University of Alberta

Nonlinear Metal-Insulator-Metal (MIM) Nanoplasmonic Waveguides
Based on Electron Tunneling for
Optical Rectification and Frequency Generation

by

Xiaoqin Lei

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Photonics and Plasmas

Department of Electrical and Computer Engineering

Spring 2013
Edmonton, Alberta

To my family and friends

Abstract

Metal-Insulator-Metal (MIM) electron tunneling diodes have recently emerged as an attractive alternative to semiconductor photodiodes for THz and optical detection due to their fast response time and relative ease of fabrication. However, current antenna-coupled MIM diode detectors are still limited by poor responsivity and low detection bandwidth due to impedance mismatch between the diode and antenna, large RC time constant of lumped MIM junctions, and narrow bandwidth of traditional antenna designs. In this thesis we address these issues by considering traveling-wave MIM detector designs which exhibit enhanced responsivity and low impedance that can be more easily matched to planar antennas. We also propose new antenna geometries based on surface modification of traditional bowtie antennas that are capable of receiving ultra-wideband THz signals. The concept of traveling-wave MIM detectors is then extended to the investigation of nonlinear MIM nanoplasmonic waveguides for on-chip single-cycle THz pulse generation, frequency conversion, as well as plasmonic switching and modulation.

Acknowledgements

This thesis would not be possible without the support of many people. I would like to express my sincerest thanks to all of you.

First and foremost, I would like to thank my supervisor Dr. Vien Van, for his tireless supervision and unwavering support throughout the course of this thesis. The countless time that Dr. Van spent on discussing problems and making suggestions help to finalize the contents of this thesis. In the past two years, he taught me invaluable skills such as deliberate way of thinking, communicating ideas and presenting results logically. His insight, enthusiasm and passion to scientific research have great impacts on my study and research.

Secondly, I would like thank Dr. Rambabu Karumudi, for his helpful discussions and encouragement in my graduate study.

Many thanks go to my peers in the Nanophotonic Research Lab Ashok Prabhu Masilamani, Alan Tsay, Marcelo Wu, Siamak Abdhollahi, Guangcan Mi, Cameron Horvath, and Daniel Bachman. I appreciate your helpful discussions and supports in my study. You are all valuable friends and help me a lot in many ways. My special gratitude goes to Pooya Taheri, for your encouragements in my study.

Life in Edmonton is warmed up by many valuable friends. I enjoy the joyful activities with all of you. Thank you for the support from members in the Taishan Group. I will never forget your kindness and the wonderful moments that we have spent together.

Lastly, I would like to thank my family and especially my dearest Molly since she will never leave me or forsake me. Big thanks go to Cherry, my dearest friend, for your great supports.

Table of Contents

List of Figures

List of Tables

List of symbols

List of Abbreviations

ABC	Absorbing boundary condition	68
ADE	Auxiliary differential equation	67
ALD	Atomic layer deposition	1
CMOS	Complementary metal-oxide-semiconductor	8
CMT	Coupled mode theory	7
CW	Continuous wave	1
CFL	Courant Friedrichs Lewy	69
EBL	Electron beam lithography	4
FDTD	Finite-difference time-domain	8
FEM	Finite element method	7
FWHM	Full width half magnitude	56
IBD	Ion-beam deposition	4
IMI	Insulator metal insulator	93
IMIMI	Insulator metal insulator metal insulator	8
I-V	Current-voltage	10
MBE	Molecular beam epitaxy	1
MIM	Metal insulator metal	1
MIMIM	Metal insulator metal insulator metal	1
MIS	Metal insulator semiconductor	8
MOM[1]	Metal oxide metal	11
MOM[2]	Method of moments	64
MOCVD	Metal-organic chemical vapour deposition	1
PML	Perfect match layer	68
SPP	Surface plasmon polariton	2
THz	Terahertz	1
TM	Transverse magnetic	8
TWPDs	Traveling wave photodiodes	22
TW-MIM	Traveling-wave metal-insulator-metal	6

UV	Ultra violet	91
WKB	Wentzel-Kramers-Brillouin	13

Chapter 1

Introduction

Electromagnetic radiation detectors are important components for scientific research and applications in the broad electromagnetic spectrum from microwave to optical frequencies. Over the last several decades, integrated optical detectors based on semiconductor diodes have attracted a lot of attention and experienced rapid development fuelled largely by the advance of the semiconductor technology. Semiconductor detectors are typically characterized by very high responsivities; however, their wavelength range of operation is limited by the material bandgap and the speed of operation is hindered by the electron transit time across the pin junction.

Metal-Insulator-Metal (MIM) diode detectors have recently emerged as an attractive alternative to semiconductor detectors. Initially demonstrated in 1966 [1], there has been renewed interest in the development of these devices for applications as rectifiers from Terahertz (THz) to infrared [2-5] thanks to recent progress in thin film fabrication technologies such as Atomic Layer Deposition (ALD). Compared to semiconductor detectors, MIM diodes are considerably easier and cheaper to fabricate since the metal and insulator layers can be deposited using ion-beam or electron beam deposition rather than more complex and costly epitaxial processes such as molecular beam epitaxy (MBE) and metal-organic chemical vapour deposition (MOCVD). In addition, due to the ultrafast electron tunneling process (in the order of 10^{-16}-10^{-15}s [6]), MIM diode detectors exhibit broadband operations for continuous wave (CW) signal detection from THz to infrared radiation [7]. However, to meet the requirements for efficient and fast detectors for ultra-wideband applications, several technical challenges relating to signal coupling efficiency and RC bandwidth limitation must be addressed. This thesis focuses on the modeling, optimization and improvement of MIM detectors for ultrafast THz and optical rectification.

At optical frequencies the MIM structure can also support and guide surface plasmon polaritons (SPPs). SPPs are electromagnetically-coupled electron oscillations that exist and propagate along a metal-dielectric interface [8]. Surface plasmon waveguides have attracted much attention recently because unlike conventional dielectric optical waveguides, these waveguides can confine and guide electromagnetic energy below the diffraction limit of light. In particular, the MIM structure can confine and guide light in the nanometer-thin insulator gap between the two metal layers. Combining with the nonlinear electron tunneling process, we show that the MIM waveguide acts as a nonlinear nanoplasmonic waveguide which can be used for optical rectification, second harmonic frequency generation as well as other nonlinear plasmonic device applications.

1.1 Background and Motivation

MIM diodes, typically consisting of a 1-3nm thick oxide layer sandwiched between two metal electrodes, exhibit nonlinear current-voltage relation under an applied potential [9-11]. This nonlinear response arises from electron tunneling through the insulator layer under a bias voltage and can be used to detect electromagnetic radiation [1]. Electromagnetic waves incident on a MIM diode will be rectified, resulting in a DC current which is proportional to the power of the incident radiation. MIM diodes are especially attractive for detecting THz radiation, which has potential applications in security, medical imaging and many other areas [12-13]. Due to the small size of an MIM structure compared to the wavelength of interest, the fraction of incident radiation that could be received by the diode is very small, so antennas are commonly used to couple electromagnetic radiation into the diode more efficiently [4, 14].

Optimal design of antenna-coupled MIM diode detectors must address two competing issues: 1) detection efficiency (or external quantum efficiency) and 2) detection bandwidth. Current MIM diodes exhibit poor detection efficiency due to the inherently low rectification efficiency of the MIM junction and the poor coupling efficiency from the antenna to the diode. The poor coupling efficiency is

primarily caused by the impedance mismatch between the antenna and the diode. Commonly used planar antennas have typical impedance values around 50Ω [15] to 100Ω [4], which are much smaller than the typical diode resistance in the KΩ-MΩ [16-17]. As a result, there is a large impedance mismatch between the antenna and the diode, resulting in poor coupling efficiency even though the received signal is enhanced by the antenna. In addition to the impedance mismatch problem, the speed of an antenna-coupled MIM diode detector is also limited by its RC time constant, as given by the antenna resistance and the diode capacitance. This time constant sets a cut-off frequency beyond which rectification cannot be achieved efficiently. For an Al-Al_2O_3-Al MIM diode with antenna impedance of 50Ω, diode size of 150 × 150nm^2 and insulator layer thickness of 3nm; the RC bandwidth is around 11THz, which prevents its use at the optical frequencies. To increase the cut-off frequency, the size of the diode has to be further reduced which presents considerable fabrication challenges. On the other hand, to increase the diode efficiency, the area of the diode must be increased, which reduces the RC bandwidth and increases the impedance mismatch with the antenna. Thus, simultaneous wide detection bandwidth and high detection efficiency are not achievable for current antenna-coupled MIM diode detectors.

The above discussion shows that, it is a challenging task to design MIM diodes with both high responsivity and wide bandwidth simultaneously. Therefore, there is strong motivation in the research of an antenna-coupled MIM-configured detector which can make use of the advantages of MIM systems while achieving wide detection bandwidth and high detection efficiency for ultra-wideband optical applications, especially for the detection of ultrafast optical pulses.

1.2 Literature Review on Antenna-coupled MIM Diodes

The first antenna-coupled MIM diode was fabricated in 1966 [1] and consisted of a point contact junction coupled to a dipole antenna. Since then much research work has been conducted on antenna-coupled MIM diodes with the aims of

improving the detector responsivity, coupling efficiency of radiation into the diode and the detection bandwidth and response time of the diode.

1.2.1 MIM diodes

Research on MIM diodes dates back to the 1960s starting with the point-contact configuration consisting of a tungsten wire attached to a planar metal surface [18]. The MIM junction was formed at the contact point with a few nanometers of native oxide as the insulator layer. Point contact diodes have a very fast response time estimated in the order of 10^{-16}-10^{-15}s [6] and are capable of detecting electromagnetic waves up to the infrared frequencies [3,7,18] by means of rectification. However, the point contact configuration is mechanically unstable and cannot be reproduced since it is not easy to control the contact area and native oxide thickness. Such a structure is therefore not suitable for commercial applications. With later development of ion-beam deposition (IBD) and electron beam lithography (EBL), the fabrication of more stable and reproducible thin-film planar MIM diodes becomes possible. In this configuration, the insulator layer is sandwiched between the two metal electrodes in a layered hierarchy. It can also be conveniently integrated onto substrates for chip-level applications as detectors in imaging systems. As a result, planar MIM diodes have become popular. The common geometry takes a square shape with a width of 100nm or less [3-4,7] to achieve fast response time and high cut-off frequency.

Experimentally, numerous material systems have been used in the fabrication of MIM diodes. These include symmetrical structures such as Ni-NiO-Ni [3,4], Al-Al_2O_3-Al [5] and asymmetrical structures such as Al-AlO_x-Pt [19]. These diodes have been demonstrated for detecting CW signals up to 10μm wavelength [3,4,20] with responsivity as high as 0.95mA/W [4,21]. Operation of a symmetrical MIM diode requires a bias voltage to achieve rectification while an asymmetrical structure can rectify without bias. Recent research in MIM diodes even suggests the use of amorphous metals (also known as metallic glass) as electrodes, exhibiting highly nonlinear current-voltage characteristics [22].

Regardless of the material system used, the primary objective in MIM diode research is to obtain highly nonlinear current-voltage response which permits detectors with high responsivity and ultrafast response time to be achieved.

1.2.2 Antennas

Typical MIM diodes have dimensions of $100 \times 100 nm^2$ which are much smaller than the wavelengths of interest. Due to the small areas of MIM diodes, coupling of radiation to the diode is very poor. To enhance the coupling, antennas are commonly connected to diodes as receiving units. Both the MIM diode structure and the planar antenna are integrated on the same substrate, with the two arms of the antenna each connected to a metal layer of the diode. Incoming radiation is captured by the antenna and subsequently transferred to the diode for rectification. With a much larger area, the antenna can greatly enhance the overall responsivity of the diode.

Depending on the polarization of the received radiation, there are two major types of antennas used: linearly polarized and circularly polarized. Early MIM diodes used linearly polarized dipole antennas since they are easy to fabricate and analyze [1,20]. Bowtie antennas are another types of linearly polarized antennas that are commonly used [4] because they have relatively wideband performance and can be easily scaled to operate at different frequencies [4,23]. To receive circularly polarized radiation, spiral antennas are the most commonly used [4]. The three types of antennas mentioned above have been well characterized at the microwave frequencies and can be easily extended into the THz and infrared regimes. Experimental studies of these antennas at the THz frequencies have shown excellent performances [4].

In addition to the above traditional antenna geometries, novel antenna designs such as sierpinski and apollonian fractal antennas [24-25] can be used to achieve ultra-wideband impedance matching to the diode. Geometrical modification techniques can also be applied to traditional antennas following basic principles in antenna design at the microwave frequencies to achieve higher

efficiency and wider bandwidth performance. In general, optimum design of diode antenna aims at providing the highest available signal power to the diode over a widest available bandwidth with a reasonably low return loss.

1.2.3 Traveling-wave diodes

The use of antennas greatly increases the received radiation power. However, not all of the available power can be coupled to the MIM diode because of mismatch between the diode and antenna impedances. The problem of impedance mismatch is tied to the device material and geometry and has no easy solution. On one hand, planar antennas made with highly conductive metals cannot have resistances as high as KΩ-MΩ [16-17]. On the other hand, commonly used fabrication techniques cannot bring the diode impedance down to the antenna's 50Ω range. There has been some research effort on lowering the diode impedance by using different oxidation methods for oxide deposition but no significant improvement has been achieved [16].

One way to overcome the impedance mismatch problem and achieve wide coupling bandwidth is by using a traveling-wave MIM (TW-MIM) diode structure [26]. In a TW-MIM diode, electromagnetic waves are rectified as they propagate along the length of the MIM structure, which serves as a type of waveguide. Since electromagnetic waves propagate in the MIM waveguide as the fundamental transverse magnetic (TM) mode, they experience only the modal characteristic impedance [26]. This characteristic impedance can be designed to be in the range of 50Ω to achieve perfect impedance matching to a typical planar antenna [26]. Also, the fundamental TM mode of the MIM waveguide has no cut-off frequency, so that the waveguide can support wave propagation from microwave up to the optical frequencies [26-29]. As a result, the TW-MIM diode structure is expected to have ultrawide electrical and optical bandwidths. Although TW-MIM diodes have been theoretically analyzed for CW signal detection [26], their performance for rectifying ultrafast optical pulses has not been investigated.

1.3 Research Objectives and Contributions

This thesis has two primary objectives. The first one is to realize improved antenna designs for receiving/radiating wideband THz radiation. The second one focuses on theoretical and numerical investigations of the MIM structure as a nonlinear nanoplasmonic waveguide for ultrafast optical applications such as rectification, second harmonic generation and frequency mixing.

This thesis consists of two main parts. The first part is motivated by the need to improve the performance of antenna-coupled MIM detectors for THz signal detection. In this part we will first provide a detailed analysis of current antenna-coupled MIM diode designs to identify their limitations. We will next focus our effort on the design of improved antenna geometries to achieve wide-bandwidth impedance matching to the MIM junction. This is accomplished by geometrical modification of traditional antennas, making use of antenna theory as well as numerical tools to aid the design optimization process. In particular, the multi-resonance technique is used to improve upon the traditional bowtie antenna design and the finite element method (FEM) [30] is used to simulate the response of the proposed antenna designs. The performance of our antenna designs in the sub-THz (1THz-2.5THz) band will be compared with those of traditional bowtie antennas.

The second part of the thesis will focus on the MIM structure as a nonlinear nanoplasmonic waveguide. Our primary effort will be devoted to the theoretical and numerical analyses of MIM plasmonic waveguides for rectifying ultrafast optical pulses and second harmonic frequency generation. Theoretically, we will model the MIM plasmonic waveguide as a one-dimensional effective medium with effective permittivity and conductivity and treat the electron tunneling in the MIM junction as an effective nonlinear conductivity. We will also develop a set of nonlinear coupled equations using the coupled-mode theory (CMT) for describing the nonlinear propagation of both CW and pulse signals. These sets of equations will also be able to approximate the amplitude behaviours of CW and pulse signals as well as the bandwidth performance of the pulse

signals. To simulate the structure we will develop a nonlinear finite-difference time-domain (FDTD) method which incorporates electron tunneling in the MIM junction as an effective conductivity in the insulator layer. Both the CMT analysis and numerical simulations will be done to investigate the characteristics of the MIM nanoplasmonic waveguide for optical rectification, frequency generation, and other nonlinear plasmonic applications. In particular, we find that the device can be used to generate single-cycle THz pulses by rectifying ultrafast optical pulses, for potential application as on-chip THz source generation. The nonlinear MIM waveguide can also be used as a plasmonic device to perform nonlinear signal processing, such as second harmonic generation, sum and difference frequency generation, as well as plasmonic logic functions.

We also extend the scope of our investigation by exploring alternative plasmonic waveguide designs, material systems and other electron tunneling junctions for realizing nonlinear plasmonic devices. Firstly, we look at the MIM waveguide providing third-order nonlinearity and find that its effects are much weaker than the second-order nonlinearity. In particular we investigate the use of long range surface plasmon modes in Insulator-Metal-Insulator-Metal-Insulator (IMIMI) waveguides for achieving improved nonlinear frequency conversion efficiency. We also study the metal-insulator-semiconductor (MIS) plasmonic waveguide as an alternative structure to the MIM structure. These structures are found to be especially promising for nonlinear plasmonic device applications due to their lower loss, high nonlinearity and compatibility with the CMOS fabrication process.

1.4 Thesis Outline

This thesis contains six chapters and is organized as follows.

Chapter 1 provides the motivation, literature background and outlines our research objective and contributions.

Chapter 2 introduces the design and analysis of current antenna-coupled MIM diode detectors, investigates their limitations and proposes improved

antenna designs for wideband applications in the sub-THz range.

Chapter 3 presents a theoretical analysis of MIM nanoplasmonic waveguides based on the Coupled Mode Theory for wave propagation in nonlinear conductive medium. The performance of the MIM structure for ultrafast optical rectification and second harmonic generation will be investigated.

Chapter 4 presents the development of a rigorous nonlinear FDTD method for simulating MIM nanoplasmonic waveguide structures. The method is used to study nonlinear pulse propagation in the MIM waveguide and the various second order nonlinear processes in the waveguide. The numerical simulations are also compared to approximate theoretical results to validate the CMT theory developed in Chapter 3.

Chapter 5 explores additional MIM plasmonic waveguide designs, material systems, and other electron tunneling junctions for realizing nonlinear nanoplasmonic applications.

Chapter 6 summarizes the contributions of this thesis and provides suggestions for future work.

Chapter 2

Analysis of Antenna-coupled MIM Diode Detectors and Improved Antenna Design

2.1 Introduction

Since their first demonstration in 1966 [1], antenna-coupled MIM diodes have been extensively researched for use as fast rectifiers and mixers of electromagnetic signals from the microwave to infrared frequencies [2-5]. The operations of these devices rely on the nonlinear current-voltage (I-V) characteristic of MIM junctions. The advantages of MIM diodes over conventional semiconductor photodiodes include their ease of fabrication, wideband operation and fast response time. Unlike semiconductor photodiodes which require expensive fabrication processes such as MBE growth, MIM diodes can be fabricated by simple low-cost chemical vapour deposition or sputtering techniques. Since MIM diodes operate based on electron tunneling, their response times are much faster than semiconductor photodiodes, whose response times are typically limited by electron recombination and transit time through the pin junction [31].

Given their advantages there has been much effort in the development of MIM diodes for various applications [2-5,32]. In particular, there has recently been increasing interest in employing antenna-coupled MIM diode detectors for THz imaging [12-13]. THz radiation (in the frequency range from 0.5THz to 2.5THz) is believed to have good penetration in many materials, especially human tissues [12-13,33]. The absorption spectra in the THz frequency provide unique fingerprints that can be used to detect and identify various chemicals of interest in medical diagnostics and security applications. THz imaging technology relies on sensitive and broadband THz detectors and it is in this arena that antenna-coupled MIM diode rectifiers can potentially provide an attractive solution in terms of

device performance and low cost. While there has been development in improving the characteristics of MIM junctions, there has not been much effort devoted to the engineering of new antenna designs optimized for broadband THz reception. Most currently used THz antennas are miniaturized versions of their traditional microwave counterparts [4,14,17], including a few popular simple designs such as dipole, bowtie and spiral antennas [4]. To make use of antenna-coupled MIM diodes as broadband THz detectors, the antennas are required to have broad receiving bandwidth in the THz range. Traditional microwave antenna designs cannot accommodate such a wide bandwidth. To address this problem, we will devote a portion of this chapter to the engineering of new THz antenna designs for the reception of broadband THz radiation.

The objective of this chapter is to investigate the characteristics of antenna-coupled MIM diode detectors, identify their constraints and propose improved antenna designs for receiving broadband THz signals. The goal of the THz antenna design is to achieve impedance matching to a load impedance of 50Ω with less than -10dB return loss over the THz band from 1THz to 2.5THz. The proposed antenna design strategy can also be scaled to higher THz and far infrared frequencies. Other related parameters such as gain profile, efficiency and radiation pattern of the proposed antenna designs will also be discussed.

The outline of the rest of this chapter is as follows. Section 2.2 provides a review of MIM diodes and their equivalent circuit modeling when connected to an antenna. The purpose is to identify the parameters related to the overall performance of antenna-coupled MIM diode detectors. Section 2.3 reviews popular THz antenna designs and proposes improvements to these designs for receiving THz radiation in the frequency range from 1THz to 2.5THz and summaries the current constraints on the performance of antenna-coupled MIM diode detectors. Concluding remarks will be given in Section 2.4.

2.2 MIM Tunnel Diodes

Metal-insulator-metal diodes, also known as metal-oxide-metal (MOM) diodes,

are a type of tunnel junctions consisting of a thin insulator layer sandwiched between two metal electrodes, as shown in Fig. 2.1(a).

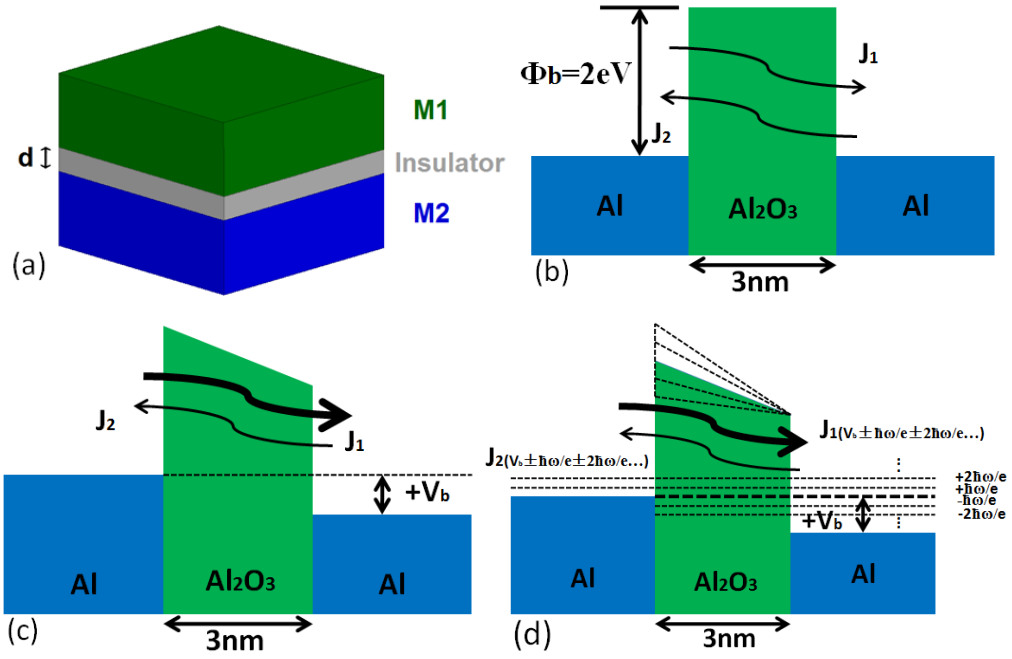


Figure 2.1 (a) Schematic of a planar MIM diode and the band diagrams of an Al-Al_2O_3-Al diode: (b) without bias voltage and a barrier height Φ_b of 2eV [32], (c) with a forward DC bias voltage V_b, and d) with a forward DC bias voltage V_b and effects from energetic photons [34-35]. J_1 and J_2 represent the amplitudes of electron tunneling current densities and the tunneling directions of electrons are indicated by the arrows.

The MIM diodes are symmetrical if the two metals are the same and asymmetrical if they are different. A symmetrical MIM diode has symmetric current-voltage characteristic and requires an applied bias voltage for rectification. An asymmetrical MIM diode does not require a bias voltage since its current-voltage characteristic is already asymmetric. In this thesis, we consider only symmetrical MIM diodes.

In order to allow electron tunneling, the metal and insulator must be chosen to create a potential barrier height separating the two metal electrodes and the thickness of the barrier must also be small to achieve high electron tunneling probability [36]. The barrier height φ between a metal and an insulator is the difference between the work function Φ_m of the metal and the electron affinity χ_{ins} of the insulator, [32]

$$\varphi = \Phi_m - \chi_{ins}. \quad (2.1)$$

Thus, a potential barrier exists when the work function of the metal is higher than the electron affinity of the insulator. When the insulator layer thickness is reduced to only a few nanometers, net electron tunneling can take place between the two metal electrodes under an applied potential across the MIM junction [36]. Field assisted electron tunneling is regarded as the dominating conduction mechanism in MIM diodes over thermionic emission effects [10]. The tunneling current exhibits a nonlinear dependence on the DC bias potential and is primarily determined by the height and width of the potential barrier.

The operation of a symmetrical MIM diode can be understood from its energy band diagram which is shown in Fig. 2.1(b). We use the example of an Al-Al_2O_3-Al junction which is a commonly used and well researched material system for MIM diodes [5,37-39]. The Al_2O_3 insulator layer has a typical thickness of 3nm [40] and a potential barrier height of about 2eV with respect to the Fermi surface of the Al metal layers [32]. Under zero applied bias (Fig. 2.1(b)), electrons have equal probability of tunneling through the barrier in both directions, so there is no net current flow through the diode. Under a DC bias voltage, a net tunneling current arises due to the unequal tunneling probabilities in the two directions induced by the potential difference, as illustrated in Fig. 2.1(c). When the MIM junction is illuminated with high frequency radiation, virtual bias states equal to multiple photon voltage steps are created around the DC bias point [34-35], resulting in modulated bias states as illustrated in Fig. 2.1(d).

Many metals have been used in the construction of MIM diodes. Besides the commonly used Al, Cr and Ni metals [3,4,32], MIM diodes with amorphous metals (metallic glass) have also been demonstrated, showing superior nonlinear current-voltage characteristics [22]. In addition, two or more insulator layers can also be used as barriers to enhance the current-voltage nonlinearity [32,41].

2.2.1 Current-voltage characteristics of MIM diodes

Theoretical analysis of the tunneling currents in MIM junctions was intensively

studied by Simmons [10,11] who developed a classical tunneling current model in MIM diodes under DC bias condition using the Wentzel-Kramers-Brillouin (WKB) approximations. In Simmons' approach, field-assisted electron tunneling is considered to be the major conduction mechanism and minor contributions from photon-assisted tunneling and thermionic effects are neglected. The model also includes the effects of image charges.

As shown in Fig. 2.1(b), when there is no applied bias voltage, electron tunneling probabilities in both directions are equal, so the net tunneling current is zero. When a positive bias voltage is applied to the right metal as shown in Fig. 2.1(c), electron tunneling from the left to the right has a higher probability than in the reverse direction. As a result a net tunneling current equal to J_1–J_2 is observed. The tunneling current densities for small signal detection under a DC bias voltage V_{bias} are given by [10]

$$J_1(V_{bias}) = \frac{J_o}{\Delta s^2}\overline{\varphi_I}\exp(-A\Delta s\sqrt{\overline{\varphi_I}}), \tag{2.2a}$$

$$J_2(V_{bias}) = \frac{J_o}{\Delta s^2}(\overline{\varphi_I} + eV_{bias})\exp[-A\Delta s\sqrt{\overline{\varphi_I} + eV_{bias}}], \tag{2.2b}$$

where

$$J_o = \frac{e}{2\pi h\beta^2}, \tag{2.2c}$$

$$A = \frac{4\pi\beta}{h}m^{\frac{1}{2}}, \tag{2.2d}$$

$$\beta \approx 1, \tag{2.2e}$$

and the net tunneling current is

$$J(V_{bias}) = J_1(V_{bias}) - J_2(V_{bias}) = \frac{J_o}{\Delta s^2}\{\overline{\varphi_I}\exp(-A\Delta s\sqrt{\overline{\varphi_I}}) - (\overline{\varphi_I} + eV_{bias})\exp[-A\Delta s\sqrt{\overline{\varphi_I} + eV_{bias}}]\}. \tag{2.2f}$$

In the above, J is the tunneling current density amplitude in units of A/cm^2, V_{bias} is the bias voltage in units of *Volts*, e is the charge of an electron, Δs is the effective barrier thickness in units of *Å*, h is the Planck's constant, m is the mass of electron, β is the integration correction factor and $\overline{\varphi_I}$ is the effective barrier

height in units of *eV*.

When the MIM junction is illuminated by high frequency radiation from THz and above, the impact of energetic photons on the device behaviour becomes significant [34] and the classical analysis of Simmons must be modified to account for the contribution from photon-assisted tunneling. The absorption of high frequency radiation generates virtual energy levels equal to multiple photon voltage steps ($n\hbar\omega/e$) around the bias point as illustrated in Fig. 2.1(d). These extra energy levels modify the energy states between the top and bottom electrodes and result in step-like bias states with a modulated current-voltage characteristic. As a result the MIM diode no longer behaves as a classical rectifier. In the semi-classical model the tunneling current is modified from the classical DC current to give [35]

$$I(V_{bias}+V_{\omega}) = I_o(V_{bias}+V_{\omega}) + I_{\omega}(V_{bias}+V_{\omega}), \tag{2.3a}$$

where the DC contribution I_o is

$$I_o(V_{bias}+V_{\omega}) = \sum_{n=-\infty}^{\infty} J_n^2(eV_{\omega}/\hbar\omega) I_{dc}(V_{bias}+n\hbar\omega/e), \tag{2.3b}$$

and the AC contribution I_{ω} is

$$I_{\omega}(V_{bias}+V_{\omega}) =$$

$$\sum_{n=-\infty}^{\infty} J_n(eV_{\omega}/\hbar\omega)[J_{n+1}(eV_{\omega}/\hbar\omega) + J_{n-1}(eV_{\omega}/\hbar\omega)] I_{dc}(V_{bias}+n\hbar\omega/e). \tag{2.3c}$$

In the above, V_{bias} is the DC bias voltage, V_{ω} is the equivalent AC voltage induced by illumination, I is the tunneling current density (I is used instead of J to differentiate from the Bessel's function J_n), $\hbar$ is the reduced Planck's constant, ω is the frequency of the radiation, and J_n is the Bessel function of order n. The total voltage across the tunnel diode is given as

$$V_{total}(t) = V_{bias} + V_{\omega}(t). \tag{2.4}$$

Fig. 2.2 compares the J-V curves obtained from Simmons' model (classical model) and the semi-classical model for an Al-Al_2O_3-Al diode with a 3nm thick Al_2O_3 layer. It is seen that the semi-classical model predicts a slightly higher and more nonlinear tunneling current density than the classical model when illuminated with an optical signal at 1.55μm wavelength.

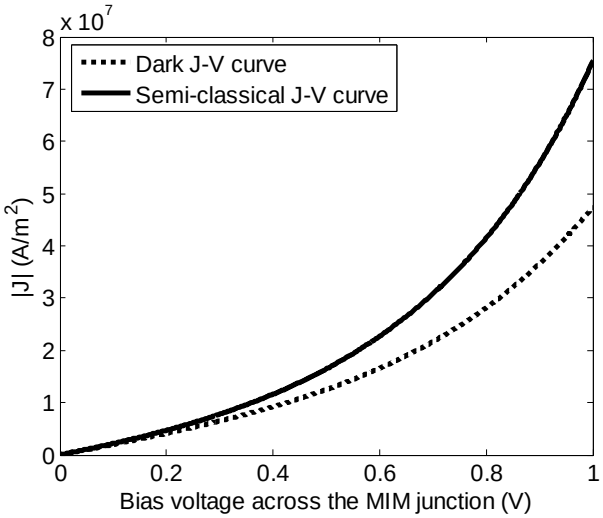


Figure 2.2 Plots of dark (given by Simmons' model in Eq. 2.2(f)) and semi-classical (illuminated at 1.55μm wavelength and peak field amplitude of 3.33×10^7 V/m) *J-V* curves of an Al-Al_2O_3-Al diode with a 3nm thick Al_2O_3 layer.

2.2.2 Circuit model of antenna-coupled MIM diode detectors

To enhance the detection efficiency of MIM diodes, an antenna is typically used to capture the incident radiation and coupled it into the MIM diode, inducing a voltage drop across the two metal electrodes. The antenna thus serves as a voltage source feeding the signal to the diode. Since the physical dimensions of a typical MIM diode are much smaller than the incident wavelength, the MIM diode can be treated as a lumped circuit connected to the antenna [42].

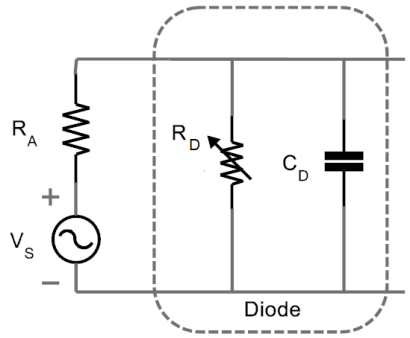


Figure 2.3 Schematic of a lumped circuit model for an antenna-coupled MIM diode detector. The diode model is indicated as inside the dashed zone.

Specifically, the MIM diode is modeled as a nonlinear resistor in parallel with a capacitor while the antenna is modeled as a voltage source with a series resistor to account for the antenna resistance. A schematic of the equivalent circuit of an antenna-coupled MIM detector is given in Fig. 2.3, where R_A is the antenna resistance, V_s is the equivalent voltage source representing incident signal captured by the antenna, R_D is the diode junction resistance and C_D is the diode capacitance.

The capacitance of the MIM junction is calculated from

$$C_D = \varepsilon_r \varepsilon_o \frac{A}{d}, \tag{2.5}$$

where ε_r is the relative permittivity of the insulator layer, ε_o is the permittivity of vacuum, A is the area of MIM junction and d is the thickness of the insulator layer. The diode resistance is obtained from the slope of the current-voltage curve at the bias point as [2]

$$R_D = \frac{dV}{dI}|(V_{bias}) \cdot \tag{2.6a}$$

For the semi-classical I-V curve, the diode resistance can be calculated using the following finite difference form [43]

$$R_D = \frac{dV}{dI}|(V_{bias}) = \frac{2\hbar\omega/e}{I(V_{bias}+\hbar\omega/e) - I(V_{bias}-\hbar\omega/e)}, \tag{2.6b}$$

where I is the tunneling current density given by Eq. 2.3a. Using the classical I-V curve of the Al-Al_2O_3-Al diode with a 3nm thick insulator layer (dotted line in Fig. 2.2), R_D is calculated to be 7.5MΩ at zero bias and 4.7MΩ when biased at 0.1V. This semi-classical model calculates the resistance from the reciprocal of the slope of the secant line between the two tunneling currents at voltages $V_{bias} \pm \hbar\omega/e$ [43]. In the semi-classical model, the diode resistance R_D exhibits frequency dependence under illumination. We calculated R_D for the Al-Al_2O_3-Al diode at a bias voltage of 0.1V and compared it to the classical value in Fig. 2.4. It is clear that the diode resistance extracted from the semi-classical model decreases as the frequency of the incident radiation increases. At higher radiation frequency, the effective tunnel barrier is reduced (as illustrated from the band diagram in Fig. 2.2(c)) and the junction becomes less resistive to electron tunneling. As a result,

increasing tunneling current density is achieved at the same bias condition leading to a lower diode resistance. Note, however, that both models give vary large diode resistances in the range of MΩs.

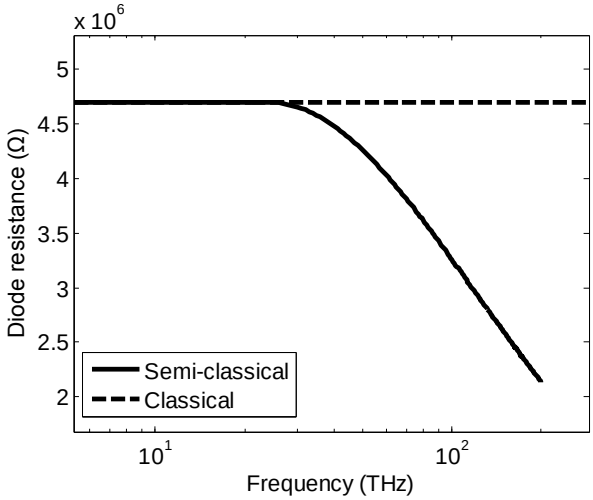


Figure 2.4 Plots of semi-classical and classical resistances (at 0.1V bias voltage) for the Al-Al_2O_3-Al diode with its current-voltage characteristic given in Fig. 2.2. The classical value is calculated to be frequency independent and has a value of 4.7MΩ.

MIM diode resistances in the range of 70KΩ-20MΩ were also reported for an Nb-NbO_x-Nb junction with a size of 500 × 500nm^2 at zero bias resistance by Berland et al. [16]. The same group [16] was later able to reduce their diode resistance to below 7.5KΩ with improved fabrication techniques. Much lower resistances were observed for Ni-NiO-Ni diodes. For example, Fumeaux et al. [4] reported a zero bias resistance of about 100Ω for a 100 × 100nm^2 Ni-NiO-Ni junction and Hobbs et al. [3] achieved a value of 150Ω for a junction of size 600 × 600nm^2. These results show that small diode resistances may be achieved by proper choice of the material system and better refinement of the fabrication techniques; although this still remains a significant challenge with current MIM fabrication technology and most MIM diodes exhibit large zero-bias resistances.

In the circuit model in Fig. 2.3, the antenna is treated as a resistor in series with a voltage source. The resistance R_A is around 50-100Ω in the microwave to infrared frequencies for a typical planar antenna made of highly conductive metals [4,15]. Analysis of the circuit in Fig. 2.3 shows that the antenna and diode work

collectively as a low pass filter with cut-off frequency

$$f_C = \frac{1}{2\pi R_A C_D}. \tag{2.7}$$

Rectification cannot be carried out efficiently for incident radiation beyond the cut-off frequency which is mainly determined by the diode capacitance. Using the relative permittivity of Al_2O_3 (4.29 at 1.55μm wavelength [44]) and the resistance of the Al-Al_2O_3-Al junction (4.7MΩ at 0.1V bias), we find that to rectify an optical signal at 1.55μm wavelength, we would require an MIM diode with an area of 20 × 20nm^2 or less. This small diode size is extremely demanding for current fabrication technology. For a typical MIM diode junction of size 100 × 100nm^2, the cut-off frequency f_C is calculated to be 25THz.

Besides the above bandwidth limitation, the low coupling efficiency from the antenna to the diode places another limitation on the performance of MIM detectors. The power coupling efficiency is defined as

$$\eta_C = \frac{P_D}{P_{in}}, \tag{2.8}$$

where P_D is the power coupled to the diode and P_{in} is the power received by the antenna. For antennas in the THz and infrared frequencies, there is a simple approximation for the power received by the antenna. For a balance-arm antenna, the power received by the antenna can be estimated from the voltage V_s induced across the two antenna arms and the antenna input resistance R_A [42],

$$P_{in} = \frac{V_s^2}{8R_A}. \tag{2.9}$$

In the circuit in Fig. 2.3, the power delivered to the diode is given by

$$P_D = \frac{|V_D|^2}{2R_D}, \tag{2.10}$$

with the voltage V_D across the diode calculated from voltage division

$$V_D = \frac{V_s Z_D}{R_A + Z_D} = \frac{V_s R_D}{R_A + R_D + j\omega R_A R_D C_D}. \tag{2.11}$$

Using Eqs. 2.9 - 2.11 in Eq. 2.8, we obtain the following expression for the frequency dependent coupling efficiency

$$\eta_C = \frac{4R_A R_D}{(R_A + R_D)^2 + \omega^2 R_A^2 R_D^2 C_D^2}. \quad (2.12a)$$

At THz, the condition of $\omega^2 R_A^2 R_D^2 C_D^2 << 1$ is satisfied and Eq. 2.12a is reduced to

$$\eta_C = \frac{4R_A R_D}{(R_A + R_D)^2} = \frac{4}{(\sqrt{\frac{R_D}{R_A}} + \sqrt{\frac{R_A}{R_D}})^2}. \quad (2.12b)$$

The above equation shows that to achieve unity coupling efficiency, we require the diode resistance to be matched to the antenna resistance: $R_D = R_A$.

Due to the large diode resistance, the impedance matching condition is difficult to achieve for most MIM junctions, resulting in poor coupling efficiency. For the Al-Al_2O_3-Al diode with current-voltage relation shown in Fig. 2.2, its frequency dependent coupling efficiency is plotted in Fig. 2.5, assuming antenna input resistance of 50Ω and 0.1V diode bias voltage. It can be seen that the coupling efficiency is extremely low.

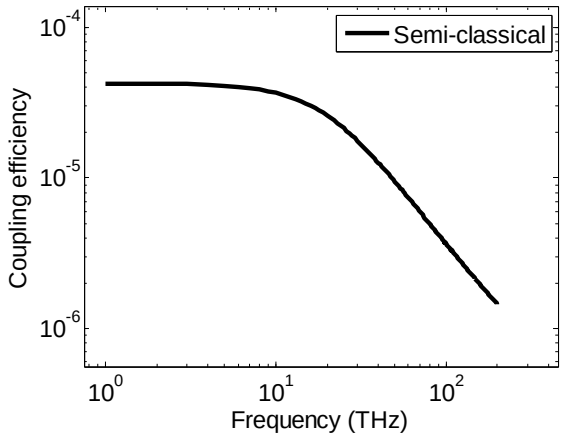


Figure 2.5 Coupling efficiency between an antenna (50Ω) and the Al-Al_2O_3-Al diode (Fig. 2.4).

The output of an MIM diode is a rectified DC current which is proportional to the incident radiation power. The detection efficiency of the MIM diode is specified by its responsivity which is defined as the rectified DC current i_{dc} divided by the radiation power P_{in} delivered to the junction [45]

$$\Re_{diode} = \frac{i_{dc}(V_{bias})}{P_{in}}. \quad (2.13)$$

Suppose the incident radiation induces an AC voltage V_ω across the diode junction.

The diode current can be expressed as a Taylor series expansion as around the bias voltage V_{bias}:

$$I(V_{bias}+V_{\omega}) = I(V_{bias}) + \frac{I^{'}(V_{bias})}{1!}V_{\omega} + \frac{I^{''}(V_{bias})}{2!}V_{\omega}^{2} + \frac{I^{'''}(V_{bias})}{3!}V_{\omega}^{3} + \ldots. \quad (2.14a)$$

Assuming the AC voltage has the form $V_{\omega} = V\cos(\omega t)$, where ω is the frequency of the incident radiation, the rectified DC current arising from the V^2 term in Eq. 2.14(a) is found to be [32]

$$I_{dc} = \frac{I^{''}(V_{bias})}{4}V^{2}. \quad (2.14b)$$

The power delivered to the MIM diode is

$$P_{diode} = V_{\omega}I_{\omega} = V_{\omega}I^{'}(V_{bias})V_{\omega} = \frac{I^{'}(V_{bias})}{2}V^{2}. \quad (2.14c)$$

The responsivity of the MIM diode given by Eq. 2.13 is thus

$$\Re_{diode} = \frac{I_{dc}}{P_{diode}} = \frac{\frac{I^{''}(V_{bias})}{4}V^{2}}{\frac{I^{'}(V_{bias})}{2}V^{2}} = \frac{1}{2}\frac{I^{''}}{I^{'}}(V_{bias}). \quad (2.15a)$$

In the semi-classical analysis, the responsivity can be computed using finite difference schemes for the derivatives $I^{'}$ and $I^{''}$ as [43]

$$\Re_{diode} = \frac{1}{2}\frac{I^{''}}{I^{'}} = \frac{I(V_{bias}+\hbar\omega/e) - 2I(V_{bias}) + I(V_{bias}-\hbar\omega/e)}{(\hbar\omega/e)[I(V_{bias}+\hbar\omega/e) - I(V_{bias}-\hbar\omega/e)]}. \quad (2.15b)$$

The semi-classical model reflects the slope change of the secant line between the two tunneling currents at voltages $V_{bias} \pm \hbar\omega/e$ (other than a continuous derivative at the bias voltage V_{bias} as used in a classical approach) [43]. The responsivity of the Al-Al_2O_3-Al junction with a 3nm-thick insulator layer is plotted as a function of the incident radiation frequency in Fig. 2.6(a). At low frequencies the diode responsivity is given by its classical value (1.84A/W). At high frequencies the responsivity is seen to approach the quantum limit [43], which is the detection efficiency of one free electron produced for each photon absorbed in the junction.

The overall responsivity of the antenna-coupled MIM detector is the product of the coupling efficiency and the diode responsivity

$$\Re_{MIM\_detector} = \eta_C \Re_{Diode}. \tag{2.16}$$

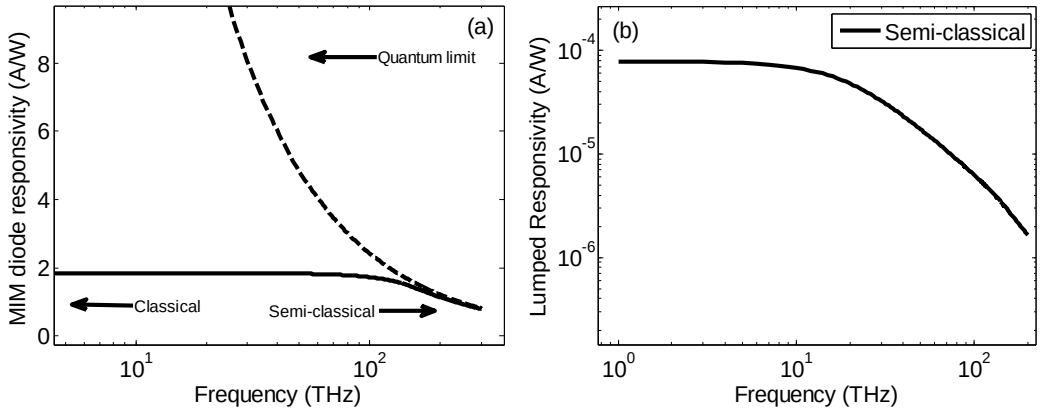


Figure 2.6 Responsivity of (a) an Al-Al_2O_3-Al diode and (b) antenna-coupled Al-Al_2O_3-Al diode. The diode has a square shape and a width of 100nm as well as a 3nm thick insulator layer.

The responsivity of the antenna-coupled Al-Al_2O_3-Al diode detector as a function of frequency is shown in Fig. 2.6(b). The downward trend of the responsivity at higher frequencies is caused by the poor coupling efficiency from the antenna to the diode (Fig. 2.5) as well as the frequency cut-off of the diode (f_c = 25THz).

2.3 Antenna Design for Wideband Terahertz Detection

To realize broadband THz detectors using antenna-coupled MIM diodes, the antennas are required to have broad receiving bandwidth in the THz range. This requirement cannot be satisfied by current THz antenna designs based on conventional dipole and bowtie structures. Thus, there is a strong motivation for the engineering of new antenna designs capable of receiving broadband THz radiation.

Antennas are commonly used to enhance the receiving efficiency in detectors such as antenna-coupled traveling-wave photodiodes (TWPDs) [46] and antenna-coupled MIM diodes [2-5]. Several parameters are used to evaluate the performance of a receiver antenna: impedance, return loss bandwidth, gain,

efficiency, and radiation patterns [15]. Antenna impedance determines the coupling efficiency from the antenna to a fixed load such as a diode detector. Coupling efficiency is commonly expressed in terms of the power return loss which is calculated as the ratio of radiated power to the incident power [15]

$$RL(dB) = 10\log_{10}\frac{P_r}{P_i}. \tag{2.17}$$

The return loss (in dB) is a negative value representing loss due to reflection and reflects the impedance matching condition between the antenna and the load. The frequency span over which the antenna return loss falls below -10dB is defined as the antenna bandwidth [15]. Antenna gain is another important parameter for receiver antennas and is defined as the ratio of the radiation intensity in a specific direction to the total power accepted by the antenna. Antenna directivity relates to its gain via the antenna efficiency, which is the ratio of the accepted power to the total incident power. For planar antennas on an integrated chip, we evaluate the antenna gain in the direction normal to the substrate surface, which is the direction of the incoming signal. The three parameters, gain, directivity and efficiency measure how efficiently an antenna is radiating or receiving in a specific direction. Antenna radiation patterns provide a visual description of its directivity. In general, antennas with wide bandwidth, high directivity and high gain are desirable for MIM diode detectors. In this section we propose and investigate a novel multi-arm THz antenna design based on surface modification of a traditional bowtie antenna with the aim of enhancing its receiving bandwidth.

2.3.1 Traditional antenna surfaces

Most traditional antennas are resonant antennas. They are designed for a specific operational frequency range around the resonant frequency, which appears as a dip in the return loss plot. At resonance, the antenna achieves the best impedance matching condition with a load. Resonant antennas are simple to design for a specific operational frequency but they normally have narrow bandwidths.

A variety of antenna surfaces has been proposed for antenna-coupled MIM

diode detectors, with the most widely used geometries being the bowtie antenna, the dipole antenna, and the spiral antenna [4]. All three antennas are balanced-arm antennas where the two metal arms are connected to the two metal electrodes of an MIM diode. These antennas are miniaturized from their microwave counterparts with well-known characteristics and established fabrication processes. Of the three popular antennas, the bowtie antenna has been the most studied [4,47] due to its simple geometry, high receiving efficiency, broadband performance, and ease of fabrication. For this reason, we will concentrate our study on the bowtie antenna design and its improvements.

Fig. 2.7(a) shows a schematic of the bowtie antenna and defines the various parameters of the structure: the neck width *D*, the total arm length *L*, the arm width *W* and the flare angle θ [15]. A thin-film bowtie antenna on a substrate receives radiation with effective wavelength λ_e determined by the substrate permittivity. The lowest order resonant bowtie antenna has its total arm length equal to the effective wavelength, $L = \lambda_e$. Increasing the arm length will shift the antenna to multi-resonance condition. The impedance of a bowtie antenna is determined primarily by its flare angle and the arm length. Fig. 2.7(b) shows the surface current distribution of a bowtie antenna. The highest surface current density is found near the neck area and is characteristic of the bowtie geometry.

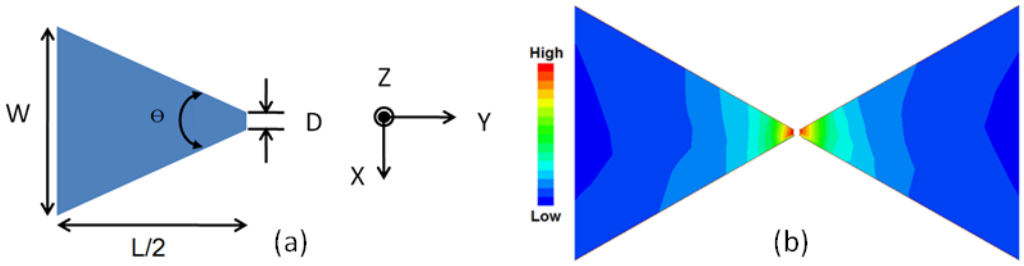


Figure 2.7 (a) Parameters (right arm not shown) : *D* is the neck width, *L* is the total arm length, *W* is the arm width and *Θ* is the arm flare angle and (b) surface current density distribution for a typical bowtie antenna with a flare angle of 60^o.

To evaluate the performance of the conventional bowtie antenna design, we simulated its return loss in the frequency range from 1THz to 2.5THz using

the finite element method (FEM). The antenna was assumed to be Al for connection to the electrodes of an Al-Al_2O_3-Al diode. The neck width *D* was chosen to be 100nm, which is the typical width of a square MIM diode. We fixed the antenna flare angle at 60^o and considered three arm lengths L of λ_e, 1.5λ_e and 2λ_e, where λ_e = 58μm is the effective wavelength at 1.5THz on a Si substrate (ε_{Si}=11.9 [44]). Assuming the antenna was connected to a typical load impedance of 50Ω, we plotted the return loss of the antenna as a function of frequency in Fig. 2.8. It is seen that as the arm length is increased, more dips appear in the return loss spectrum due to higher-order resonances. These dips effectively widen the total antenna bandwidth (-10dB bandwidth). For example, the bandwidth of the structure with 2λ_e arm length is seen to increase by nearly three times that of the structure with λ_e arm length. Inspite of this, all three antennas still do not satisfy the requirement of a single broad bandwidth exceeding 1THz.

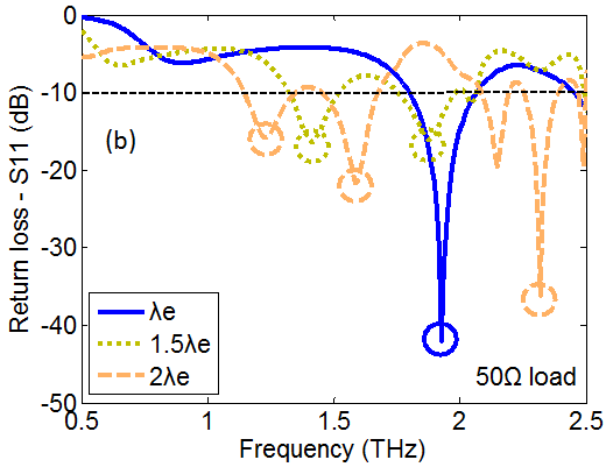


Figure 2.8 Return loss responses of the bowtie antennas with a flare angle of 60^o and arm length of λ_e, 1.5λ_e and 2λ_e (λ_e = 58μm).

2.3.2 Improved antenna design

Surface current distribution patterns affect antenna characteristics and performance, especially the antenna bandwidth. One way to enhance the antenna bandwidth is by achieving proper surface current redistribution via surface engineering or geometrical modification. In some situations, a change of the

antenna surface will result in new resonant frequencies and multiple resonances can be tuned and merged to form a wider effective bandwidth.

There have been attempts to geometrically enhance the performance of bowtie antennas. One example is a bowtie antenna with reconfigurable radiation patterns by connecting multiple bowtie antennas into a ceiling-fan configuration [48]. With proper excitation of individual bowtie antennas and choosing the appropriate number of units, the new antenna could be controlled to have its radiation profiles tuned towards specific directions. In another example a wired bowtie antenna is reconfigured by connecting multiple dipole antennas into the shape of a rake with a flare angle separating two adjacent dipole antennas [49]. With proper adjustment of the flare angles, the new antenna was demonstrated to have less reflection and enhanced transmitted fields compared to a single dipole antenna [49]. In either example, a new antenna is constructed from an array arrangement of multiple unit antennas. The enhanced performance of the new antenna was caused by the array effect rather than the geometrical change of a single antenna.

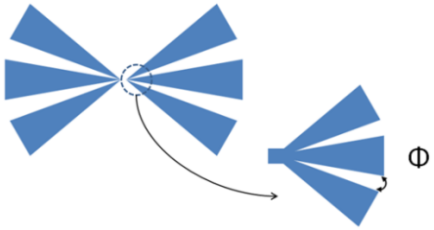


Figure 2.9 Schematic of a triple-arm bowtie antenna, each smaller arm has a flare angle of 20^{o} and the spacing angle between two smaller arms is defined as the split angle Φ.

Our proposed antenna design is based on engineering the surface current redistribution of the bowtie antenna to enhance the antenna bandwidth. Our design, which we call the split-arm configuration and shown in Fig. 2.9, has the shape of a multi-arm bowtie antenna with a split angle Φ separating two adjacent arms. For example, instead of a single-arm bowtie antenna with a flare angle of 60^{o}, the new antenna can be constructed with three smaller arms having a flare angle of 20^{o} and a split angle Φ between two arms. Each arm is characterized by

the same parameters of a conventional bowtie antenna. Since the arms are connected together, they receive the same excitation simultaneously.

In the split-arm configuration, the split angle Φ, the flare angle θ and the number of arms control the degree of surface current redistribution. These parameters can be adjusted to realize wideband antenna performance. Below we summarize the results of our study on the effects of the number of arms and the split angles on the antenna bandwidth performance. More detailed study of other antenna parameters is presented in Appendix A.

(a) Effect of the number of bowtie arms

Fixing the antenna arm length at L = $1.5\lambda_e$ and the split angle Φ at 15^o, for each flare angle θ of 15^o, 20^o and 30^o, we increased the number of arms until a total effective flare angle of 60^o was reached.

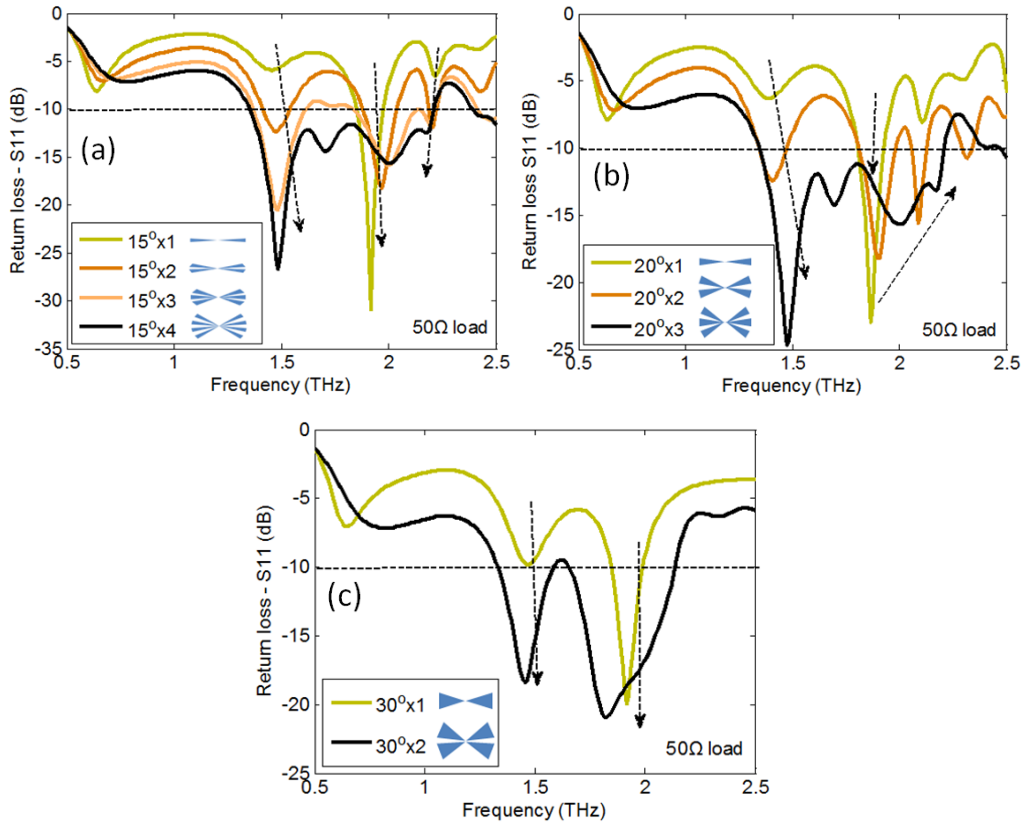


Figure 2.10 Plots of arm assembly effects on the antenna return loss bandwidth performance (with a 50Ω load) for multi-arm bowtie antennas with unit arm flare angle of (a) 15^o, (b) 20^o and (c) 30^o.

Fig. 2.10 plots the return loss bandwidth for each flare angle with increasing number of arms. All three cases show that as the number of arms is increased, the return loss bandwidth becomes broadened due to improved impedance matching to the load (assumed to be 50Ω). If we regard the split-arm antenna as multiple single-arm antennas in parallel, each equivalent to a load, then it is clear that increasing the number of arms (or loads) in parallel has the effect of reducing the total load impedance. This can also be seen from the surface current distributions in Fig. 2.11 for antennas with a 15^o flare angle and number of arms increased from 1 to 4. It is seen that the surface current density in the feeding neck area increases with the number of arms. Since the total power is constant, the high current density implies that lower impedance is achieved.

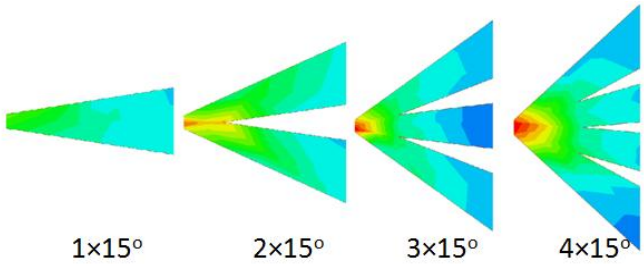


Figure 2.11 Surface current distributions for multi-arm bowtie antennas 1×15^o, 2×15^o, 3×15^o, and 4×15^o with $1.5\lambda_e$ effective arm length at 1.5THz.

(b) Effect of the split angle

Fixing the antenna arm length at $L = 1.5\lambda_e$ and the number of arms at 4, for each flare angle θ of 15^o, 20^o and 30^o we varied the split angle Φ. Fig. 2.12 shows the return loss bandwidth for different values of the split angle. It is seen that as the split angle is increased, a broader return loss bandwidth is achieved. As shown in Fig. 2.13 for the antenna with a fixed flare angle of 15^o, increasing the split angle from 5^o to 15^o leads to an increase in the total surface current density at the feeding neck area, thereby lowering the antenna impedance at constant power.

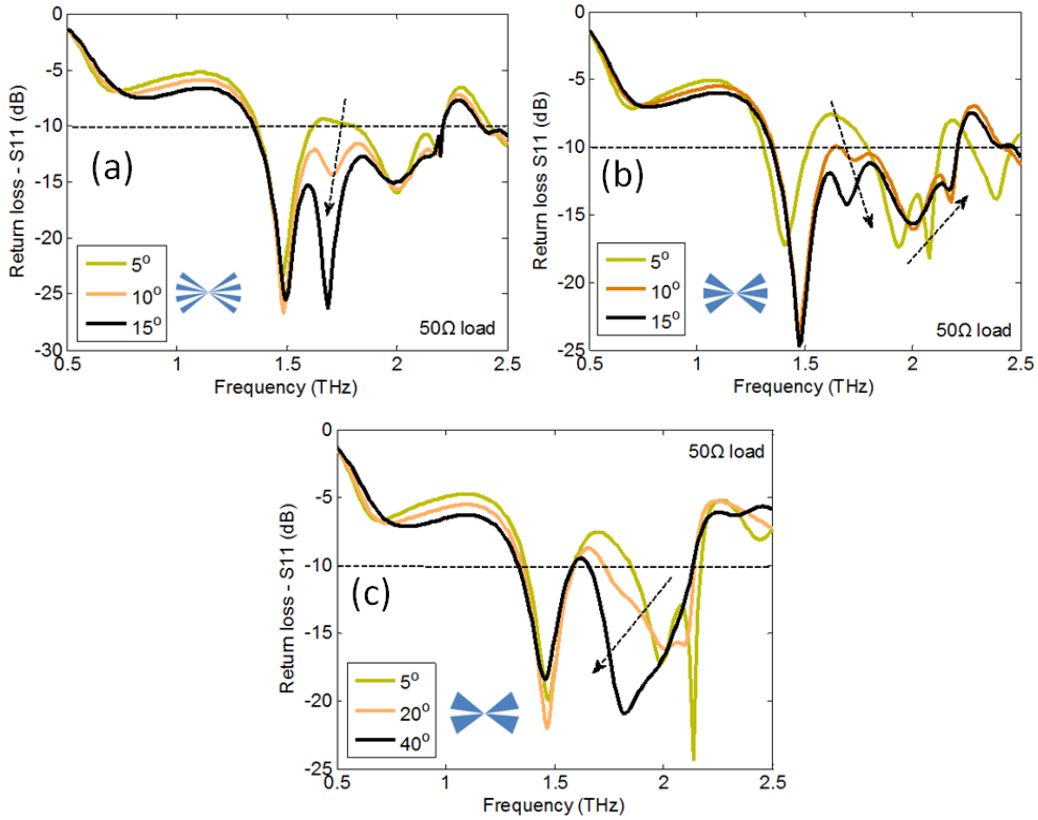


Figure 2.12 Plots of return loss bandwidth (with a 50Ω load) for multi-arm bowtie antennas with unit arm flare angle of (a) 15°, (b) 20° and (c) 30° with different split angles.

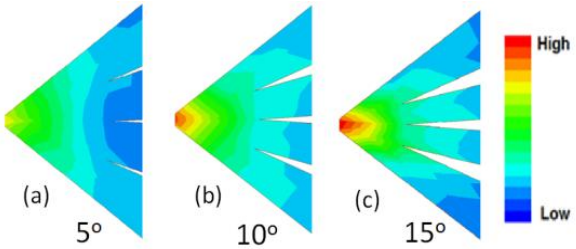


Figure 2.13 Surface current distributions for multi-arm bowtie antennas 4×15° at split angles of (a) 5°, (b) 10°, and (c) 15°.

(c) Summary of split-arm antenna performance

For an effective flare angle of 60° and an antenna arm length of $1.5\lambda_e$, there is a combination of flare angle, number of arms, and split angle that will give optimum antenna performance. A comparison of the different combinations will

help us choose the optimum design to achieve the desired antenna characteristics. For the effective arm length of $1.5\lambda_e$, we compare three different configurations: 4×15^o, 3×20^o and 2×30^o at their optimum split angles. Comparisons are made in terms of the antenna impedance and return loss bandwidth as given in Fig. 2.15 (similar results for the λ_e and $2\lambda_e$ cases and other antenna parameters will be given in Appendix A).

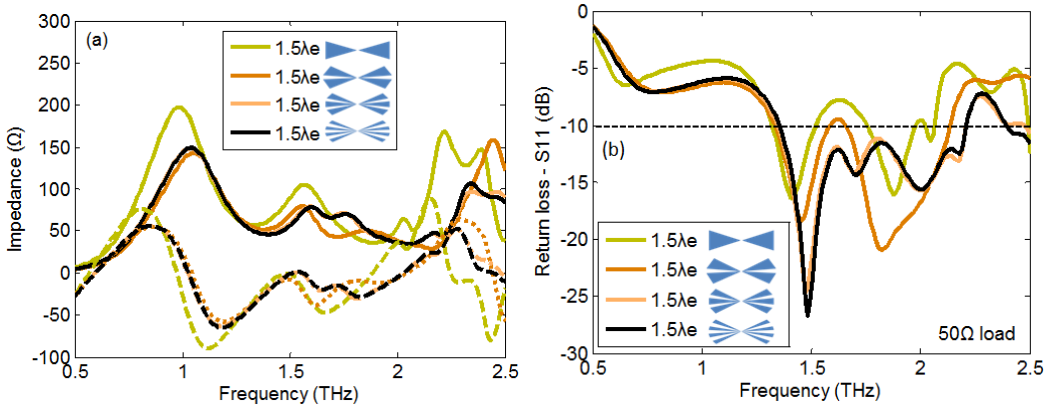


Figure 2.14 Plots of (a) reactance (dashed), resistance (solid) and (b) return loss bandwidth S11for Bowtie, 2×30^o, 3×20^o and 4×15^o multi-arm antennas.

Multi-arm antennas generally exhibit lower reactance and resistance compared to traditional single-arm bowtie antennas. This is primarily a consequence of the parallel connection of the multi-arms. Also, the impedance frequency response of a multi-arm antenna has less fluctuation compared to the response of a traditional bowtie antenna. As the impedance response flattens out, the effective antenna bandwidth also increases.

For a given effective flare angle and antenna arm length, there is a combination of flare angle and split angle that gives the optimum antenna performance. Table 2.1 summarizes the optimum designs and bandwidths for antennas with 60^o effective flare angle and arm lengths of λ_e, 1.5 λ_e and $2\lambda_e$. Of particular note is the $15^o\times4$ antenna design with L = $2\lambda_e$, which has a bandwidth exceeding 1THz and can thus support wideband THz applications.

Table 2.1 Total or (single-resonance) return-loss bandwidths for the multi-arm bowtie antennas in the frequency range from 1THz to 2.5THz and with optimum split angels.

Antenna	Bowtie	$30^o \times 2$	$20^o \times 3$	$15^o \times 4$	Increase %
λ_e	316GHz (266GHz)	544GHz (341GHz) $\Phi = 40^o$	545GHz (335GHz) $\Phi = 15^o$	**650GHz** **(650GHz)** $\Phi = 15^o$	105% (144%)
$1.5\lambda_e$	476GHz (225GHz)	731GHz (479GHz) $\Phi = 35^o$	894GHz **(865GHz)** $\Phi = 15^o$	**968GHz** (858GHz) $\Phi = 15^o$	103% (284%)
$2\lambda_e$	706GHz (226GHz)	876GHz (349GHz) $\Phi = 25^o$	1243GHz (1039GHz) $\Phi = 15^o$	**1297GHz** **(1051GHz)** $\Phi = 10^o$	84% (365%)

In summary, the results of our study suggest that with a proper choice of the number of arms, flare angle and split angle, multi-arm antennas can be realized for wideband operation. These antenna designs can also be scaled to operate at higher frequencies into the infrared. For example, Fumeaux et al. [4] demonstrated miniaturized antennas operating at 30THz and Zhong et al. [50] studied the characteristics of plasmonic bowtie antenna at 1.55μm wavelength. Additionally, antennas with higher order resonances can be used in the upper THz frequency range instead of antennas with a single resonance. At higher frequencies, antennas with higher order resonances are not required to obtain a smaller physical size which makes fabrication easier.

2.4 Summary

In conclusion, our analysis in this chapter shows that the overall performance of an antenna-coupled MIM diode detector is predominantly limited by the MIM junction in terms of its high impedance and low responsivity. With the new multi-arm antenna geometry, we were able to design antennas for wideband operation when connected to a 50Ω load. The proposed antenna geometry can provide a wide single-resonance bandwidth of over 1THz, which is a significant

improvement compared to traditional bowtie antennas. On the other hand, impedance mismatch between the antenna and the diode remains an unresolved problem. It is not likely that planar antennas can achieve the high impedance of current diodes, and it is difficult to fabricate diodes with low junction resistance. In addition, the cut-off frequency of antenna-coupled MIM diodes is still restricted by the requirement of small diode junction areas while the responsivity of the detector is restricted by the poor coupling efficiency from the antenna to the diode. In the next chapter we investigate MIM diode detectors based on the traveling-wave structure that allows both high responsivity and wide bandwidth to be achieved.

Chapter 3

Theoretical Analysis of MIM Nanoplasmonic Waveguides[1]

3.1 Introduction

In the previous chapter we saw that the impedance mismatch between an MIM diode junction and the antenna is currently a major challenge which prevents the simultaneous achievement of high responsivity and wide bandwidth performance. One way to overcome this problem is to use the traveling-wave design, which is commonly employed for semiconductor photodiodes to enhance their electrical bandwidth [51-52]. Traveling-wave MIM (TW-MIM) detectors have been theoretically studied by Grover et al. [26] to overcome the optical bandwidth limitation of lumped MIM diodes for CW rectification. In a TW-MIM detector, the MIM structure acts as a waveguide which supports the propagation of both the detected optical signal as well as the rectified signal. The TW-MIM diode structure offers three advantages over the lumped MIM junctions:

(1) Electromagnetic waves propagate in a TW-MIM structure in certain modes and experience modal characteristic impedance instead of the lumped impedance of the MIM junction. The modal characteristic impedance can be designed to match with the impedance of a typical antenna (e.g., 50Ω) [4,15,26] by proper choice of the waveguide dimensions [26]. The ability to achieve impedance matching greatly enhances the coupling efficiency from the antenna to the MIM waveguide.

(2) The MIM waveguide supports wave propagation in the fundamental TM mode with no cutoff condition [26,29]. As a result, the TW-MIM detector has extremely wide detection bandwidth from DC up to the infrared frequencies.

[1] Part of this work has been published by X. Lei et al. in IEEE Photonics Conference (IPC) 2012, pp. 953-954, Sept. 2012 and presented in IEEE Photonics Conference 2012, Nanoplasmonic II session, Sep 23-27, 2012, Burlingame, CA, USA.

(3) As the optical signal propagates along the MIM waveguide, it is also being rectified by the MIM junction. The increase in the interaction length results in higher detector responsivity compared to lumped MIM diodes [26].

As mentioned above, TW-MIM detectors have been investigated for overcoming the RC-limited detection bandwidth of lumped MIM diodes, thereby enabling the rectification of CW signals up to the optical frequencies. In this chapter we investigate the use of the traveling-wave structure for broadening the *electrical* bandwidth of MIM detectors. In other words we are interested in the potential application of TW-MIM detectors in ultrafast optical pulse rectification leading to single-cycle THz pulse generation. With the ultrafast electron tunneling time in the order of 10^{-16}-10^{-15}s [6], TW-MIM detectors can potentially provide extremely broad electrical bandwidths unmatched by semiconductor traveling-wave photodetectors (TWPDs), whose response is limited by the electron transit time [51-52].

In our analysis of the TW-MIM structure for ultrafast pulse rectification, we treat the MIM structure as a nonlinear metallic waveguide with an effective nonlinear conductivity arising from the nonlinear current-voltage characteristic of the MIM junction. At optical frequencies, with suitable choice of the metal, the MIM waveguide can also support surface plasmon polariton (SPP) propagation [8]. In this case the TW-MIM structure acts as a nonlinear nanoplasmonic waveguide which can be used for novel applications such as frequency conversion on a plasmonic chip.

This chapter focuses on the theoretical analysis of nonlinear MIM nanoplasmonic waveguides in which the nonlinearity arises from electron tunneling. The coupled mode theory (CMT) given in [53-56] is modified and used to analyze nonlinear CW rectification, CW second-order harmonic generation, CW frequency mixing, pulse propagation, pulse rectification and frequency mixing in the MIM nanoplasmonic waveguide. We choose the Al-Al_2O_3-Al material system for the MIM waveguide structure since the current-voltage

characteristics of this junction are well characterized [38-39], as presented in chapter 2. However, our method of analysis can readily be applied to MIM nanoplasmonic waveguides composed of other material systems.

The rest of this chapter is organized as follows. Section 3.2 provides an overview of linear MIM nanoplasmonic waveguides and their characteristics. Section 3.3 incorporates the nonlinear conductive model derived from electron tunneling into the MIM waveguide analysis. Section 3.4 presents analytical results for CW signal rectification, second harmonic generation and frequency mixing in the MIM waveguide. In Section 3.5, the coupled mode theory is used to analyze nonlinear pulse propagation and frequency mixing in the MIM nanoplasmonic waveguide. Concluding remarks will be given in Section 3.6.

3.2 Linear Propagation Characteristics of MIM Nanoplasmonic Waveguides

At optical frequencies, electromagnetic waves propagating along a metal-dielectric interface can excite surface plasmon polaritons (SPPs). SPPs result from the coupling of electromagnetic waves to electron plasma (or electron gas) oscillations at the metal surface, which propagates along the metal-dielectric interface as a surface mode [27,57]. Various types of plasmonic waveguides have been proposed, such as long range SPP waveguides [58], dielectric loaded SPP waveguides [27], and metal-insulator-metal (MIM) waveguides [28]. However, among these structures only the MIM waveguide can support electromagnetic wave propagation at the deep sub-wavelength scale [27,29] since the mode is primarily confined within the insulator layer which can be reduced to a few nanometers thick.

With such a thin insulator layer, electron tunneling begins to occur between the two metal layers, giving rise to an effective nonlinear conductivity in the MIM waveguide which can be used for optical rectification and frequency generation. In this section we will discuss the linear propagation characteristics of MIM nanoplasmonic waveguides.

3.2.1 Dispersion characteristics of MIM nanoplasmonic waveguides

Surface plasmon polariton (SPP) propagation at a metal-dielectric interface is illustrated in Fig. 3.1(a) as coupled electromagnetic field to electron oscillations. The SPP field propagates in the x direction while decays evanescently away from the metal-dielectric interface in the z direction. At the interface of the two half spaces (metal and dielectric), the dispersion relation of the propagation of SPPs is governed by [29]

$$\beta = k_o \sqrt{\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}}, \tag{3.1}$$

where k_o is the free space wave number, ε_d is the relative permittivity of the dielectric material, ε_m is the relative permittivity of the metal material, and β is the propagation constant of the SPP field propagating along the metal-dielectric interface.

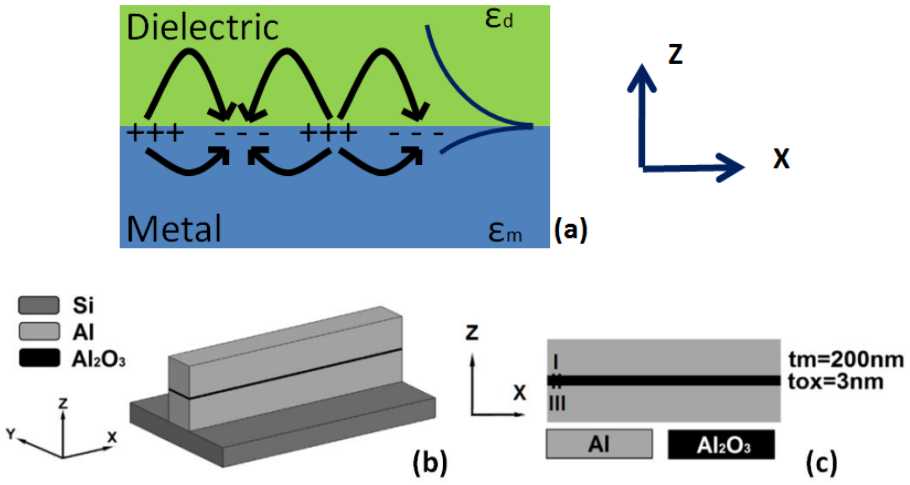


Figure 3.1 (a) Illustration of excited surface plasmon polariton (SPP) at a metal-dielectric interface, (b) 3D and (c) 2D (X-Z plane cut) schematics of the Al-Al_2O_3-Al nanoplasmonic waveguide with 200nm thick Al layers on both sides and 3nm thick Al_2O_3 layer.

A MIM nanoplasmonic waveguide is constructed with an insulator layer sandwiched between two metal layers and it supports the propagation of coupled

SPPs from two metal-dielectric interfaces. The linear propagation characteristics of an MIM nanoplasmonic waveguide are given by its dispersion and modal field profiles. Fig. 3.1(b) shows a schematic of an Al-Al_2O_3-Al MIM nanoplasmonic waveguide with 200nm thick metal layers and 3nm thick Al_2O_3 layer. Since the waveguide width is much larger than the insulator gap thickness, the field distribution of the fundamental mode along the y-direction can be considered to be constant. This approximation allows us to reduce the 3D waveguide structure in Fig. 3.1(b) to a 2D slab waveguide shown in Fig. 3.1(c).

We are interested in the fundamental transverse magnetic (TM) mode with the electric field E_z directed across the insulator gap to drive the electron tunneling. The transverse magnetic field H_y of the waveguide mode is given by the Helmholtz equation

$$\frac{\partial^2 H_y}{\partial z^2} + (k_o^2 \varepsilon_r - \beta^2) H_y = 0, \tag{3.2a}$$

where ε_r is the relative permittivities of the three waveguide layers, and β is the waveguide propagation constant. By solving Eq. 3.2a in the different layers and applying the continuity boundary conditions for the tangential field components E_x and H_y, we obtain the dispersion relation [29]

$$e^{-4k_1 d} = \left(\frac{k_1/\varepsilon_1 + k_2/\varepsilon_2}{k_1/\varepsilon_1 - k_2/\varepsilon_2}\right)\left(\frac{k_1/\varepsilon_1 + k_3/\varepsilon_3}{k_1/\varepsilon_1 - k_3/\varepsilon_3}\right). \tag{3.2b}$$

In the above equation, d is the thickness of the insulator layer, and k_i is the transverse (z-direction) propagation constant in layer i (as denoted in Fig. 3.1(c)) and is related to the propagation constant β via $k_i^2 = \beta^2 - k_o^2 \varepsilon_i$.

We solved Eq. 3.2b for the propagation constant of the Al-Al_2O_3-Al plasmonic waveguide with parameters d = 3nm, $\varepsilon(Al_2O_3)$ = 4.29 [44] and ε(Al) taken from [59-60]. Fig. 3.2 plots the phase velocity and the propagation length of the waveguide at frequencies from microwave to infrared. (The propagation length is defined as $L_p = 1/\alpha$ (attenuation constant).) The plot shows strong frequency dispersion in both the phase velocity and propagation loss. The rapidly decreasing propagation length with increasing frequency indicates that the plasmonic waveguide becomes substantially more lossy at higher frequencies.

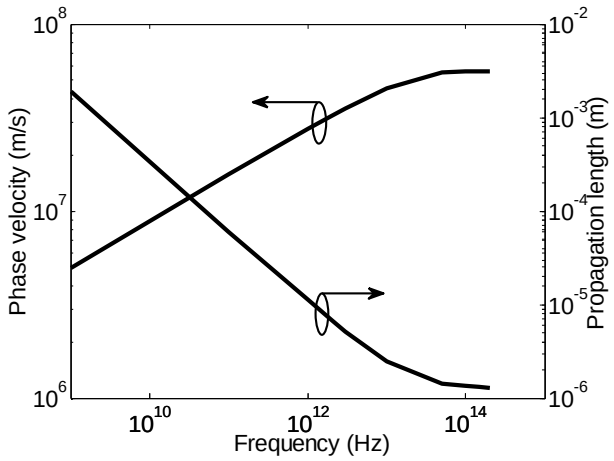


Figure 3.2 Plots of phase velocity (left) and propagation length (right) dispersion characteristics of the Al-Al_2O_3 (3nm)-Al nanoplasmonic waveguide from microwave to infrared [61].

The H_y field distribution of the fundamental TM mode obtained from the solution of Eq. 3.2a and is plotted in Fig. 3.3(a) at the 1.55 μm wavelength. From the solution for H_y, we can also determine the field profiles for the electric field components E_z and E_x from Maxwell's equations. The results are shown in Figs. 3.3(b) and (c). The field components E_z and H_y are observed to have symmetric field profiles with respect to the center of the insulator layer while the E_x component is antisymmetric. Note also that the E_z field is strongly confined in the insulator gap, yielding a maximum voltage drop across the potential barrier which enhances the electron tunneling probability.

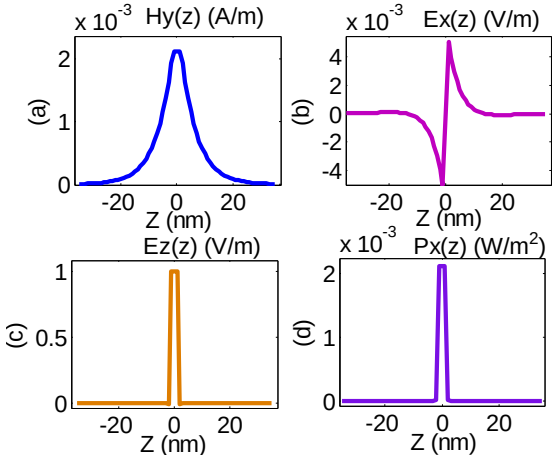


Figure 3.3 Transverse z-direction profiles in the Al-Al_2O_3-Al MIM nanoplasmonic waveguide (Fig. 3.1) for field components: (a) H_y, (b) E_x, (c) E_z, and (d) power distribution $P_x(z)$ at the wavelength of 1.55μm. The plots are zoomed to show the fields in waveguide central area.

Another interesting aspect of the MIM plasmonic waveguide is the large power confinement. Since the fields are relatively constant across the waveguide width (the y-direction), we can calculate the transverse power distribution of the waveguide from

$$P_x(z) = w\frac{1}{2}\int_{-\infty}^{+\infty} E_z^*(z)H_y(z)dz, \tag{3.3}$$

where w is the width of the waveguide. Fig. 3.3(d) shows the transverse power distribution of the Al-Al_2O_3-Al waveguide at the 1.55μm wavelength. It is clear that the power is strongly confined within the insulator layer.

The power confinement factor of the MIM waveguide is defined as the ratio of the power confined in the insulator gap to the total power in the waveguide [62]

$$F\% = \frac{P_{gap}}{P_{total}} \times 100\%. \tag{3.4}$$

Using the above formula we determined the power confinement in the Al-Al_2O_3-Al waveguide to be more than 95% over the frequency range from microwave to the optical frequencies. The strong power confinement and approximately constant E_z field distribution across the insulator gap allow us to treat nonlinear wave propagation in the MIM nanoplasmonic waveguide as a one-dimensional problem in our theoretical analysis in Section 3.3.

3.2.2 Antenna coupling efficiency to MIM nanoplasmonic waveguides

To achieve high coupling efficiency from an antenna to the MIM waveguide, we need to match the impedance of the antenna to the characteristic impedance of the MIM waveguide instead of its junction resistance. As we will show in this section, the MIM waveguide has much lower characteristic impedance than the junction resistance, thereby enabling close to 100% antenna coupling efficiency to be achieved.

In analogy to transmission lines, the characteristic impedance of the MIM nanoplasmonic waveguide can be calculated from its equivalent voltage V and

surface current density I [63]. Since the MIM waveguide supports only the fundamental TM mode, its equivalent voltage and current density can be obtained from direct integration of the transverse field components. The characteristic impedance is calculated from [63]

$$Z_{MIM} = \frac{V}{I} = \left| \frac{\int_{M1}^{M2} E_z dz}{\int_{-\infty}^{+\infty} H_y dl} \right|. \tag{3.5}$$

As seen in Fig. 3.3(c), the transverse electric field E_z exists mainly inside the insulator layer and is constant over this region, so we can approximate the top integral in Eq. 3.5 by $V = E_z d$. Similarly, we can approximate H_y to be constant over the waveguide width w and zero outside, which allows the integral over H_y to be approximated by $I = H_y w$. In addition, E_z and H_y are related by $E_z = \beta H_y / \omega \varepsilon_d \varepsilon_o$, where ε_d is the relative permittivity of the insulator. Eq. 3.5 thus gives

$$Z_{MIM} = \frac{\beta_{TM} d}{\omega \varepsilon_d \varepsilon_O w}. \tag{3.6}$$

For the Al-Al_2O_3-Al nanoplasmonic waveguide with 3nm thick insulator, the characteristic impedances for the frequency span from sub-THz to optical frequencies for two waveguide widths of 50nm and 100nm are plotted in Fig. 3.4(a). The plot shows that the characteristic impedance of the MIM nanoplasmonic waveguide is much smaller than the MIM junction resistance (which is in the KΩ-MΩ range [16-17]). More importantly, it can be designed to perfectly match with the impedance of a planar antenna (typically 50Ω) by proper choice of the waveguide width.

The coupling efficiency from an antenna to the MIM waveguide is given by

$$\eta_C = \frac{4 R_A Z_{MIM}}{(R_A + Z_{MIM})^2}. \tag{3.7}$$

The coupling efficiency from a 50Ω antenna to the Al-Al_2O_3-Al nanoplasmonic waveguide is plotted in Fig. 3.4(b) for two waveguide widths of 50nm and 100nm. Very higher coupling efficiency exceeding 92% is achieved for the 50nm wide

waveguide over the frequency span from microwave to the optical frequencies.

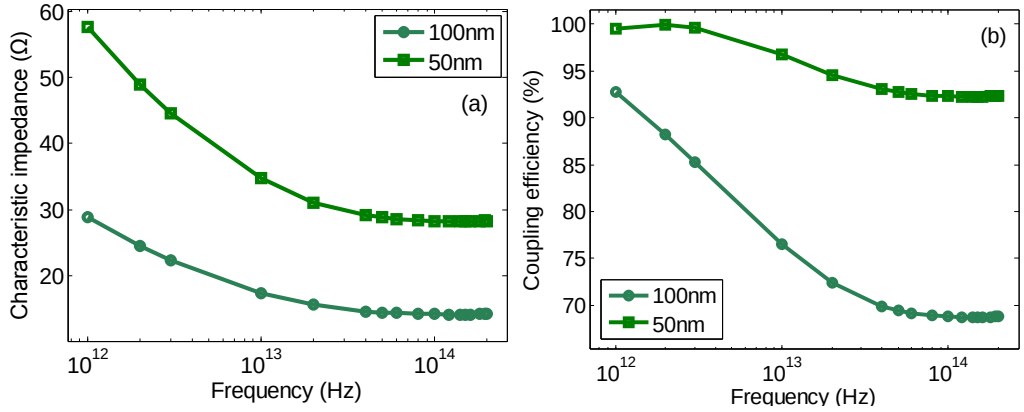


Figure 3.4 Plots of (a) characteristic impedance of an Al-Al_2O_3-Al MIM nanoplasmonic waveguide with waveguide width of 100nm (and 50nm) and insulator layer thickness of 3nm and (b) coupling efficiency from an antenna (50Ω) to the Al-Al_2O_3-Al nanoplasmonic waveguide with waveguide width of 100nm (and 50nm) from sub-THz to the infrared.

Besides impedance matching, the MIM nanoplasmonic waveguide is also required to achieve mode matching with the antenna. From the field profiles in Fig. 3.3(b), the TM mode is seen to have an anti-symmetric longitudinal electric field (E_x) distribution [26], which matches with the anti-symmetric current distributions of a balanced-arm antenna. As a result, the MIM nanoplasmonic waveguide also has good mode matching with the receiver antenna [26].

3.3 Nonlinear Conductivity Model of MIM Waveguides in the Presence of Electron Tunneling

To study wave propagation in an MIM nanoplasmonic waveguide in the presence of electron tunneling, we model the nonlinear tunneling process by prescribing an effective nonlinear conductivity to the insulator layer. Due to the ultrafast electron tunneling time in the order of 10^{-16}-10^{-15}s [6], the nonlinear conductive process can be considered to be instantaneous. From the nonlinear J-V curve of the MIM tunneling junction, we derive a nonlinear constitutive relation for the current density J_z as a function of the transverse electric field E_z in the insulator as

$$J(E_z) = \sigma_1 E_z + \sigma_2 E_z^2, \tag{3.8}$$

where σ_1 (Ω^{-1}/m) represents the linear ohmic loss and σ_2 (Ω^{-1}/V) is the differential conductivity responsible for second-order nonlinear processes such as rectification and second harmonic generation. These conductivity parameters are computed by applying central finite-differencing schemes to the semi-classical $J(V)$ curve around the bias point V_{bias} as:

$$\sigma_1 = d\frac{dJ}{dV_\omega} = \frac{J(V_{bias} + \hbar\omega/e) - J(V_{bias} - \hbar\omega/e)}{2(\hbar\omega/ed)}, \tag{3.9}$$

$$\sigma_2 = \frac{d^2}{2}\frac{d^2J}{dV_\omega^2} = \frac{J(V_{bias} + \hbar\omega/e) - 2J(V_{bias}) + J(V_{bias} - \hbar\omega/e)}{2(\hbar\omega/ed)^2}. \tag{3.10}$$

In the above, V_ω is the equivalent AC voltage induced by the illumination and is given by $V_\omega = E_z d$, assuming constant E_z field across the insulator gap d.

Using the semi-classical J-V curve for an Al-Al_2O_3-Al junction with 3nm, 2.5nm and 2nm thick insulator gaps, we computed the conductivity parameters σ_1 and σ_2 for different junction bias voltages at a fixed illumination wavelength of 1.55μm. The results are plotted in Fig. 3.5, which indicates that higher nonlinear conductivity σ_2 can be achieved by using a higher bias voltage, although the linear ohmic loss σ_1 also increases with the bias.

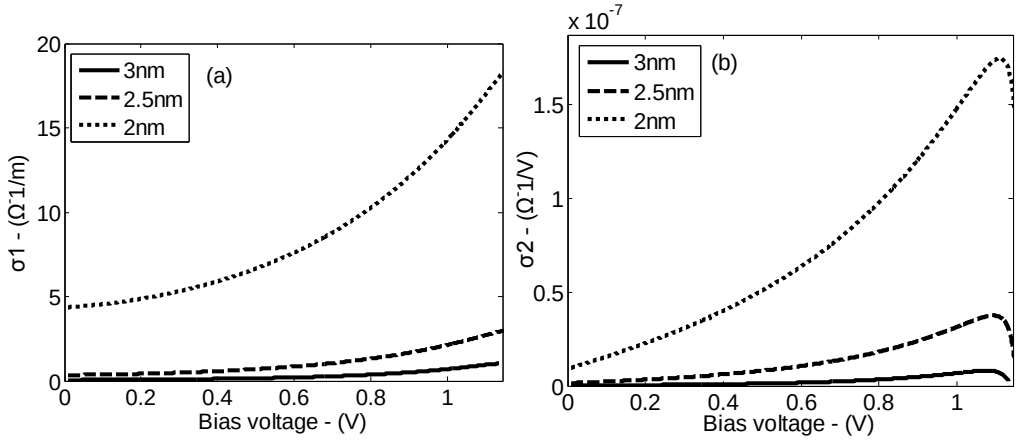


Figure 3.5 Plots of the conductivity parameters (a) σ_1, and (b) σ_2 at different bias voltages for an Al-Al_2O_3-Al MIM diode with a barrier width of 3nm, 2.5nm and 2nm and barrier height of 2eV.

As the junction gap thickness decreases, both σ_1 and σ_2 increase and demonstrate similar change trends. Thus, we focus our analysis on a junction gap thickness of

3nm. At a bias voltage of 0.1V and junction gap thickness of 3nm, the conductivity values are $\sigma_1 = 4\times10^{-2}\Omega^{-1}$/m and $\sigma_2 = 3.5\times10^{-10}\ \Omega^{-1}$/V. The quadratic J-V curve fit using Eq. 3.8 is plotted in Fig. 3.6, showing good agreement with the semi-classical J-V curve of the MIM junction.

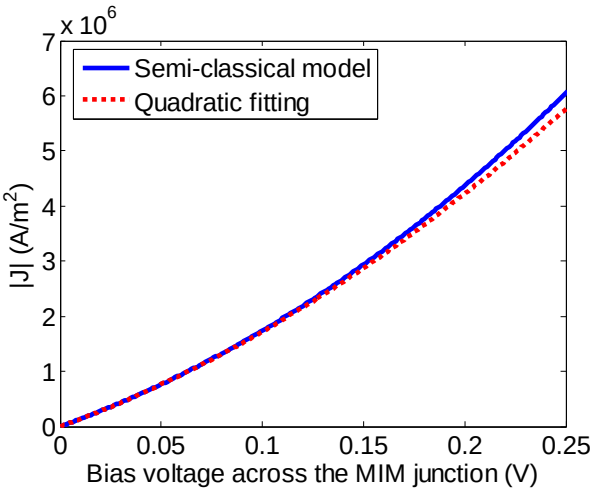


Figure 3.6 Plots of the semi-classical (illuminated) *J-V* characteristics of an Al-Al_2O_3-Al MIM diode with a barrier width of 3nm and barrier height of 2eV (Fig. 2.2) and the *J-V* quadratic fit around the bias voltage of 0.1V.

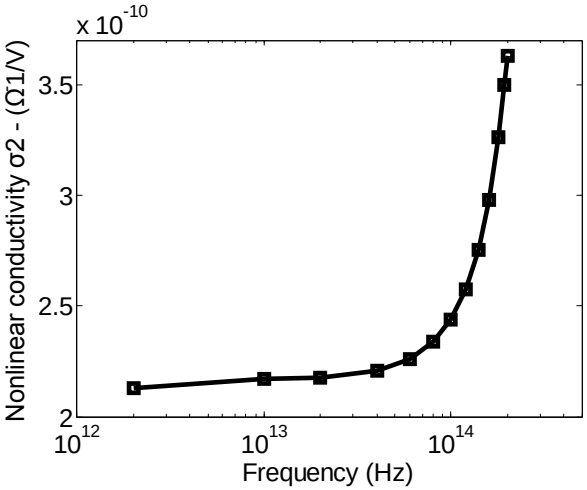


Figure 3.7 Plot of the nonlinear conductivity σ_2 at different frequencies and a bias voltage of 0.1V of an Al-Al_2O_3-Al MIM diode with a barrier width of 3nm and barrier height of 2eV.

In addition to the bias voltage dependence, we also study the frequency dependence of the second-order nonlinear conductivity σ_2 by calculating its values at different illumination wavelengths at a fixed bias voltage of 0.1V. Fig. 3.7

shows the result for the frequency span from 2THz to 200THz. It is seen that the nonlinear conductivity increases only slightly over a large frequency range from THz to infrared.

The nonlinear constitutive relation in Eq. 3.8 will be used to model the electron tunneling process across the MIM junction in all our subsequent theoretical and numerical analyses of nonlinear wave propagation in MIM plasmonic waveguides.

3.4 CW Propagation in Nonlinear MIM Nanoplasmonic Waveguides

In this section we present a theoretical analysis of nonlinear wave propagation in MIM nanoplasmonic waveguides based on the nonlinear conductivity model derived in the previous section. Using the Coupled Mode Theory (CMT), we investigate the second-order nonlinear processes of optical rectification and second harmonic generation under continuous wave (CW) excitation. The analytical derivations given in the rest of this chapter is based on works done by Kothari [54] and Yariv and Yeh [53].

Rectification of a CW signal in a TW-MIM detector has been studied by Grover et al. [26] using a distributed circuit model of the waveguide. Assuming that rectification causes no depletion of the input signal, the authors derived an expression for the rectified DC current I_{dc} in the TW-MIM detector as

$$I_{dc}(L) = \frac{\Re_{diode} V_{\omega}^2 W}{2R_{D/S}} \left(\frac{1 - e^{-2\alpha L}}{2\alpha} \right). \tag{3.11}$$

In the above, V_{ω} is the equivalent AC voltage of the illuminating signal at the input of the MIM waveguide, R_{diode} is the responsivity of the lumped MIM diode and is given by Eq. 2.15b, $R_{D/S}$ is the resistance per unit area of the MIM junction, α is the attenuation constant of the waveguide, W and L are the width and length of the waveguide, respectively. The responsivity of the detector is given by

$$\Re_{TW-MIM} = \frac{I_{dc}(L)}{P_{optical}}, \tag{3.12}$$

where $P_{optical}$ is the input optical power. The above analysis is based on the distributed lumped circuit model of the TW-MIM detector, which treats the MIM waveguide as an array of lumped MIM diodes connected in series. Our following analysis will be more rigorous in that it is based on nonlinear wave propagation in the MIM waveguide and will be derived from the coupled wave equations.

3.4.1 Nonlinear CW wave propagation and rectification

Since the mode in the MIM waveguide is a quasi transverse electromagnetic (quasi-TEM) mode with the transverse electric field E_z strongly confined inside the gap layer, we can approximate the waveguide as a one-dimensional (1D) medium with an effective permittivity ε_R and conductivity σ_e. The wave equation for propagation in the 1D nonlinear medium is given by

$$\frac{\partial^2 E}{\partial x^2} = \frac{\varepsilon_R}{c^2}\frac{\partial^2 E}{\partial t^2} + \mu_o \sigma_e \frac{\partial E}{\partial t} + \mu_o \sigma_2 \frac{\partial E^2}{\partial t}, \tag{3.13}$$

where c is the speed of light in vacuum. The effective permittivity ε_R and linear conductivity σ_e can be computed from the propagation constant γ of the MIM waveguide, which is given by

$$\gamma^2 = -k_o^2(\varepsilon_R - j\varepsilon_I) = -k_o^2(\varepsilon_R - j\frac{\sigma_e}{\omega\varepsilon_o}), \tag{3.14}$$

where k_o is the free space wave number.

In the 1D effective medium, the field is a plane wave propagating in the x direction. For CW wave propagation, the field components are expressed as

$$\tilde{E}(x,t) = \frac{1}{2}\{E(x)e^{j(\beta x - \omega t)} + c.c.\}, \tag{3.15}$$

where $E(x)$ represents the slowly varying field envelope of the wave along the propagation direction and β is the real part of the propagation constant γ. Substituting Eq. 3.15 into Eq. 3.13 and retaining only fields with $\exp(-j\omega_1 t)$ dependence (assuming an input CW frequency of ω_1), we obtain the equation for the field envelope (detailed analysis is provided in Appendix B)

$$\frac{\partial^2 E_1}{\partial x^2} + 2\frac{\partial E_1}{\partial x}(j\beta_1) + E_1(j\beta_1)^2 \cdot$$

$$= \frac{\varepsilon_{1R}}{c^2} E_1(-j\omega_1)^2 - j\omega_1\mu_o\sigma_{1e}E_1 - j\omega\mu_o\sigma_2 E_2 E_1^* e^{j(\beta_2 - 2\beta_1)x}, \quad (3.16)$$

Using the slowly varying envelope approximation, the second order derivative can be neglected and Eq. 3.16 is simplified into

$$\frac{\partial E_1}{\partial x} = -\frac{\omega_1\mu_o\sigma_{1e}}{2\beta_1} E_1 - \frac{\omega_1\mu_o\sigma_2}{2\beta_1} E_2 E_1^* e^{j(\beta_2 - 2\beta_1)x}. \quad (3.17)$$

With $E_2 << E_1$, the term which contributes to the input signal due to the second order nonlinear conductive effect is negligible compared to the input signal. We have the condition that

$$|\sigma_{1e}E_1| >> |\sigma_2 E_2 E_1^*|. \quad (3.18)$$

Eq. 3.17 is further simplified into

$$\frac{\partial E_1}{\partial x} = -\frac{\omega_1\mu_o\sigma_{1e}}{2\beta_1} E_1. \quad (3.19)$$

Letting

$$\alpha_1 = \frac{\omega_1\mu_o\sigma_{1e}}{2\beta_1}, \quad (3.20a)$$

$$E_1 = \frac{A_1}{\sqrt{n_1}} \propto \frac{\sqrt{P_1}}{\sqrt{n_1}}, \quad (3.20b)$$

where P_1 is the peak power of field E_1, Eq. 3.19 becomes

$$\frac{\partial A_1}{\partial x} = -\alpha_1 A_1, \quad (3.21)$$

which has the solution

$$A_1 = C_1 e^{-\alpha_1 x}, \quad (3.22)$$

where C_1 is an arbitrary constant. The solution for $E_1(x)$ is then

$$E_1(x) = E_1(0)e^{-\alpha_1 x}. \quad (3.23)$$

The tunneling current density given by Eq. 3.8 has a DC component of

$$J_{dc}(x) = \frac{\sigma_2}{2} |E_1(x)|^2. \quad (3.24)$$

The total rectified DC current I_{dc} is obtained by integrating Eq. 3.24 over the waveguide length

$$I_{dc}(x) = \frac{W\sigma_2}{2}\int_0^L |E_1(x)|^2 dx, \tag{3.25}$$

where W is the width and L is the length of the MIM waveguide. Putting Eq. 3.23 into Eq. 3.25 and using the approximation $E(0)=V_o/d$, where V_o is the peak voltage drop across the insulator gap induced by the incident signal at the input end of the waveguide, the total rectified DC current is obtained as (replace E_1 by E and α_{1e} by α_e)

$$I_{dc}(L) = \frac{W\sigma_2 V_o^2}{2d^2}\left(\frac{1-e^{-2\alpha_e L}}{2\alpha_e}\right), \tag{3.26}$$

where α_e is the effective attenuation constant which also takes into account loss due to electron tunneling. Comparing Eq. 3.26 with Eq. 3.11, we find that they give the same exponential dependence of the rectified current on the propagation distance. We also notice that they depend on the same MIM junction parameters. Relating Eq. 2.6b, Eq. 2.15b, Eq. 3.9 and Eq. 3.10, we find that the diode responsivity $\Re_{diode}$ and per unit area resistance $R_{D/S}$ are related to the nonlinear conductivity parameters σ_1 and σ_2 as

$$\Re_{diode} = \frac{1}{2}\frac{I''}{I'} = \frac{1}{2}\frac{AJ''}{AJ'} = \frac{1}{2}\frac{J''}{J'} = \frac{1}{2}\frac{2\sigma_2/d^2}{\sigma_1/d} = \frac{\sigma_2}{d\sigma_1}. \tag{3.27a}$$

$$R_{D/S} = \frac{1}{AI'} = \frac{1}{J'} = \frac{1}{\sigma_1/d} = \frac{d}{\sigma_1}, \tag{3.27b}$$

where d is the thickness of the diode insulator layer and A is diode junction area. Putting Eqs. 3.27a and 3.27b into Eq. 3.11b, we obtain

$$I_{dc}(L) = \frac{\frac{\sigma_2}{d\sigma_1}V_\omega^2 W}{2\frac{d}{\sigma_1}}\left(\frac{1-e^{-2\alpha L}}{2\alpha}\right) = \frac{W\sigma_2 V_\omega^2}{2d^2}\left(\frac{1-e^{-2\alpha L}}{2\alpha}\right). \tag{3.28}$$

Comparing Eq. 3.28 and Eq. 3.26, we find that they are identical except for the use of the effective waveguide attenuation constants α and α_e where α_e takes into account loss due to electron tunneling while α doesn't.

Assuming a typical planar antenna impedance of 50Ω, we compare the responsivity of the MIM waveguide detector with a lumped MIM diode detector. Both devices have a gap thickness of 3nm. The waveguide width is 100nm while the MIM diode has an area of 100×100 nm^2. The responsivities are plotted in Fig. 3.8 assuming unity coupling efficiency and a waveguide length of $3L_p$ at the respective frequency. The analytical result computed using Eq. 3.11 is also plotted for comparison.

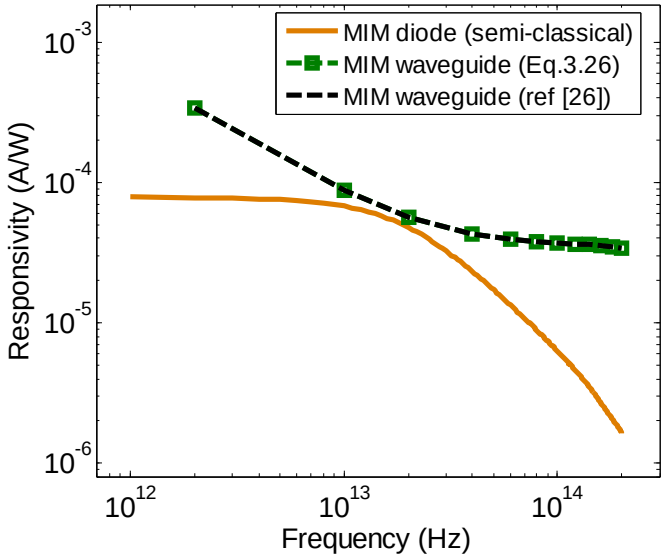


Figure 3.8 Responsivity of the antenna-coupled Al-Al_2O_3-Al waveguide detector with a 3nm Al_2O_3 layer and 100nm waveguide width at unity coupling efficiency from an antenna with impedance of 50Ω and a waveguide length of $3L_p$ at the respective frequency. (Results for the lumped Al-Al_2O_3-Al diode with an area of 100×100 nm^2 and results from [26] are also provided as reference.)

Compared to the lumped MIM diode, the MIM waveguide exhibits much higher responsivity at frequencies below and above the corner frequency around 20THz. Near the corner frequency, which is where the semi-classical model for the diode resistance approaches the classical model as seen from Fig. 2.4, the responsivities of both devices approach each other. This corner frequency is also close to the cut-off frequency of the lumped diode (25THz) where its responsivity begins to drop rapidly. We also note that the responsivity curve of the MIM waveguide slowly levels off at high frequencies. This result indicates that the MIM waveguide detector experiences no frequency cut-off and the slightly

decreasing responsivity at high frequencies is caused by increased waveguide loss. It is also clear that the MIM waveguide detector can still have very high responsivity at the optical frequencies while the lumped device cannot.

We also notice that result obtained by Grover et al. agrees well with our analysis. The agreement of these two results indicates that loss caused by electron tunneling is negligible as compared to metal absorption. In this case, the traveling-wave analysis is as valid as the distributed circuit modeling.

3.4.2 Two wave coupling in nonlinear MIM waveguides

In this section we derive the coupled mode equations for the mixing of two waves at frequencies ω_2 and ω_3 to generate two new waves at the sum and difference frequencies, $\omega_1 = \omega_3 - \omega_2$ and $\omega_4 = \omega_3 + \omega_2$. The total electric fields at the four frequencies involved are

$$\widetilde{E}_1(x,t) = \widetilde{E}_{\omega 3-\omega 2}(x,t) = \frac{1}{2}\{E_1(x)e^{j(\beta_1 x-\omega_1 t)} + c.c.\}, \tag{3.29a}$$

$$\widetilde{E}_2(x,t) = \widetilde{E}_{\omega 2}(x,t) = \frac{1}{2}\{E_2(x)e^{j(\beta_2 x-\omega_2 t)} + c.c.\}, \tag{3.29b}$$

$$\widetilde{E}_3(x,t) = \widetilde{E}_{\omega 3}(x,t) = \frac{1}{2}\{E_3(x)e^{j(\beta_3 x-\omega_3 t)} + c.c.\}, \tag{3.29c}$$

$$\widetilde{E}_4(x,t) = \widetilde{E}_{\omega 4}(x,t) = \frac{1}{2}\{E_4(x)e^{j(\beta_4 x-\omega_4 t)} + c.c.\}, \tag{3.29d}$$

where $E_i(x)$ (i=1 to 4) represents the slowly varying envelope of the i frequency component. Substituting Eqs. 3.29 into Eq.3.13 and using the conditions (with $E_4 << E_2$ and $E_1 << E_3$, the terms which contribute to the input signals due to second order nonlinear conductive effects are negligible compared to the input signals)

$$|\sigma_{2e}E_2| >> |\sigma_2 E_3 E_1^*| \; or \; |\sigma_2 E_4 E_3^*|, \tag{3.30a}$$

$$|\sigma_{3e}E_3| >> |\sigma_2 E_4 E_2^*| \; or \; |\sigma_2 E_1 E_2|, \tag{3.30b}$$

we obtain a set of coupled equations (detailed derivations in Appendix B):

$$\frac{\partial E_1}{\partial x} = -\frac{\omega_1\mu_o\sigma_{1e}}{2\beta_1}E_1 - \frac{\omega_1\mu_o\sigma_2}{2\beta_1}E_3E_2^*e^{j(\beta_3-\beta_2-\beta_1)x}, \tag{3.31a}$$

$$\frac{\partial E_2}{\partial x} = -\frac{\omega_2 \mu_o \sigma_{2e}}{2\beta_2} E_2, \tag{3.31b}$$

$$\frac{\partial E_3}{\partial x} = -\frac{\omega_3 \mu_o \sigma_{3e}}{2\beta_3} E_3, \tag{3.31c}$$

$$\frac{\partial E_4}{\partial x} = -\frac{\omega_4 \mu_o \sigma_{4e}}{2\beta_4} E_4 - \frac{\omega_4 \mu_o \sigma_2}{2\beta_4} E_2 E_3 e^{j(\beta_2+\beta_3-\beta_4)x}. \tag{3.31d}$$

Defining the parameters

$$\alpha_i = \frac{\omega_i \mu_o \sigma_{ie}}{2\beta_i}, \tag{3.32a}$$

$$\kappa_1 = \frac{\omega_1 \mu_o \sigma_2}{2\beta_1} \frac{\sqrt{n_1}}{\sqrt{n_2 n_3}}, \tag{3.32b}$$

$$\kappa_4 = \frac{\omega_4 \mu_o \sigma_2}{2\beta_4} \frac{\sqrt{n_4}}{\sqrt{n_2 n_3}}, \tag{3.32c}$$

and the coordinate transformation

$$E_i = \frac{A_i}{\sqrt{n_i}} \propto \frac{\sqrt{P_i}}{\sqrt{n_i}}, \tag{3.32d}$$

where P_i is the peak power of each field and n is the effective index, Eqs. 3.31 become

$$\frac{\partial A_1}{\partial x} = -\alpha_1 A_1 - \kappa_1 A_3 A_2^* e^{j(\beta_3-\beta_2-\beta_1)x}, \tag{3.33a}$$

$$\frac{\partial A_2}{\partial x} = -\alpha_2 A_2, \tag{3.33b}$$

$$\frac{\partial A_3}{\partial x} = -\alpha_3 A_3, \tag{3.33c}$$

$$\frac{\partial A_4}{\partial x} = -\alpha_4 A_4 - \kappa_4 A_2 A_3 e^{j(\beta_2+\beta_3-\beta_4)x}. \tag{3.33d}$$

To solve the above equations, we first obtain the solutions for A_2 and A_3 in Eqs. 3.33(b-c):

$$E_2(x) = E_2(0) e^{-\alpha_2 x}, \tag{3.34a}$$

$$E_3(x) = E_3(0) e^{-\alpha_3 x}, \tag{3.34b}$$

We then substitute the above solutions into Eqs. 3.33a and 3.33d for A_1 and A_4.

Following the approach in Appendix B and the use of integrating factors, the solutions for A_1 and A_4 with the initial conditions $E_1(0) = 0$ and $E_4(0) = 0$ are

$$E_1(x) = \frac{\sqrt{n_2 n_3}\kappa_1 E_2(0)E_3(0)}{\sqrt{n_1}[j(\beta_3 - \beta_2 - \beta_1) - \alpha_2 - \alpha_3 + \alpha_1]}\{1 - e^{[j(\beta_3 - \beta_2 - \beta_1) - \alpha_2 - \alpha_3 + \alpha_1]x}\}e^{-\alpha_1 x}, \tag{3.35a}$$

$$E_4(x) = \frac{\sqrt{n_2 n_3}\kappa_4 E_2(0)E_3(0)}{\sqrt{n_4}[j(\beta_3 + \beta_2 - \beta_4) - \alpha_2 - \alpha_3 + \alpha_4]}\{1 - e^{[j(\beta_3 + \beta_2 - \beta_4) - \alpha_2 - \alpha_3 + \alpha_4]x}\}e^{-\alpha_4 x}. \tag{3.35b}$$

The above equations show that the generated E_1 and E_4 waves have an exponential dependence on the propagation distance via the loss coefficients α_i. The factor involving β_i gives the effect of phase mismatch. Below we consider separately the cases of rectification, second harmonic generation, and sum and difference frequency generation.

(a) Rectification and second harmonic generation

For the cases of optical rectification and second harmonic generation we set $\omega_2 = \omega_3 = \omega$ and $\omega_4 = 2\omega$ in Eqs. 3.34b and 3.35b to obtain

$$E_2(x) = E_2(0)e^{-\alpha_2 x}, \tag{3.36a}$$

$$E_4(x) = \frac{n_2 \kappa_4 E_2^2(0)}{\sqrt{n_4}[j(2\beta_2 - \beta_4) - 2\alpha_2 + \alpha_4]}\{1 - e^{[j(2\beta_2 - \beta_4) - 2\alpha_2 + \alpha_4]x}\}e^{-\alpha_4 x}. \tag{3.36b}$$

The above equations describe the amplitudes of the input CW and second harmonic waves as functions of the MIM waveguide length. It is also clear that the amplitude of the second harmonic signal has a quadratic dependence on the field amplitude of the input signal. For an input CW signal at wavelength $\lambda_2 = 1.55\mu m$ and peak field amplitude $E_2(0) = 3.33\times10^7$V/m (average power of 0.362mW), the field amplitudes of the fundamental and second harmonic signals are plotted as functions of the propagation distance in Fig. 3.9. It is seen that the

fundamental wave decays monotonically, which is mainly due to waveguide loss. The second harmonic signal reaches a maximum value before decreasing exponentially. This behaviour is due to the competing effects of frequency conversion and waveguide loss. Due to the very short propagation distance in the MIM waveguide, the effect of phase mismatching between the ω and 2ω waves can be neglected from our analysis. At short propagation distance, the conversion rate is higher than the waveguide loss and the second harmonic wave increases rapidly. As the distance increases, the fundamental wave becomes attenuated, the conversion rate decreases rapidly so the second harmonic wave also begins to decrease. The plot in Fig. 3.9(b) indicates that there is an optimum waveguide length at which maximum second harmonic generation efficiency can be achieved.

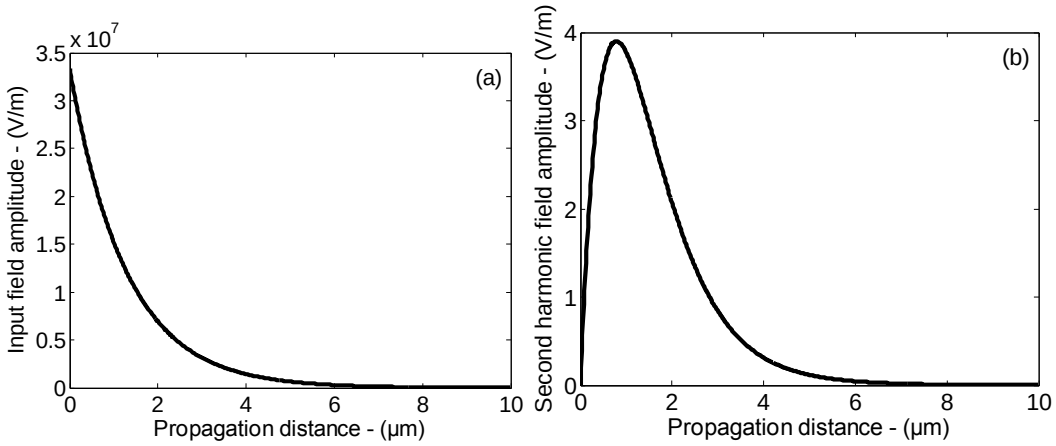


Figure 3.9 Theoretical plots of field amplitudes of (a) input signal, and (b) second harmonic signal as a function of propagation distance. The input CW signal has peak field amplitude of 3.33×10^{7}V/m at 1.55μm wavelength.

The generation of second harmonic frequency based on second-order nonlinear conductivity is similar to second harmonic generation in a medium with nonlinear susceptibility $\chi^{(2)}$. In the former case the new frequency component arises from a nonlinear tunneling current while in the latter case it arises from a nonlinear polarization current. In the $\chi^{(2)}$ medium, the polarization current has an amplitude of

$$| J_p | = | \frac{\partial P}{\partial t} | = | \frac{\partial(\varepsilon_o \chi^{(2)} E^2)}{\partial t} | = 2\omega\varepsilon_o \chi^{(2)} E^2, \qquad (3.37a)$$

Typical bulk media such as crystalline quartz have $\chi^{(2)}$ of 10^{-12}m/V [64], while some particularly good crystals such as GaAs can exhibit large $\chi^{(2)}$ in the order of 10^{-10}m/V [65]. Assuming a typical value of $\chi^{(2)} = 10^{-12}$m/V, the nonlinear polarization current induced by an incident CW signal at the 1.55μm wavelength and peak field amplitude of 3.33×10^{7}V/m is

$$| J_p |= 2\omega\varepsilon_o\chi^{(2)}E^2 = 2.4\times10^{7}\, A/m^2 . \quad (3.37b)$$

In the MIM waveguide, the nonlinear tunneling current is modeled as

$$J_{NL} = \sigma_2 E^2 . \quad (3.37c)$$

Assuming the calculated $\sigma_2 = 3.5\times10^{-10}\ \Omega^{-1}$/V from Section 3.3 at $V_{bias} = 0.1$V, we obtain for the same incident CW wave at 1.55μm and peak field amplitude 3.33×10^{7}V/m

$$| J_{NL} |= \sigma_2 E^2 = 0.4\times10^{6}\, A/m^2 . \quad (3.38)$$

The above results indicate that in general the $\chi^{(2)}$ nonlinear effect is stronger by two orders of magnitude than the second order nonlinear conductive effect in the MIM waveguide [66]. Note, however, that we can enhance the nonlinear conductive effect by biasing the MIM junction at higher voltages.

(b) Sum and difference frequency generation

To investigate the behaviour of the MIM waveguide for sum and difference frequency generation, we assumed incident wavelengths λ_2 = 1.55μm and λ_3 = 1.5μm with field amplitudes $E_2(0) = 1.18\times10^{7}$V/m and $E_3(0) = 1.18\times10^{7}$V/m (so that the maximum voltage drop across the gap layer will not exceed 0.1V which is the bias voltage for the waveguide). We calculated the field amplitudes of the different frequency components and plotted them in Fig. 3.10. The two input CW signals (Fig. 3.10(a)) decay monotonically due to waveguide loss while the frequency difference and frequency sum signals achieve peak field amplitudes at optimum waveguide lengths (Fig. 3.10(b)). These behaviours are similar to that of the second harmonic generation. We also note that the higher amplitude achieved by the frequency difference field is due to the lower waveguide loss at lower

frequencies.

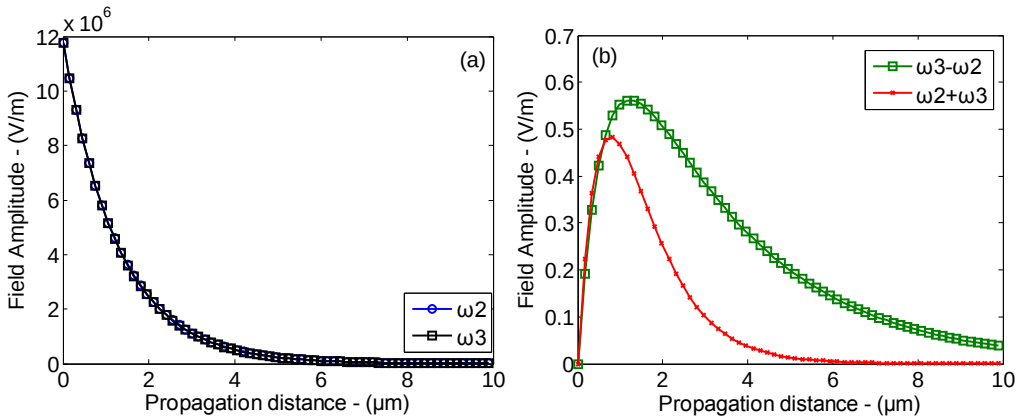


Figure 3.10 Theoretical plots of the amplitudes of (a) input signals ω_2 and ω_3, and (b) frequency mixing products ω_3-ω_2 and ω_2+ω_3 as a function of propagation distance in the MIM waveguide with 3nm thick insulator layer. The incident CW signals have wavelengths λ_2 = 1.55μm and λ_3 = 1.5μm as well as peak field amplitudes $E_2(0) = 1.18\times10^7$V/m and $E_3(0)=1.18\times10^7$V/m.

3.5 Pulse Propagation in a Nonlinear MIM Nanoplasmonic Waveguide

In this section we present a theoretical analysis of nonlinear pulse rectification, second harmonic generation and frequency mixing in the MIM nanoplasmonic waveguide using the CMT theory. The coupled wave equations are similar to those in the CW case but also take into account the time dependence of the pulse envelopes.

3.5.1 Nonlinear pulse rectification and second harmonic generation

We consider an input pulse signal $E_1(x, t)$ at frequency ω_1 to the MIM waveguide. Due to second order nonlinearity, two new pulses will be generated: $E_0(x, t)$ at the baseband and $E_2(x, t)$ at the second harmonic frequency $\omega_2 = 2\omega_1$. The electric fields of the three frequency components can be represented as

$$\tilde{E}_0(x,t) = E_0(x,t)e^{-\alpha_o x}, \tag{3.39a}$$

$$\widetilde{E}_1(x,t)=\frac{1}{2}\{E_1(x,t)e^{j(\beta_1 x-\omega_1 t)}+c.c\}, \tag{3.39b}$$

$$\widetilde{E}_2(x,t)=\frac{1}{2}\{E_2(x,t)e^{j(\beta_2 x-\omega_2 t)}+c.c.\}, \tag{3.39c}$$

where $E_i(x,t)$ is the slowly varying pulse envelope of the i frequency component. By substituting Eqs. 3.39 into Eq. 3.13 and matching fields at each frequency ω_i, we obtain the following coupled wave equations for the different pulses propagating in the MIM waveguide (detailed derivation is given in Appendix C):

$$\frac{\partial E_0}{\partial x}=[\alpha_0^2 E_0-\mu_o\sigma_{0e}\frac{\partial E_0}{\partial t}-\mu_o\sigma_2(\frac{\partial|E_1|^2}{\partial t}+\frac{\partial|E_2|^2}{\partial t})e^{\alpha_0 x}]\frac{1}{2\alpha_0} \tag{3.40a}$$

$$\frac{\partial E_1}{\partial x}=\{(\frac{-2j\omega_1\varepsilon_{1R}}{c^2}+\mu_o\sigma_{1e})\frac{\partial E_1}{\partial t}-j\omega_1\mu_o\sigma_{1e}E_1$$
$$+\mu_o\sigma_2[(\frac{\partial E_2}{\partial t}E_1^*+E_2\frac{\partial E_1^*}{\partial t})e^{j(\beta_2-2\beta_1)x}-j\omega_1 E_2E_1^*e^{j(\beta_2-2\beta_1)x}]\}\frac{1}{2j\beta_1}, \tag{3.40b}$$

$$\frac{\partial E_2}{\partial x}=\{(\frac{-2j\omega_2\varepsilon_{2R}}{c^2}+\mu_o\sigma_{2e})\frac{\partial E_2}{\partial t}-j\omega_2\mu_o\sigma_{2e}E_2$$
$$+2\mu_o\sigma_2E_1[\frac{\partial E_1}{\partial t}e^{j(2\beta_1-\beta_2)x}+E_1(-j\omega_1)e^{j(2\beta_1-\beta_2)x}]\}\frac{1}{2j\beta_2}, \tag{3.40c}$$

We next apply the following transformation to the retarded frame [53] of reference

$$A_i(t,x)=A_i(T_i(x,t),x)=A_i(t-\frac{x}{v_i},x), \tag{3.41a}$$

$$\frac{\partial}{\partial x}=\frac{\partial}{\partial x}+\frac{\partial T_i}{\partial x}\frac{\partial}{\partial T_i}=\frac{\partial}{\partial x}-\frac{1}{v_i}\frac{\partial}{\partial T_i}, \tag{3.41b}$$

$$\frac{\partial}{\partial t}=\frac{\partial T_i}{\partial t}\frac{\partial}{\partial T_i}=\frac{\partial}{\partial T_i}. \tag{3.41c}$$

In the above v_i is the group velocity of wave at frequency ω_i. Defining the parameters

$$\alpha_1=\frac{\mu_o\sigma_{1e}}{2j\beta_1}, \tag{3.42a}$$

$$\alpha_2=\frac{\mu_o\sigma_{2e}}{2j\beta_2}, \tag{3.42b}$$

$$\kappa_1 = \frac{\mu_o \sigma_2}{2j\beta_1 \sqrt{n_2}}, \tag{3.42c}$$

$$\kappa_2 = \frac{\mu_o \sigma_2 \sqrt{n_2}}{2j\beta_2 n_1}, \tag{3.42d}$$

and the coordinate transformation

$$E_i = \frac{A_i}{\sqrt{n_i}} \propto \frac{\sqrt{P_i}}{\sqrt{n_i}}, \tag{3.42e}$$

where P_i is the peak power of each field and n_i is the effective index of each field, we apply the transformations in Eqs. 3.42(a-e) to Eqs. 3.40(a-c) to obtain (with T_i replaced by t in the final expressions)

$$\frac{\partial A_0}{\partial x} - \frac{1}{v_0}\frac{\partial A_0}{\partial t} =$$

$$[\alpha_0^2 A_0 - \mu_o \sigma_{0e} \frac{\partial A_0}{\partial t} - \mu_o \sigma_2 (\frac{\partial |A_1|^2}{\partial t}\frac{\sqrt{n_0}}{n_1} + \frac{\partial |A_2|^2}{\partial t}\frac{\sqrt{n_0}}{n_2}) e^{\alpha_0 x}] \frac{1}{2\alpha_0}, \tag{3.43a}$$

$$\frac{\partial A_1}{\partial x} - \frac{1}{v_1}\frac{\partial A_1}{\partial t} + \frac{\omega_1 \varepsilon_{1R}}{\beta_1 c^2}\frac{\partial A_1}{\partial t} =$$

$$\alpha_1 (\frac{\partial A_1}{\partial t} - j\omega_1 A_1) + \kappa_1 (\frac{\partial A_2}{\partial t} A_1^* + A_2 \frac{\partial A_1^*}{\partial t} - j\omega_1 A_2 A_1^*) e^{j(\beta_2 - 2\beta_1)x}, \tag{3.43b}$$

$$\frac{\partial A_2}{\partial x} - \frac{1}{v_2}\frac{\partial A_2}{\partial t} + \frac{\omega_2 \varepsilon_{2R}}{\beta_2 c^2}\frac{\partial A_2}{\partial t} =$$

$$\alpha_2 (\frac{\partial A_2}{\partial t} - j\omega_2 A_2) + 2\kappa_2 (A_1 \frac{\partial A_1}{\partial t} - j\omega_1 A_1^2) e^{j(2\beta_1 - \beta_2)x}. \tag{3.43c}$$

The above equations describe the propagation of the envelopes of the three pulses in the MIM waveguide. They can be solved numerically using the finite difference method [67].

Assuming an input Gaussian pulse with carrier wavelength $\lambda = 1.55\mu m$, FWHM pulse width of 1ps and peak field amplitude of $1.18 \times 10^7 V/m$, we computed the solutions for the three pulses and plotted them in Fig. 3.11. Fig. 3.11(a) shows the normalized temporal shapes of the rectified, ω and 2ω pulses. The rectified pulse has the shape of the derivative of the input pulse envelope, yielding a single-cycle THz pulse, which is characteristic of a second-order

rectification process [68]. The second-order harmonic pulse has the same Gaussian shape as the input pulse, which is similar to the second harmonic generation process based on $\chi^{(2)}$ nonlinearity [66]. This characteristic is also predicted by the simplified CW coupled mode analysis in Section 3.5.3.

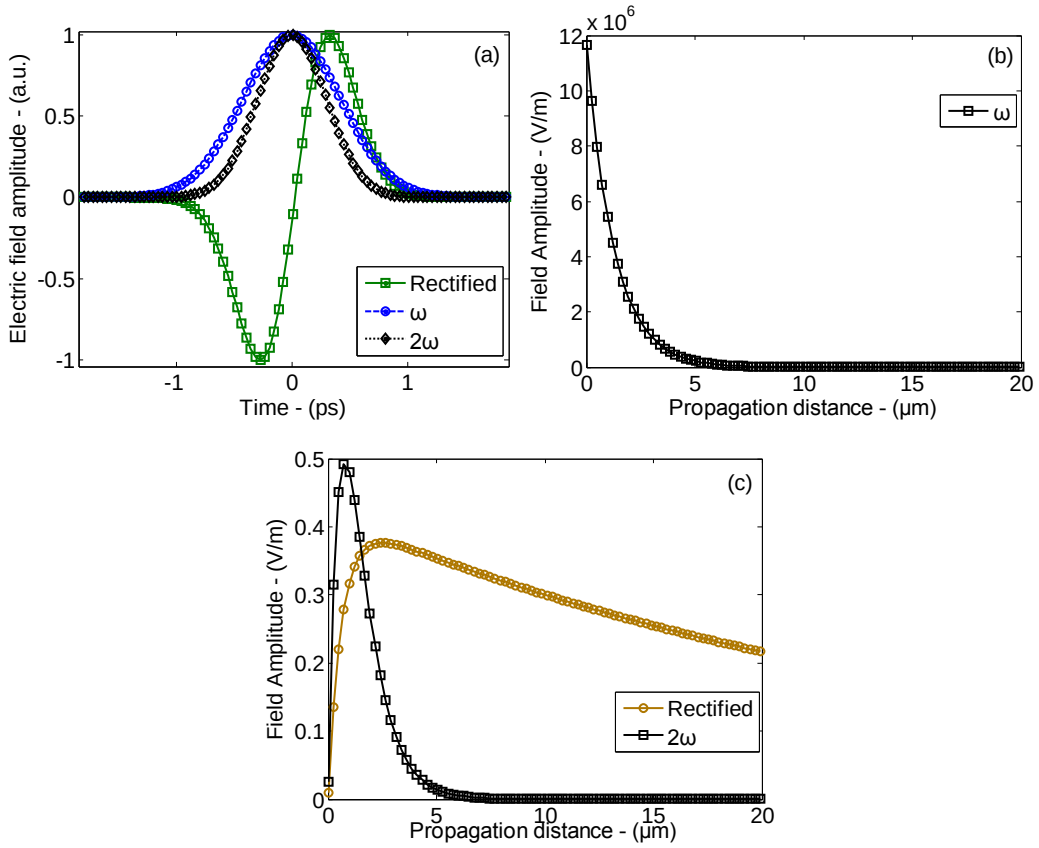


Figure 3.11 Plots of (a) Temporal responses for the rectified, ω and 2ω pulses, (b) peak electric field amplitude for the ω pulse and (c) the rectified and 2ω pulses as functions of propagation distance in the MIM waveguide with a 3nm thick insulator layer. An input 1ps Gaussian pulse with carrier wavelength $\lambda = 1.55\mu m$ and peak field amplitude of 1.18×10^{7}V/m is assumed.

Fig. 3.11(b) and 3.11(c) plot the peak field amplitudes of the rectified, ω and 2ω pulses as a function of the propagation distance, indicating the optimum lengths where the rectified and 2ω pulses achieve their maxima. These curves are also similar to the CW results case and demonstrate the competing effects of frequency conversion and waveguide loss. The plots show that the rectified pulse has a much longer optimum length than that of the 2ω pulse due to much lower

loss at the baseband frequencies. However, the rectified pulse reaches a lower peak amplitude than the 2ω pulse due to the strong velocity mismatch between the baseband and optical signals (the ω and 2ω signals have less velocity mismatch, as can be seen in Fig. 3.2). The above results show that the MIM nanoplasmonic waveguide can potentially be used for detecting ultrafast optical pulses, as well as generating single-cycle THz pulses and second harmonic pulses.

3.5.2 Nonlinear pulse mixing

We consider next the case of frequency sum and difference pulse generations for two input pulses modulating carrier frequencies ω_2 and ω_3 ($\omega_3 > \omega_2$). The electric fields of the two input pulses (E_2 and E_3), the frequency difference pulse (E_1), and the frequency sum pulse (E_4)) are given by

$$\tilde{E}_1(x,t) = \tilde{E}_{\omega 3-\omega 2}(x,t) = \frac{1}{2}\{E_1(x)e^{j(\beta_1 x-\omega_1 t)} + c.c.\}, \tag{3.44a}$$

$$\tilde{E}_2(x,t) = \tilde{E}_{\omega 2}(x,t) = \frac{1}{2}\{E_2(x)e^{j(\beta_2 x-\omega_2 t)} + c.c.\}, \tag{3.44b}$$

$$\tilde{E}_3(x,t) = \tilde{E}_{\omega 3}(x,t) = \frac{1}{2}\{E_3(x)e^{j(\beta_3 x-\omega_3 t)} + c.c.\}, \tag{3.44c}$$

$$\tilde{E}_4(x,t) = \tilde{E}_{\omega 4}(x,t) = \frac{1}{2}\{E_4(x)e^{j(\beta_4 x-\omega_4 t)} + c.c.\}, \tag{3.44d}$$

Follow similar approach given in Section 3.5.1 and described in more details in Appendix C, we obtain the following coupled mode equations for the field envelopes of the different pulses:

$$\frac{\partial A_1}{\partial x} - \frac{1}{v_1}\frac{\partial A_1}{\partial t} + \frac{\omega_1 \varepsilon_{1R}}{\beta_1 c^2}\frac{\partial A_1}{\partial t} =$$

$$\alpha_1\left(\frac{\partial A_1}{\partial t} - j\omega_1\right)A_1 + \kappa_1\left(\frac{\partial A_2^*}{\partial t}A_3 + A_2^*\frac{\partial A_3}{\partial t} - j\omega_1 A_2^* A_3\right)e^{j(-\beta_2 x+\beta_3 x-\beta_1 x)}, \tag{3.45a}$$

$$\frac{\partial A_2}{\partial x}-\frac{1}{v_2}\frac{\partial A_2}{\partial t}+\frac{\omega_2\varepsilon_{2R}}{\beta_2 c^2}\frac{\partial A_2}{\partial t}=$$

$$\alpha_2(\frac{\partial A_2}{\partial t}-j\omega_2)A_2+\kappa_{21}(\frac{\partial A_4}{\partial t}A_3^*+A_4\frac{\partial A_3^*}{\partial t}-j\omega_2 A_3^* A_4)e^{j(\beta_4 x-\beta_2 x-\beta_3 x)}$$

$$+\kappa_{22}(\frac{\partial A_3}{\partial t}A_1^*+A_3\frac{\partial A_1^*}{\partial t}-j\omega_2 A_1^* A_3)e^{j(\beta_3 x-\beta_1 x-\beta_2 x)}, \quad (3.45b)$$

$$\frac{\partial A_3}{\partial x}-\frac{1}{v_3}\frac{\partial A_3}{\partial t}+\frac{\omega_3\varepsilon_{3R}}{\beta_3 c^2}\frac{\partial A_3}{\partial t}=$$

$$\alpha_3(\frac{\partial A_3}{\partial t}-j\omega_3)A_3+\kappa_{31}(\frac{\partial A_2}{\partial t}A_1+A_2\frac{\partial A_1}{\partial t}-j\omega_3 A_1 A_2)e^{j(\beta_1 x+\beta_2 x-\beta_3 x)}$$

$$+\kappa_{32}(\frac{\partial A_4}{\partial t}A_2^*+A_4\frac{\partial A_2^*}{\partial t}-j\omega_3 A_2^* A_4)e^{j(\beta_4 x-\beta_2 x-\beta_3 x)}, \quad (3.45c)$$

$$\frac{\partial A_4}{\partial x}-\frac{1}{v_4}\frac{\partial A_4}{\partial t}+\frac{\omega_4\varepsilon_{4R}}{\beta_4 c^2}\frac{\partial A_4}{\partial t}=$$

$$\alpha_4(\frac{\partial A_4}{\partial t}-j\omega_4)A_4+\kappa_4(\frac{\partial A_2}{\partial t}A_3+A_2\frac{\partial A_3}{\partial t}-j\omega_4 A_2 A_3)e^{j(\beta_2 x+\beta_3 x-\beta_4 x)}, \quad (3.45d)$$

where

$$\alpha_i=\frac{\mu_o\sigma_{ie}}{2j\beta_i}, \quad (3.46a)$$

$$\kappa_1=\frac{\mu_o\sigma_2\sqrt{n_1}}{2j\beta_1\sqrt{n_2 n_3}}, \quad (3.46b)$$

$$\kappa_{21}=\frac{\mu_o\sigma_2\sqrt{n_2}}{2j\beta_2\sqrt{n_3 n_4}}, \quad (3.46c)$$

$$\kappa_{22}=\frac{\mu_o\sigma_2\sqrt{n_2}}{2j\beta_2\sqrt{n_1 n_3}}, \quad (3.46d)$$

$$\kappa_{31}=\frac{\mu_o\sigma_2\sqrt{n_3}}{2j\beta_3\sqrt{n_1 n_2}}, \quad (3.46e)$$

$$\kappa_{32}=\frac{\mu_o\sigma_2\sqrt{n_3}}{2j\beta_3\sqrt{n_2 n_4}}, \quad (3.46f)$$

$$\kappa_4=\frac{\mu_o\sigma_2\sqrt{n_4}}{2j\beta_2\sqrt{n_2 n_3}}. \quad (3.46g)$$

We solved the above equations using the finite-difference method for the case of

two 1ps input Gaussian pulses at λ_2 = 1.55μm and λ_3 = 1.5μm. The two pulses were assumed to have peak field amplitude of 1.18×10^7V/m and were coincident in time.

Fig. 3.12(a) shows the normalized temporal responses of the two input pulses and the sum and difference frequency pulses. The sum and difference frequency pulses have the same Gaussian shape as the input pulses, which is similar to the frequency mixing processes based on χ^2 nonlinearity [66]. However, their pulse widths are found to be 0.707 times smaller than the width of the two input pulses. This bandwidth broadening is a result of the second order conductive processes and will be further investigated in the next section.

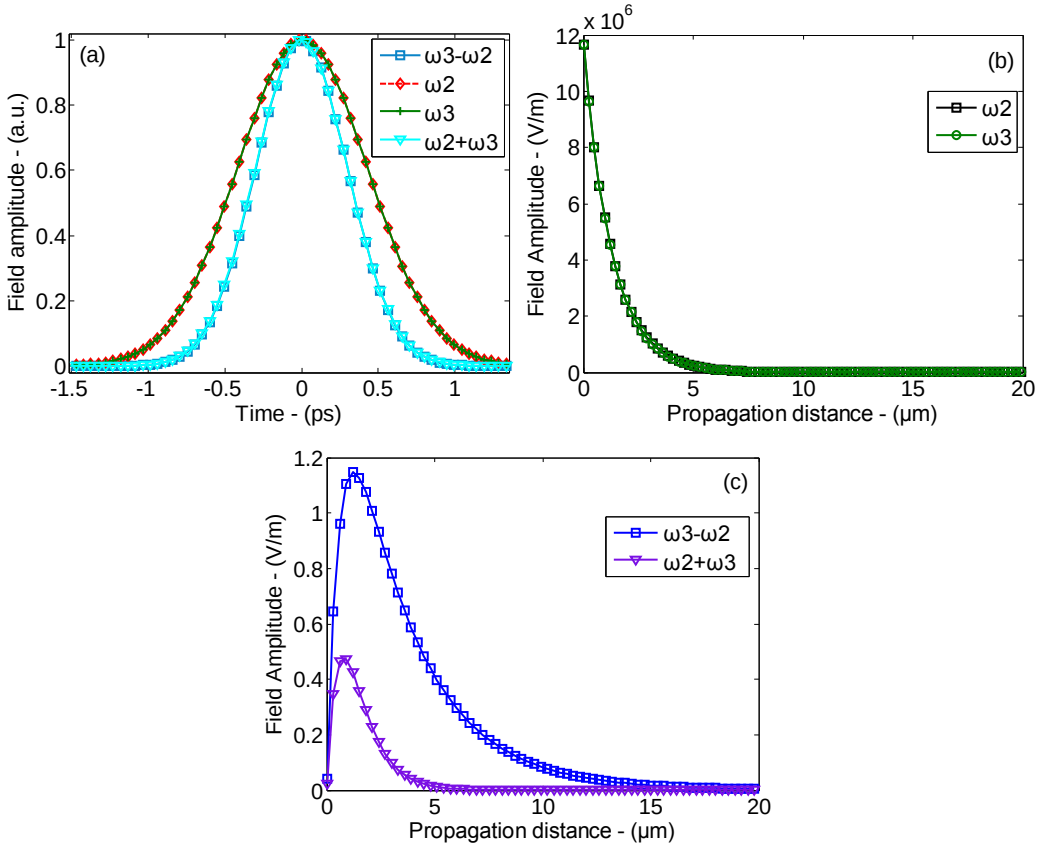


Figure 3.12 Plots of (a) Temporal responses for the ω_3-ω_2, ω_2, ω_3, and ω_2+ω_3 pulses, (b) electric field amplitudes for the ω_2 and ω_3 and (c) the ω_3-ω_2 and ω_2+ω_3 pulses as functions of propagation distance in the MIM waveguide with a 3nm insulator layer. Two 1ps Gaussian pulses with carrier wavelength λ_2 = 1.55μm and λ_3 = 1.5μm and peak field amplitude of 1.18×10^7V/m are assumed.

Fig. 3.12(b) and 3.12(c) plot the peak field amplitudes of the four pulses as

a function of the propagation distance, indicating optimum lengths where the sum and difference frequency pulses achieve their maxima. The frequency difference pulse has longer optimum length than the frequency sum pulse due to much lower loss at the lower frequencies. The above analytical results show that the MIM nanoplasmonic waveguide can potentially be used to generate pulses at different carrier frequencies and pulse widths.

3.5.3 Pulse bandwidth analysis in MIM plasmonic waveguides

MIM nanoplasmonic waveguides exhibit strong dispersion from microwave to the optical frequencies. However, over a narrow frequency range the propagation constant changes only slightly. For example, over a frequency span of 2.5THz around the 1.55μm wavelength, the MIM waveguide has a maximum group index change of only 0.0075% and a maximum change of 0.08% in the attenuation constant. The effect of this dispersion is negligible over a propagation length of 1.27μm. Similar analysis at the second harmonic wavelength (0.775μm) also shows negligible dispersion over a THz pulse bandwidth. As a result, the MIM waveguide can support ultrafast optical pulse propagation with duration shorter than a picosecond, with negligible distortion on the pulse shape.

Based on the above analysis, we can use the CMT equations for CW signals (Eqs. 3.34 and 3.35) to estimate the bandwidth change of nonlinear pulse propagation in the MIM waveguide. A Gaussian pulse envelope can be expressed as

$$E_i(t, x=0) = E_i e^{-\frac{\rho t^2}{\tau_i^2}}, \tag{3.47a}$$

$$\rho = 4\log_e(2), \tag{3.47b}$$

where τ_i is the FWHM width of the Gaussian pulse at carrier frequency ω_i. By substituting Eq. 3.47a into Eqs. 3.34 and 3.35, we obtain

$$E_1(x,t) \propto E_2(0) e^{-\frac{\rho t^2}{\tau_2^2}} E_3(0) e^{-\frac{\rho t^2}{\tau_3^2}} = E_2(0) E_3(0) e^{-\frac{\rho t^2}{\tau_2^2 \tau_3^2/(\tau_2^2+\tau_3^2)}}, \tag{3.48a}$$

$$E_2(x,t) \propto E_2(0)e^{-\frac{\rho t^2}{\tau_2^2}}, \tag{3.48b}$$

$$E_3(x,t) \propto E_3(0)e^{-\frac{\rho t^2}{\tau_3^2}}, \tag{3.48c}$$

$$E_4(x,t) \propto E_2(0)e^{-\frac{\rho t^2}{\tau_2^2}} E_3(0)e^{-\frac{\rho t^2}{\tau_3^2}} = E_2(0)E_3(0)e^{-\frac{\rho t^2}{\tau_2^2\tau_3^2/(\tau_2^2+\tau_3^2)}}. \tag{3.48d}$$

The above expressions show that the input pulses at frequencies ω_2 and ω_3 maintain their pulse widths (τ_2 and τ_3, respectively) as they propagate in the MIM nanoplasmonic waveguide. On the other hand, the frequency difference and the frequency sum pulses have their pulse widths given by $\frac{\tau_2\tau_3}{\sqrt{\tau_2^2+\tau_3^2}}$. If the two input pulses have equal pulse widths of $\tau_2 = \tau_3 = \tau$, then the frequency difference and the frequency sum pulses experience pulse width shrinking by a factor of $\sqrt{2}$ (or bandwidth expansion factor of $\sqrt{2}$). This prediction agrees with the observation in Section 3.5.2 for nonlinear pulses mixing where the frequency difference and frequency sum pulses are observed to have a bandwidth expansion factor of $\sqrt{2}$.

3.6 Summary

In this section we presented approximate theoretical analysis of nonlinear wave propagation in MIM nanoplasmonic waveguides based on the coupled wave equations in a 1D effective medium with second order nonlinear conductivity. Analysis showed that the MIM waveguide can rectify CW signals from sub-THz to the infrared frequencies with much higher responsivity than lumped MIM diodes. The MIM waveguide was also shown to be capable of rectifying ultrafast optical pulses leading to the generation of single cycle THz pulses, making the device potentially useful for on-chip THz source generation. In addition, second order nonlinear processes such as second harmonic frequency generation and sum and difference frequency generation were also analytically investigated. In contrast to conventional second order processes in $\chi^{(2)}$ materials, frequency

conversion in the MIM waveguide was found to be primarily limited by waveguide loss rather than velocity mismatch. In the next chapter we will perform rigorous numerical analysis of nonlinear MIM nanoplasmonic waveguides to validate the analytical results obtained in this chapter.

Chapter 4

FDTD Investigation of Nonlinear MIM Nanoplasmonic[2] Waveguides

4.1 Introduction

In chapter 3 we theoretically investigated the characteristics of MIM nanoplasmonic waveguides for optical rectification and frequency generation using the Coupled Mode Theory (CMT). The CMT analysis is only approximate since it assumes wave propagation in an effective 1D medium and considers nonlinear interaction between only three waves. For a more accurate analysis, we use a numerical method to obtain rigorous solutions of Maxwell's equations in the nonlinear MIM waveguide in this chapter. Many numerical methods have been extensively used in the study of electromagnetic wave propagation in waveguides. Among them, the finite element method (FEM) [30], the finite-difference time-domain method (FDTD) [67] and the method of moments (MOM) [69] are three powerful tools in electromagnetic computations.

The FDTD method, first introduced by Yee in 1966, provides direct time-domain solution to Maxwell's equations in the differential form. The numerical solution requires a spatial discretization of the volumetric region containing the materials and structures which interact with the electromagnetic waves. With the increasing availability of computational power and memory, the FDTD method has become a powerful and widely-used method in computational electromagnetics since it is simple to implement and rigorously solves Maxwell's equations without making any assumptions or approximations [69]. As a result the FDTD solution captures all the physics associated with electromagnetic field interaction with the material in real time, which is especially important for

[2] Part of this work has been published by X. Lei et al. in IEEE Photonics Conference (IPC) 2012, pp. 953-954, Sept. 2012 and presented in IEEE Photonics Conference 2012, Nanoplasmonic II session, Sep 23-27, 2012, Burlingame, CA, USA. Part of this work will appear in the Journal of Optics Communications.

nonlinear wave propagation analysis. For this reason we choose to analyze the nonlinear MIM nanoplasmonic waveguide using the FDTD method.

In this chapter we will first present the formulation and implementation of a nonlinear FDTD scheme for simulating CW and pulse propagation in a MIM plasmonic waveguide. The performance of the device for frequency conversion and rectification of CW and pulsed signals will be numerically investigated. The numerical results are also used to validate the approximate results obtained from the Coupled Mode Theory in chapter 3.

The outline of the rest of this chapter is as follows. Section 4.2 presents a formulation of the nonlinear FDTD method for simulating MIM nanoplasmonic waveguides in the presence of electron tunneling. Section 4.3 investigates the performance of the MIM nanoplasmonic waveguide for rectifying CW radiation from THz to infrared frequencies. In Section 4.4, we study the impulse response of the MIM nanoplasmonic waveguide and its electrical bandwidth performance for optical pulse detection. In Section 4.5, we demonstrate the performance of the device for ultrafast pulse rectification and second harmonic generation. In Section 4.6, two optical pulses modulating different carrier frequencies are used as excitation sources for the investigation of nonlinear pulses mixing in the MIM nanoplasmonic waveguide. Section 4.7 presents additional applications of nonlinear MIM nanoplasmonic waveguides. Concluding remarks are given in Section 4.8.

4.2 FDTD Method for Simulating MIM Nanoplasmonic Waveguides

The MIM nanoplasmonic waveguides we are interested in simulating have very thin insulator thickness (3-5nm) and much larger waveguide width (50–100nm). Since the fields of the fundamental TM mode are approximately constant along the waveguide width (the y-direction in Fig. 3.1(a)), we can approximate the 3D MIM waveguide structure by a 2D slab waveguide as shown in Fig. 3.1(b). The FDTD method is applied to solve for the fields E_x, E_z and H_y in the metal layers

and the nonlinear conductive insulator gap as described below. Other numerical issues such as stability, absorbing boundary conditions and field marching procedure of the FDTD algorithm are given in this section and Appendix D.

In a 2D medium with permittivity ε and permeability μ, Maxwell's equations for TM waves with field components E_x, H_y, E_z are

$$\mu \frac{\partial H_y}{\partial t} = -(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}), \tag{4.1a}$$

$$\varepsilon \frac{\partial E_x}{\partial t} = -\frac{\partial H_y}{\partial z} - J_x, \tag{4.1b}$$

$$\varepsilon \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} - J_z. \tag{4.1c}$$

The current densities J_x and J_z represent either the ohmic currents in the metal layers or the tunneling current in the insulator gap. The above equations are discretized using Finite Difference formulas on a rectangular grid of size Δx and Δz. We use Yee's method of field assignments where the E and H field components are offset from each other by half a grid size and half a time step. Each Yee's cell for the TM polarization is shown in Fig. 4.1. The discrete equations are:

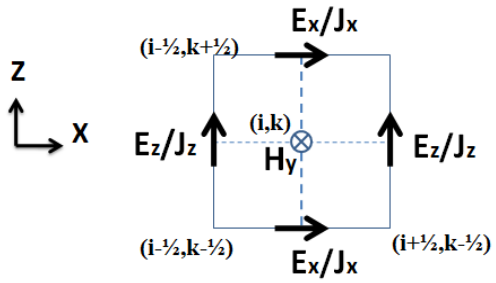


Figure 4.1 Schematic of a 2D TM Yee-cell with fields E_x, E_z, H_y and cell size Δx by Δz.

$$\mu(i,k)\frac{H_y^{n+1/2}(i,k) - H_y^{n-1/2}(i,k)}{\Delta t} = -[\frac{E_x^n(i,k+\frac{1}{2}) - E_x^n(i,k-\frac{1}{2})}{\Delta z} - \frac{E_z^n(i+\frac{1}{2},k) - E_z^n(i-\frac{1}{2},k)}{\Delta x}], \tag{4.2a}$$

$$\varepsilon(i,k+\frac{1}{2})\frac{E_x^{n+1}(i,k+\frac{1}{2}) - E_x^n(i,k+\frac{1}{2})}{\Delta t} = -\frac{H_y^{n+1/2}(i,k+1) - H_y^{n+1/2}(i,k)}{\Delta z} - J_x^{n+\frac{1}{2}}(i,k+\frac{1}{2}), \tag{4.2b}$$

$$\varepsilon(i+\frac{1}{2},k)\frac{E_z^{n+1}(i+\frac{1}{2},k)-E_z^{n}(i+\frac{1}{2},k)}{\Delta t}=\frac{H_y^{n+1/2}(i+\frac{3}{2},k)-H_y^{n+1/2}(i+\frac{3}{2},k)}{\Delta x}-J_z^{n+\frac{1}{2}}(i+\frac{1}{2},k),$$

(4.2c)

The above equations are then solved for the most recent field components, $E_x(n+1)$, $E_z(n+1)$ and $H_y(n+1/2)$, in terms of fields at the previous time steps. These equations are used in the FDTD algorithm to march the TM fields in time from an initial field distribution.

We now describe how the current densities J_x and J_z in Eqs. 4.2 are computed. Since metals are highly dispersive media, we incorporate the frequency dependent permittivity of the metal layers into the FDTD equations using the Auxiliary Differential Equation (ADE) method [69].

We assume that the metal layers have relative complex permittivity ε_r obeying the Drude model [69]

$$\varepsilon_r(\omega)=\varepsilon_\infty-\frac{\omega_p^2}{\omega(\omega+j\Gamma_e)},\tag{4.3}$$

where $\varepsilon_\infty \sim 1$ is the relative permittivity at infinite frequency, ω_p is the plasma frequency, Γ_e is the relaxation coefficient and ω is the angular frequency. In the frequency domain, the constitutive relation for the electric flux density is

$$D_{(\omega)}=\varepsilon_r(\omega)\varepsilon_o E_{(\omega)}=\varepsilon_\infty\varepsilon_o E_{(\omega)}-\frac{\omega_p^2}{\omega(\omega+j\Gamma_e)}\varepsilon_o E_{(\omega)}.\tag{4.4a}$$

We define the polarization P as

$$P_{(\omega)}=-\frac{\omega_p^2}{\omega(\omega+j\Gamma_e)}\varepsilon_o E_{(\omega)},\tag{4.4b}$$

which can be rearranged to give

$$P_{(\omega)}(-j\omega)^2+P_{(\omega)}(-j\omega)\Gamma_e=\omega_p^2\varepsilon_o E_{(\omega)}.\tag{4.4c}$$

Taking the inverse Fourier transform of the above equation, we obtain

$$\frac{\partial^2 P}{\partial t^2}+\Gamma_e\frac{\partial P}{\partial t}=\varepsilon_o\omega_p^2 E,\tag{4.5}$$

Next defining the polarization current

$$J_p = \frac{\partial P}{\partial t}, \tag{4.6}$$

Eq. 4.5 becomes

$$\frac{\partial J_P}{\partial t} + \Gamma_e J_P = \varepsilon_o \omega_p^2 E, \tag{4.7}$$

The above equation is the Auxiliary Differential Equation relating the current in a dispersive medium to the electric field. In the metal layers of the MIM waveguide, the currents J_x and J_z are related to the fields E_x and E_z by

$$\varepsilon_o \omega_p^2 E_z = \frac{\partial J_z}{\partial t} + \Gamma_e J_z, \tag{4.8a}$$

$$\varepsilon_o \omega_p^2 E_x = \frac{\partial J_x}{\partial t} + \Gamma_e J_x. \tag{4.8b}$$

The above equations are discretized using the Finite Difference method and solved for the current components $J_x(n+1/2)$ and $J_z(n+1/2)$ that are used in Eqs. 4.2.

In the insulator layer the current density J_z represents the electron tunneling current while J_x is set to zero. Treating the insulator layer as a medium with nonlinear conductivity, the constitutive relation for the current density J_z as a function of the transverse electric field E_z across the insulator gap is given by

$$J_z(E_z) = \sigma_1 E_z + \sigma_2 E_z^2. \tag{4.9}$$

The conductivities σ_1 and σ_2 are obtained from the J-V curve of the MIM junction at a given bias voltage and illumination frequency.

In the FDTD simulations, absorbing boundary conditions (ABCs) are used to terminate the computational domain. Many ABCs have been developed for FDTD simulations, with the Perfectly Matched Layer (PML) technique being the most popular [69] and the most effective for preventing numerical back reflections into the computational domain. For this reason, we choose the PML absorbing boundary conditions for our FDTD simulations. Detailed formulation of the PML boundary condition over the MIM plasmonic waveguide structure is given in Appendix D.

In a standard linear FDTD simulation, the maximum time step Δt that can

be used is determined by the well-known Courant-Friedrichs-Lewy (CFL) stability condition [69]:

$$\Delta t_{\max} = \frac{\sqrt{\varepsilon_r}}{c\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta z)^2}}}. \tag{4.10}$$

where c is the speed of light in vacuum and ε_r is the smallest relative permittivity value in the computational domain.

The above CFL condition, however, is typically not valid for nonlinear FDTD algorithms such as our method. To determine a suitable time step for our algorithm, we performed simulations of the MIM structure with decreasing time steps from the maximum CFL value until the change in the output field amplitude was observed to be below 10^{-8} compared to the previous simulation run. The final time step chosen for our simulations is about 70% of the maximum CFL condition in vacuum ($\varepsilon_r = 1$).

$$\Delta t = \frac{0.7}{c\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta z)^2}}}. \tag{4.11}$$

In our nonlinear FDTD simulation, the grid dimensions are chosen as Δx = 3nm and Δz = 1.5nm. The MIM waveguide has a width of 403nm and length of 20μm. The simulation time step is then calculated using Eq. 4.11.

The remainder of this chapter is devoted to presenting and discussing the results of our numerical investigation of nonlinear Al-Al_2O_3-Al MIM plasmonic waveguides. The waveguide structure consists of a 3nm Al_2O_3 layer sandwiched between two Al layers of 200nm thickness. The relative permittivity of Al_2O_3 is 4.29 while the permittivity of Al is given by the Drude model with parameters ε_∞ = 1, $\omega_p = 3.57\times10^{16}$ Hz and $\Gamma_e = 1.979\times10^{14}$ Hz [59-60]. We also assume conductivity parameters $\sigma_1 = 4\times10^{-2}\Omega^{-1}$/m and $\sigma_2 = 3.5\times10^{-10}\Omega^{-1}$/V for the Al-$Al_2O_3$-Al junction biased at 0.1V. In the nonlinear FDTD simulations, the grid dimensions are chosen to be Δx = 3nm and Δz = 1.5nm, and the time step is set at 0.7 of the maximum CFL value for vacuum.

4.3 Rectification of Continuous-wave Signals

We first investigated the performance of the MIM nanoplasmonic waveguide for rectifying CW signals by exciting it with a sinusoidal optical signal at 1.55μm wavelength. The average power of the signal was set to be 0.362mW, which corresponds to a peak E_z field strength of 3.33×10^7V/m, yielding a peak voltage drop of 0.1V across the insulator layer. Fig. 4.2 shows the distributions of the E_x, E_z and H_y fields in the MIM waveguide. The transverse electric field E_z is observed to be strongly localized and enhanced within the insulator layer, which allows efficient rectification to be achieved over a short propagation distance. The longitudinal electric field E_x and the magnetic field H_y are less strongly confined in the waveguide, exhibiting more gradual decay into the metal layers. In addition, the power distribution in the waveguide is also highly localized within the insulator layer.

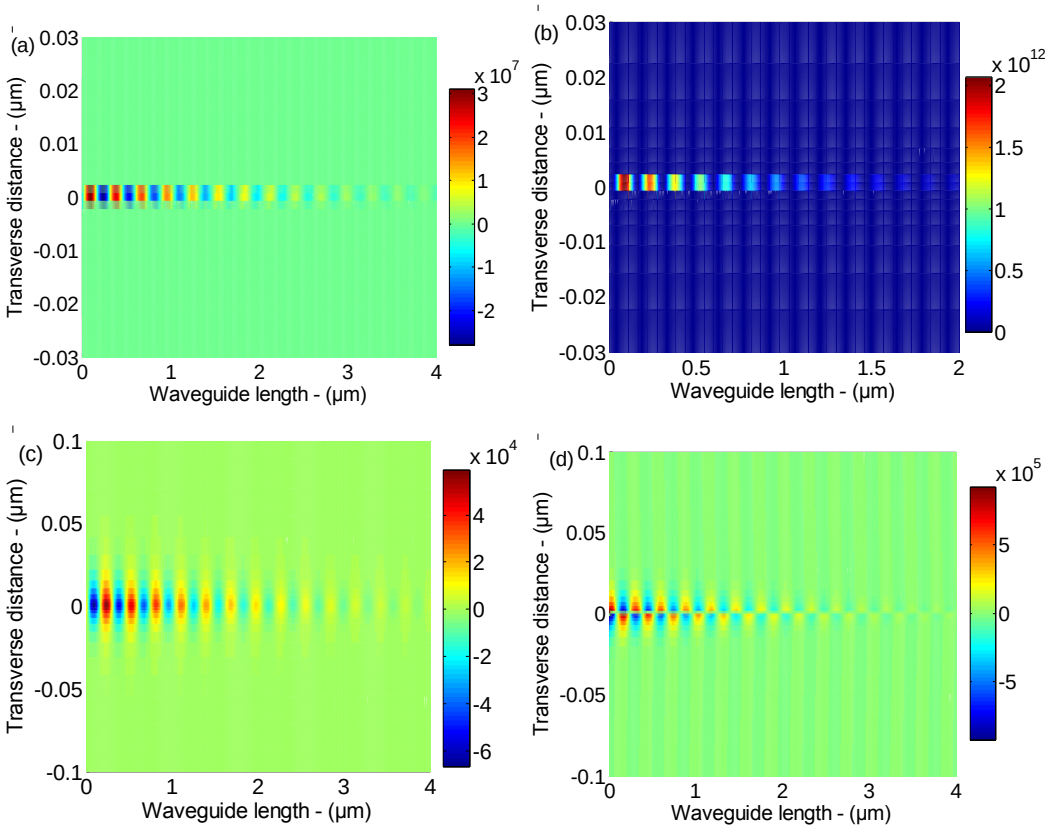


Figure 4.2 FDTD simulated field and power distributions (at 1.55μm wavelength) in the MIM nanoplasmonic waveguide: (a) E_z (V/m), (b) P_x (W/m^2), (c)H_y (A/m) and (d) E_x (V/m).

Fig. 4.3 shows a plot of the E_z field amplitude of the input CW signal along the propagation direction along with the total rectified DC current. The total rectified current is obtained by integrating the tunneling current density over the length of propagation,

$$I_{dc}(x) = \frac{w\sigma_2}{2}\int_0^L |E_z(x)|^2 dx, \tag{4.12}$$

where the MIM waveguide width w is assumed to be 100nm. It is observed from the plot that the E_z field amplitude decays exponentially with the propagation distance as a result of both absorption in the metal layers and optical power depletion due to rectification. The simulated propagation length L_p is determined to be 1.27μm, which agrees with the value obtained from the propagation constant β of the MIM waveguide. The decay in the E_z field causes the rectified DC current to initially increase with the propagating distance before gradually becoming saturated. This saturating characteristic suggests that there is an optimum length for achieving maximum responsivity in the MIM waveguide rectifier. From the plot in Fig. 4.3, the optimum length can be taken to be around 3.8μm, or roughly about three times the propagation length of the plasmonic waveguide, $3L_p$.

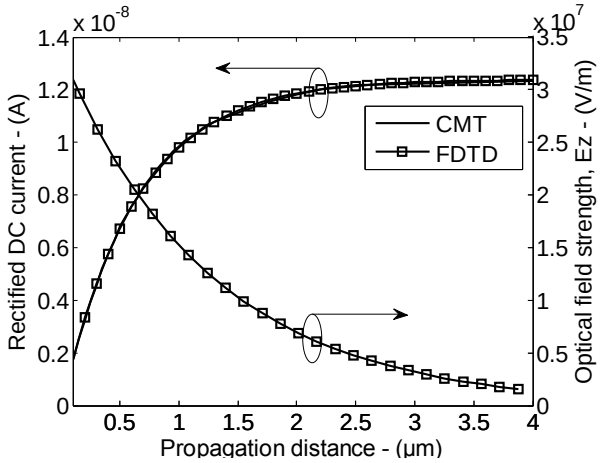


Figure 4.3 Plot of the rectified DC current and the transverse optical field strength E_z along the propagation direction of the MIM nanoplasmonic waveguide for an input CW optical signal at 1.55μm wavelength and peak field amplitude of 3.33×10^7V/m.

We next investigated the frequency dependence of the MIM waveguide rectifier over the frequency range from Terahertz to the infrared. We varied the

frequency of the input CW signal from 2THz to 200THz (around 1.55μm wavelength) and performed simulation to determine the rectified current. The average power of the input CW signal was kept at 0.362mW. Regardless of the frequency of the CW signal, the rectified DC current was observed to saturate at a propagation distance around $3L_p$. To quantify the results, we calculated the responsivity of the MIM waveguide rectifier assuming unity coupling efficiency from the antenna to the nanoplasmonic waveguide. The frequency dependence of the responsivity is plotted in Fig. 4.4. The theoretical response obtained in Chapter 3 is also plotted (black dashed line), which shows very agreement with the simulation results.

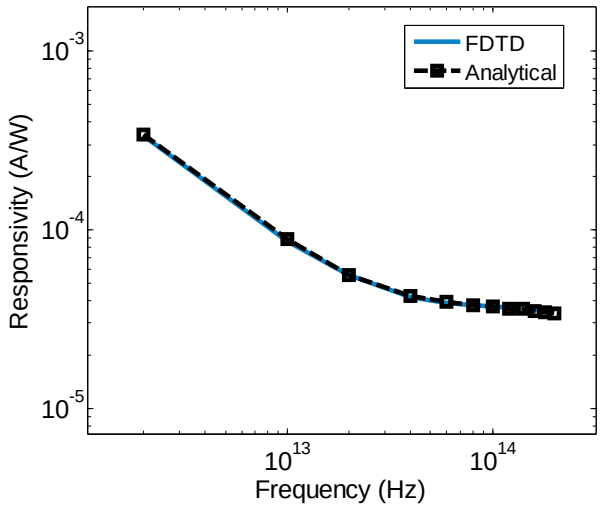


Figure 4.4 Plot of frequency response responsivity of the Al-Al_2O_3-Al nanoplasmonic waveguide detector with a 3nm insulator layer. Also given are analytical results obtained in chapter 3.

Overall, three key points can be drawn from the plot in Fig. 4.4. First, the plot shows that the MIM waveguide rectifier has an extremely broad optical bandwidth for detecting CW signals from THz to the optical frequencies. This broad bandwidth is due to the broadband nature of the MIM waveguide and the fast electron tunneling time of the MIM junction. Second, the responsivity of the device does decrease with increasing frequency, but this is a result of the higher propagation loss of the MIM at shorter wavelengths and not due to some inherent cut off condition of the MIM waveguide or the rectifying process. And third, our approximate theoretical analysis based on nonlinear wave propagation in an

effective 1D medium gives fairly accurate results as validated by the FDTD simulations.

4.4 Impulse Response of the MIM Waveguide Detector

The optical bandwidth of the MIM waveguide detector, as shown by the plot in Fig. 4.4, gives the range of CW radiation frequencies that can be detected or rectified by the device. If the illuminating signal is modulated by a pulse envelop, then it is also important to know the electrical bandwidth of the detector, which specifies how fast a pulse can be detected or rectified. For THz applications, the electrical bandwidth of the detector must be in the THz range. This is beyond the capability of most semiconductor photodiodes, whose electrical bandwidths are limited by the slow electron transit time (10^{-11}s [52]) across the depletion region.

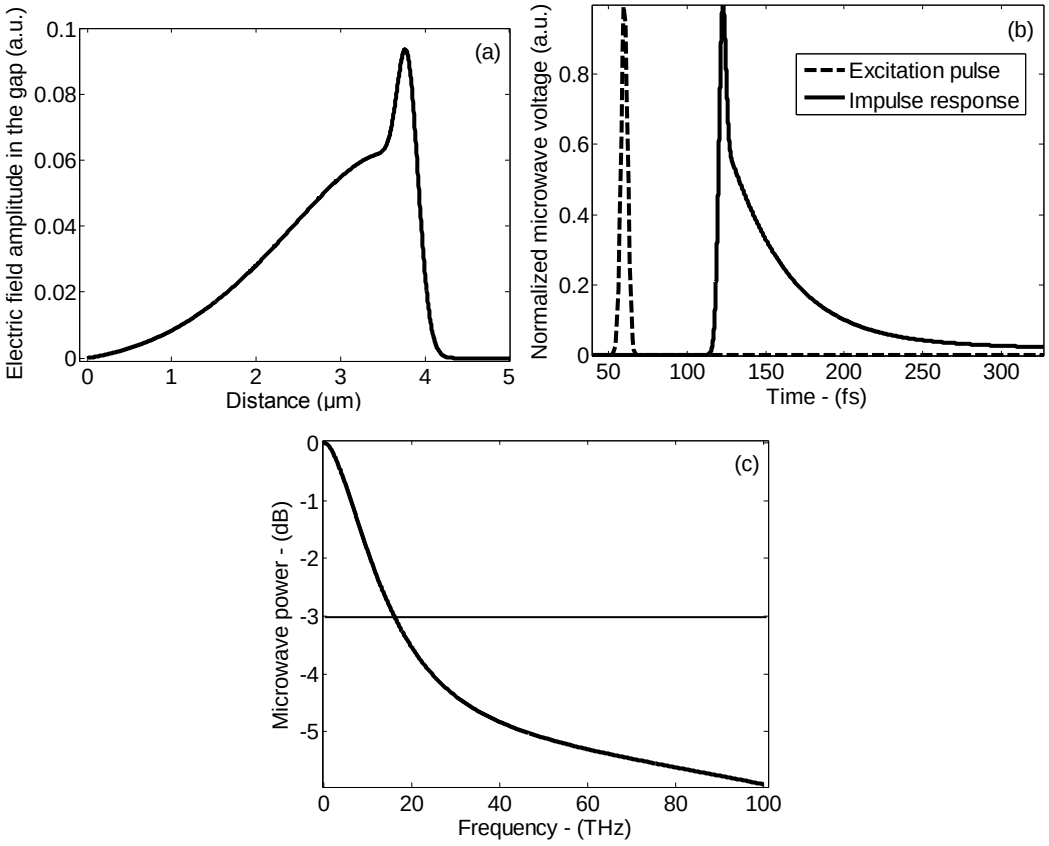


Fig. 4.5 Plots of (a) instantaneous transverse E_z field profile of the rectified pulse as a function of propagation distance, (b) voltage impulse response (at a distance of 1.27μm) and (c) frequency impulse response of the MIM nanoplasmonic waveguide after simulation time of 150fs [61].

The electrical bandwidth of a diode can be determined from the Fourier transform of its impulse response. To determine the impulse response of the MIM waveguide rectifier, we simulated the structure by exciting it with a very short optical pulse of 5fs duration. We recorded the field propagating along the waveguide as well as the temporal response of the voltage across the insulator gap at a propagation distance of L_p (at 1.55μm wavelength). Fig. 4.5(a) shows a snap shot of the instantaneous field distribution along the waveguide after 150fs. The long tail of the pulse is due to waveguide dispersion, which has the effect of broadening the pulse. Physically, the long tail is caused by frequency components which propagate at slower phase velocities. The temporal response of the rectified voltage pulse at a propagation length of L_p (1.55μm wavelength) is shown in Fig. 4.5(b) along with the input pulse for reference. Again strong waveguide dispersion is manifested in the long tail of the rectified pulse.

The Fourier transform of the impulse response is plotted in Fig. 4.5(c). The 3dB bandwidth of the spectral response is determined to be 16THz, indicating that the MIM waveguide can rectify ultrafast optical pulses with durations as short as 65fs. Since the electron tunneling process is assumed to be instantaneous in our simulations, it is not the limiting factor for the response of the MIM waveguide detector. Instead, the electrical bandwidth of the device is solely caused by the strong dispersion of the MIM waveguide, which causes some frequency components of the pulse to travel more slowly than others, resulting in the temporal pulse spread seen in Fig. 4.5(b). To achieve even broader electrical bandwidth, one can minimize the waveguide dispersion by proper design of the waveguide dimensions and using a less dispersive metal.

4.5 Rectification and Second Harmonic Generation of Optical Pulses

In the CMT analysis in chapter 3, we showed that when an MIM nanoplasmonic waveguide is excited with a pulsed signal at a carrier frequency of ω, single-cycle THz and second harmonic pulses are generated from the second order nonlinear

processes. To verify the theoretical analysis, we perform rigorous numerical investigation of the MIM nanoplasmonic waveguide under pulsed excitation in this section. These numerical simulations serve as important validation of the CMT results. They also provide more accurate assessment of the performance of the MIM waveguide for single-cycle THz and second harmonic pulses generation, especially for very short pulses where the effects of strong waveguide dispersion are not accounted for in the CMT analysis. In addition, we will also investigate the effects of the carrier frequency and pulse width on the efficiencies of the rectification and frequency conversion processes.

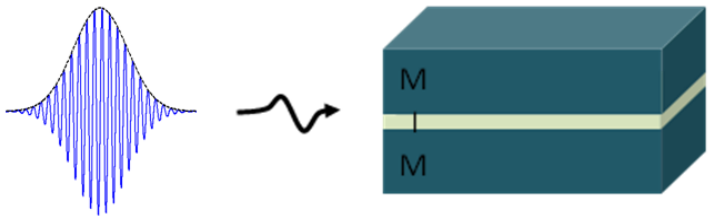


Fig. 4.6 Illustration of the MIM nanoplasmonic waveguide excited by a 1ps Gaussian pulse with carrier wavelength of 1.55μm and peak power of 0.65mW and biased at 0.1V.

In the simulation the MIM waveguide was excited by a 1ps Gaussian pulse modulating a 1.55μm wavelength optical carrier, as illustrated in Fig. 4.6. The peak power of the input pulse was set to 0.65mW, which corresponds to a peak voltage drop of 0.1V across the insulator gap. We also assumed a bias voltage of 0.1V is applied to the MIM waveguide. We recorded the temporal variation of the electric field E_z across the insulator gap after a propagation distance of L_p (at 1.55μm wavelength). We then performed Fourier transform on the E_z field and filtered out the different harmonic components. Taking the inverse Fourier transform of the harmonic components gives their time domain waveforms.

Fig. 4.7(a) and 4.7(b) show the temporal and spectral shapes of the rectified pulse and the 2ω pulse as compared to the input pulse. Several key observations can be made from the plots. Firstly, we observe that the input pulse maintains its Gaussian shape and FWHM bandwidth as it propagates along the MIM waveguide due to the low amplitude and phase dispersion near the 1.55μm wavelength. The simulation verifies that the MIM nanoplasmonic waveguide can

support ultrafast pulse propagation with negligible distortion. Secondly, the 2ω pulse is also found to have a Gaussian shape but its FWHM bandwidth is increased by a factor of $\sqrt{2}$ compared to the input pulse. This bandwidth expansion is a result of the second order nonlinear conductive process which is theoretically analyzed in Section 3.5.3. Thirdly, the rectified pulse has the shape of the derivative of the input Gaussian pulse, which yields a single-cycle THz pulse. However, there is more pronounced difference between the simulation and CMT results for the rectified pulse than the results for ω and 2ω pulses. The reason for the deviation can be attributed to the strong waveguide dispersion at the baseband frequencies which is not adequately taken into account in the CMT analysis. This example highlights a limitation of our theoretical model.

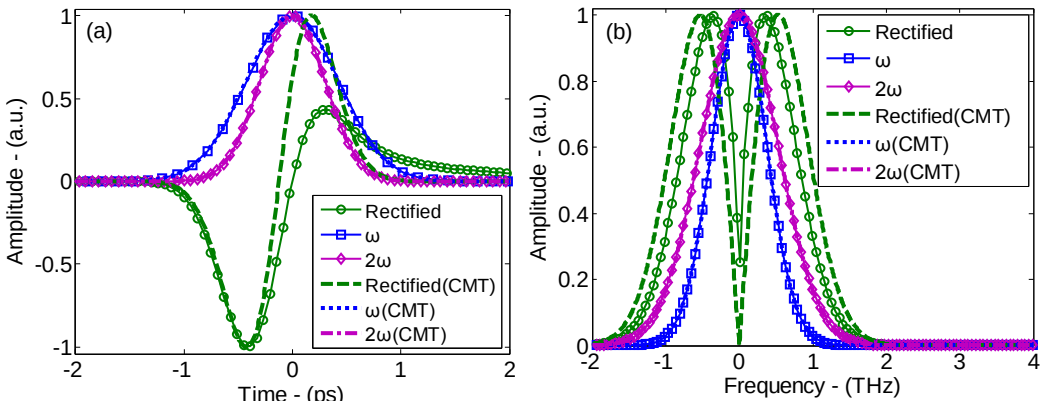


Figure 4.7 Plots of (a) temporal and (b) spectral shapes of the rectified, 2ω and input pulses after a propagation distance of 1.27μm in the MIM nanoplasmonic waveguide. The input 1ps Gaussian pulse modulates a carrier wavelength of 1.55μm and has peak power of 0.65mW. Results from the CMT analysis in chapter 3 are also given for comparison [61].

Fig. 4.8 shows the normalized peak field amplitudes of the three pulses as functions of the propagation distance, indicating the optimum waveguide lengths where the rectified and 2ω pulses achieve their maxima. It is seen that the rectified pulse has a longer optimum length than the 2ω pulse due to much lower loss at lower frequencies. The plot in Fig. 4.11 can be used to determine the optimum MIM waveguide length for achieving maximum conversion efficiency in on-chip generation of single cycle THz pulse or 2ω pulse.

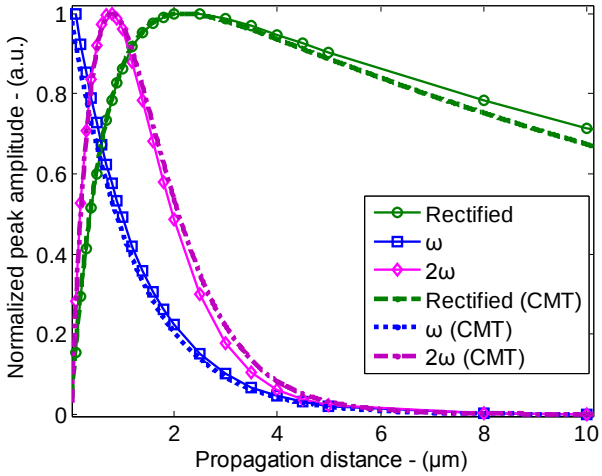


Figure 4.8 Plot of normalized peak field amplitudes of the rectified, 2ω and input pulses as functions of the propagation distance in the MIM nanoplasmonic waveguide [61].

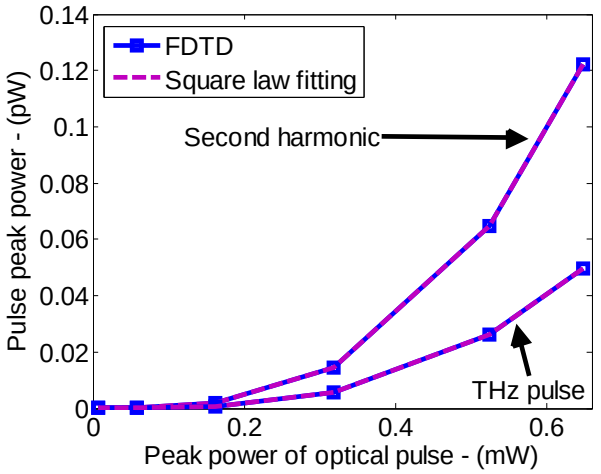


Figure 4.9 Peak powers of both the rectified and 2ω pulses as functions of the peak power of the input pulse [61].

In Fig. 4.9, we plotted the peak power of the rectified and 2ω pulses as functions of the peak power of the input pulse. The peak power of both the rectified and 2ω pulses show a quadratic dependence on the peak power of the input pulse. This power dependence is typical of second order processes and is also observed in traveling-wave photodetectors [51]. We also observe that the 2ω pulse has a higher peak power than the rectified pulse, indicating that the generation of the 2ω pulse is more efficient. This result is caused by the strong waveguide dispersion (between THz and ω frequencies) which effectively reduces the conversion efficiency.

To further study the characteristics of the MIM waveguide for THz and second-harmonic pulse generation, we performed simulations of the structure for different input carrier frequencies and input pulse widths. These results reflect the optical and electrical bandwidth performance of the device and are summarized below.

4.5.1 Effect of input carrier frequency (optical bandwidth)

We varied the carrier frequency of the input pulse from 10THz to 200THz while keeping the input pulse width constant at 1ps and pulse peak power at 0.65mW.

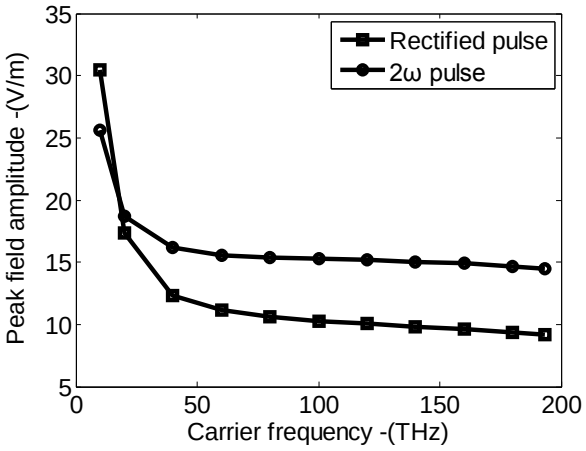


Figure 4.10 Plot of peak field amplitudes of both the rectified and 2ω pulses as functions of carrier frequency. The input pulse has a FWHM width of 1ps.

Fig. 4.10 shows the peak amplitudes of the rectified and 2ω pulses as functions of the input carrier frequency. It can be seen that the peak field amplitudes of both the rectified and 2ω pulses increase with decreasing carrier frequencies due to lower waveguide loss at lower frequencies. The simulation results indicate that the frequency conversion processes in the MIM waveguide are dominated by loss rather than phase mismatch, as is commonly the case in $\chi^{(2)}$ nonlinear processes. The result is consistent with the frequency-dependent responsivity obtained for CW rectification in Fig. 4.4.

4.5.2 Effect of input pulse width (electrical bandwidth)

We also studied the effect of the input pulse width on the performance of the MIM waveguide. We fixed the carrier wavelength at 1.55μm and peak pulse power at 0.65mW, and varied the input pulse width from 100fs to 1ps. Figs. 4.11(a) and 4.11(b) show the peak field amplitudes and the optimum waveguide lengths, respectively, of the rectified and 2ω pulses as functions of the input pulse width.

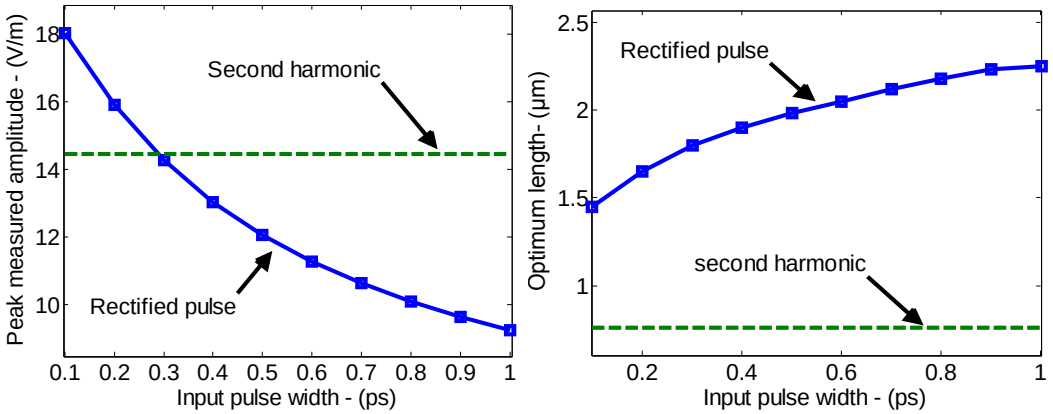


Figure 4.11 Plot of (a) peak field amplitudes and (b) optimum lengths of the rectified and 2ω pulses as functions of input pulse FWHM width in the MIM nanoplasmonic waveguide.

We observe that the peak field of the rectified pulse increases slightly with decreasing input pulse width while the 2ω pulse stays constant. In addition, the optimum propagation length of the rectified pulse decreases slightly with decreasing input pulse width while the 2ω pulse stays constant. The behavior of the rectified pulse is caused by peak spectral shifting towards higher frequencies with decreasing input pulse width while the peak spectral of the 2ω pulse stays at the 2ω frequency. As the peak spectral of the rectified pulse shifts towards higher frequencies, higher effective loss will be experienced by the rectified pulse and thus the optimum propagation length is reduced. In addition, we notice that the peak field amplitude of the rectified pulse from Eq. 3.43a indicates that

$$E_{rectified} \propto \frac{\partial |E_{\omega}|^2}{\partial t}. \qquad (\text{since } \frac{\partial |E_{\omega}|^2}{\partial t} >> \frac{\partial |E_{2\omega}|^2}{\partial t}) \qquad (4.13a)$$

For an input Gaussian pulse with envelope of

$$E_{\omega}(t) = Ee^{-\frac{\rho t^2}{\tau_1^2}}, \tag{4.13b}$$

$$\rho = 4\log_e(2), \tag{4.13c}$$

where τ_1 is the FWHM pulse width of the input Gaussian pulse. Eq. 4.13a can be rearranged to give

$$E_{rectified} \propto \frac{|E_{\omega}|^2}{\tau_1^2} \propto \frac{P_{\omega}}{\tau_1^2}, \tag{4.13d}$$

where P_{ω} is the peak power of the input pulse. The above equation indicates that the peak field amplitude of the rectified pulse is inversely proportional to the square of the input pulse width (while the input peak power P_{ω} is unchanged), indicating higher peak field amplitude can be achieved at shorter input pulse width. In addition, the velocity mismatch between the rectified and optical pulses (the ω and 2ω pulses) will be improved as the peak spectral of the rectified pulse shifts towards higher frequencies. However, as the input pulse width decreases, the rectified pulse experiences higher loss. These two effects play together and result in the slowly increasing peak field amplitude of the rectified pulse with decreasing input pulse width. On the other hand, the peak field amplitude of the 2ω pulse from Eq. 3.43c indicates that

$$E_{2\omega} \propto \omega|E_{\omega}|^2 \propto \omega P_{\omega}, \tag{4.14a}$$

since the condition

$$\frac{|E_{\omega}|^2}{\tau_1^2} t << \omega|E_{\omega}|^2, \tag{4.14b}$$

is valid for a 1ps Gaussian pulse modulating a 1.55μm optical carrier and propagates a length of L_p = 1.27μm with propagation time t << 1ps. As a result, both the peak field and optimum propagation length of the 2ω pulse stays nearly constant at different input pulse width since the input peak field amplitude (or input peak power P_{ω}) is unchanged during the simulation and the loss condition of the 2ω pulse stays nearly constant at the 2ω frequency.

4.6 Nonlinear Frequency Mixing of Optical Pulses

In chapter 3 we demonstrated using CMT analysis that frequency sum and difference generation can also be achieved with the second-order nonlinear conductive process in MIM nanoplasmonic waveguides. Here we further investigate these nonlinear frequency mixing processes using numerical simulations, which allow a more detailed and accurate analysis to be obtained. The aim here is to explore the potential use of MIM nanoplasmonic waveguides for nonlinear optical signal processing such as frequency generation, frequency mixing and photonic logic operations.

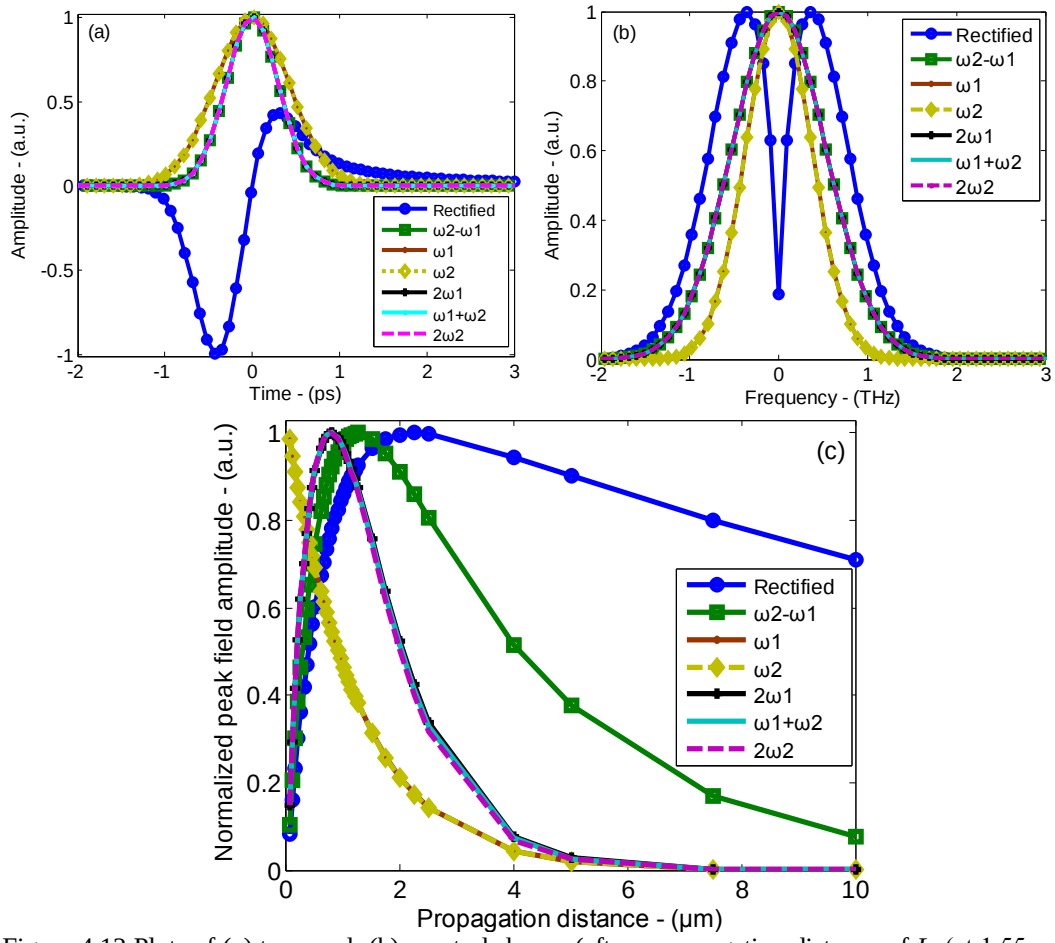


Figure 4.12 Plots of (a) temporal, (b) spectral shapes (after a propagation distance of L_p (at 1.55μm wavelength)) and (c) normalized peak field amplitudes as functions of propagation distance of the rectified, frequency difference, frequency sum, and second harmonic pulses as well as the input 1ps Gaussian pulses with carrier wavelengths of 1.55μm and 1.5μm in the MIM nanoplasmonic waveguide. Similar results are observed in the CMT analysis given in chapter 3 (Fig. 3.12).

In the simulation we excited the MIM nanoplasmonic waveguide with two 1ps Gaussian pulses at carrier wavelengths λ_1 = 1.55μm and λ_2 = 1.5μm. We assumed the two pulses are coincident in time and have the same peak power of 0.08mW. Figs. 4.12(a) and 4.12(b) show the temporal and spectral shapes of the various frequency components generated after a propagation length of L_p (at 1.55μm wavelength) in the waveguide. In addition to the rectified and second harmonic pulses at frequencies $2\omega_1$ and $2\omega_2$, two new pulses appeared at the sum and difference of the input pulse carrier frequencies. These pulses have the same temporal shapes as the input pulses, although their bandwidths are expanded by a factor of $\sqrt{2}$ compared to the bandwidth of the two input pulses. Fig. 4.12(c) shows the normalized peak field amplitudes of all the frequency components as functions of propagation distance in the MIM waveguide. The plots indicate the optimum lengths where the various pulses achieve their maxima. The optimum length is mainly determined by the loss experienced by the different frequency components. Since the waveguide loss decreases with frequency, we have the result that the rectified pulse has the longest optimum length followed by the frequency difference pulse, frequency sum pulse and the two second harmonic pulses.

To further understand the dependence of the sum and difference frequency conversion efficiencies on the powers of the input pulses, we investigated the case where the peak powers of the two input pulses were varied while maintaining the total input power constant. In the simulation we kept the total peak power of the two pulses, P_{total} = P1 + P2, to 0.16mW but varied the ratio P1:P2 from 10%:90% to 90%:10%. The two pulses had the same pulse width of 1ps. Fig. 4.13 shows the peak field amplitudes of the rectified, sum and difference pulses vs. the power ratio P1:P2. It is seen that the peak of the rectified pulse does not change appreciably with respect to the power ratio of the two input pulses, while the peaks of the sum and difference pulses show a quadratic dependence. These behaviours can be explained as follows. The rectified, frequency sum and difference pulses arise from the nonlinear current element. In case of the rectified pulse, Eq. 4.13d is modified to give

$$E_{rectified} \propto \frac{|E_{\omega 1}|^2}{\tau_1^2} + \frac{|E_{\omega 2}|^2}{\tau_2^2} \propto \frac{P_1}{\tau_1^2} + \frac{P_2}{\tau_2^2} = \frac{P_1 + P_2}{\tau^2} = \frac{P_{total}}{\tau^2}. \qquad (4.15a)$$

Since the two input pulses have the same pulse width τ, the peak field of the rectified pulse given by Eq. 4.15a will be constant since the total peak power P1+P2 is kept constant. On the other hand, Eq. 4.14a is modified according to Eq. 3.45a and 3.45d for the sum and difference frequency cases as

$$E_{\omega 2 \pm \omega 1} \propto (\omega 2 \pm \omega 1)|E_{\omega 1} E_{\omega 2}| \propto (\omega 2 \pm \omega 1)\sqrt{P_1}\sqrt{P_2}, \qquad (4.15b)$$

which depends quadratically on the power ratio P1/P2.

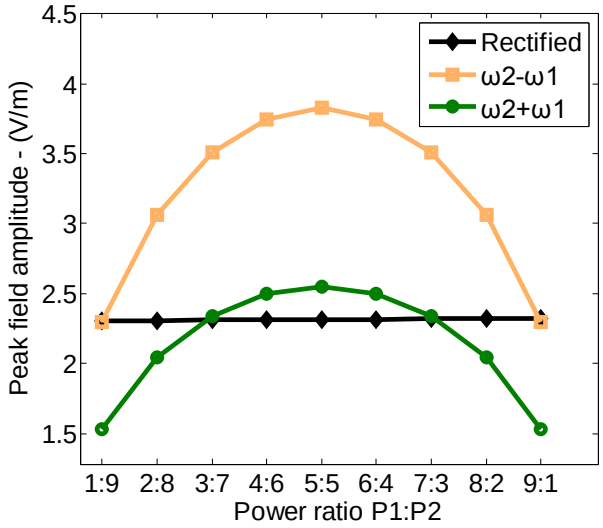


Figure 4.13 Plots of peak field amplitudes of the rectified, frequency difference and frequency sum pulses as functions of the power ratio P1:P2 in the MIM nanoplasmonic waveguide.

4.7 Applications of Nonlinear MIM Nanoplasmonic Waveguides

Our numerical and analytical studies of the MIM waveguide suggests that it can be used for on-chip rectification of ultrafast optical pulses, leading to the generation of single-cycle THz pulses. The structure can also be used for second harmonic frequency generation and frequency mixing of optical pulses. Below we propose two additional applications of MIM nanoplasmonic waveguides based on the nonlinear conductivity derived from electron tunneling.

1) All-optical plasmonic AND gate

As illustrated in Fig. 4.14 the MIM waveguide can be used to perform the logical AND function between two optical pulses I_1 and I_2 at carrier frequencies ω_1 and ω_2. The output O is a pulse at the sum frequency $\omega_1+\omega_2$ (or alternatively, at the difference frequency $\omega_1-\omega_2$). When both or one of the input pulses have weak intensities below the input OFF threshold (0), the sum frequency $\omega_1+\omega_2$ pulse will have weak intensities below the output OFF threshold (0). Only when both the input pulses have high intensities above the input ON threshold (1) can the sum frequency $\omega_1+\omega_2$ pulse have high intensities above the output ON threshold (1). The AND logical operation of the device is visually illustrated in Fig. 4.14(b).

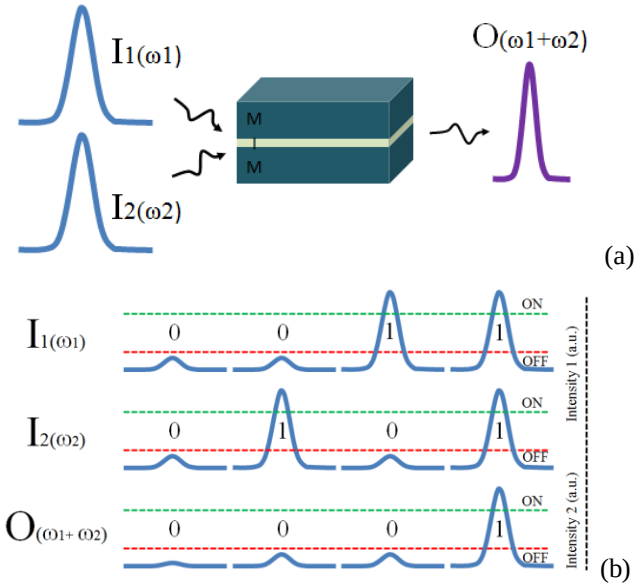


Figure 4.14 (a) schematic of 2 input pulses fed to the MIM waveguide and (b) illustration of all-optical logical AND operation using two input pulses at carrier frequencies ω_1 and ω_2 as the inputs and the frequency sum product at $\omega_1+\omega_2$ as the output.

2) Plasmonic switch and modulator

As illustrated in Fig. 4.15(a), the MIM waveguide can be used to realize a plasmonic switch by electrically controlling the bias voltage applied to the waveguide. Both the input pulses are set to have high intensities above the input ON threshold (1). The output pulse O at the sum frequency $\omega_1+\omega_2$ exists only

when the bias voltage is turned on.

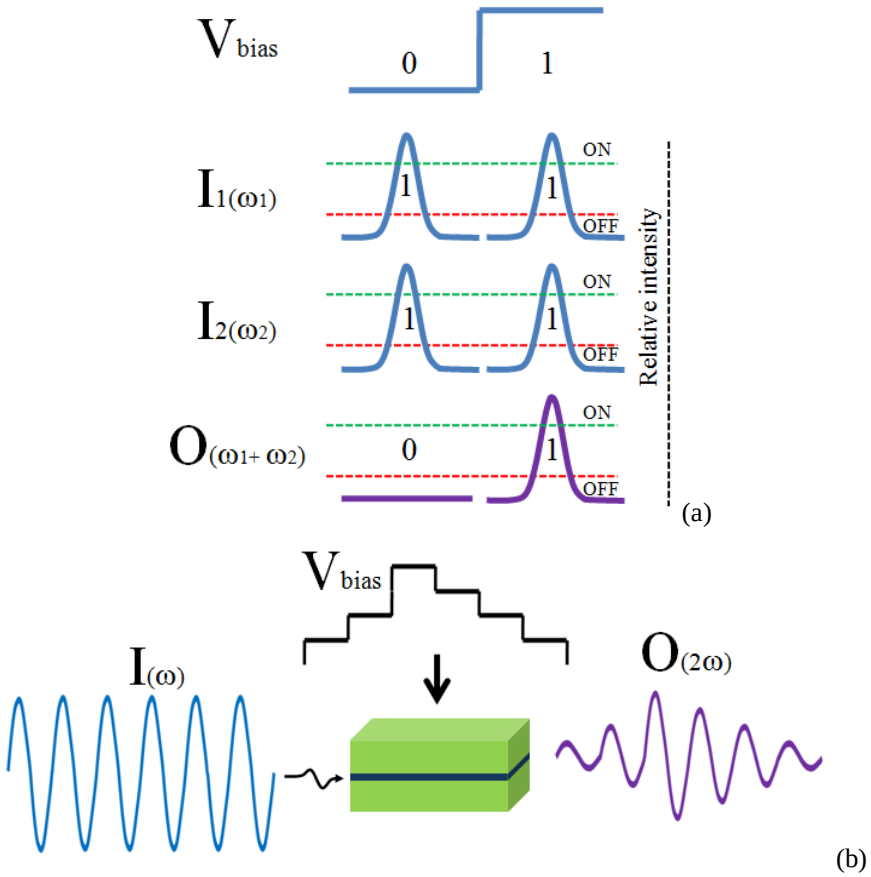


Figure 4.15 Illustrations of (a) plasmonic switching using two input pulses at carrier frequencies ω_1 and ω_2 as the inputs and the frequency adding product at $\omega_1+\omega_2$ as the output. The ON and OFF of the bias voltage applied to the MIM waveguide realizes a plasmonic switch and (b) plasmonic modulator using an CW signal at frequency ω as the input and the generated second harmonic CW signal at 2ω frequency as the output. The change in the bias voltage modulates the output plasmonic signal.

Similarly, we can control the bias voltage to realize a plasmonic modulator as illustrated in Fig. 4.15(b). Analytical results in chapter 3 demonstrate that the nonlinear conductivity σ_2 is a function of bias voltage while the peak intensity (conversion efficiency) of the second harmonic CW signal 2ω is function of σ_2. As a result, varying the bias voltage can modulate the intensity of the CW signal 2ω and thus electrically modulate a plasmonic signal. It is worth noticing that the speed of the modulator will be limited by the capacitive effect of the MIM waveguide rather than the electron tunneling time.

4.8 Summary

In this chapter we developed a nonlinear FDTD method for simulating nonlinear pulse propagation in MIM nanoplasmonic waveguide. Numerical simulations showed that the MIM nanoplasmonic waveguide can support ultrafast optical pulse propagation with pulse width as short as 100fs with negligible distortion. Nonlinear conducting process due to electron tunneling in the MIM junction can be used for optical rectification of ultrafast optical pulses and single-cycle THz pulse generation. Due to the fast electron tunneling time, the MIM waveguide has an ultra-wide electrical bandwidth of 16THz, which enables the waveguide to rectify optical pulses with duration as short as 65fs. We observed that the electrical bandwidth of the MIM nanoplasmonic waveguide is mainly restricted by the strong waveguide dispersion at the baseband frequencies. It could be improved by using dispersion compensation design or alternative material systems with reduced dispersion at the baseband.

We also explored the use of nonlinear conductivity in the MIM plasmonic waveguide for second harmonic pulse generation and optical pulse mixing. For example we showed that the device can be used to generating pulses at sum and difference frequencies with adjustable pulse width. These results demonstrate that the MIM nanoplasmonic waveguide can potentially be used for nonlinear plasmonic signal processing, such as frequency conversion and mixing, plasmonic switching and modulation, as well as basic all-optical logic operations.

Chapter 5

Further Analysis and Design Considerations of Nonlinear Nanoplasmonic Waveguides[3]

In chapters 3 and 4 we theoretically and numerically investigated wave propagation in MIM nanoplasmonic waveguides with second-order nonlinear conductivity arising from electron tunneling through the MIM junction. The device was shown to have ultra-wide electrical and optical bandwidths for optical rectification, second harmonic generation and frequency mixing. In this chapter we explore additional MIM waveguide designs, material systems, electron tunneling junctions and nonlinear conductivities that may also be used to realize nonlinear nanoplasmonic devices.

Firstly, we will look at the possibility of achieving third-order nonlinear conductivity in the MIM waveguide. In principle, the nonlinear current-voltage characteristic of the MIM junction can be modeled by a constitutive relation which includes higher-order conductivity terms. In particular we will expand the nonlinear conductivity model of the MIM junction to include the third-order term, which would give rise to third harmonic generation and four-wave mixing. We will investigate the efficiency of these third-order processes in the MIM waveguide compared to the second-order processes already analyzed in previous chapters.

Secondly, we will discuss other material systems that may be used to form the MIM plasmonic waveguide, and highlight the advantages of our choice of the Al-Al_2O_3-Al material system for the device.

Thirdly we will look at alternative MIM plasmonic waveguide designs that may provide lower propagation loss, thereby potentially improving the frequency conversion efficiencies. In particular we will look at the possibility of using long range plasmonic modes in MIM waveguides. The MIM waveguide used in the

[3] Part of this work has been presented by X. Lei et al. in IEEE Photonics Conference 2012, Nanoplasmonic II session, Sep 23-27, 2012, Burlingame, CA, USA.

previous chapters is designed to have thick metal layers (200nm thickness) to achieve high field confinement within the insulator gap. Theoretical analysis shows that the waveguide supports a short range mode characterized by a very short propagation distance due to the high metal absorption. If the thickness of the metal layers is reduced, the MIM nanoplasmonic waveguide can also support a long range mode with much longer propagation length. We will investigate the performance of the MIM plasmonic waveguide operating in the long range mode and compare it to the short range mode.

Lastly, we will consider other electron tunneling junctions that can be incorporated into a plasmonic waveguide structure for ultrafast nonlinear applications. An example is the metal-insulator-semiconductor (MIS) nanoplasmonic waveguide, which incorporates the MIS tunneling junction [70] into a hybrid plasmonic waveguide. We will theoretically analyze the characteristics of the MIS plasmonic waveguide for ultrafast nonlinear applications.

The rest of this chapter is outlined as follows: Section 5.1 presents the third-order nonlinear conductivity modelling in MIM nanoplasmonic waveguides. Section 5.2 investigates alternative material systems for MIM nanoplasmonic waveguides and discusses the advantages for choosing Al metal. Section 5.3 investigates variations of the MIM nanoplasmonic waveguide structure which support long range modes. In Section 5.4, we study the MIS nanoplasmonic waveguide for ultrafast nonlinear applications. Concluding remarks will be given in Section 5.5.

5.1 Third Order Nonlinear Conductivity in MIM Plasmonic Waveguides

In our analysis of nonlinear effects in MIM plasmonic waveguides, we model the nonlinear current-voltage relation of the MIM junction using a Taylor series expansion of the J-V curve at the bias point. In our analytical and numerical studies in chapters 3 and 4, we truncated the Taylor series to second-order because

we were primarily interested in second-order nonlinear processes such as rectification and second-order harmonic generation in the MIM waveguide. In theory the MIM junction possesses an infinite number of higher-order nonlinear conductivities, although the effects of higher-order terms are much smaller. It should also be noted that if the MIM junction has zero bias voltage, the second-order nonlinear term will be absent and the lowest nonlinear term would be the third-order term.

In this section we investigate the third-order nonlinear conductivity that can be achieved in the MIM junction. For this purpose, we expand the J-V curve of the MIM junction to third order as

$$J(E_z) = \sigma_1 E_z + \sigma_2 E_z^2 + \sigma_3 E_z^3, \tag{5.1}$$

where σ_3 is the third order nonlinear conductivity, with units of $\Omega^{-1}\text{m/V}^2$, which is responsible for third-order processes such as third harmonic generation and four-wave mixing. The value of σ_3 can be obtained by applying a finite differencing scheme to the $J(V)$ curve around the bias point V_{bias} as

$$\sigma_3 = \frac{d^3}{6}\frac{d^3 J}{dV_\omega^3} = \frac{J(V_{bias} + 2\hbar\omega/e) - 2J(V_{bias} + \hbar\omega/e) + 2J(V_{bias} - \hbar\omega/e) - J(V_{bias} - 2\hbar\omega/e)}{12(\hbar\omega/ed)^3}. \tag{5.2}$$

For the Al-Al_2O_3-Al junction with a 3nm thick Al_2O_3 layer, we obtained $\sigma_3 = 5\times10^{-19}\ \Omega^{-1}\text{m/V}^2$ at V_{bias}= 0.1V. The linear and second-order conductivities are the same as those values used in the previous chapters, $\sigma_1=4\times10^{-2}\Omega^{-1}/\text{m}$ and $\sigma_2=3.5\times10^{-10}\Omega^{-1}/\text{V}$.

We simulated nonlinear pulse propagation in the Al-Al_2O_3-Al waveguide with the nonlinear conductivity model of Eq. 5.1 and the conductivity parameter values at 0.1V bias. The waveguide was excited with a 1ps Gaussian pulse at carrier wavelength of 1.55μm and peak power of 0.65mW. In addition to the rectified and 2ω pulses, we also observed a pulse generated at the 3ω frequency. Fig. 5.1 compares the peak field amplitudes of the 2ω and 3ω pulses as a function of the propagation distance, with the plot of the 3ω pulse scaled by a factor of 100. It is clear that the peak field amplitude of the 3ω pulse is smaller than that of the 2ω pulse by two orders of magnitude. In addition, the 3ω pulse has a shorter

propagation length than the 2ω pulse. The above results indicate that the third-order nonlinearity in the MIM waveguide is much smaller compared to the second-order nonlinearity, and we are justified in using only a second-order nonlinear conductivity in the waveguide analysis.

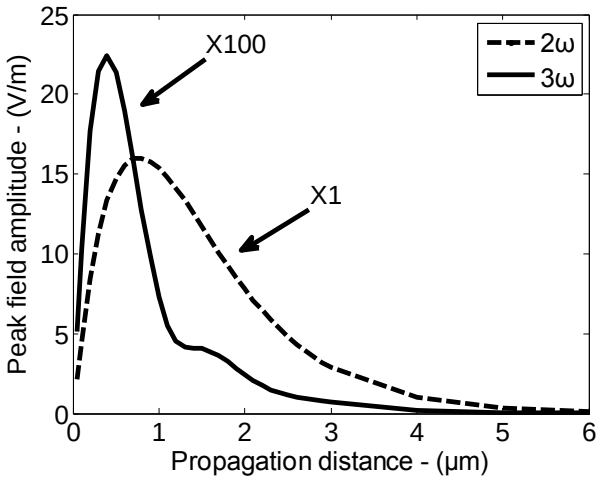


Figure 5.1 Peak electric field amplitudes for the 2ω and 3ω (scaled by a factor of 100) pulses as functions of propagation distance in the Al-Al_2O_3-Al waveguide.

5.2 Alternative Material Systems for MIM Nanoplasmonic Waveguides

Many metals such as Al, Cr, Au, Pt and Nb [3-5] have been used as metal electrodes for MIM tunnel junctions and many materials such as SiO_2, NiO, Cr_2O_3 and Al_2O_3 [3-5] have been used as insulator layers.

For plasmonic waveguide applications, we require metals with low loss and permittivity with negative real part. A large negative permittivity value leads to less field penetration into the metal and thus lower propagation loss and stronger field confinement in the insulator gap. With the metal permittivity modeled using the Drude's model, a high plasma frequency is needed to achieve negative real permittivity at optical frequencies. Both Ag and Au have high plasma frequencies and low loss at the optical frequencies. Additionally, structures made of Ag or Au are relatively easy to fabricate. Also, Au is a chemically stable metal [71-72]. As a result, Au and Ag are often used in plasmonic waveguides.

MIM plasmonic waveguides based on Au- poly-methyl-methacrylate (PMMA)-Au [73] and Ag-PMMA-Ag [74] material systems have been experimentally demonstrated.

The incorporation of an electron tunneling junction into the MIM plasmonic waveguide allows nonlinear conductive effects to be achieved for ultrafast optical rectification and broadband frequency conversion. However, the metal-insulator material system must be chosen properly to provide a potential barrier for electron tunneling as well as low-loss field confinement. Table 5.1 provides a comparison of three possible material systems for nonlinear MIM waveguide design.

Examination of the table shows that there are several advantages for choosing the Al-Al_2O_3-Al material system. First, Al is a potentially attractive metal for plasmonic waveguides since it has higher plasma frequency than Ag and Au. The real part of its permittivity remains negative into the deep UV (200nm wavelength) [71,75], thereby providing a much wider plasmonic bandwidth. Furthermore, Al also has lower loss than Ag and Au at shorter wavelengths, which is desirable for frequency up-conversion. As a result, Al is a better plasmonic material than Ag or Au at higher frequencies such as the blue and the ultra violet (UV) [75].

Table 5.1 Comparisons of Ag, Al and Au MIM material systems.

	Ag	Al	Au
Relative permittivity (1.55μm)	-128.7+3.44i [76]	-840+137i [59-60]	-132+12.65i [44]
Plasma frequency (eV)	9.2	12.7	8.9
Barrier height (eV)	2.26	2.28	3.1
M-Al_2O_3-M propagation length (at 1.55 μm wavelength) (μm)	2.41	1.27	0.71

At the visible and infrared frequencies, although Al has slightly higher absorption coefficient than Ag and Au, the large negative part of its permittivity helps reduce field penetration into the metal, which results in lower effective propagation loss in the MIM waveguide. For example, as shown in Table 5.1, at

the 1.55μm wavelength an Al-Al_2O_3(3nm)-Al waveguide with a 3nm insulator thickness has a propagation distance of 1.27μm as compared to 2.41μm for Ag-Al_2O_3-Ag and 0.71μm for Au-Al_2O_3-Au waveguides with the same insulator thickness. Thus Al is better than Au but worse than Ag in terms MIM propagation length at the infrared wavelengths.

In terms of the tunneling junction properties, Al (4.28eV [71]) has a work function comparable to Ag (4.26eV [71]) but lower than Au (5.1eV [71]) by about 0.9eV. With the same Al_2O_3 layer (electron affinity of 2eV [77]), MIM junctions made with Au have a potential barrier which is too high leading to low tunneling probability. On the other hand, with metals having lower work functions (such as Mg: 3.1eV [78]), the barrier height is too low to eliminate thermionic current [79]. Al, with intermediate work function, can provide both high tunneling current and minimize thermionic effects.

An additional advantage that Al has over Au and Ag metals is in terms of fabrication. Al oxidizes easily under atmospheric condition and forms a stable Al_2O_3 with a typical thickness of 2.5-3nm [40]. The thickness of this native oxide is within the optimum range for electron tunneling. Thus the Al-Al_2O_3-Al junction can be cheaply and easily fabricated using a standard metal deposition process and native oxide growth. Finally it should be noted that Al is compatible with the complementary-metal-oxide-semiconductor (CMOS) fabrication technology while Ag and Au are not [75,80]. This process compatibility allows MIM nanoplasmonic waveguides to be integrated with CMOS electronics on the same chip using current semiconductor technology.

5.3 MIM Nanoplasmonic Waveguides Operating in Long Range Modes

Metal-Insulator-Metal plasmonic waveguides, in which the thickness of the metal layers exceeds the skin depth, can support only short range modes characterized by short propagation lengths. As the thickness of the metal layer decreases, the MIM waveguide evolves into an Insulator-Metal-Insulator-Metal-Insulator

(IMIMI) waveguide which can support long range modes with much longer propagation lengths. The schematics of the MIM and IMIMI waveguides are shown in Fig. 5.2.

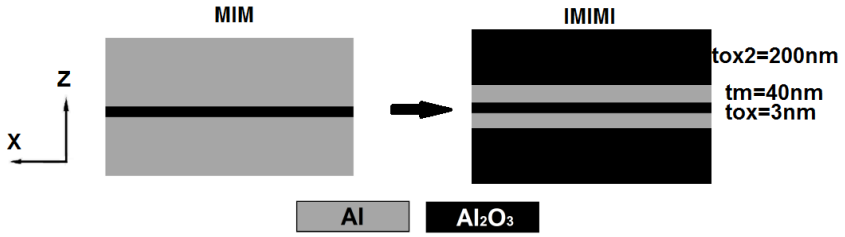


Figure 5.2 Schematics of the MIM (Fig. 3.1(b)) and IMIMI plasmonic waveguides.

The IMIMI waveguide can be treated as two Insulator-Metal-Insulator (IMI) waveguides stacked in parallel. Each IMI waveguide supports a long-range antisymmetric mode and a short-range symmetric mode [81] as illustrated in Fig. 5.3(a) and 5.3(b).

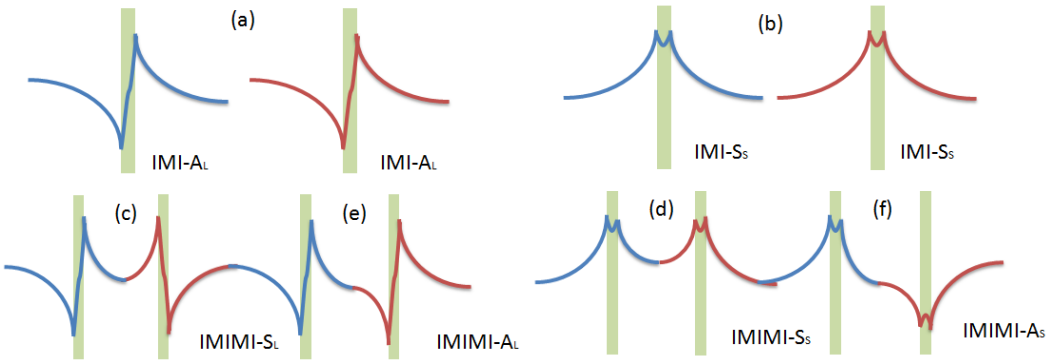


Figure 5.3 Illustrations of mode formation in the IMIMI waveguide in terms of the transverse electric E_z field profile. IMI: (a) Two A_L modes, (b) two S_S modes, IMIMI: (c) S_L, (d) S_S, (e) A_L, and (f) A_S (Figures are reproduced from Fig. 2 in [81]).

Here the mode symmetry refers to the symmetry of the transverse electric field E_z with respect to the waveguide centre. When the two IMI waveguides are brought together, their modes are merged into four coupled plasmonic modes: a) symmetric long range (S_L), b) antisymmetric long range (A_L), c) symmetric short range (S_S), and d) antisymmetric short range (A_S) as illustrated in Figs. 5.3(c-f). The two short range modes have similar propagation distances as the MIM waveguide with thick metal layers, while the two long range modes have much

longer propagation lengths. The long propagation lengths of the long range modes are due to weaker fields inside the metal layers, which lead to lower ohmic loss. The characteristics of the four modes in an IMIMI waveguide have been studied in detail in [81]. Here, we are interested in evaluating the IMIMI structure for nonlinear waveguide applications.

The dispersion equation for the antisymmetric modes in an IMIMI waveguide is given by [81] (assuming the same insulator material for the three insulator layers and the same metal material for the two metal layers)

$$(\frac{k_{ox}\varepsilon_m}{k_m\varepsilon_{ox}})\tanh(k_{ox}\frac{a}{2}) = -[\frac{k_m \sinh(k_m b)/\varepsilon_m + k_{ox}\cosh(k_m b)/\varepsilon_{ox}}{k_m \cosh(k_m b)/\varepsilon_m + k_{ox}\sinh(k_m b)/\varepsilon_{ox}}], \tag{5.3}$$

and for the symmetric modes *tanh* in Eq. 5.3 is replaced by *coth*. In the above, k is the transverse propagation constant in the insulator (k_{ox}) or metal (k_m) layers, ε is the relative permittivity of the insulator (ε_{ox}) or metal (ε_m) layers, a is the thickness of the center insulator layer and b is the thickness of the metal layers.

We consider an IMIMI structure composed of Al metal layers and a center Al_2O_3 insulator layer of 3nm thickness. For simplicity we also assume the outer insulator layers to be Al_2O_3. We vary the thickness of the metal layers of the IMIMI waveguide and calculate the propagation constants for the antisymmetric and symmetric modes at 1.55μm wavelength using Eq. 5.3.

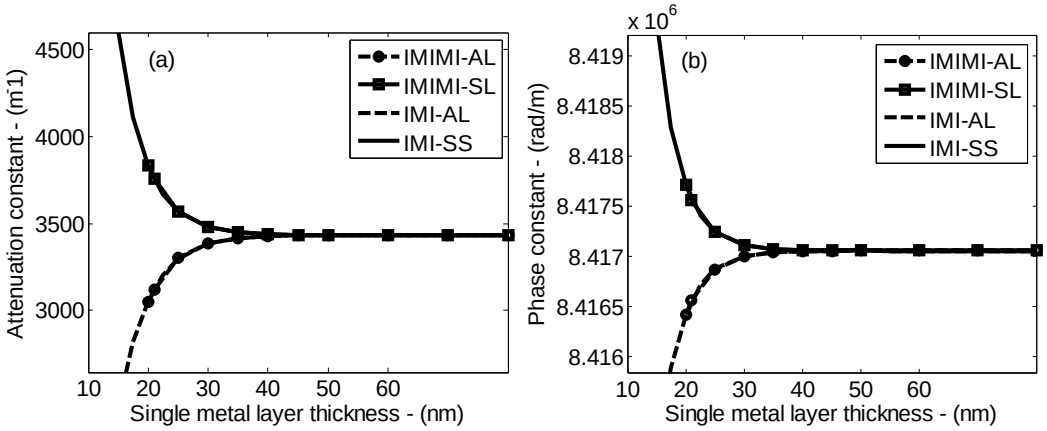


Figure 5.4 Calculated (a) attenuation constants and (b) phase constants for the antisymmetric (A_L) and symmetric (S_L) modes at the 1.55μm wavelength for the IMIMI waveguide at different single metal layer thickness. The attenuation constants and phase constants of the two modes from the IMI waveguide with the same total metal thickness are also given for reference.

The attenuation constants and phase constants obtained are plotted in Figs. 5.4(a) and 5.4(b), respectively. On the same plots, we also show the attenuation constants and phase constants of the two modes for the IMI waveguide with the same metal thickness. It is seen that the two long range modes converge to the same mode when the metal thickness is greater than 40nm. In addition, due to the ultra-thin centre insulator layer, the two long-range modes behave similar to the odd and even modes of an IMI waveguide.

The two long range modes are found to exist only when the thickness of the metal layer is reduced to below about 80nm. With the metal thickness set at 40nm, we calculated the propagation lengths of the two long range modes at the 1.55μm wavelength to be: a) 292μm for A_L and b) 291μm for S_L. These values are much higher than that of the MIM plasmonic waveguide (1.27μm). Their propagation lengths are further increased if the thickness of the metal layers can be further reduced, as indicated by Fig. 5.4(a). In Fig. 5.5 we plotted the transverse electric field E_z profiles for the two long range modes. It is seen that the field for both long range modes resides mainly in the outer insulator layers and only a weak field exists in the center insulator gap. Thus the long range modes have poor field confinement, which is a trade-off for the long propagation lengths.

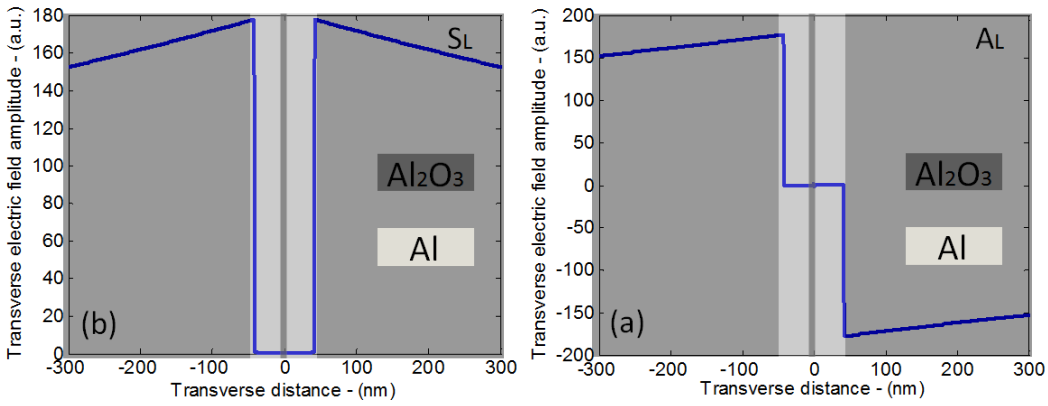


Figure 5.5 Transverse electric E_z field profiles (at 1.55μm wavelength) for the IMIMI waveguide: (a) S_L and (b) A_L with 40nm thick metal layers and a 3nm thick center insulator layer.

We also calculated the responsivity (at a propagation distance of $3L_P$) of the IMIMI waveguide operating at the two long-range modes for rectifying CW

signal at the 1.55μm wavelength and average power of 0.362mW using Eq. 3.26 and Eq. 3.12. The results are compared with the MIM structure in Table 5.2. It is apparent that both long-range modes provide low responsivity due to the extremely low power confinement.

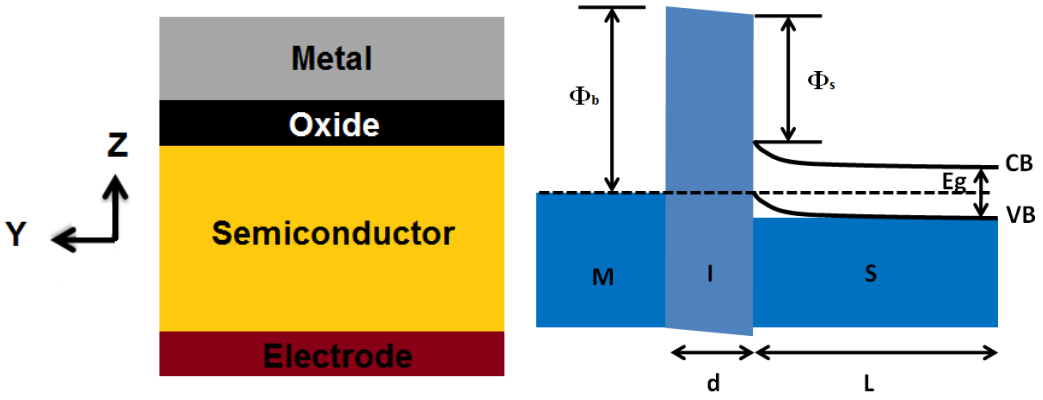
Table 5.2 Responsivity comparison of MIM and IMIMI waveguides for rectifying CW signal at 1.55μm wavelength.

Type	MIM	$IMI_1MI_2(A_L)$	$IMI_1MI_2(S_L)$
Dimensions	I=3nm, M=200nm	I_1=3nm,I_2=200nm, M=40nm	
Confinement (%)	98.6	1.43×10^{-5}	8.26×10^{-3}
Attenuation constant (1/m)	7.85×10^{5}	3.42×10^{3}	3.43×10^{3}
Phase constant (rad/m)	2.18×10^{7}	8.42×10^{6}	8.42×10^{6}
Peak gap voltage drop (V)	0.1	~0	9.14×10^{-4}
L_p (μm)	1.27	292	291
Rectified DC current (A)	1.24×10^{-8}	~0	2.36×10^{-10}
Responsivity (A/W)	3.43×10^{-5}	~0	7.18×10^{-7}

5.4 Metal-Insulator-Semiconductor Nanoplasmonic Waveguides

Besides MIM junctions, another type of electron tunneling junctions that can support surface plasmon polaritons is the Metal-Insulator-Semiconductor (MIS) junction. Fig. 5.6(a) shows a schematic of a MIS junction, which consists of a thin insulator layer sandwiched between a metal layer and a thick semiconductor layer. A typical band diagram of the MIS tunnel junction is given in Fig. 5.6(b). MIS diodes are commonly used as photodetectors which produce a photocurrent when illuminated by radiation with energy above the semiconductor bandgap. The insulator layer thickness in these devices is typically greater than 5nm while the semiconductor layer has a thickness of hundreds of nanometers [82], which is determined by the absorption length of the incoming radiation. For a junction with insulator thickness greater than 5nm, electron tunneling across the potential barrier is greatly inhibited. As a result the total current in the MIS diode is dominated by the photocurrent. When the thickness of the insulator is reduced to

below 3nm, electron tunneling current becomes more significant. Since the temporal response of the photocurrent is determined by the electron transit time across the semiconductor layer (or more precisely, across the depletion region at the semiconductor-insulator interface), it is much slower than the tunneling current. Typical MIS photodiodes have a response time in the order of 10^{-11}s [83], and are thus not suitable for ultrafast applications.



(a) (b)

Figure 5.6 (a) Schematic of a MIS junction and (b) typical band diagram of a MIS junction with insulator layer thickness d and metal-insulator barrier height of Φ_b.

For ultrafast THz and second harmonic pulse generation, the MIS junction must be operated in the regime where the tunneling current dominates, so the insulator layer must be kept thin. In addition, for low loss plasmonic propagation, the guided radiation must have energy below the semiconductor bandgap. This effectively precludes the generation of the photocurrent due to single photon absorption. However, there is photocurrent generated due to two-photon absorption. To simplify our analysis, we do not include the effects from two-photon absorption which will be considered in a future study. In here, we assumed that the I-V characteristic of the junction is solely due to electron tunneling.

We considered an MIS plasmonic waveguide consisting of a 3nm layer of SiO_2 sandwiched between a 200nm Au layer and a 350nm thick Si layer. The indexes of SiO_2 and Si are 1.44 and 3.48, respectively, at the 1.55μm wavelength. Since the bandgap of Si is 1.1eV [84], the waveguide can support plasmonic propagation in the telecom wavelength range. MIS waveguides based on Au-SiO_2-

Si material system have been extensively studied recently [85]. We calculated its mode characteristics at the 1.55μm wavelength and plotted the transverse electric field E_z and magnetic field H_y in Fig. 5.7. Due to the index discontinuities at the Si-SiO_2 and SiO_2-metal interfaces, the electric field E_z is seen to be strongly confined and enhanced within the insulator layer [85]. This field enhancement creates a large voltage drop across the insulator gap leading to a high electron tunneling probability. The MIS waveguide can achieve a propagation length of 42μm, which is much longer than the 1.27 μm of the Al-Al_2O_3-Al MIM plasmonic waveguide. However, as shown in Fig. 5.7(b), the magnetic field in the MIS waveguide penetrates deeply into the semiconductor layer. As a result, the power confinement inside the insulator layer is calculated to be only 5.1%.

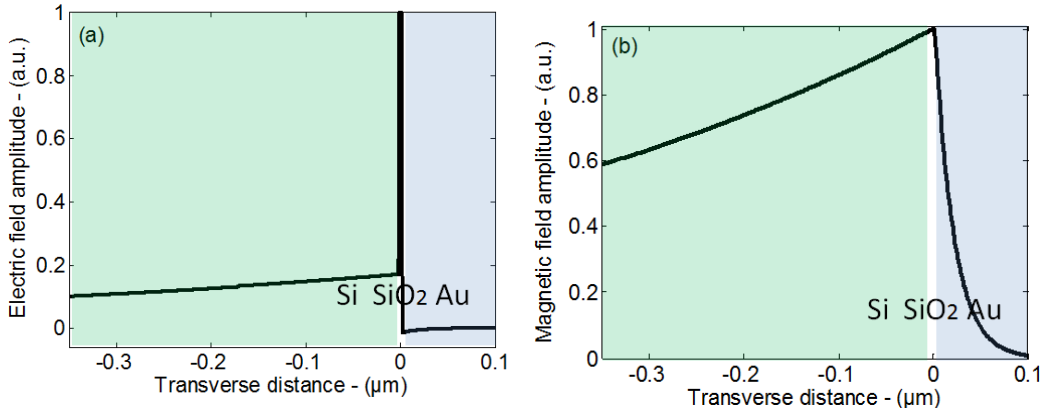


Figure 5.7 Normalized (a) transverse electric field E_z and (b) transverse magnetic field H_y profiles for the Au-SiO_2-Si waveguide at 1.55μm wavelength with an insulator layer thickness of 3nm.

The band diagram of the Au-SiO_2-Si tunneling junction is shown in Fig. 5.8(a). With a positive DC bias voltage applied to the Au layer, the dark tunneling current of the Au-SiO_2-Si junction could be calculated using the following equations [86]

$$J(V_{bias}) = A^* T^2 \exp(-\alpha_T d\sqrt{q\phi_T}) \exp(\frac{-q\phi_B}{kT})[\exp(\frac{qV_{bias}}{\eta kT}) - 1], \quad (5.4)$$

where A^* is the effective Richardson constant, T is the temperature, q is the electron charge, d is the thickness of the junction, k is the Boltzmann constant, η is

ideality factor, $\exp(-\alpha_T d\sqrt{q\phi_T})$ is the tunneling probability, Φ_B is the metal-insulator barrier height and Φ_T is the insulator-semiconductor barrier height. Assuming an ideality factor of 1 and using Eq.5.4 and junction parameters in Fig. 5.8(a), we calculated the dark tunneling current of the Au-SiO_2-Si junction at different bias voltages and plotted the results in Fig. 5.8(b). The nonlinear conductivity parameters σ_1 and σ_2 computed from Eq. 3.9 and Eq. 3.10 and are plotted in Fig. 5.8(c) and Fig. 5.8(d). It is seen that σ_2 achieves a peak value of $8.5\times10^{-13}\Omega^{-1}$/V when biased at around 1V.

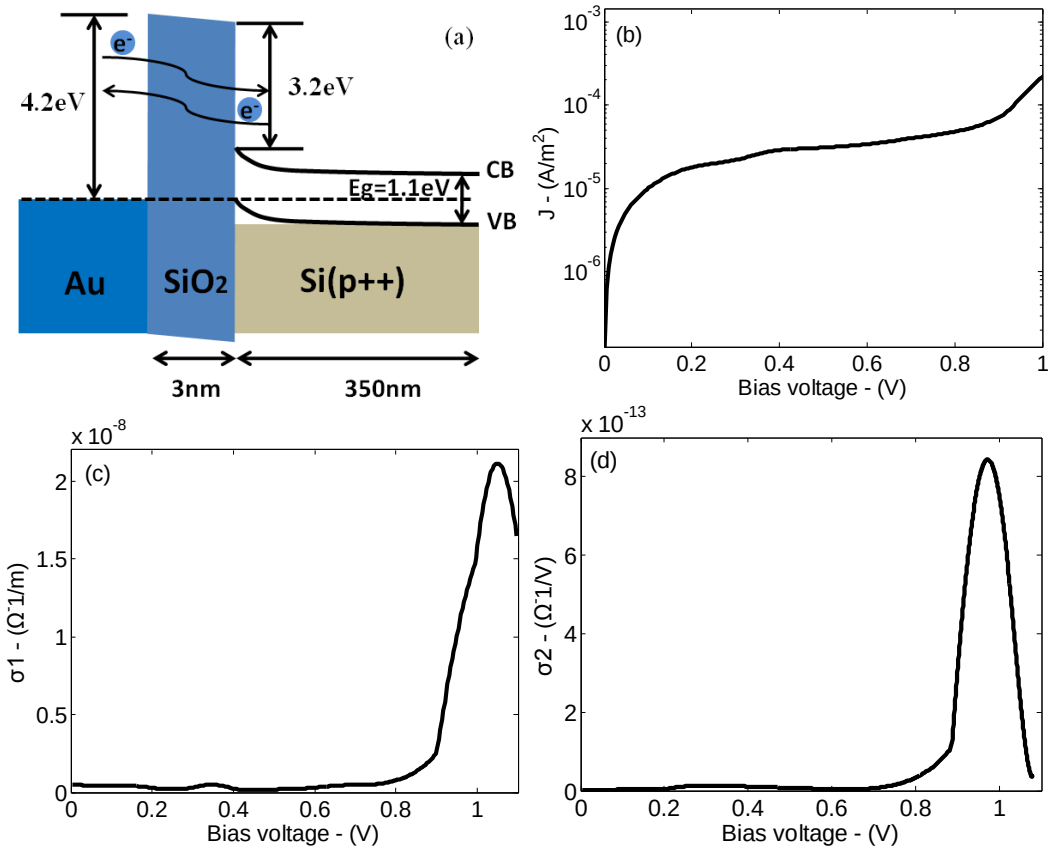


Figure 5.8 Plots of (a) band diagram of an Au-SiO_2-Si junction with 3nm thick SiO_2 layer and 350nm thick Si layer, (b) calculated dark tunneling current of an Au-SiO_2-Si junction with a barrier width of 3nm, (c) nonlinear conductivity σ_1 and (d) σ_2 for the MIS junction at different bias voltages.

We next calculated the responsivity of the MIS waveguide for rectifying CW signal at the 1.55μm wavelength. The MIS waveguide length was assumed to

be $3L_P$ and the average input power was 0.362mW with 1V bias voltage (optimum bias voltage). The responsivity and other parameters are given in Table 5.3. It is clear that the MIS plasmonic waveguide has a lower responsivity (operates at optimum condition with 1V bias voltage) than the MIM waveguide (operates at optimum condition with 1.1V bias voltage) due to lower power confinement and lower effective second-order nonlinear conductivity at the bias voltage of 1V voltage (optimum bias voltage).

Table 5.3 Optimum responsivity of the Au-SiO_2-Si plasmonic waveguide for CW signal rectification. (An incident CW signal at 1.55μm wavelength and average power of 0.362mW is assumed. The results from the Al-Al_2O_3-Al plasmonic waveguide are given as a reference.)

Type	MIS	MIM
Dimensions	I=3nm,M=200nm,S=350nm	I=3nm, M=200nm
Confinement %	5.1	98.6
Attenuation constant (1/m)	2.37×10^{4}	7.85×10^{5}
Phase constant (rad/m)	1.42×10^{7}	2.18×10^{7}
Peak gap voltage drop (V)	1.34×10^{-3}	0.1
L_p (μm)	42.1	1.27
Optimum bias voltage (V)	1	1.1
Responsivity (A/W)	5.5×10^{-10}	8.00×10^{-4}

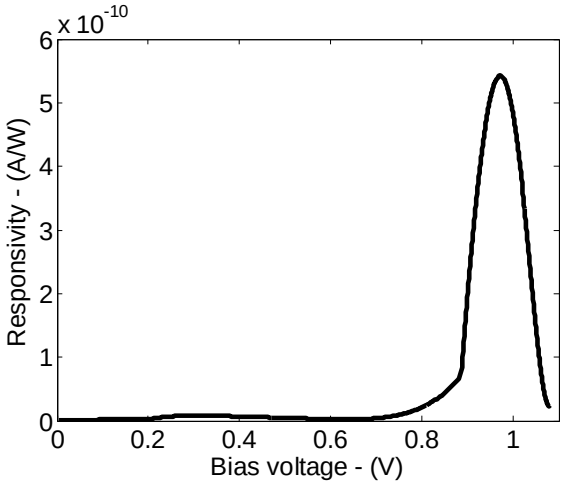


Figure 5.9 Plot of the responsivity of the Au-SiO_2-Si plasmonic waveguide with 3nm thick SiO_2 layer for rectifying CW signal of 1.55μm wavelength and average power of 0.362mW at different bias voltages.

In Fig. 5.9 we plotted the calculated responsivity vs. bias voltage of the Au-SiO_2-Si plasmonic waveguide. A maximum responsivity of $5.5x10^{-10}$ A/W can be achieved when the bias voltage is set at 0.97V. This value is still much lower than the achievable responsivity of an MIM waveguide. However, even though the MIS waveguide shows worse performance than the MIM waveguide, there are a few positive features for this type of devices that should not be neglected. Firstly, it is worth noticing that the tunneling current density of the Au-SiO_2-Si junction under forward-bias condition increases exponentially when the insulator gap thickness is reduced to below 2.8nm [86]. For example, with a gap thickness of 2.2nm, the tunneling current density (under the same bias condition) of the MIS junction is increased by a factor of 10^6 [86], which could greatly improve the second-order nonlinear conductivity of MIS plasmonic waveguides. Secondly, the much longer propagation length of the MIS plasmonic waveguide (compared to MIM waveguides with the same insulator thickness) helps to improve the efficiency of the nonlinear frequency conversion processes. Thirdly, the large dimensions of the MIS waveguides make their fabrication easier and also allow them to be better integrated with other on-chip structures such as silicon photonic waveguides. For these reasons, MIS plasmonic waveguides are still attractive candidates for on-chip nonlinear plasmonic applications that should be further explored.

5.5 Summary

In this chapter we provided further analysis of nonlinear MIM waveguides and investigated alternative waveguide designs, material systems and tunneling junctions. We found that while the MIM junction can provide third-order nonlinear conductivity, its effect is much smaller than the second-order nonlinearity. As a result, the quadratic conductivity model is found to be sufficient for investigating nonlinear processes in MIM waveguides. We also investigated the use of long range modes in IMIMI plasmonic waveguides for improving the nonlinear frequency conversion efficiencies. We found that although the

propagation lengths of these long range modes are a few orders of magnitude larger than those of the short range modes, most of the field exists outside the structure and very weak field develops across the insulator gap. As a result the effective nonlinearity is actually weaker than the more strongly confined short range modes with optimum metal thickness of 200nm for achieving maximum responsivity.

We also explored other material systems for implementing the MIM waveguide structure. Compared to commonly used plasmonic metals such as Ag and Au, Al is found to be a better material for nonlinear plasmonic applications due to its high plasma frequency, relatively low loss, suitable work function, easy fabrication and compatibility with CMOS technology.

Finally we also investigated MIS plasmonic waveguides as an alternative structure to MIM junctions. An Au-SiO_2-Si MIS waveguide with a 3nm thick SiO_2 gap can achieve a propagation length 30 times that of an MIM waveguide. Although the responsivity of the MIS waveguide for rectifying CW signals at the 1.55μm wavelength is poorer than that of the MIM waveguide, the long propagation length and large dimensions of the waveguide make it an attractive alternative platform for realizing nonlinear plasmonic devices that can be integrated with other photonic components.

Chapter 6

Conclusion and Future Work

6.1 Summary of Research

This thesis focuses on the analytical and numerical investigation of MIM nanoplasmonic waveguides in the presence of electron tunneling for ultrafast and broadband nonlinear optical applications such as rectification, second-order harmonic generation and frequency mixing. We also investigated improved antenna designs for achieving ultra-wideband efficient coupling of THz radiation to MIM waveguides. Below we summarize the key findings of the research.

In chapter 2 the performance of antenna-coupled lumped MIM diodes for THz detection was analyzed in detail. We found that antenna-coupled MIM diodes generally have poor detection efficiency and restricted bandwidth operation due to two main causes:

(i) MIM junctions typically have large impedance in the kΩ-MΩ range, leading to a large impedance mismatch between the diode and a typical planar antenna with impedance of 50Ω. This results in poor coupling of radiation to the MIM diode.

(ii) The bandwidth of the MIM diode is restricted by the RC time constant as determined by the antenna impedance and the junction capacitance. The RC time constant sets a cut-off frequency beyond which rectification cannot be carried out efficiently.

In addition, antenna-coupled MIM detectors commonly use the standard bowtie antenna design which cannot provide the wide bandwidth required for coupling THz radiation to the diode. Using geometrical modification technique, we propose a new THz antenna design based on the split-arm bowtie geometry that can receive THz pulse as short as 1ps duration (up to 1THz bandwidth). The design principle behind the proposed antenna is to achieve proper surface current

distribution in the bowtie arms, via the adjustment of the number of arms, flare angles and split angles. The split-arm configuration is equivalent to connecting multiple bowtie antennas in parallel with reduced overall antenna impedance and enhanced bandwidth. The new design enables ultrafast THz pulses to be efficiently coupled to the MIM detector or radiated out from the MIM detector.

In chapter 3, we theoretically investigated nonlinear MIM nanoplasmonic waveguides in the presence of electron tunneling across the MIM junction. The structure behaves as a traveling-wave detector when used to rectify continuous-wave and pulsed signals. Modeling electron tunneling as an effective nonlinear conductivity in the insulator layer, we developed equations based on the Coupled Mode Theory for analyzing optical rectification and second harmonic generation in the MIM waveguide under both CW and pulsed excitations. Our theoretical analysis showed that as a traveling-wave detector, the MIM plasmonic waveguide provides three key advantages over the lumped MIM detector: (a) broad electrical and optical bandwidths by overcoming the RC time constant of the MIM junction, (b) improved responsivity for THz detection by allowing the signal to propagate along the waveguide thereby increasing the nonlinear interaction length, and (c) low characteristic impedance that can be better matched to typical antenna impedance of 50Ω, allowing detector coupling efficiency near unity to be achieved.

In chapter 4, we developed a nonlinear FDTD algorithm for performing rigorous numerical analysis of MIM nanoplasmonic waveguides. The FDTD method incorporates the nonlinear conductive constitutive relation of the MIM junction and accounts for metal dispersion via the Auxiliary Differential Equation method. The simulation results not only served to validate the approximate theoretical analysis based on the Coupled Mode Theory, but also enabled more accurate analysis of dispersive effects associated with nonlinear pulse propagation in the MIM waveguides.

Both theoretical and numerical analysis of the MIM plasmonic waveguide under pulsed excitation revealed important characteristics of the structure which could lead to novel applications. Specifically, we demonstrated that the structure

can rectify ultrafast optical pulses yielding single-cycle THz pulses. As a result, the device has potential application for on-chip THz source generation. Other applications of the MIM structure include second harmonic pulse generation, frequency mixing of two optical pulses which could have applications in all-optical logic operations, and plasmonic switching and modulation.

In chapter 5, we expanded the scope of our analysis beyond second-order nonlinear conductivity in MIM structures and explored alternative plasmonic waveguide designs, material systems and tunneling junctions. Firstly, we investigated the strength of the 3^{rd} order nonlinear conductivity in the MIM junction and found that its effect is much smaller than that of the second-order conductivity. Secondly, we compared the advantages and disadvantages of using Al as the plasmonic material against other commonly used metals such as Au and Ag. We concluded that overall the Al-Al_2O_3-Al structure provides better nonlinear device performance as well as compatibility with the CMOS process. Thirdly, we considered the possibility of using long range modes in MIM waveguides for enhancing the nonlinear interaction length, but found that the overall responsivity is still lower than the short range MIM mode due to poorer field confinement. Lastly, we investigated nonlinear plasmonic waveguides based on the MIS tunneling junction and found that these structures have lower nonlinear conductivity than MIM structures. Nevertheless, the low loss and ease of integration with photonic waveguides make these structures attractive candidates for nonlinear plasmonic device applications.

6.2 Major Contributions of This Work

The major contributions of this work are as follows:

1) Proposal of nonlinear MIM nanoplasmonic waveguides based on electron tunneling: We exploit the nonlinear current-voltage relation of an MIM tunneling junction to incorporate ultrafast, second-order nonlinear conductivity into the plasmonic waveguide. The structure enables second-order frequency conversion processes such as optical rectification, second harmonic generation,

and sum and difference frequency generations to be achieved on a nanoplasmonic platform. We also explored nonlinear MIM waveguides based on long range surface plasmon modes and nonlinear plasmonic waveguides based on Metal-Insulator-Semiconductor tunneling junctions.

2) Derivation and solution of the Coupled Mode Theory equations for wave couplings in a nonlinear conductive medium under continuous-wave and pulse excitations.

3) Development and implementation of a nonlinear Finite Difference Time Domain method for simulating plasmonic waveguides with nonlinear conductivity. The method allows accurate simulation of more complex nonlinear plasmonic devices based on electron tunneling junctions such as MIM and MIS junctions.

4) Proposal and analysis of broadband THz antennas based on the split-arm bowtie design. We showed that the new antenna design could achieve a -10dB return loss bandwidth of more than 1THz and could thus be used to receive/radiate ultrafast pulses.

6.3 Future work

This work theoretically and numerically demonstrated the feasibility and achievable performance of MIM nanoplasmonic waveguides for nonlinear frequency conversion and signal processing applications. Future work could focus on the practical realization and further optimization of the device, as suggested below.

1) Device optimization: Research work can focus improving in and out coupling of radiation to the MIM plasmonic waveguide. Instead of using coupling antennas, the MIM waveguide may be embedded in an optical or plasmonic waveguide that can guide light into and out of the nonlinear MIM section. To further improve the electrical bandwidth of the device, dispersion compensation technique can be applied to design waveguides with lower dispersion. To reduce waveguide loss and improve nonlinear efficiency, parallel arrays of MIM

waveguides can be considered. The use of amorphous metals can potentially be used to reduce ohmic loss.

2) Device fabrication: The Al-Al_2O_3-Al plasmonic waveguides studied in this work can be fabricated using standard thin-film deposition and electron beam lithography processes. The device can be fabricated on a Silicon-on-Insulator substrate which is a popular platform for silicon photonic devices. The Al metal layers could be deposited by RF sputtering while the insulator layer can be grown by native oxidation in a controlled environment to achieve the desired insulator thickness. Electron beam lithography can be used to pattern the lateral dimension of the MIM waveguide as well as the split-arm bowtie antennas.

3) Experimental demonstration of practical applications: The ultimate test of the usefulness of the proposed nonlinear MIM nanoplasmonic waveguides would be to demonstrate their practical applications for on-chip single-cycle THz pulse generation, second harmonic generation, as well as plasmonic switching and modulation.

4) Rigorous 3D FDTD simulations: the nonlinear FDTD algorithm can be extended to handle three-dimensional structures for more rigorous and realistic simulations of MIM nanoplasmonic waveguides.

References

[1] J. W. Dees, “Detection and harmonic generation in the sub-millimeter wavelength region,” *Microwave Journal*, vol. 9, pp. 48-55, Sept. 1966.

[2] K. Choi, F. Yesilkoy, G. Ryu, S. Cho, N. Goldsman, M. Dagenais, and M. Peckerar, "A Focused Asymmetric Metal–Insulator–Metal Tunneling Diode: Fabrication, DC Characteristics and RF Rectification Analysis," *IEEE Transactions on Electron Devices*, vol. 28, no. 10, pp. 3519-3528, Oct. 2011.

[3] P.C. Hobbs, R. B. Laibowitz, and F. R. Libsch, “Ni-NiO-Ni tunnel junctions for terahertz and infrared detection,” *Applied Optics*, vol.44, no.32, pp.6813-6822, Nov. 2005.

[4] C. Fumeaux, W. Herrmann, F. K. Kneubühl, and H. Rothuizen, “Nanometer thin-film Ni–NiO–Ni diodes for detection and mixing of 30 THz radiation,” *Infrared Physics Technology*, vol. 39, pp. 123–183, Apr. 1998.

[5] B. Rakos, H. Yang, J. Bean, G. H. Bernstein, P. Fay, and W. Porod, “Investigation of antenna-coupled MOM diodes for infrared sensor applications,” *Nonequilibrium Carrier Dynamics in Semiconductors, Springer Proceedings of the 14th International Conference*, vol.110, pp.105-108, 2006.

[6] M. Nagae, “Response time of metal-insulator-metal tunnel junctions,” *Japanese Journal of Applied Physics*, vol. 11, pp. 1611-1621, 1972.

[7] A. Sanchez, C. F. Davis, Jr., K. C. Liu, and A. Javan “The MOM tunneling diode: Theoretical estimate of its performance at microwave and infrared frequencies,” *Journal of Applied Physics*, vol. 49, pp. 5270-5277, Oct. 1978.

[8] R. H. Ritchie, “Plasma losses by fast electrons in thin films,” *Physics Review Letters*, vol. 106, no. 5, pp. 874–881, 1957.

[9] J. Cohen, "Tunnel emission into vacuum. II,” *Applied Physics Letter*, vol. 1, pp. 61-62, Oct. 1962.

[10] J.G. Simmons, “Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film,” *Journal of Applied Physics*, vol. 34, pp. 1793-1803, 1963.

[11] J.G. Simmons, “Electric tunnel effect between dissimilar electrodes separated by a thin insulating film,” *Journal of Applied Physics*, vol. 34, pp. 2581-2590, 1963.

[12] B. Hu and M. Nuss, “Imaging with terahertz waves,” *Optics Letters*, vol. 20, pp. 1716–1718, Aug. 1995.

[13] P.H. Siegel, "Terahertz technology," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 910-928, Mar. 2002.

[14] P.C. Hobbs, R. B. Laibowitz, F. R. Libsch, N. C. LaBianca, and P. P. Chiniwalla, “Effecient waveguide-integrated tunnel junction detectors at 1.6μm,” *Optics Express*, vol. 15, pp. 16376-16389, 2007.

[15] A. Balanis, *Antenna Theory*. New York: John Wiley, 1989.

[16] B. Berland, Photovoltic Technologgies Beyond the Horizon: Optical Rectenna Solar Cell, NREL, Feb. 2003.

[17] S. Krishnan, H. La Rosa, E. Stefanakos, S. Bhansali, and K. Buckle, “Design and development of batch fabricatable metal-insulator-metal diode and microstrip slot antenna as rectenna elements,” *Sensors and Actuators A*, vol. 142, no. 1, pp. 40–47, 2008.

[18] L. O. Hocker, D. R. Sokoloff, V. Daneu, A. Szoke, and A. Javan, "Frequency Mixing in the Infrared and Far-Infrared Using a Metal-to-Metal Point Contact Diode," *Applied Physics Letters*, vol. 12, no. 12, pp. 401-402, Jun. 1968.

[19] J. Bean, A. Weeks, and G. Boreman, “Performance optimization of antenna-coupled Al/AlO_X/Pt tunnel diode infrared detectors,” *IEEE Journal of Quantum Electronics*, vol. 47, no. 1, pp. 126-135, Jan. 2011.

[20] J. A. Bean, B. Tiwari, G. H. Bernstein, P. Fay, and W. Porod, “Thermal infrared detection using dipole antenna-coupled metal-oxide-metal diodes,” *Journal of Vacuum Science & Technology B*, vol. 27, no. 1, pp. 11-14, Jan. 2009.

[21] B. Tiwari, J. A. Bean, G. Szakmány, G. H. Bernstein, P. Fay, and W. Porod, “Controlled etching and regrowth of tunnel oxide for antenna-coupled metal-oxide-metal diodes,” *Journal of Vacuum Science & Technology B*: *Microelectronics and Nanometer Structures*, vol. 27, no. 5, pp. 2153-2160, Sep. 2009.

[22] E. W. Cowell, N. Alimardani, C.C. Knutson, J. F. Conley, D. A. Keszler, B. J. Gibbons, and J. F. Wager, "Advancing MIM Electronics: Amorphous Metal Electrodes," *Advanced Materials*, vol. 23, pp. 74–78, Oct. 2010.

[23] P. J. Schuck, D. P. Fromm, A. Sundaramurthy, G. S. Kino, and W. E. Moerner, "Improving the Mismatch between Light and Nanoscale Objects with Gold

Bowtie Nanoantennas," *Physical Review Letters*, vol. 94, no. 1, pp. 17402-17405, Jan. 2005.

[24] C. Puente, J. Romeu, R. Pous, X. Garcia and F. Benitez, "Fractal multiband antenna based on the Sierpinski gasket," *Electronics Letters*, vol. 32, no. 1, pp. 1-2, Jan. 1996.

[25] J. C. Liu, C.Y. Wu, D.C. Chang and C.Y. Liu, "Relationship between Sierpinski gasket and Apollonian packing monopole antennas," *Electronics Letters*, vol. 42, no. 15, pp. 847-848, Jul. 2006.

[26] S. Grover, O. Dmitriyeva, M. J. Estes, and G. Moddel, "Traveling-Wave Metal/Insulator/Metal Diodes for Improved Infrared Bandwidth and Efficiency of Antenna-Coupled Rectifiers," *IEEE Transaction on Nanotechnology*, vol. 9, no. 6, pp. 716-722, Nov. 2010.

[27] L. Liu, Z. Han and S. He, "Novel surface plasmon waveguide for high integration," *Optics Express*, vol. 13, no. 17, pp. 6645-6650, Aug. 2005.

[28] I.D. Rukhlenko, M. Premaratne, and G. P. Agrawal, "Nonlinear propagation in silicon-based plasmonic waveguides from the standpoint of applications," *Optics Express*, vol.19, no. 1, pp.206-217, Jan. 2011.

[29] S. A. Maier. Plasmonics: Fundamentals and Applications. Chapter 2. Springer, 1st edition, May. 2007.

[30] T.Y Yang. Finite Element Structural Analysis. Prentice-Hall, Inc, Englewood, NJ, 1986.

[31] C.H. Lin and C. W. Liu, "Metal-Insulator-Semiconductor Photodetectors," *Sensors*, vol. 10, no. 10, pp. 8797-8826, Sept. 2010.

[32] B. J. Eliasson, "Metal-Insulator-Metal Diodes For Solar Energy Conversion," University of Colorado at Boulder, Boulder, PhD Thesis, 2001.

[33] D. Saeedkia, "Terahertz photoconductive antennas: Principles and applications," *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pp. 3326-3328, Apr. 2011.

[34] P. K. Tien, and J. P. Gordon, "Multiphoton Process Observed in the Interaction of Microwave Fields with the Tunneling between Superconductor Films," *Physical Review*, vol. 129, no. 2, pp. 647-651, Jan. 1963.

[35] J. R. Tucker, and M. J. Feldman, "Quantum detection at millimeter wavelengths," *Reviews of Modern Physics*, vol. 57, no. 4, pp. 1055–1113, Oct. 1985.

[36] J. C. Fisher and I. Giaever, "Tunneling Through Thin Insulating Layers," *Journal of Applied Physics*, vol. 32, no. 2, pp. 172-177, Feb. 1961.

[37] T.G. Shands, and J. A. Woody, "Metal-insulator-metal junctions as surface sources of intermodulation," RADC Technical Report RADC-TR-83-31, Feb. 1983.

[38] P. K. Dubey, "Tunnel Effect in Al-Al_2O_3-Al Junctions: Inclusion of Short-Range Forces," *Journal of Applied Physics*, vol. 42, no. 6, pp. 2534-2542, May. 1971.

[39] V. Soti and Y. Ahmadizadeh, "Electric capacitance of Al/Al2O3/Al tunnel junctions," *Advanced Studies in Theoretical Physics*, vol. 5, no. 8, pp. 379-386, 2011.

[40] C. Langhammer, M. Schwind , B. Kasemo and I. Zoric, "Localized Surface Plasmon Resonances in Aluminum nanodisks," *Nano Letters*, vol. 8, no. 5, pp. 1461-1471, May. 2008.

[41] P. Maraghechi1, A. Foroughi-Abari, K. Cadien, and A. Y. Elezzabi, "Enhanced rectifying response from metal-insulator-insulator-metal junctions," *Applied Physics Letters*, vol. 99, no. 25, pp. 253503-1-4, Dec. 2011.

[42] A. Sanchez, C. F. Davis, K. C. Liu, and A. Javan, "The MOM tunneling diode: Theoretical estimate of its performance at microwave and infrared frequencies," *Journal of Applied Physics*, vol. 49, no. 10, pp. 5270-5277, Jan. 1978.

[43] S. Grover, and G. Moddel, "Applicability of Metal/Insulator/Metal (MIM) Diodes to Solar Rectennas," *IEEE Journal of Photovoltaics*, vol. 1, no. 1, pp. 78-83, Jul. 2011.

[44] E. D. Palik, and G. Ghosh, Handbook of Optical Constants of Solids, first ed., Academic Press, New York, 1985.

[45] J. R. Tucker and M. F. Millea, "Photon detection in nonlinear tunneling devices," *Applied Physics Letters*, vol. 31, no. 7, pp. 611-613, Jul. 1978.

[46] A. Malcoci, A. Stöhr, A. Sauerwald, S. Schulz, D. Jäger, "Waveguide and antenna coupled traveling-wave 1.55-μm photodetectors for optical (sub)millimeter-wave generation", *Microwave and Terahertz Photonics*, Proc. in SPIE Photonics Europe, Strasbourg, France, pp. 202-209, 2004.

[47] S Krishnan, Design, Fabrication and Characterization of Thin-Film MIM Diodes for Rectenna Array, University of South Florida, Ph.D Thesis, 2004.

[48] G. Monti, L. Corchia, and L. Tarricone, "Planar Bowtie Antenna with a Reconfigurable Radiation Pattern," *Progress in Electromagnetics Research C*, vol. 28, pp. 61-70, Mar. 2012.

[49] A. A. Lestari, A.G. Yarovoy, and L.P. Ligthart, "Adaptation capabilities of a wire bow-tie antenna for ground penetrating radar," *IEEE Antennas and Propagation Society International Symposium*, vol. 2, no. 1, pp. 564-567, 2001.

[50] W. Zhong, Y. Wang, R. He and X. Zhou, "Investigation of plasmonics resonance infrared bowtie metal antenna," *Applied Physics B: Lasers and Optics*, vol. 105, no. 2, pp. 231-237, 2011.

[51] Y. J. Chiu, S. Z. Zhang, S. B. Fleischer, J. E. Bowers, and U. K. Mishra, "GaAs-based, 1.55 μm high speed, high saturation power, low-temperature grown GaAs p-i-n photodetector," *Electronic Letters*, vol. 34, no.12, pp.1253-1255, Jun. 1998.

[52] K. S. Giboney, M. J. W. Rodwell, and J. E. Bowers, "Traveling-wave photodetectors," *IEEE Photonics Technology Letters*, vol. 4, no. 12, pp. 1363-1365, Dec. 1992.

[53] A. Yariv and P. Yeh, "*Optical waves in crystal*," Chapter 12, Wiley, 1984.

[54] N. C. Kothari, "Nonlinear light propagation in a two-component heterogeneous medium," *Journal of Optical Society of America B*, vol. 5, no. 11, pp. 2348-2356, Nov. 1988.

[55] K. Mori, Y. Tamaki, M. Obara, and K. Midorikawa, "Second-harmonic generation of femtosecond high-intensity Ti:sapphire laser pulses," *Journal of Applied Physics*, vol. 83, no. 6, pp. 2915-2919, Dec. 1997.

[56] T. Zhang, M. Yonemura, L. B. Sharma, H. Daido, Y. Kato, "Second-harmonic generation with traveling-wave pulses," *Applied Physics B*: *Lasers and Optics*, vol. 63, pp. 237-242, Feb. 1996.

[57] C. J. Powell and J. B. Swan, "Origin of the Characteristic Electron Energy Losses in Aluminum," *Physics Review Letters*, vol. 115, pp. 869-875, Aug. 1959.

[58] A. Degiron and D. R. Smith, "Nonlinear long-range plasmonic waveguides," *Physical Review A*, vol. 82, no. 3, pp. 033812(1-5), Sept. 2010.

[59] M. A. Ordal, R. J. Bell, R. W. Alexander, Jr., L. A. Newquist, and M. R. Querry, "Optical properties of Al, Fe, Ti, Ta, W, and Mo at submillimeter wavelengths," *Applied Optics*, vol. 27, no. 6, pp. 1203-1209, Mar. 1988.

[60] M. A. Ordal, R. J. Bell, R. W. Alexander, Jr., L. L. Long, and M. R. Querry, "Optical properties of fourteen metals in the infrared and far-infrared: Al, Co, Cu,

Au, Fe, Pb, Mo,Ni, Pd, Pt, Ag, Ti, V, and W," *Applied Optics*, vol. 24, no. 24, pp. 4493–4499, Dec. 1985.

[61] X. Q. Lei, and V. Van, "Nonlinear MIM nanoplasmonic waveguide based on electron tunneling for ultrafast optical pulse rectification," *Photonics Conference (IPC), 2012 IEEE*, pp. 953-954, Sept. 2012.

[62] R. Zia, M. D. Selker, P. B. Catrysse, and M. L. Brongersma, "Geometries and materials for subwavelength surface plasmon modes," *Journal of Optical Society of America A*, vol. 21, no. 12, pp. 2442-2446, Dec. 2004.

[63] G. Veronis and S. Fana, "Bends and splitters in metal-dielectric-metal subwavelength plasmonic waveguides," *Applied Physics Letters*, vol. 87, pp. 131102(1-3), Sept. 2005.

[64] F. J. Rodriguez, F. X. Wang, and M. Kauranen, "Calibration of the second-order nonlinear optical susceptibility of surface and bulk of glass," *Optics Express*, vol. 16, no. 12, pp. 8704-8710, Jun. 2008.

[65] M. Savanier, A. Andronico, A. Lema ître, E. Galopin, C. Manquest, I. Favero, S. Ducci, and G. Leo, "Large second-harmonic generation at 1.55 μmin oxidized AlGaAs waveguides," *Optics Letters*, vol. 36, no. 15, pp. 2955-2957, Aug. 2011.

[66] J. B. Han, P. Abolghasem, D.P. Kang, B. J. Bijlani, and A. S. Helmy, "Difference-frequency generation in AlGaAs Bragg reflection waveguides," *Optics Letters*, vol. 35, no. 14, pp. 2334-2336, Jul. 2010.

[67] K. S. Yee, "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media," *IEEE Transaction on Antennas Propagation*. vol. 14, no. 3, pp. 302-307, May. 1966.

[68] Z. P. Wang, "Generation of Terahertz Radiation via Nonlinear Optical Methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 1, no. 1, pp. 1-5, Nov. 2010.

[69] A. Taflove and S. C. Hagness: Computational Electrodynamics: The Finite-Difference Time-Domain Mtehod, pp. 361-368. Artech House, Inc., Norwood, MA, 2003.

[70] M. A. Green and J. Shewchun, "Capacitance properties of MIS tunnel diodes," *Journal of Applied Physics*, vol. 46, no. 12, pp. 5185-5190, Dec. 1975.

[71] P. R. West, S. Ishii, G. V. Naik, N. K. Emani,V. M. Shalaev, and A. Boltasseva, "Searching for better plasmonic materials," *Laser & Photonics Reviews*, vol. 4, no. 6, pp. 795-808, Nov. 2010.

[72] H. R. Park, J. M. Park, M. S. Kim, and M. H. Lee, "A waveguide-typed plasmonic mode converter," *Optics Express*, vol. 20, no. 17, pp. 18636-18645, Aug. 2012.

[73] Z. Han, A. Y. Elezzabi, and V. Van, "Experimental realization of subwavelength plasmonic slot waveguides on a silicon platform," Optics Letters, vol. 35, no. 4, pp. 502-504, Feb. 2010.

[74] Z. Han, A. Y. Elezzabi, and V. Van, "Wideband Y-splitter and aperture-assisted coupler based on sub-diffraction confined plasmonic slot waveguides," *Applied Physics Letters*, vol. 96, pp. 131106(1-3), Mar. 2010.

[75] L. Tang, S. Latif, and D.A.B. Miller, "Plasmonic device in silicon CMOS," *Electronics Letters*, vol.45, no.13, pp.706-708, Jun. 2009.

[76] V. E. Babicheva and A. V. Lavrinenko, "Plasmonic Modulator Optimized by Patterning of Active Layer and Tuning Permittivity," arXiv:1202.6559v1, 2012.

[77] Y. C. Yeo, T. j. King, and C. M. Hu, "Metal-dielectric band alignment and its implications for metal gate complementary metal-oxide-semiconductor technology," *Journal of Applied Physics*, vol. 92, no. 12, pp. 7266-7271, Dec. 2002.

[78] S. K. Gupta, A. K. Kapil, C. M. Singal, and V. K. Srivastava, "Measurement of the work function of some metals using internal voltage in MIM structures," *Journal of Applied Physics*, vol. 50, no. 4, pp. 2852-2855, Apr. 1979.

[79] T. O'Regan, M. Chin, C. Tan, and A. Birdwell, "Modeling, Fabrication, and Electrical Testing of Metal-Insulator-Metal Diode," Army Research Laboratory, ARL-TN-0464, USA, Dec. 2011.

[80] A. V. Krasavin and A. V. Zayats, "Silicon-based plasmonic waveguides," *Optics Express*, vol. 18, no. 11, pp. 11791-11799, May. 2010.

[81] D. Woolf, M. Loncar, and F. Capasso, "The forces from coupled surface plasmon polaritons in planar waveguides," *Optics Express*, vol. 17, no. 22, pp. 19996-20011, Oct. 2009.

[82] S. Chand, and S. Bala, "Simulation studies of current transport in metal–insulator–semiconductor Schottky barrier diodes," *Physica B: Condensed Matter*, vol. 390, no. 1–2, pp. 179-184, Mar. 2007.

[83] K. S. Giboney, M. J. W. Rodwell, and J. E. Bowers, "Traveling-wave photodetectors," *IEEE Photonics Technology Letters*, vol. 4, no. 12, pp. 1363-1365, Dec. 1992.

[84] M. J. Weber, Handbook of Optical Materials, CRC Press, 2003.

[85] M. Wu, Z. H. Han, and V. Van, "Conductor-gap-silicon plasmonic waveguides and passive components at subwavelength scale," *Optics Express*, vol. 18, no. 11, pp. 11728-11736, May. 2010.

[86] M. A. Green, F. D. King, and J. Shewchun, "Minority carrier MIS tunnel diodes and their application to electron- and photo-voltaic energy conversion-I. Theory," *Solid-State Electronics*, vol. 17, no. 6, pp. 551-561, Jun. 1974.

Appendix A

Simulation Results for the Multi-arm Bowtie Antennas

This Appendix provides additional results for our multi-arm bowtie antenna designs that are not given in chapter 2. Three antenna arm length investigated for our designs are λ_e, $1.5\lambda_e$ and $2\lambda_e$, where λ_e (58μm) is the effective wavelength at 1.5THz on a Si substrate (ε_{Si} =11.9 [44]). In here, we will present results for the different antennas and compare with the traditional single-arm bowtie antenna.

A.1 Bowtie Antenna with a 60^o Flare Angle

The gain profiles, radiation efficiency and radiation patterns of the bowtie antenna with a flare angle of 60^o are given in Fig. A.1, showing good antenna gain and radiation efficiency.

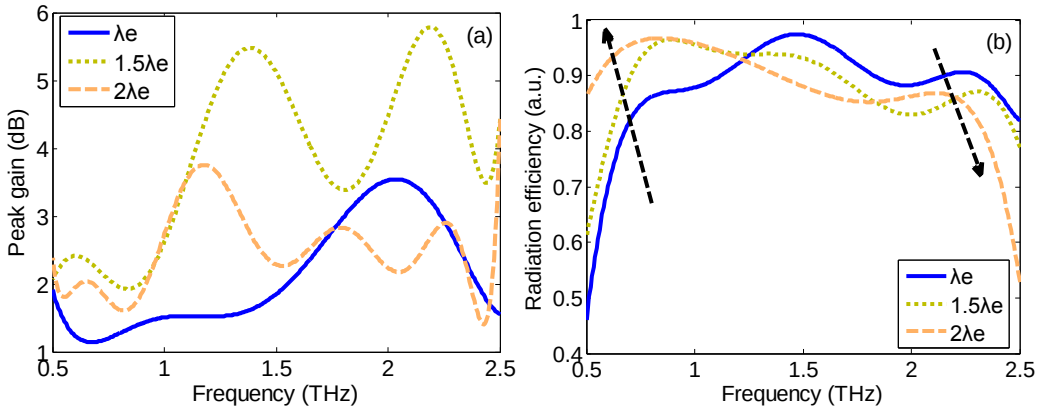


Figure A.1 Plots of (a) frequency dependent peak gain and (b) radiation efficiency of the bowtie antenna with a flare angle of 60^o at arm length of λ_e, $1.5\lambda_e$ and $2\lambda_e$.

Antenna radiation patterns provide good visualizations of the directivity. Fig. A.2 plots the radiation patterns (x-z and y-z plane cuts) of the bowtie antennas with the three different arm lengths at frequencies 0.8THz, 1.5THz and 2THz. The 0^o angle corresponds to the positive Z direction. The radiation patterns

are observed to gradually change with increasing frequency and the bowtie antenna favours receiving from the direction normal to substrate surface.

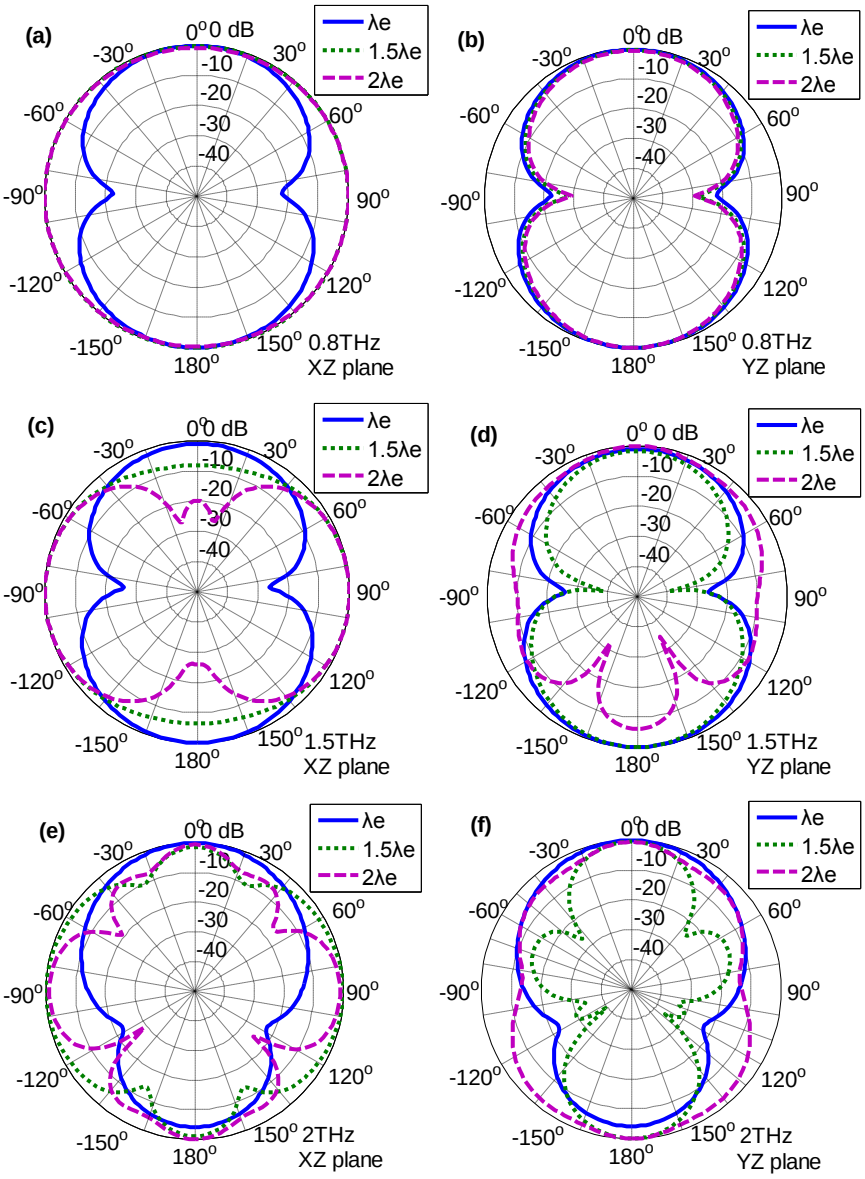


Figure A.2 Radiation patterns of the bowtie antenna with a flare angle of 60° at arm length of λ_e, $1.5\lambda_e$ and $2\lambda_e$ (λ_e=58μm at 1.5THz on Si substrate) at frequencies of 0.8THz, 1.5THz and 2THz. The radiation patterns were observed at the XZ and YZ planes.

A.2 Multi-arm Bowtie Antenna

We next present results for the proposed multi-arm bowtie antennas in terms of antenna impedance, S11, peak gain, efficiency, and radiation patters.

A.2.1 Antenna arm length of λ_e

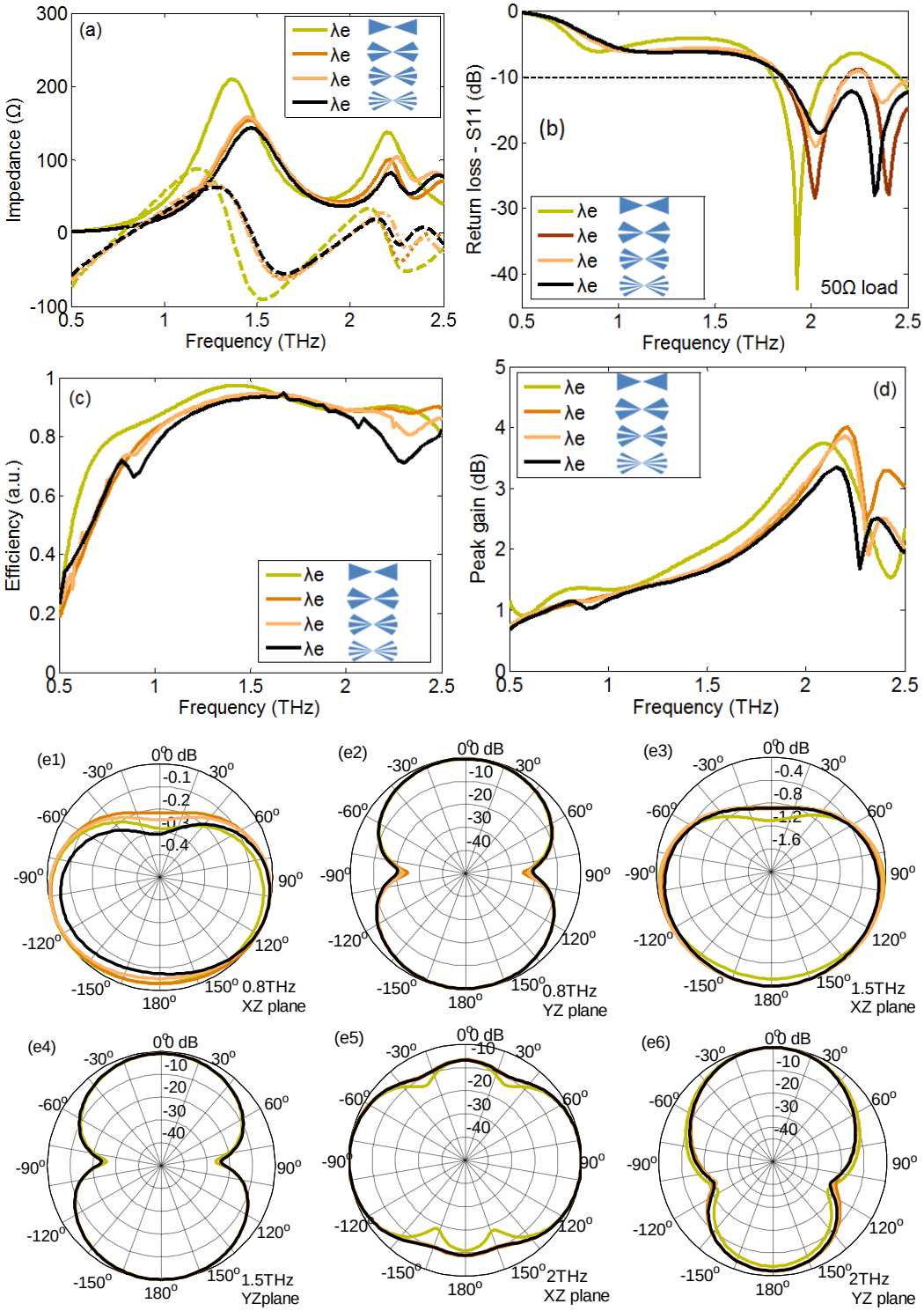


Figure A.3 Plots of (a) reactance (dashed), resistance (solid), (b) S11, (c) efficiency, (d) peak gain, and (e) radiation patterns: 1) XZ plane/0.8THz, 2) YZ plane/0.8THz, 3) XZ plane/1.5THz, 4) YZ plane/1.5THz, 5) XZ plane/2THz and 6) YZ plane/ 2THz for Bowtie, 2×30°, 3×20° and 4×15° with λe arm length.

A.2.2 Antenna arm length of $1.5\lambda_e$

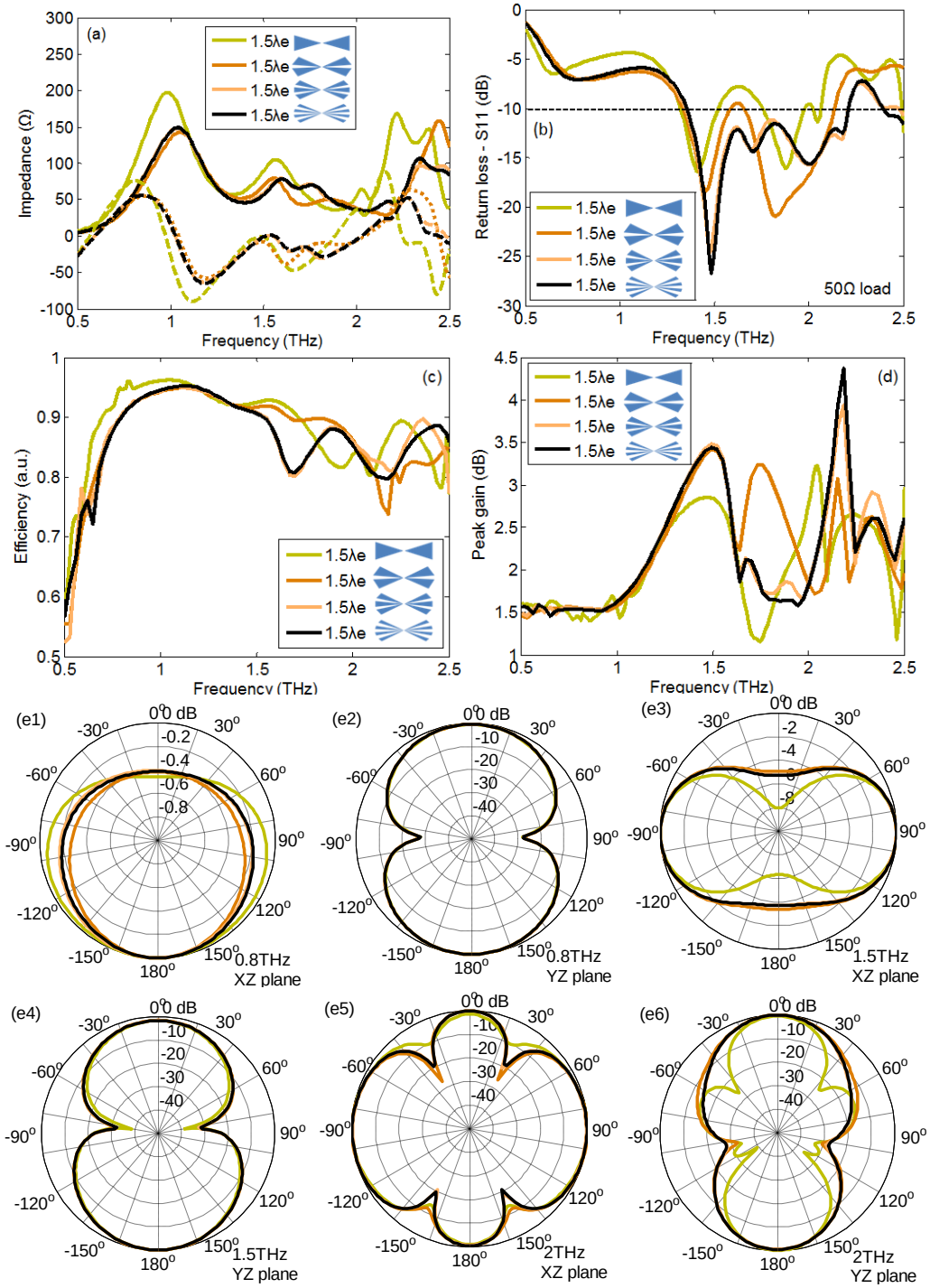


Figure A.4 Plots of (a) reactance (dashed), resistance (solid), (b) S11, (c) efficiency, (d) peak gain, and (e) radiation patterns: 1) XZ plane/0.8THz, 2) YZ plane/0.8THz, 3) XZ plane/1.5THz, 4) YZ plane/1.5THz, 5) XZ plane/2THz and 6) YZ plane/ 2THz for Bowtie, 2×30°, 3×20° and 4×15° with 1.5λe arm length.

A.2.3 Antenna arm length of $2\lambda_e$

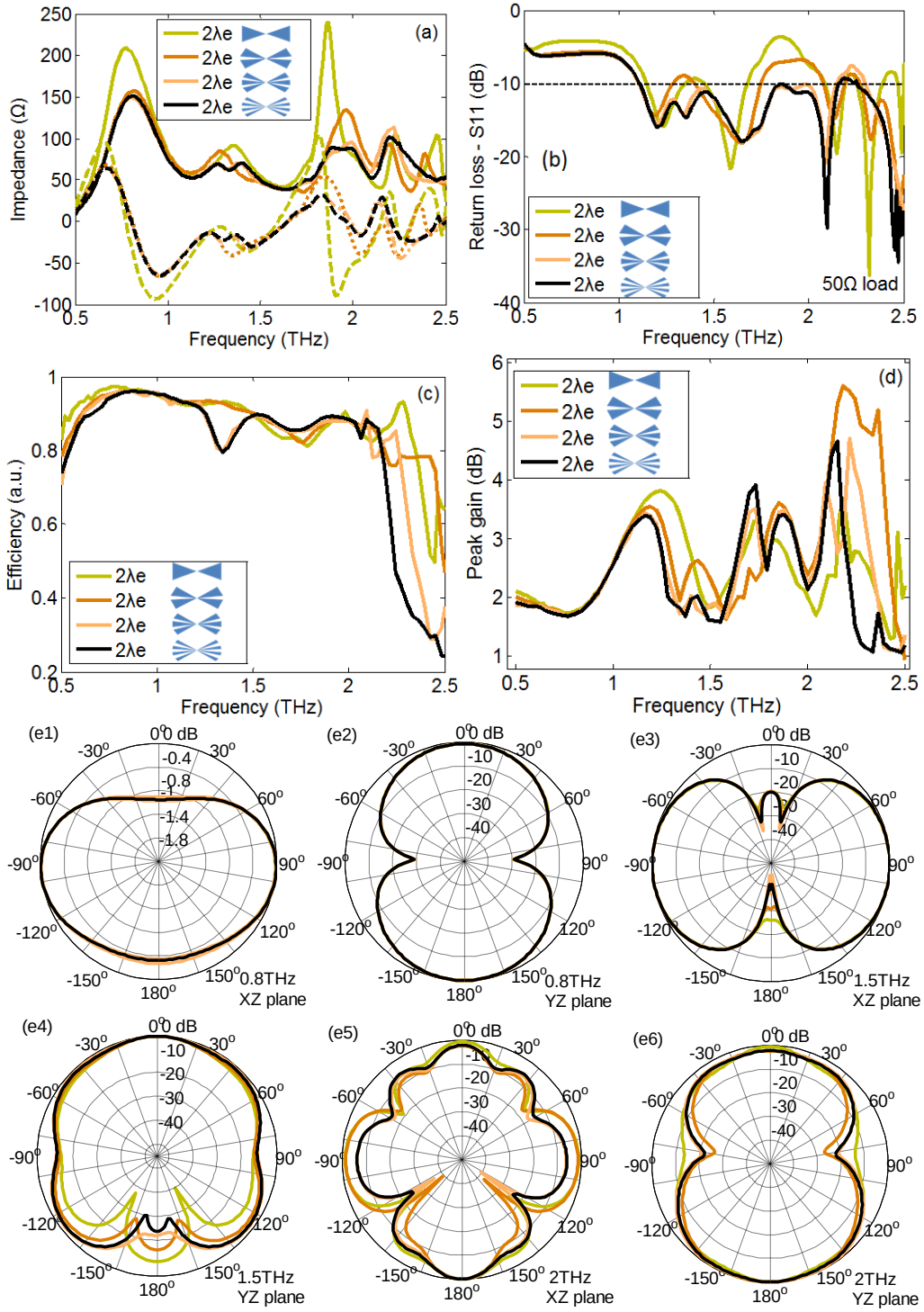


Figure A.5 Plots of (a) reactance (dashed), resistance (solid), (b) S11, (c) efficiency, (d) peak gain, and (e) radiation patterns: 1) XZ plane/0.8THz, 2) YZ plane/0.8THz, 3) XZ plane/1.5THz, 4) YZ plane/1.5THz, 5) XZ plane/2THz and 6) YZ plane/ 2THz for Bowtie, 2×30°, 3×20° and 4×15° with 2λe arm length.

In the multi-arm configuration, the antenna generally demonstrates lower reactance and resistance compared to the traditional bowtie antenna. This is primarily a consequence of the parallel connection of the multi-arms. Also, the impedance frequency response of the multi-arm antenna has less fluctuation compared to the traditional bowtie antenna. As the impedance curve flattens out, the antenna bandwidth increases which is evident from the total or single-resonance bandwidths. The new antenna design also demonstrates very high gain and efficiency. In summary, similar results are identified in the antenna impedance, return loss, gain and efficiency for different antenna arm lengths. The most distinctive behavior is observed in the antenna radiation patterns. At shorter effective arm length, the radiation patterns of different multi-arm antennas are fairly similar to one another. As the effective arm length is increased, their radiation patterns diverge. This behavior is caused by effectively broadened arm width at longer arm length, especially for the four-arm antenna with large total split and flare angles.

Appendix B

Nonlinear CW Wave Propagation in MIM Nanoplasmonic Waveguides

Derivations in Appendix B, and C are based on works done by Kothari [54] and Yariv and Yeh [53]. With high power and modal field confinements in the insulator layer, the MIM nanoplasmonic waveguide structure could be treated as a one-dimensional (1D) effective medium [54] with effective permittivity ε_R and conductivity σ_e. We analyze the 1D effective medium using the coupled mode theory (CMT) by assuming linearly polarized waves propagating in the *x* direction. For generalized cases, the total electric field in the waveguide is the sum of all field components as

$$E = \sum_1^i \tilde{\tilde{E}}_i \,. \tag{B.1}$$

where $\tilde{\tilde{E}}_i$ is the electric field of the component *i*. The wave equation for propagation in the 1D nonlinear medium is given by

$$\frac{\partial^2 E}{\partial x^2} = \frac{\varepsilon_R}{c^2}\frac{\partial^2 E}{\partial t^2} + \mu_o \sigma_e \frac{\partial E}{\partial t} + \mu_o \sigma_2 \frac{\partial E^2}{\partial t}, \tag{B.2}$$

We use the coupled mode analysis to study nonlinear CW waves mixing (assuming two input CW waves $E_2(x)$ and $E_3(x)$) in the MIM waveguide and investigate frequency difference $E_1(x)$ ($\omega_1 = \omega_3 - \omega_2$) (with $\omega_3 > \omega_2$) and frequency sum generations $E_4(x)$ ($\omega_4 = \omega_3 + \omega_2$). The total electric field includes the following components

$$\tilde{\tilde{E}}_1(x,t) = \tilde{\tilde{E}}_{\omega 3-\omega 2}(x,t) = \frac{1}{2}\{E_1(x)e^{j(\beta_1 x-\omega_1 t)} + c.c.\}, \tag{B.3a}$$

$$\tilde{\tilde{E}}_2(x,t) = \tilde{\tilde{E}}_{\omega 2}(x,t) = \frac{1}{2}\{E_2(x)e^{j(\beta_2 x-\omega_2 t)} + c.c.\}, \tag{B.3b}$$

$$\tilde{\tilde{E}}_3(x,t) = \tilde{\tilde{E}}_{\omega 3}(x,t) = \frac{1}{2}\{E_3(x)e^{j(\beta_3 x-\omega_3 t)} + c.c.\}, \tag{B.3c}$$

$$\tilde{\tilde{E}}_4(x,t) = \tilde{\tilde{E}}_{\omega 4}(x,t) = \frac{1}{2}\{E_4(x)e^{j(\beta_4 x-\omega_4 t)} + c.c.\}, \tag{B.3d}$$

where $E_i(x)$ represents the slowly varying envelopes of the field components along the propagation direction due to both metal loss and rectification. Substituting Eqs. B.3 into Eq. B.2, forcing the frequency matching condition and expanding the derivatives, we obtain four coupled wave equations

$$\begin{aligned}
&\frac{\partial^2 E_1}{\partial x^2}e^{j(\beta_1 x-\omega_1 t)} + 2\frac{\partial E_1}{\partial x}(j\beta_1)e^{j(\beta_1 x-\omega_1 t)} + E_1(j\beta_1)^2 e^{j(\beta_1 x-\omega_1 t)} \\
&= \frac{\varepsilon_{1R}}{c^2}[E_1(-j\omega_1)^2 e^{j(\beta_1 x-\omega_1 t)}] \\
&+ \mu_o\sigma_{1e}[E_1(-j\omega_1)e^{j(\beta_1 x-\omega_1 t)}] + \mu_o\sigma_2[E_3E_2^* j(\omega_2-\omega_3)e^{j(\beta_3 x-\omega_3 t-\beta_2 x+\omega_2 t)}],
\end{aligned} \tag{B.4a}$$

$$\begin{aligned}
&\frac{\partial^2 E_2}{\partial x^2}e^{j(\beta_2 x-\omega_2 t)} + 2\frac{\partial E_2}{\partial x}(j\beta_2)e^{j(\beta_2 x-\omega_2 t)} + E_2(j\beta_2)^2 e^{j(\beta_2 x-\omega_2 t)} \\
&= \frac{\varepsilon_{2R}}{c^2}[E_2(-j\omega_2)^2 e^{j(\beta_2 x-\omega_2 t)}] \\
&+ \mu_o\sigma_{2e}[E_2(-j\omega_2)e^{j(\beta_2 x-\omega_2 t)}] + \mu_o\sigma_2[E_3E_1^* j(-\omega_3+\omega_1)e^{j(\beta_3 x-\omega_3 t-\beta_1 x+\omega_1 t)}] \\
&+ \mu_o\sigma_2[E_4E_3^* j(-\omega_4+\omega_3)e^{j(\beta_4 x-\omega_4 t-\beta_3 x+\omega_3 t)}],
\end{aligned} \tag{B.4b}$$

$$\begin{aligned}
&\frac{\partial^2 E_3}{\partial x^2}e^{j(\beta_3 x-\omega_3 t)} + 2\frac{\partial E_3}{\partial x}(j\beta_3)e^{j(\beta_3 x-\omega_3 t)} + E_3(j\beta_3)^2 e^{j(\beta_3 x-\omega_3 t)} \\
&= \frac{\varepsilon_{3R}}{c^2}[E_3(-j\omega_3)^2 e^{j(\beta_3 x-\omega_3 t)}] \\
&+ \mu_o\sigma_{3e}[E_3(-j\omega_3)e^{j(\beta_3 x-\omega_3 t)}] + \mu_o\sigma_2[E_4E_2^* j(-\omega_4+\omega_2)e^{j(\beta_4 x-\omega_4 t-\beta_2 x+\omega_2 t)}] \\
&+ \mu_o\sigma_2[E_1E_2 j(-\omega_1-\omega_2)e^{j(\beta_1 x-\omega_1 t+\beta_2 x-\omega_2 t)}],
\end{aligned} \tag{B.4c}$$

$$\begin{aligned}
&\frac{\partial^2 E_4}{\partial x^2}e^{j(\beta_4 x-\omega_4 t)} + 2\frac{\partial E_4}{\partial x}(j\beta_4)e^{j(\beta_4 x-\omega_4 t)} + E_4(j\beta_4)^2 e^{j(\beta_4 x-\omega_4 t)} \\
&= \frac{\varepsilon_{4R}}{c^2}[E_4(-j\omega_4)^2 e^{j(\beta_4 x-\omega_4 t)}] \\
&+ \mu_o\sigma_{4e}[E_4(-j\omega_4)e^{j(\beta_4 x-\omega_4 t)}] + \mu_o\sigma_2[E_2E_3 j(-\omega_2-\omega_3)e^{j(\beta_2 x-\omega_2 t+\beta_3 x-\omega_3 t)}].
\end{aligned} \tag{B.4d}$$

The asterisk represents complex conjugate the above equations are simplified to be

$$\frac{\partial^2 E_1}{\partial x^2} + 2\frac{\partial E_1}{\partial x}(j\beta_1) + E_1(j\beta_1)^2$$
$$= \frac{\varepsilon_{1R}}{c^2} E_1(-j\omega_1)^2 + \mu_o\sigma_{1e}E_1(-j\omega_1) + \mu_o\sigma_2 E_3E_2^* j(\omega_2 - \omega_3)e^{j(\beta_3-\beta_2-\beta_1)x},$$

(B.5a)

$$\frac{\partial^2 E_2}{\partial x^2} + 2\frac{\partial E_2}{\partial x}(j\beta_2) + E_2(j\beta_2)^2$$
$$= \frac{\varepsilon_{2R}}{c^2} E_2(-j\omega_2)^2 + \mu_o\sigma_{2e}E_2(-j\omega_2) + \mu_o\sigma_2 E_3E_1^* j(-\omega_3 + \omega_1)e^{j(\beta_3-\beta_1-\beta_2)x}$$
$$+ \mu_o\sigma_2 E_4E_3^* j(-\omega_4 + \omega_3)e^{j(\beta_4-\beta_3-\beta_2)x}$$

(B.5b)

$$\frac{\partial^2 E_3}{\partial x^2} + 2\frac{\partial E_3}{\partial x}(j\beta_3) + E_3(j\beta_3)^2$$
$$= \frac{\varepsilon_{3R}}{c^2} E_3(-j\omega_3)^2 + \mu_o\sigma_{3e}E_3(-j\omega_3) + \mu_o\sigma_2 E_4E_2^* j(-\omega_4 + \omega_2)e^{j(\beta_4-\beta_2-\beta_3)x}$$
$$+ \mu_o\sigma_2 E_1E_2 j(-\omega_1 - \omega_2)e^{j(\beta_1+\beta_2-\beta_3)x},$$

(B.5c)

$$\frac{\partial^2 E_4}{\partial x^2} + 2\frac{\partial E_4}{\partial x}(j\beta_4) + E_4(j\beta_4)^2$$
$$= \frac{\varepsilon_{4R}}{c^2} E_4(-j\omega_4)^2 + \mu_o\sigma_{4e}E_4(-j\omega_4) + \mu_o\sigma_2 E_2E_3 j(-\omega_2 - \omega_3)e^{j(\beta_2+\beta_3-\beta_4)x},$$

(B.5d)

Using the constitutive relation in Eq. B.3 and the condition that

$$\frac{\partial^2 E_i}{\partial x^2} << 2\frac{\partial E_i}{\partial x}(j\beta_i), \tag{B.6}$$

Eqs. B.5 are further simplified into

$$\frac{\partial E_1}{\partial x} = -\frac{\mu_o\sigma_{1e}\omega_1}{2\beta_1}E_1 - \frac{\mu_o\sigma_2\omega_1}{2\beta_1}E_3E_2^* e^{j(\beta_3-\beta_2-\beta_1)x}, \tag{B.7a}$$

$$\frac{\partial E_2}{\partial x}$$
$$= -\frac{\mu_o \sigma_{2e} \omega_2}{2\beta_2} E_2 - \frac{\mu_o \sigma_2 \omega_2}{2\beta_2} E_3 E_1^* e^{j(\beta_3 - \beta_1 - \beta_2)x} - \frac{\mu_o \sigma_2 \omega_2}{2\beta_2} E_4 E_3^* e^{j(\beta_4 - \beta_3 - \beta_2)x},$$

(B.7b)

$$\frac{\partial E_3}{\partial x}$$
$$= -\frac{\mu_o \sigma_{3e} \omega_3}{2\beta_3} E_3 - \frac{\mu_o \sigma_2 \omega_3}{2\beta_3} E_4 E_2^* e^{j(\beta_4 - \beta_2 - \beta_3)x} - \frac{\mu_o \sigma_2 \omega_3}{2\beta_3} E_1 E_2 e^{j(\beta_1 + \beta_2 - \beta_3)x},$$

(B.7c)

$$\frac{\partial E_4}{\partial x} = -\frac{\mu_o \sigma_{4e} \omega_4}{2\beta_4} E_4 - \frac{\mu_o \sigma_2 \omega_4}{2\beta_4} E_2 E_3 e^{j(\beta_2 + \beta_3 - \beta_4)x}. \quad \text{(B.7d)}$$

To obtain simplified analytical solutions to the above equations, we notice that the following condition can be applied (with $E_2 >> E_1$ and $E_3 >> E_4$, the terms which contribute to the input signals due to second order nonlinear conductive effects are negligible compared to the input signals).

$$|\sigma_{2e} E_2| >> |\sigma_2 E_3 E_1^*| \; or \; |\sigma_2 E_4 E_3^*|, \quad \text{(B.8a)}$$

$$|\sigma_{3e} E_3| >> |\sigma_2 E_4 E_2^*| \; or \; |\sigma_2 E_1 E_2|. \quad \text{(B.8b)}$$

Eqs. B.7 are rearranged as

$$\frac{\partial E_1}{\partial x} = -\frac{\mu_o \sigma_{1e} \omega_1}{2\beta_1} E_1 - \frac{\mu_o \sigma_2 \omega_1}{2\beta_1} E_3 E_2^* e^{j(\beta_3 - \beta_2 - \beta_1)x}, \quad \text{(B.9a)}$$

$$\frac{\partial E_2}{\partial x} = -\frac{\mu_o \sigma_{2e} \omega_2}{2\beta_2} E_2, \quad \text{(B.9b)}$$

$$\frac{\partial E_3}{\partial x} = -\frac{\mu_o \sigma_{3e} \omega_3}{2\beta_3} E_3, \quad \text{(B.9c)}$$

$$\frac{\partial E_4}{\partial x} = -\frac{\mu_o \sigma_{4e} \omega_4}{2\beta_4} E_4 - \frac{\mu_o \sigma_2 \omega_4}{2\beta_4} E_2 E_3 e^{j(\beta_2 + \beta_3 - \beta_4)x}. \quad \text{(B.9d)}$$

With the coordinate transformation

$$E_i = \frac{A_i}{\sqrt{n_i}} \propto \frac{\sqrt{P_i}}{\sqrt{n_i}}, \quad \text{(B.10)}$$

Eqs. B.9 become

$$\frac{\partial A_1}{\partial x} = -\frac{\mu_o \sigma_{1e} \omega_1}{2\beta_1} A_1 - \frac{\mu_o \sigma_2 \omega_1}{2\beta_1} \frac{\sqrt{n_1}}{\sqrt{n_2 n_3}} A_3 A_2^* e^{j(\beta_3 - \beta_2 - \beta_1)x}, \tag{B.11a}$$

$$\frac{\partial A_2}{\partial x} = -\frac{\mu_o \sigma_{2e} \omega_2}{2\beta_2} A_2, \tag{B.11b}$$

$$\frac{\partial A_3}{\partial x} = -\frac{\mu_o \sigma_{3e} \omega_3}{2\beta_3} A_3, \tag{B.11c}$$

$$\frac{\partial A_4}{\partial x} = -\frac{\mu_o \sigma_{4e} \omega_4}{2\beta_4} A_4 - \frac{\mu_o \sigma_2 \omega_4}{2\beta_4} \frac{\sqrt{n_4}}{\sqrt{n_2 n_3}} A_2 A_3 e^{j(\beta_2 + \beta_3 - \beta_4)x}. \tag{B.11d}$$

Defining the following coefficients

$$\alpha_i = \frac{\omega_i \mu_o \sigma_{ie}}{2\beta_i}, \tag{B.12a}$$

$$\kappa_1 = \frac{\mu_o \sigma_2 \omega_1}{2\beta_1} \frac{\sqrt{n_1}}{\sqrt{n_2 n_3}}, \tag{B.12b}$$

$$\kappa_4 = \frac{\mu_o \sigma_2 \omega_4}{2\beta_4} \frac{\sqrt{n_4}}{\sqrt{n_2 n_3}}. \tag{B.12c}$$

Eqs. B.11 become

$$\frac{\partial A_1}{\partial x} = -\alpha_1 A_1 - \kappa_1 A_3 A_2^* e^{j(\beta_3 - \beta_2 - \beta_1)x}, \tag{B.13a}$$

$$\frac{\partial A_2}{\partial x} = -\alpha_2 A_2, \tag{B.13b}$$

$$\frac{\partial A_3}{\partial x} = -\alpha_3 A_3, \tag{B.13c}$$

$$\frac{\partial A_4}{\partial x} = -\alpha_4 A_4 - \kappa_4 A_2 A_3 e^{j(\beta_2 + \beta_3 - \beta_4)x}. \tag{B.13d}$$

Analytical solution to Eqs. B.13a and B.13b can be obtained by integrating on both sides of

$$\frac{\partial A_2}{A_2} = -\alpha_2 \partial x, \tag{B.14a}$$

$$\frac{\partial A_3}{A_3} = -\alpha_3 \partial x, \tag{B.14b}$$

to give

$$A_2 = C_2 e^{-\alpha_2 x}, \tag{B.15a}$$

$$A_3 = C_3 e^{-\alpha_3 x}, \tag{B.15b}$$

where C_i is an arbitrary constant. Applying the condition at $x = 0$ that

$$A_2(0) = C_2, \tag{B.16a}$$

$$A_3(0) = C_3, \tag{B.16b}$$

we obtain

$$A_2(x) = A_2(0) e^{-\alpha_2 x}, \tag{B.17a}$$

$$A_3(x) = A_3(0) e^{-\alpha_3 x}. \tag{B.17b}$$

Substituting Eqs. B.17 into Eqs. B.13a and B.13d yields

$$\frac{\partial A_1}{\partial x} + \alpha_1 A_1 = -\kappa_1 A_2(0) A_3(0) e^{[j(\beta_3 - \beta_2 - \beta_1) - \alpha_2 - \alpha_3]x}, \tag{B.18a}$$

$$\frac{\partial A_4}{\partial x} + \alpha_4 A_4 = -\kappa_4 A_2(0) A_3(0) e^{[j(\beta_2 + \beta_3 - \beta_4) - \alpha_2 - \alpha_3]x}. \tag{B.18b}$$

Solution to the above equations can be obtained by using the method of integrating factors which is illustrated as follows for a general differential equation of the form

$$\frac{dy}{dx} + p(x) y = q(x) \cdot \tag{B.19a}$$

The integration factor F is defined as

$$F = e^{\int P(x) dx}. \tag{B.19b}$$

Multiply both sides of Eq. B.19a by the integration factor F gives

$$\frac{dy}{dx} e^{\int p(x) dx} + p(x) y e^{\int p(x) dx} = q(x) e^{\int p(x) dx}. \tag{B.19c}$$

With the relation that

$$\frac{dy}{dx} e^{\int p(x) dx} + p(x) y e^{\int p(x) dx} = \frac{d}{dx}[e^{\int p(x) dx} y], \tag{B.19d}$$

Eq. B.19c becomes

$$\frac{d}{dx}[e^{\int p(x)dx} y] = q(x)e^{\int p(x)dx}. \tag{B.19e}$$

Integrating on both sides of Eq. B.19e gives

$$e^{\int p(x)dx} y + C = \int q(x)e^{\int p(x)dx} dx, \tag{B.19f}$$

where C is an arbitrary constant. Then y is obtained from

$$y = \frac{\int q(x)e^{\int p(x)dx} dx - C}{e^{\int p(x)dx}}. \tag{B.19g}$$

Using the above approach, Eqs. B.18 can be solved to give

$$A_1(x) = \frac{\int -\kappa_1 A_2(0)A_3(0)e^{[j(\beta_3-\beta_2-\beta_1)-\alpha_2-\alpha_3]x} e^{\alpha_1 x} dx - C_1}{e^{\alpha_1 x}}, \tag{B.20a}$$

$$A_4(x) = \frac{\int -\kappa_4 A_2(0)A_3(0)e^{[j(\beta_2+\beta_3-\beta_4)-\alpha_2-\alpha_3]x} e^{\alpha_4 x} dx - C_4}{e^{\alpha_4 x}}. \tag{B.20b}$$

And subsequently as

$$A_1(x) = \frac{-\kappa_1 A_2(0)A_3(0)\dfrac{e^{[j(\beta_3-\beta_2-\beta_1)-\alpha_2-\alpha_3+\alpha_1]x}}{[j(\beta_3-\beta_2-\beta_1)-\alpha_2-\alpha_3+\alpha_1]} - C_1}{e^{\alpha_1 x}}, \tag{B.21a}$$

$$A_4(x) = \frac{-\kappa_2 A_2(0)A_3(0)\dfrac{e^{[j(\beta_2+\beta_3-\beta_4)-\alpha_2-\alpha_3+\alpha_4]x}}{[j(\beta_2+\beta_3-\beta_4)-\alpha_2-\alpha_3+\alpha_4]} - C_4}{e^{\alpha_4 x}}. \tag{B.21b}$$

Using the conditions that $A_1(0)=0$ and $A_4(0)=0$, we obtain

$$C_1 = -\frac{\kappa_1 A_2(0)A_3(0)}{[j(\beta_3-\beta_2-\beta_1)-\alpha_2-\alpha_3+\alpha_1]}, \tag{B.22a}$$

$$C_4 = -\frac{\kappa_2 A_2(0)A_3(0)}{[j(\beta_2+\beta_3-\beta_4)-\alpha_2-\alpha_3+\alpha_4]}. \tag{B.22b}$$

With the reverse coordinate transformation, analytical solutions to the four field components are given as

$$E_1(x) = \frac{\sqrt{n_2 n_3}\kappa_1 E_2(0)E_3(0)}{\sqrt{n_1}[j(\beta_3-\beta_2-\beta_1)-\alpha_2-\alpha_3+\alpha_1]}\{1-e^{[j(\beta_3-\beta_2-\beta_1)-\alpha_2-\alpha_3+\alpha_1]x}\}e^{-\alpha_1 x}, \tag{B.23a}$$

$$E_2(x) = E_2(0)e^{-\alpha_2 x}, \tag{B.23b}$$

$$E_3(x) = E_3(0)e^{-\alpha_3 x}, \tag{B.23c}$$

$$E_4(x) = \frac{\sqrt{n_2 n_3}\kappa_4 E_2(0)E_3(0)}{\sqrt{n_4}[j(\beta_3+\beta_2-\beta_4)-\alpha_2-\alpha_3+\alpha_4]}\{1-e^{[j(\beta_3+\beta_2-\beta_4)-\alpha_2-\alpha_3+\alpha_4]x}\}e^{-\alpha_4 x}. \tag{B.23d}$$

The magnitudes of the above equations describe the amplitude behaviours of the input CW waves and the frequency difference and sum waves as they propagate in the MIM waveguide.

When the input signals are pulse signals, the above results can also be used to obtain simplified analytical solutions for the bandwidth performance of nonlinear pulse propagation in the MIM nanoplasmonic waveguide. A Gaussian pulse can be expressed as

$$E_i(t) = E_i e^{-\frac{\rho t^2}{\tau_i^2}}, \tag{B.24a}$$

$$\rho = 4\log_e(2). \tag{B.24b}$$

where τ_i is the full width half magnitude (FWHM) pulse width of the input Gaussian pulse *i*. By substituting Eq. B.24 into Eqs. B.25, we obtain the normalized envelopes of the output components as

$$|E_{U1}(t)| = e^{-\frac{\rho t^2}{\tau_2^2}} e^{-\frac{\rho t^2}{\tau_3^2}} = e^{-\frac{\rho t^2}{\tau_2^2\tau_3^2/(\tau_2^2+\tau_3^2)}} = e^{-\frac{\rho t^2}{\tau_1^2}}, \tag{B.25a}$$

$$|E_{U2}(t)| = e^{-\frac{\rho t^2}{\tau_2^2}}, \tag{B.25b}$$

$$|E_{U3}(t)| = e^{-\frac{\rho t^2}{\tau_3^2}}, \tag{B.25c}$$

$$|E_{U4}(t)| = e^{-\frac{\rho t^2}{\tau_2^2}} e^{-\frac{\rho t^2}{\tau_3^2}} = e^{-\frac{\rho t^2}{\tau_2^2\tau_3^2/(\tau_2^2+\tau_3^2)}} = e^{-\frac{\rho t^2}{\tau_4^2}}, \tag{B.25d}$$

The above expressions indicate the FWHM pulse widths of the output components.

Appendix C

Nonlinear Pulse Rectification, Second Harmonic Generation and Two Pulses Mixing

In the case of one excitation optical pulse, the total electric field contains field components of the rectified, input and second harmonic pulses which are given as

$$\tilde{E}_0(x,t) = E_0(x,t)e^{-\alpha_o x}, \tag{C.1a}$$

$$\tilde{E}_i(x,t) = \frac{1}{2}\{E_i(x,t)e^{j(\beta_i x-\omega_i t)} + c.c\}, \tag{C.1b}$$

where $E_i(x,t)$ represents the slowly varying envelopes of the different pulse components. Substituting Eqs. C.1 into Eq. B.2 and forcing the frequency matching condition, we obtain the propagating equations for the different fields as

$$\begin{aligned}
&(\frac{\partial^2 E_0}{\partial x^2} - 2\alpha_0 \frac{\partial E_0}{\partial x} + \alpha_0^2 E_0)e^{-\alpha_0 x} \\
&= \frac{\varepsilon_{0R}}{c^2}\frac{\partial^2 E_0}{\partial t^2}e^{-\alpha_0 x} + \mu_o\sigma_{0e}\frac{\partial E_0}{\partial t}e^{-\alpha_0 x} + \mu_o\sigma_2(\frac{\partial|E_1|^2}{\partial t} + \frac{\partial|E_2|^2}{\partial t}),
\end{aligned} \tag{C.2a}$$

$$\begin{aligned}
&\frac{\partial^2 E_1}{\partial x^2}e^{j(\beta_1 x-\omega_1 t)} + 2\frac{\partial E_1}{\partial x}(j\beta_1)e^{j(\beta_1 x-\omega_1 t)} + E_1(j\beta_1)^2 e^{j(\beta_1 x-\omega_1 t)} \\
&= \frac{\varepsilon_{1R}}{c^2}[\frac{\partial^2 E_1}{\partial t^2}e^{j(\beta_1 x-\omega_1 t)} + 2\frac{\partial E_1}{\partial t}(-j\omega_1)e^{j(\beta_1 x-\omega_1 t)} + E_1(-j\omega_1)^2 e^{j(\beta_1 x-\omega_1 t)}] \\
&+ \mu_o\sigma_{1e}[\frac{\partial E_1}{\partial t}e^{j(\beta_1 x-\omega_1 t)} + E_1(-j\omega_1)e^{j(\beta_1 x-\omega_1 t)}] \\
&+ \mu_o\sigma_2[(\frac{\partial E_2}{\partial t}E_1^* + A_2\frac{\partial E_1^*}{\partial t})e^{j(\beta_2 x-\omega_2 t-k_1 x+\omega_1 t)} + E_2E_1^* j(-\omega_1)e^{j(\beta_2 x-\omega_2 t-k_1 x+\omega_1 t)}],
\end{aligned} \tag{C.2b}$$

$$\begin{aligned}
&\frac{\partial^2 E_2}{\partial x^2}e^{j(\beta_2 x-\omega_2 t)} + 2\frac{\partial E_2}{\partial x}(j\beta_2)e^{j(\beta_2 x-\omega_2 t)} + E_2(j\beta_2)^2 e^{j(\beta_2 x-\omega_2 t)} \\
&= \frac{\varepsilon_{2R}}{c^2}[\frac{\partial^2 E_2}{\partial t^2}e^{j(\beta_2 x-\omega_2 t)} + 2\frac{\partial E_2}{\partial t}(-j\omega_2)e^{j(\beta_2 x-\omega_2 t)} + E_2(-j\omega_2)^2 e^{j(\beta_2 x-\omega_2 t)}] \\
&+ \mu_o\sigma_2 2E_1[\frac{\partial E_1}{\partial t}e^{j(2\beta_1 x-2\omega_1 t)} + E_1(-j\omega_1)e^{j(2\beta_1 x-2\omega_1 t)}].
\end{aligned} \tag{C.2c}$$

With the slow envelope approximation and follow the approach given in Appendix B, Eqs. C.2 can be simplified as

$$\frac{\partial E_0}{\partial x} = [\alpha_0^2 E_0 - \mu_o \sigma_{0e} \frac{\partial E_0}{\partial t} - \mu_o \sigma_2 (\frac{\partial |E_1|^2}{\partial t} + \frac{\partial |E_2|^2}{\partial t}) e^{\alpha_0 x}] \frac{1}{2\alpha_0}, \tag{C.3a}$$

$$\frac{\partial E_1}{\partial x} = [(\frac{-2j\omega_1 \varepsilon_{1R}}{c^2} + \mu_o \sigma_{1e}) \frac{\partial E_1}{\partial t} - j\omega_1 \mu_o \sigma_{1e} E_1,$$

$$+ \mu_o \sigma_2 (\frac{\partial E_2}{\partial t} E_1^* + E_2 \frac{\partial E_1^*}{\partial t} - j\omega_1 E_2 E_1^*) e^{j(\beta_2 - 2\beta_1)x}] \frac{1}{2j\beta_1}, \tag{C.3b}$$

$$\frac{\partial E_2}{\partial x} = \{(\frac{-2j\omega_2 \varepsilon_{2R}}{c^2} + \mu_o \sigma_{2e}) \frac{\partial E_2}{\partial t} - j\omega_2 \mu_o \sigma_{2e} E_2,$$

$$+ 2\mu_o \sigma_2 E_1 [\frac{\partial E_1}{\partial t} e^{j(2\beta_1 x - \beta_2 x)} + E_1 (-j\omega_1) e^{j(2\beta_1 - \beta_2)x}]\} \frac{1}{2j\beta_2}. \tag{C.3c}$$

The transformation to the retarded frame [68] of reference is defined as

$$E_i(t,x) = E_i(t - \frac{x}{v_i}, x). \tag{C.4}$$

Using the above transformation and the coordinate transformation in Eq. B.10, and replacing T by t in the final expressions, we obtain

$$\frac{\partial A_0}{\partial x} - \frac{1}{v_0} \frac{\partial A_0}{\partial t} =$$

$$[\alpha_0^2 A_0 - \mu_o \sigma_{0e} \frac{\partial A_0}{\partial t} - \mu_o \sigma_2 (\frac{\partial |A_1|^2}{\partial t} \frac{\sqrt{n_0}}{\sqrt{n_1 n_1}} + \frac{\partial |A_2|^2}{\partial t} \frac{\sqrt{n_0}}{\sqrt{n_2 n_2}}) e^{\alpha_0 x}] \frac{1}{2\alpha_0}, \tag{C.5a}$$

$$\frac{\partial A_1}{\partial x} - \frac{1}{v_1} \frac{\partial A_1}{\partial t} = [(\frac{-2j\omega_1 \varepsilon_{1R}}{c^2} + \mu_o \sigma_{1e}) \frac{\partial A_1}{\partial t} - j\omega_1 \mu_o \sigma_{1e} A_1$$

$$+ \mu_o \sigma_2 (\frac{\partial A_2}{\partial t} A_1^* + A_2 \frac{\partial A_1^*}{\partial t} - j\omega_1 A_2 A_1^*) \frac{\sqrt{n_1}}{\sqrt{n_2 n_1}} e^{j(\beta_2 - 2\beta_1)x}] \frac{1}{2j\beta_1}, \tag{C.5b}$$

$$\frac{\partial A_2}{\partial x} - \frac{1}{v_2} \frac{\partial A_2}{\partial t} = [(\frac{-2j\omega_2 \varepsilon_{2R}}{c^2} + \mu_o \sigma_{2e}) \frac{\partial A_2}{\partial t} - j\omega_2 \mu_o \sigma_{2e} A_2,$$

$$+ 2\mu_o \sigma_2 A_1 (\frac{\partial A_1}{\partial t} - j\omega_1 A_1) \frac{\sqrt{n_2}}{\sqrt{n_1 n_1}} e^{j(2\beta_1 - \beta_2)x}] \frac{1}{2j\beta_2}. \tag{C.5c}$$

In the above, the asterisk represents complex conjugate. The above equations describe the propagation of the envelopes of the three pulses in the MIM

waveguide. They can be solved using the finite difference time domain method (FDTD) [67].

For the case with two input pulses (E_2 and E_3), the major field components in the mixing process are identified as the two fundamentals (E_2 and E_3), two second harmonics (E_4 and E_6), frequency difference (E_1), frequency sum (E_5) and the rectified (E_0). They are represented by

$$\tilde{\tilde{E}}_0(x,t) = E_0(x,t)e^{-\alpha_0 x}, \tag{C.6a}$$

$$\tilde{\tilde{E}}_i(x,t) = \frac{1}{2}\{E_i(x,t)e^{j(\beta_i x-\omega_i t)} + c.c.\}, \tag{C.6b}$$

where i=1:6. Substituting the Eqs. C.6 into Eq. B.2, forcing the frequency matching condition and using similar approach as for the second harmonic generation case, we obtain the propagating equations for the different field components in the MIM waveguide. For simplification, we only consider dominated frequency terms for the following equations.

$$\frac{\partial A_0}{\partial x} - \frac{1}{v_0}\frac{\partial A_0}{\partial t} = [\alpha_0^2 A_0 - \mu_o\sigma_{0e}\frac{\partial A_0}{\partial t}$$
$$- \mu_o\sigma_2(\frac{\partial|A_1|^2}{\partial t}\frac{\sqrt{n_0}}{\sqrt{n_1 n_1}} + \frac{\partial|A_2|^2}{\partial t}\frac{\sqrt{n_0}}{\sqrt{n_2 n_2}} + \frac{\partial|A_3|^2}{\partial t}\frac{\sqrt{n_0}}{\sqrt{n_3 n_3}} +$$
$$\frac{\partial|A_4|^2}{\partial t}\frac{\sqrt{n_0}}{\sqrt{n_4 n_4}} + \frac{\partial|A_5|^2}{\partial t}\frac{\sqrt{n_0}}{\sqrt{n_5 n_5}} + \frac{\partial|A_6|^2}{\partial t}\frac{\sqrt{n_0}}{\sqrt{n_6 n_6}})e^{\alpha_0 x}]\frac{1}{2\alpha_0}, \tag{C.7a}$$

$$\frac{\partial A_1}{\partial x} - \frac{1}{v_1}\frac{\partial A_1}{\partial t} = [(\frac{-2j\omega_1\varepsilon_{1R}}{c^2} + \mu_o\sigma_{1e})\frac{\partial A_1}{\partial t} - j\omega_1\mu_o\sigma_{1e}A_1$$
$$+ \mu_o\sigma_2(\frac{\partial A_2^*}{\partial t}A_3 + A_2^*\frac{\partial A_3}{\partial t} - j\omega_1 A_2^* A_3)\frac{\sqrt{n_1}}{\sqrt{n_2 n_3}}e^{j(-\beta_2+\beta_3-\beta_1)x}]\frac{1}{2j\beta_1}, \tag{C.7b}$$

$$\frac{\partial A_2}{\partial x} - \frac{1}{v_2}\frac{\partial A_2}{\partial t} = [(\frac{-2j\omega_2\varepsilon_{2R}}{c^2} + \mu_o\sigma_{2e})\frac{\partial A_2}{\partial t} - j\omega_2\mu_o\sigma_{2e}A_2$$
$$+ \mu_o\sigma_2(\frac{\partial A_4}{\partial t}A_2^* + A_4\frac{\partial A_2^*}{\partial t} - j\omega_2 A_2^* A_4)\frac{\sqrt{n_2}}{\sqrt{n_2 n_4}}e^{j(\beta_4-2\beta_2)x}]\frac{1}{2j\beta_2}, \tag{C.7c}$$

$$\frac{\partial A_3}{\partial x}-\frac{1}{v_3}\frac{\partial A_3}{\partial t}=[(\frac{-2j\omega_3\varepsilon_{3R}}{c^2}+\mu_o\sigma_{3e})\frac{\partial A_3}{\partial t}-j\omega_3\mu_o\sigma_{3e}A_3$$

$$+\mu_o\sigma_2(\frac{\partial A_6}{\partial t}A_3^*+A_6\frac{\partial A_3^*}{\partial t}-j\omega_3A_3^*A_6)\frac{\sqrt{n_3}}{\sqrt{n_3n_6}}e^{j(\beta_6-2\beta_3)x}]\frac{1}{2j\beta_3}, \quad \text{(C.7d)}$$

$$\frac{\partial A_4}{\partial x}-\frac{1}{v_4}\frac{\partial A_4}{\partial t}=[(\frac{-2j\omega_4\varepsilon_{4R}}{c^2}+\mu_o\sigma_{4e})\frac{\partial A_4}{\partial t}-j\omega_4\mu_o\sigma_{4e}A_4$$

$$+\mu_o\sigma_2(2A_2\frac{\partial A_2}{\partial t}-j\omega_4A_2{}^2)\frac{\sqrt{n_4}}{\sqrt{n_2n_2}}e^{j(2\beta_2-\beta_4)x}]\frac{1}{2j\beta_4}, \quad \text{(C.7e)}$$

$$\frac{\partial A_5}{\partial x}-\frac{1}{v_5}\frac{\partial A_5}{\partial t}=[(\frac{-2j\omega_5\varepsilon_{5R}}{c^2}+\mu_o\sigma_{5e})\frac{\partial A_5}{\partial t}-j\omega_5\mu_o\sigma_{5e}A_5$$

$$+\mu_o\sigma_2(\frac{\partial A_2}{\partial t}A_3+A_2\frac{\partial A_3}{\partial t}-j\omega_5A_2A_3)\frac{\sqrt{n_5}}{\sqrt{n_2n_3}}e^{j(\beta_2+\beta_3-\beta_5)x}]\frac{1}{2j\beta_5}, \quad \text{(C.7f)}$$

$$\frac{\partial A_6}{\partial x}-\frac{1}{v_6}\frac{\partial A_6}{\partial t}=[(\frac{-2j\omega_6\varepsilon_{6R}}{c^2}+\mu_o\sigma_{6e})\frac{\partial A_6}{\partial t}-j\omega_6\mu_o\sigma_{6e}A_6$$

$$+\mu_o\sigma_2(2A_3\frac{\partial A_3}{\partial t}-j\omega_6A_3{}^2)\frac{\sqrt{n_6}}{\sqrt{n_3n_3}}e^{j(2\beta_3-\beta_6)x}]\frac{1}{2j\beta_6}. \quad \text{(C.7g)}$$

The above equations can be solved using the finite-difference time-domain method for both the spatial and time dependences.

Appendix D

Absorbing Boundary Condition for FDTD Simulation

With limited computational resource such as time and computer memory, the FDTD simulation must be restricted to a reasonable numerical size. As a result, the simulation domain needs to be truncated and terminated with a boundary condition which absorbs all outgoing electromagnetic waves as if the structure extends beyond the FDTD simulation domain. In addition, negligible reflection is required for the boundary condition.

Different types of absorbing boundary conditions have been proposed for the FDTD algorithm. Among them, the perfect match layer (PML) technique is the most popular [69]. It achieves low reflection based on impedance matching between the simulation domain and the artificial lossy PML material which gradually attenuate any outgoing fields. To achieve impedance matching and introduce loss, the properties of PML material must satisfy the following governing equation [69]

$$\frac{\sigma(\omega)}{\varepsilon_o \varepsilon_r(\omega)} = \frac{\sigma^*(\omega)}{\mu_o \mu_r(\omega)}, \tag{D.1}$$

where σ is the artificial electric conductivity, σ^* is the artificial magnetic conductivity, ε_r is the relative complex permittivity, and μ_r is the relative complex permeability. Furthermore, the MIM waveguide structure must be extended into the PML layer to maintain its structure continuity for field propagation. With the superimposition of the waveguide structure and the PML layer, the PML materials are assigned the same permittivity and permeability as the MIM waveguide structure. The ADE approach is also used for taking care of the PML materials. Since the same attenuation rate is needed for both Al and Al_2O_3 layers, Eq. D.1 in the Al_2O_3 layer can be used to determine the attenuation rate.

$$\frac{\sigma}{\varepsilon_o \varepsilon_r}(Al_2O_3) = \frac{\sigma^*}{\mu_o \mu_r}(Al_2O_3) = CC, \tag{D.2a}$$

while for the Al layer

$$\frac{\sigma(\omega)}{\varepsilon_o \varepsilon_r(\omega)} = \frac{\sigma^*}{\mu_o \mu_r} = CC\,, \tag{D.2b}$$

where CC is a constant used to determine the attenuation rate. The optimum value of the maximum conductivity is calculated as [69]

$$\sigma_{\max} = (n+1)\log(\frac{1}{R})\varepsilon_o \varepsilon_r \frac{C}{2L}\cdot \tag{D.3}$$

In the above, *n* is the order number which determines the conductivity profile inside the PML layer, *R* is the expected level of reflection, *C* is the speed of light in vacuum and *L* is the total thickness of the PML layer. Generally, the artificial conductivity in the PML layer is gradually turned on from the PML-FDTD boundary. As the thickness of the PML layer increases, the conductivity increases. Maximum conductivity appears at the maximum PML thickness. The conductivity profile in the PML layer is governed by

$$\sigma(d) = \sigma_{\max}(\frac{d}{L})^n\cdot \tag{D.4}$$

From Eq. D.2a and Eq. D.3, we are able to calculate σ_{max} and σ_{max}^*. We also denote the following constitutive relation for the loss introduced by the artificial conductivity in the PML material

$$P = \sigma(\omega)E = CC\varepsilon_o \varepsilon_r(\omega)E\,, \tag{D.5}$$

and in the ADE form

$$P = CC(\varepsilon_\infty \varepsilon_o E - \frac{\omega_p^2}{\omega(\omega - i\Gamma_e)}\varepsilon_o E) = CC(\varepsilon_\infty \varepsilon_o E + J)\,, \tag{D.6}$$

where *J* has the same expression as the polarization current of Al. The TM mode fields update equations in the PML region will be (based on the TM mode equations given in Chapter 4)

$$\mu_o \frac{\partial H_y}{\partial t} = -(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} + \sigma^* H_y)\,, \tag{D.7a}$$

$$\varepsilon_r \varepsilon_o \frac{\partial E_x}{\partial t} = -(\frac{\partial H_y}{\partial z} + J_x + P_x)\,, \tag{D.7b}$$

$$\varepsilon_r \varepsilon_o \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} - J_z - P_z\,, \tag{D.7c}$$

$$P_x = CC(\varepsilon_\infty \varepsilon_o E_x + J_x), \tag{D.7d}$$

$$P_z = CC(\varepsilon_{z\infty} \varepsilon_o E_z + J_z), \tag{D.7e}$$

$$\varepsilon_o \omega_p^2 E_z = \frac{\partial J_z}{\partial t} + \Gamma_e J_z, \tag{D.7f}$$

$$\varepsilon_o \omega_p^2 E_x = \frac{\partial J_x}{\partial t} + \Gamma_e J_x. \tag{D.7g}$$

Following this procedure, we are able to implement effective absorbing boundary condition for the MIM nanoplasmonic waveguide.

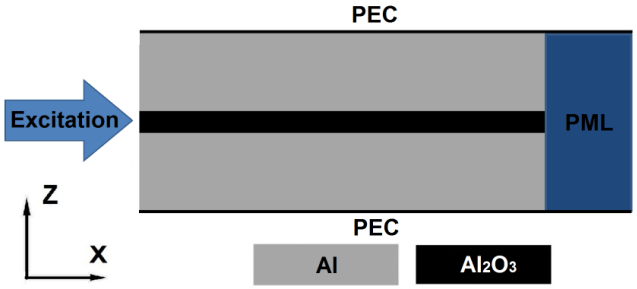


Figure D.1 Schematic of the 2D FDTD simulation domain for the MIM waveguide.

In the simulation, the excitation source is injected from the left end of the waveguide and propagates towards the right end as illustrated in Fig. D.1. Since the field propagates in parallel with the sidewalls, we only need to terminate the right end side of the waveguide with PML materials. With thick metal layers, the two sidewalls can be terminated with perfect electric conductor (PEC) boundary conditions.