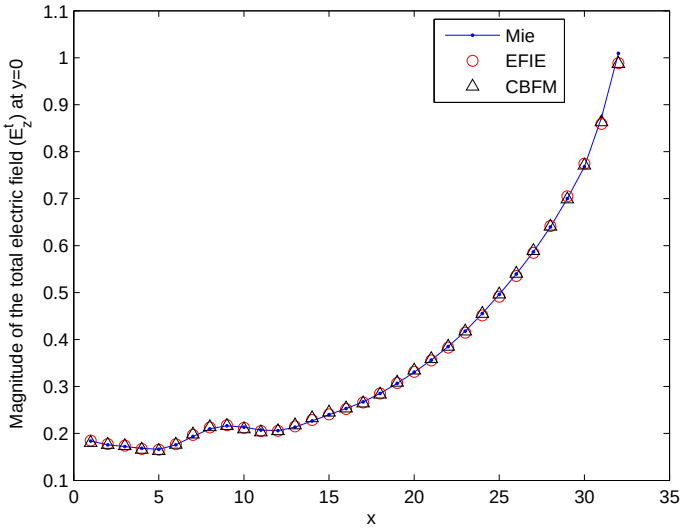

Figure 6.18: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 5 at $z = y = 0$ computed using the Mie series, EFIE-MoM and CBFM.*

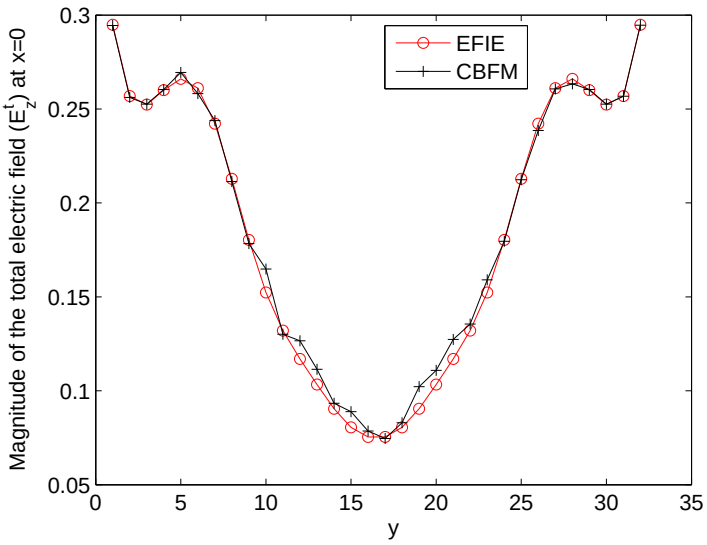

Figure 6.19: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 6 at $z = x = 0$ computed using the EFIE-MoM and CBFM.*

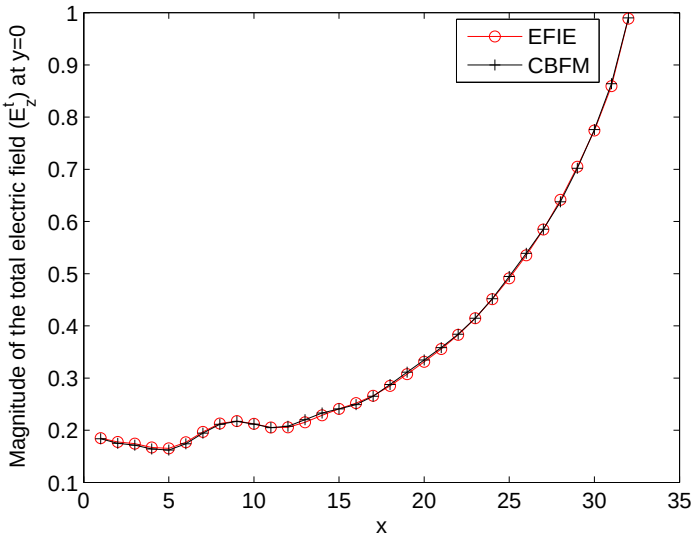

Figure 6.20: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 6 at $z = y = 0$ computed using the EFIE-MoM and CBFM.*

For test case 2, the CBFM results show excellent agreement with the analytical Mie series and EFIE-MoM results. For test case 3, the size of the object is increased. In this case, the results are not accurate especially at $z = x = 0$ in Figure (6.13). In test case 4, we increase the buffer width to 2Δ and the normalised error comes down. This is evident in Figure (6.15). With a buffer width of 3Δ, in test case 5 results are more accurate as shown in Figure (6.17) and (6.18). In Figures (6.19) and (6.20), we see excellent agreement between the CBFM result and the EFIE-MoM solution.

These results confirm that the CBFM can be tuned for the electromagnetic analysis of homogeneous dielectric cylinders. In Section (6.3.3), infinite inhomogeneous dielectric cylinders with various dielectric properties and sizes are analysed using the CBFM.

6.3.3 Infinite Inhomogeneous Circular Cylinder

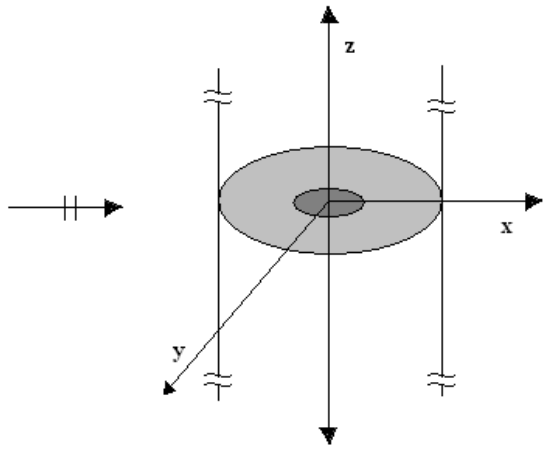

Figure 6.21: *An infinite inhomogeneous cylinder whose electrical properties are invariant in the z-direction. Shaded region is the cross-section of the cylinder at $z = 0$. The radius of the inner cylinder is denoted as r_i and that of the outer cylinder is denoted as r_o. A plane wave in the $+x$-direction with frequency f impinges the object normally.*

A number of infinite inhomogeneous cylinders of circular cross-section like that shown in Figure (6.21) are impinged by a TM plane wave of frequency f. The outer radius of the object is denoted r_o and the inner radius is denoted by r_i. The dielectric properties of the outer and inner cylinders are ϵ_{ro}, ϵ_{ri} and σ_o ans σ_i, respectively. Dielectric properties are varied gradually for each test case. The interaction matrix Z is created using Equations (6.3.2) and (6.3.3). In all cases, the linear equation (6.3.1) is iteratively solved using the GMRES method with a tolerance $\delta_g \leq 1e-05$. The CBFM with $M = 16$ patches is applied to find the total electric field $\mathbf{E}^t$ produced on the object. A buffer size of Δ_b is used. Table (6.2) summarises results from all the test cases.

Test case	R	f (MHz)	ϵ_{r1}	σ_{r1}	ϵ_{r2}	σ_{r2}	N	Δ_b	κ	δ
1	0.5λ	300	5	0.8	8	2.8	576	1Δ	51.328	0.0048
2	0.5λ	300	5	0.8	8	2.8	576	3Δ	51.318	0.0039
3	λ	800	5	0.8	8	2.8	2304	1Δ	104.487	0.0472
4	λ	800	5	0.8	8	2.8	2304	3Δ	107.344	0.0668
5	λ	300	3	0.8	5	2.8	1296	3Δ	36.154	0.0243
6	λ	300	3	0.8	5	2.8	1296	5Δ	36.352	0.0175
7	λ	300	3	0.8	5	2.8	1296	6Δ	36.090	0.0141

Table 6.2: *Summary of results from various test cases involving scattering from inhomogeneous circular infinite cylinders with various dielectric parameters. R is the radius of the object, f is the frequency of the incident field, D_λ is the number of discretisation cells used per wavelength, Δ_b is the buffer width used, N_c is the size of the reduced matrix Z_c, κ is the condition number of Z_c and δ is the normalised error in satisfying 6.3.1.*

In Figures (6.26)-(6.41) the magnitude of E_z^t computed using the CBFM is compared with solutions computed using the Mie series and EFIE-MoM. It can be observed that as the buffer size Δ_b is increased, CBFM solutions gets closer to the EFIE solution. This in turn confirms the fact that the CBFM can be tuned to achieve accurate solutions for electromagnetic scattering problems involving inhomogeneous objects. If the object gets larger, one can keep M constant considering the available RAM size on the computer and increase the buffer size appropriately. Note that this does not add any significant burden to the over all computation time.

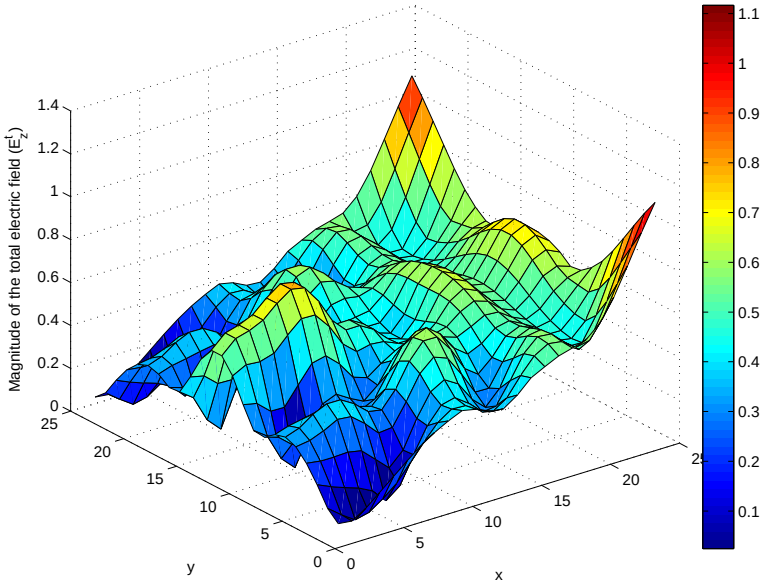

Figure 6.22: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 1 at $z = 0$ computed using the EFIE.*

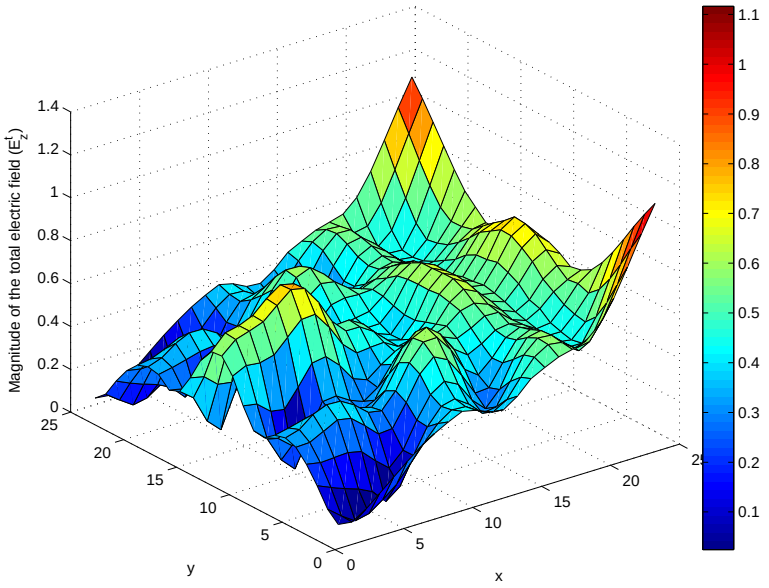

Figure 6.23: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 1 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

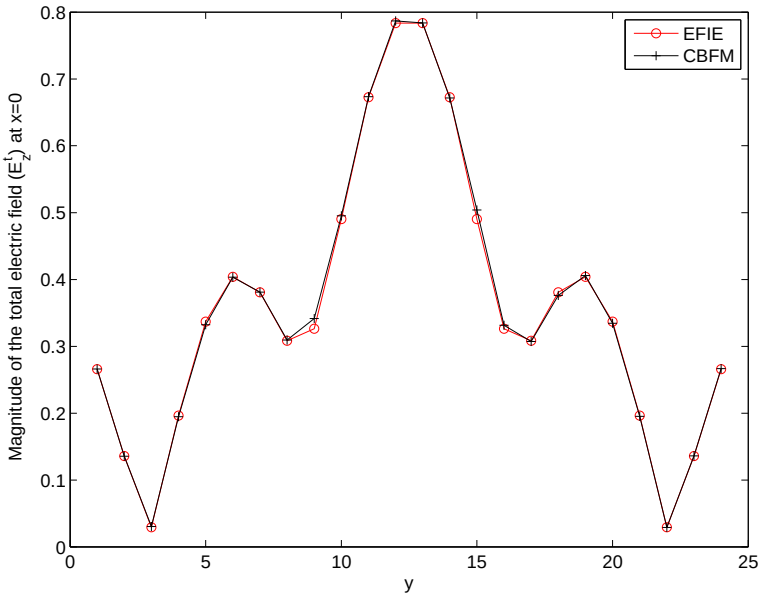

Figure 6.24: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 1 at $z = x = 0$ computed using the EFIE-MoM and CBFM.*

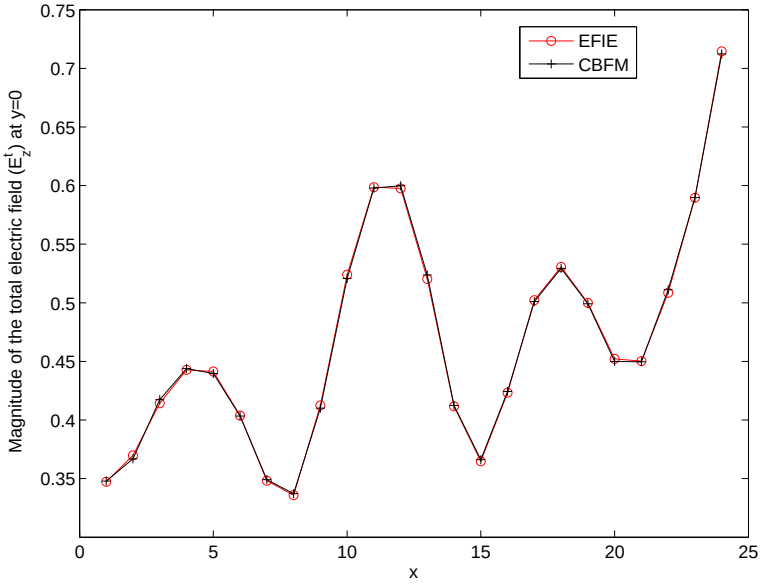

Figure 6.25: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 1 at $z = y = 0$ computed using the EFIE-MoM and CBFM.*

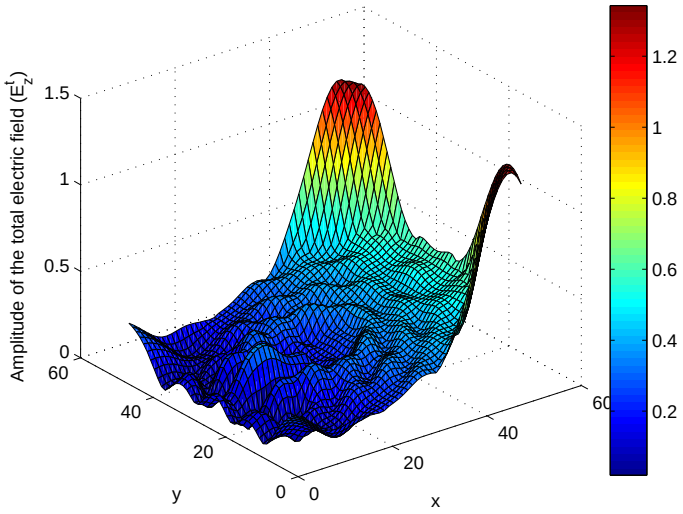

Figure 6.26: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = 0$ computed using the EFIE.*

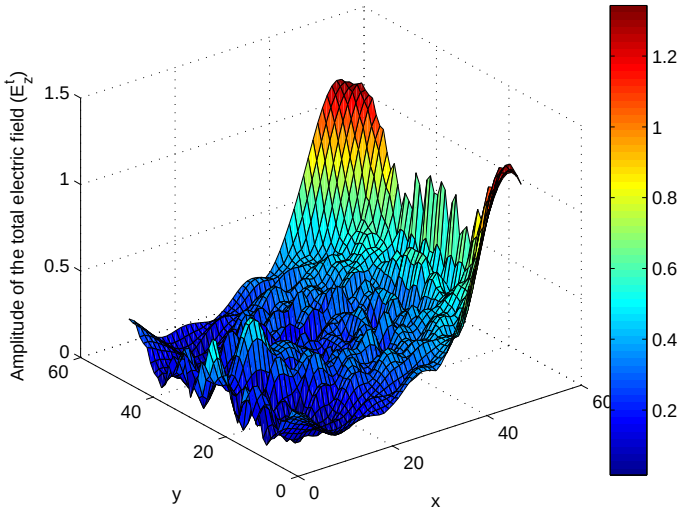

Figure 6.27: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

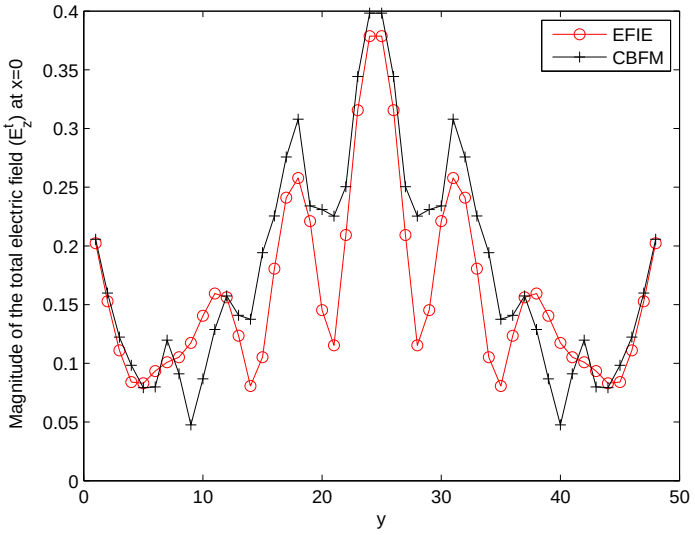

Figure 6.28: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM.*

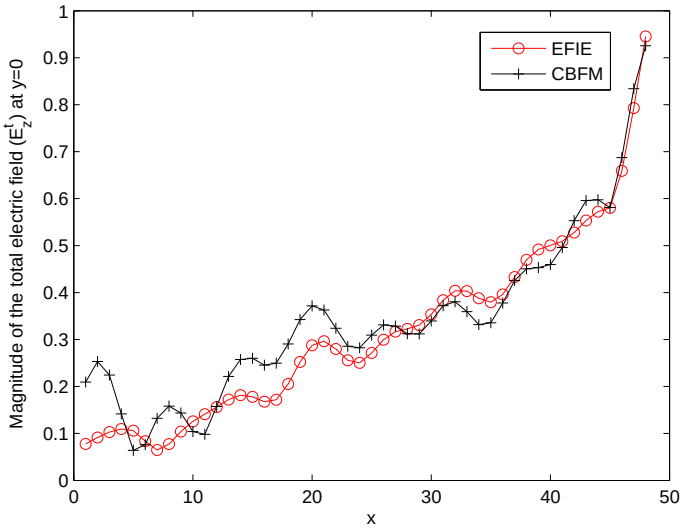

Figure 6.29: *A comparison of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM.*

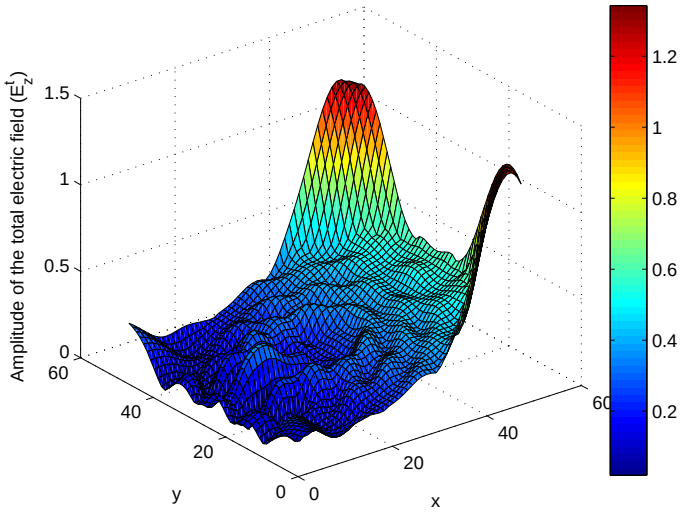

Figure 6.30: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 4 at $z = 0$ computed using the EFIE.*

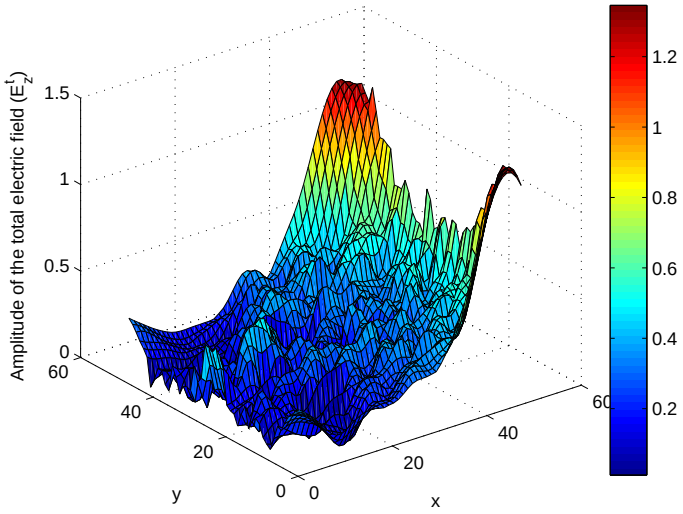

Figure 6.31: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 4 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

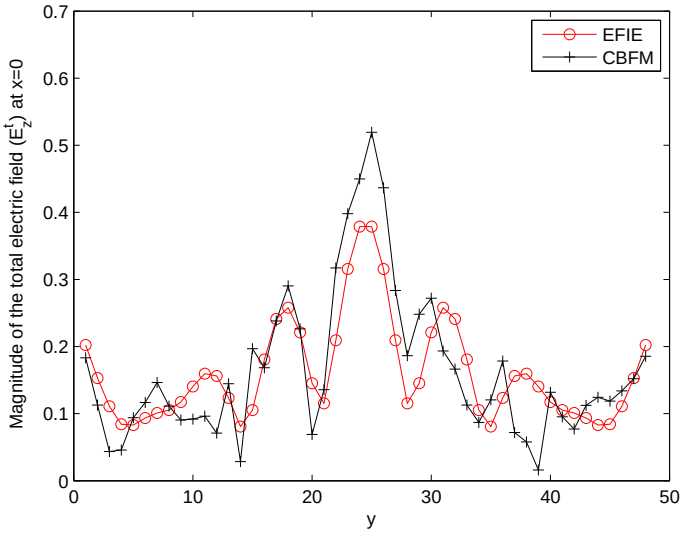

Figure 6.32: *A comparison of the magnitude of E_z^t on the object in test case 4 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM.*

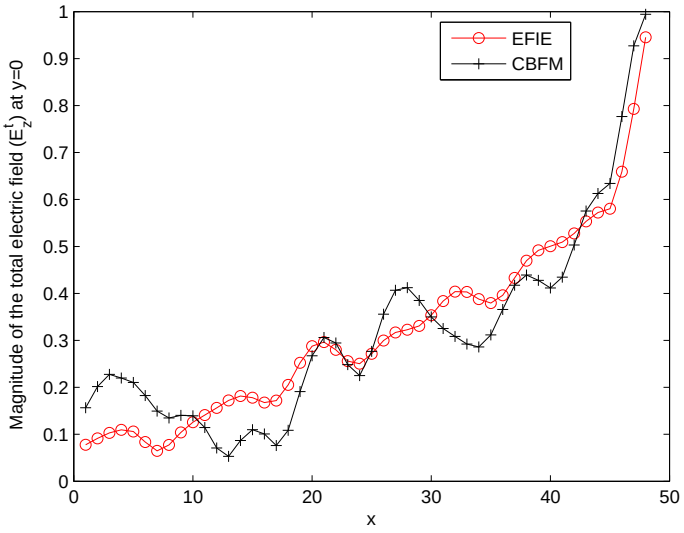

Figure 6.33: *A comparison of the magnitude of E_z^t on the object in test case 4 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM.*

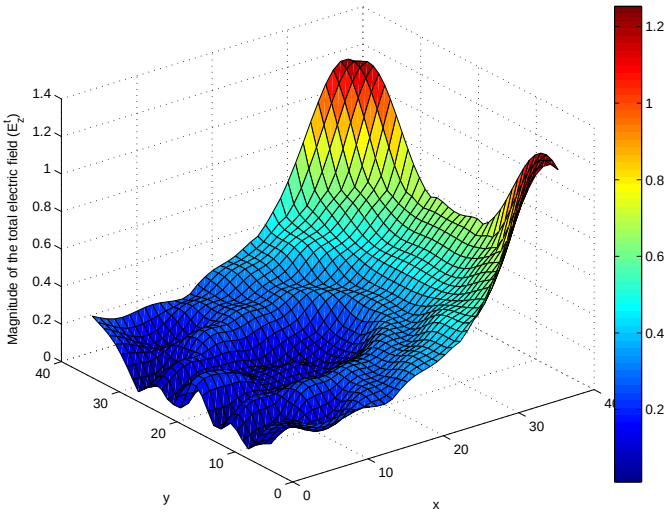

Figure 6.34: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 5 at $z = 0$ computed using the EFIE.*

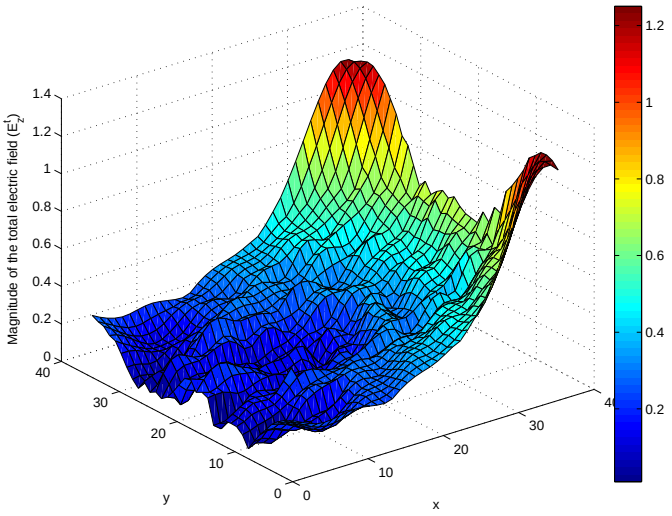

Figure 6.35: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 5 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

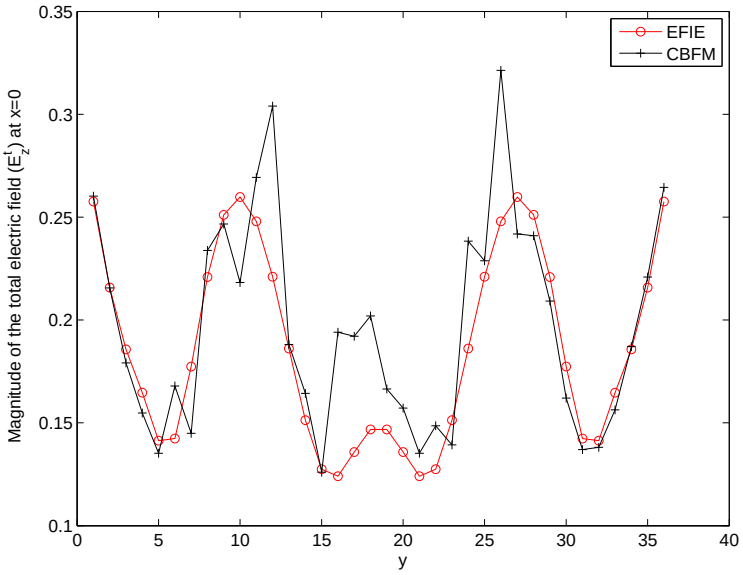

Figure 6.36: *A comparison of the magnitude of E_z^t on the object in test case 5 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM.*

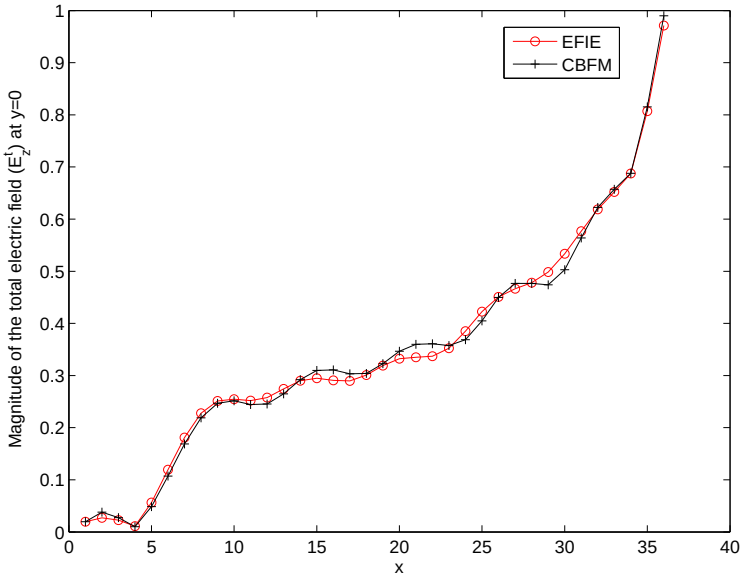

Figure 6.37: *A comparison of the magnitude of E_z^t on the object in test case 5 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM.*

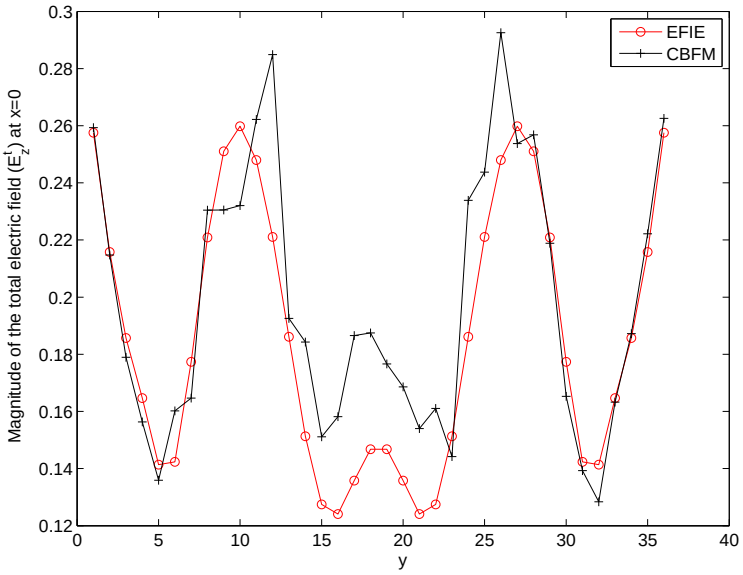

Figure 6.38: *A comparison of the magnitude of E_z^t on the object in test case 6 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM. A buffer width of $\Delta_b = 5\Delta$ is used.*

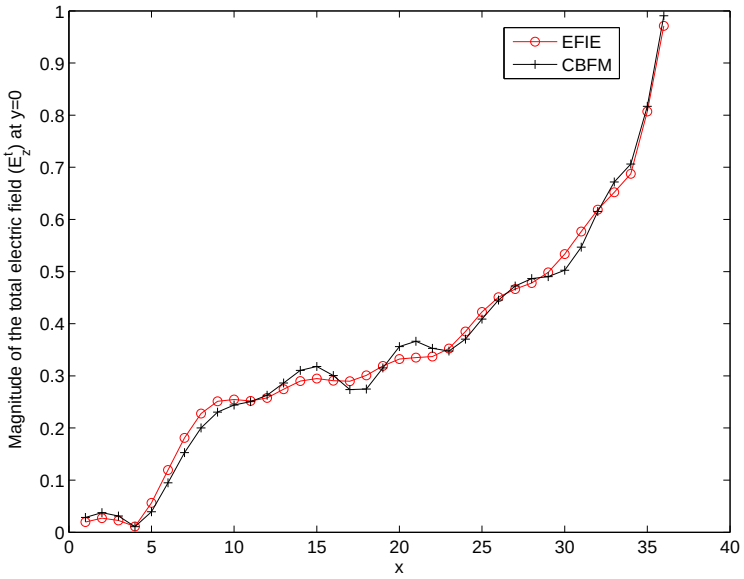

Figure 6.39: *A comparison of the magnitude of E_z^t on the object in test case 6 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM. A buffer width of $\Delta_b = 5\Delta$ is used.*

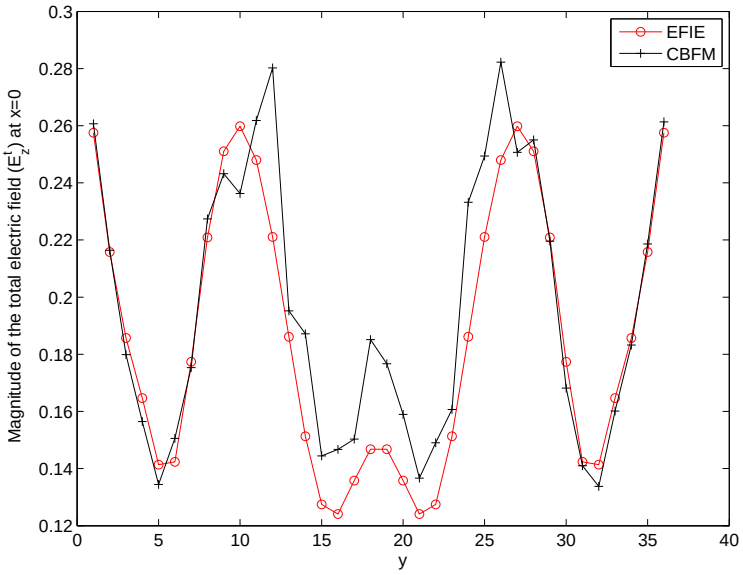

Figure 6.40: *A comparison of the magnitude of E_z^t on the object in test case 7 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM. A buffer width of $\Delta_b = 6\Delta$ is used.*

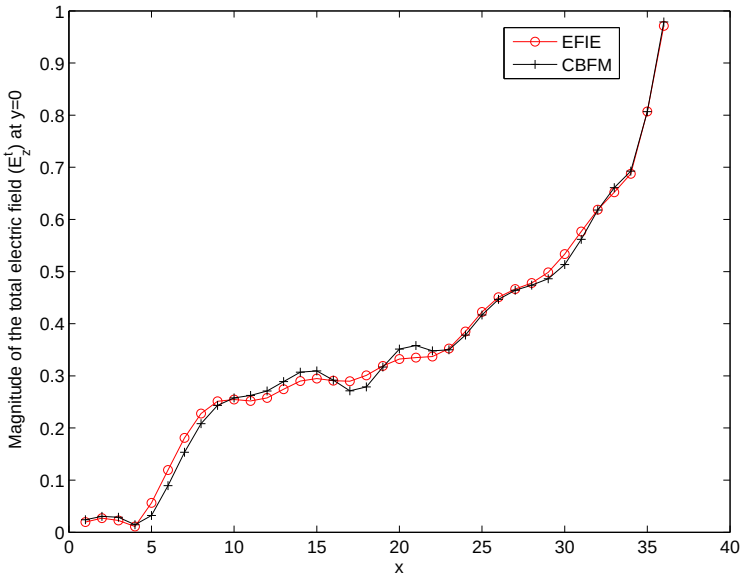

Figure 6.41: *A comparison of the magnitude of E_z^t on the object in test case 7 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM. A buffer width of $\Delta_b = 6\Delta$ is used.*

Table (6.2) shows that the CBFM can be very well tuned to achieve accuracy for DD based computation using the CBFM. It is clearly seen in Figures (6.26)-(6.41) that as the buffer width is increased the CBFM solutions get closer to the EFIE-MoM solutions.

6.3.4 Infinite Square Cylinder

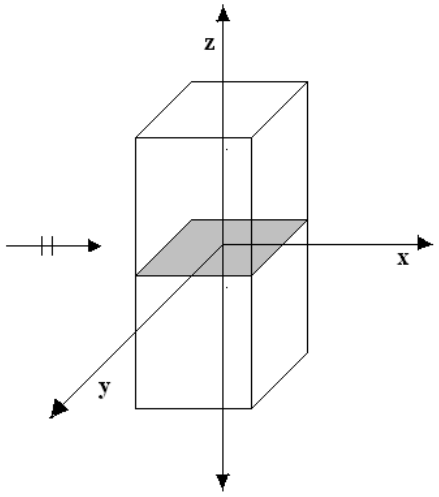

Figure 6.42: *An infinite homogeneous square cylinder whose electrical properties are invariant in the z-direction. The shaded region is the cross-section of the cylinder at $z = 0$. The width of the cylinder is denoted as a. A plane wave in the $+x$-direction with frequency f illuminates the object normal to the surface of the cylinder.*

An infinite homogeneous dielectric square cylinder with side $a = 2\lambda$ and $\epsilon_r = 5 - \jmath 0.8$ is analysed in this section. The object is impinged with a normally incident TM plane wave at frequency $f = 300$ MHz. The CBFM is used to compute the total electric field E_z^t at the cross-section of the object at $z = 0$. The number of discretisation cells used per wavelength is $D_\lambda = 10$. The cross-section at $z = 0$ is discretised into N cells using the pulse basis functions. The MoM matrix equation resulting from

the EFIE-MoM formulation is iteratively solved using the GMRES method with a tolerance $\delta^i \leq 1e-5$.

The CBFM with $M = 16$ is applied to find the scattered field $\mathbf{E}^t$ produced on the object. A buffer size of Δ_b is used in all test cases. The magnitude of $\mathbf{E}^t$ computed with the GMRES method is compared with the CBFM solution in the following figures.

Test case	a	f (MHz)	ϵ_r	σ	N	Δ_b	κ	δ
1	λ	300	5	0.8	576	1Δ	39.3323	0.0079
2	λ	300	5	0.8	576	2Δ	39.3596	0.0071
3	λ	300	5	0.8	576	3Δ	39.3831	0.0062
4	2λ	300	5	0.8	2304	1Δ	84.5024	0.0704
5	2λ	300	5	0.8	2304	2Δ	84.8191	0.0717
6	2λ	300	5	0.8	2304	3Δ	84.8696	0.0667
7	2λ	300	5	0.8	2304	6Δ	83.6334	0.0534
7	2λ	300	5	0.8	2304	6Δ	87.3086	0.0400

Table 6.3: *Summary of results from various test cases involving scattering from homogeneous square infinite cylinders with various dielectric properties and sizes. a is the width of the object, f is the frequency of the incident field, D_λ is the number of discretisation cells used per wavelength, Δ_b is the buffer width used, $N_c = 256$ is the size of the reduced matrix Z_c, κ is the condition number of Z_c and δ is the normalised error in satisfying 6.3.1.*

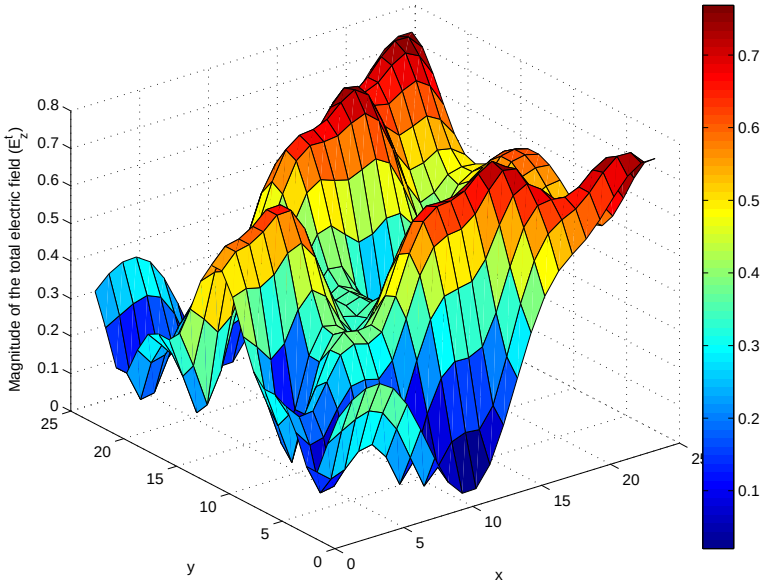

Figure 6.43: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = 0$ computed using the EFIE.*

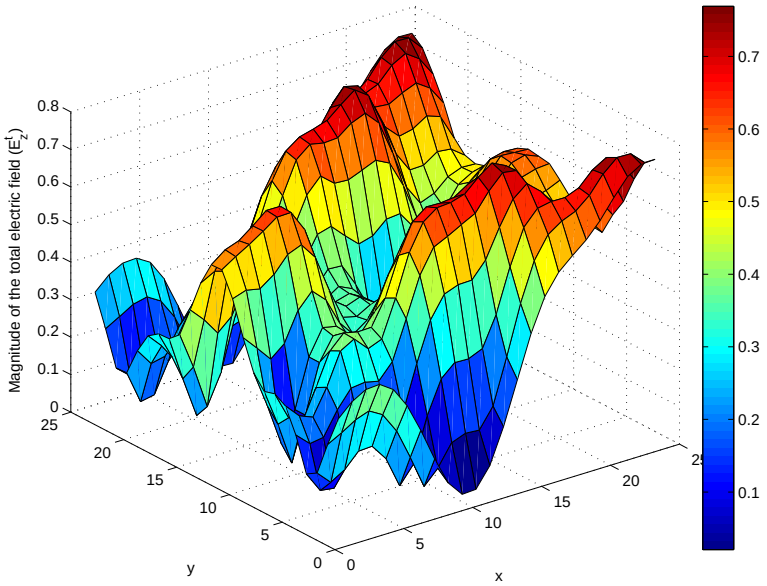

Figure 6.44: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 3 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

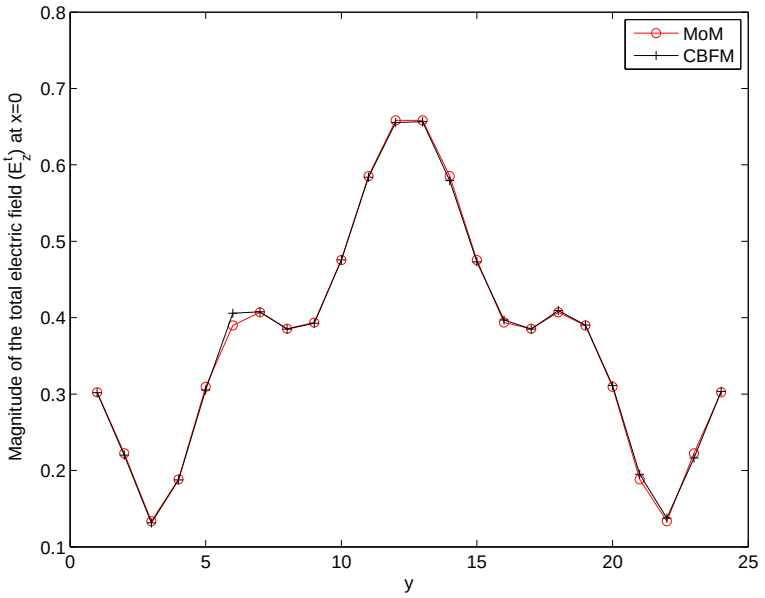

Figure 6.45: *A comparison of the magnitude of E_z^t on the object in test case 3 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM.*

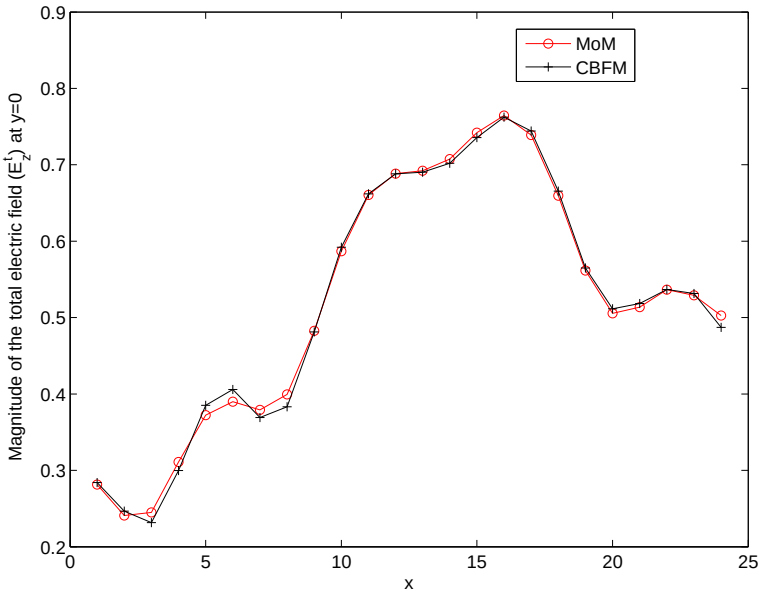

Figure 6.46: *A comparison of the magnitude of E_z^t on the object in test case 3 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM.*

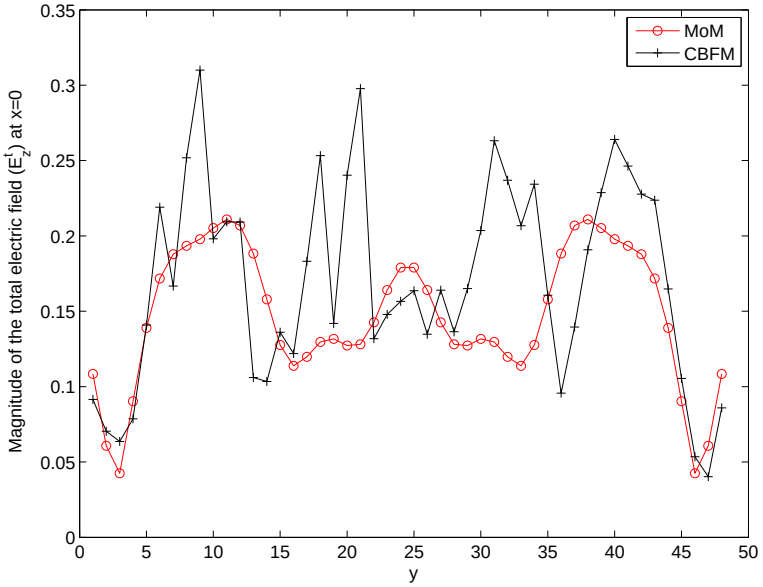

Figure 6.47: *A comparison of the magnitude of E_z^t on the object in test case 7 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM.*

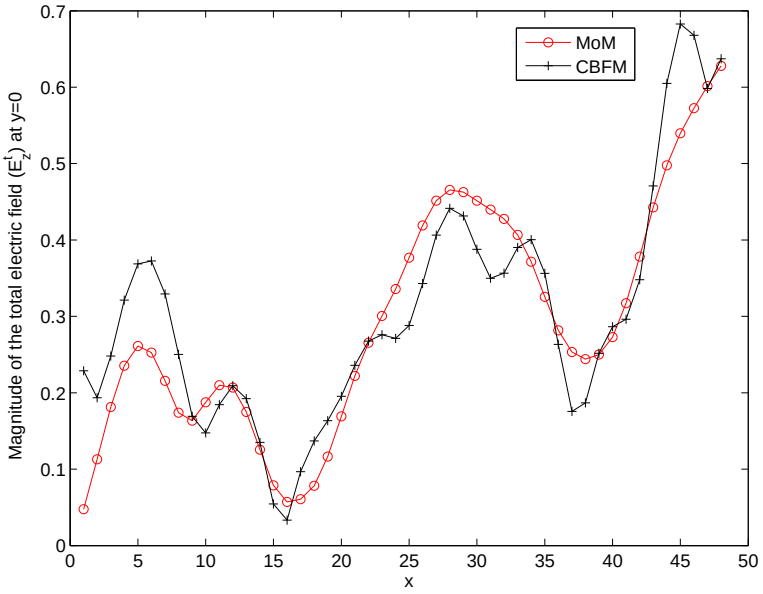

Figure 6.48: *A comparison of the magnitude of E_z^t on the object in test case 7 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM.*

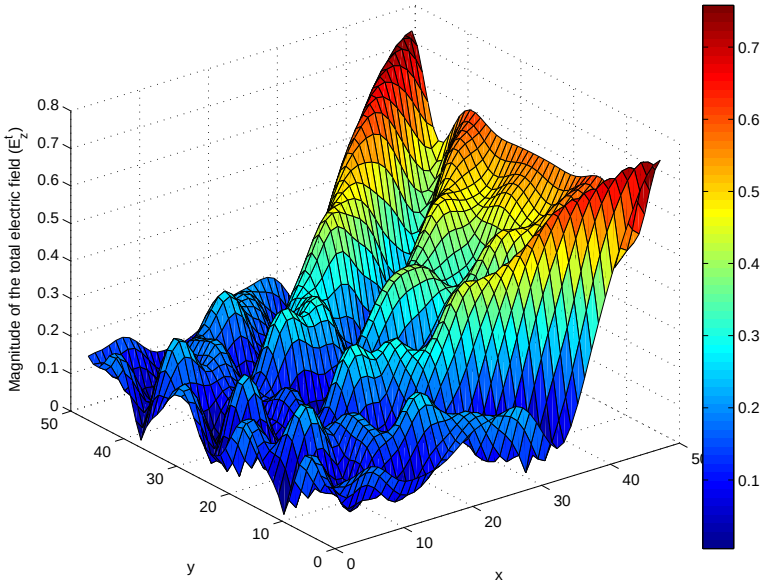

Figure 6.49: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 8 at $z = 0$ computed using the EFIE.*

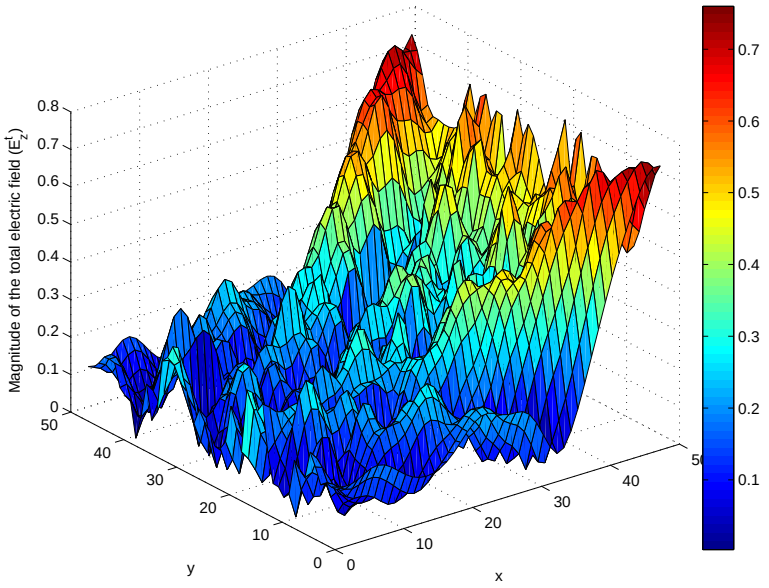

Figure 6.50: *A surface plot of the magnitude of E_z^t on the cross-section of the object in test case 8 at $z = 0$ computed using the CBFM with $M = 16$ patches.*

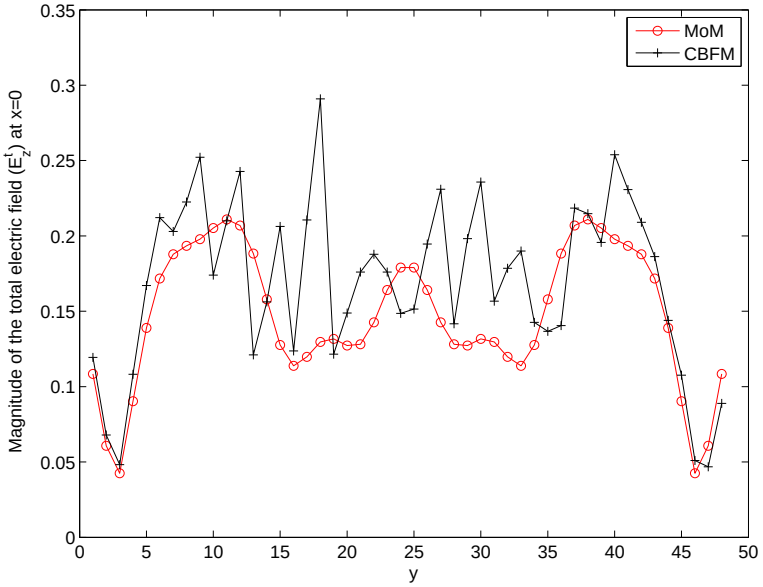

Figure 6.51: *A comparison of the magnitude of E_z^t on the object in test case 8 at $z = 0$ and $x = 0$ computed using the EFIE-MoM and CBFM.*

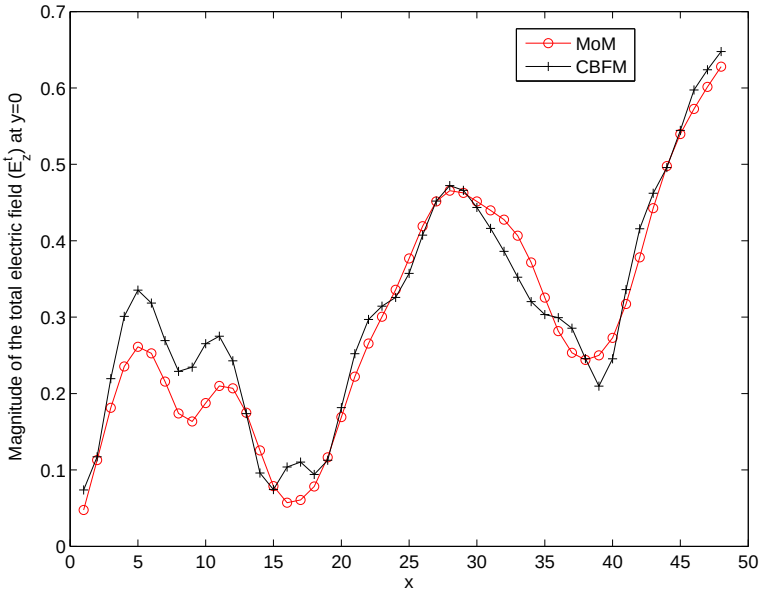

Figure 6.52: *A comparison of the magnitude of E_z^t on the object in test case 8 at $z = 0$ and $y = 0$ computed using the EFIE-MoM and CBFM.*

Test cases in Table (6.3) show that the CBFM can be used to analyse homogeneous

square infinite cylinders with various dielectric properties and sizes. In Figures (6.47) and (6.48) even though a buffer width of $\Delta_b = 6\Delta = a/8$ is used we are unable to achieve good accuracy. But it can be clearly seen in Figures (6.49) - (6.52) that further increase in the buffer width is improving the accuracy.

6.4 Chapter Summary

Results in Sections (6.3.1) - (6.3.4) show that the CBFM can be successfully used for the analysis of dielectric objects. The only disadvantage in this method is that the required buffer width keeps on increasing as the electrical size of the object increases. Total electric field $\mathbf{E}_z^t$ computed on the cross-section of test cases using the CBFM show good agreement with solutions obtained using the Mie series and MoM-EFIE approach. Our aim in this chapter was to confirm that the CBFM can be used to analyse scattering problems involving dielectric objects of any size even if the computational resources are limited. That is, one can use CBFM even if an object's MoM discretisation results in a matrix equation that cannot be loaded to the RAM of a desktop computer. In our test cases, we fixed the size of the reduced matrix equation that needs to be solved during the CBFM. The size of this matrix equation depends only on the the number of imaginary patches created to decompose the object.

It was seen in all test cases that one has to increase the buffer width Δ_b as the object becomes electrically large, either due to the physical size or with an increase in dielectric constant. Even though this was easily done in most of the test cases presented here, it will be advantageous to avoid this requirement as it would sometimes require buffer sizes of the order of the magnitude of the patch width. In the next chapter, a modification of the original CBFM is done avoiding the use of buffer regions.

Chapter 7

The CBFM Tailored for the EM Analysis of Dielectric Objects

In this chapter, we propose a novel modification of the CBFM by using a new type of characteristic basis function. The new CBFs will serve as a basis set that does not depend on the angle of incidence. They are computed using a spectrum of plane-waves in several possible directions of incidence. The new CBFs are created for each patch treating them as separate isolated objects. An investigation is done into using the new CBFs to avoid the use of buffer regions. The new CBFs span a subspace for representing the unknown total field. Results confirm that the method is suitable for the EM analysis of dielectric objects and is free from the shortcomings of previous versions of the CBFM.

7.1 Introduction

In this new version of the CBFM, which is specially tailored for scattering problems involving dielectric objects, initially, we do not include buffer regions as in conventional versions of the CBFM [46]. Instead, while decomposing the object, smaller patches (strips) of the size of the order $\lambda/10$ are sandwiched in between bigger patches. The bigger patches are a few times wider than the sandwiched patches. For small objects,

we would be able to keep all or most of the CBFs in all patches and in that case all patches can be of the same width.

A set of Macro-domain Basis Functions [62] are defined on each patch in the dielectric object. To derive a macro-domain basis set that does not depend on the angle of incidence of the incident electric field, each patch is illuminated by a number of plane-waves at different angles of incidence as shown in Figure (7.1). A number of marco-domain basis functions will be generated for all patches. When creating the macro-domain basis functions, each patch is considered as an individual object. Use of these MBFs is expected to be a helping factor in truncating the edge effects.

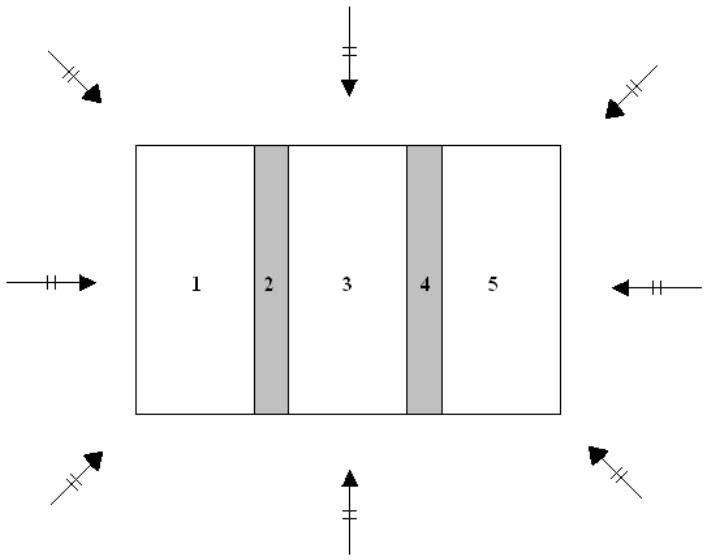

Figure 7.1: *Cross-section of an infinite rectangular dielectric cylinder at $z = 0$. N_θ plane-waves normal to the the surfaces illuminate the cylinder with a uniform angular interval of $d\theta = 360/N_\theta$ between them.*

Each plane wave impinges the object at intervals $d\theta = 360/N_\theta$ degrees. Depending on the availability of computational resources, dielectric properties, required accuracy, etc., one can flexibly choose the number of incident waves used in the procedure. N_θ macro basis functions are computed for each block. The set of solutions of the

individual sub-sectional basis functions belonging to a given patch are merged into MBFs for that patch.

We use a large number of plane-waves to create MBFs that are invariant of the direction of propagation of the electric field. The MBFs will have redundant information. The SVD will be applied to the complete set of MBFs to down-select the number of basis functions and to remove the redundancy. This also helps in maintaining the condition number κ of the reduced matrix low. A threshold will be set for the normalised singular values of each patch and only a small number of macro basis functions will be retained to represent the unknown field. A dominant sub-set of these MBFs are retained after the SVD. These MBFs will be used as the CBFs for individual blocks. These new CBFs enable us to construct a reduced linear equation of lower order than that in Chapter (6).

The decomposition of the object geometry into patches is somewhat arbitrary. There is no limitation on the number and size of the patches, except that the sandwiched blocks (ie., patch 2 and 4 in Figure (7.1)) are kept thinner compared to other patches. This enables us to retain all the MBFs in these patches as CBFs because the degrees of freedom of the sandwiched patches would be lower than other patches. Around the boundaries of the imaginary edges this is expected to be advantageous. Another advantage of this method is that it enables us to solve for multiple excitations using the same reduced matrix. Only the RHS of the reduced system needs to be computed for a new excitation. This results in a significant time-saving. Moreover, it is possible to generate the CBFs by using a sparse representation of the impedance matrix. In addition, this version of the CBFM is highly parallelisable and suitable for shared memory parallel computations.

7.2 Numerical Procedure

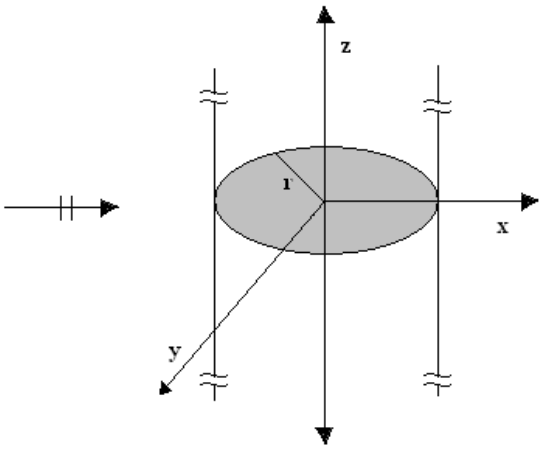

Figure 7.2: *An infinite homogeneous cylinder whose electrical properties are invariant in the z-direction. The shaded region is the cross-section of the cylinder at $z = 0$. The radius of the cylinder is denoted as r. A plane wave in the $+x$-direction with frequency f illuminates the object normal to the surface of the cylinder.*

As in the original CBFM, the new procedure begins by dividing the geometry of the object into patches, say M in number. To illustrate the method, consider a cross-section of the infinite circular dielectric object in Figure (7.2) divided into $M = 3$ patches as shown in Figure (7.3). The matrix equation resulting from the Method of Moment (MoM) solution method of the EFIE [49] is denoted as:

$$Z\mathbf{J} = \mathbf{V} \tag{7.2.1}$$

Z is the $N \times N$ matrix containing the coefficients of the interaction between different cells in the object given by Equations (6.3.2) and (6.3.3) [49, 44]. $\mathbf{V}$ is the excitation field vector of size N. $\mathbf{J}$ is the unknown solution vector of size N that contains information related to the total electric field on the object. The set of macro-domain basis functions that belong to the same patch are grouped together to

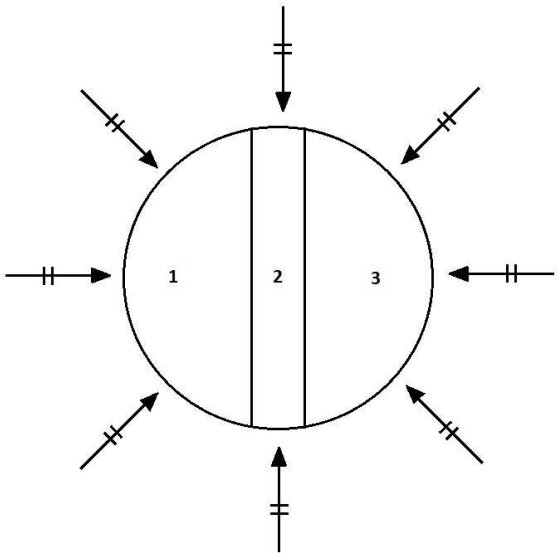

Figure 7.3: *Cross-section of the infinite circular cylinder in Figure (7.2) at $z = 0$ illuminated by N_θ plane waves at equally spaced discrete angles around the object. The object is divided into $M = 3$ blocks.*

create M^2 sub-matrices of Z. The M diagonal matrices represent the self-impedance sub-matrices and the off-diagonal blocks correspond to the mutual impedance of the patches. Matrix elements that belong to each patch will be blocked appropriately as shown in Figure (7.4). The decomposed parts of the impedance matrices are denoted $\tilde{Z}_i$, where i denotes the i^{th} block in Figure (7.4) containing interaction elements related to the i^{th} patch. N_i is the number of pulse basis functions in patch i. When the objects under investigation are large compared to the available RAM size, the sub-matrices shall be stored in the main memory.

7.2.1 Creation of Macro-domain Basis Functions

To create macro-domain basis functions for all patches, the object in Figure (7.2) will be impinged by N_θ normally incident TM plane-waves of frequency f at $d\theta$ equally spaced angles around them. Macro MoM matrices corresponding to block i

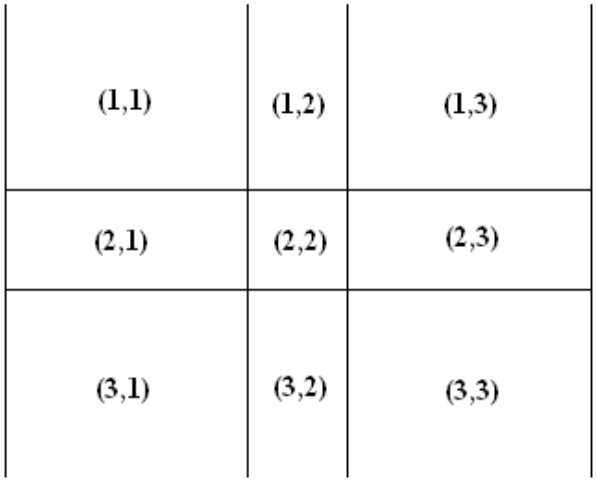

Figure 7.4: *The coefficient matrix Z divided into 9 blocks grouping MoM matrix elements belonging to 3 patches in Figure (7.3). Indices (i, j) denote MoM elemets from patch i and j.*

in Figure (7.4) are given by:

$$\tilde{Z}_{ij}\tilde{\mathbf{J}}^{(i,\theta)} = \tilde{\mathbf{V}}^{(i,\theta)} \qquad \text{for } \theta = 1, 2, \ldots, N_\theta \qquad \text{if } (i = j) \tag{7.2.2}$$

Block matrix equations in (7.2.2) are used for computing N_θ macro basis functions for each patch by inverting Equation (7.2.2):

$$\tilde{\mathbf{J}}^{(i,\theta)} = \tilde{Z}_{ij}^{-1}\tilde{\mathbf{V}}^{(i,\theta)} \qquad \text{for } \theta = 1, 2, \ldots, N_\theta \qquad \text{if } (i = j) \tag{7.2.3}$$

Macro basis functions $\tilde{J}^{(i,\theta)}$ for each patch for all angles of incidence of $\tilde{\mathbf{V}}^{(i,\theta)}$ are computed. If there are N_1, $N_2, \ldots, N_M$ cells in each patch, inverting Equation (7.2.2) would give M sets of macro basis functions. We arrange the macro basis functions of the i^{th} block in a matrix form.

$$\tilde{J}^{(i)} = \begin{pmatrix} J_1^{(i,1)} & J_1^{(i,2)} & \dots & J_1^{(i,N_\theta)} \\ J_2^{(i,1)} & J_2^{(i,2)} & & J_2^{(i,N_\theta)} \\ \vdots & \vdots & \ddots & \vdots \\ J_{Ni}^{(i,1)} & \dots & \dots & J_{Ni}^{(i,N_\theta)} \end{pmatrix} \tag{7.2.4}$$

Due to the presence of the inversion operation in Equation (7.2.3) this would turn out to be a computationally expensive method. However, one can approximate the inversion with a vector-vector division using the sparse representation of $Z_i i$, it will be summarised in Section (7.2.2).

7.2.2 Sparse Representation

Macro basis functions created in Section (7.2.1) will only be used for the testing procedure and thus there is no need to invert the matrix equations in (7.2.2). Instead, one can use the sparse representation of the macro impedance matrices to approximate $\tilde{J}^{(i,\theta)}$ for each patch for all angles of incidence of the $\tilde{\mathbf{V}}^{(i,\theta)}$ given by:

$$\tilde{\mathbf{J}}^{(i,\theta)} = \frac{\tilde{\mathbf{V}}_{ii}^{(i,\theta)}}{\tilde{Z}_{ii}^{i}} \qquad \text{for} \;\; \theta = 1, 2, \dots, N_\theta \qquad \text{if} \;\; (i = j) \tag{7.2.5}$$

However, in our work, we will construct macro-basis function by inverting (7.2.2). Sparse representations will not be used in test cases, from our observation it is seen that using sparse representations produce results that are close to the exact results but not that satisfactory. Since our aim is the development of the method and speed-up is not a concern at this stage. The sparse representation would reduce the computation time by several times as matrix inversion that costs $\mathcal{O}(\tilde{N}^3)$ is substituted with $\mathcal{O}(\tilde{N})$ vector-vector divisions, where $\tilde{N}$ is the size of sub-matrices.

7.2.3 Creation of the Characteristic Basis Functions

The number of plane waves used to generate the macro basis functions sometimes exceeds the degrees of freedom (DoF) associated with each block. It is desirable to remove the redundancy in the macro basis functions also to improve the condition number of the reduced matrix equation. A new set of characteristic basis functions that are linear combinations of the significant macro basis functions will be computed using the SVD procedure. The SVD is applied to the M sets of macro basis functions separately. Applying the SVD procedure to 7.2.4 results in:

$$J^{(i)} = U^{(i)} S^{(i)} V^{(i)H} \tag{7.2.6}$$

$$U^{(i)} = (u_1, u_2, \ldots, u_n) \in \mathbb{C}^{N_i \times N_i} \tag{7.2.7}$$

$$V^{(i)} = (v_1, v_2, \ldots, v_n) \in \mathbb{C}^{N_\theta \times N_\theta} \tag{7.2.8}$$

$$S^{(i)} = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{N_i \times N_\theta} \quad p = \min(N_i, N_\theta) \tag{7.2.9}$$

$U^{(i)}$ and $V^{(i)H}$ are the column orthogonal matrices. $S^{(i)}$ is a diagonal matrix containing singular values of $J^{(i)}$ given by:

$$\sigma_1 \geq \sigma_2 \geq \ldots \sigma_p \geq 0$$

Now the singular values are normalised with respect to the maximum σ_1 and termed σ_j^N given by:

$$\sigma_j^N = \frac{\sigma_j}{\sigma_1} \quad \text{for} \quad j = 1, 2, \ldots, p \tag{7.2.10}$$

The normalised singular values will span over several orders of magnitude. A thresholding is done on the normalised singular values with respect to the maximum (σ_1^N). Macro basis functions with normalised singular values below this threshold will be discarded to avoid redundancy. Say, S_i normalised singular values are retained for each block, then the first S_i columns of $U^{(i)}$ denoted $C^{(i)}$ in (7.2.11) are used as the characteristic basis functions for block i.

$$C^{(i)} = \begin{pmatrix} c_1^{(i,1)} & c_2^{(i,1)} & \dots & c_{S_i}^{(i,1)} \\ c_1^{(i,2)} & c_2^{(i,2)} & \dots & c_{S_i}^{(i,2)} \\ \vdots & \vdots & \ddots & \vdots \\ c_1^{(i,N_i)} & c_2^{(i,N_i)} & \dots & c_{S_i}^{(i,N_i)} \end{pmatrix} \tag{7.2.11}$$

Note that in patch i, the maximum number of linearly independent CBFs will be:

$$S_i^{max} = min(N_i, N_\theta) \tag{7.2.12}$$

Following the above mentioned procedure, we construct $K = \sum_{i=1}^{M} S_i$ primary basis functions. These CBFs can be used to create a reduced matrix of size $K \times K$ using the Galerkin method.

7.2.4 Computation of the reduced matrix

J, the solution to the entire problem in Equation (7.2.1), can be expressed as a linear combination of the CBFs as in Equation (7.2.13). Here $\alpha_k^{(i)}$, for $i = 1, 2, \dots, M$ and $k = 1, 2, \dots, S_i$ are the K unknown expansion coefficients to be determined using a reduced matrix,. The next step is to generate the reduced $K \times K$ matrix equation for the unknown complex coefficients $\alpha_k^{(i)}$ by using the Galerkin method. This procedure results in a reduced linear equation given by:

$$Z^c_{K\times K}\alpha_{K\times 1} = V^c_{K\times 1} \tag{7.2.14}$$

$$\begin{pmatrix} \\ \mathbf{J} \\ \\ \end{pmatrix} = \sum_{k=1}^{S_1} \alpha_k^{(1)} \begin{pmatrix} \left(c_k^{(1)}\right) \\ (0) \\ \vdots \\ (0) \end{pmatrix} + \sum_{k=1}^{S_2} \alpha_k^{(2)} \begin{pmatrix} (0) \\ \left(c_k^{(2)}\right) \\ \vdots \\ (0) \end{pmatrix} + \ldots + \sum_{k=1}^{S_M} \alpha_k^{(M)} \begin{pmatrix} (0) \\ (0) \\ \vdots \\ \left(c_k^{(M)}\right) \end{pmatrix} \tag{7.2.13}$$

where,

$$Z^c = \begin{pmatrix} \left\langle c_1^{(1)t} \tilde{Z}_{11}, c_1^{(1)} \right\rangle & \left\langle c_1^{(1)t} \tilde{Z}_{12}, c_2^{(1)} \right\rangle & \ldots & \left\langle c_1^{(1)t} \tilde{Z}_{1S_M}, c_{S_M}^{(M)} \right\rangle \\ \left\langle c_2^{(1)t} \tilde{Z}_{21}, c_1^{(1)} \right\rangle & \left\langle c_2^{(1)t} \tilde{Z}_{21}, c_2^{(1)} \right\rangle & \ldots & \left\langle c_2^{(1)t} \tilde{Z}_{2S_M}, c_{S_M}^{(M)} \right\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \left\langle c_{Sm}^{(M)t} \tilde{Z}_{M1}, c_1^{(1)} \right\rangle & \left\langle c_{Sm}^{(M)t} \tilde{Z}_{M2}, c_2^{(1)} \right\rangle & \ldots & \left\langle c_{Sm}^{(M)t} \tilde{Z}_{MS_M}, c_{S_M}^{(M)} \right\rangle \end{pmatrix} \tag{7.2.15}$$

$$V^c = \left(\left\langle c_1^{(1)t}, \tilde{\mathbf{V}}_\mathbf{1} \right\rangle \quad \left\langle c_2^{(1)t}, \tilde{\mathbf{V}}_\mathbf{2} \right\rangle \quad \ldots \quad \left\langle c_{S_M}^{(1)t}, \tilde{\mathbf{V}}_\mathbf{M} \right\rangle \right)^T \tag{7.2.16}$$

Here, $Z_{ij},\ i \neq j$ is the coupling matrix linking patches i and j, Z_{ii} is the self coupling matrix of patch i. Note that Z^c is diagonally symmetric and this property saves time in its creation and memory space in its storage. The reduced matrix, Z^c has to be computed only once for the object under investigation. Also, if there are some changes in the geometry or electrical constitution of one or more patches, the corresponding elements in Z^c can computed and replaced.

7.3 Numerical Results

A number of infinite cylindrical objects with various cross-sectional shapes are considered as test cases. Both homogeneous and inhomogeneous objects are tested. The objects to be analysed are discretised with $D_\lambda = 10$ pulse basis functions per wavelength and the number of cells N is given by Equation (3.8.1) to discretise the cross-section of the object at $z = 0$. The MoM-EFIE formulation in Section (3.7.2) is used to discretise the object and the MoM matrix is generated. The improved CBFM with M patches is implemented. In Section (7.3.1) a validation of the method is done by comparing the CBFM results with the *Mie Series* [32] solution for a homogeneous infinite cylinder. The incident field E^i is a TM plane wave at frequency f in the $+x$ direction with a polarisation parallel to the axis of the cylinder [32]. The field components excited by the TM plane waves are E_z,H_x and H_y. Throughout this work, we compute only the electric field component E_z.

7.3.1 Validation of the Method, Circular cylinder

In order to validate the method, a homogeneous infinite dielectric cylinder with circular cross-section as in Figure (7.2) is analysed in this section. The object is impinged with a normally incident TM plane wave at frequency $f = 300$ MHz. The radius of the cylinder is $r = 0.5\lambda$ and the relative permittivity is $\epsilon_r = 2 - \jmath 0.8$. The object is discretised into $N = 225$ cells. We compute the total electric field E_z^t on the object using the Mie series. A surface plot of the magnitude of E_z^t on the cross-section at $z = 0$ computed using the Mie series is shown in Figure (7.5). The matrix equation resulting from the EFIE-MoM formulation is iteratively solved using the GMRES method with a tolerance $\leq 1e-05$. This iterative solution is plotted in Figure (7.6).

Three patches with $N_1 = 105, N_2 = 15$ and $N_3 = 105$ pulse basis functions

each are used for the improved CBFM procedure. $N_\theta = 20$ plane-waves are used to create N_θ MBFs for each block. The normalised singular values σ^N of matrices $J^{(i)}$ obtained using the SVD are plotted in Figure (7.7). A threshold $\sigma_t^N = 1e-4$ is set and $S_1 = 13, S_2 = 8$ and $S_3 = 13$ CBFs are retained in each block. The Galerkin procedure in Equation (7.2.15) results in a reduced matrix $Z^c_{34\times34}$. Figure (7.8) shows a surface plot of the magnitude of E_z^t computed using the CBFM. Figure (7.9) shows a comparison of all the three solutions at $z = x = 0$, while in Figure (7.10) a comparison is made at $z = y = 0$.

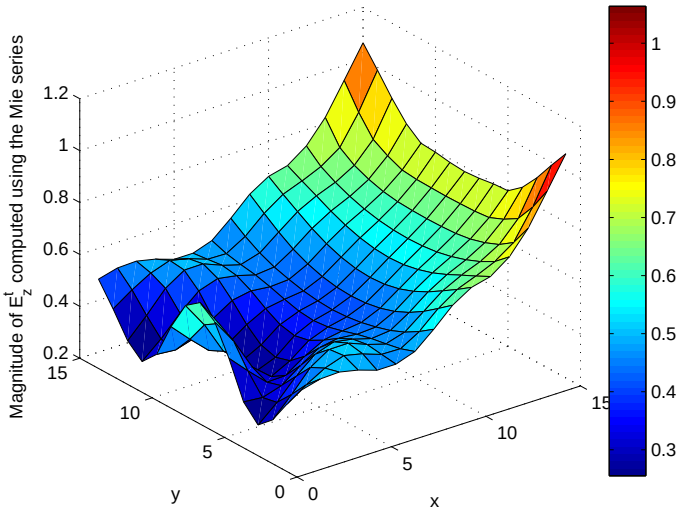

Figure 7.5: *A surface plot of the magnitude of E_z^t on the object at $z = 0$ computed using the Mie series.*

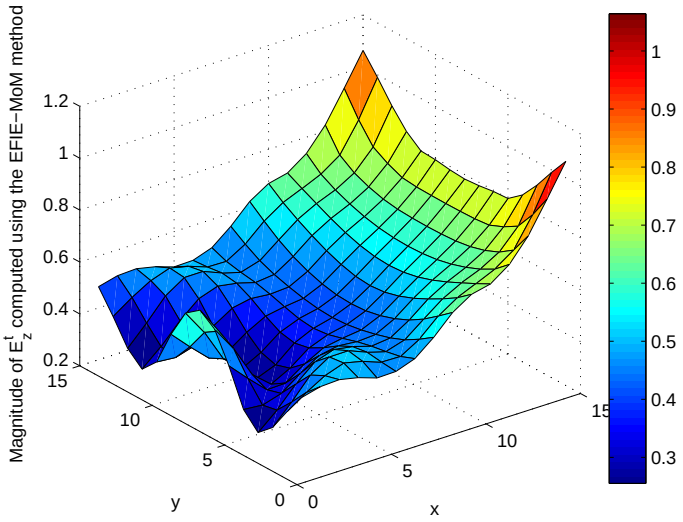

Figure 7.6: *A surface plot of the magnitude of E_z^t on the object at $z = 0$ computed using the EFIE-MoM approach.*

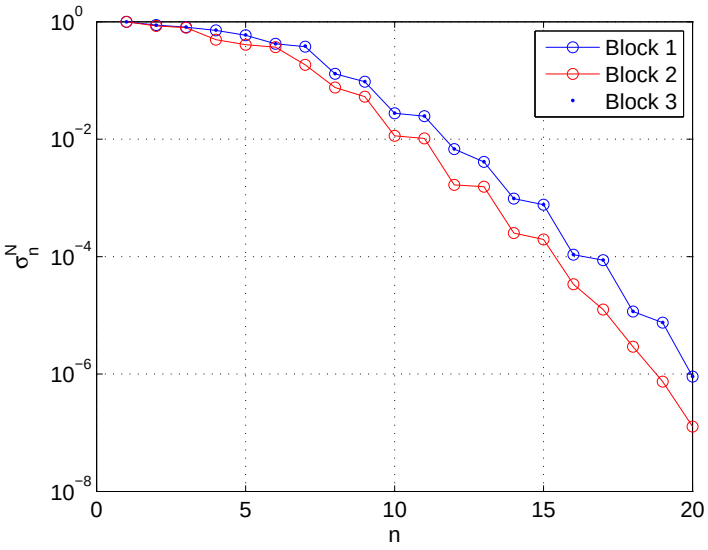

Figure 7.7: *On the log y-axis is the normalised singular values σ_j^N of all the 3 blocks, $N_\theta = 20$.*

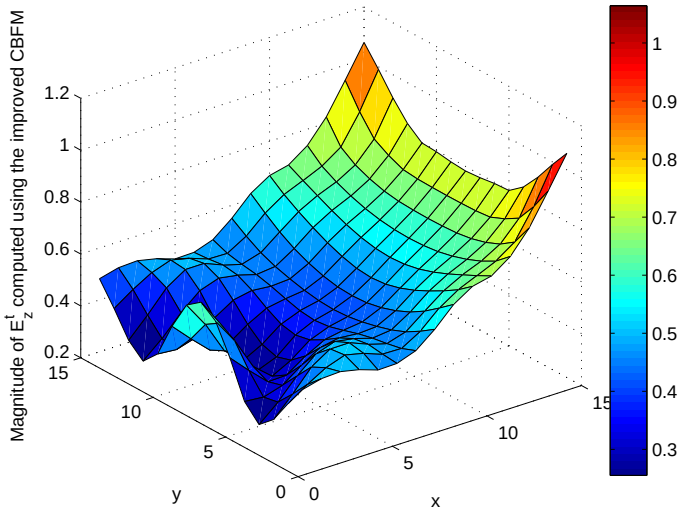

Figure 7.8: *A surface plot of the magnitude of E_z^t on the object at $z = 0$ computed using the CBFM with $M = 3$ patches.*

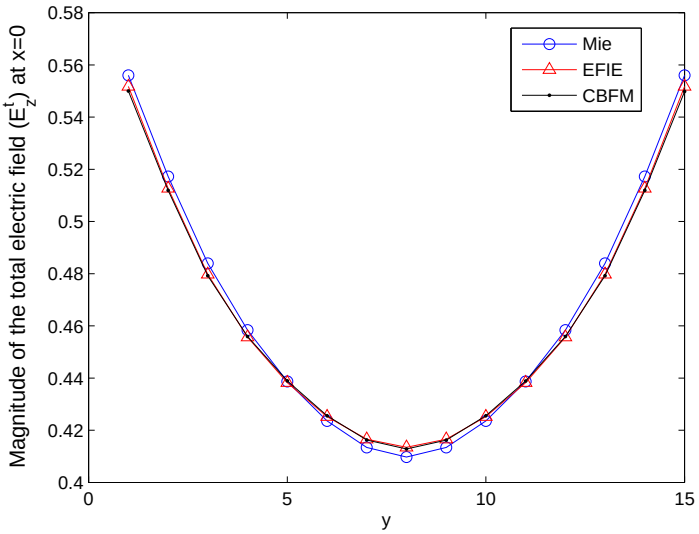

Figure 7.9: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the Mie series, EFIE-MoM and the CBFM.*

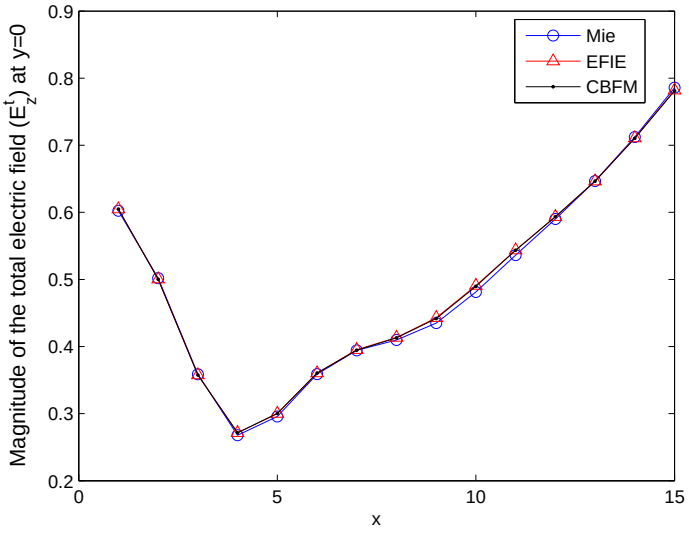

Figure 7.10: *A comparison of the magnitude of the E_z^t on the object at $z = y = 0$ computed using the Mie series, EFIE-MoM and the CBFM.*

Figures (7.5)-(7.10) show that the total electric field E_z^t computed using the Improved CBFM matches very well with the analytical Mie series solution. In the next sections, a number of objects with different electrical properties and sizes are analysed in detail.

7.3.2 Infinite Inhomogeneous Cylinders

Two inhomogeneous infinite dielectric cylinders with circular cross-section are analysed in this section. The incident fields are normally incident TM plane waves at frequency $f = 300$ MHz. The inhomogeneity in the objects are due to a coaxial cylinder inside the object as shown in Figure (6.21). The radius of the outer cylinder is denoted r_o while that of the inner cylinder is $r_i = 0.2r_o$. The dielectric profiles of the inner and outer cylinders are $\epsilon_{ro} = 2$, $\epsilon_{ri} = 20$, $\sigma_o = 0.8$ and $\sigma_i = 0.8$, respectively. D_λ pulse basis functions per wavelength are used to discretise the cross-section of the object at $z = 0$. This results in N discretisation points in total.

Test Case 2A

An inhomogeneous infinite dielectric cylinder of circular cross-section with $r_o = 0.5\lambda$ and $r_i = 0.2r_o$ is used as the object. Three patches with $N_1 = 90, N_2 = 45$ and $N_3 = 90$ pulse basis functions each are used for the Improved CBFM procedure. N_θ plane-waves are used to create N_θ MBFs for each block. The normalised singular values σ_n^N of the matrix $J^{(i)}$ from various test cases, computed using the SVD, are plotted in figures below. After the SVD, S_i CBFs are selected in each block. The Galerkin procedure in (7.2.15) results in a reduced matrix $Z^c_{K\times K}$ which is solved using LU decomposition. Table (7.1) summarises results from several test cases where N_θ and S_i are varied and the error in satisfying Equation (7.1), denoted by δ is computed.

The matrix equation resulting from the EFIE-MoM formulation is iteratively solved using the GMRES method with a tolerance $\leq 1e-05$. This iterative solution is plotted following figures. Results from a number of relevant test cases in Table (7.1) are plotted in Figures (7.11)-(7.25). Figures showing surface plots of the magnitude of E_z^t computed using the CBFM are shown in these figures. Other figures show a comparison of the magnitude of the CBFM and EFIE solutions at $z = x = 0$ and $z = y = 0$.

Test case	N_θ	S_i			δ
		S_1	S_2	S_3	
1	20	15	15	15	0.0135
2		17	15	17	0.0101
3		20	14	20	0.0075
4		20	15	20	0.0072
5		20	20	20	0.0072
6	30	17	15	17	0.0101
7		27	27	27	0.0026
8		27	16	27	0.0026
9		30	16	30	0.0014
10		30	30	30	0.0012
11	36	17	15	17	0.0101
12		25	25	25	0.0031
13		30	30	30	0.0012
14		33	33	33	0.0009
15		36	36	36	0.0013

Table 7.1: *A Summary of the results from test case 2A involving scattering from infinite inhomogeneous circular cylinder. The frequency of the incident field* $f =$ 300MHz. *The radius of the object is* $R = 0.5\lambda$. N_θ *is the number of plane waves used to create the macro basis functions.* S_i *is the number of CBFs used from block* i *and* δ *is the normalised error in satisfying Equation (7.2.1).*

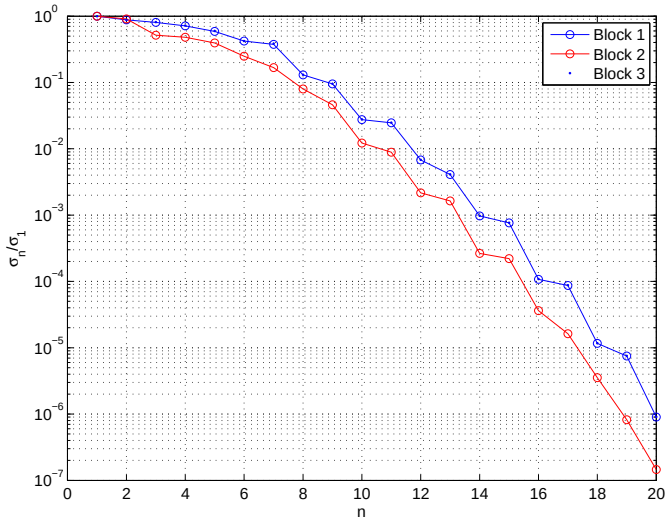

Figure 7.11: *On the log y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks in TC 2(A) (1-5),* $N_\theta = 20$.

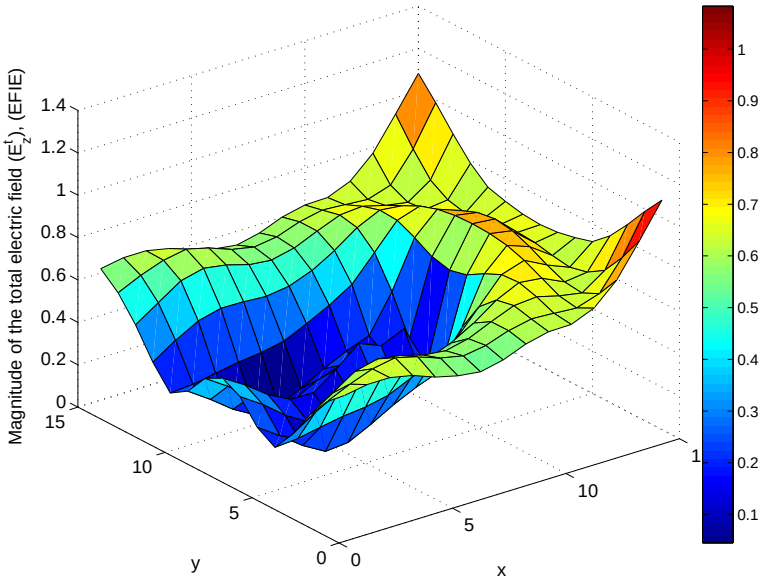

Figure 7.12: *A surface plot of the magnitude of E_z^t on the object in test case 2A computed using the EFIE-MoM approach.*

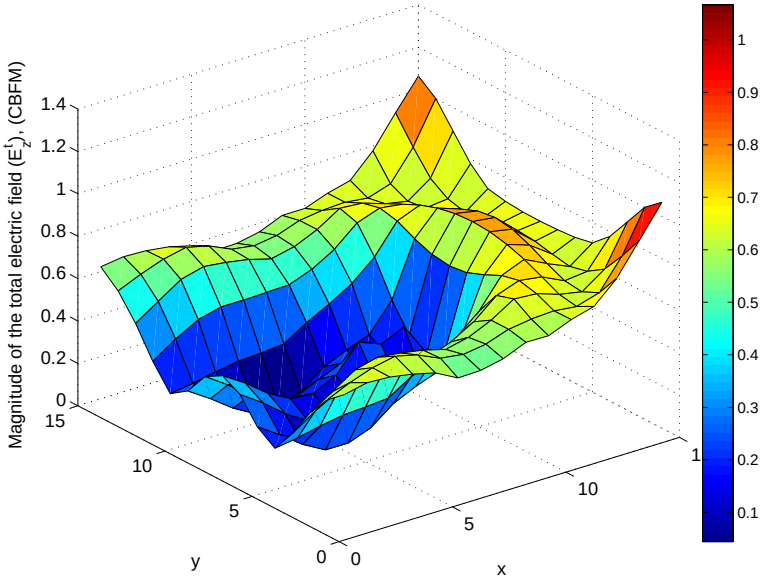

Figure 7.13: *A surface plot of the magnitude of E_z^t computed using the CBFM with parameters as in test case 2A(1) in Table (7.1).*

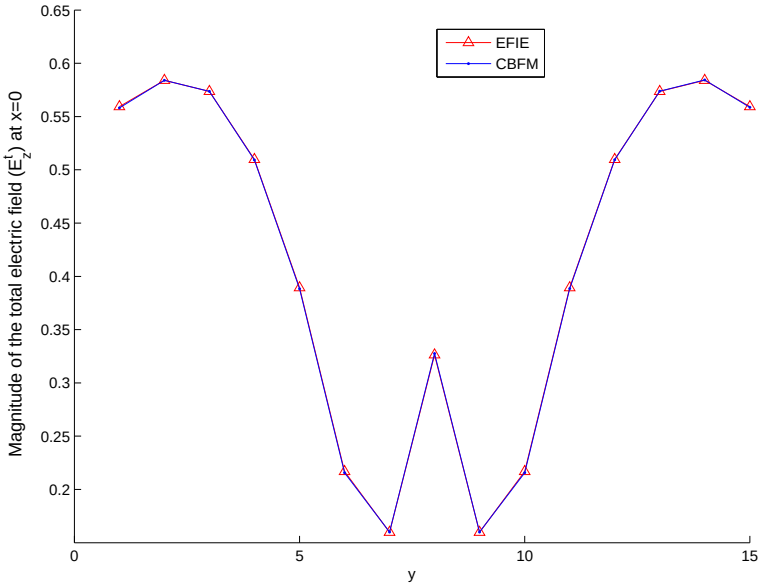

Figure 7.14: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM approach and the CBFM with parameters as in test case 2A(1) in Table (7.1).*

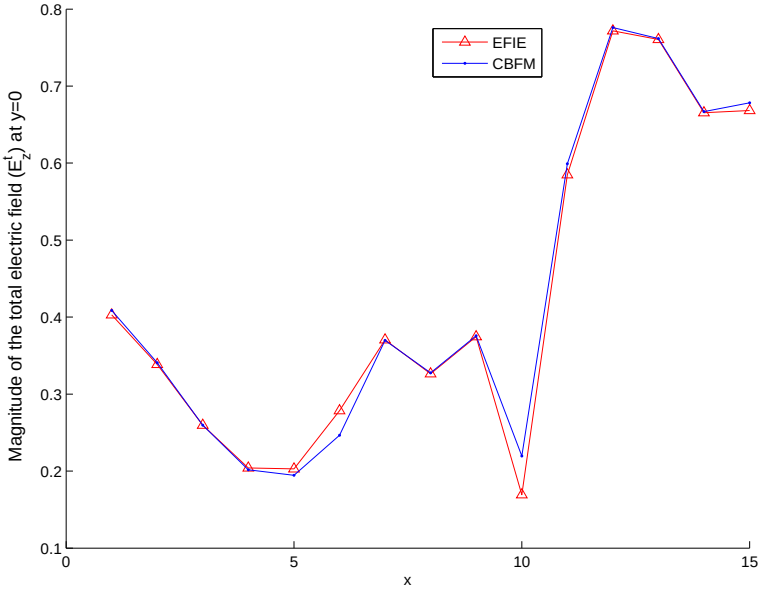

Figure 7.15: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM approach and the CBFM with parameters as in test case 2A(1) in Table (7.1).*

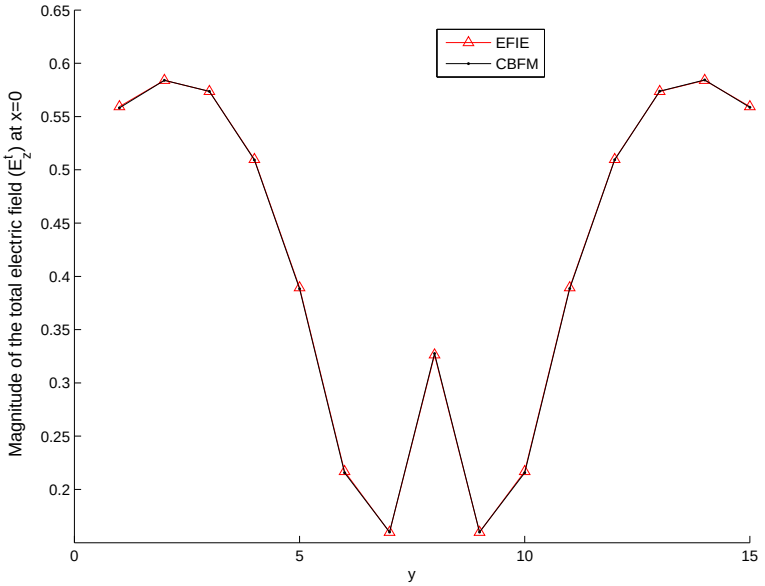

Figure 7.16: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(4) in Table (7.1).*

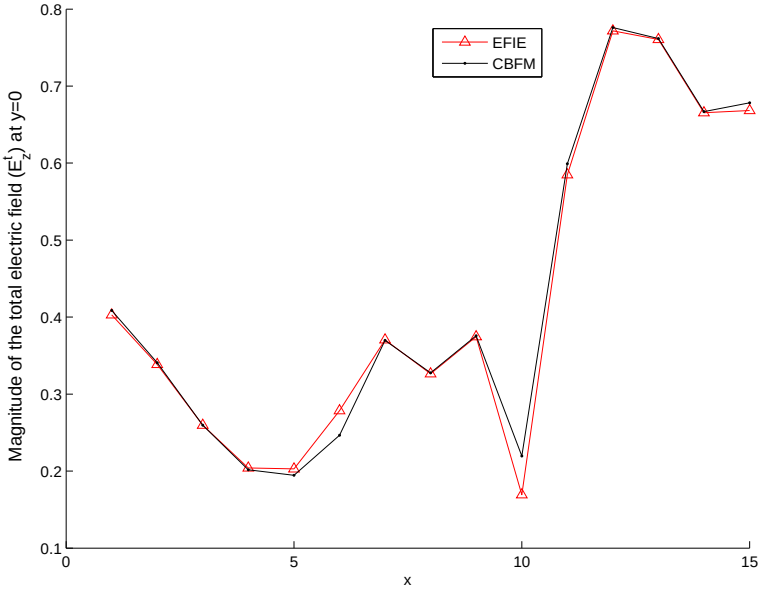

Figure 7.17: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(4) in Table (7.1).*

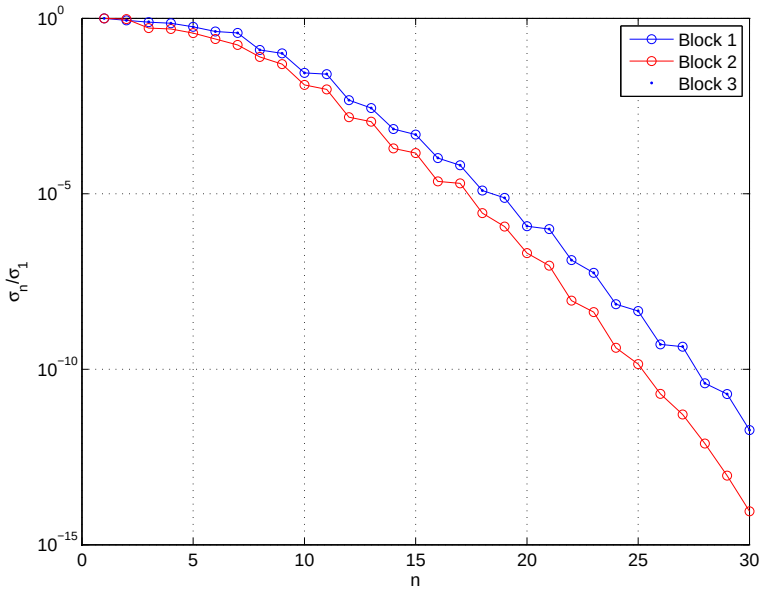

Figure 7.18: *On the log y-axis is the normalised singular values σ_n^N of all the 3 blocks blocks in TC 2(A) (6-10), $N_\theta = 30$.*

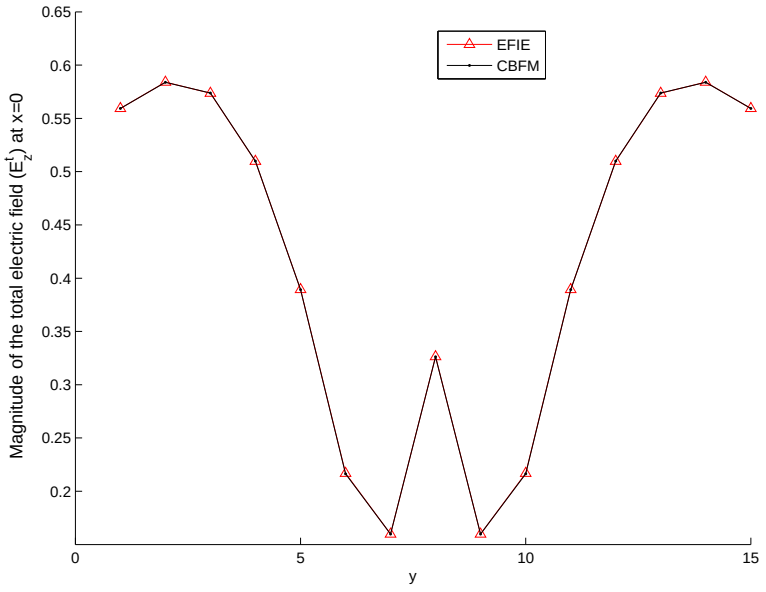

Figure 7.19: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(8) in Table (7.1).*

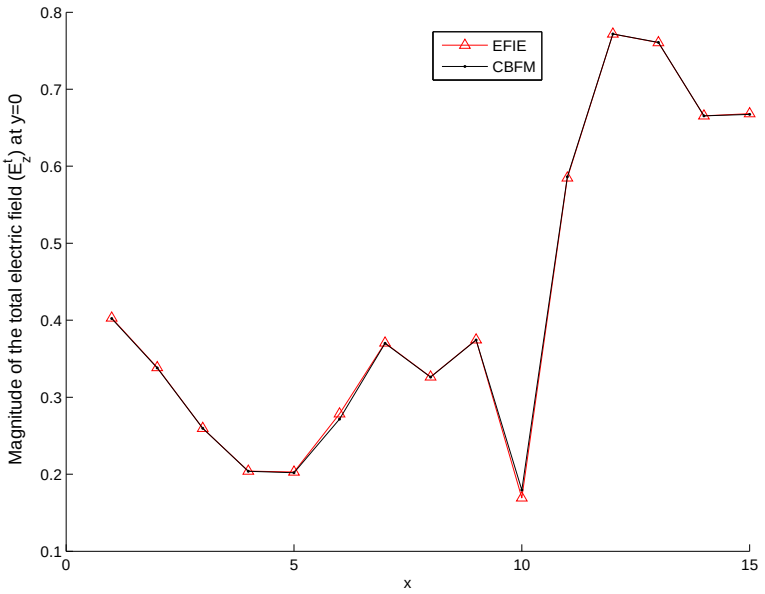

Figure 7.20: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(8) in Table (7.1).*

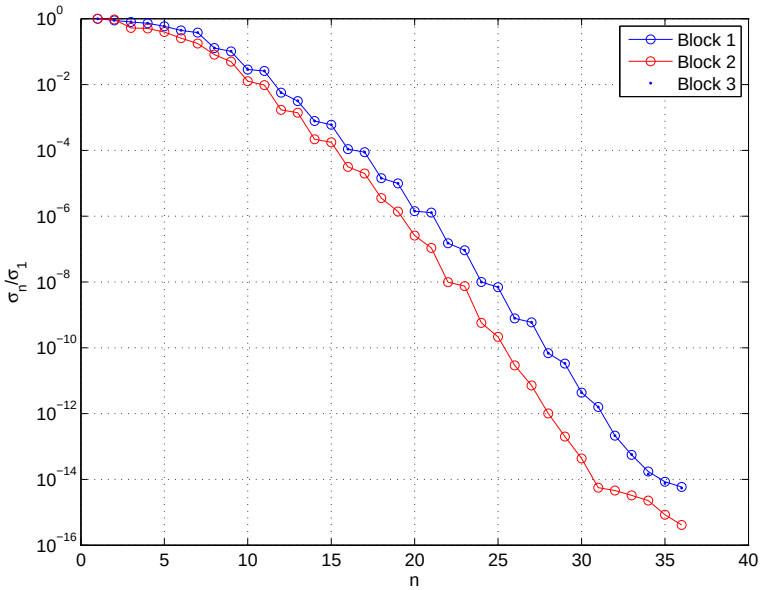

Figure 7.21: *On the y-axis is the normalised singular values σ_n^N of all the 3 blocks blocks in TC 2(A) (11-15), $N_\theta = 30$.*

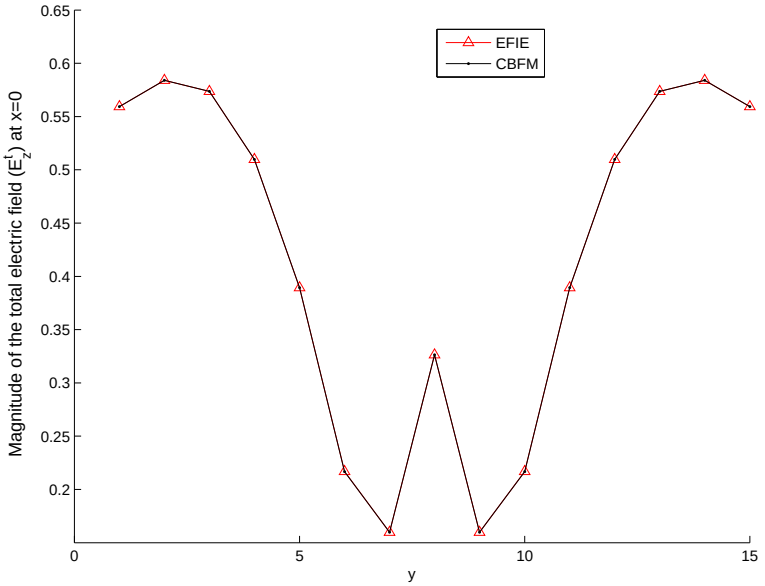

Figure 7.22: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(12) in Table (7.1).*

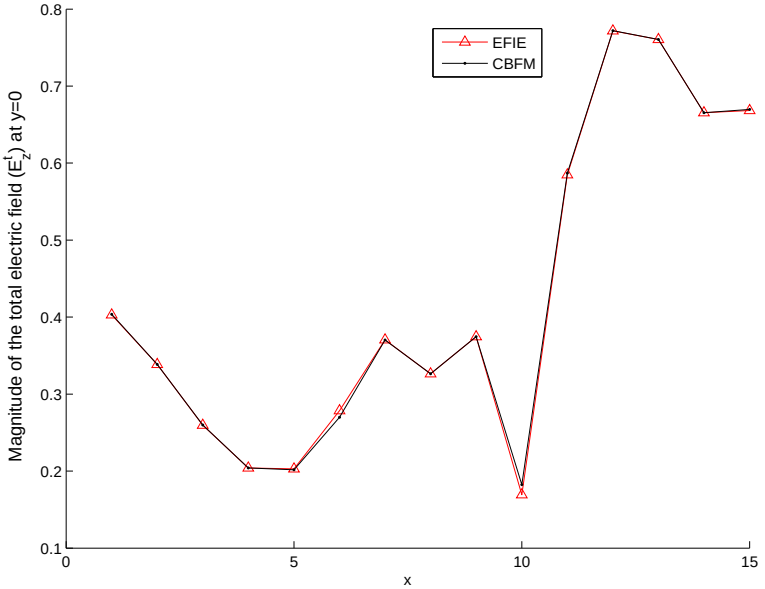

Figure 7.23: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(12) in Table (7.1).*

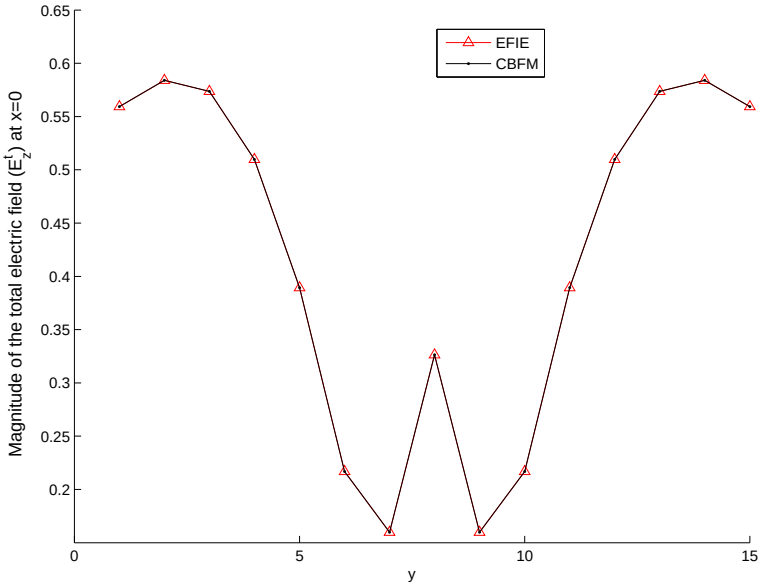

Figure 7.24: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(14) in Table (7.1).*

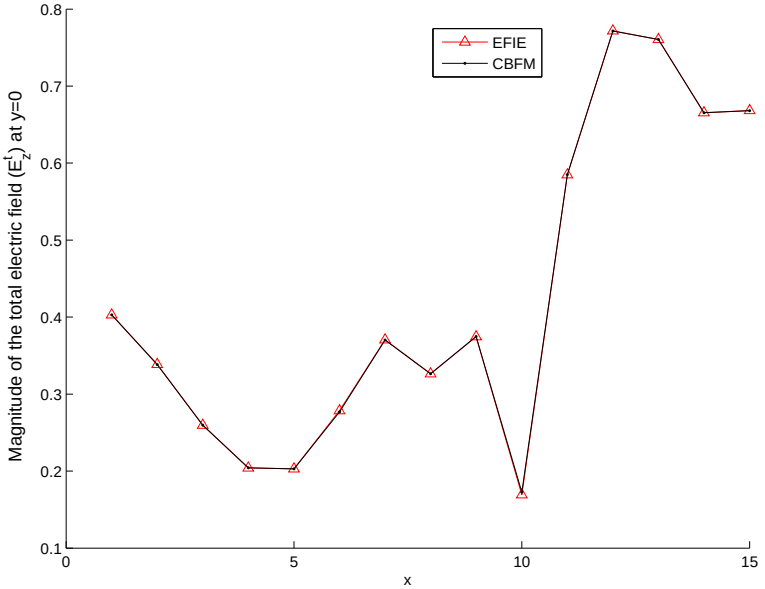

Figure 7.25: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(14) in Table (7.1).*

Figures (7.12) - (7.25) show that the Improved CBFM can be used to compute electromagnetic fields produced by scattering from infinite inhomogeneous dielectric cylinders with circular cross-section. Various test cases show that increasing N_θ always helps in achieving better accuracy. It is clear from test cases $2A(7), 2A(8), 2A(14)$ and $2A(15)$ in Table (7.1) that increasing S_i does not always guarantee better accuracy, while it can sometimes deteriorate the accuracy as seen in $2A(14)$ and $2A(15)$.

Test Case 2B

An inhomogeneous infinite dielectric cylinder with a circular cross-section is analysed in this section. The inner radius of the object is $r_o = \lambda$ and the outer radius $r_i = 0.2r_o$. Three patches with $N_1 = 348, N_2 = 145$ and $N_3 = 348$ pulse basis functions are used for the Improved CBFM procedure. N_θ plane-waves are used to create N_θ MBFs for each block. The normalised singular values σ_j^N of the matrix $J^{(i)}$ computed using the SVD are plotted in the following figures. After the SVD, S_i CBFs are selected in each block. The Galerkin procedure in (7.2.15) results in a reduced matrix $Z^c_{K\times K}$ which is solved using the LU decomposition. Table (7.2) summarises results from several test cases where N_θ and S_i are varied and the error in satisfying Equation (7.2.1), denoted by δ, is computed.

The matrix equation resulting from the EFIE-MoM formulation is iteratively solved using the GMRES method with a tolerance $\leq 1e-05$. This iterative solution is plotted in the figures below. Results from a number of relevant test cases in Table (7.2) are plotted in Figures (7.26)-(7.38). Figures showing surface plots of the magnitude of E_z^t computed using the CBFM are shown. Other figures show a comparison of the magnitude of the CBFM and EFIE solutions at $z = x = 0$ and $z = y = 0$.

Test case	N_θ	S_i			δ
		S_1	S_2	S_3	
1	20	17	17	17	0.0131
2		18	18	18	0.0121
3		18	16	18	0.0122
4		20	18	20	0.0115
5		20	20	20	0.0117
6	30	23	21	23	0.0085
7		18	16	18	0.0110
8		20	18	20	0.0104
9		20	20	20	0.0099
10		30	30	30	0.0070
11	36	17	17	17	0.0128
12		20	20	20	0.0099
13		30	30	30	0.0062
14		36	30	36	0.0048
15		36	36	36	0.0047

Table 7.2: *A Summary of the results from test case 2B involving scattering from infinite inhomogeneous circular cylinder. The frequency of the incident field is* $f =$ 300MHz. *The radius of the object is* $R = \lambda$. N_θ *is the number of incident fields used to create the macro basis functions,* S_i *is the number of singular values retained in block* i *and* δ *is the normalised error in satisfying Equation (7.2.1).*

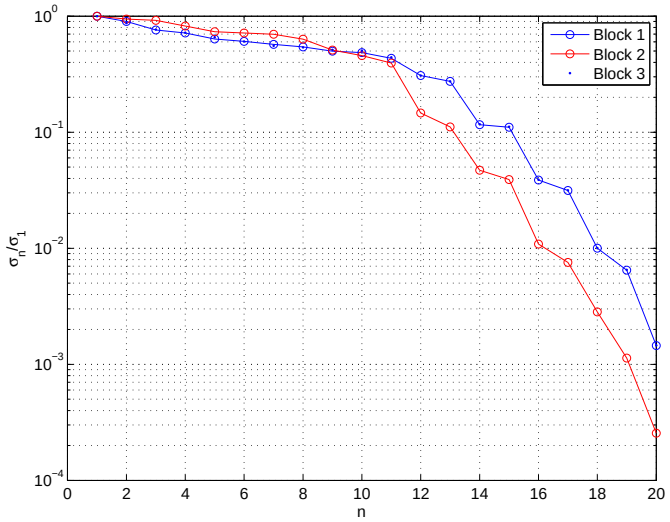

Figure 7.26: *On the log y-axis is the normalised singular values* σ_n^N *of all of the* 3 *blocks,* $N_\theta = 20$.

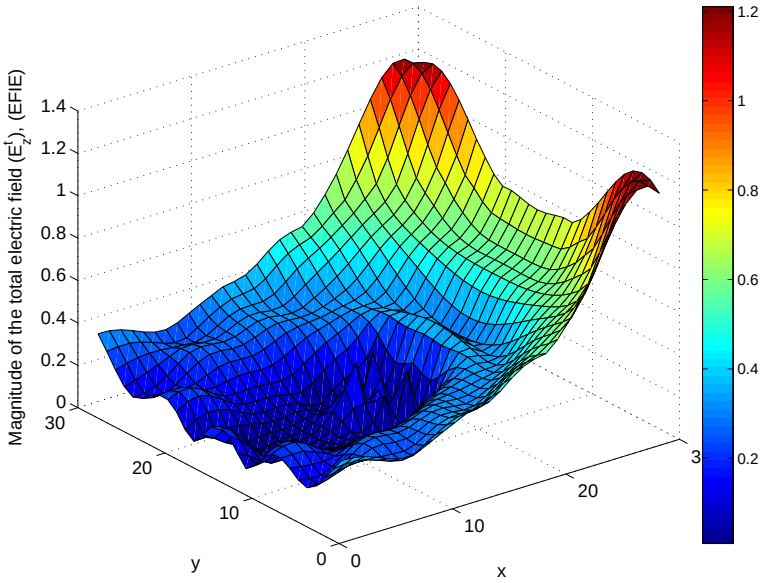

Figure 7.27: *A surface plot of the magnitude of E_z^t computed using the EFIE-MoM for test case 2B.*

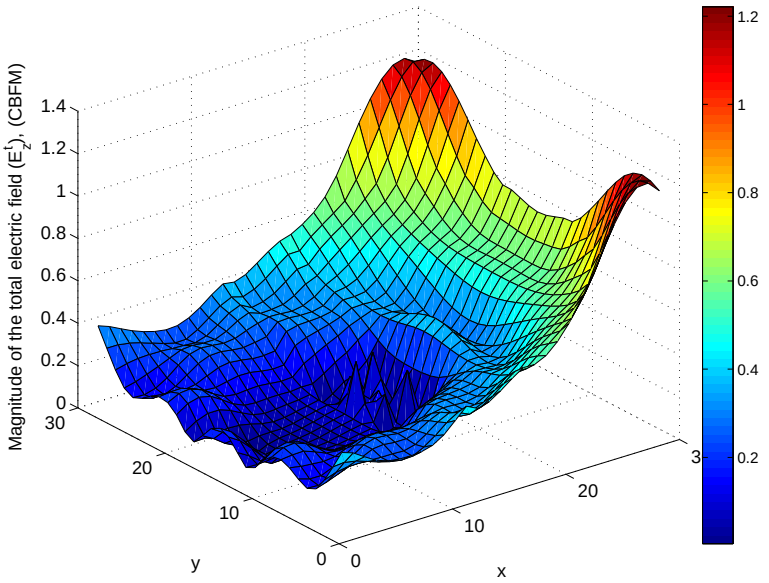

Figure 7.28: *A surface plot of the magnitude of E_z^t computed using the CBFM with parameters as in test case 2B(1) in Table (7.2).*

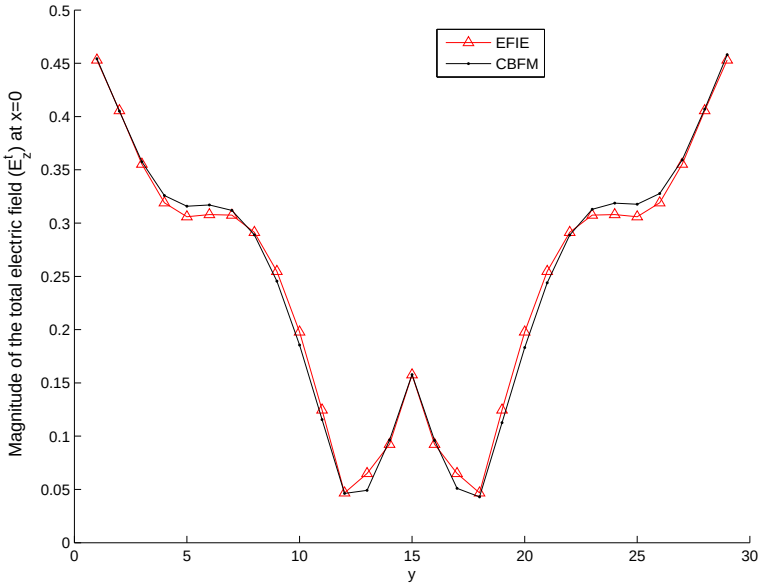

Figure 7.29: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(1) in Table (7.2).*

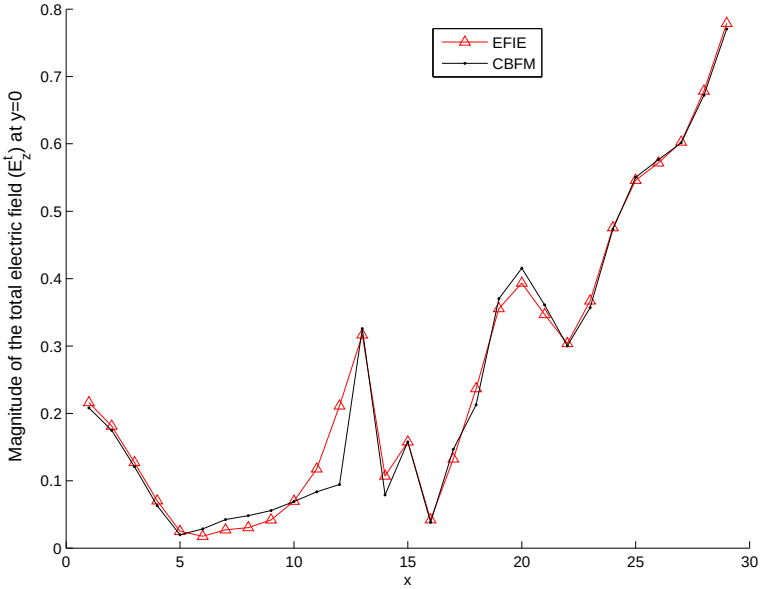

Figure 7.30: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2A(1) in Table (7.2).*

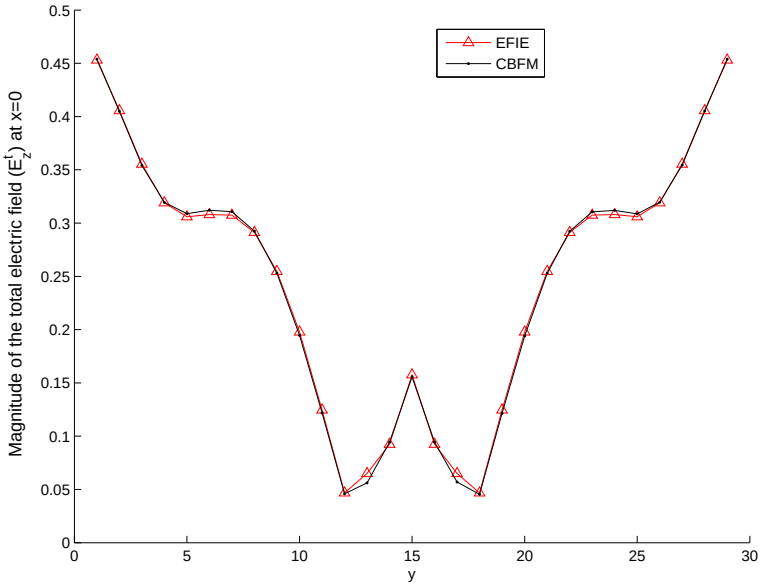

Figure 7.31: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(4) in Table (7.2).*

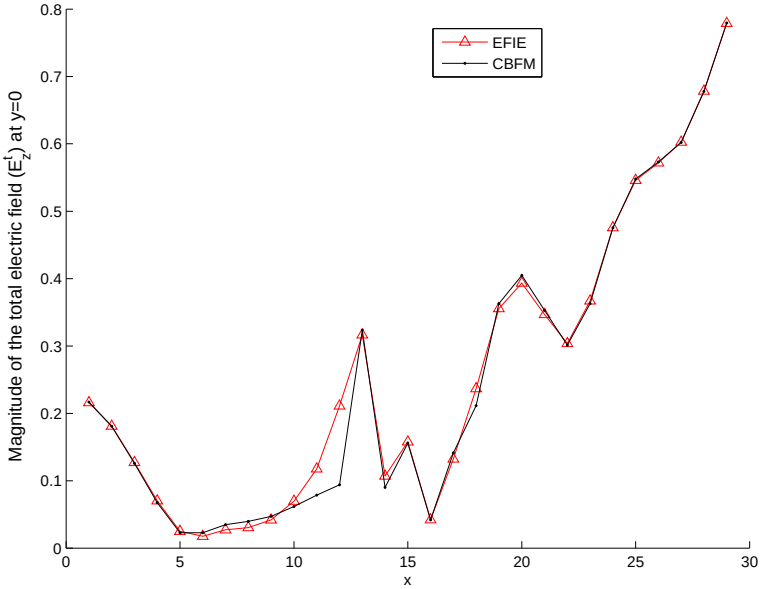

Figure 7.32: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(4) in Table (7.2).*

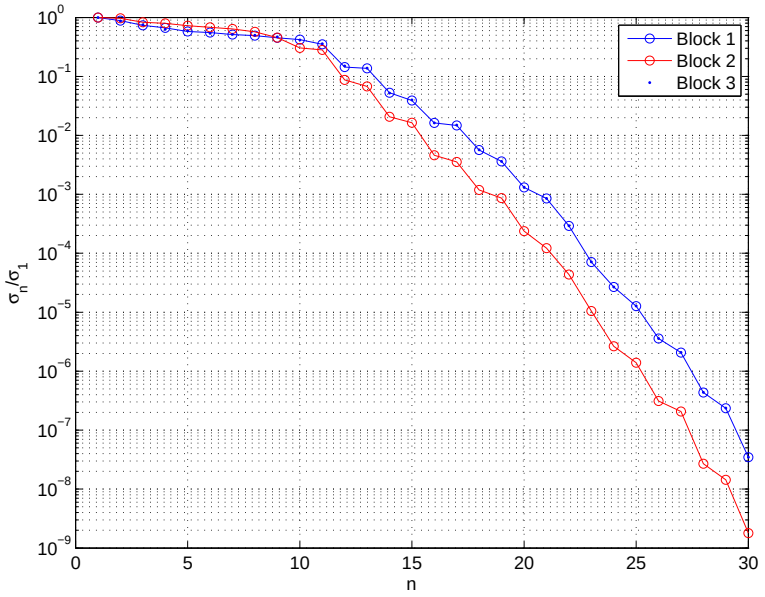

Figure 7.33: *On the log y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks,* $N_\theta = 20$.

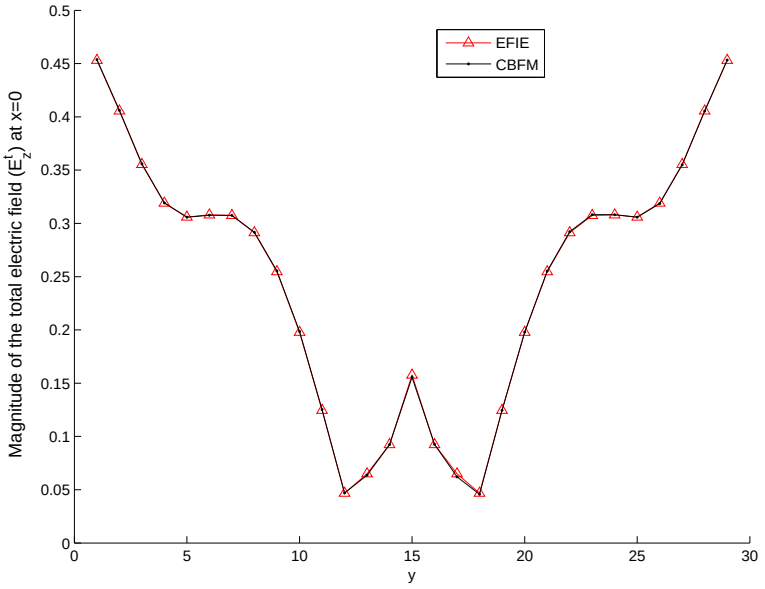

Figure 7.34: *A comparison of the magnitude of* E_z^t *on the object at* $z = x = 0$ *computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(9) in Table (7.2).*

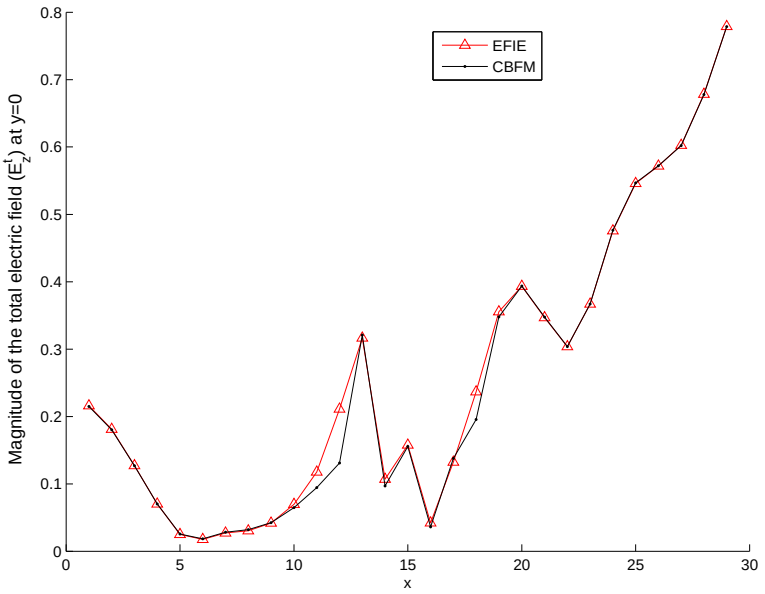

Figure 7.35: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(9) in Table (7.2).*

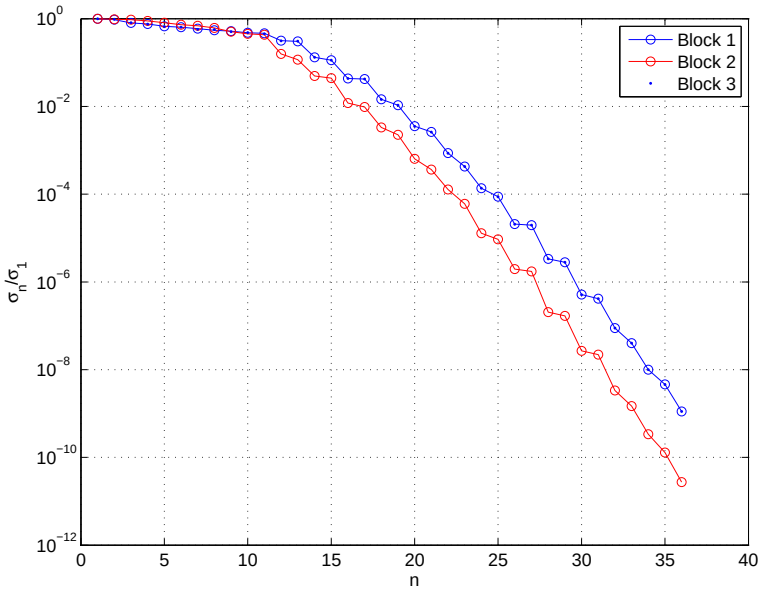

Figure 7.36: *On the log y-axis is the normalised singular values σ_n^N of all the 3 blocks, $N_\theta = 20$.*

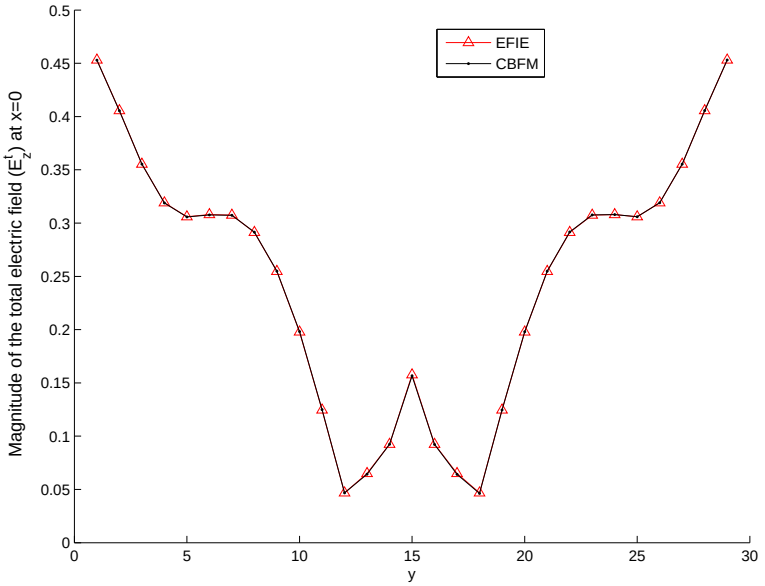

Figure 7.37: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(13) in Table (7.2).*

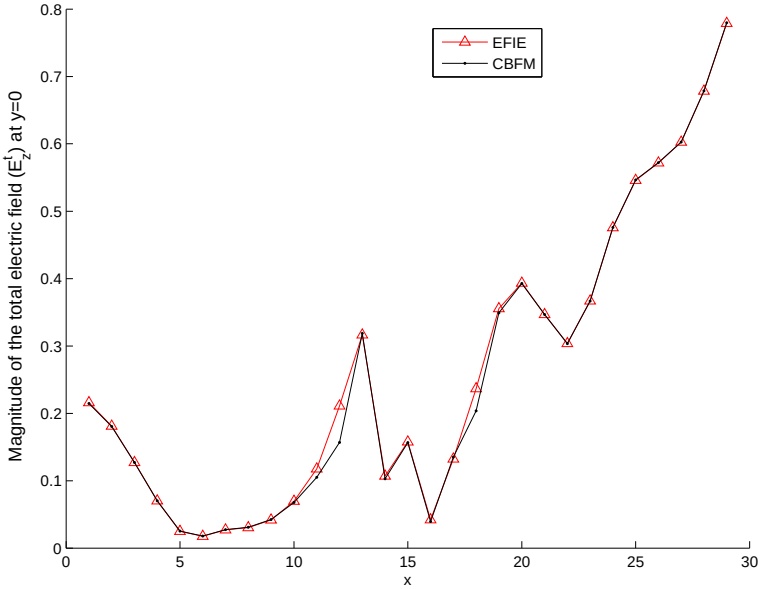

Figure 7.38: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2B(13) in Table (7.2).*

Table (7.2) shows that δ is reasonably small in all test cases. A close examination of the figures shows that compared to the EFIE-MoM results, the CBFM results show slight disagreements in the total electric field at $y = 0$, close to the boundary of the inhomogeneity. This draws attention to the fact that the width of the middle patch in 2B was less than the diameter of the inner cylinder. It is desirable to increase the width of the middle patch such that an unnecessary decomposition of the inhomogeneity is avoided. In test case 2C a few experiments are done by increasing the width of patch 2, such that the inhomogeneity is enclosed inside it.

Test Case 2C

The same object as in test case 2B is used in this section. The inhomogeneous infinite dielectric cylinder with circular cross-section has an outer radius $r_o = \lambda$ and inner radius $r_i = 0.2r_o$. The difference in this test case is that the middle patch is made wider, Three patches with $N_1 = 309, N_2 = 203$ and $N_3 = 309$ pulse basis functions each are used for the Improved CBFM procedure. N_θ plane-waves are used to create N_θ MBFs for each block. Normalised singular values σ_j^N of the matrix $J^{(i)}$ computed using the SVD is plotted in following Figures (7.39) and (7.42). After the SVD S_i CBFs are selected in each block. The Galerkin procedure in (7.2.15) results in a reduced matrix $Z^c_{K\times K}$. Table (7.3) summarises results from several test cases where N_θ and S_i are varied and the error in satisfying Equation (7.1), denoted by δ is computed.

The impedance matrix resulting from the EFIE-MoM formulation is iteratively solved using the GMRES method with a tolerance $\leq 1e-05$. Results from test cases 2C(10) and 2C(8) in Table (7.3) are plotted in Figures (7.40)-(7.43).

Test case	N_θ	S_i			κ	δ
		S_1	S_2	S_3		
1	20	17	17	17	38.1247	0.0086
2		18	18	18	38.1159	0.0081
3		18	17	18	38.1829	0.0081
4		20	18	20	38.7016	0.0115
5		20	20	20	64.7244	0.0072
6	30	20	20	20	60.5514	0.0058
7		25	25	25	99.6466	0.0033
8		30	30	30	115.8138	0.0023
9	40	30	30	30	115.2604	0.0022
10		36	36	36	118.5709	0.0014
11		40	40	40	115.0110	0.0012

Table 7.3: *A Summary of the results from test case 2C involving scattering from infinite inhomogeneous circular cylinder. The frequency of the incident field* $f = 300MHz$*, the radius of the object is* $R = \lambda$*.* N_θ *is the number of incident fields used to create the macro basis functions,* S_i *is the number of singular values retained in block* i *and* δ *is the normalised error in satisfying Equation (7.2.1).*

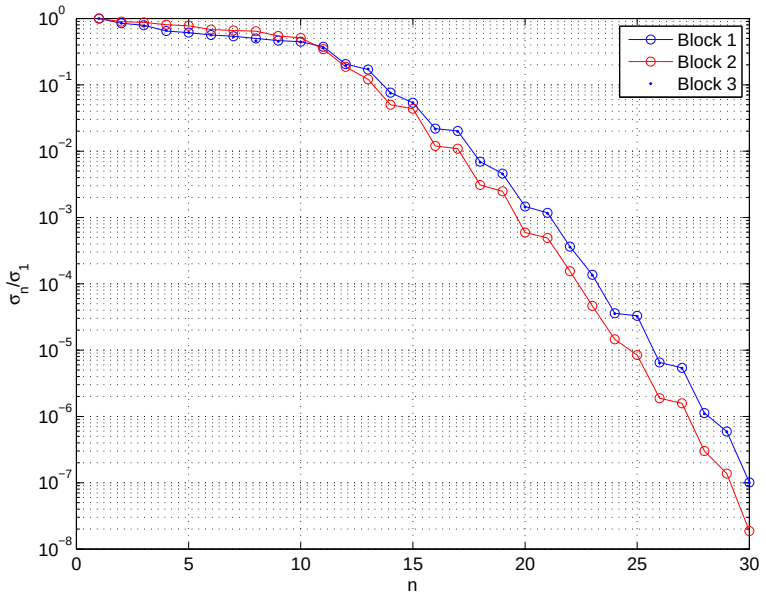

Figure 7.39: *On the log y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks in test case 2C,* $N_\theta = 20$.

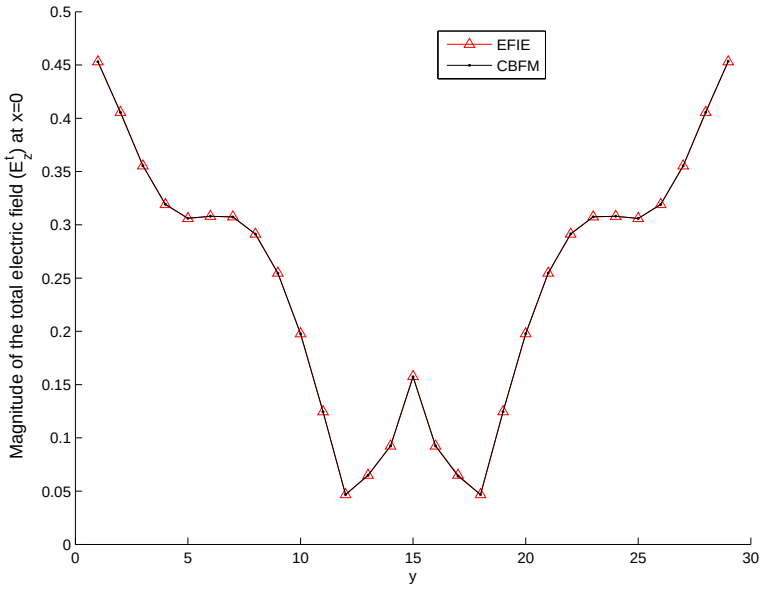

Figure 7.40: *A comparison of the magnitude of* E_z^t *on the object at* $z = x = 0$ *computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(8) in Table (7.2).*

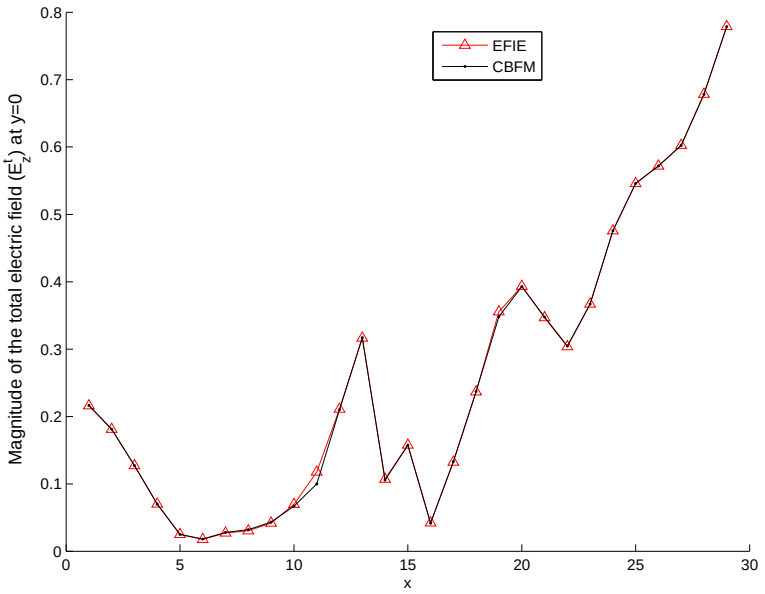

Figure 7.41: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(8) in Table (7.2).*

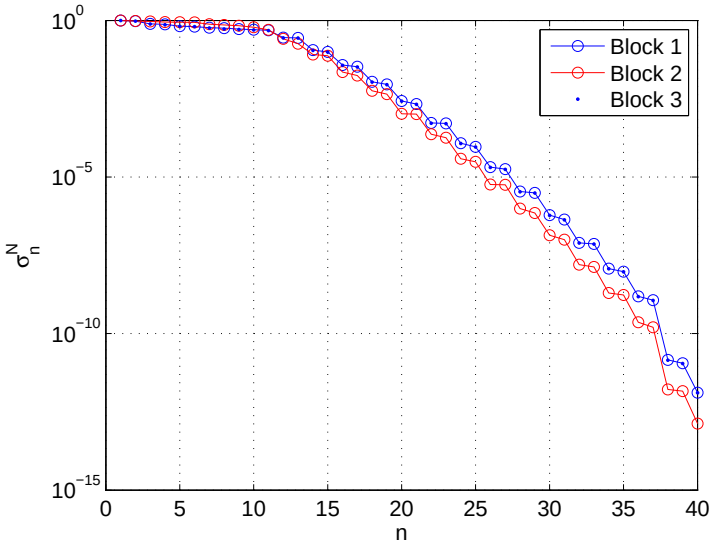

Figure 7.42: *On the log y-axis is the normalised singular values σ_n^N of all the 3 blocks in test case 2C, $N_\theta = 40$.*

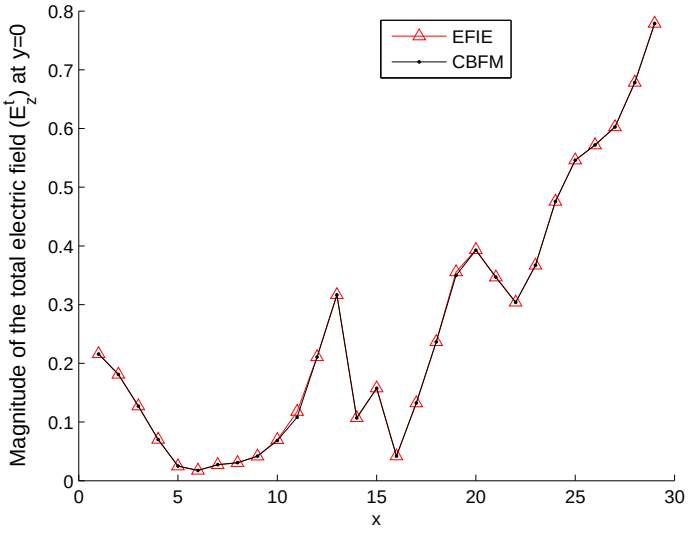

Figure 7.43: *A comparison of the magnitude of E_z^t on the object at $z = y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(10) in Table (7.2).*

Figure (7.43) shows that the Improved CBFM can be used to compute EM fields produced by the scattering of infinite inhomogeneous dielectric cylinders with circular cross-section. Various test cases show that increasing N_θ always helps in achieving better accuracy. Note that we used all the possible CBFs in test case 2C(10). Even-though σ_{40}^N for all the blocks are of several orders of magnitude smaller than σ_1^N, including all the CBFs helped. This is due to the fact that by keeping all CBFs we enrich the sub-space such that it can handle scattered waves in more directions near the inhomogeneity. Excellent agreement with iterative solution of the EFIE-MoM formulation is obtained for infinite inhomogeneous cylinders. Test case shows that care must be taken in decomposing patches when inhomogeneity exists. In test case $2B$, width of patch 2 was shorter than the inhomogeneity and this is increased in test case 2C. This is shown to have increased the accuracy of the solution around the boundary of the inhomogeneity.

7.3.3 Infinite Square Cylinder

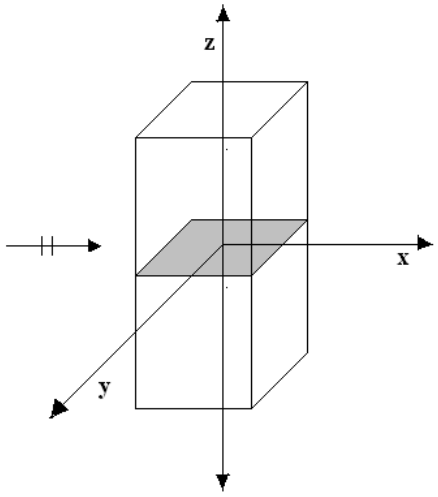

Figure 7.44: *An infinite homogeneous square cylinder whose electrical properties are invariant in the z-direction. The shaded region is the cross-section of the cylinder at $z = 0$. The width of the cylinder is denoted as a. A plane wave in the $+x$-direction with frequency f illuminates the object normal to the surface of the cylinder.*

Infinite homogeneous dielectric square cylinders with different electrical properties and sizes are analysed in this section. To create MBFs, the objects are illuminated with a normally incident TM plane wave at frequency f. The number of discretisation cells used per wavelength $D_\lambda = 10$. The object is discretised into N cells using pulse basis functions. The Improved CBFM is used to compute the total electric field E_z^t at the cross-section of the object at $z = 0$.

Test case 3A

An infinite homogeneous dielectric square cylinder with side $a = 2\lambda$ and $\epsilon_r = 2 - \jmath 0.8$ is analysed in this section. The object is illuminated with a normally incident TM plane wave of frequency $f = 300$ MHz. The cross-section of the object is discretised

into $N = 841$ cells. The MoM equation resulting from the EFIE-MoM formulation in Section (3.7.2) is directly solved using the GMRES method with a tolerance $\delta \leq 1e{-}5$.

To decompose the object, $N_1 = 348, N_2 = 145$ and $N_3 = 348$ pulse basis functions each are grouped together into three patches for the CBFM procedure. $N_\theta = 20, 30$ and 36 plane waves are used to create N_θ MBFs for each patch. The normalised singular values of the matrix $J^{(i)}$ computed using the SVD are plotted in Figure (7.45) and (7.50). Table (7.4) shows a number of test cases. S_i is the number of CBFs used in each patch for the computation of E_z^t. The Galerkin procedure in (7.2.15) results in a reduced matrix $Z_{K\times K}^c$ which is inverted using the LU decomposition. The condition number κ of the reduced matrix is also tabulated. Figure (7.46) - (7.52) show surface plots of the magnitude of E_z^t computed using the CBFM.

Test case	N_θ	S_i			κ	δ
		S_1	S_2	S_3		
1	20	18	18	18	08.0191	0.0051
2		19	19	19	10.0770	0.0037
3		20	20	20	10.0489	0.0036
4	30	20	20	20	7.6039	0.0017
5		20	15	20	7.7327	0.0038
6		25	25	25	7.6059	3.29e-4
7	36	20	20	20	7.5967	0.0014
8		20	18	20	7.5962	0.0017
9		30	30	30	7.6254	5.93e-5

Table 7.4: *A Summary of the results from test case 3A involving scattering from infinite inhomogeneous circular cylinder. The frequency of the incident field* $f =$ 300MHz*, the width the object is* $a = 2\lambda$*.* N_θ *is the number of incident fields used to create the macro basis functions,* S_i *is the number of singular values retained in block* i *and* δ *is the normalised error in satisfying Equation (7.2.1).*

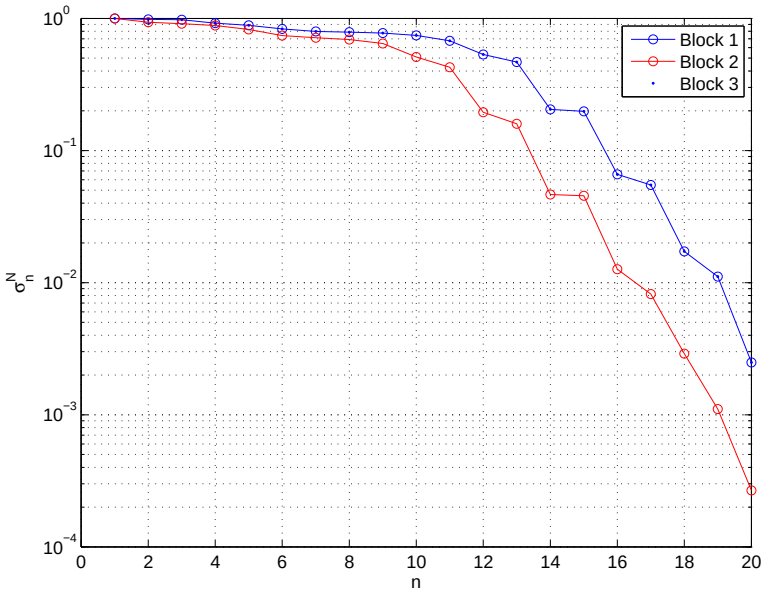

Figure 7.45: *On the log y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks,* $N_\theta = 20$.

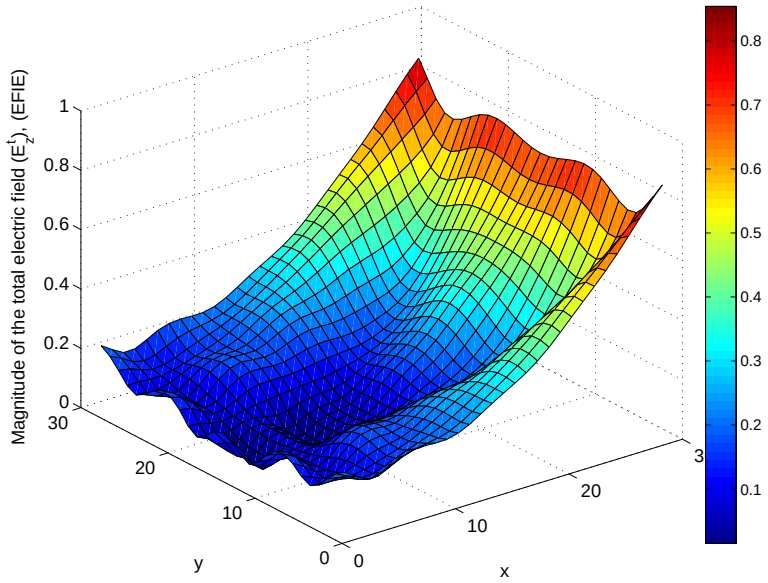

Figure 7.46: *A surface plot of the magnitude of* E_z^t *on the object at* $z = 0$ *in test case 3A computed using the EFIE-MoM approach.*

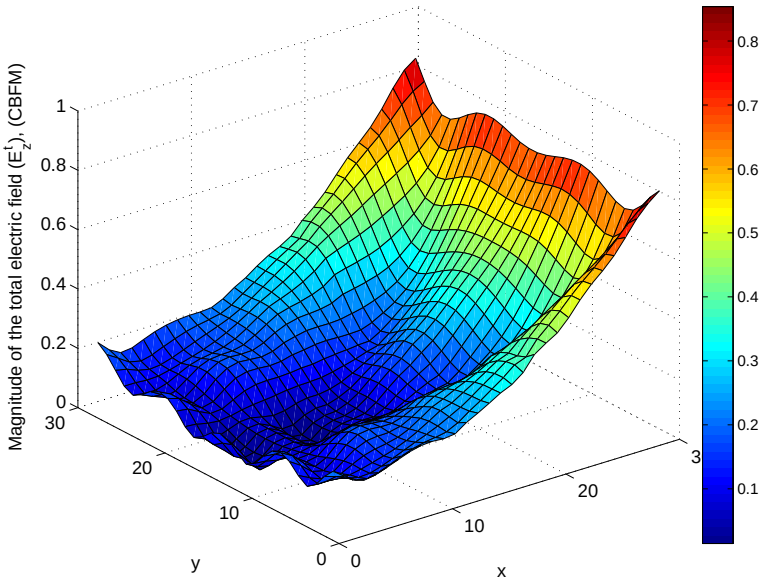

Figure 7.47: *A surface plot of the magnitude of E_z^t on the object at $z = 0$ computed using the Improved CBFM with parameters as in test case 3A(3).*

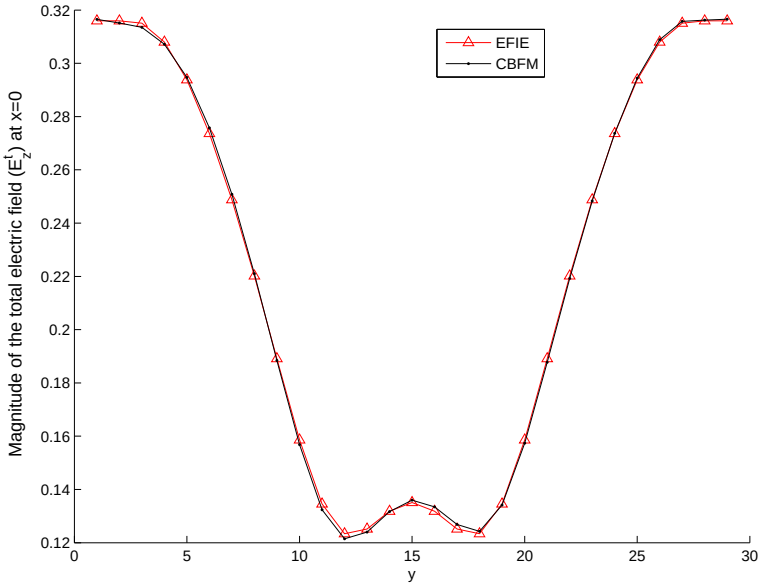

Figure 7.48: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 3A(3).*

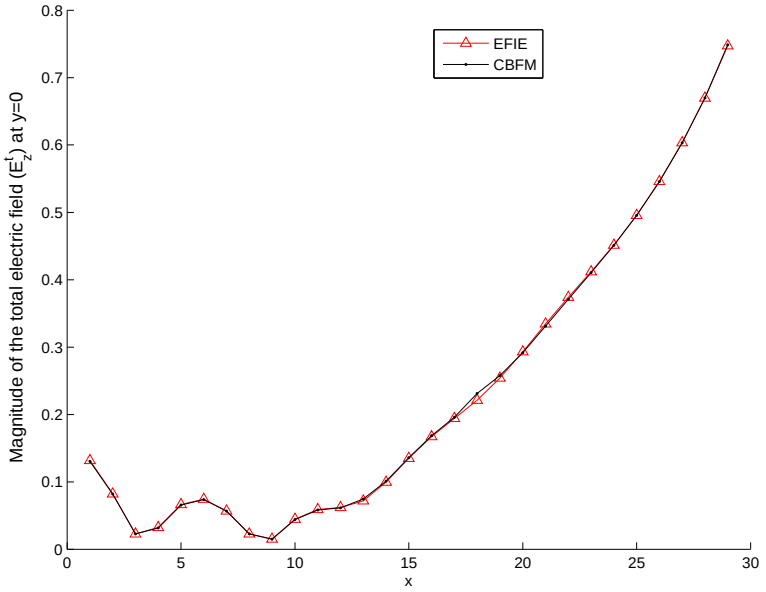

Figure 7.49: *A comparison of the magnitude of the E_z^t on the object at $z = y = 0$ computed using the, EFIE-MoM and the CBFM with parameters as in test case 3A(3).*

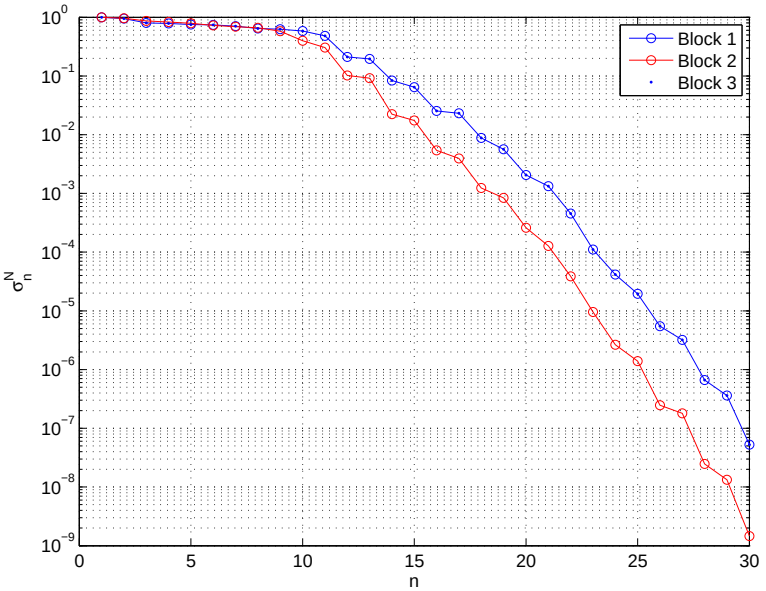

Figure 7.50: *On the log y-axis is the normalised singular values σ_n^N of all the 3 blocks, $N_\theta = 30$.*

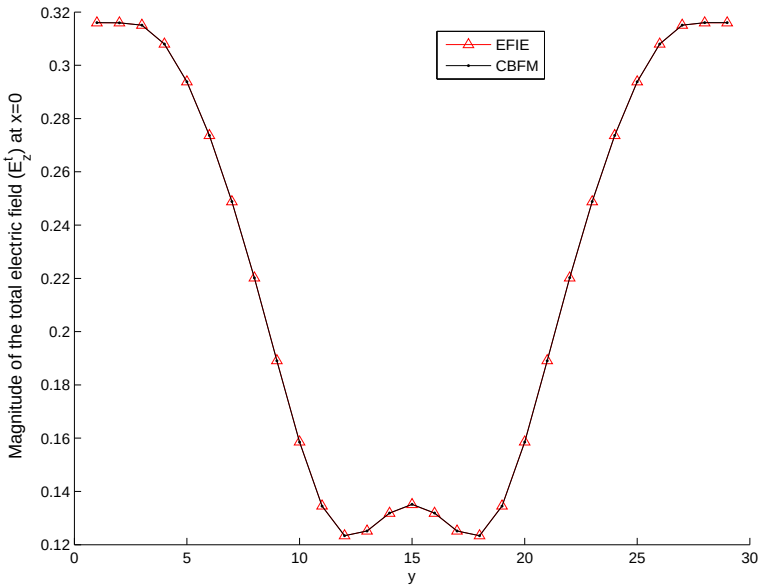

Figure 7.51: *A comparison of the magnitude of E_z^t on the object at $z = x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 3A(6).*

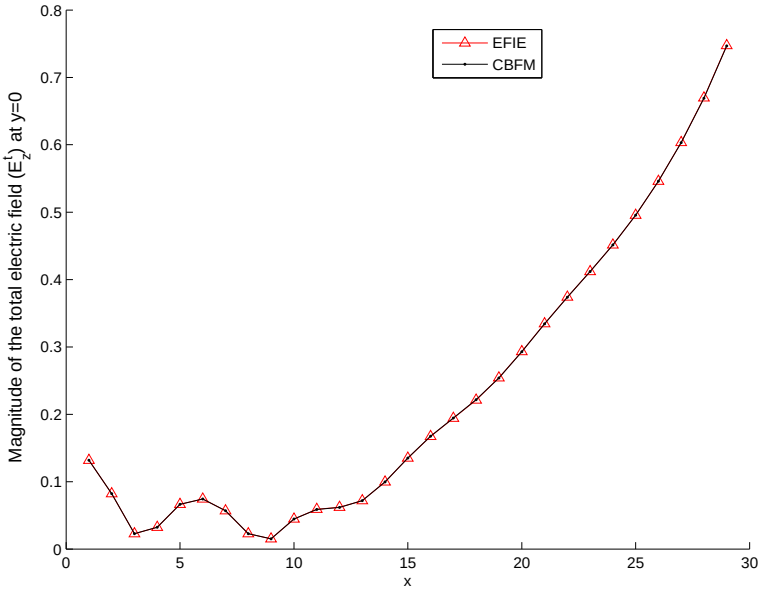

Figure 7.52: *A comparison of the magnitude of the E_z^t on the object at $z = y = 0$ computed using the, EFIE-MoM and the CBFM with parameters as in test case 3A(6).*

Figures (7.46)-(7.52) show that the Improved CBFM can be used to compute electromagnetic fields produced by the scattering of infinite homogeneous dielectric cylinders with square cross-section. Various test cases show that increasing N_θ always helps in achieving better accuracy.

Test case 3B

An infinite homogeneous dielectric cylinder of square cross-section with side $a = 4\lambda$ and $\epsilon_r = 2 - \jmath 0.8$ is analysed in this section. The object is illuminated with a normally incident TM plane wave at frequency $f = 300$ MHz. The cross-section of the object is discretised into $N = 3481$ cells. The MoM matrix resulting from the EFIE-MoM formulation in Section (3.7.2) is directly solved using the GMRES method with a tolerance $\delta \leq 1e-5$.

To decompose the object $N_1 = 1593$, $N_2 = 295$ and $N_3 = 1593$ pulse basis functions each are grouped together into three patches for the CBFM procedure. $N_\theta = 36$ plane waves are used to create MBFs for each patch. Normalised singular values of the matrix $J^{(i)}$ computed using the SVD is plotted in Figure (7.53). Table (7.5) shows a number of test cases with S_i number of CBFs used in each patch for the computation of E_z^t. The Galerkin procedure in (7.2.15) results in a reduced matrix $Z^c_{K\times K}$ which is inverted using the LU decomposition. The condition number κ of the reduced matrix is also tabulated.

Test case	N_θ	S_i			κ	δ
		S_1	S_2	S_3		
1	36	25	25	25	22.1127	0.0454
2		30	30	30	12.3546	0.0044
3		36	36	36	11.3728	5.87e-4

Table 7.5: *A Summary of the results from test case 3B involving scattering from infinite inhomogeneous circular cylinder. The frequency of the incident field* $f =$ 300MHz, *the width the object is* $a = 4\lambda$. N_θ *is the number of incident fields used to create the macro basis functions,* S_i *is the number of singular values retained in block* i *and* δ *is the normalised error in satisfying Equation (7.2.1).*

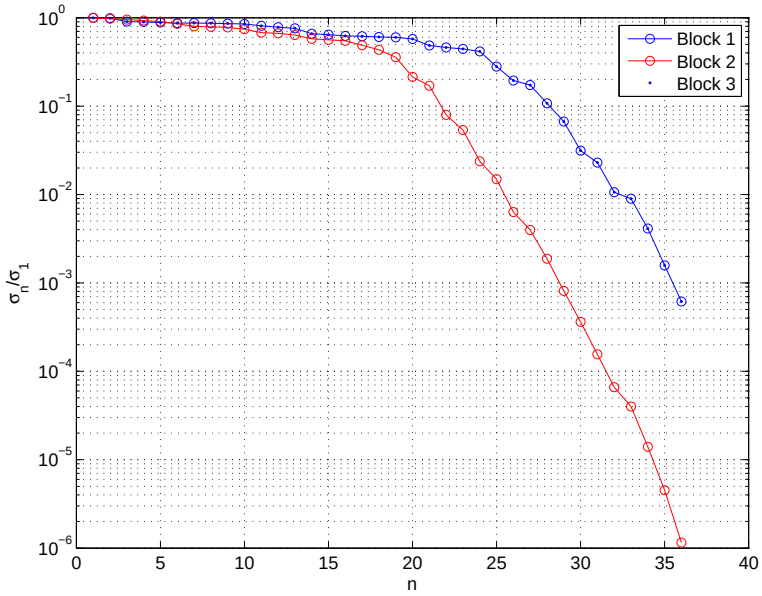

Figure 7.53: *On the log y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks.*

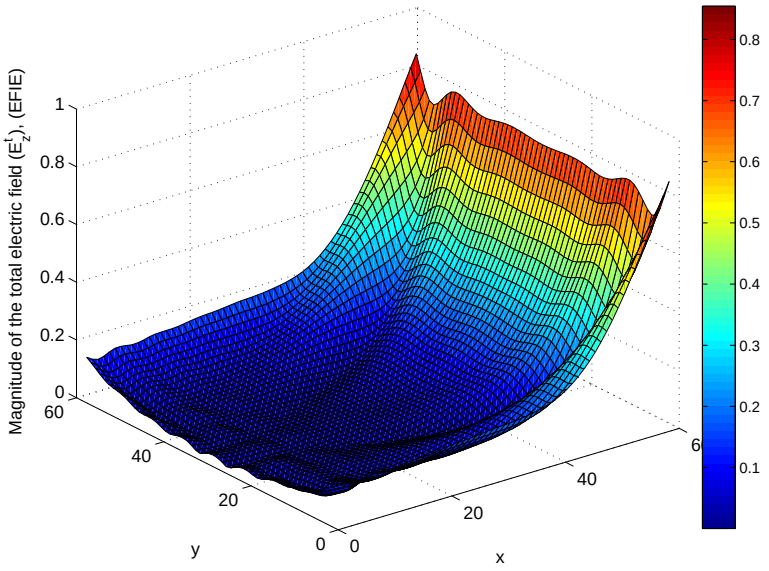

Figure 7.54: *A surface plot of the magnitude of E_z^t on the object computed using the EFIE-MoM approach in test case 3B.*

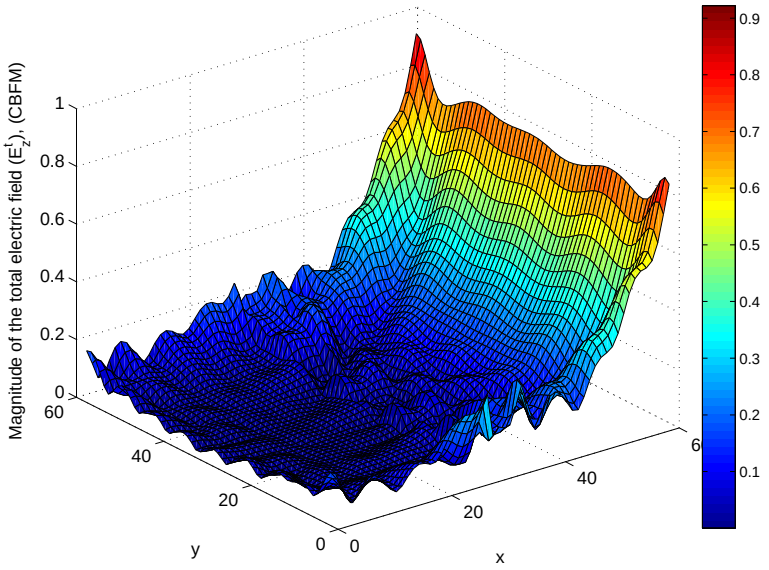

Figure 7.55: *A surface plot of the magnitude of E_z^t on the object computed using the CBFM with parameters as in test case 3B(1).*

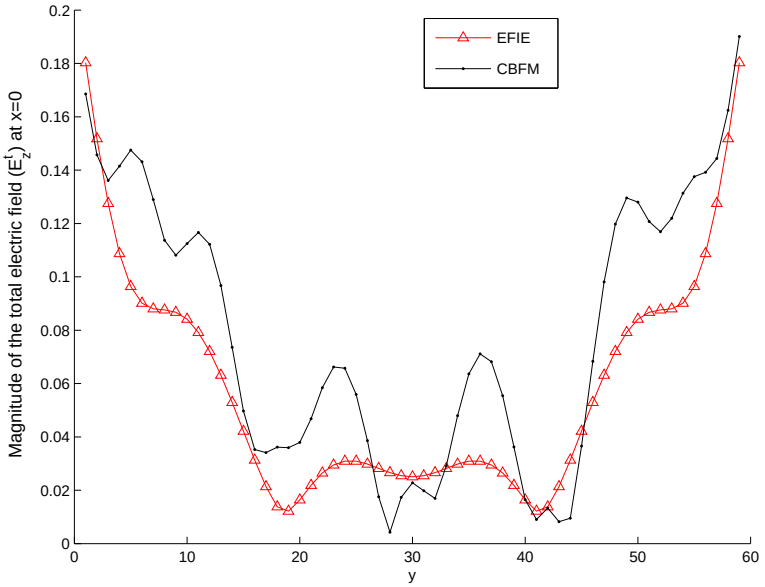

Figure 7.56: *A comparison of the magnitude of E_z^t on the object at $x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 3B(1).*

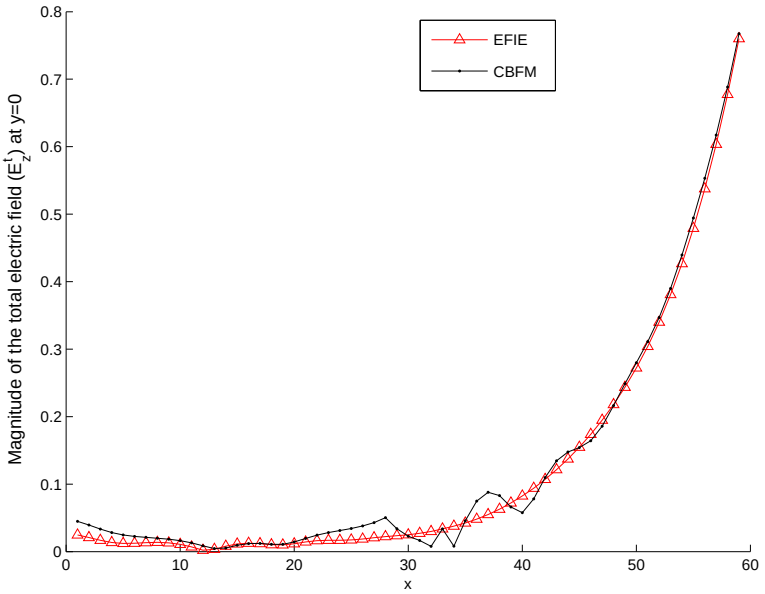

Figure 7.57: *A comparison of the magnitude of the E_z^t on the object at $y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 3B(1).*

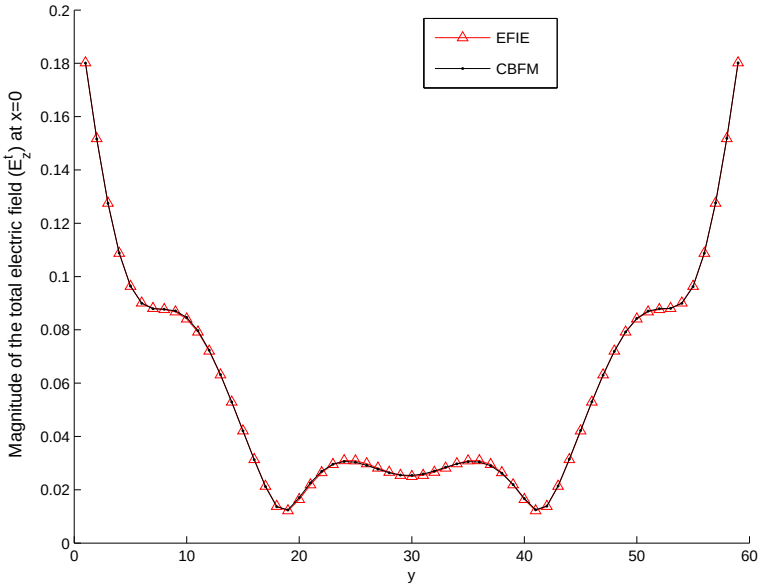

Figure 7.58: *A comparison of the magnitude of E_z^t on the object at $x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 3B(3).*

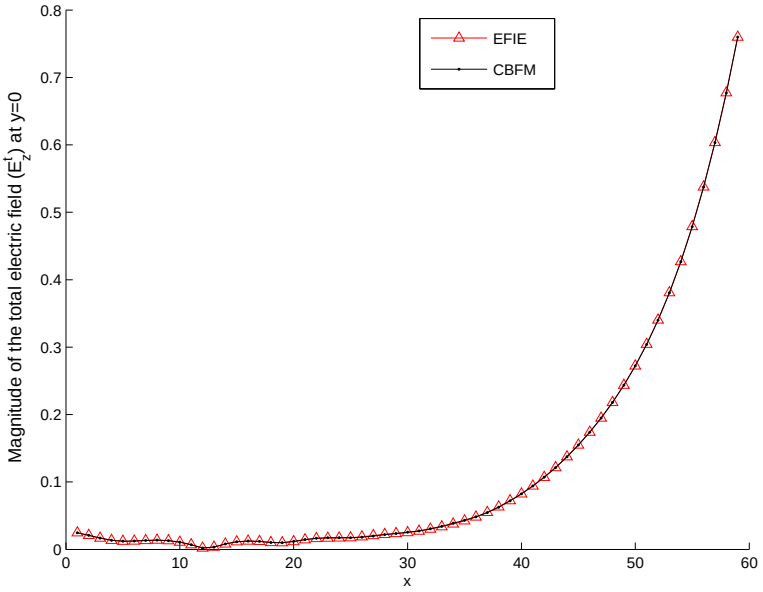

Figure 7.59: *Comparison of the magnitude of the E_z^t on the object at $y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 3B(3).*

In test case 3B(1) we observed that using only 25 CBFs in each block out of the available 36 does not produce good enough result. But using all the possible CBFs produces excellent agreement with the the iterative results computed using the EFIE-MoM method. This confirms that in the case of a large objects, the Improved CBFM is capable of producing accurate results.

Test case 3C

An infinite homogeneous dielectric cylinder of square cross-section with side $a = 4\lambda$ and $\epsilon_r = 4 - \jmath 0.8$ is analysed in this section. The object is illuminated with a normally incident TM plane wave at frequency $f = 300$ MHz. The cross-section of the object is discretised into $N = 6561$ cells. The MoM matrix resulting from the EFIE-MoM formulation is directly solved using the GMRES method with a tolerance $\delta \leq 1e-5$.

Three blocks with $N_1 = 3078, N_2 = 450$ and $N_3 = 3078$ pulse basis functions each grouped together into three blocks for the CBFM procedure. $N_\theta = 36$ plane waves are used to create N_θ CBFs for each patch. Normalised singular values of the matrix $J^{(i)}$ computed using the SVD is plotted in Figure (7.60). Table (7.6) shows 3 test cases with S_i number of CBFs used in each patch for the computation of E_z^t. The Galerkin procedure in Equation (7.2.15) results in a reduced matrix $Z^c_{K\times K}$ which is inverted using the LU decomposition. The condition number κ of the reduced matrix is also tabulated.

Test case 3C was implemented with MATLAB Version 7.2.0.283 (R2006a) running on a Dell Precision 670 workstation with 2 Intel Xeon dual-core processors and 4GB RAM with Fedora Core 7 OS, only one core was used. The creation of MoM elements took a little less than 24 hours while the generation of CBFs took 6 hours. Generation of the reduced matrix in test case 3C(3) took only 35 minutes.

Test case	N_θ	S_i			κ	δ
		S_1	S_2	S_3		
1	60	36	36	36	50.0062	0.0091
2		45	35	35	60.1421	0.0031
3		60	60	60	73.0110	0.0016

Table 7.6: *A Summary of the results from test case 3C involving scattering from infinite inhomogeneous circular cylinder. The frequency of the incident field* $f =$ 300MHz, *the width the object is* $a = 4\lambda$. N_θ *is the number of incident fields used to create the macro basis functions,* S_i *is the number of singular values retained in block* i *and* δ *is the normalised error in satisfying Equation (7.2.1).*

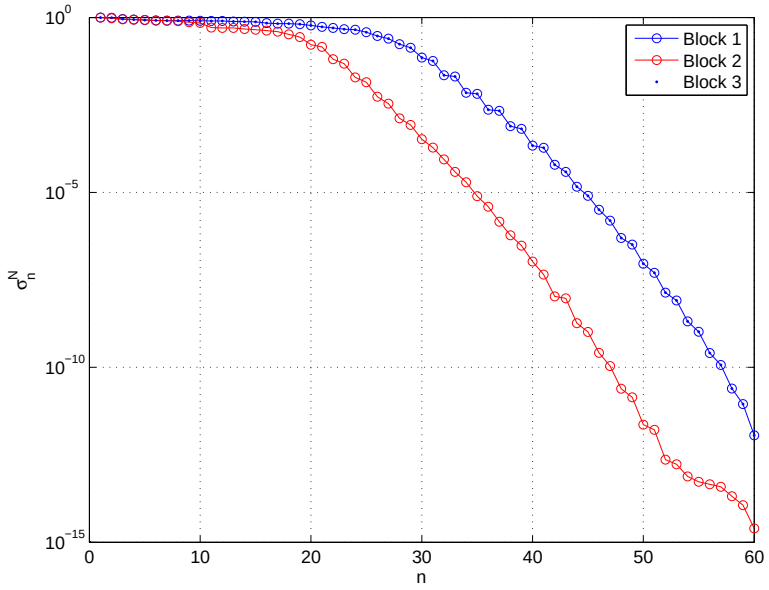

Figure 7.60: *On the log y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks,* $N_\theta = 60$.

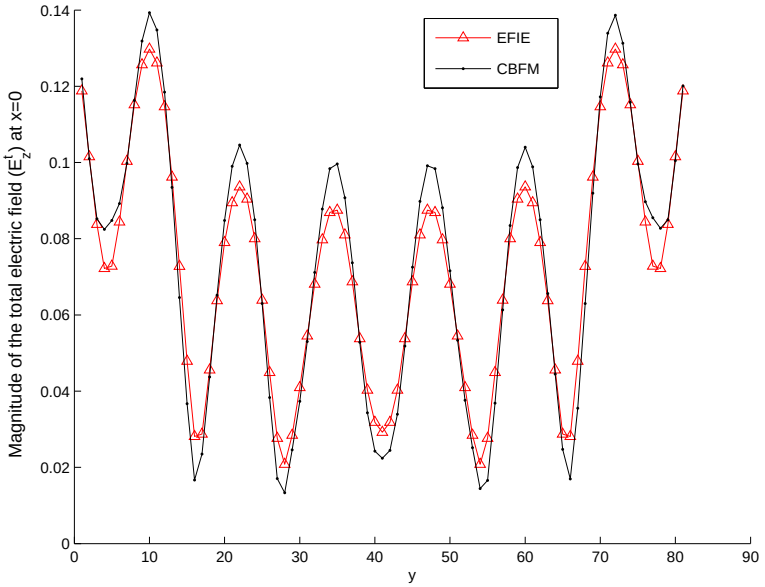

Figure 7.61: *Comparison of the magnitude of E_z^t on the object at $x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(2).*

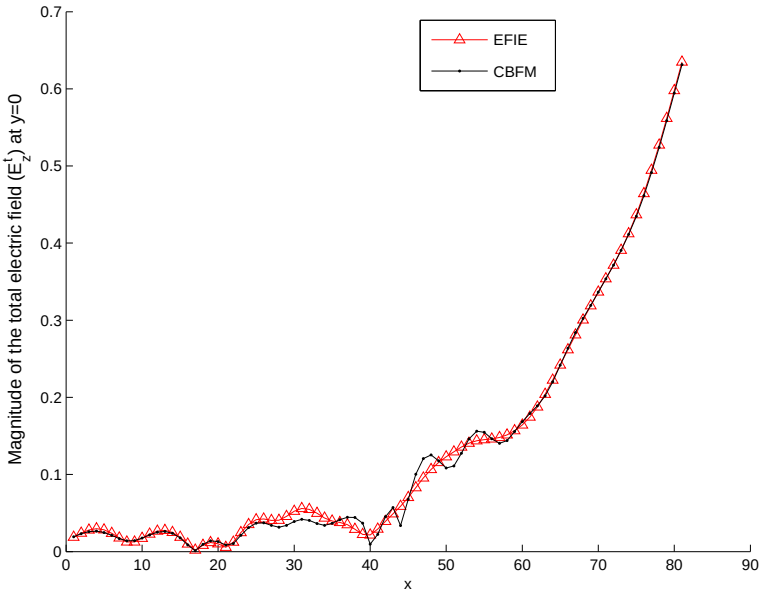

Figure 7.62: *Comparison of the magnitude of the E_z^t on the object at $y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(2).*

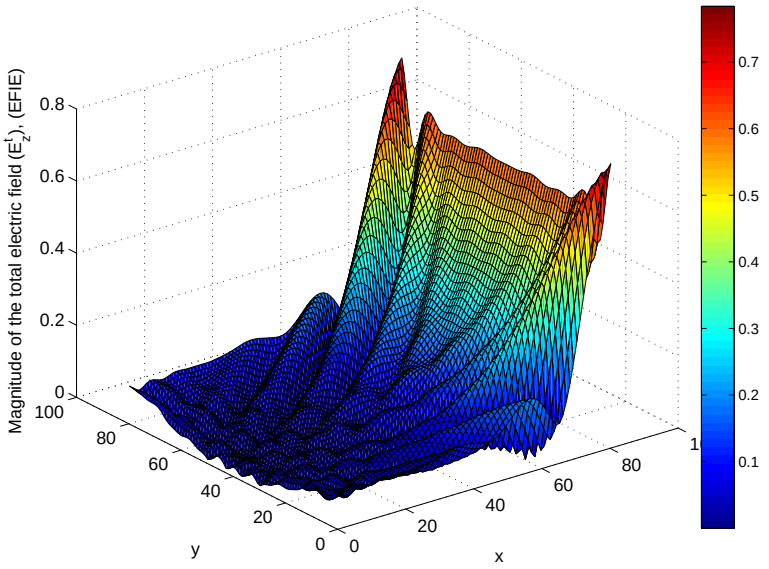

Figure 7.63: *Surface plot of the magnitude of E_z^t on the object computed using the EFIE-MoM approach in test case 2C(3).*

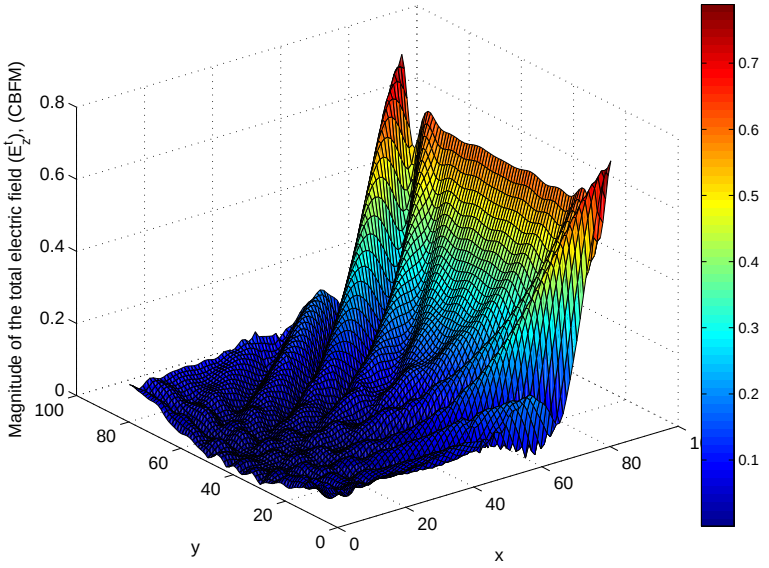

Figure 7.64: *Surface plot of the magnitude of E_z^t on the object computed using the CBFM with parameters as in test case 2C(3).*

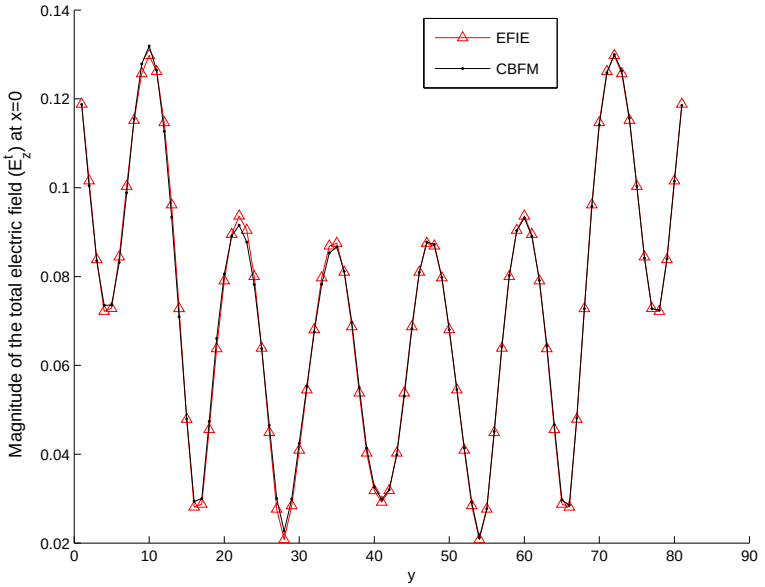

Figure 7.65: *Comparison of the magnitude of E_z^t on the object at $x = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(3).*

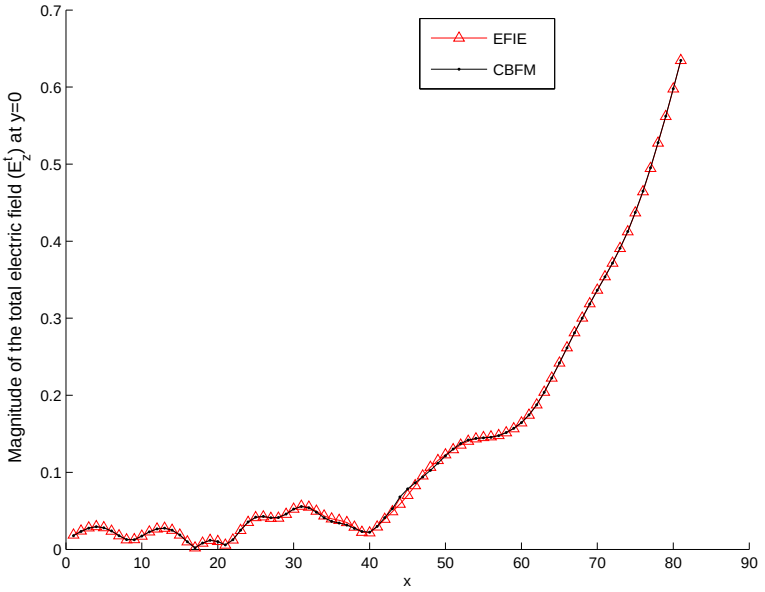

Figure 7.66: *Comparison of the magnitude of the E_z^t on the object at $y = 0$ computed using the EFIE-MoM and the CBFM with parameters as in test case 2C(3).*

Test case 3C confirms that, when dielectric constant increases, the increased wave-effects inside the object requires a higher N_θ. Even though σ_1^N and σ_{60}^N differ by several times in magnitude, we had to retain all the possible CBFs. Clearly as shown in Figures (7.65) and (7.66) the Improved CBFM provides accurate results.

7.4 The efficiency of the CBFM for E^is at an angle θ to the x-axis

It is desirable to verify that the efficiency of the new CBF method is invariant with the direction of propagation of the incident electric filed. Consider a number of plane waves E_θ^i that illuminate the object, where, θ is the angle between the $+x$-axis and the direction of propagation of the plane waves (PWs). We use N_θ plane waves to compute S_i CBFs for each patch.

An infinite homogeneous dielectric square cylinder with side $a = \lambda$ and $\epsilon_r = 2 - \jmath 0.8$ is analysed in this section. The cross-section of the object is discretised into $N = 225$ cells. The MoM equation resulting from the EFIE-MoM formulation in Section (3.7.2) is directly solved using the GMRES method with a tolerance $\delta \leq 1e{-}5$. To decompose the object, $N_1 = N_2 = N_3 = 75$ pulse basis functions are grouped together into three patches for the CBFM procedure. $N_\theta = 36$ plane waves are used to create N_θ MBFs for each patch. The normalised singular values of the matrix $J^{(i)}$ computed using the SVD are plotted in Figure (7.67). $S_i = 15$ is the number of CBFs used in each patch for the computation of E_z^t. The Galerkin procedure in Equation (7.2.15) results in a reduced matrix $Z_{45\times 45}^c$ which is inverted using the LU decomposition.

The object is illuminated with four incident TM plane waves of frequency $f =$ 800 MHz at $\theta = 0, 90, 180$ and 270. Figures (7.68) and (7.69) show that the improved

CBFM is efficient in producing accurate results irrespective of the angle of incidence of the incident field.

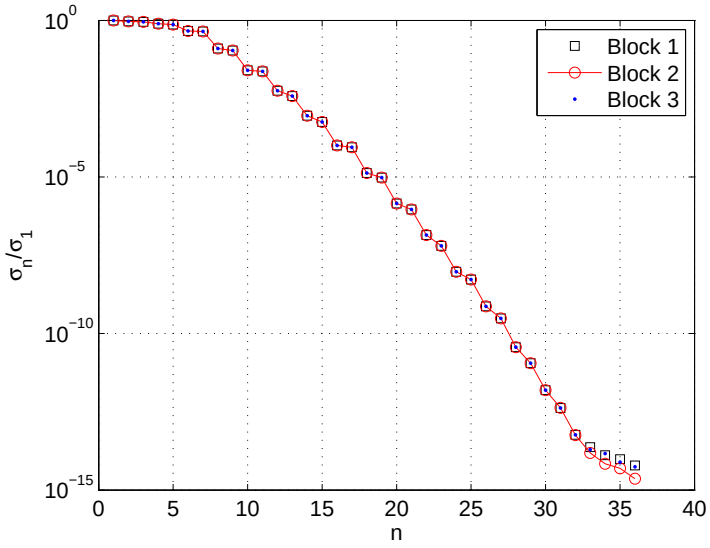

Figure 7.67: *On the y-axis is the normalised singular values* σ_n^N *of all the* 3 *blocks,* $N_\theta = 36$.

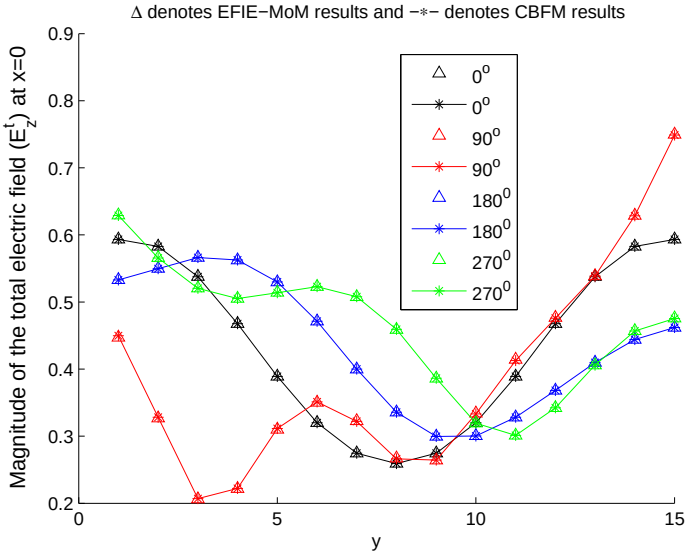

Figure 7.68: *Comparison of the magnitude of* E_z^t *on the object at* $x = 0$ *produced by the incident field* E^i *at* θ^o *computed using the EFIE-MoM and the CBFM.*

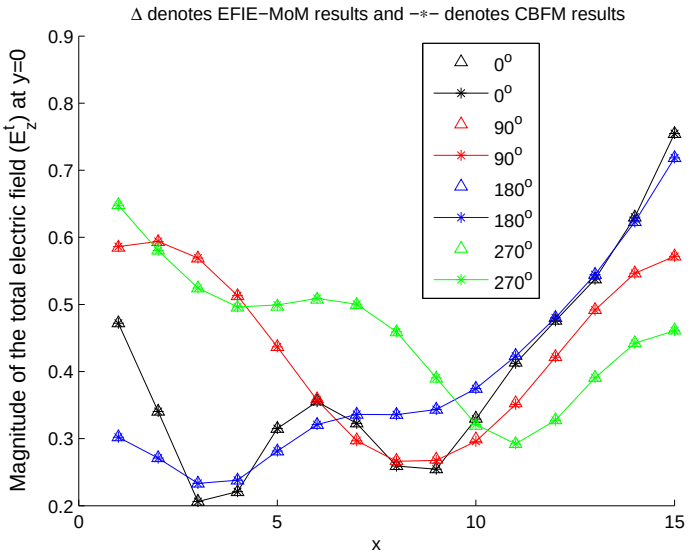

Figure 7.69: *Comparison of the magnitude of E_z^t on the object at $y = 0$ produced by the incident field E^i at θ^o computed using the EFIE-MoM and the CBFM.*

7.5 Conclusion

Several infinite dielectric objects with various cross-sectional types, sizes and electrical properties were analysed using the Improved CBFM procedure. Results show excellent agreement with results obtained using analytical and EFIE-MoM methods. The use of the new kind of CBFs that are invariant of the prorogation of the incident electromagnetic field is shown to be helpful in truncating edge effects.

When the object being analysed enclose inhomogeneities inside, special care must be taken by not breaking the inhomogeneity such that they go into different patches. If the decomposition is improper like this, the Improved CBFM would sometimes result in slight edge effects at the boundaries of the inhomogeneity. This can also be avoided by increasing N_θ, but if a proper decomposition can be done as stated above, that would be a better option.

It was observed in test cases 2C that sometimes one has to use more CBFs even though σ_1^N and σ_i^N (cut-off singular value) differ by several times in magnitude. Even in this case to get accurate results that overlies the EFIE-MoM solution, we had to use all the CBFs created. In any case the maximum number of CBFs $K = MN_\theta$ can be kept lower than M^2 as in the case of the original CBFM. Results presented in this chapter confirm that this Improved CBFM has all the capabilities to tune results to achieve excellent accuracy if required.

Chapter 8

Conclusion

The work presented in this dissertation evolved from an investigation to find domain decomposition methods suitable for the efficient electromagnetic analysis of dielectric objects. Several methods that are successful for perfect electrical conductors were investigated during the course of this work. The goal was to develop algorithms that could be used for the implementation of EM scattering analysis software that could handle electrically large dielectric objects.

8.1 Findings and Recommendations

In Chapter 5, variants of the forward backward method were applied for the EM analysis of dielectric objects. Investigations showed that the method is not suitable for computations for dielectric objects. Small changes in the electrical properties or size of the object resulted in unpredictable instability in the results. Changes made in the size of the sub-regions used for the decomposition of objects also resulted in instabilities. The most advanced method in this class of method was presented in Chapter 5. It is termed the Buffered Adaptive Strip Forward Backward (BASFB) Method and was successful to an extent. This method requires buffer strips of a width equal to that of the strips used in each block. Sometimes the required buffer width

had to be even greater than the strip width. The instabilities in this method come because of the consideration of the strips as individual objects. Although the buffer regions help us overcome this, as the dielectric constant increases, there are greater wave-effects inside the object and the buffer size required to truncate these spurious effects gets larger. The FBM is semi-iterative in nature and lacks a global nature in its algorithm. That is, in each iteration, individual blocks are solved at different steps to find the currents in that block. This is one reason behind the unpredictable behaviour of solutions obtained via FBMs. This is not desirable and motivated us to look for better methods.

A sub-domain multilevel method termed the Characteristic Basis Function Method (CBFM) was investigated for dielectric objects. Again, the idea was to break up the object into pieces. However, isolated solutions are merged into macro-domain basis functions. These MBFs are fit into a single global MoM system that considers the physics of the problem. This new MoM system is a compressed or reduced linear equation that is much smaller in size compared to the original MoM matrix equation. The reduced matrix equation is solved using a direct matrix inversion method. This global operation solves the problem faced by the FBM. In Chapter 6, the CBFM was used for the EM analysis of dielectric objects. The CBFM method in Chapter 6 used two kinds of MBFs for the ith decomposed patch of the object, the primary characteristic basis functions and secondary characteristic basis functions. The former account for the self interaction inside patch i and the later account for the mutual coupling of patch i with all the remaining patches. The method used a similar buffering scheme to that used in the FBM. This buffering scheme has been shown to be very useful in truncating the edge effect created by the decomposition of the object. However, the computation involved in computing all secondary basis functions significantly increases the over all operation count. In addition, although the method is shown to be

suitable for the EM analysis of scattering problems of dielectrics, the need for wide buffers for electrically large objects is disadvantageous. In Chapter 7 the CBFM is further developed to rectify the aforementioned disadvantages.

In Chapter 7, we propose a novel modification of the CBFM by using a new type of characteristic basis function. The new CBFs serve as a basis set that does not depend on the direction of propagation of the incident field. The CBFs are computed using a spectrum of plane-waves in several possible directions of incidence. The new CBFs are created for each patch treating them as separate isolated objects. The use of new CBFs helps in truncating spurious edge effects. The new CBFs span a subspace for representing the unknown total field. In this new version of the CBFM, which is specially tailored for scattering problems involving dielectric objects, we did not include buffer regions as in the conventional versions of the CBFM [46]. For small objects, we would be able to keep all or nearly all of the CBFs for all patches. In that case, all patches can be of the same width. For electrically larger objects, some patches were kept smaller compared to the bigger ones such that we can keep all the available CBFs in these patches. Use of these advanced CBFs is shown to be a helping factor in truncating the edge effects. This is due to the use of the CBFs that are invariant with the direction of propagation waves inside the object.

8.2 Future Study

The novel CBFM algorithm in Chapter 7 was developed for efficient implementation on desktop computers. Software developed using this algorithm will be able to do electromagnetic scattering simulations of large objects on desktop computers. Some of the most time-consuming parts of the algorithm, like generation of MBFs, SVD, etc. were successfully implemented using OpenMP [56, 9]. This showed several times speed

up and agreed with Gustafson's [47] law. Further work will be carried out with the full implementation of the new CBFM for shared memory parallel desktop computing. Graphic Processing Units (GPUs) can also be used for the efficient implementation of these algorithms. Work in this direction will further facilitate development of software for the EM analysis of large objects on desktop computers much faster than contemporary EM solvers in market.

Bibliography

[1] W. E. Arnoldi. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quart. Appl. Math*, 9:17–29, 1951.

[2] Bijilash Babu. A comparison of the forward backward method and the characteristic basis function method for the electromagnetic analysis of dielectric objects. *In preparation.*

[3] C. A. Balanis. *Advanced Engineering Electromagnetics.* John Wiley and sons, 1989.

[4] C. A. Balanis. *Antenna theory Analysis and design.* Wiley-Interscience; 3 edition, 2005.

[5] Jean-Pierre Berenger. *Perfect Matched Layer (PML) for Computational Electromagnetics.* Synthesis Lectures on Computational Electromagnetics, Springer.

[6] C. Brennan, B. Babu, M. Condon, and M. Mullen. A novel iterative solution of the three dimensional electric field integral equation. *IEEE Transactions on Antennas and Propagation*, 52(10):2781–2784, 2008.

[7] C. Brennan, P. Cullen, and M. Condon. A novel iterative solution of the three dimensional electric field integral equation. *IEEE Transactions on Antennas and Propagation*, 52(10):2781–2784, 2004.

[8] F. Catedra, E. Garcia, C. Delgado, F.S. de Adana, and R. Mittra. Development of an efficient rigorous technique based on the combination of CBFM and MLFMA to solve very large electromagnetic problems. In *Electromagnetics in Advanced Applications, 2007. ICEAA 2007. International Conference on*, pages 579 –582, 2007.

[9] B. Chapman, G. Jost, and R. van der Pas. *Using OpenMP: Portable Shared Memory Parallel Programming*. The MIT Press, 2007.

[10] R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: a pedestrian prescription. *IEEE Transactions on Antennas and Propagation*, 35(3):7–12, 1993.

[11] C. Delgado, F. Catedra, and R. Mittra. A numerically efficient technique for orthogonalizing the basis functions arising in the solution of electromagnetic scattering problems using the CBFM. In *Antennas and Propagation Society International Symposium, 2007 IEEE*, pages 3608 –3611, 2007.

[12] C. Delgado, M.F. Catedra, and R. Mittra. Application of the characteristic basis function method utilizing a class of basis and testing functions defined on nurbs patches. *Antennas and Propagation, IEEE Transactions on*, 56(3):784–791, 2008.

[13] C. Delgado, E. Garcia, F. Catedra, and R. Mittra. Generation of characteristic basis functions defined over large surfaces by using a multilevel approach. *Antennas and Propagation, IEEE Transactions on*, 57(4):1299–1301, 2009.

[14] G. Goulb and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996.

[15] D H Griffel. *Applied Functional Analysis*. Wiley, New York, 1981.

[16] R. F. Harrington. *Time-Harmonic Electromagnetic Fields*. McGraw-Hill, 1961.

[17] S. Hay, F. Cooray, J. O'Sullivan, Neng-Tien Huang, and R. Mittra. Numerical and experimental studies of a dual-polarized planar connected-array antenna for the australian square kilometer array pathfinder. In *Antennas and Propagation Society International Symposium, 2009. APSURSI '09. IEEE*, pages 1–4, 2009.

[18] M. R Hestenes and E. Stiefel. Method of conjugate gradient for solving linear equations. *Journal of Research of the National Bureau of Standards*, 49:409–439, 1952.

[19] Jan S. Hesthaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer.

[20] D. Holliday. Forward-backward: A new method for computing low-grazing angle scattering. *IEEE Trans. Antennas and Propagat.*, 44:722–729, 1996.

[21] L. Hu, L. W. Li, and R. Mittra. Electromagnetic scattering by finite periodic arrays using the characteristic basis function and adaptive integral methods. *Antennas and Propagation, IEEE Transactions on*, PP(99), 2010.

[22] M. Hurst. An efficient approach to moment method analysis of large antenna arrays. *IEEE AP-S/URSI Int Symp, URSI Digest, Salt Lake City*, page 170, 2000.

[23] J D Jackson. *Classical Electrodynamics*. Wiley Interscience, New York, 1975.

[24] Leo C. Kempel John L. Volakis, Arindam Chatterjee. *Finite Element Method Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications*. IEEE Press Series on Electromagnetic Wave Theory ,Wiley-IEEE Press, 1998.

[25] D. A. Kapp and G. S. Brown. A new numerical method for rough surface scattering calculations. *IEEE Trans. Antennas and Propagat.*, 44:711–721, 1996.

[26] S. Kapur and D.E. Long. IES^3: Efficient electrostatic and electromagnetic simulation. *Computational Science Engineering, IEEE*, 5(4):60–67, 1998.

[27] E.F Knott, J. F. Shaeffer, and M. T. Tuley. *Radar Cross Section*. Society for Industrial and Applied Mathematics,Philadelphia, PA, USA, 2003.

[28] R.G. Kouyoumjian. The geometrical theory of diffraction and its application, numerical and asymptotic techniques in electromagnetics, R. Mittra, ed. *Berlin, Heidelberg, New York: Springer-Verlag*, 1975.

[29] J. D. Krauss. *Antennas, 2nd Edition*. McGraw-Hill Companies, 1988.

[30] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45:225–280, 1950.

[31] J. Laviada, M. R. Pino, F. Las-Heras, and R. Mittra. Efficient calculation of the reduced matrix in the characteristic basis functions method. *IEEE Antennas and Propagation Society International Symposium, AP-S 2008.*, ISBN: 978-1-4244-2041-4:1–4, 2008.

[32] D. E. Lawrence and K. Sarabandi. Acoustic and electromagnetic wave interaction: Analytical formulation for acousto-electromagnetic scattering behavior of a dielectric cylinder. *IEEE Transactions on Antennas and Propagation*, 49,10:1382, 2001.

[33] E. Lucente, A. Monorchio, and R. Mittra. An iteration-free mom approach based on excitation independent characteristic basis functions for solving large multi-scale electromagnetic scattering problems. *Antennas and Propagation, IEEE Transactions on*, 56(4):999–1007, 2008.

[34] Rob Maaskant. *Analysiws of Large Antenna Systems*. PhD thesis, Technische Universiteit Eindhoven, 2010.

[35] J. C. Maxwell. A treatise on electricity and magnetism. *Clarendon Press Series*, 1873.

[36] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia.

[37] R. Mittra. A look at some challenging problems in computational electromagnetics. *Antennas and Propagation Magazine, IEEE*, 46(5):18–32, oct. 2004.

[38] R. Mittra. A novel domain decomposition technique for solving very large problems in frequency and time domains. In *Antennas and Propagation, 2007. EuCAP 2007. The Second European Conference on*, pages 1–4, 2007.

[39] R. Mittra, V. V. S. Prakash, and J. Yeo. Some novel techniques for efficient analysis of large arrays and frequency selective radomes. In *Antennas and Propagation, 2003. (ICAP 2003). Twelfth International Conference (Conf. Publ. No. 491)*, volume 2, pages 462–465 vol.2, 31 2003.

[40] Raj Mittra. *Private Communication*.

[41] Peter Monk. *Finite Element Methods for Maxwell's Equations*. Oxford University Press, USA, 2003.

[42] G. Ogucu, R. Mittra, and Kai Du. An interpolation algorithm to reduce the reduced matrix fill-time in CBFM. *Antennas and Wireless Propagation Letters, IEEE*, 8:457–460, 2009.

[43] O. Ozgun, R. Mittra, and M. Kuzuoglu. Parallelized characteristic basis finite element method (CBFEM-MPI)-a non-iterative domain decomposition algorithm

for electromagnetic scattering problems. *Journal of Computational Physics*, 228:2225–2238, 2009.

[44] A. F. Peterson, S. L. Ray, and R. Mittra. *Computational Methods for Electromagnetics*. Wiley-IEEE Press Series on Electromagnetic Wave Theory, New York, 1997.

[45] V. V. S. Prakash. RCS computation over a frequency band using the characteristic basis and model order reduction method. In *Antennas and Propagation Society International Symposium, 2003. IEEE*, volume 4, pages 89–92 vol.4, 22-27 2003.

[46] V. V. S. Prakash and R. Mittra. Characteristic basis function method: a new technique for efficient solution of method of moments matrix equations. *Microwave and Optical Technology Letters*, 36:95–100, 2003.

[47] M. J. Quinn. *Parallel Programming in C with MPI and OpenMP*. McGraw-Hill, 2003.

[48] S. Rao, D. Wilton, and A. Glisson. Electromagnetic scattering by surfaces of arbitrary shape. *Antennas and Propagation, IEEE Transactions on*, 30(3):409 – 418, 1982.

[49] J. H. Richmond. Scattering by a dielectric cylinder of arbitrary cross-section shape. *IEEE Transactions on Antennas and Propagation*, 13:334–341, 1965.

[50] J. H. Richmond. Te-wave scattering by a dielectric cylinder of arbitrary cross-section shape. *IEEE Transactions on Antennas and Propagation*, 14:460–464, 1966.

[51] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics,Philadelphia, PA, USA, 2003.

[52] T. Sarkar. A note on the choice weighting functions in the method of moments. *Antennas and Propagation, IEEE Transactions on*, 33(4):436–441, 1985.

[53] T. Sarkar, A. Djordjevic, and E. Arvas. On the choice of expansion and weighting functions in the numerical solution of operator equations. *Antennas and Propagation, IEEE Transactions on*, 33(9):988–996, 1985.

[54] B. Smith, P. Bjorstad, and W. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, 2004.

[55] J. Song, C. C. Lu, and W.C. Chew. Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects. *IEEE Trans Antennas Propagat*, 45:1488–1493, 1997.

[56] The OpenMP API specification for parallel programming. http://openmp.org.

[57] Gilbert Strang. *Introduction to Linear Algebra.* Brooks Cole, 4th edition, New York, 2005.

[58] Gilbert Strang. *Introduction to Applied Mathematics.* Wellesley Cambridge Press, 2007.

[59] J. A. Stratton. *Electromagnetic Theory.* McGraw-Hill Companies; 1st edition, New York, 1941.

[60] Y. F. Sun, C. H. Chad, R. Mittra, and L. Tsang. Characteristic basis function method for solving large problems arising in dense medium scattering. *Antennas and Propagation Society International Symposium*, pages 1068–1071, 2003.

[61] E. Suter and J. Mosig. A subdomain multilevel approach for thc mom analysis of large planar antennas. *Microwave and Optical Technology Letters*, 26:270–277, 2000.

[62] R. Suter and J. Mosig. A subdomain multilevel approach for the mom analysis of large antennas. *Micro. Opt Tech Letter*, 35(3):7–12, 2001.

[63] A. Taflove. Review of the formulation and applications of the finite-difference time-domain method for numerical modeling of electromagnetic wave interactions with arbitrary structures. *Wave Motion*, 10:547–582, 1998.

[64] G. Tiberi, E. Lucente, R. Mittra, and A. Monorchio. Characteristic basis function method (CBFM) for analyzing em scattering by large conducting structures with apertures. In *Antennas and Propagation Society International Symposium, 2008. AP-S 2008. IEEE*, pages 1–4, 2008.

[65] L. N. Trefethen and D. Bau. *Numerical Linear Algebra.* SIAM, Philadelphia, 1997.

[66] Bill Trowbridge. Integral equations in electromagnetics. *Intl. Journal of Numerical Modelling: Electronic networks, devices, and fields*, 9:3–17, 1996.

[67] P. Y. Ufimtsev. Elementary edge waves and the physical theory of diffraction. *Electromagnetics*, 11:125–159, 1991.

[68] J. L. Volakis and L. C. Kempel. Electromagnetics: computational methods and considerations. *IEEE computational science and engineering*, 2:42–57, 1995.

[69] J.C. West. Integral equation formulation for iterative calculation of scattering from lossy rough surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 38(4):1609–1615, 2000.

[70] J.C. West and J.M. Sturm. On iterative approaches for electromagnetic rough-surface scattering problems. *IEEE Transactions on Antennas and Propagation*, 47(8):1281–1288, 1999.

[71] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, Cambridge, UK, 2002.

[72] K. S. Yee. Numerical solution of initial boundary value problems involving maxwells equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14:302–307, 1966.