ADVANCES IN THE ADJOINT VARIABLE METHOD FOR TIME-DOMAIN ELECTROMAGNETIC SIMULATIONS
ADVANCES IN THE ADJOINT VARIABLE METHOD FOR
TIME-DOMAIN ELECTROMAGNETIC SIMULATIONS

By
YU ZHANG, M. A. Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
© Copyright by Yu Zhang, May 2015
Title: Advances in the adjoint variable method for time-domain electromagnetic simulations

Author: Yu Zhang
B. Sc. in Electrical Engineering (University of Manitoba)
M. A. Sc. (McMaster University)

Supervisor: Mohamed H. Bakr, Professor
Department of Electrical and Computer Engineering
B. Sc., M. Sc. (Cairo University)
Ph. D. (McMaster University)
P. Eng. (Province of Ontario, Canada)

NUMBER OF PAGES: xv, 134
Abstract

This thesis covers recent advances in the adjoint variable method for sensitivity estimations through time-domain electromagnetic simulations. It considers both frequency-independent and frequency-dependent response functions, and at the same time, provides a novel adjoint treatment for addressing dispersive sensitivity parameters in the material constitutive relation. With this proposed adjoint technique, response sensitivities with respect to all $N$ sensitivity parameters can be computed through at most one extra simulations regardless of the value of $N$. This thesis also extends the existing adjoint technique to estimate all $N^2$ second-order sensitivity entries in the response Hessian matrix through $N$ additional simulations. All adjoint sensitivity techniques presented in this thesis are numerically validated through various practical examples. Comparison shows that our produced adjoint results agree with those produced through central finite difference approximations or through exact analytical approaches.
Contents

Contents iii

List of Figures vi

Acknowledgments xiv

1 Introduction 1

1.1 Adjoint variable method for dispersive mediums 2

1.2 The second-order adjoint variable method 4

1.3 Outline of the thesis 5

1.4 Summary of contributions 7

2 The Adjoint Variable Method for Frequency-Independent Constitutive Parameters 15

2.1 The wave equation in the discretized space 16

2.2 Sensitivity parameters and the perturbed wave equation 21

2.3 The response function and its sensitivity gradient 30
2.4 The adjoint sensitivity expression ........................................... 33
2.5 The solution of the adjoint problem ........................................... 36
2.6 Algorithm implementation and numerical examples ......................... 40
  2.6.1 Example: the One-Dimensional Dielectric Slab ......................... 42
  2.6.2 The Anti-Reflection Dielectric Coating ................................. 44
2.7 Summary ................................................................................. 47

3 Adjoint Sensitivity Analysis for Frequency-Dependent Response Functions
  3.1 The response function for scattering-parameters ............................. 53
  3.2 The expression of the complex adjoint sensitivity ......................... 55
  3.3 The self-adjoint formula .......................................................... 58
  3.4 Algorithm implementation and numerical examples ....................... 60
    3.4.1 A multilayer anti-reflection dielectric coating ....................... 62
    3.4.2 A dielectric resonator antenna with matching insert ................ 66
  3.5 Summary ................................................................................. 69

4 The Adjoint Variable Method for Frequency-Dependent Constitutive Parameters
  4.1 The system matrix for dispersive mediums .................................. 73
    4.1.1 The Lorentz dispersion ....................................................... 74
    4.1.2 The Drude dispersion ....................................................... 75
Contents

4.1.3 The Debye dispersion and the conductivity term .......................... 76
4.1.4 Multiple dispersion poles .................................................. 77
4.2 The solution of the dispersive adjoint equation ............................. 79
4.3 Algorithm implementation and numerical examples ....................... 82
  4.3.1 Thin Lorentz slab in a parallel-plate waveguide ....................... 83
  4.3.2 Tunneling through a metamaterial slab ................................. 86
  4.3.3 The add-drop coupler with split-ring resonator ........................ 88
  4.3.4 Gold split ring with normal plane-wave incidence ................... 90
4.4 Summary ................................................................. 94

5 The Second-Order Adjoint Variable Method for

  Time-Domain Methods .................................................. 97
  5.1 The second-order response sensitivities ................................ 98
  5.2 The second-order adjoint variable method ................................ 99
  5.3 Algorithm implementation and numerical examples ..................... 104
    5.3.1 A multilayer anti-reflection dielectric coating .................... 105
    5.3.2 A dispersive discontinuity .......................................... 109
    5.3.3 A dielectric resonator filter ....................................... 112
    5.3.4 A dielectric resonator antenna with matching insert .......... 115
  5.4 Summary ................................................................. 118

6 Conclusions ............................................................... 122
List of Figures

2.1 Illustration of cell interpolation over a Yee cell in Cartesian system, where all electric fields are computed at the center of each cell edge. With cell-interpolation, the field components at the geometric center of the cell (marked by ) are approximated using

\[ E_\kappa = \frac{1}{4}(E_{\kappa,1} + E_{\kappa,2} + E_{\kappa,3} + E_{\kappa,4}) \] for \( \kappa \in \{x, y, z\} \). ....................... 22

2.2 E-plane view of a parallel plate waveguide loaded with a dielectric slab. The sensitivity parameters of interest are the thickness and the permittivity of the loaded slab \( p = [D \quad \epsilon]^T \). For illustration purpose, the computational volume, as shown in figure, is meshed in a way that the cell sizes is uniform along each respective direction. ... 24

2.3 Illustration of step-wised on-grid perturbations of the dielectric slab in a parallel-plate waveguide. The perturbation surfaces are shown by dashed line. Cells within volume of perturbations are overlayed with colored circles. ................................. 25
2.4 Illustration of the one-to-one field mapping. In all cases shown in Figure 2.3, the entries of $\tilde{\mathbf{E}}$ computed at the cells marked by circles in Figure 2.3 are approximated using the entries of $\mathbf{E}$ computed at the cells marked by squares. ........................................... 28

2.5 The adjoint sensitivities of the energy function (2.18) with respective to the slab thickness $D$ (a) and the relative permittivity of the slab $\epsilon_r$ (b). ................................................................. 44

2.6 The six-layer anti-reflection dielectric coating. The incident region is semiconductor with relative permittivity $\epsilon_a = 12.25$, while the target region is free space with $\epsilon_b = 1.0$. The relative permittivity of each dielectric slab is given by $\epsilon_1 = 13.5$, $\epsilon_2 = 8.41$, $\epsilon_3 = 4.86$, $\epsilon_4 = 2.52$, $\epsilon_5 = 1.46$, and $\epsilon_6 = 1.08$. The nominal width of each slab can be obtained using $w_i = z_i - z_{i-1} = \pi c_0 / 2 \sqrt{\epsilon_i \omega_0}$ for $i = 1, 2, ..., 6$, where $\omega_0$ is the center frequency of the pass-band. ........... 45

2.7 The adjoint sensitivities of the energy function (2.18) with respective to the thickness $w_1$ (a) and the relative permittivity $\epsilon_{r1}$ (b) of the first dielectric layer in the anti-reflection structure. ............... 47
2.8 The adjoint sensitivities of the energy function (2.18) with respective to the thickness $w_3$ (a) and the relative permittivity $\epsilon_{r,3}$ (b) of the third dielectric layer in the anti-reflection structure; the adjoint sensitivities of the energy function (2.18) with respective to the thickness $w_5$ (c) and the relative permittivity $\epsilon_{r,5}$ (d) of the fifth dielectric layer in the anti-reflection structure. .......... 48

3.1 A diagram of a 6-layer dielectric anti-reflection structure whose outer regions have permittivity of $\epsilon_a = 12.25$ and $\epsilon_b = 1.0$ (free space). The relative permittivity of each dielectric slab is given by $\epsilon_1 = 13.5$, $\epsilon_2 = 8.41$, $\epsilon_3 = 4.86$, $\epsilon_4 = 2.52$, $\epsilon_5 = 1.46$, and $\epsilon_6 = 1.08$. Their nominal widths satisfy $w_i = z_i - z_{i-1} = \pi c_0 / (2 \sqrt{\epsilon_i \omega_0})$ for $i = 1, 2, ..., 6$, where $\omega_0$ is the center frequency of the passband. .......... 63

3.2 The adjoint sensitivities of the $S_{11}$ parameter with respect to the thickness $w_1$ (a) and the relative permittivity $\epsilon_1$ (b) of the first layered medium. ......................................................... 64

3.3 The adjoint sensitivities of the $S_{11}$ parameter with respect to the thickness $w_3$ (a) and the relative permittivity $\epsilon_3$ (b) of the third layered medium; the adjoint sensitivities of the $S_{11}$ parameter with respect to the thickness $w_5$ (c) and the relative permittivity $\epsilon_5$ (d) of the fifth layered medium. .................. 65
3.4 The dielectric-resonator antenna (DRA) with a thin layer of dielectric insert. ................................................................. 66

3.5 The adjoint sensitivities of the $S_{11}$ parameter with respect to the permittivity $\varepsilon_{r,\text{insert}}$ (a) and the length $d_1$ (b) of the dielectric insert; the adjoint sensitivities of the $S_{11}$ parameter with respect to the length $d_2$ (c) and the height $h_2$ (d) of the dielectric resonator. .... 67

3.6 The adjoint sensitivities of the $S_{11}$ parameter with respect to the width $w_1$ of the dielectric insert. .................................................. 68

4.1 Parallel-plate waveguide loaded with a thin Lorentz slab. ............... 83

4.2 Adjoint sensitivities of the scattering parameters with respect to the thickness $L$ (a) and the high-frequency dielectric constant $\varepsilon_\infty$ (b) of the thin slab. ................................................................. 84

4.3 Adjoint sensitivities of the scattering parameters with respect to the resonant frequency $\omega_p$ of the second Lorentz dispersion pole (a), the damping frequency $\zeta$ (b) and the dispersion coefficient $\epsilon_p$ (c) of the third Lorentz dispersion pole of the thin slab. ....................... 85

4.4 180° bent structure loaded with a metamaterial slab. ......................... 86

4.5 Adjoint sensitivities of the scattering parameters with respect to the thickness $D$ (a), the plasma frequency $\omega_p$ (b), and the collision time constant $\tau_c$ of the metamaterial slab. ......................... 87
List of Figures

4.6 The add-drop coupler with a SRR. The colored region represents silver, while uncolored region is air. ............................................. 89

4.7 Adjoint sensitivities of the scattering parameters with respect to the edge $p_1$ (a) and the edge $p_2$ (b) of the add-drop coupler. ............... 90

4.8 Adjoint sensitivities of the scattering parameters with respect to the edge $p_3$, the edge $p_5$ (b), the edge $p_7$ (c), and the edge $d$ (bottom) of the add-drop coupler. ............................................. 91

4.9 Gold split ring with normal plane-wave incidence (a); cross section of the gold split ring with all edges labeled (b). ......................... 92

4.10 Adjoint sensitivities of the scattering parameters with respect to the edge $p_1$ (a), the edge $p_3$ (b), the edge $p_4$, and the edge $p_5$ of the gold split ring. .................................................. 93
5.1 Illustration of changes in the system matrix $M$ and the field solution $E$ due to shape perturbations on a rectangle discontinuity (shown by dashed region) in an FDTD grid. (a) Perturbing the shape parameter $p_i$ updates the field solution to be $E + \Delta_i^+ E$; (b) perturbing the shape parameter $p_j$ updates the field solution to be $E + \Delta_j^+ E$; (c) Simultaneously perturbing both $p_i$ and $p_j$ updates the field solution to be $E + \Delta_{ij}^{++} E$; (d) affected cells corresponding to non-zero entries in $\Delta_{ij}^{++} M - \Delta_i^+ M$; (e) affected cells corresponding to non-zero entries in $\Delta_{ij}^{++} M - \Delta_j^+ M$; (f) affected cells corresponding to non-zero entries in $\Delta_{ij}^{++} M$. 100

5.2 A diagram of a 6-layer dielectric anti-reflection structure whose outer regions have permittivity of $\epsilon_a = 12.25$ and $\epsilon_b = 1.0$ (free space). The permittivity of each dielectric slab is given by $\epsilon_1 = 13.5, \epsilon_2 = 8.41, \epsilon_3 = 4.86, \epsilon_4 = 2.52, \epsilon_5 = 1.46,$ and $\epsilon_6 = 1.08$. Their nominal widths satisfy $w_i = z_i - z_{i-1} = \pi c_0/(2\sqrt{\epsilon\omega_0})$ for $i = 1, 2, ..., 6$, where $\omega_0$ is the center frequency of the passband. 106

5.3 The second-order sensitivity of the $S_{11}$ parameter with respect to permittivities $\epsilon_2, \epsilon_3$ (a) and permittivities $\epsilon_3, \epsilon_6$ (b) at the nominal state of the AR filter. 107
5.4 The second-order sensitivity of the $S_{11}$ parameter with respect to positions $z_1$ and $z_6$ (a), positions $z_3$ and $z_4$ (b), the position $z_4$ (c), positions $z_4$ and $z_6$ (d) at the nominal state of the AR filter. .............. 108

5.5 The $H$-plane view of the rectangular waveguide loaded with a rectangular lossy discontinuity following the Drude dispersion profile. . . 109

5.6 The second-order sensitivity of the $S_{11}$ parameter with respect to edges $p_1$ and $p_2$ (a), edges $p_1$ and $p_3$ (b) at the nominal state of the lossy discontinuity. ................................. 110

5.7 The second-order sensitivity of the $S_{11}$ parameter with respect to edge $p_1$ (a), edges $p_2$ and $p_3$ (b), edges $p_2$ and $p_4$ (c), the edge $p_4$ (d) at the nominal state of the lossy discontinuity. ................. 111

5.8 The $H$-plane view of the of the dielectric resonator filter composed of three rectangular dielectric posts. ............................... 112

5.9 The second-order sensitivity of the $S_{11}$ parameter with respect to the permittivity $\varepsilon_b$ (a), permittivities $\varepsilon_a$ and $\varepsilon_b$ (b) at the nominal state of the dielectric resonator filter. ......................... 113

5.10 The second-order sensitivity of the $S_{11}$ parameter with respect to permittivities $\varepsilon_a$ and $\varepsilon_c$ (a), dimensions $w_a$ and $w_b$ (b), the dimension $w_b$ (c), dimensions $l_a$ and $l_c$ (d) at the nominal state of the dielectric resonator filter. ............................... 114
5.11 The dielectric resonator antenna (DRA) loaded with a thin layer of dielectric matching insert. ........................................ 115

5.12 The second-order sensitivity of the $S_{11}$ parameter with respect to dimensions $d_1$ and $h_2$ (a), dimensions $d_1$ and $w_2$ (b) at the nominal state of the dielectric resonator antenna. ....................... 116

5.13 The second-order sensitivity of the $S_{11}$ parameter with respect to dimensions $h_1$ and $w_2$ (a), dimensions $h_1$ and $d_2$ (b), dimensions $w_1$ and $d_2$ (c), dimensions $w_1$ and $h_2$ (d) at the nominal state of the dielectric resonator antenna. ............................... 117
Acknowledgments

At this very end of my graduate study, I feel so pleasurable thinking back at all those critical moments during my past school years and recalling all those people who have had shaped my personality through various ways.

I have had the fortune of having Mohamed Bakr as my thesis advisor. I am grateful for his trust, his patience, and his encouragement. Without many of his insightful ideas, the completion of this thesis would not be as smooth as it gets.

I also want to express my sincere gratitude to all members in my supervisory committee: Natalia Nikolova, Yaser Haddara, Chih-Hung Chen, Shiva Kumar, and Omar Ramahi (Waterloo). I appreciate them immensely for reviewing this thesis with great deliberation. Their constructive comments are very valuable for the enhancement of this work.

I have had a wonderful experience working at Blackberry Ltd. in the middle of my graduate study. This has provided me many valuable insights towards the design, prototyping, and testing of new generation mobile antennas. Therefore, I deeply appreciate Shirook Ali (Blackberry) for giving me this opportunity to grow.
Finally, I want to thank my colleagues Mohamed Negm, Osman Ahmed, Laleh Kalantari, Ahmed Elsharabasy, and Mahmoud Taha for their collaborations over various projects, as well as the friendship that I have experienced from them.
Chapter 1

Introduction

The adjoint variable method (AVM) is an efficient sensitivity approach for parameter-dependent differential-algebraic equations (DAE) [1]. By studying an adjoint system of the original DAE system, the response sensitivities with respect to all involved parameters can be evaluated simultaneously regardless of their number. Because of its mathematical simplicity and superior computational efficiency, the classical AVM technique has attracted great attention in many fields of engineering and science wherever design and optimization are encountered [2–4].

In a given design problem, the AVM technique provides an efficient and unified framework for estimating the first-order response sensitivities in the considered
dynamic system. The estimated adjoint sensitivities, in general, give designers information on how design variations will impact the overall system performance. In microwave engineering, the efficiency provided by the AVM technique has made it widely adopted in many gradient-based optimization techniques for designing microwave circuits [5–8]. In the area of computational electromagnetics, the AVM technique has been applied in the gradient-based optimization of high frequency electromagnetic structures [9–11].

Over the last decade, various forms of the AVM approach have been developed for various electromagnetic computational techniques, both in the time and frequency domains, see for example [12–26]. Using these adjoint-based sensitivity analysis techniques, the response sensitivity, required by gradient-based optimization techniques, can be computed with at most one extra simulation. If the same procedure is performed by a conventional finite-difference (FD) approximation, at least $N$ extra simulations would be required where $N$ is the number of designable parameters. Because of the superior efficiency, AVM techniques have been implemented in several commercial high-frequency simulation software packages including Ansoft HFSS since version 13, CST Microwave Studio since version 2010.06, and recently FEKO Suite.

1.1 Adjoint variable method for dispersive mediums

One limitation of the existing AVM approach [12–17, 19–21, 24–26] is that it can only handle frequency-independent parameters in the time-domain simulations. In elec-
electromagnetic problems, these parameters are usually associated with materials whose permittivity and electric conductivities are constant. This constant approximation is accurate for treating low-dispersion dielectrics within a narrow frequency band. However, most media in the real world exhibit strong frequency-dispersion and additional computational effort is required to model the associated dispersion parameters through time-domain simulations. Some well-developed computational techniques for treating dispersive media include the auxiliary differential equation (ADE) method [27–30], the $Z$-transform methods [31], and methods based on the discrete convolution of the dispersion relation [32]. These computational treatments for modeling dispersive media, in general, complicates the sensitivity analysis, and thus make the existing AVM approaches inapplicable.

To address this limitation, this thesis generalizes the current AVM approach to a wider range of materials and proposes an extended framework for treating frequency-dependent material parameters associated with linear dispersive media. The constitutive relation of our considered dispersive medium may contain the Lorentz dispersion pole, the Drude dispersion pole, the Debye dispersion pole, or a linear summation of them. Our proposed AVM generalization does not produce additional computational overhead as compared to all existing AVM treatments.

It is worth mentioning that, prior to our research, the AVM approach for dispersive medium has also been partially addressed in [25]. However, it was only limited to transmission-line matrix (TLM) based solvers [33]. Our proposed dispersive AVM approach extends [25] in three aspects. First, our AVM approach is applicable to
all standalone grid-based time-domain solvers, such as those based on the popular finite-difference time-domain (FDTD) method [34]. Second, our AVM approach does not require modifying the code of the simulation engine, as long as the time-domain field solutions computed at the desired points can be exported after the simulation. Third, our AVM approach provides unified treatment for materials containing multiple linear dispersion poles.

1.2 The second-order adjoint variable method

During the course of this research, we have discovered that, with some modification, the AVM technique could also be extended to supply second-order response sensitivities (often written in the form of a Hessian matrix) through the time-domain electromagnetic simulations. These second-order sensitivity information is particularly useful in a wide range of problems where knowing only the first-order response sensitivity is inadequate. For example, in many large-scale optimization problems when Newton-type methods are used, both the first- and the second-order sensitivity information are required [35]. The optimization techniques requiring Hessian matrices have, in general, a faster rate of convergence than those using only gradient information.

According to the literature, the second-order AVM approach has been investigated in mechanical system design [36, 37], fluid dynamics [38], and high-frequency
circuit applications [39–42]. Recently, this second-order approach is developed for
time-domain TLM based electromagnetic simulators [43].

This thesis extends the existing second-order AVM technique developed in [43],
which only applies to TLM solvers, to accommodate all other grid-based time-domain
solvers, including solvers based on the widely-used FDTD method. Our proposed
second-order AVM technique poses the same degree of efficiency as comparing to
the first-order AVM approach. With our proposed adjoint approach, the complete
$N^2$ second-order sensitivity entries in the response Hessian matrix can be estimated
through at most $N + 1$ extra simulations, where $N$ is the number of involved sensi-
tivity parameters. At the same time, all $N$ first-order sensitivities are also estimated
as a by-product. Our approach applies to problems involving both dispersive and
non-dispersive materials. With the obtained second-order sensitivity information, the
second-order Taylor expansion of the system response can be achieved in an accurate
and efficient way. This could greatly benefit a number of electromagnetic problems
involving optimization, device modeling, tolerance, yield analysis, and the solution of
inverse problems.

1.3 Outline of the thesis

This thesis covers the entire theory of the adjoint variable method developed for
time-domain electromagnetic simulations. We organize this theory into the following
five chapters.
In Chapter 2, we develop and introduce all the mathematical notations which are used throughout the thesis. All preliminary knowledge in understanding the adjoint variable method is covered. We also discuss the simplest form of the adjoint variable method in the time-domain simulations, in which all considered sensitivity parameter are associated with perfect dielectrics.

In Chapter 3, we consider the adjoint treatment for the most important frequency-dependent response function in computational electromagnetics—the scattering parameter. We also cover the sensitivity treatment for problems which are self-adjoint, where in this case, all adjoint simulations are not required.

In Chapter 4, we generalize the existing adjoint variable method to handle linear dispersive media utilized in the time-domain electromagnetic simulations. We provide a unified adjoint treatment for Drude dispersion, Lorentz dispersion, Debye dispersion, and dispersive relations composed of mixed linear dispersion poles. Our proposed technique does not yield additional computational or memory overhead.

In Chapter 5, we extend the existing adjoint variable method to estimate all second-order sensitivities in the response Hessian matrix through time-domain electromagnetic simulations. With this second-order adjoint approach, the number of required simulations for computing the response Hessian is reduced from $O(N^2)$ to $O(N)$, when compared with the conventional finite-difference techniques.

In Chapter 6, we draw conclusions for this research, and provide suggestions for the future development.
1.4 Summary of contributions

The theory covered in this thesis generalizes the existing adjoint variable method developed for time-domain electromagnetic simulations from the following two aspects:

(1) extends the adjoint variable method to handle sensitivity parameters associated with linear dispersive materials;

(2) extends the adjoint variable method to compute all second-order sensitivity entries in the response Hessian matrix.

Our collective research on the adjoint variable method has resulted in the following three journal publications.


The following articles are presented in refereed conferences.


References


Chapter 2

The Adjoint Variable Method for Frequency-Independent Constitutive Parameters

This chapter investigates the simplest form of the adjoint variable method for sensitivity estimation in time-domain electromagnetic simulations. In this scenario, all sensitivity parameters are associated with homogeneous dielectric mediums, whose constitutive parameters are independent of frequency.

Before stepping into the main topic, this chapter begins with our compact formulation of the wave equation describing the field computation in a discretized computation volume. It then covers all necessary preliminary concepts required for understanding the sensitivity analysis through on-grid perturbation. Several topics are discussed here explaining how the nominal wave equation is perturbed in a meshed
computation domain as a result of perturbing an arbitrary sensitivity parameter. The chapter also introduces the one-to-one field mapping procedure for approximating the perturbed field solution from the nominal wave equation. These topics, altogether, are crucial to the adjoint variable method developed in this thesis.

The adjoint-based approach covered in this chapter was initially proposed in [1] and [2], wherein, only the sensitivity parameters associated with metallic discontinuities are considered. Shortly after this early development, the scheme of the centralized adjoint variable method was then proposed in [3] and [4]. This new approach enhances the accuracy of the approximated adjoint sensitivities by one order without extra computational effort. In the research reported in [5] and [6], the adjoint framework is extended to handling homogeneous dielectric medium in time-domain simulations. In this chapter, we summarizes the above mentioned sensitivity techniques together and reformulate the previously published theories in a more unified and understandable way using our compact discrete-space formulation of the wave equation. This new formulation of the adjoint variable method allows mathematical extension to handle discontinuities with more generalized material properties as will be shown in the following chapters.

2.1 The wave equation in the discretized space

We begin our discussion from Maxwell’s curl equations for non-magnetized media

\[ \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}, \quad (2.1) \]
Differentiating both sides of (2.1) with respect to time and then substituting the time-derivative of \( \mathbf{H} \) in the resultant equation using the left-hand side of (2.2), we obtain the following second-order wave equation expressed in terms of the electric fields \( \mathbf{E} \)

\[
\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t}.
\]

(2.2)

The electric flux intensity \( \mathbf{D} \) relates to the electric field \( \mathbf{E} \) through the constitutive relation. For a simple dielectric medium, we have the following linear relation

\[
\mathbf{D} = \bar{\varepsilon} \mathbf{E},
\]

(2.4)

where \( \bar{\varepsilon} \) is known as the permittivity tensor.

To numerically solve Maxwell’s equations with limited computing resources, we often focus our consideration within a bounded volume in space with prescribed boundary conditions. This bounded volume is known as the computational volume and it is denoted here by \( \Theta \). In order to compute its enclosed electromagnetic fields, we partition \( \Theta \) into \( M \) cells using a set of orthogonal mesh lines. Through this approach, we can sequentially compute each localized field solution by considering one cell at a time. To minimize the numerical dispersion and improve the computational accuracy, the largest cell size must be less than one tenth of the smallest wavelength. At the \( j \)th labeled cell, we denote the localized time-dependent solutions of the electric field components as \( E_x^{(j)} \), \( E_y^{(j)} \), and \( E_z^{(j)} \). Further, we define \( \mathbf{E} \), given below, as the vector containing the time-dependent solution of the electric field components computed at
all cells within the entire computational volume $\Theta$

\[
\mathbf{E} = \begin{bmatrix} \mathbf{E}^{(1)^T} & \mathbf{E}^{(2)^T} & \ldots & \mathbf{E}^{(M)^T} \end{bmatrix}^T, \tag{2.5}
\]

where

\[
\mathbf{E}^{(j)} = \begin{bmatrix} E_x^{(j)} & E_z^{(j)} & E_z^{(j)} \end{bmatrix}^T, \quad j \in \{1, 2, \ldots, M\}. \tag{2.6}
\]

It should also be noted that all solution entries in $\mathbf{E}$ are continuous functions of time which are integrable and differentiable with respect to time.

Once the vector of time-dependent field solutions $\mathbf{E}$ is defined, we can then rewrite (2.3) in a vector-matrix form as follows\footnote{We denote matrices using uppercase alphabets with slanted bold font face, for example, $\mathbf{K}$ and $\mathbf{M}$. Vectors whose entries are time-dependent functions are denoted using uppercase alphabets with regular bold font face, for example, $\mathbf{E}$ and $\mathbf{J}$. All other vectors whose entries are scalars are denoted using lowercase alphabets with slanted bold font face, for example, $\mathbf{p}$.}

\[
\mathbf{K} \mathbf{E} + \mathbf{M} \mathbf{E} = \frac{d\mathbf{J}}{dt}. \tag{2.7}
\]

The first matrix-vector multiplication term $\mathbf{K} \mathbf{E}$ in the left-hand side of (2.7) contains all time-independent operations of the wave equations. The scalar matrix $\mathbf{K} \in \mathbb{R}^{3M \times 3M}$, the discretized form of the double-curl operator $\nabla \times \mu_0^{-1} \nabla \times (\cdot)$, is a symmetric matrix and can be derived using the spacial finite-difference discretization. All entries in the matrix $\mathbf{K}$ depend on the topology of the grid and the imposed boundary conditions on $\Theta$. For problems considered in this thesis, all entries in this matrix are assumed independent of perturbations in the considered sensitivity parameters. At the $j$th cell, the $x$-polarized expression of the matrix-vector product $\mathbf{K} \mathbf{E}$ is given by
\[(\mathbf{K} \mathbf{E})^{(j)}_x = \frac{1}{\mu_0} \left( \frac{D_{yy} E^{(j)}_x}{\Delta y^2} + \frac{D_{zz} E^{(j)}_x}{\Delta z^2} - \frac{D_{yx} E^{(j)}_x}{\Delta y \Delta x} - \frac{D_{zx} E^{(j)}_x}{\Delta z \Delta x} \right), \tag{2.8} \]

where \(\Delta x, \Delta y,\) and \(\Delta z\) are the cell sizes of the \(j\)th cell in the \(x, y,\) and \(z\) directions, respectively. It should be noted that equation (2.7) fully supports a graded mesh, and the values of \(\Delta x, \Delta y,\) and \(\Delta z\) may vary from cell to cell. For a better readability, we drop the superscripts on \(\Delta (u^o_x), \Delta (u^o_y),\) and \(\Delta (u^o_z),\) and use \(\Delta x, \Delta y,\) and \(\Delta z\) to denote the cell sizes of each considered cell.

If the \(j\)th cell is centered at \((x_0, y_0, z_0)\) in space and \(E^{(x_0, y_0, z_0)}_x\) is an alternative notation for \(E^{(j)}_x\), then the second-order spatial operators in (2.8) are defined by, for example,

\[D_{yy} E^{(x_0, y_0, z_0)}_x = E^{(x_0, y_0+\Delta y, z_0)}_x + E^{(x_0, y_0-\Delta y, z_0)}_x - 2E^{(x_0, y_0, z_0)}_x \tag{2.9}\]

and

\[D_{yx} E^{(x_0, y_0, z_0)}_x = \frac{1}{4} E^{(x_0+\Delta x, y_0+\Delta y, z_0)}_x + \frac{1}{4} E^{(x_0-\Delta x, y_0-\Delta y, z_0)}_x - \frac{1}{4} E^{(x_0-\Delta x, y_0+\Delta y, z_0)}_x - \frac{1}{4} E^{(x_0+\Delta x, y_0-\Delta y, z_0)}_x, \tag{2.10}\]

where the field expressions \(E^{(x_0, y_0 \pm \Delta y, z_0)}_x\) and \(E^{(x_0 \pm \Delta x, y_0 \pm \Delta y, z_0)}_x\) in (2.9) and (2.10) denote the respective solution entries in \(\mathbf{E}\) computed at cells neighbouring the \(j\)th cell in the respective directions. Similar expressions can be derived for all other difference operators in (2.8). Using the cyclic substitution \(x \rightarrow y \rightarrow z \rightarrow x,\) one can obtain similar expressions for \((\mathbf{K} \mathbf{E})^{(j)}_y\) and \((\mathbf{K} \mathbf{E})^{(j)}_z\) from (2.8). It should be clear from the previous discussion that all entries in the matrix \(\mathbf{K}\) are scalar constants.

Unlike the first matrix-vector multiplication term \(\mathbf{K} \mathbf{E},\) the second matrix-vector
multiplication term $M E$ in the left-hand side of (2.7) describes all time-dependent operations of the wave equations. Here, the operator $M$, defines a mapping between the electric field vector $E$ to another vector containing second-order time-derivatives of respective electric flux density $D$ over the same discretized volume. For a linear medium, where different waves at any particular point in the medium can be superposed, the correspondent operator $M$ must also be linear and bounded. In this case, we commonly treat the operation of $M$ on $E$, as if $E$ is pre-multiplied by a square matrix also denoted by $M$. The resultant vector $\frac{\partial^2 D}{\partial t^2} = M E$ contains the expressions of the second-order time-derivative of the electric flux density $D$ computed at each cell for all polarization directions. The entries of the matrix $M$, which can be derived from the constitutive relation for the electric field, are time-dependent operators describing the wave behavior at each localized cell.

For an isotropic computational volume, the matrix $M$ is a diagonal matrix whose diagonal entries are denoted by

$$M = \begin{bmatrix}
M^{(1)} & 0 & \ldots & 0 \\
0 & M^{(2)} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & M^{(M)}
\end{bmatrix}$$

If the $j$th cell is filled with a perfect dielectric material with constant permittivity $\epsilon^{(j)}$, then the expression of the respective diagonal entries in $M$ is given by

$$M^{(j)}_{x/y/z} \cdot f = \epsilon^{(j)} \frac{d^2 f}{dt^2}, \quad (2.11)$$
where the symbol $f$ denotes the operand of $M^{(j)}_{x/y/z}$ resulting from matrix-vector multiplication in $ME$.

The vector term $d\mathbf{J}/dt$ in the right-hand side of (2.7) is the excitation vector of the wave equation. It contains all first-order time derivatives of the current density sources $\mathbf{J}$ imposed on each cell at each polarization direction.

After defining the system matrices $K$ and $M$, the original wave equation (2.3) is re-expressed in a compact vector-matrix form (2.7) in terms of the electric field solution $\mathbf{E}$. Our discrete formulation of the wave equation presumes that each solution entry in $\mathbf{E}$ is computed at the center of each localized cell for all polarization directions. This is the default discretization scheme in transmission-line matrix (TLM) based solvers, if the symmetrical condensed node (SCN) is utilized [7]. However, for Yee-based finite-difference time-domain (FDTD) solvers [8], all three electric field components are computed at three different edges of their localized cell, rather than at the center of the same cell. In this case, an additional field-averaging process, known as the cell interpolation, must be applied at each cell to co-locate these offset field values to the geometric center of each considered cell [9]. This cell-averaging procedure is illustrated in Figure 2.1.

2.2 Sensitivity parameters and the perturbed wave equation

For a mathematical system with a set of designable parameters $p$, sensitivity analysis provides a quantitative study of how likely a given system response function $F(p)$
Figure 2.1 Illustration of cell interpolation over a Yee cell in Cartesian system, where all electric fields are computed at the center of each cell edge. With cell-interpolation, the field components at the geometric center of the cell (marked by $\bullet$) are approximated using $E_\kappa = \frac{1}{4}(E_{\kappa,1} + E_{\kappa,2} + E_{\kappa,3} + E_{\kappa,4})$ for $\kappa \in \{x,y,z\}$. 

Ph.D. Thesis — Yu Zhang
McMaster University (2015)
would change if any parameter contained in $p$ is changed by a small amount. We refer to the parameters contained in $p$ as sensitivity parameters. In case of time-domain electromagnetic simulations, these sensitivity parameters are usually associated with the dimensions or material properties of a given structure. If the numerical technique for computing the field response $F$ is governed by (2.7), any perturbation in $p$ would result in a perturbation in the matrix $M$ by a differential amount $\Delta M$. This perturbation would, in turn, affect the field solution through the following perturbed wave equation

$$K\tilde{E} + (M + \Delta M)\tilde{E} = \frac{dJ}{dt},$$

(2.13)

where $\tilde{E}$ denotes the field solution of this perturbed wave equation.

Consider a simple example where we have a thin homogeneous dielectric slab placed within an air-filled parallel plate waveguide as shown in Figure 2.2. Our interested sensitivity parameters are the thickness $D$ and the permittivity $\epsilon$ of the thin slab, $p = [D \; \epsilon]^T$. For this simple example, we consider four types of perturbations, as illustrated in Figure 2.3.

The first type of perturbation, shown in Figure 2.3(a), extends the thickness of the slab $D$ (the sensitivity parameter $p_1$), by an increment $\Delta z$ along the outward direction. We refer to the extended volume, indicated in Figure 2.3(a), as the volume of forward perturbation $\Psi^+_p$. As a result of this perturbation, the permittivity of the cells within $\Psi^+_p$ changes from its nominal value $\epsilon_0$ to $\epsilon$. Consequently, in the corresponding perturbed wave equation (2.13), the expression of the non-zero entries
Dielectric Slab, $\epsilon$

**Figure 2.2** E-plane view of a parallel plate waveguide loaded with a dielectric slab. The sensitivity parameters of interest are the thickness and the permittivity of the loaded slab $p = [D \; \epsilon]^T$. For illustration purpose, the computational volume, as shown in figure, is meshed in a way that the cell sizes is uniform along each respective direction.

contained in the difference matrix $\Delta M$ follows

$$
\Delta M_{x/y/z}^{(j)} \cdot f = (\epsilon - \epsilon_0) \frac{d^2 f}{dt^2}, \quad j \in \Psi^+_{p_1}.
$$

These non-zeros entries of $\Delta M$ are only local to field entries computed at cells within $\Psi^+_{p_1}$.

Conversely, for the second type of perturbation, as illustrated on Figure 2.3(b), the thickness of the dielectric slab is squeezed by the same amount $\Delta z$. As a result of this backward perturbation in $p_1$, the permittivity of the cells within the perturbed volume would decrease from its nominal value $\epsilon$ to $\epsilon_0$. We refer to this volume, indicated in Figure 2.3(b), as the volume of backward perturbation $\Psi^-_{p_1}$. In this case,
Figure 2.3  Illustration of step-wised on-grid perturbations of the dielectric slab in a parallel-plate waveguide. The perturbation surfaces are shown by dashed line. Cells within volume of perturbations are overlayed with colored circles ●.
the non-zero entries contained in $\Delta M$ would take the following expression

$$\Delta M_{x/y/z}^{(j)} \cdot f = (\epsilon_0 - \epsilon) \frac{d^2 f}{dt^2} \quad j \in \Psi^{-}_{p_1}.$$  

(2.15)

Similar to the forward case, these non-zeros entries in $\Delta M$ apply only to field entries computed at cells within $\Psi^{-}_{p_1}$.

The third and fourth types of perturbation, illustrated in Figure 2.3(c), both modify the permittivity of the dielectric slab $\epsilon$ (the sensitivity parameter $p_2$). As $p_2$ is perturbed, the value of the permittivity on cells within the slab volume changes from $\epsilon$ to $\epsilon \pm \Delta \epsilon$, depending on the direction of perturbations. In these two cases, the volumes of both the forward and the backward perturbation, $\Psi^{-}_{p_2}$ and $\Psi^{+}_{p_2}$, refer to the same volume occupied by the dielectric slab. As the consequence of these material perturbations, the non-zeros entries in $\Delta M$ take the following expression

$$\Delta M_{x/y/z}^{(j)} \cdot f = \Delta \epsilon \frac{d^2 f}{dt^2} \quad j \in \Psi^{\pm}_{p_2}.$$  

(2.16)

These non-zeros entries in $\Delta M$ correspond only to cells within $\Psi^{\pm}_{p_2}$ —the volume of the dielectric slab.

The above example illustrates how perturbing a geometry or a material parameter affect the system matrices in the wave equation. In general, perturbing a sensitivity parameter related to the size of a discontinuity along its forward direction modifies entries in $M$ related to cells just outside its perturbation surface. Perturbing the same shape parameter along its backward direction modifies entries in $M$ related to cells just inside its perturbation surface. If a material parameter is perturbed, then all entries in $M$ related to cells within the discontinuity are modified.
Next, we focus our attention on the perturbed field solution \( \tilde{\mathbf{E}} \) within the volume of perturbations \( \psi_{\pm}^{p_i} \) for each perturbation scenario illustrated in Figure 2.3. From perturbation theory, it is well-known that the field distribution in a perturbed problem does not differ significantly from that in the original unperturbed problem, if the perturbation is small, and thus it is possible to obtain a low-order approximation of the perturbed field solution \( \tilde{\mathbf{E}} \) in terms of the unperturbed field solution \( \mathbf{E} \) computed from (2.7). In the case of on-grid perturbations as illustrated in Figure 2.3, a reasonable approximation of the perturbed field solution \( \tilde{\mathbf{E}} \) within the volume of perturbations \( \psi_{\pm}^{p_i} \) can be obtained through a simple procedure called the one-to-one field mapping [1]. Here we illustrate this approximation procedure in Figure 2.4 for the same dielectric slab example shown in Figure 2.2.

Consider the first perturbation scenario shown in Figure 2.3(a), where the dielectric slab is extended along the \( z \)-axis by an incremental volume \( \Psi_{p_1}^+ \). If the cell-size \( \Delta z \) is made much smaller than the nominal thickness of the dielectric slab \( \Delta z \ll D \), we can approximate the solution entries in \( \tilde{\mathbf{E}} \) computed at the cells inside \( \Psi_{p_1}^+ \) (see cells marked by circles in Figure 2.3(a)) using the respective entries in \( \mathbf{E} \) computed at the cells adjacent to the same perturbation surface inside the dielectric slab before the shape is perturbed, and these cells for the unperturbed structure are labeled by squares in Figure 2.4(a). We also notice that these cells labeled with squares in Figure 2.4(a) are within the volume \( \Psi_{p_1}^- \). In this case, we say that the solution entries in \( \tilde{\mathbf{E}} \) computed at cells within \( \Psi_{p_1}^+ \) is mapped to the solution entries of \( \mathbf{E} \) computed
Figure 2.4 Illustration of the one-to-one field mapping. In all cases shown in Figure 2.3, the entries of $\tilde{E}$ computed at the cells marked by circles $\circ$ in Figure 2.3 are approximated using the entries of $E$ computed at the cells marked by squares $\blacksquare$. 

(a) Perturbed fields in $\Psi^+_{p_1}$ is approximated using nominal fields in $\Psi^-_{p_1}$;

(b) Perturbed fields in $\Psi^-_{p_1}$ is approximated using nominal fields in $\Psi^+_{p_1}$;

(c) Perturbed fields in $\Psi^\pm_{p_2}$ is approximated using nominal fields in $\Psi^\mp_{p_2}$. 
at the cells within \( \Psi_{p_1}^- \).

Consider the opposite scenario as shown in Figure 2.3(b), where the dielectric slab is squeezed by another incremental volume \( \Psi_{p_1}^- \). A similar approximation procedure can be made through the mapping between the solution entries of \( \tilde{\mathbf{E}} \) computed at cells adjacent to the perturbation surface outside the perturbed dielectric slab (see cells labelled with circles in Figure 2.3(b)) and the respective entries in \( \mathbf{E} \) computed at the cells adjacent to the same perturbation surface outside the dielectric slab before the shape is perturbed (see cells labelled with squares in Figure 2.4(b)). In this case, the field entries of the perturbed solution \( \tilde{\mathbf{E}} \) computed at the cells within \( \Psi_{p_1}^- \) are mapped to the respective field entries in the nominal solution \( \mathbf{E} \) computed at the cells within \( \Psi_{p_1}^+ \).

In the last two perturbation scenarios illustrated in Figure 2.3(c), the resultant perturbation volumes for both the forward and the backward perturbations, \( \Psi_{p_2}^+ \) and \( \Psi_{p_2}^- \), refer to the same volume which encloses all cells within the dielectric slab. In this case, the mapping procedure developed for shape perturbations is still valid, provided that the amount of perturbation in material parameters (\( \Delta \varepsilon \), for example) is much smaller then the nominal value of the considered material parameter. In this case, the field entries of the perturbed solution \( \tilde{\mathbf{E}} \) computed at cells with \( \Psi_{p_2}^{\pm} \) is mapped to the respective field entries in the nominal solution \( \mathbf{E} \) computed at cells within \( \Psi_{p_2}^{\mp} \). This mapping procedure is illustrated in Figure 2.4(c).

Based on our experience, the one-to-one field mapping supplies a good approximation for the solution of the perturbed wave equations. The degree of the approxima-
tion accuracy is heavily affected by the grid step size, and the degree of approximation accuracy improves as the mesh is made finer. Later in this chapter, we shall apply this mapping procedure to approximate perturbed field solutions required for computing the adjoint sensitivities.

2.3 The response function and its sensitivity gradient

The type of response function \( F \), whose sensitivities we are interested in, is a scalar valued function of the form: [10–12]

\[
F(\mathbf{E}, p) = \int_0^{T_{\text{max}}} \phi(\mathbf{E}) \, dt, \tag{2.17}
\]

where the integral kernel \( \phi \) is a scalar valued functional of the field solution \( \mathbf{E} \). Its expression may take various forms depending on the type of the response that we are considering. If we are interested in the total \( y \)-polarized electric energy transmitted through a bounded port \( \Omega \), say perpendicular to the \( z \)-axis, then the choice of the response function \( F \) would be the following expression [10]

\[
F = \frac{1}{\Delta z_p} \int_0^{T_{\text{max}}} \sum_{j \in \Omega} \frac{1}{2} v^{(j)} E_y^{(j)}^2 \, dt, \tag{2.18}
\]

where \( v^{(j)} \) denotes the volume of the \( j \)th cell and \( \Delta z_p \) denotes the longitudinal cell size of the port plane (or the thickness of the port plane). With the above response function \( F \), we can write the following expression for the integral kernel \( \phi \)

\[
\phi(\mathbf{E}) = \frac{1}{2\Delta z_p} \sum_{j \in \Omega} v^{(j)} E_y^{(j)}^2, \tag{2.19}
\]
and \( \phi \) is not dependent on the field solutions computed within \( \Omega \).

With other prescribed choices of \( \phi \), the response function \( F \) may represent the scattering parameter at a given frequency over a port plane, or the value of a field component at a given instantaneous time. We shall address these topics in Chapter 3.

For a given response function \( F \), we refer to the gradient of \( F \) with respect to the vector of sensitivity parameters \( p \), given below, as the response sensitivity of \( F \):

\[
\frac{\partial F}{\partial p} = \begin{bmatrix} \frac{\partial F}{\partial p_1} & \frac{\partial F}{\partial p_2} & \cdots & \frac{\partial F}{\partial p_N} \end{bmatrix}^T.
\] (2.20)

Applying the chain rule on (2.17), the derivative of \( F \) with respect to the \( i \)th sensitivity parameter \( p_i \) can be expressed as [2]

\[
\frac{\partial F}{\partial p_i} = \frac{\partial^e F}{\partial p_i} + \int_0^{T_{\text{max}}} \left( \frac{\partial \phi}{\partial \mathbf{E}} \right)_i \frac{\partial \mathbf{E}}{\partial p_i} \, dt,
\] (2.21)

where the first derivative term in (2.21), denoted by \( \partial^e F/\partial p_i \), represents the explicit dependence of \( F \) on the \( i \)th sensitivity parameter \( p_i \). This term, for most practical choices of \( F \), is zero, and thus, only the integral term in the right-hand side of (2.21) contributes to the sensitivity expression.

In the second term of the sensitivity expression (2.21), the row vector \((\partial \phi/\partial \mathbf{E})^T\) in the integrand, as the result of the partial differentiation, takes the following expression

\[
\left( \frac{\partial \phi}{\partial \mathbf{E}} \right)^T = \begin{bmatrix} \frac{\partial \phi}{\partial E_x^{(1)}} & \frac{\partial \phi}{\partial E_y^{(1)}} & \frac{\partial \phi}{\partial E_z^{(1)}} & \cdots & \frac{\partial \phi}{\partial E_z^{(M)}} \end{bmatrix}.
\] (2.22)

If \( \phi \) is a higher order function of \( \mathbf{E} \), for example (2.19), then the field solution \( \mathbf{E} \) must be known to compute the vector \((\partial \phi/\partial \mathbf{E})^T\).

The second differentiation term \( \partial \mathbf{E}/\partial p_i \) in the integrand of (2.21) results in a
column vector with the following expression

\[ \frac{\partial \mathbf{E}}{\partial p_i} = \begin{bmatrix} \frac{\partial E_x^{(1)}}{\partial p_i} & \frac{\partial E_y^{(1)}}{\partial p_i} & \frac{\partial E_z^{(1)}}{\partial p_i} & \cdots & \frac{\partial E_z^{(M)}}{\partial p_i} \end{bmatrix}^T. \]  

(2.23)

To evaluate this vector term \( \partial \mathbf{E}/\partial p_i \), the analytical expression of \( \mathbf{E} \) as a function of \( p_i \) must be known. Unfortunately, this information is usually unavailable for practical electromagnetic problems.

In time-domain electromagnetic simulations, the conventional approach for computing the response sensitivities (2.21) is through finite-difference (FD) approximation formulas. If the problem requires second-order accuracy in the sensitivity approximation, the following central finite-difference (CFD) formula can be employed:

\[ \frac{\partial F}{\partial p_i} \approx \frac{F(p_i + \Delta p_i) - F(p_i - \Delta p_i)}{2\Delta p_i}, \]  

(2.24)

where \( \Delta p_i \) denotes the amount of imposed perturbation on \( p_i \).

The CFD formula (2.24) provides a simple and direct approach for sensitivity estimation, which is readily applicable to all types of numerical computational techniques. However, the CFD formula (2.24) requires that the function value of \( F \) is known beforehand at two different perturbed states of \( p_i \). The amount of computation required for computing \( F(p_i + \Delta p_i) \) and \( F(p_i - \Delta p_i) \) in (2.24) involves solving the respective perturbed wave equations through two additional time-domain simulations. If a problem considers \( N \) sensitivity parameters \( (p \in \mathbb{R}^N) \), then with the CFD approach, the total number of required simulations for computing all \( N \) first-order sensitivities (2.20) would be \( 2N \). For a large scaled problem with numerous sensitivity parameters, this approach is computationally prohibitive. In this scenario, the
adjoint-based sensitivity technique must be employed to enhance the computational efficiency.

2.4 The adjoint sensitivity expression

Adjoint-based sensitivity analysis, in general, involves solving a direct problem and its correspondent adjoint problems. The direct problem is governed by the wave equation in the discrete domain (2.7). Any modification in the sensitivity parameters contained in $p$ would modify the system matrices in (2.7), and thus results in a perturbation in the field solution $E$.

To simplify the derivation, we consider one sensitivity parameter at a time, say $p_i \in p$, and perturb its nominal value by an amount $\Delta p_i$ along its forward direction. We thus obtain the following perturbed wave equation

$$ K(E + \Delta E) + (M + \Delta M)(E + \Delta E) = \frac{dJ}{dt}. $$  

(2.25)

Here, the equation (2.25) is mathematically the same as (2.13) with $\Delta E = \tilde{E} - E$.

Subtracting (2.7) from (2.25) and rearranging the resultant equation, we have

$$ K\Delta E + \tilde{M}\Delta E = -R, $$  

(2.26)

where this new matrix $\tilde{M} = M + \Delta M$ denotes the perturbed system matrix of $M$.

The residual vector $R$, (2.26), with the expression given below, depends only on the nominal field solution $E$

$$ R = \Delta ME. $$  

(2.27)
In the above expression of the residual vector (2.27), the matrix perturbation $\Delta M$, defined in (2.14), is a diagonal matrix if the computational domain is isotropic. Its non-zero entries correspond to solution entries of $\mathbf{E}$ computed within $\Psi_{p_i}^+$. 

Multiplying both sides of (2.26) by a yet-to-be determined variable $\hat{E}$—another time-dependent vector, and integrating both sides of the resultant equation with respect to the complete simulation time, equation (2.26) becomes

$$
\int_0^{T_{\text{max}}} \hat{E}^T (K \Delta \mathbf{E}) \, dt + \int_0^{T_{\text{max}}} \hat{E}^T (\tilde{M} \Delta \mathbf{E}) \, dt = -\int_0^{T_{\text{max}}} \hat{E}^T \mathbf{R} \, dt.
$$

(2.28)

For the first term in the left-hand side of (2.28), the integrand $\hat{E}^T (K \Delta \mathbf{E})$ is mathematically equivalent to $(K^T \hat{\mathbf{E}})^T \Delta \mathbf{E}$ through

$$
\int_0^{T_{\text{max}}} \hat{E}^T (K \Delta \mathbf{E}) \, dt = \int_0^{T_{\text{max}}} (K^T \hat{\mathbf{E}})^T \Delta \mathbf{E} \, dt.
$$

(2.29)

For the second term in the left-hand side of (2.28) where the matrix $\tilde{M}$ contains time-dependent operations on $\Delta \mathbf{E}$, our treatment is to define a new matrix of operators denoted as $\hat{\tilde{M}}$ such that

$$
\int_0^{T_{\text{max}}} \hat{E}^T (\tilde{M} \Delta \mathbf{E}) \, dt = \int_0^{T_{\text{max}}} (\hat{\tilde{M}}^T \hat{\mathbf{E}})^T \Delta \mathbf{E} \, dt.
$$

(2.30)

The expression of $\hat{\tilde{M}}$ will be discussed in the next section.

Substituting (2.29) and (2.30) back into (2.28), it follows that

$$
\int_0^{T_{\text{max}}} (K^T \hat{\mathbf{E}} + \hat{\tilde{M}}^T \hat{\mathbf{E}})^T \Delta \mathbf{E} \, dt = -\int_0^{T_{\text{max}}} \hat{E}^T \mathbf{R} \, dt.
$$

(2.31)

Now, we can uniquely specify the expression of the vector $\hat{\mathbf{E}}$ by imposing the following equality condition.
\[ K^T \hat{E} + \hat{M}^T \hat{E} = \frac{\partial \phi}{\partial E}, \]  

(2.32)

where \( \phi \), as defined in (2.17), is a functional of the nominal field solution \( E \).

Finally, by substituting (2.32) back into (2.31) and comparing the resultant equation with the analytical expression of the response sensitivity given in (2.21), we can write the following adjoint expression of the response sensitivity with respect to the sensitivity parameter \( p \) [2]

\[
\frac{\partial F}{\partial p_i} \approx \frac{\partial^e F}{\partial p_i} - 1 \Delta p_i \int_0^{T_{max}} \hat{E}^T R \, dt.
\]

(2.33)

The above adjoint sensitivity expression only requires entries of \( R \) and \( \hat{E} \) computed at cells within \( \Psi_p^+ \). This is because entries of \( R \) are zero outside \( \Psi_p^+ \).

To compute the residual vector \( R \), we first solve the field entries of \( E \) computed at cells within \( \Psi_p^+ \), and then solve for non-zeros entries of \( R \) by using (2.27). To compute the adjoint vector \( \hat{E} \), we must solve the following wave equation

\[
K^T \hat{E} + \hat{M}^T \hat{E} = \frac{d \hat{J}}{dt},
\]

(2.34)

where the vector \( \hat{J} \), derived from (2.32), takes the following expression

\[
\hat{J} = \int_0^{T_{max}} \frac{\partial \phi}{\partial E} \, dt.
\]

(2.35)

This new equation (2.34), whose solution is the adjoint vector \( \hat{E} \), takes the same form as the original wave equation(2.7). We thereby refer to (2.34) as the adjoint wave equation, and refer to the vector \( \hat{J} \) as the adjoint current source.

Recall that the adjoint vector \( \hat{E} \) and the residual vector \( R \) in (2.33) are obtained...
by perturbing $p_i$ by an amount $\Delta p_i$ along its forward direction. Here, we add a superscript + to these two vectors, as $\hat{\mathbf{E}}^+$ and $\mathbf{R}^+$, to indicate this perturbation along the forward direction. Following a similar mathematical procedure, one can obtain another set of $\hat{\mathbf{E}}$ and $\mathbf{R}$ by the same sensitivity parameter $p_i$, by perturbing it along the opposite direction for the same amount. We denote these new resultant vectors as $\hat{\mathbf{E}}^-$ and $\mathbf{R}^-$, where the superscript − indicates that these vectors are resulted from a backward perturbation. With these two sets of vectors, involving both forward and backward perturbations, we can write the following central adjoint expression for the response sensitivity with respect to the $i$th sensitivity parameter $p_i$ as [4]

$$
\frac{\partial F}{\partial p_i} \approx \frac{\partial \psi}{\partial p_i} = \frac{1}{\Delta p_i} \int_0^{T_{\max}} \frac{\hat{\mathbf{E}}^+ \mathbf{R}^+ + \hat{\mathbf{E}}^- \mathbf{R}^-}{2} \, dt. \tag{2.36}
$$

This central adjoint expression (2.36), in general, improves the accuracy of the adjoint sensitivity estimation by an order, however, it requires to solve two adjoint equations for two different perturbation directions of $p_i$.

### 2.5 The solution of the adjoint problem

The adjoint formulas, given in (2.33) and (2.36), provide accurate estimations of the response sensitivity with respect to a given sensitivity parameter $p_i \in \mathbf{p}$, using the field entries of $\hat{\mathbf{E}}$ computed within $\Psi_{p_i}$. However, there are still two issues that remain unaddressed through our sensitivity deviation.
First, the system matrix $\hat{M}$, defined in (2.30), is still unknown in the adjoint equation given in (2.34). New computational techniques must be developed to solve the adjoint vector $\hat{E}$ from the adjoint equation (2.34).

Second, the adjoint equation (2.34) is dependent on $p_i$ and its respective directions of perturbation. To compute the response sensitivities with respect to all $N$ sensitivity parameters contained in $p$, the forward adjoint formula (2.33) would require to solve $N$ additional adjoint equations for each sensitivity parameter. With the central adjoint formula (2.36), which considers both forward and backward perturbation directions, a total of $2N$ adjoint equations must be solved. Thus, our derived adjoint sensitivity formulas given in (2.33) or (2.36) would not offer any computational advantages as compared to the CFD approximation formula (2.24).

In this section, we only address these two issues by considering a simplified case, in which the computational volume contains only perfect dielectric mediums with constant permittivities. At the $j$th cell, we denote its associated nominal permittivity as $\epsilon^{(j)}$ and its new permittivity due to the perturbation on $p_i$ as $\tilde{\epsilon}^{(j)}$. In this scenario, the non-zero entries in the system matrix of the perturbed wave equation (2.25) $\hat{M}$ take the same form as those for $M$ given in (2.12) as

$$\hat{M}^{(j)}_{x/y/z} \cdot f = \tilde{\epsilon}^{(j)} \frac{d^2 f}{dt^2},$$  \hspace{1cm} (2.37)

and we have $\epsilon^{(j)} = \tilde{\epsilon}^{(j)}$ if the $j$th cell is not within volumes of perturbation.

Then we can rewrite the integrand in the left-hand side of (2.30), in a summation form of element-wised multiplications, as
\[
\int_0^{T_{\text{max}}} \mathbf{E}^T (\mathbf{M} \Delta \mathbf{E}) \, dt = \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \epsilon^{(j)} (d^2 \mathbf{E}^{(j)})^T \frac{d^2 \Delta \mathbf{E}^{(j)}}{dt^2} \, dt,
\]  
(2.38)

where \( \mathbf{E}^{(j)} = \left[ \hat{E}_x^{(j)}, \hat{E}_y^{(j)}, \hat{E}_z^{(j)} \right]^T \) and \( \Delta \mathbf{E}^{(j)} = \left[ \Delta E_x^{(j)} \Delta E_y^{(j)} \Delta E_z^{(j)} \right]^T \), denote the respective field entries in \( \hat{\mathbf{E}} \) or \( \Delta \mathbf{E} \) computed at the \( j \)th cell.

We then integrate the right-hand side of (2.38) by parts twice. It follows that

\[
\int_0^{T_{\text{max}}} \mathbf{E}^T (\mathbf{M} \Delta \mathbf{E}) \, dt = \sum_{j=1}^{M} \epsilon^{(j)} \left( \frac{d \mathbf{E}^{(j)}}{dt} \right)^T \Delta \mathbf{E}^{(j)} \bigg|_0^{T_{\text{max}}} - \sum_{j=1}^{M} \epsilon^{(j)} (d^2 \mathbf{E}^{(j)})^T \Delta \mathbf{E}^{(j)} \bigg|_0^{T_{\text{max}}} + \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \epsilon^{(j)} \left( \frac{d^2 \mathbf{E}^{(j)}}{dt^2} \right)^T \Delta \mathbf{E}^{(j)} \, dt.
\]  
(2.39)

The first two terms in the right-hand side of (2.39) vanish, if we simultaneously prescribe \( \hat{\mathbf{E}} \) and \( \Delta \mathbf{E} \) with the following four initial/terminal conditions

\[
\Delta \mathbf{E} = 0 \quad \frac{d \Delta \mathbf{E}}{dt} = 0 \quad \text{at } t = 0,
\]

\[
\hat{\mathbf{E}} = 0 \quad \frac{d \hat{\mathbf{E}}}{dt} = 0 \quad \text{at } t = T_{\text{max}}.
\]  
(2.40)

Without the first two terms, the right-hand side of (2.39) becomes

\[
\int_0^{T_{\text{max}}} (\hat{\mathbf{E}}^T \mathbf{V})(\mathbf{M} \Delta \mathbf{E}) \, dt = \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \epsilon^{(j)} \left( \frac{d^2 \mathbf{E}^{(j)}}{dt^2} \right)^T \Delta \mathbf{E}^{(j)} \, dt
\]

\[
= \int_0^{T_{\text{max}}} (\mathbf{M} \hat{\mathbf{E}})^T \Delta \mathbf{E} \, dt.
\]  
(2.41)

Comparing the right-hand sides of (2.41) and (2.30), we can conclude that if the considered computational volume is filled with only perfect dielectric mediums, we have \( \hat{\mathbf{M}} = \mathbf{M} \). We also notice that, in (2.34), the matrices \( \mathbf{K} \) and \( \mathbf{M} \) are both symmetrical, if the computational domain is isotropic. Then under these conditions,
equation (2.34) is equivalent to the following

\[ K\hat{E} + \tilde{M}\hat{E} = \frac{d\hat{J}}{dt}, \]  

This new adjoint equation (2.42) takes the same form as the perturbed wave equation given in (2.13) with \( \tilde{M} = M + \Delta M \), thus the same numerical techniques for solving the original wave equation (2.7), for example through an FDTD simulation, can be used here to solve (2.42).

Further, in our derivation of (2.42), we have prescribed \( \hat{E} \) with terminal conditions rather than with initial conditions, see (2.40). As a consequence, we can only obtain \( \hat{E} \) by solving (2.42) reversely with respect to time starting from \( t = T_{\text{max}} \). This time-reversal computation, however, does not require any modification in the computational technique used for the forward-time computation. This is because in the left-hand side of (2.42), the only involved time-dependent operation is the second-order time-differentiation, and this time-differentiation, contained in the matrix \( \tilde{M} \) given by (2.37), is invariant if the time is reversed. In the right-hand side of (2.42), differentiating \( \hat{J} \) backward in time is the same as differentiating \( -\hat{J} \) forward in time. Therefore, solving (2.42) backward in time is equivalent to solving the following alternative wave equation forward in time

\[ K\hat{E} + \tilde{M}\hat{E} = \frac{d\hat{J}}{dt}, \]  

provided that all entries in the imposed current source \( \hat{J} \) satisfies

\[ \hat{J}(t) = -\hat{J}(T_{\text{max}} - t). \]  

So far, our derived adjoint equation given in (2.42) is still dependent on the
perturbed system matrix $\tilde{M}$, and solving (2.42) produces an exact solution of the adjoint vector $\tilde{E}$. However, one may notice that the adjoint sensitivity formulas, given in (2.33) or (2.36), only make uses of the solution entries in $\tilde{E}$ which are computed within $\Psi_{u\pm}$. Then in this case, if only entries of $\tilde{E}$ within $\Psi_{u\pm}$ are concerned, we can further approximate (2.42) using its corresponding unperturbed system matrix $M$ as

$$K\tilde{E} + M\tilde{E} \approx \frac{d\tilde{J}}{dt}.$$  \hspace{1cm} (2.45)

Once (2.45) is solved backward in time, one can apply the one-to-one mapping procedure (illustrated in Figure 2.4) to obtain accurate approximations of $\tilde{E}$ for field entries computed at cells within $\Psi_{u\pm}$.

Now, we have approximated the adjoint equation in a way that is independent of the sensitivity parameter $p_i$, see (2.45). If a problem considers $N$ sensitivity parameters contained in $p$ simultaneously, we only need to solve (2.45) once, and store the field components contained within $\Psi_{u\pm}$ for each sensitivity parameters $p_i \in p$. Then with the procedure of one-to-one field mapping, one can obtain an accurate approximation of required entries in $\tilde{E}$ for estimating adjoint sensitivities with respect to each sensitivity parameter.

### 2.6 Algorithm implementation and numerical examples

The adjoint variable method introduced in this chapter, in general, requires two independent simulations for estimating the response sensitivities with respective to all sensitivity parameters contained in $p$, regardless of their number. It can be efficiently
implemented for any grid-based time-domain electromagnetic solvers without the need of modifying the simulation code, as long as the solver allows exciting user-customized source signal. The other requirement for the solver is that for each sensitivity parameter, say $p_i$, the time-domain field solutions at cells within the respective volumes of perturbations $\Psi_{p_i}^\pm$, can be exported. Here, we summarize this adjoint-based sensitivity procedure into the following four simple steps:

**Step 1. Parametrization**

For each sensitivity parameter, say $p_i$, we first identify its volumes of perturbations $\Psi_{p_i}^\pm$ it would produce due to on-grid perturbations. Then we group and label each cell within $\Psi_{p_i}^\pm$.

**Step 2. Direct analysis**

Perform the time-domain simulation over the nominal structure by imposing the current density source $\mathbf{J}$ specified by the considered problem. For each sensitivity parameter, say $p_i$, we export the time-domain field solution computed at cells within $\Psi_{p_i}^\pm$. Here, we refer the vector of solution entries computed at cells within $\Psi_{p_i}^+$ as $\mathbf{E}_{p_i}^+$, and the vector of the solution entries computed at cells within $\Psi_{p_i}^-$ as $\mathbf{E}_{p_i}^-$.

**Step 3. Adjoint analysis**

Re-simulate the domain by imposing another current source $\tilde{\mathbf{J}}$. Here, the entries of $\tilde{\mathbf{J}}$ can be computed through (2.35) and then (2.44). For each sensi-
tivity parameter, say $p_i$, we export the time-domain field solution computed at cells within $\Psi_{p_i}^\pm$ from the simulation. Here, we refer the vector of solution entries computed at cells within $\Psi_{p_i}^-$ as $\hat{\mathbf{E}}_{p_i}^-$, and the vector of solution entries at cells within $\Psi_{p_i}^+$ as $\hat{\mathbf{E}}_{p_i}^+$ (notice that the sign changes on superscripts as the result of one-to-one field mapping).

**Step 4. Post-processing**

For each sensitivity parameter, say $p_i$, we first compute the non-zero entries in the residual vectors $\mathbf{R}^\pm$ from $\mathbf{E}^\pm_{p_i}$ using (2.27). Here, we refer the vectors of these entries as $\mathbf{R}^\pm_{p_i}$. Then the adjoint response sensitivity can be computed through (2.36) by using $\mathbf{R}^\pm_{p_i}$ and $\hat{\mathbf{E}}^\pm_{p_i}$.

### 2.6.1 Example: the One-Dimensional Dielectric Slab

We consider the dielectric slab problem shown in Figure 2.2. Recall that our considered sensitivity parameters are the thickness $D$ and the permittivity $\epsilon$ of the dielectric slab loaded in the waveguide. Here, we further specify that the nominal permittivity of the dielectric slab is $\epsilon = 3.0\epsilon_0$. The nominal width of the slab $D$ may take any of the 20 values evenly spaced between $0.3\lambda_0$ to $\lambda_0$, where $\lambda_0$ is the wavelength for waves travelling inside the dielectric slab. Then, for a given slab thickness $D$ and a known excitation source, we are able to evaluate the energy function (2.18) using the field solution computed over the output port of the waveguide.
In our analysis, the electric field solution of this slab problem is solved numerically using the one-dimensional FDTD algorithm. Our considered computational volume has a length of 30 mm terminated by Mur’s absorbing boundary conditions. In our discretization setup, all mesh lines, shown by vertical grid lines in Figure 2.2, are smoothly distributed along the direction of wave propagation. The separation between adjacent mesh lines varies between 0.01 mm to 0.08 mm and the mesh lines are more condensed near the material interfaces than those within homogeneous regions. The cross section area of the computational volume is $5 \text{ mm}^2$. We choose the FDTD time step to be $\Delta t = c_0/\Delta h$, where $c_0$ denotes the speed of light in free space and $\Delta h$ denotes the smallest mesh size.

Our choice of the excitation source is a modulated Gaussian pulse centered at 10 GHz with a bandwidth of 8 GHz. Its peak amplitude is $1.0 \text{ A} \cdot \text{m}^2$. This excitation source is uniformly placed across the excitation plane 1.5 mm away from the left boundary and 4.5 mm away from the dielectric slab. On the other side of the dielectric slab, the plane of the output port, where the energy function is evaluated, is placed 2.5 mm away from the right boundary. Under these settings, we computed the field solution of this slab structure in discrete time for a total of 40000 time steps.

In our sensitivity analysis, our target is to compute the sensitivity of the energy function (2.18) for 20 different nominal values of the slab thickness. The produced sensitivity results are given in Figure 2.5. In the figure, the sensitivity results labeled with FAVM or BAVM, respectively, are computed using the adjoint formula (2.33) by perturbing the considered sensitivity parameters in either the forward or the backward
Figure 2.5  The adjoint sensitivities of the energy function (2.18) with respective
to the slab thickness $D$ (a) and the relative permittivity of the slab $\epsilon_r$ (b).

directions. The sensitivity result labeled with CAVM are computed using the central
adjoint formula (2.36) which simultaneous involves both perturbation directions. The
comparison shows that all adjoint sensitivity results accurately match those computed
using the CFD approximations.

2.6.2 The Anti-Reflection Dielectric Coating

In the previous example, where only two sensitivity parameters are considered, the
adjoint variable method computes the required response sensitivities through one ex-
tra adjoint simulation for each given nominal slab thickness. If the same sensitivity
Figure 2.6 The six-layer anti-reflection dielectric coating. The incident region is semiconductor with relative permittivity $\varepsilon_a = 12.25$, while the target region is free space with $\varepsilon_b = 1.0$. The relative permittivity of each dielectric slab is given by $\varepsilon_1 = 13.5$, $\varepsilon_2 = 8.41$, $\varepsilon_3 = 4.86$, $\varepsilon_4 = 2.52$, $\varepsilon_5 = 1.46$, and $\varepsilon_6 = 1.08$. The nominal width of each slab can be obtained using $w_i = z_i - z_{i-1} = \pi c_0 / 2\sqrt{\varepsilon_i \omega_0}$ for $i = 1, 2, ..., 6$, where $\omega_0$ is the center frequency of the pass-band.

procedure is perform using the CFD approximation, the amount of required computational effort would be increased by a factor of four. The advantage of the adjoint based technique can become more pronounced if the number of sensitivity parameters increases.

Let us look at the next anti-reflection coating example where the number of considered sensitivity parameters is increased to twelve. This multi-layered structure, as shown in Figure 2.6, is optimized for plane wave travelling normal to the plane of the material interface [13, 14]. It matches the wave impedance of a high permittivity region with $\varepsilon_a = 12.25$ (semiconductor) to the impedance of a region with low permittivity $\varepsilon_b = 1.0$ (free space) over a bandwidth of $1.7\omega_0$, where the $\omega_0$ is the central
angular frequency of the pass band. The nominal thickness of each dielectric layer equals quarter of the wavelength for waves at the center frequency $\omega_0$ in each layered medium.

We use the same one-dimensional FDTD code used in the previous example to compute the field solutions of this example. In our simulation setup, the total computational volume has a length of 50 mm truncated by Mur’s absorbing boundary conditions. The mesh sizes vary between 0.001 mm to 0.01 mm, and they are smoothly distributed along the direction of wave propagation. The cross section area of the computational volume is $5 \text{ mm}^2$. The FDTD time step is $\Delta t = c_0/\Delta h$, where $\Delta h$ denotes the smallest mesh size.

The excitation source is a modulated Gaussian pulse centered at 10 GHz with a bandwidth of 19 GHz and a peak amplitude of $1.0 \text{ A} \cdot \text{m}^2$. This excitation source is uniformly placed across the excitation plane 1.5 mm away from the left boundary and 4.5 mm away from the leftmost layer of this multilayer structure. At the low-permittivity region, the plane of the output port where the energy function is evaluated is place at 2.5 mm away from the right absorbing boundary. With these simulation settings, the energy function is evaluated over the output port for a total of 80000 time steps.

In our sensitivity analysis, our considered sensitivity parameters are the width $w_i$ and the permittivity $\epsilon_i$ of each layered medium for $i = 1, 2, \ldots, 6$, as indicated in Figure 2.6. To demonstrate the accuracy of the adjoint sensitivity technique, we computed the adjoint response sensitivities at 40 different nominal values of $w_6$, and
these values are evenly spaced between $0.1\lambda_0$ and $0.4\lambda_0$ where $\lambda_0$ is the wavelength for waves in the 6th layered medium. A subset of our computed adjoint sensitivity results are plotted in Figure 2.7 and Figure 2.8. The comparison shows that the adjoint sensitivity technique yields accurate estimation of the response sensitivity. At the same time, the adjoint technique also reduces the number required extra simulation from 24 to 1 for each nominal value of $w_6$.

2.7 Summary

This chapter introduced the simplest form of the adjoint variable method through
Figure 2.8  The adjoint sensitivities of the energy function (2.18) with respective to the thickness $w_3$ (a) and the relative permittivity $\epsilon_{r3}$ (b) of the third dielectric layer in the anti-reflection structure; the adjoint sensitivities of the energy function (2.18) with respective to the thickness $w_5$ (c) and the relative permittivity $\epsilon_{r5}$ (d) of the fifth dielectric layer in the anti-reflection structure.
time-domain electromagnetic simulations for sensitivity estimation with respect to frequency-independent constitutive parameters. The presented sensitivity technique is very easy to implement for all grid-based time-domain solvers, as long as the time-domain field solution near the vicinity of each perturbation surfaces can be exported during each simulation session. This sensitivity procedure is illustrated through two one-dimensional examples computed using the MATLAB. Good agreement is observed with results obtained through the CFD approximation. The advantage of this adjoint-based sensitivity technique, is in the computational efficiency rather than its accuracy if the target problem requires to consider multiple sensitivity parameters simultaneously.

References


In the previous chapter, we have discussed the adjoint-based sensitivity approach for estimating the response sensitivities of energy functions [1–3] through time-domain electromagnetic simulations. The resultant adjoint sensitivities are real valued scalars which are independent of frequency. This chapter considers another type of response functions that are dependent on frequency. In this scenario, the resultant adjoint sensitivities are often expressed in the complex domain. To illustrate the treatment for this type of response functions, we focus on the scattering parameters—one of the most important response quantifier in microwave engineering.
The adjoint treatment for scattering-parameters covered in this chapter was initially proposed in [4] and [5], which at that time, can only be applied in conjunction with the Transmission-line Matrix (TLM) based time-domain electromagnetic solvers [6]. In the subsequent research reported in [7–9], this adjoint treatment gradually became applicable to all other grid-based time-domain solvers, such as those based on the widely used finite-difference time-domain (FDTD) method [10]. Moreover, it was also noted in [5] that if all involved media in the simulation volume are reciprocal (which is the case for most problems), then the adjoint simulation is redundant to the original simulation, and is thus not required.

3.1 The response function for scattering-parameters

The scattering parameter (always complex valued) is a mathematical construct which describes how energy is spread within a given multi-port electromagnetic structure for a particular frequency at the steady state.

To compute the full scattering matrix of an electromagnetic structure containing $K$ ports, $K$ simulations are required with one of the ports being excited while the rest of the ports are terminated with matched loads. If the excitation is imposed at the $q$th port with a weighted modal distribution, then the expression of the $S\_{pq}$ parameter takes the following form [11]

$$S\_{pq} = \sqrt{\frac{Z\_q(j\omega) F\_{p\_q}(j\omega)}{Z\_p(j\omega) F\_q(j\omega)}} \quad (\text{if } p \neq q),$$

(3.1)
or for the case that $q$ and $p$ simultaneously refer to a same port

$$S_{pq} = \frac{F_{p/q}(j\omega) - \bar{F}_q(j\omega)}{\bar{F}_q(j\omega)} \quad \text{(if } p = q), \quad (3.2)$$

In (3.1) and (3.2), $Z_p$ and $Z_q$, respectively, denote the characteristic impedances at the two ports. The complex-valued term $\bar{F}_q$ denotes the weighted spectral sum of the equivalent voltage wave subject to the propagation mode over the plane of the $q$th port. This term, usually known beforehand, can be obtained by studying the field response of a reference structure with a matched load attached onto the $q$th port [7]:

$$\bar{F}_q(j\omega) = \frac{1}{\Delta z_q} \int_0^{T_{\text{max}}} \bar{w}_q^T \bar{E}_q e^{-j\omega t} dt. \quad (3.3)$$

In (3.3), the vector $\bar{E}_q$ contains the time-domain field solutions computed at the cells on the plane of the $q$th port. The over-line on $\bar{E}_q$ indicates that this field solution is computed from the reference simulation. The scalar $\Delta z_q$ denotes the thickness (or the longitudinal cell size) of the considered port plane. The vector $\bar{w}_q$ contains normalized modal weighting coefficients for each field entry computed at the cells on the port plane of the $q$th port. The dot-product between $\bar{w}_q$ and $\bar{E}_q$ reflects a spatial integration over the volume occupied by the cells on the considered port plane.

Similarly, the term $F_{p/q}$ in (3.2) and (3.1) is the weighted spectral sum of the total voltage wave due to the incident wave imposed at the $q$th port subject to the propagation mode over the plane of the $p$th port. To compute $F_{p/q}$ from the time-domain electromagnetic simulation, the following equation can be used

$$\bar{F}_{p/q}(j\omega) = \frac{1}{\Delta z_p} \int_0^{T_{\text{max}}} \bar{w}_p^T \bar{E}_p e^{-j\omega t} dt. \quad (3.4)$$
Here, the vector $\mathbf{E}_p$, a subset $\mathbf{E}$, contains the field solutions computed at the cells on the plane of the $p$th port from the nominal simulation. The scalar $\Delta z_p$ denotes the thickness (or the longitudinal cell size) of the $p$th port. The vector $\mathbf{w}_p$ contains normalized modal weighting coefficients for each field entry computed at the cells on the port plane of the $p$th port.

### 3.2 The expression of the complex adjoint sensitivity

In either (3.1) or (3.2), the port impedances $Z_p$, $Z_q$, and the spectral sum of the incident voltage wave $\bar{F}_q$ are not subjected to design changes. Thus, we can write the derivative of the scattering parameter $S_{pq}$ with respect to the $i$th sensitivity parameter $p_i \in \mathbf{p}$ as

$$\frac{\partial S_{pq}}{\partial p_i} = \sqrt{\frac{Z_q}{Z_p} \frac{1}{F_q}} \frac{\partial F_{p|q}}{\partial p_i},$$

(3.5)

and this sensitivity expression is fully determined by the differentiation term $\partial F_{p|q}/\partial p_i$.

At a given frequency, the response function $F_{p|q}$ and its sensitivity $\partial F_{p|q}/\partial p_i$ are complex values. To compute this complex sensitivity, our treatment is to decompose $F_{p|q}$ into its real and the imaginary parts as

$$F_{p|q} = F_{\text{Re}} + jF_{\text{Im}},$$

(3.6)

then we can study its derivative with respect to a given sensitivity parameter $p_i$ by considering the derivatives of two separate real valued response functions as

$$\frac{\partial F_{p|q}}{\partial p_i} = \frac{\partial F_{\text{Re}}}{\partial p_i} + j \frac{\partial F_{\text{Im}}}{\partial p_i}.$$  

(3.7)
At the frequency $\omega_0$, we apply the Euler’s identity on (3.4), then we can obtain the real and the imaginary parts of $F_{\omega_0}^{p|q}$ as

$$
F_{\omega_0}^{\text{Re}} = \frac{1}{\Delta z_p} \int_0^{T_{\text{max}}} w_p^T E_p \cos(\omega_0 t) dt, \quad (3.8)
$$

$$
F_{\omega_0}^{\text{Im}} = \frac{1}{\Delta z_p} \int_0^{T_{\text{max}}} w_p^T E_p \sin(-\omega_0 t) dt. \quad (3.9)
$$

Consider only the real part $F_{\omega_0}^{\text{Re}}$, and compare (3.8) with (2.17), we can immediately write the following kernel function $\phi$ as

$$
\phi_{\omega_0}^{\text{Re}}(E) = \frac{w_p^T E_p}{\Delta z_p} \cos(\omega_0 t). \quad (3.10)
$$

Here, the subscript Re and the superscript $\omega_0$ on $\phi_{\omega_0}^{\text{Re}}$ emphasize that this expression of $\phi$ is used for computing the adjoint sensitivity $\partial F_{\omega_0}^{\text{Re}} / \partial p_i$.

To obtain the adjoint sensitivity expression for $F_{\omega_0}^{\text{Re}}$, we first derive its corresponding adjoint current source $\hat{J}_{\omega_0}^{\text{Re}}$ using (3.10) and (2.35) to get

$$
\hat{J}_{\omega_0}^{\text{Re},p} = \frac{\sin(\omega_0 t)}{\omega_0 \Delta z_p} w_p \quad \text{and} \quad \hat{J}_{\omega_0}^{\text{Re},p} \subseteq \hat{J}_{\omega_0}^{\text{Re}}, \quad (3.11)
$$

where the subscript $p$ on $\hat{J}_{\omega_0}^{\text{Re},p}$ indicates that entries of $\hat{J}_{\omega_0}^{\text{Re}}$ are nonzero only at cells over the plane of the $p$th port. Here, we also denote the resultant vectors of adjoint field solutions due to the perturbation on $p_i$ as $\tilde{E}_{\omega_0}^{\text{Re}}$.

With a similar procedure applied to the imaginary part $F_{\omega_0}^{\text{Im}}$ given in (3.9), we may also obtain another expression of the adjoint current source for $\hat{J}_{\omega_0}^{\text{Im}}$ as

$$
\hat{J}_{\omega_0}^{\text{Im},p} = \frac{\cos(\omega_0 t)}{\omega_0 \Delta z_p} w_p \quad \text{and} \quad \hat{J}_{\omega_0}^{\text{Im},p} \subseteq \hat{J}_{\omega_0}^{\text{Im}}, \quad (3.12)
$$

It is worth to note that this adjoint current source $\hat{J}_{\omega_0}^{\text{Im},p}$ in (3.12) is in phase quadr-
rature with the \( \hat{\mathbf{J}}^{\omega_0}_{\text{Re, } p} \) given in (3.11). Therefore, the resultant vector of adjoint field solution \( \hat{\mathbf{E}}^{\omega_0}_{\text{Im}} \) can be obtained by shifting correspondent entries in \( \hat{\mathbf{E}}^{\omega_0}_{\text{Re}} \) forward in time by \( \pi/(2\omega_0) \) as
\[
\hat{\mathbf{E}}^{\omega_0}_{\text{Im}}(t) = \hat{\mathbf{E}}^{\omega_0}_{\text{Re}}(t - \frac{\pi}{2\omega_0}).
\]
(3.13)

Then using (2.33) as a guide, we obtain the following forward adjoint sensitivity expression for \( F_{p|q} \) at the frequency \( \omega_0 \) with respective to the sensitivity parameter \( p_i \) as
\[
\frac{\partial F_{p|q}^{\omega_0}}{\partial p_i} \approx -\frac{1}{\Delta p_i} \int_0^{T_{\text{max}}} \hat{\mathbf{E}}^{\omega_0}_{\text{Re}} \dot{\mathbf{R}} dt - j \frac{1}{\Delta p_i} \int_0^{T_{\text{max}}} \hat{\mathbf{E}}^{\omega_0}_{\text{Im}} \dot{\mathbf{R}} dt,
\]
(3.14)

where the vector \( \mathbf{R} \), known as the residual vector, can be computed through the nominal field solution \( \mathbf{E} \) using (2.27). Once the above adjoint sensitivity is obtained, we can then compute the sensitivities for the scattering parameter \( S_{pq} \) at the frequency \( \omega_0 \) using (3.5).

In the above derivation, both the adjoint current source \( \hat{\mathbf{J}}^{\omega_0}_{\text{Re}} \) and its phased equivalence \( \hat{\mathbf{J}}^{\omega_0}_{\text{Im}} \) caries a monochromatic signal at the frequency \( \omega_0 \). For time-harmonic media, we may expect that the resultant adjoint field solutions \( \hat{\mathbf{E}}^{\omega_0}_{\text{Re}} \) and \( \hat{\mathbf{E}}^{\omega_0}_{\text{Im}} \) are also monochromatic. If a problem requires to study the sensitivities for scattering parameters at multiple frequencies, one simple treatment is to excite the adjoint system with a wideband signal \( \hat{g}(t) \) as opposed to the sinusoidal signal \( \sin(\omega_0 t) \) used in (3.11) and (3.12) as
\[
\hat{\mathbf{J}}_p = \hat{g}(t) \frac{1}{\omega_0 \Delta z_p} \mathbf{w}_p \quad \text{and} \quad \hat{\mathbf{J}}_p \subseteq \hat{\mathbf{J}},
\]
(3.15)
where $\hat{\mathbf{J}}$ is non-zero only at cells on the plane of the $p$th port.

If the excitation vector $\hat{\mathbf{J}}_p$ is excited at the $p$th port in the adjoint simulation, then at the frequency $\omega_0$, the associated adjoint fields $\hat{\mathbf{E}}_{\text{Re}}^{\omega_0}$ and $\hat{\mathbf{E}}_{\text{Im}}^{\omega_0}$ can be extracted from the resultant wideband adjoint solution $\hat{\mathbf{E}}$ in means of spectral analysis through the following two formulas for each respective entry [4, 7]

$$
\hat{E}_{\text{Re}}^{\omega_0} = \frac{|\hat{E}(t)|_{\omega_0}}{|\hat{g}(t)|_{\omega_0}} \sin \left( \omega_0 t + \angle_{\omega_0} \hat{E}(t) - \angle_{\omega_0} \hat{g}(t) \right)
$$

$$
\hat{E}_{\text{Im}}^{\omega_0} = \frac{|\hat{E}(t)|_{\omega_0}}{|\hat{g}(t)|_{\omega_0}} \cos \left( \omega_0 t + \angle_{\omega_0} \hat{E}(t) - \angle_{\omega_0} \hat{g}(t) \right)
$$

where $|\hat{g}(t)|_{\omega_0}$ and $\angle_{\omega_0} \hat{g}(t)$ denotes the spectral amplitude and phase of the excitation signal $\hat{g}(t)$ computed at the frequency $\omega_0$; $|\hat{E}(t)|_{\omega_0}$ and $\angle_{\omega_0} \hat{E}(t)$ denotes the spectral amplitude and phase of the correspondent entry contained in $\hat{\mathbf{E}}$ computed at the frequency $\omega_0$.

### 3.3 The self-adjoint formula

We notice that the adjoint excitation signal $\hat{g}(t)$ in the expression of the adjoint current source (3.15) can be any arbitrary function of time, as long as it has smooth variations and covers the band of frequencies that we are interested in. Thereby, by prescribing the expression of $\hat{g}(t)$, it is possible for us to recast the resultant adjoint equation (2.42) into a form that is identical to the original wave equation (2.7). In this case, the adjoint field solution $\hat{\mathbf{E}}$ can be computed from the original wave equation.
(2.7) without solving addition adjoint equations. To achieve this goal, for cells on the plane of the $p$th port, the enclosed entries in the excitation vector $\mathbf{J}$ defined in (2.7) must also be in a form given by

$$\mathbf{J}_p = g(t)\mathbf{w}_p,$$  \hspace{1cm} (3.18)

where the excitation signal $g(t)$ is another customizable function of time, and excitation entries in $\mathbf{J}$ are nonzero only for those cells on the plane of the $p$th port.

From Section 2.5, we know that solving the adjoint field solution $\hat{\mathbf{E}}$ from the adjoint equation (2.45) in a backward time sequence using the excitation source $\hat{\mathbf{J}}$ is equivalent to solve $\check{\mathbf{E}}$ from the following wave equation by exciting another current source vector $\check{\mathbf{J}}$ forward in time as

$$\mathbf{K}\check{\mathbf{E}} + \mathbf{M}\check{\mathbf{E}} = \frac{d\check{\mathbf{J}}}{dt}, \hspace{1cm} (3.19)$$

where $\check{\mathbf{J}}$ relates to $\hat{\mathbf{J}}$ through

$$\check{\mathbf{J}}(t) = -\hat{\mathbf{J}}(T_{\text{max}} - t), \hspace{1cm} (3.20)$$

and $\hat{\mathbf{E}}$ relates to $\check{\mathbf{E}}$ through

$$\hat{\mathbf{E}}(t) = \check{\mathbf{E}}(T_{\text{max}} - t). \hspace{1cm} (3.21)$$

In (3.19), we select $\check{\mathbf{J}}$ equal to $\mathbf{J}$ given in (3.18). In this case, the equation (3.19) becomes identical to (2.7). Then on the plane of the $p$th port, we can write the following expression for the adjoint source vector $\hat{\mathbf{J}}$

$$\hat{\mathbf{J}}_p(t) = -\hat{\mathbf{J}}_p(T_{\text{max}} - t) = -g(T_{\text{max}} - t)\mathbf{w}_p. \hspace{1cm} (3.22)$$

Comparing the expressions of $\hat{\mathbf{J}}$ in (3.22) and (3.15), we obtain the following expres-
sion for the adjoint excitation signal \( \hat{g} \)

\[
\hat{g}(t) = -\Delta z \omega_0 g(T_{\text{max}} - t).
\]  

(3.23)

Using the expression of \( \hat{g}(t) \) given in (3.23) and applying the following two identities

\[
|f(T_{\text{max}} - t)|_{\omega_0} = |f(t)|_{\omega_0},
\]

(3.24)

\[
\angle \omega_0 f(T_{\text{max}} - t) = -\omega_0 T_{\text{max}} - \angle \omega_0 f(t),
\]

(3.25)

we can rewrite (3.16) and (3.17) in the following form [5]

\[
\hat{E}_{\text{Re}} = \frac{|E(t)|_{\omega_0}}{\Delta z \omega_0 |g(t)|_{\omega_0}} \sin \left( \omega_0 t + \angle \omega_0 E(t) - \angle \omega_0 g(t) \right),
\]

(3.26)

\[
\hat{E}_{\text{Im}} = \frac{|E(t)|_{\omega_0}}{\Delta z \omega_0 |g(t)|_{\omega_0}} \cos \left( \omega_0 t + \angle \omega_0 E(t) - \angle \omega_0 g(t) \right).
\]

(3.27)

Here, we notice that the real and imaginary parts of the adjoint field solution given in (3.26) and (3.27) only depend on the field solution \( E \) computed from the original wave equation (2.7) and its respective excitation signal \( g(t) \) given in (3.18). Therefore, in this scenario, there is no need to solve any additional adjoint wave equations. Once each adjoint field entry contained in \( \hat{E}_{\text{Re}} \) and \( \hat{E}_{\text{Im}} \) is computed using (3.26) and (3.27), we can then proceed to compute the sensitivities of the scattering parameter \( S_{pq} \) using (3.14) followed by (3.5).

### 3.4 Algorithm implementation and numerical examples

The theory of self-adjoint sensitivity discussed in this chapter requires no additional simulations, if only the sensitivities for scattering parameters are considered. For
$K$-port electromagnetic device, these adjoint sensitivities are computed at the same time when we compute all $K^2$ response entries in the scattering matrix from required $K$ simulations. Here, we summarize this self-adjoint sensitivity procedure into the follow three simple steps:

**Step 1. Parametrization**

For each sensitivity parameter, say $p_i$, we first identify its volumes of perturbations $\Psi_p^{\pm}$. Then, we group and label each cell within $\Psi_p^{\pm}$. This parametrization step is the same as that introduced in Section 2.6.

**Step 2. Direct analysis**

Perform $K$ simulations each with only one of the $K$ ports being excited using the excitation signal $g(t)$, while the rest of the ports are terminated with matched loads. For each simulation, if the excitation is imposed on the $q$th port, we first acquire the electric field solutions computed over all ports. Next, we use these field information to compute $\bar{F}_q(j\omega)$ and $F_{p|q}(j\omega)$ with (3.3) and (3.4) for $p, q = \{1, 2, ..., K\}$, and then compute the respective scattering parameter $S_{pq}$ using either (3.1) or (3.2). In this step, we also store the field solutions computed at cells within each volumes of perturbations determined from the Step 1.

**Step 3. Post-processing**
To compute the sensitivity for the $S_{pq}$ parameter with respect to the sensitivity parameter $p_i$ at the frequency $\omega_0$, we first compute the amplitude and phase of the excitation signal $g(t)$ at the frequency $\omega_0$ using the Fourier transform. We then perform the same spectral procedure on each time-dependent field solutions computed within the respective volumes of perturbation $\Psi_{p_i}^\pm$. With these field information expressed in the complex domain, we are then able to compute the real and the imaginary of the required adjoint vector using (3.26) and (3.27). Finally, $\partial S_{pq}/\partial p_i$ can be computed through (3.14) followed by (3.5). This step can be repeated for all other considered frequencies, and for all other choices of $p$ and $q$ in $S_{pq}$.

3.4.1 A multilayer anti-reflection dielectric coating

Our first illustrative example is the multi-layer anti-reflection (AR) filter shown in Figure 3.1. This multi-layer structure is optimized for plane waves whose incidence is normal to the plane of the material interface [12, 13]. It matches the wave impedance of a high permittivity region with $\epsilon_a = 12.25$ to the impedance of another region with low permittivity $\epsilon_b = 1.0$ (free space) over a bandwidth of $1.7\omega_0$, where $\omega_0$ is the central angular frequency of the passband. The nominal thickness of each dielectric layer equals quarter of the wavelength for waves within each layered medium at the frequency $\omega_0$.

The field solution of this passive structure is solved through an one-dimensional
Figure 3.1  A digram of a 6-layer dielectric anti-reflection structure whose outer regions have permittivity of $\epsilon_a = 12.25$ and $\epsilon_b = 1.0$ (free space). The relative permittivity of each dielectric slab is given by $\epsilon_1 = 13.5$, $\epsilon_2 = 8.41$, $\epsilon_3 = 4.86$, $\epsilon_4 = 2.52$, $\epsilon_5 = 1.46$, and $\epsilon_6 = 1.08$. Their nominal widths satisfy $w_i = z_i - z_{i-1} = \pi c_0/(2\sqrt{\epsilon_i \omega_0})$ for $i = 1, 2, ..., 6$, where $\omega_0$ is the center frequency of the passband.

FDTD solver. In our simulation setup, the total computational volume has a length of 50 mm truncated by Mur’s absorbing boundary conditions. The mesh sizes vary between 0.001 mm to 0.01 mm, and they are smoothly distributed along the direction of wave propagation. The cross section area of the computational volume is fixed at 5 mm$^2$. The FDTD time step is $\Delta t = c_0/\Delta h$, where $c_0$ denotes the speed of light in free space and $\Delta h$ denotes the smallest mesh size. The structure is excited with a wideband plane wave from its high permittivity side covering frequencies from 0 to $2\omega_0$.

In our sensitivity analysis, a total of 12 sensitivity parameters are considered, and these parameters includes the permittivity of each dielectric slabs $\epsilon_i$ and the thicknesses of each layer $w_i$ for $i = 1, 2, ..., 6$ as shown in Figure 3.1. Our considered
Figure 3.2 The adjoint sensitivities of the $S_{11}$ parameter with respect to the thickness $w_1$ (a) and the relative permittivity $\epsilon_1$ (b) of the first layered medium.

objective function is the return loss (the complex $S_{11}$ parameter) of this multilayer structure. Using our proposed self-adjoint technique, all 12 sensitivities are computed through only one simulation. A small subset of our produced sensitivity results are shown in Figure 3.2 and Figure 3.3, while all other obtained results pose the same degree of accuracy. The comparison shows that our adjoint results match those exact sensitivity results obtained through analytical approach [12]. If the same sensitivity procedure is repeated using the conventional central finite difference approximations, a total of 24 extra simulations would be required.
Figure 3.3 The adjoint sensitivities of the $S_{11}$ parameter with respect to the thickness $w_3$ (a) and the relative permittivity $\epsilon_3$ (b) of the third layered medium; the adjoint sensitivities of the $S_{11}$ parameter with respect to the thickness $w_5$ (c) and the relative permittivity $\epsilon_5$ (d) of the fifth layered medium.
3.4.2 A dielectric resonator antenna with matching insert

Our next illustrative example is a three-dimensional dielectric resonator antenna (DRA) [14, 15], see Figure 3.4. In our antenna configuration, the dielectric radiator is fed through a microstrip line printed on top of a substrate with relative permittivity 3.0 and a thickness of 0.762 mm. The width of the microstrip line is 1.9 mm. To achieve a wide-band coupling between the stripline and the dielectric radiator, a thin layer of high-permittivity insert is placed between the stripline and the dielectric radiator [15].

To obtain the scattering-parameters over a wide band of frequencies, we simulate this antenna in the time domain using the openEMS software package [16]. In our simulation setup, the excitation plane is placed 16.0 mm away from the radiator and its source signal is a Gaussian-modulated sinusoid centered at 15 GHz with a
Figure 3.5  The adjoint sensitivities of the $S_{11}$ parameter with respect to the permittivity $\varepsilon_{r,\text{insert}}$ (a) and the length $d_1$ (b) of the dielectric insert; the adjoint sensitivities of the $S_{11}$ parameter with respect to the length $d_2$ (c) and the height $h_2$ (d) of the dielectric resonator.
bandwidth of 10 GHz. The smallest mesh size that we use to model the antenna is
\[ \Delta h = 0.15 \text{ mm} \]. The time increment for the simulation is given by
\[ \Delta t = \Delta h / (2c_0) \].

To reduce the amount of computational effort, only half of the antenna structure
shown in Figure 3.4 is simulated through the utilization a symmetry plane.

In our sensitivity analysis, a total of 8 sensitivity parameters are considered.
These includes the width \( w_1 \), the height \( h_1 \), the length \( d_1 \), and the relative permittivity
\( \varepsilon_{r, \text{insert}} \) of the matching insert, as well as the width \( w_2 \), the height \( h_2 \), the length
\( d_2 \), and the relative permittivity \( \varepsilon_{r, \text{DRA}} \) of the dielectric resonator. The nominal
values of these considered parameters are given in the inset of Figure 3.4. With the
self-adjoint sensitivity technique, all 8 sensitivities of the return loss parameter are
computed through only one simulation. A subset of the obtained sensitivity results

Figure 3.6 The adjoint sensitivities of the \( S_{11} \) parameter with respect to the
width \( w_1 \) of the dielectric insert.
is shown in Figure 3.5 and Figure 3.6. All results including the ones not shown show good match with the central finite-difference (CFD) approximation which requires 12 extra simulations.

3.5 Summary

This chapter presented a simple treatment for handling frequency-dependent objective functions in the adjoint sensitivity analysis. The introduced technique is able to supply sensitivities of scattering parameters over a wide band of frequencies through at most one extra time-domain simulations. The sensitivity procedure is illustrated through two numerical electromagnetic examples. Good agreements are observed with results obtained through either the analytical approaches or the CFD approximations.

References


Chapter 4

The Adjoint Variable Method for Frequency-Dependent Constitutive Parameters

In Chapter 2, we discussed the adjoint technique for sensitivity computations concerning with perfect dielectric media whose permittivities are real and do not change with frequencies. This treatment, in general, produces accurate sensitivity estimation for problems involving constant approximations in modeling material permittivities.

In real world, as far as causality is concerned, all materials are lossy and their associated losses cannot vanish for all frequencies. The materials whose permittivities are functions of frequencies are generally known as dispersive materials. In frequency domain, we can simply account for the dispersive behavior of these materials through the multiplication in the time-harmonic constitutive relation $\mathcal{D}(s) = \varepsilon(s)\mathcal{E}(s)$. How-
ever, in the time-domain computational techniques, this multiplication of harmonic functions reflects a convolution in time. Therefore, additional effort is required to compute the time-dependent field solutions for frequency-dependent materials, and this, in turn, complicates in the sensitivity analysis.

In this chapter, our objective is to generalize the adjoint treatment developed in Chapter 2 to accommodate linear dispersive materials whose electric constitutive relations contain linear dispersion poles, such as those for Drude dispersion, Lorentz dispersion, Debye dispersion, or a mixture of them. Through our deviation, we show that the same sensitivity procedure developed in Chapter 2 and Chapter 3 can also be applied for treating sensitivity parameters associated with dispersive media. Our procedure is illustrated numerically through various practical examples using the MATLAB.

The theory covered in this chapter is initially proposed in [1] for transmission-line matrix (TLM) based solvers [2] and also in [3] for finite-difference time-domain (FDTD) based solvers [4]. In [5] and [6], this adjoint procedure is applied to study the sensitivities for plasmonic devices.

4.1 The system matrix for dispersive mediums

The derivation procedure for the dispersive adjoint expression is very similar to that for the non-dispersive media as discussed in Chapter 2, despite the fact that the
expression of the system matrix $M$ in the wave equation (2.7) involves time convolutions.

### 4.1.1 The Lorentz dispersion

Without loss of generality, we consider an arbitrary $y$-polarized field solution entry, say $E_{y}^{(j)} \in \mathbf{E}$, computed at the $j$th cell in an isotropic computational domain. We first investigate the case when the electric constitutive relation at the considered cell contains only one Lorentz dispersion pole $\epsilon_L(s)$

$$D_y^{(j)}(s) = \epsilon E_y^{(j)}(s) + \epsilon_L(s) E_y^{(j)}(s)$$

$$= \epsilon E_y^{(j)}(s) + \frac{\alpha}{s^2 + \gamma s + \beta} E_y^{(j)}(s), \quad (4.1)$$

where $D_y^{(j)}$ denotes the respective electric flux density of $E_y^{(j)}$ computed at the same cell; the scalar constants $\epsilon$, $\alpha$, $\beta$, and $\gamma$ ($\gamma^2 < 4\beta$) describes the Lorentz dispersion at the considered cell.

Multiplying both sides of (4.1) by $s^2$ and defining a new scalar constant $\omega_d$ such that $\omega_d = \sqrt{\beta - \gamma^2 / 4}$, we have

$$s^2 D_y^{(j)}(s) = \epsilon s^2 E_y^{(j)}(s) + \alpha E_y^{(j)}(s) - \frac{\alpha(\omega_d^2 - \gamma^2 / 4)}{(s + \gamma/2)^2 + \omega_d^2} E_y^{(j)}(s)$$

$$- \frac{\alpha \gamma (s + \gamma/2)}{(s + \gamma/2)^2 + \omega_d^2} E_y^{(j)}(s). \quad (4.2)$$

Transforming (4.2) back to the time domain, we obtain the following time-domain expression of the second-order time-derivative of the flux density $D_y^{(j)}$

$$\frac{\partial^2}{dt^2} D_y^{(j)}(t) = \epsilon \frac{\partial^2}{dt^2} E_y^{(j)}(t) + \alpha E_y^{(j)}(t) + \chi_L(t) * E_y^{(j)}(t), \quad (4.3)$$
where the operator $\ast$ denotes forward convolution in time and the time-dependent expressions of $\chi_L(t)$ is given by

$$
\chi_L(t) = -\alpha \left( \frac{4\omega_d^2 - \gamma^2}{4\omega_d} \sin(\omega_d t) + \gamma \cos(\omega_d t) \right) \exp \left( -\frac{\gamma t}{2} \right) u(t),
$$

and $u(t)$ denotes the unit step function in the time domain.

In this case where the constitutive relation contains only a single Lorentz pole, using (4.3) as a guide, we expect that the respective diagonal entries of $M$ in the wave equation (2.7) takes the form:

$$
M^{(\j)\cdot f} = \epsilon \frac{\partial^2}{\partial t^2} f + \alpha f + \chi_L \ast f,
$$

where $f$ represents an arbitrary second-order differentiable function of time.

### 4.1.2 The Drude dispersion

Next, we consider another case when the electric constitutive relation imposed at the considered cell follows a Drude dispersive relation described by $\epsilon_D(s)$ as

$$
D_y^{(\j)}(s) = \epsilon E_y^{(\j)}(s) + \epsilon_D(s)E_y^{(\j)}(s)
= \epsilon E_y^{(\j)}(s) + \frac{\alpha}{s^2 + \gamma s} E_y^{(\j)}(s).
$$

Following the same treatment as for the Lorentz dispersive relation, we multiply both side of (4.6) by $s^2$, and we obtain the following

$$
s^2 D_y^{(\j)}(s) = \epsilon s^2 E_y^{(\j)}(s) + \alpha E_y^{(\j)}(s) - \frac{\alpha \gamma}{s + \gamma} E_y^{(\j)}(s).
$$

Transforming (4.7) back in the time domain, we then have

$$
\frac{\partial^2}{\partial t^2} D_y^{(\j)}(t) = \epsilon \frac{\partial^2}{\partial t^2} E_y^{(\j)}(t) + \alpha E_y^{(\j)}(t) + \chi_D(t) \ast E_y^{(\j)}(t),
$$
where, in this case, the expression of $\chi_D(t)$ is given by

$$\chi_D(t) = -\alpha \gamma \exp(-\gamma t) u(t). \quad (4.9)$$

We can thus write the expression for the respective diagonal entries of $M$ as

$$M_y^{(j)} \cdot f = \epsilon \frac{\partial^2}{\partial t^2} f + \alpha f + \chi_D * f. \quad (4.10)$$

### 4.1.3 The Debye dispersion and the conductivity term

Now, consider the case when the electric constitutive relation imposed at the same field entry in the considered cell contains only a Debye dispersion pole $\epsilon_B(s)$ as

$$D_y^{(j)}(s) = \epsilon E_y^{(j)}(s) + \epsilon_B(s) E_y^{(j)}(s)$$

$$= \epsilon E_y^{(j)}(s) + \frac{\alpha}{s + \beta} E_y^{(j)}(s). \quad (4.11)$$

Multiply both side of (4.11) by $s^2$, and perform the partial fraction decomposition on the right-hand side of the resultant equation. It follows that

$$s^2 D_y^{(j)}(s) = \epsilon s^2 E_y^{(j)}(s) + \alpha s E_y^{(j)}(s) - \alpha \beta E_y^{(j)}(s) + \frac{\alpha \beta^2}{s + \beta} E_y^{(j)}(s). \quad (4.12)$$

Transforming (4.7) back in the time domain, we get

$$\frac{\partial^2}{\partial t^2} D_i(t) = \epsilon \frac{\partial^2}{\partial t^2} E_i(t) + \alpha \frac{\partial}{\partial t} E_i(t) + \alpha \beta E_i(t) + \chi_B(t) * E_i(t), \quad (4.13)$$

and the expression of $\chi_B(t)$ is given by

$$\chi_B(t) = \alpha \beta^2 \exp(-\beta t) u(t). \quad (4.14)$$

We thus obtain the following expression for the respective diagonal entry on $M$

$$M_y^{(j)} \cdot f = \epsilon \frac{\partial^2}{\partial t^2} f + \alpha \frac{\partial}{\partial t} f + \alpha \beta f + \chi_B * f. \quad (4.15)$$
Now, consider a special scenario where the constitutive relation for the field entry \( E^{(j)}_y \) describes a lossy dielectrics with a constant electric conductivity \( \alpha \) as

\[
D^{(j)}_y(s) = \varepsilon E^{(j)}_y(s) + \varepsilon_\sigma(s)E^{(j)}_y(s) \\
= \varepsilon E^{(j)}_y(s) + \frac{\alpha}{s} E^{(j)}_y(s). 
\]

(4.16)

In this case, by setting the value of \( \beta \) to zero in (4.11), we could obtain the following expression for the respective diagonal entry of \( M \) as

\[
M^{(j)}_y \cdot f = \varepsilon \frac{\partial^2}{\partial t^2} f + \alpha \frac{\partial}{\partial t} f. 
\]

(4.17)

Our derived expression (4.17) agrees with the ones published in [7] and [8], but instead, through a different approach.

4.1.4 Multiple dispersion poles

So far, we have considered three different scenarios when the electric constitutive relation imposed at the \( j \)th cell contains only one linear dispersion pole. For all investigated cases, we can write the following generic expression for the diagonal entries of \( M \)

\[
M^{(j)}_{x/y/z} \cdot f = \kappa_2 \frac{\partial^2}{\partial t^2} f + \kappa_1 \frac{\partial}{\partial t} f - \kappa_0 f + \chi(t) * f, 
\]

(4.18)

and the expressions of \( \kappa_2, \kappa_1, \kappa_0, \) and \( \chi(t) \), which are local to the \( j \)th cell, are summarized in the Table 4.1.

For a generalized situation, in which the constitutive relation imposed on the \( j \)th cell contains multiple linear dispersion poles as
Table 4.1 Dispersion Coefficients in the System Matrix $M$

<table>
<thead>
<tr>
<th>dispersion term</th>
<th>$\kappa_2$</th>
<th>$\kappa_1$</th>
<th>$\kappa_0$</th>
<th>$\chi(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>scalar term</td>
<td>$\epsilon$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>conductivity term</td>
<td>$\epsilon(s) = \frac{\alpha}{s}$</td>
<td>0</td>
<td>$\alpha$</td>
<td>0</td>
</tr>
<tr>
<td>Debye pole</td>
<td>$\epsilon_B(s) = \frac{\alpha}{s + \beta}$</td>
<td>0</td>
<td>$\alpha$</td>
<td>$-\alpha\beta$ exp($-\beta t$)$u(t)$</td>
</tr>
<tr>
<td>Drude pole</td>
<td>$\epsilon_D(s) = \frac{\alpha}{s^2 + \gamma s}$</td>
<td>0</td>
<td>0</td>
<td>$\alpha\gamma$ exp($-\gamma t$)$u(t)$</td>
</tr>
<tr>
<td>Lorentz pole</td>
<td>$\epsilon_L(s) = \frac{\alpha}{s^2 + \gamma s + \beta}$</td>
<td>0</td>
<td>0</td>
<td>$\alpha$ $-\alpha\frac{4\gamma^2 - \gamma^2}{4\omega_d} \sin(\omega_d t) \exp\left(-\frac{\gamma t}{2}\right)u(t)$ $-\alpha\gamma\cos(\omega_d t) \exp\left(-\frac{\gamma t}{2}\right)u(t)$</td>
</tr>
</tbody>
</table>

$D_y^{(j)}(s) = \left(\epsilon + \epsilon_\sigma(s) + \sum_{k_B=1}^{K_B} \epsilon_{B,k_B}(s) + \sum_{k_D=1}^{K_D} \epsilon_{D,k_D}(s) + \sum_{k_L=1}^{K_L} \epsilon_{L,k_L}(s)\right)E_y^{(j)}(s)$, \hspace{1cm} (4.19)

then through the principle of superposition, the resulted system matrix $M$ would still take the same form as (4.18), except that the coefficients $\kappa_2$, $\kappa_1$, $\kappa_0$, and $\chi(t)$ in (4.18) are determined by

$$\kappa_{0/1/2} = \sum_{k=1}^{K} \kappa_{0/1/2,k} \hspace{1cm} \text{and} \hspace{1cm} \chi = \sum_{k=1}^{K} \chi_k, \hspace{1cm} (4.20)$$

where $K = K_B + K_D + K_L + 1$ is the total number of poles contained in the constitutive relation (4.19); the values of $\kappa_{0,k}$, $\kappa_{1,k}$, $\kappa_{2,k}$, and the expression of $\chi_k$, which can be computed from Table 4.1, denotes the linear contribution due to the $k$th dispersion
4.2 The solution of the dispersive adjoint equation

Once we obtained the expression of the dispersive system matrix $M$, the sensitivity derivation is line-by-line identical to those discussed in Section 2.4 (see (2.25) to (2.36)). Therefore the same sensitivity expressions (2.33) or (2.36) can be used to compute the adjoint sensitivity associated with dispersive mediums. The only difference is that the dispersive expression of $M$ given in (4.18) is used instead for the computation of the residual vector $R$, see (2.27).

To solve the dispersive adjoint equation (2.34), we remove the transpose operations on diagonal matrices $K$ and $\hat{M}$ (2.34) and apply the one-to-one field mapping procedure. Then the adjoint equation (2.34) can be approximated as

$$\hat{K}\hat{E} + \hat{M}\hat{E} \approx \frac{d\hat{J}}{dt},$$

(4.21)

and the system matrix $\hat{M}$ in the adjoint equation must satisfy (see the treatment in (2.30))

$$\int_0^{T_{\text{max}}} (\hat{M}^T \hat{E})^T \Delta E \, dt = \int_0^{T_{\text{max}}} \hat{E}^T (M \Delta E) \, dt.$$  

(4.22)

Using the expression of $M$ given in (4.18), we may rewrite the matrix-vector multiplications in the right-hand side integral term in the summation form of element-wised multiplications as
\[
\int_0^{T_{\text{max}}} \hat{E}^T(M \Delta E) \, dt = \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \kappa_2^{(j)} \hat{E}^{(j)T} \frac{d^2 \Delta E^{(j)}}{dt^2} \, dt + \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \kappa_1^{(j)} \hat{E}^{(j)T} \frac{d \Delta E^{(j)}}{dt} \, dt \\
+ \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \kappa_0^{(j)} \hat{E}^{(j)T} \Delta E^{(j)} \, dt + \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \hat{E}^{(j)T} \left( \chi^{(j)}(t) \ast \Delta E^{(j)} \right) \, dt \tag{4.23}
\]

where the superscript \((j)\) on \(\kappa_0^{(j)}, \kappa_1^{(j)}, \kappa_2^{(j)}\), and of \(\chi^{(j)}\) indicates that these dispersion coefficients are local to the \(j\)th cell in the computational volume.

Following exactly the same mathematical procedure given by (2.38) to (2.41), we can rewrite the first integral term in the right-hand side of (4.23) as

\[
\int_0^{T_{\text{max}}} \sum_{j=1}^{M} \kappa_2^{(j)} \hat{E}^{(j)T} \frac{d^2 \Delta E^{(j)}}{dt^2} \, dt = \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \left( \kappa_2^{(j)} \frac{d^2 \hat{E}^{(j)}}{dt^2} \right)^T \Delta E^{(j)} \, dt \tag{4.24}
\]
as long as the terminal conditions specified in (2.40) simultaneously hold.

Considering now the second integral term in the right-hand side of (4.23) and integrate it by parts. With the conditions given in (2.40), we have

\[
\int_0^{T_{\text{max}}} \sum_{j=1}^{M} \kappa_1^{(j)} \hat{E}^{(j)T} \frac{d \Delta E^{(j)}}{dt} \, dt = \sum_{j=1}^{M} \kappa_1^{(j)} \hat{E}^{(j)T} \Delta E^{(j)} \left[ \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \kappa_1^{(j)} \frac{d \hat{E}^{(j)}}{dt} \right]^T \Delta E^{(j)} \, dt \\
= - \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \left( -\kappa_1^{(j)} \frac{d \hat{E}^{(j)}}{dt} \right)^T \Delta E^{(j)} \, dt. \tag{4.25}
\]

On the last term of right-hand side of (4.23), if \(\chi^{(j)}(t)\) is only nonzero within the closed interval \([0, T_{\text{max}}]\) at all cells, we can further rewrite the convolution operation in the integrand as

\[
\int_0^{T_{\text{max}}} \sum_{j=1}^{M} \hat{E}^{(j)} \left( \chi^{(j)}(t) \ast \Delta E^{(j)} \right) \, dt = \sum_{j=1}^{M} \hat{E}^{(j)T} \int_0^{T_{\text{max}}} \chi^{(j)}(\tau) \Delta E^{(j)}(t - \tau) \, d\tau \, dt
\]
\[
\sum_{j=1}^{M} \int_0^{T_{\text{max}}} \Delta E^{(j)T} \int_0^{T_{\text{max}}} \chi^{(j)}(\tau) \hat{E}^{(j)}(\tau - t) d\tau \, dt
\]
\[
\int_0^{T_{\text{max}}} \sum_{j=1}^{M} \left( \chi^{(j)}(t) \ast \hat{E}^{(j)}(t) \right)^T \Delta E^{(j)} \, dt, \quad (4.26)
\]
where the operation \( f_1 \ast f_2 \) in (4.26) denotes backward time-convolution as:
\[
f_1 \ast f_2 = \int_t^{T_{\text{max}}} f_1(\tau) f_2(\tau - t) d\tau. \quad (4.27)
\]

Substitute the first, second, and last integration terms in the right-hand side of (4.23) with their respective expressions given in (4.24), (4.25), and (4.26), we have
\[
\int_0^{T_{\text{max}}} \hat{E}^T (M \Delta E) \, dt = \int_0^{T_{\text{max}}} \sum_{j=1}^{M} \left( \kappa_2^{(j)} \frac{d^2 \hat{E}^{(j)}}{dt^2} - \kappa_1^{(j)} \frac{d \hat{E}^{(j)}}{dt} + \kappa_0^{(j)} \hat{E}^{(j)} \right)
\]
\[
+ \chi^{(j)}(t) \ast \hat{E}^{(j)} \Delta E^{(j)} \, dt \quad (4.28)
\]
Comparing the resultant new expression (4.28) with (4.22), we then obtain the following generic expression for the diagonal entry of the system matrix \( \hat{M} \) in the dispersive adjoint equation (4.21) as
\[
\hat{M}^{(j)}_{x/y/z} \cdot f = \kappa_2^{(j)} \frac{d^2 f}{dt^2} - \kappa_1^{(j)} \frac{d f}{dt} + \kappa_0^{(j)} f + \chi^{(j)}(t) \ast f. \quad (4.29)
\]
Thus if the adjoint dispersive equation (4.21) is solved backward in time starting at \( t = T_{\text{max}} \), then the same simulation technique used for solving the respective direct problem (2.7) can be applied here to solve the adjoint field solution \( \hat{E} \). This treatment is the same as that for solving the non-dispersive adjoint equation discussed in Section 2.5.
4.3 Algorithm implementation and numerical examples

The algorithm for computing the adjoint sensitivities for dispersive mediums is the same as that for non-dispersive media discussed in Section 2.6. It requires two simulations for sensitivity estimation with respective to all sensitivity parameters, regardless of their number. Due to the dispersive poles in the constitutive relation (4.19), the computation of the residual vector \( \mathbf{R} \) involves time-convolution between the field solution \( \mathbf{E} \) and the material-dependent function \( \chi(t) \), see (4.18). These convolution operations, however, can be efficiently implemented with recursive approaches, thus their contributed computational overhead is insignificant.

The adjoint vector \( \hat{\mathbf{E}} \) can be computed using the same time-domain solver as the one used in the original problem. Similar to the case for non-dispersive medium, the adjoint vector \( \hat{\mathbf{E}} \) is only evaluated for the nominal structure. For each sensitivity parameter, we approximate corresponding entries in the perturbed adjoint vector \( \hat{\mathbf{E}} \) through the procedure of one-to-one field mapping. The final estimated adjoint sensitivities are then computed using either (2.33) or (2.36).

The theory summarized this chapter is numerically validated through four practical examples involving non-magnetized plasma and Lorentzian materials. In all numerical examples, the time-domain field solutions are computed through the FDTD algorithms and the dispersive behavior of materials are treated with the auxiliary differential equation (ADE) technique [9].
4.3.1 Thin Lorentz slab in a parallel-plate waveguide

In our first illustrating example, we consider a thin Lorentzian slab in a parallel-plate waveguide illuminated by a plane wave with zero incidence angle, see Figure 4.1. Our adjoint technique is applied to estimate the sensitivity of scattering-parameters with respect to geometrical and material properties of the inserted dispersive slab. The waveguide has a width of 3.75 mm and is excited with a wideband Gaussian pulse centered at 50.0 GHz with a bandwidth of 70.0 GHz. The dispersive Lorentzian slab has a width $L = 1.5$ mm and contains three Lorentz poles in the constitutive relation:

$$D(s) = \left( \epsilon_0 \epsilon_\infty + \sum_{k=1}^{3} \frac{\epsilon_p \epsilon_0 \omega_p^2}{s^2 + \zeta_s + \omega_p^2} \right) E(s).$$

(4.30)

Here, all dispersion parameters contained in (4.30) are summarized in the inset of Figure 4.1.

<table>
<thead>
<tr>
<th>Pole</th>
<th>$\omega_p$</th>
<th>$\zeta$</th>
<th>$\epsilon_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30 GHz</td>
<td>3 GHz</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>40 GHz</td>
<td>8 GHz</td>
<td>0.3</td>
</tr>
<tr>
<td>3</td>
<td>50 GHz</td>
<td>5 GHz</td>
<td>0.9</td>
</tr>
<tr>
<td>$\epsilon_\infty = 1.5$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The field responses of this electromagnetic structure are computed through the FDTD algorithm. For accurate modeling, a uniform spatial discretization of $\Delta h = 75.0 \, \mu m$ is utilized. In our analysis, a total of 11 parameters are considered. These parameters include the thickness of the Lorentzian slab $L$, the high-frequency dielectric constant of the slab $\epsilon_\infty$, and the dispersion parameters $\omega_p$, $\zeta$, and $\epsilon_p$ for each of the three dispersion poles.

Using the dispersive sensitivity procedure developed in this chapter, the adjoint sensitivities of the scattering parameters ($S_{11}$ and $S_{21}$) with respect to all variables are evaluated using only one extra FDTD simulation. A subset of the adjoint sensitivity results are shown on Figure 4.2 and Figure 4.3. All other results that are not...
Figure 4.3  Adjoint sensitivities of the scattering parameters with respect to the resonant frequency $\omega_p$ of the second Lorentz dispersion pole (a), the damping frequency $\zeta$ (b) and the dispersion coefficient $\epsilon_p$ (c) of the third Lorentz dispersion pole of the thin slab.
shown maintain the same degree of accuracy. Our produced results demonstrate good accuracy in comparison with those obtain through central finite-difference (CFD) approximations.

4.3.2 Tunneling through a metamaterial slab

In the second example, the Drude model is utilized to model a slab of metamaterial (see Figure 4.4). The metamaterial is able to reroute the electromagnetic waves by an angle of 180° with minimal reflection [10]. Inside this device, the electromagnetic wave is coupled from a parallel-plate waveguide of width $a$ to another through the thin slab of metamaterial with width $D = 0.1a$. The material contains a single Drude pole with plasma frequency $\omega_p = 2c_0/a$ and a collision time constant $\tau_c = 20/\omega_p$ as

$$D(s) = \left(\epsilon_0\epsilon_\infty + \frac{\epsilon_0\omega_p^2}{s^2 + s/\tau_c}\right)E(s), \quad (4.31)$$
Figure 4.5 Adjoint sensitivities of the scattering parameters with respect to the thickness $D$ (a), the plasma frequency $\omega_p$ (b), and the collision time constant $\tau_c$ of the metamaterial slab.
where $c_0$ denotes the speed of light in free space. The optimal tunneling performance at abrupt bands occurs at specified wavelength at which the total permittivity vanishes [10].

This electromagnetic device is simulated through a two-dimensional FDTD solver with three nonzero field components $E_x$, $E_z$, and $H_y$. The input port is excited with a wideband Gaussian signal centered at $\omega_0 = 2c_0/a$. In our simulations, we choose $\Delta h = 0.01a$ to be the cell size for all cells and $\Delta t = 0.5\Delta h/c_0$ to be time step. The adjoint scattering parameters with respect to three considered parameters $D$, $\omega_p$, and $\tau_e$ are given in Figure 4.2. Good accuracy is observed when they are compared with those sensitivities approximated through conventional CFD approach.

4.3.3 The add-drop coupler with split-ring resonator

In this numerical example, the adjoint sensitivity approach is demonstrated through a compact add-drop coupler consisting of two air-filled parallel metal-insulator-metal (MIM) plasmonic waveguides [11] (see Figure 4.6). The two waveguides are 50.0 nm wide and are separated by a distance of 510.0 nm. The coupling of the two waveguides is enhanced through an insulator-filled square ring resonator (SRR) centered between the waveguides. The square ring of the SRR has a width 50.0 nm and its outer edge is 470.0 nm long. The metal utilized is silver, and we model it using the following Drude dispersive relation

$$D(s) = \left( \epsilon_0 \epsilon_\infty + \frac{\epsilon_0 \omega_p^2}{s^2 - v_e s} \right) E(s),$$

(4.32)
where the associated dispersion parameters are given by $\epsilon_{\infty} = 3.7$, $\omega_p = 1.38 \times 10^{16}$ rad/s, and $v_c = 27.3$ THz.

We excite this plasmonic device with a wideband uniform line source across port 1. The field responses are computed through the two-dimensional FDTD formulae. In our simulation, we choose $\Delta h = 2.5$ nm as the mesh size for all cells and $\Delta t = c_0/(2\Delta h)$ as the FDTD time step. Our considered sensitivity parameters are the edges of the SRR structure $\mathbf{p} = [p_1 \ldots p_8]^T$.

Our resultant adjoint sensitivities with respect to parameters $p_1, p_2, p_3, p_5, p_7,$ and $p_8$ are shown in Figure 4.7 and Figure 4.8, respectively. The comparison shows that the AVM results agree well with those approximated through the CFD approach.
4.3.4 Gold split ring with normal plane-wave incidence

In this example, we investigate the adjoint sensitivity with respect to shape parameters of a gold split ring due to normal incidence of the plane wave (see Figure 4.9). The gold split ring has a thickness of 30.0 nm and is modeled by a single Drude pole as

\[
D(s) = \left( \epsilon_0 \epsilon_\infty + \frac{\epsilon_0 \omega_p^2}{s^2 + s/\tau_c} \right) E(s),
\]

(4.33)

where the dispersion parameters are given by \( \epsilon_\infty = 1.0, \omega_p = 1.367 \times 10^{16} \text{ rad/s}, \) and \( \tau_c = -1.544 \times 10^{-14} \text{ s}. \) Beneath the gold resonator, there is a glass substrate with a refractive index of 1.46. The vector of sensitivity parameters \( p = [p_1 \ldots p_8]^T \)
Figure 4.8  Adjoint sensitivities of the scattering parameters with respect to the edge $p_3$, the edge $p_5$ (b), the edge $p_7$ (c), and the edge $d$ (bottom) of the add-drop coupler.
represents the edge lengths of the gold structure. Its nominal values are given by
\[ \mathbf{p} = [200.0 \ 200.0 \ 120.0 \ 70.0 \ 65.0 \ 120.0 \ 65.0 \ 200.0]^T \text{ in nano-meters.} \]
At 200.0 THz, the structure exhibits a pronounced fundamental LC resonance, which would lead to a negative magnetic permeability [12].

In our FDTD simulation, we choose the mesh size \( \Delta h = 5.0 \text{ nm} \) for all cells with an FDTD time step \( \Delta t = \Delta h/2c_0 \). The discretized structure is excited with a plane wave normal to the surface of the gold split ring. Perfect electric conducting boundary conditions is imposed to mimic the effect of the periodic boundary conditions for the normal incidence case. The excitation signal is a wideband Gaussian pulse centered at 250.0 THz with a bandwidth of 200.0 THz.
Figure 4.10  Adjoint sensitivities of the scattering parameters with respect to the edge $p_1$ (a), the edge $p_3$ (b), the edge $p_4$, and the edge $p_5$ of the gold split ring.
Our adjoint sensitivity results of scattering parameters with respect to four shape parameters are given in Figure 4.10. All other results that are not shown pose the same degree of accuracy. The comparison shows that our adjoint sensitivity results matches well with those obtained through CFD approximations.

4.4 Summary

This chapter discussed the adjoint sensitivity technique for treating frequency-dependent materials through time-domain electromagnetic simulations. We have demonstrated that the same adjoint procedure developed in Chapter 2 or Chapter 3 can also be utilized to compute the sensitivities associated with dispersive materials, once the non-zero entries in the residual vectors are computed using our derived formulas. Our adjoint treatment for dispersive materials is accurate and does no yield additional computational overhead. The sensitivity procedure is numerically illustrated through the sensitivity analysis on the scattering parameters for various electromagnetic devices.

References


Chapter 5

The Second-Order Adjoint Variable Method for Time-Domain Methods

In all previous chapters, we have shown that the adjoint variable method (AVM) is able to supply accurate approximation of response sensitivities with respect to parameters regardless of their number. With some modification, the same adjoint technique could also be extended to supply the second-order sensitivity (or Hessian) matrix of a given response function with the same degree of efficiency. Second-order sensitivity information is useful in a wide range of engineering problems where knowing only the first-order response sensitivity is inadequate. For example, in many large-scale optimization problems when Newton-type methods are used, both the first- and the second-order sensitivity information are required. The class of optimization
techniques requiring Hessian matrices, in general, has a faster rate of convergence than Hessian-free optimization techniques [1].

So far, the second-order adjoint sensitivities have been applied in many areas of engineering including mechanical system design [2, 3], fluid dynamics [4], and narrow band analysis of microwave filters [5–8]. Recently, a second-order adjoint sensitivity technique were developed for wideband time-domain transmission-line matrix (TLM) based electromagnetic simulations [9].

In this chapter, we cover the computation of second-order adjoint sensitivities through the popular finite-difference time-domain (FDTD) solvers. We show that with the adjoint technique, the complete $N^2$ sensitivity entries in the Hessian matrix of the response function can be estimated using at most $N + 1$ extra simulations, if $N$ is the total number of considered sensitivity parameters. If the same procedure is performed using the central finite-difference (CFD) approximation, the number of required simulations would be $O(N^2)$.

5.1 The second-order response sensitivities

We consider response functions of the form given by (2.17) in Chapter 2. Using the chain rule, the second-order derivative of (2.17) with respect to two arbitrary sensitivity parameters $p_i, p_j \in p$ can be written as [9]

$$\frac{\partial^2 F}{\partial p_i \partial p_j} = \frac{\partial^{[e]}F}{\partial p_i \partial p_j} + \int_0^{T_{\max}} \left( \frac{\partial \mathbf{E}}{\partial p_i} \right)^T \frac{\partial^2 \phi}{\partial \mathbf{E}^2} \frac{\partial \mathbf{E}}{\partial p_j} dt + \int_0^{T_{\max}} \left( \frac{\partial \phi}{\partial \mathbf{E}} \right)^T \frac{\partial^2 \mathbf{E}}{\partial p_i \partial p_j} dt, \quad (5.1)$$
where the first term on the left-hand-side of (5.1) denotes the explicit second order derivative of $F$. If the objective function $F$ computes the scattering parameter of the considered electromagnetic devices, then the first two terms on the right-hand-side of (5.1) reduce to zeros [9].

Conventionally, the second-order derivative (5.1) is computed numerically through the finite-difference approximation techniques. This approach requires the response function to be evaluated at perturbed states of $p$. The most commonly used central finite-difference expression is given by [10]

$$\frac{\partial^2 F}{\partial p_i \partial p_j} = \frac{1}{4\Delta p_i \Delta p_j} \begin{pmatrix} F(p_i + \Delta p_i, p_j + \Delta p_j) \\ -F(p_i + \Delta p_i, p_j - \Delta p_j) \\ -F(p_i - \Delta p_i, p_j + \Delta p_j) \\ +F(p_i - \Delta p_i, p_j - \Delta p_j) \end{pmatrix}. \tag{5.2}$$

With (5.2), the computation of all $N^2$ second-order sensitivity entries in the Hessian matrix, in general, requires $O(N^2)$ extra simulations. This cost, even for a small $N$, is computationally prohibitive.

## 5.2 The second-order adjoint variable method

For the sake of simplicity, we assume that our considered computational domain is isotropic and non-magnetized. Consider two arbitrary parameters $p_i$ and $p_j$ in $p$, and $i \neq j$. We first perturb the system matrix $M$ in the wave equation (2.7) on the forward direction of $p_i$ by an incremental amount $\Delta p_i$ (see Figure 5.1(a)). Due to this
Figure 5.1 Illustration of changes in the system matrix $M$ and the field solution $E$ due to shape perturbations on a rectangle discontinuity (shown by dashed region) in an FDTD grid. (a) Perturbing the shape parameter $p_i$ updates the field solution to be $E + \Delta_i E$; (b) perturbing the shape parameter $p_j$ updates the field solution to be $E + \Delta_j E$; (c) Simultaneously perturbing both $p_i$ and $p_j$ updates the field solution to be $E + \Delta_{ij} E$; (d) affected cells corresponding to non-zero entries in $\Delta_{ij} M - \Delta_i M$; (f) affected cells corresponding to non-zero entries in $\Delta_{ij} M - \Delta_j M$; (e) affected cells corresponding to non-zero entries in $\Delta_{ij} M$. 

100
perturbation, the matrix $M$ becomes $M + \Delta^+_i M$, where the symbol $\Delta^+_i$ denotes the differential changes due to the perturbation on the positive (or forward) direction of $p_i$. As a result, the field solution changes from $E$ to $E + \Delta^+_i E$ accordingly and (2.7) becomes

$$K(E + \Delta^+_i E) + (M + \Delta^+_i M)(E + \Delta^+_i E) = \frac{\partial J}{\partial t}. \quad (5.3)$$

Similarly, we perturb the matrix $M$ on the forward direction of $p_j$ by another incremental amount $\Delta p_j$ (see Figure 5.1(b)). We then have

$$K(E + \Delta^+_j E) + (M + \Delta^+_j M)(E + \Delta^+_j E) = \frac{\partial J}{\partial t}. \quad (5.4)$$

Further, we perturb the matrix $M$ simultaneously on the forward directions of $p_i$ and $p_j$ by amounts $\Delta p_i$ and $\Delta p_j$ (see Figure 5.1(c)). As such, we obtain the third perturbed wave equations as

$$K(E + \Delta^{++}_{ij} E) + (M + \Delta^{++}_{ij} M)(E + \Delta^{++}_{ij} E) = \frac{\partial J}{\partial t}. \quad (5.5)$$

Adding (2.7) to (5.5) and then subtracting (5.3) and (5.4), we have

$$K \Delta^{2} E + (M + \Delta^{++}_{ij} M) \Delta^{2} E = -R, \quad (5.6)$$

where

$$\Delta^{2} E = \Delta^{++}_{ij} E - \Delta^{+}_i E - \Delta^{+}_j E \approx \frac{\partial^{2} E}{\partial p_i \partial p_j} \Delta p_i \Delta p_j. \quad (5.7)$$

If the technique of one-to-one field mapping is applied, then (5.6) can be further approximated as

$$K \Delta^{2} E + M \Delta^{2} E \approx -R. \quad (5.8)$$

In (5.6) or (5.8), the expression of the residual vector $R$ is given as
where the field solution $\mathbf{E}$ is computed by simulating the nominal unperturbed structure; the field solution $\mathbf{E} + \Delta_j^+ \mathbf{E}$ is computed by simulating the perturbed structure on the forward direction of $p_i$ (see Figure 5.1(a)); The field solution $\mathbf{E} + \Delta_j^+ \mathbf{E}$ is computed by simulating the perturbed structure on $p_j$ (see Figure 5.1(b)). We also note that all non-zero entries in $\Delta_i^+ \mathbf{M} - \Delta_i^+ \mathbf{M}$, $\Delta_i^+ \mathbf{M} - \Delta_j^+ \mathbf{M}$, and $\Delta_i^+ \mathbf{M}$ are located at the vicinity of each perturbation surface, thus only fields within these volume are stored for sensitivity computation during simulations, see Figure 5.1(e–f).

Consider the case when $p_i$ and $p_j$ refer to a same sensitivity parameter $i = j$. Our treatment is to perturb $\mathbf{M}$ on the backward direction of $p_i$, and we get

$$K(\mathbf{E} + \Delta_i^- \mathbf{E}) + (\mathbf{M} + \Delta_i^- \mathbf{M})(\mathbf{E} + \Delta_i^- \mathbf{E}) = \frac{\partial J}{\partial t}. \quad (5.10)$$

Here, the superscript ‘−’ on $\Delta_i^-$ denotes differential changes due to a backward perturbation on $p_i$.

Adding (5.10) to (5.3) and then subtracting (2.7) twice, we obtain the following expression for the residual vector $\mathbf{R}$

$$\mathbf{R} = \Delta_i^+ \mathbf{M} (\mathbf{E} + \Delta_i^+ \mathbf{E}) + \Delta_i^- \mathbf{M} (\mathbf{E} + \Delta_i^- \mathbf{E}), \quad (5.11)$$

and in this case, the expression of $\Delta^2 \mathbf{E}$ in (5.8) is given by

$$\Delta^2 \mathbf{E} = \Delta_i^+ \mathbf{E} + \Delta_i^- \mathbf{E} \approx \frac{\partial^2 \mathbf{E}}{\partial p_i^2} \Delta p_i^2. \quad (5.12)$$

We note that the expression of the residual vector $\mathbf{R}$ given in (5.11) involves another
perturbed field solution \( \mathbf{E} + \Delta \mathbf{E} \) resulting from perturbing \( p_i \) along its backward direction. Instead of solving this new solution through another simulation, one can alternatively obtain an accurate approximation of this field solution by extrapolating the perturbed field solution \( \mathbf{E} + \Delta \mathbf{E} \) along the direction of the nominal field solution \( \mathbf{E} \) in means of the one-to-one field mapping procedure.

Following a similar adjoint procedure to that discussed in the Chapter 2, we multiply both sides of (5.8) by an yet-to-be determined adjoint vector \( \hat{\mathbf{E}} \). Integrating over the complete simulation time, we get

\[
\int_0^{T_{\text{max}}} \hat{\mathbf{E}}^T \left( \mathbf{K} \Delta^2 \mathbf{E} + \mathbf{M} \Delta^2 \mathbf{E} \right) = - \int_0^{T_{\text{max}}} \hat{\mathbf{E}}^T \mathbf{R} \, dt. \tag{5.13}
\]

With the same treatment as from (4.23) to (4.26), we can rewrite (5.13) as

\[
\int_0^{T_{\text{max}}} \left( \mathbf{K} \hat{\mathbf{E}} + \hat{\mathbf{M}} \hat{\mathbf{E}} \right)^T \Delta^2 \mathbf{E} = - \int_0^{T_{\text{max}}} \hat{\mathbf{E}}^T \mathbf{R} \, dt, \tag{5.14}
\]

provided that the terminal conditions given in (2.40) simultaneously hold and that the expression of \( \hat{\mathbf{M}} \) is given in (4.29).

Comparing left-hand side of (5.14) with the last integral term on the right-hand side of (5.1), we can thus write the following second-order adjoint sensitivity expression

\[
\frac{\partial^2 F}{\partial p_i \partial p_j} = \frac{\partial^{(e)}^2 F}{\partial p_i \partial p_j} + \int_0^{T_{\text{max}}} \left( \frac{\partial \mathbf{E}}{\partial p_j} \right)^T \frac{\partial^2 \phi}{\partial \mathbf{E}^2} \frac{\partial \mathbf{E}}{\partial p_i} \, dt - \frac{1}{\Delta p_i \Delta p_j} \int_0^{T_{\text{max}}} \hat{\mathbf{E}}^T \mathbf{R} \, dt, \tag{5.15}
\]

provided that the adjoint vector \( \hat{\mathbf{E}} \) is computed from the equation (2.45). This is also the same adjoint equation used for computing first-order adjoint sensitivities.
5.3 Algorithm implementation and numerical examples

The second-order adjoint variable method discussed in this chapter, in general, requires \( N + 1 \) extra simulations for the computation of \( N^2 \) sensitivity entries in the Hessian matrix. It can be efficiently implemented for any grid-based time-domain electromagnetic solvers, provided that the field solutions at the vicinity of each perturbation surface can be exported from required simulations. Here, we summarize the entire sensitivity procedure into the following four steps:

**Step 1. Parametrization**

For each sensitivity parameter, say \( p_i \), we first identify its volumes of perturbations \( \Psi^\pm \) it would produce as a result of on-grid perturbations. Then, we group and label each cell within \( \Psi^\pm_{p_i} \).

**Step 2. Direct analysis**

Perform \( N \) independent time-domain simulations on perturbed structures along each perturbation direction of \( \mathbf{p} \). For each sensitivity parameter, say \( p_i \), we export the time-domain field solution computed at cells within \( \Psi^\pm_{p_i} \). Using (5.9) or (5.11), we compute all non-zero entries in the residual vector \( \mathbf{R} \).

**Step 3. Adjoint analysis**
Re-simulate the nominal structure by imposing the adjoint current source \( \dot{\mathbf{J}} \) in a backward time-sequence. For each sensitivity parameter, say \( p_i \), we then export the time-domain field solution computed at cells within \( \Psi_{p_i}^\pm \) from the simulation.

**Step 4. Post-processing**

Using (5.15), each sensitivity entry in the Hessian matrix can be computed on-the-fly using the non-zero entries in the residual vector \( \mathbf{R} \) computed from step 2 and the their corresponding entries in the adjoint vector \( \dot{\mathbf{E}} \) computed from step 3.

The theory presented in this chapter is validated through four numerical examples involving both dispersive and non-dispersive discontinuities. For the first three examples, our in-house Yee-based finite-difference time-domain (FDTD) solver is used to compute all necessary field solutions required for computing adjoint sensitivities. For the last example, the required field solutions are computed using openEMS, a free and open-source equivalent-circuit (EC) FDTD software package [11].

5.3.1 **A multilayer anti-reflection dielectric coating**

Multilayer dielectric structures have been widely used in the design of bandpass filters, dielectric mirrors, optical interference filters, and broadband terminations for many forms of transmission lines [12]. Here, we consider a 6-layer broadband anti-re-
**Figure 5.2** A diagram of a 6-layer dielectric anti-reflection structure whose outer regions have permittivity of $\epsilon_a = 12.25$ and $\epsilon_b = 1.0$ (free space). The permittivity of each dielectric slab is given by $\epsilon_1 = 13.5$, $\epsilon_2 = 8.41$, $\epsilon_3 = 4.86$, $\epsilon_4 = 2.52$, $\epsilon_5 = 1.46$, and $\epsilon_6 = 1.08$. Their nominal widths satisfy $w_i = z_i - z_{i-1} = \pi \epsilon_0 / (2 \sqrt{\epsilon_i \omega_0})$ for $i = 1, 2, ..., 6$, where $\omega_0$ is the center frequency of the passband.
(a) (b)

Figure 5.3 The second-order sensitivity of the $S_{11}$ parameter with respect to permittivities $\epsilon_2, \epsilon_3$ (a) and permittivities $\epsilon_3, \epsilon_6$ (b) at the nominal state of the AR filter.

Using the adjoint approach discussed in this chapter, all 78 independent second-order sensitivities are efficiently computed through only 12 extra simulations. Notice that the Hessian matrix is symmetric thus reducing the number of independent sensitivity entries. A subset of our result is shown in Figure 5.3 and Figure 5.4. Comparison shows that the adjoint sensitivities match with the exact sensitivity results computed through the analytical approach [12] for all parameters. If the same sensitivity procedure is performed using the CFD approximation through (5.2), a total of 288 extra simulations would be required.
Figure 5.4 The second-order sensitivity of the $S_{11}$ parameter with respect to positions $z_1$ and $z_6$ (a), positions $z_3$ and $z_4$ (b), the position $z_4$ (c), positions $z_4$ and $z_6$ (d) at the nominal state of the AR filter.
5.3.2 A dispersive discontinuity

Many electromagnetic problems deal with artificial or biological materials whose constitutive parameters are frequency-dependent, see for example [14, 15]. In FDTD simulations, these types of materials are usually treated through the auxiliary differential equation (ADE) method [16] or the Z-transform method [17]. Here, we consider a simplified example involving such a material to illustrate our adjoint sensitivity technique. In our example, a dispersive dielectric discontinuity is placed inside a rectangular waveguide (see Figure 5.5). The dielectric discontinuity is modeled using a Drude dispersive pole whose parameters are shown in the inset of Figure 5.5.

To accurately model this structure in the time domain, a uniform mesh size of
$\frac{\partial^2 |S_{11}|}{\partial p_1 \partial p_2} \text{ (mm}^{-2}\text{)}$

Figure 5.6  The second-order sensitivity of the $S_{11}$ parameter with respect to edges $p_1$ and $p_2$ (a), edges $p_1$ and $p_3$ (b) at the nominal state of the lossy discontinuity.

$\Delta h = 75.0\mu m$ and an FDTD time step $\Delta t = \Delta h/(2c_0)$ are utilized for the simulation through our two-dimensional FDTD solver. The ADE-FDTD technique [16] is employed to model the dispersive behavior of the dielectric discontinuity. The whole structure is excited by a wideband Gaussian pulse with a $TE_{10}$ mode profile across the input port. The source signal covers frequencies from 40 GHz to 80 GHz. For the sensitivity analysis, 4 shape parameters are considered. These parameters are the four bounding edges of the dispersive dielectric discontinuity, see Figure 5.5. Using only 4 extra simulations, the adjoint technique computes all 10 independent second-order sensitivities. A subset of our computed results are shown in Figure 5.6 and Figure 5.7.
Figure 5.7  The second-order sensitivity of the $S_{11}$ parameter with respect to edge $p_1$ (a), edges $p_2$ and $p_3$ (b), edges $p_2$ and $p_4$ (c), the edge $p_4$ (d) at the nominal state of the lossy discontinuity.
Our results show reasonable agreement for all parameters with those obtained using the time-intensive CFD approximations which requires 32 extra simulations.

5.3.3 A dielectric resonator filter

In this example, our adjoint sensitivity approach is applied to an ultra-sensitive dielectric resonator inside a rectangular waveguide [18, 19]. Due to its simple tuning mechanism, this type of filters is often employed in base-station transceivers [18]. Here, our considered dielectric resonator filter contains three ceramic blocks ($\epsilon_r = 38.5$) mounted on one edge of a rectangular waveguide (see Figure 5.8). The whole structure is simulated using a two-dimensional FDTD solver with a uniform mesh size $\Delta h = 0.29$ mm and an FDTD time step $\Delta t = \Delta h/(2c_0)$. The excitation plane is placed $200\Delta h$ away form the first dielectric discontinuity. The source signal follows...
the dominant $TE_{10}$ spatial profile with a central frequency of 7 GHz and a bandwidth of 2 GHz. We consider 9 sensitivity parameters including the permittivity $\epsilon_r$, the width $w$, and the length $l$ for each of the three ceramic post. The second-order sensitivity results of the return loss (the $S_{11}$ parameter) are computed using 9 extra simulations. A subset of those results is shown in Figure 5.9 and Figure 5.10. All results show good match with those obtained through the CFD approximations which require a total of 162 extra simulations.

Figure 5.9 The second-order sensitivity of the $S_{11}$ parameter with respect to the permittivity $\epsilon_b$ (a), permittivities $\epsilon_a$ and $\epsilon_b$ (b) at the nominal state of the dielectric resonator filter.
Figure 5.10  The second-order sensitivity of the $S_{11}$ parameter with respect to permittivities $\epsilon_a$ and $\epsilon_c$ (a), dimensions $w_a$ and $w_b$ (b), the dimension $w_b$ (c), dimensions $l_a$ and $l_c$ (d) at the nominal state of the dielectric resonator filter.
5.3.4 A dielectric resonator antenna with matching insert

Our last numerical example is a three-dimensional dielectric resonator antenna (DRA) [20, 21] (see Figure 5.11). In this antenna configuration, the dielectric radiator is fed through a microstrip line printed on top of a substrate with relative permittivity 3.0 and a thickness of 0.762 mm. The width of the microstrip line is 1.9 mm. To achieve a wideband coupling between the stripline and the dielectric radiator, a thin layer of high-permittivity insert is placed between the stripline and the dielectric radiator [21]. To obtain the antenna response over a wide band of frequencies, we simulate this antenna in the time domain using the openEMS software package [11]. The excitation plane is placed 16.0 mm away from the radiator and its source signal is a Gaussian-modulated sinusoid centered at 15 GHz with bandwidth of 10 GHz. The smallest mesh size that we use to model the antenna is $\Delta h = 0.15$ mm. The FDTD
time step for the simulation is $\Delta t = \Delta h/(2c_0)$. To reduce the amount of computational effort, only half of the antenna structure shown in Figure 5.11 is simulated through the utilization of a symmetry plane.

In the sensitivity analysis, a total of 6 sensitivity parameters are considered. These include the width $w_1$, the height $h_1$, and the length $d_1$ of the matching insert, as well as the width $w_2$, the height $h_2$, and the length $d_2$ of the dielectric resonator. The nominal values of these considered parameters are given in the insert of Figure 5.11. With the introduced adjoint sensitivity approach, all 21 independent sensitivity entries in the response Hessian matrix are estimated through only 6 extra simulations. A subset

**Figure 5.12** The second-order sensitivity of the $S_{11}$ parameter with respect to dimensions $d_1$ and $h_2$ (a), dimensions $d_1$ and $w_2$ (b) at the nominal state of the dielectric resonator antenna.
The second-order sensitivity of the $S_{11}$ parameter with respect to dimensions $h_1$ and $w_2$ (a), dimensions $h_1$ and $d_2$ (b), dimensions $w_1$ and $d_2$ (c), dimensions $w_1$ and $h_2$ (d) at the nominal state of the dielectric resonator antenna.

Figure 5.13
of our obtained results are shown in Figure 5.12, Figure 5.13, and Figure ??_. All results (including the ones not shown) show good match when compared with the CFD approximation which requires 72 extra simulations.

5.4 Summary

This chapter discussed a general and efficient approach for estimating the second-order sensitivities of electromagnetic devices from their simulated transient field solutions. With the adjoint approach, the number of simulations required for performing sensitivity analysis is reduced from $O(N^2)$ to $O(N)$ when compared with the finite-difference techniques. The accuracy of the computed adjoint sensitivities has been verified through a number of numerical examples. The computed sensitivity solutions could greatly benefit a number of electromagnetic problems involving optimization, device modeling, tolerance, and yield analysis.

References


Chapter 6

Conclusions

This thesis summarized the complete theory of the adjoint variable method developed for sensitivity estimations through time-domain electromagnetic simulations. Comparing with conventional methods, our considered adjoint-based sensitivity techniques, in general, reduce the amount of required computational effort by at least a factor of $N$, if $N$ is the total number of involved sensitivity parameters. The developed numerical algorithms are readily applicable to any grid-based time-domain electromagnetic simulators, and can be directly applied to the design and optimization of microwave, terahertz, photonic, and plasmonic devices.
In Chapter 2, we discussed the type of the wave equations that we are considering in our adjoint sensitivity analysis. Some preliminary knowledge in understanding the adjoint variable method are covered. This chapter also considered the simplest form of the adjoint variable method, in which all involved sensitivity parameters are associated with perfect dielectrics. The produced adjoint results accurately match those obtained through the finite-difference approximations.

In Chapter 3, we considered the treatment for frequency-dependent response functions in the adjoint sensitivity analysis. With the proposed self-adjoint sensitivity technique, the complex-valued sensitivities of scattering parameters can be computed over a wide band of frequencies without additional adjoint simulations. The introduced sensitivity procedure is illustrated through two numerical electromagnetic examples. The produced adjoint results agree with those obtained through either the analytical approaches or through the central finite-difference approximations.

In Chapter 4, we generalized the adjoint variable method to handle frequency-dependent sensitivity parameters for linear dispersive materials. The constitutive relations in our considered materials may contain Debye, Drude, Lorentz dispersion poles, or a linear combination of them. Through our mathematical derivation, we have demonstrated that the same adjoint sensitivity procedures developed in Chapter 2 and Chapter 3 can be utilized to treat sensitivity parameters associated with dispersive materials without contributing additional computational overhead. The accuracy of the produced adjoint results has been validated through a number of numerical examples using the auxiliary differential equation (ADE) enabled finite-differ-
In Chapter 5, we extended the existing adjoint variable method to supply an accurate estimation of response Hessian matrix in a same degree of efficiency. With this second-order adjoint approach, the number of required simulations for estimating the response Hessian is reduced from $O(N^2)$ to $O(N)$ when compared with the finite-difference techniques. The accuracy of the computed adjoint sensitivities has been verified through a number of numerical examples. The superior efficiency of this second-order adjoint variable method opens up a new field of device optimization though the utilization of the Newton’s method.

From the experience gained during this research, the author of this thesis suggests the following research directions to be addressed in future developments:

1. extend the adjoint variable method to accommodate cylindrical and spherical grids;

2. extend the adjoint variable method to handle anisotropic and nonlinear sensitivity parameters associated with semiconductors or metamaterials;

3. extend the adjoint variable method to handle some critical nano-plasmonic applications, such as the nano-scaled radiation-matter interaction, nano-pillars, and nano-rodes;
(4) utilize the second-order adjoint variable method to accelerate the design and optimization process for electromagnetic devices.
Bibliography


