Design, Modeling And Simulation Of Nanoscale Optoelectronic Devices:

Semiconductor Nano-Lasers And Plasmonic Waveguides

By

Debin Li

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved February 2012 by the
Graduate Supervisory Committee:

Cun-Zheng Ning, Chair
Yong-Hang Zhang
Constantine A. Balanis
Hongbin Yu

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

This thesis summarizes the research work carried out on design, modeling and simulation of semiconductor nanophotonic devices. The research includes design of nanowire (NW) lasers, modeling of active plasmonic waveguides, design of plasmonic nano-lasers, and design of all-semiconductor plasmonic systems. For the NW part, a comparative study of electrical injection in the longitudinal p-i-n and coaxial p-n core-shell NWs was performed. It is found that high density carriers can be efficiently injected into and confined in the core-shell structure. The required bias voltage and doping concentrations in the core-shell structure are smaller than those in the longitudinal p-i-n structure. A new device structure with core-shell configuration at the p and n contact regions for electrically driven single NW laser was proposed. Through a comprehensive design trade-off between threshold gain and threshold voltage, room temperature lasing has been proved in the laser with low threshold current and large output efficiency. For the plasmonic part, the propagation of surface plasmon polariton (SPP) in a metal-semiconductor-metal structure where semiconductor is highly excited to have an optical gain was investigated. It is shown that near the resonance the SPP mode experiences an unexpected giant modal gain that is 1000 times of the material gain in the semiconductor and the corresponding confinement factor is as high as 10^5. The physical origin of the giant modal gain is

the slowing down of the average energy propagation in the structure. Secondly, SPP modes lasing in a metal-insulator-semiconductor multi-layer structure was investigated. It is shown that the lasing threshold can be reduced by structural optimization. A specific design example was optimized using AlGaAs/GaAs/AlGaAs single quantum well sandwiched between silver layers. This cavity has a physical volume of $1.5\times10^{-4}\ \lambda_0^3$ which is the smallest nanolaser reported so far. Finally, the all-semiconductor based plasmonics was studied. It is found that InAs is superior to other common semiconductors for plasmonic application in mid-infrared range. A plasmonic system made of InAs, GaSb and AlSb layers, consisting of a plasmonic source, waveguide and detector was proposed. This on-chip integrated system is realizable in a single epitaxial growth process.

TABLE OF CONTENTS

Page

LIST OF TABLES……………………………………………………………vii

LIST OF FIGURES…………………………………………………………..viii

CHAPTER

1. INTRODUCTION 1

1.1. Electrically driven nanowire lasers 2

1.2. Active surface plasmon polariton waveguide 4

1.3. Plasmonic nano-lasers 6

1.4. All-semiconductor plasmonics 8

2. ELECTRICAL INJECTION IN NANOWIRES 11

2.1. Longitudinal and core-shell device structures 11

2.2. Comparative study of electrical injection efficiency 15

2.2.1. Metal contacts 15

2.2.2. Injection efficiency verses bias voltage 17

2.2.3. Injection efficiency verses doping density.................................... 22

2.2.4. Injection efficiency verses band edge alignment 26

2.3. Conclusion 31

CHAPTER Page

3. ELECTRICALLY DRIVEN SINGLE NANOWIRE LASER 33

3.1. A proposal of device structure 33

3.2. Electrical injection efficiency 35

3.3. Mode analysis and laser design... 38

3.4. Laser performance.. 45

3.5. Conclusion 46

4. METAL-SEMICONDUCTOR SURFACE PLASMONIC WAVEGUIDES .. 48

4.1. Fundamental physics of surface plasmon polaritons 48

4.2. Planar surface plasmon waveguides with optical gain 54

4.2.1. Modal gain in active plasmonic waveguide................................. 54

4.2.2. Explanation of giant modal gain.. 60

4.2.3. Feasibility of giant modal gain 64

4.3. Confinement factor in the planar surface plasmon waveguides 68

4.3.1. Derivation of confinement factors in a plasmonic waveguide...... 70

4.3.2. Enhancement of confinement factors... 74

4.3.3. Gain dependence of confinement factors..................................... 77

CHAPTER Page

4.4. Conclusion 79

5. METAL-SEMICONDUCTOR-METAL NANO-LASER 81

5.1. Metal-semiconductor-metal nano-cavity 81

5.2. Parametric dependence of loss and gain 84

5.2.1. Mode analysis 86

5.2.2. Dependence on middle layer thickness 88

5.2.3. Dependence on in-plane sizes 90

5.2.4. Dependence on metal layer thickness 95

5.2.5. Effects of insulating layer 98

5.3. Toward the smallest semiconductor laser 101

5.4. Conclusion 106

6. ALL-SEMICONDUCTOR PLASMONICS.. 108

6.1. Interband and intraband transition in a semiconductor 108

6.2. Plasmonic features of InAs heterostructures... 110

6.3. All-semiconductor active plasmonic system 120

6.4. Conclusion 126

CHAPTER Page

7. SUMMARY 127

ACKNOWLEDGEMENT 135

REFERENCES 136

APPENDIX

A. DERIVATION OF CONFINEMENT FACTOR 144

B. DETERMINATION OF LASING THRESHOLD 146

LIST OF TABLES

Table Page

1. Bandgap Narrowing Coefficients A, B and C……………….................13

2. Material Parameters at 300 K………………………………………………15

3. Material Parameters at Room Temperature…………………………………35

4. Effective Widths in InAs/GaSb Structure at 3 μm..…………………………120

LIST OF FIGURES

Figure Page

1. Device structures of simulation…………………………...……………...15

2. Energy band profile of a p-i-n structure with metal contacts………………17

3. Carrier density in different NWs as a function of anode voltage…………...19

4. Band lineup of GaAs/InAs/GaAs p-i-n structure along the z-axis and InAs/GaAs core-shell structure along the diameter…………………………19

5. Schematics and carrier profiles of GaAs/InAs p-n and GaAs/InAs/GaAs p-n-p heterojunctions……………………………………………………………21

6. Electron and hole density profiles of the InAs/GaAs core-shell and the GaAs/InAs/GaAs p-i-n structures…………………………………………...22

7. Carrier density at the center of NWs as a function of doping density………25

8. Normalized charge density and valence band profile along the diameter of InAs/GaAs core-shell structure with different doping densities…………….26

9. Electron and hole density at the center of core region of InP/GaAs and GaAs/GaAs core-shell structure as a function of bias………………………28

10. Valence band lineups of InP/GaAs and GaAs/GaAs core-shell structures at 3 V bias, with the corresponding hole QFLs………………………………….28

11. Electron and hole density at the center of core region of InAs/GaAs and InP/GaAs core-shell structure as a function of core region radius …………30

Figure Page

12. Electron and hole density at the center of core region of InAs/GaAs and InP/GaAs core-shell structure as a function of shell thickness…..…………..31

13. A proposed device structure of an electrically-driven nanowire laser with core-shell structure at p and n contact region……………………………….34

14. Electron and hole concentration along z axis for different overlapping length between the contact regions and the NW…………………………………...37

15. Electron and hole concentration along z axis for different NW length……..38

16. Normalized $|E|^2$ profile of a HE_{11p} mode along the radius of a CdS and a CdSe NW across ZnO, AAO and GaN region……………………………………40

17. Spectrum of HE_{11p} modes in CdS and CdSe NWs………………………….41

18. Carrier concentration along the length of CdS and CdSe NWs for different biases……………………………………………………………………...43

19. Threshold voltage and threshold gain in CdS and CdSe with various overlapping length and NW length…………………………………………..44

20. L-I curve of the CdS and the CdSe laser……………………………………46

21. Real and imaginary parts of the dielectric function of silver………………..51

22. Geometry for SPP propagation at a single interface between a metal and a semiconductor…………………………………………………………….52

Figure Page

23. Geometry of a three-layer structure consisting of a thin mid-layer semiconductor sandwiched between two infinite metal layers……………...54

24. The imaginary part of k_z as a function of photon energy……………………56

25. Real and imaginary parts of k_z versus photon energy……………………….59

26. Group velocity, modal gain and energy velocity vs. photon energy………...62

27. Distribution of two electric field components E_x and E_z in the MSM waveguide at 3 different photon energies…………………………………...63

28. Average energy velocity and modal gain versus photon energy of an anti-symmetric SPP mode…………………………………………………...64

29. Bandgap energy versus refractive index of some II-VI materials in both cubic and hexagonal structures at 77 K and 300 K………………………………..67

30. Real and imaginary part of k_z for the fundamental TM mode of silver/$Zn_{0.8}Cd_{0.2}Se$/silver waveguide………………………………………..68

31. Confinement factors in semiconductor and metal layer as a function of photon energy for two different gain levels in a MSM waveguide…………………76

32. Modal gain from semiconductor, modal loss from metal and net modal gain as a function of photon energy………………………………………………76

33. Confinement factor as a function of imaginary part of dielectric constant at different photon energies in the MSM waveguide…………………………..79

Figure Page

34. Schematic of and the intensity spectrum of a MSM cavity…………………82

35. Linewidth and peak height of the intensity spectrum in a MSM cavity…….84

36. Intensity spectrum of MSM cavities with various W and L………………...87

37. Real part of the propagation wavevector of some high order modes in the MSM cavities with various W and L………………………………………..87

38. Intensity spectrum of MSM cavities with various h_s………………………..88

39. Threshold gain of consecutive TM_{01m}, TM_{02m} and TM_{03m} modes in the cavity with different core layer thickness……………………………………………..90

40. Threshold gain of consecutive TM_{01m} modes for different in-plane sizes…..91

41. The minimum threshold gain of the cavity with different core layer thicknesses and in-plane sizes…………………………………………………93

42. Intensity spectra and threshold gain of each mode in the cubic core cavities with various size of core region……………………………………………..95

43. Energy profiles in the x-z plane, the percentage of the energy leaked, the minimum threshold gain, and the position of the mode as a function of metal layer thickness for W=L=50 nm……………………………………………97

44. Energy profiles in the x-z plane, the percentage of the energy leaked, the minimum threshold gain, and the position of the mode as a function of metal layer thickness for W=L=100 nm…………………………………………..98

Figure Page

45. Minimum threshold gain as a function of insulating layer and core layer thickness……………………………………………………………………101

46. Schematic, mode spectrum, near field and far field characteristics of an optimized MISIM nanolaser………………………………………………104

47. Schematic, mode spectrum, near field and far field characteristics of a smaller MISIM nanolaser…………………………………………….............105

48. Plasmonic features of InAs………………………………………………...112

49. Electron density as a function of bias in a GaSb/InAs/GaSb structure…….113

50. Electron density profile in a GaSb/InAs/GaSb structure with different InAs doping densities and anode biases…………………………………………114

51. Real and imaginary part of propagation wavevector k_z vs. wavelength of a SPP mode in an InAs/GaSb bi-layer structure……………………………..116

52. Quality factors vs. wavelength of a SPP mode in an InAs/GaSb structure, an Ag/GaSb structure, and an Au/GaSb structure……………………………..117

53. Normalized energy density and power flux profile across the interface of an InAs/GaSb structure at 3 and 5 μm………………………………………...119

54. All-semiconductor active plasmonic system………………………………122

55. Electron and hole density profiles along x axis……………………………124

56. Modal pattern of a SPP mode in the plasmonic system……………………126

1. Introduction

Nanoscale-optoelectronics is an area of active research dealing with the interaction of light with matter at submicron dimensions. Nanoscale optoelectronic devices have the potential advantages of miniaturizing integrated photonic circuits, expanding the bandwidth, increasing the switching speed, and providing low loss photonic components. This technology has the potential applications in telecommunication, computing, detection and imaging in biology and medicine, laser diodes, and solar cells. Researchers are interested in various types of nano-optoelectronic structures - e.g. nanoparticles, nanowires (NWs), nanotubes (NTs), nanobelts and nanostrips based on various compound semiconductor materials and other non-semiconductor materials. Extremely compact devices have been realized including lasers, optical waveguides, LEDs, and sensors. Nano-optoelectronics adds a new dimension to nanoscale science and technology, provides challenges for fundamental research, and creates opportunities for new technology.

As optoelectronic devices shrink to the nanoscale, many new challenges arise. Two types of structures, NWs and plasmonic structures, have attracted a huge amount of attention in the research of nanophotonics. My research is focused on exploring new features and phenomena based on these two types of

structures. This thesis then consists of two parts: the first is on electrical injection properties in NWs and electrically driven NW lasers, and the second is about nano-scale surface plasmon polariton (SPP) waveguides, cavities, and lasers, and all-semiconductor nano-scale plasmonic systems.

1.1. Electrically driven nanowire lasers

As a typical example of nanoscale structures, semiconductor NWs have attracted a great deal of attention for their potential applications in nano-optoelectronics, especially NW lasers over the past decade. Very often, the Coulomb interaction potential from charge centers may not be totally screened by the surrounding electrons on the device scales of interest. Therefore, some electrostatic properties like depletion widths or electric displacement vectors may change a lot compared to the bulk situation. In fact, the depletion width for NTs varies exponentially with inverse doping and there is a very long-range (logarithmic) tail in the charge distribution, extending over the entire tube [1]. However, quantum confinement due to device size down to several nanometers is not considered in this research because, for nano-optoelectronic applications, the device dimensions are usually tens or hundreds of nanometers, which implies the effect of size quantization can be reasonably ignored.

For electrically driven optoelectronic devices, metals that usually provide Ohmic contact in bulk devices are used as electrodes. At the size scales mentioned

above, the metal contacts have some novel properties. It was reported that contact resistivity in NWs increases as the radius decrease and is more significant for three-dimensional (3D) metal contacts than for one-dimensional (1D) metal contacts. The underlying cause for this size effect is identified as the strong fringing field effects [2]. For quasi-one-dimensional (Q1D) structures such as NTs and NWs, side contacts with metals only lead to weak band realignment, in contrast to bulk metal-semiconductor contacts. The Schottky barriers are much reduced compared with the bulk limit, and should facilitate the formation of good contacts [3].

Many electrically driven nano-scale optoelectronic devices were fabricated, such as cross-bar [4], axial [5] or radial [6] geometries, but only LED function has been demonstrated in most of these structures [5,6]. Even though lasing action has been demonstrated in NWs of various semiconductors using optical pumping [4], [7-12], electrical injection lasing is still extremely difficult to achieve, with only a few of demonstrations so far [4, 13-17]. The first electrically driven lasing was observed in a single CdS NW on a p-type Si substrate at 8 K [4]. Thereafter, random lasing by electrical pumping was achieved in various NWs [13-16]. Most recently, electrically pumped Fabry-Perot type waveguide lasing from laser diodes that consist of p-type ZnO NWs and n-type ZnO thin films was demonstrated at room temperature [17]. One of the fundamental reasons for the

unsatisfactory state of affairs in electrical injection NW lasers is the lack of proper understanding of electrical behavior in these structures. It is clear from published studies [2,18] that the electrical behavior of NW lasers is sufficiently different from bulk structures where electrical injection is well understood. For example, it is not clear which NW configuration (longitudinal p-i-n vs. core-shell) is the most efficient, how to choose a combination of materials (for various regions and metals), and what structures are most feasible for material growth and laser device fabrication. All these questions will be answered in this thesis.

1.2. Active surface plasmon polariton waveguide

Usually, nano-scale optoelectronic devices have sizes smaller than the wavelength of light in vacuum, or even in the medium. Here, strong wave-guiding and large variations of the field on small scales are important. Simple size reduction based on the conventional approach to optoelectronic devices leads to poor light confinement in subwavelength waveguides and devices. Two approaches are being studied to remedy the situation: One is the use of semiconductor NWs in air, providing a large contrast in refractive indices and allowing much smaller devices to be made. Another attractive approach is to use SPPs occurring at metal/dielectric interfaces. Electromagnetic waves traveling along the structure will become well-localized at frequencies close to the SPP

resonance frequency, and the light thus will get confined and guided. Structures with various metal and dielectric material configurations that can sustain SPPs are widely investigated, especially for planar metal-dielectric-metal (MDM) structures [19-28]. The dispersion relations and propagation characteristics of different modes in this so-called plasmonic waveguide have been extensively studied.

Semiconductor nanostructures integrated with metallic structures are known to be able to guide or confine optical waves in dimensions much smaller than the wavelength involved or the diffraction limit, and thus are a natural choice for making nanophotonic devices such as nano-waveguides and nanolasers. The loss in metals has been the main roadblock, leading to diminishing propagation length in a waveguide or insurmountable threshold gain for a nanolaser. Thus it is quite natural to consider integrating metals with semiconductor gain media to compensate the metal loss, and even to provide net gain for an active device. More and more attention has been paid to the situation where dielectric materials have an optical gain, such as nano-spheres [29], nano-strips [30], nano-gratings [31,32], metal-dielectric structures [33-38], and MDM waveguides [39-41]. Among various structures that support SPP modes, the metal-semiconductor-metal (MSM) structure is the canonical example and has attracted a great deal of attention for applications in active nanophotonic devices.

Achieving net gain in plasmonic structures is important for active nanophotonic devices. It was shown [42] that a large net modal gain is possible in a semiconductor-metal core-shell structure, however, the frequency range of the positive modal gain is near the cut-off frequency which is well-below the SPP resonance. Since the effective wavelength near the cut-off is very long, the wavelength reduction or compression, a key promise of plasmonic devices, is not possible far below the resonance. Compensation of metal loss for a propagating mode in a MSM structure was studied at wavelengths around 1.5 microns [39], which is far below the SPP resonance. Since the effective wavelength is shortest near the SPP resonance, it is highly desirable to achieve a net gain near the SPP resonance. In that scenario, one would be able to achieve maximum reduction of the wavelength (and device size) and a positive gain simultaneously, making the smallest active devices possible. However, so far no theory or experiment has shown the net optical gain in a MSM waveguide that can significantly over-compensate the metal loss near the SPP resonance. The study of coupling between optical gain materials and metallic structures is still an exciting new frontier.

1.3. Plasmonic nano-lasers

Metallic cavity nanolasers [43-54] with ever shrinking cavity sizes down to

tens of nanometers have attracted a great deal of attention recently due to the progress in nano-fabrication technology and in the general understanding of semiconductor-metal interactions at the nanoscale. In the past couple of years efforts to use metals to form the nano-laser resonator have allowed both the overall size of the laser to be reduced to smaller than the wavelength of light, and the optical mode dimension to be reduced below the diffraction limit. Some of the most promising forms of metal-based nano-lasers have semiconductor gain media as their basis. Semiconductor gain media can be electrically pumped efficiently, which will be of major importance for applications. Metals such as Ag and Au have been widely adopted in the past several years in the design and fabrication of nano-scale semiconductor lasers. With metal cladding layers, a semiconductor laser with a physical cavity volume of 0.019 λ_0^3 has been fabricated [51], and a design for the nanolaser with a cavity volume smaller than 0.01 λ_0^3 has been proposed [53], where λ_0 is the corresponding wavelength of the cavity mode in vacuum. One of the reasons for using metals in nano-scale lasers is to take the advantage of the very small modal volume of SPP modes formed at the interface of metals and dielectric materials. A metallic laser lasing at SPP modes is generally called a plasmonic laser. Plasmonic nano-lasers will likely become technologically significant, particularly when the laser and mode size can be reduced far below the diffraction limit.

However, only a few papers have shown semiconductor nanolasers lasing in SPP modes. These include reports of lasing at low temperature (≤10 K) [45, 48], at room temperature but in a very large device, such as a 100-μm-long distributed feedback (DFB) gap plasmon laser [44] and a plasmonic laser utilizing total internal reflection with a physical volume of 2×10^{4} λ_0^3 [52]. The main limit is the large intrinsic modal loss of SPP modes. Although efforts have been made to compensate the modal loss by incorporating gain media in various structures that support SPP modes [33-41], the material gain required to overcompensate the loss is often large and not achievable in common semiconductors, especially near the SPP resonance [42, 55, 56]. A spaser-based nanolaser was reported with a volume of only 1.74×10^{-4} λ_0^3, but such structure is not compatible with electrical injection. Several questions remain unanswered: Is it practically possible to operate a nanolaser near the SPP resonance at room temperature under electrical injection? Can the required laser threshold be further reduced through realistic consideration and optimization, and if so, to what extent? Ultimatly, what is the smallest practically achievable size for SPP lasers, especially under electrical injection? All these questions will be answered in this thesis.

1.4. All-semiconductor plasmonics

Metal plays an important role in plasmonics and metamaterials, which have

impacted many fields of research and applications, as mentioned above. However, great challenges remain to fully realize many promised potentials, as described in the following. First, common metals such as gold or silver have plasmon resonances in blue or deep ultra-violate wavelength ranges. There are no available metals whose plasmon resonances are in the near or mid-infrared (MIR) wavelength range (say from 1 to 10 microns), which is an extremely important wavelength range for detection and sensing [57-59]. Second, it is highly desirable for many applications to integrate plasmonic structures with gain materials or with other dielectric materials. These applications include active metamaterials or active plasmonic systems containing gain sections. But there is an intrinsic incompatibility of low-quality metal deposition with high-quality epitaxial growth of semiconductors or dielectrics. As a result, many intrinsic plasmonic properties can be masked by the poor metal quality or poor semiconductor-metal interfaces. Third, large metal loss is still a key problem for many plasmonic and metamaterial applications. In addition, the plasmonic resonance frequency (or wavelength) is fixed for a given metal. There is no tenability, though the development of such a capability could benefit many applications. Thus it is important for plasmonic applications to look into other alternatives to metals such as highly doped semiconductors.

Semiconductor plasmonics using doped semiconductors has been studied for

the past several years [57-67]. Recently, plasmonic features and waveguides of doped silicon have been studied at the mid and far-infrared ranges [57-59]. Plasmonic metamaterials with a negative real part of the effective dielectric constant have been realized at the MIR range [61, 62]. Heavily doped n-InAs has been used as the cladding layers in cascade lasers operating at 3.1~12 μm [63-66]. The carrier concentration (doping density) in the semiconductor was reported to be as high as 2.8×10^{22} cm^{-3} [67] which is on the same order of magnitude of the electron density in metals. While metallic properties of highly doped semiconductors have been investigated from the near infrared to far infrared wavelength ranges, their applications as entire systems have not been explored. The goal of this research is to explore the plasmonic properties of a set of doped semiconductors which are suitable for the fabrication of a plasmonic system and to design such an on-chip system that can generate, transmit, guide and detect the SPP signals.

2. Electrical injection in nanowires

In this chapter, a numerical simulation was carried out to achieve a more complete understanding of the electrical injection in different NW configurations, and further more to show a guide for the design of electrically injected heterostructure NW devices [68].

2.1. Longitudinal and core-shell device structures

The materials used in the simulation are GaAs, InAs and InP. To carry out the simulation, a 3D numerical device simulator Silvaco ATLAS [69] was used, and Poisson and drift-diffusion equations were solved using Fermi-Dirac statistics. All the simulations were done at room temperature. Shockley-Reed-Hall (SRH) and Auger recombination models were used. The SRH recombination rate is given in the following expression [70]:

$$R_{SRH} = \frac{pn - n_i^2}{\tau_p \left[n + n_i \exp(\frac{E_t}{kT}) \right] + \tau_n \left[p + n_i \exp(-\frac{E_t}{kT}) \right]} \qquad (2\text{-}1)$$

where E_t is the difference between the trap energy level and the intrinsic Fermi level, n_i is the intrinsic carrier density; τ_n and τ_p are the concentration-dependent electron and hole lifetimes:

$$\tau_n = \frac{\tau_{n0}}{A_n + B_n(\frac{N_t}{N_{SRHn}}) + C_n(\frac{N_t}{N_{SRHn}})^{E_n}} \qquad (2\text{-}2a)$$

$$\tau_p = \frac{\tau_{p0}}{A_p + B_p(\frac{N_t}{N_{SRHp}}) + C_p(\frac{N_t}{N_{SRHp}})^{E_p}} \tag{2-2b}$$

Here, N_t is the local (total) impurity concentration; τ_{n0} and τ_{p0} are electron and hole lifetimes in undoped materials, and are assumed to be 1 ns throughout the work. Other parameters such as $A_{n(p)}$, $B_{n(p)}$, $C_{n(p)}$, $E_{n(p)}$, and so on are the default value in ATLAS. Auger recombination is given by the standard model as [71]:

$$R_{Auger} = Aug_n(pn^2 - nn_i^2) + Aug_p(np^2 - pn_i^2) \tag{2-3}$$

where the default value of $Aug_{n(p)}$ can be found in ATLAS. The radiative recombination is not included in this work, since our purpose is to see how large carrier concentration can be achieved for optical emission. Bandgap narrowing due to heavy doping is also considered [72]. The bandgap change can be written in the form:

$$\Delta E_g = A \times N^{1/3} + B \times N^{1/4} + C \times N^{1/2} \tag{2-4}$$

where N is doping concentration. Parameters A, B and C for materials used in this work are listed in table 2-1.

Table 2-1. Bandgap Narrowing Coefficients *A*, *B* and *C*. [72]

	A (10^{-9})	B (10^{-7})	C (10^{-12})
n-GaAs	16.5	2.39	91.4
p-GaAs	9.83	3.90	3.90
n-InP	17.2	2.62	98.4
p-InP	10.3	4.43	3.38
n-InAs	14.0	1.97	57.9
p-InAs	8.34	2.91	4.53

The importance of metal contacts has been well addressed above for electrical injection purpose in NWs. It is usually used the conventional strategy of heavily doping the semiconductor to obtain Ohmic contacts for electronic devices. But since it is trivial to control doping density in the NWs, Ohmic contact is not very easy to realize by heavy doping. Therefore, other strategies should be used to achieve high efficiency injection in NWs. Firstly, some basic issues concerning about the simulation must be clarified. The electron and hole current densities at the metal-semiconductor interface are given by [73]

$$J_n = \frac{A_n^* T^2}{N_C}(n - n_0)\exp\left(\frac{\Delta\phi_b}{kT}\right) \quad (2\text{-}5a)$$

$$J_p = \frac{A_p^* T^2}{N_V}(p - p_0)\exp\left(\frac{\Delta\phi_b}{kT}\right) \quad (2\text{-}5b)$$

where A_n^* and A_p^* are the Richardson constants for electrons and holes, N_C and N_V are the conduction and valence band effective density of states, n_0 and p_0 are

the equilibrium electron and hole concentrations at the interface, and $\Delta\phi_b$ represents the barrier lowering. Electron tunneling current can be taken into account through typical Tsu-Esaki tunneling theory. However, for my case, the tunneling current is found to be a very small amount of the total current and has little influence on the steady carrier densities. To avoid time-consuming calculation while keeping enough accuracy, tunneling current is thus not included in this work.

The device structures used in the simulation are shown in Fig. 2-1. The length of NWs is fixed at 10 μm, and the metal contacts of anode and cathode are to be determined. The radius of p-i-n NW, R, and the inner radius of core-shell NW R_i are equal to 50 nm, and the outer radius of core-shell NW R_o is 70 nm. The length of i region is 1 μm. Both n and p regions are assumed to be uniformly doped, and have the same doping density unless specified otherwise. To be specific, the entire structure is assumed to be surrounded by air. Further material parameters [74] used in the work are listed in table 2-2.

Table 2-2. Material Parameters at 300 K. [74]

	electron affinity χ_e (eV)	bandgap Eg (eV)	electron mobility μ_e (cm^2/V·s)[a]	hole mobility μ_h (cm^2/V·s)[a]
GaAs	4.07	1.424	8500	400
InP	4.38	1.344	5400	200
InAs	4.9	0.354	40000	250

[a] Values in the undoped material

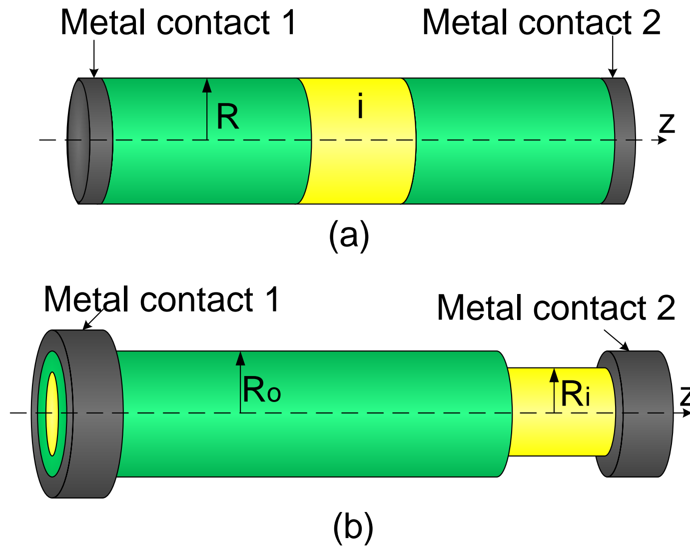

Fig. 2-1 Device structures of simulation. (a) longitudinal p-i-n NW with radius R and (b) coaxial core-shell p-n NW with inner radius R_i and outer radius R_o. Metal 1 and 2 are two contacts.

2.2. Comparative study of electrical injection efficiency

2.2.1. Metal contacts

The first thing is to determine the metal contacts. As mentioned before, it is

not controllable to form Ohmic contacts by heavy doping. To avoid Schottky contacts, metals with proper workfunctions are used. In this case, thermal carriers are injected directly into the semiconductor, which is illustrated in Fig. 2-2. In the figure, a typical band profile of p-i-n structure with two metal contacts is shown, together with the quasi Fermi level (QFL) in each layer. Ideally, if the cathode Fermi level E_{fc} is higher than the conduction band of n region and the anode Fermi level E_{fa} is lower than the valence band of p region, electrons and holes can be injected into the semiconductor mainly through thermionic emission. Since metal Fermi level E_f and metal work function W_m has the relation: $W_m = E_0 - E_f$, where E_0 is the electron energy in vacuum, the problem now is to find metals that can realize thermionic emission for anode and cathode respectively. Take a GaAs/InAs/GaAs p-i-n structure as an example. According to the electron affinity and bandgap of GaAs in table 2-2, the workfunction of anode metal $W_{ma} \geq \chi_e + E_g = 5.494$ eV, and the workfunction of cathode metal $W_{mc} \leq \chi_e = 4.07$ eV. In fact, W_{ma} can be slightly smaller than the calculated minimum value and W_{mc} can be slightly larger than the calculated maximum value without reducing much concentration in NWs. As a result, it is found that titanium (Ti) for cathode and palladium (Pd) for anode are suitable for this case, and will be used in the following work.

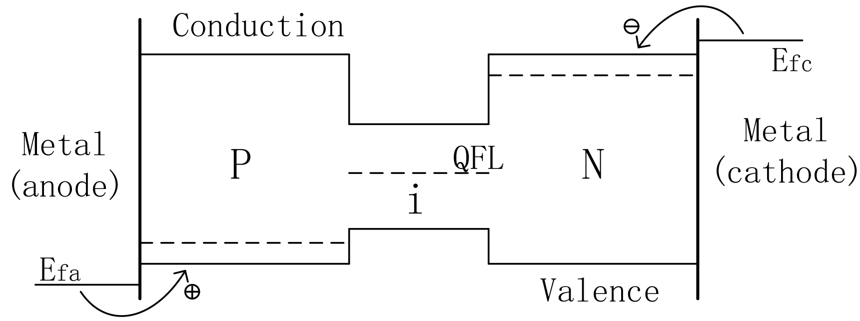

Fig. 2-2. Energy band profile of a typical p-i-n structure with metal contacts. The injections of electrons from cathode and holes from anode are schematically shown as well.

2.2.2. Injection efficiency verses bias voltage

Now consider a longitudinal p-i-n NW with GaAs p and n region and InAs i region and a coaxial core-shell NW with an n-type InAs core and a p-type GaAs shell. Anode bias V_a was swept from 0 to 5 V, and the steady state carrier densities at the center of i region of p-i-n structure and that at the center of core region of core-shell structure are shown in Fig. 2-3. Both electron and hole densities in core-shell structure are nearly two orders of magnitude larger than those in p-i-n structure when bias is larger than 1 V. The results show that core-shell p-n NW has higher injection efficiency than p-i-n NW even without i region. To understand this difference, we note that electrons and holes have to go through the n and p region respectively before they arrive at i region in p-i-n NW. Recombination occurs in n and p region which are usually several times as long as the diffusion length (~1 μm), and thus the carriers injected from electrodes will be

reduced. In the core-shell NW, the electrode is directly connected to the core region, allowing a direct injection of electrons from contact to semiconductor, so that no injected electrons are consumed by recombination. On the other hand, the shell thickness is usually much smaller than the diffusion length, reducing holes recombination in the shell as well. Therefore, the injection efficiency is significantly improved. To display more differences between the two types of the structures, we show in Fig. 2-4 the band lineup along the *z*-axis of p-i-n NW and along the diameter direction of the core-shell NW at 3 V bias. The difference between the electron QFL and the hole QFL represents the external voltage. As we see, part of the bias drops occurs in both the p and n region of the p-i-n structure, leading to a less efficient use of the applied voltage, while the forward voltage in the core region of core-shell structure is much greater. Therefore, a higher bias is required for the longitudinal structure than for the radial structure to obtain the same injected carrier density.

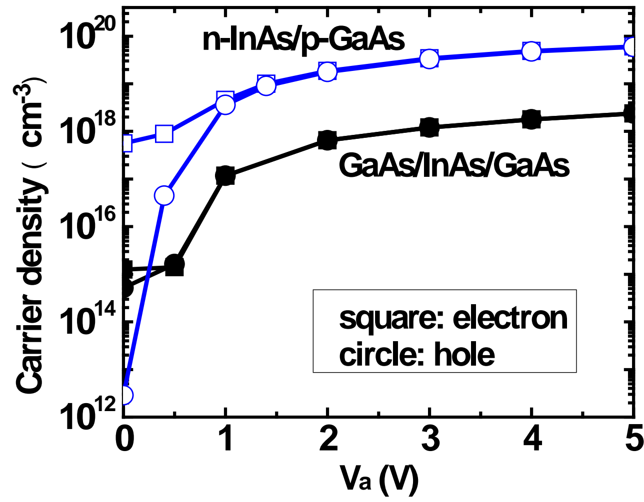

Fig. 2-3. Electrons (squares) and holes (circles) densities at the center of i region of the GaAs/InAs/GaAs p-i-n structure and those at the center of core region of the InAs/GaAs core-shell structure, as a function of anode voltage. Doping density is 1×10^{18} cm^{-3}.

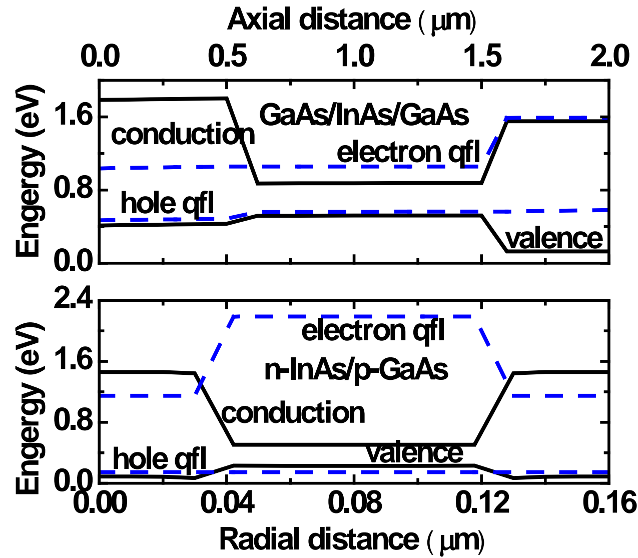

Fig. 2-4. Band lineup of GaAs/InAs/GaAs p-i-n structure along the z-axis (top) and InAs/GaAs core-shell structure along the diameter (bottom): The bias is equal to 3 V with the corresponding quasi-Fermi levels shown by the dashed curves. Doping density is 1×10^{18} cm^{-3}.

To gain a better understanding of the core-shell structure, we consider a simple GaAs/InAs p-n heterojunction shown in Fig. 2-5(a). Because the Fermi Level of metal electrode is much higher than the conduction band edge of the n-InAs region, electrons can be injected efficiently into the n region, and then confined due to the electron barrier at the heterojunction interface, as seen in Fig. 2-5(c). However, the hole density (from Fig. 2-5(c)) drops dramatically near the cathode contact and also at heterojunction interface. In order to obtain a uniform carrier density throughout the n region, a structure with two p-GaAs regions can be used, as shown in Fig. 2-5(b). In this structure, holes are injected from both p regions into n region at the front, and electrons are injected from the end. Therefore, we can see from Fig. 2-5(d) that not only the uniformity is improved but also the injection densities are increased. If such p regions are added around the periphery of the n region, a core-shell structure is formed (assuming a circular cross section). This is why a core-shell structure is a favorable injection configuration.

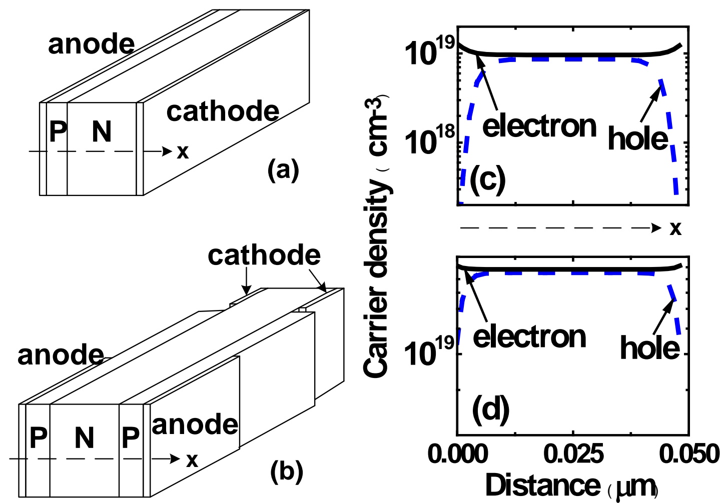

Fig. 2-5. Schematics of GaAs/InAs p-n heterojunction (a) and GaAs/InAs/GaAs p-n-p double heterojunction (b), the thicknesses (in x-axis) of n and p region are 50 nm and 20 nm respectively. Carrier profile along *x* axis in the n region of GaAs/InAs (c) and GaAs/InAs/GaAs heterojunction (d) under 3 V bias. Doping density is 1×10^{18} cm^{-3}.

The uniformity of injected carriers is very important for lasing application. Fig. 2-6 shows the carrier profiles in the cross section of the structures in Fig. 2-4. The carrier concentration is plotted in logarithm scale, and different colors represent different concentrations. Both the electrons and holes are uniformly distributed throughout the i region in the p-i-n structure, while the carrier distribution in core-shell structure shows less uniformity, especially at the core and shell interface where pn junction formed. Beyond that pn region, however, the carrier density is uniform enough in the core region, the difference between the maximum and minimum value of the density is much less than one order of

magnitude.

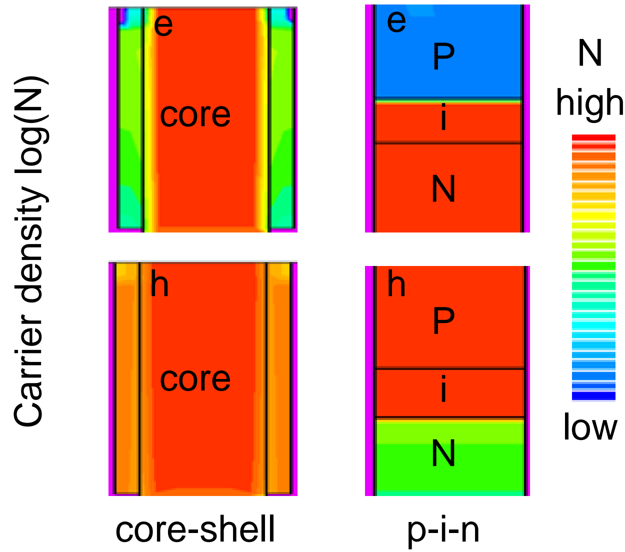

Fig. 2-6. Electrons (top) and holes (bottom) density profiles of the InAs/GaAs core-shell structure and of the GaAs/InAs/GaAs p-i-n structure. The bias is 3 V and doping density is 1×10^{18} cm^{-3}.

2.2.3. Injection efficiency verses doping density

Doping is important for electrically driven optoelectronic devices. Generally, heavy doping is required in traditional laser diodes to obtain high pumping level. However, it is not easy to achieve and to control doping density in nanoscale devices; hence a careful examination of its influence on the injected carrier concentration is important. Fig. 2-7 shows the carrier densities at the center of i region and the core region of the same structures as studied in Fig. 2-4. For the p-i-n structure, carrier density in i region increases monotonically with doping density, as expected in a normal bulk p-i-n diode. However, carrier density in the

core region of the core-shell structure is already at a very high level for a very low doping level and maintains at the same high level over a wide doping range. This is because the fully depleted feature of the core region at low doping level. Fig. 2-8(a) shows the total (or net) charge density normalized to its maximum value along the diameter for different doping concentrations at zero bias. When doping density is less than about 1×10^{17} cm^{-3}, the core region is fully depleted (Fig. 2-8(a)), leaving a relatively lower hole barrier in the core, as shown in Fig. 2-8(b). There is always an electron well in the core due to the n-type doping. Therefore, both electrons and holes can be easily injected into the core region at a high level. As doping is larger than about 1×10^{19} cm^{-3}, however, the hole density falls dramatically. This abnormal behavior of hole density can be explained by the above-mentioned relation between the depletion width and hole potential as well. When doping density is high enough, for example larger than about 1×10^{19} cm^{-3}, the depletion width in the core region becomes very small compared to the core diameter. As a result, the hole barrier is so high that it is difficult for holes to be injected from the shell to the core. Since larger n-type doping leads to deeper electron well, the electron density keeps increasing as core doping increased. The situation is the same if the polarity between core and shell is reversed, i. e., to a p-type core and n-type shell. In this case, the core becomes an electron barrier at sufficiently high doping.

Another doping configuration is the asymmetric doping. For the case that core doping is lower than shell, the carrier concentration is shown in Fig. 2-7 (green symbols) with shell doping fixed at 1×10^{18} cm^{-3}. Both the electron and hole densities remain at a high level and the value is almost the same as that of symmetric doping with doping concentration 1×10^{18} cm^{-3}. This is because of the low hole barrier at low core doping, and carrier density in the core is thus mainly dependent on the shell doping. On the other hand, for low shell doping situation, the carrier densities (red symbols in Fig. 2-7 with core doped at 1×10^{18} cm^{-3}) can keep as high as their symmetrically doped counterpart. That means the injected level in the core is still high enough even though shell doping is not very high. This result is very helpful for core-shell structure, since it is usually not easy to get a heavily doped p-type shell.

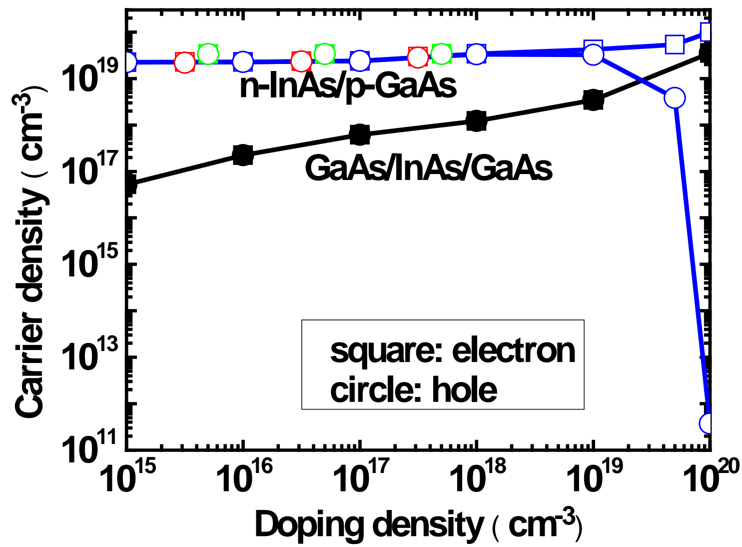

Fig. 2-7. Electrons (squares) and holes (circles) density at the center of i region of GaAs/InAs/GaAs p-i-n structure and at the center of core region of InAs/GaAs core-shell structure at 3 V bias, as a function of doping concentration. The carrier densities of asymmetrically doped InAs/GaAs core-shell strucure for lower core doping (green symbols) and for lower shell doping (red symbols) are also shown under the same bias.

The advantages of the core-shell structure are obvious from Fig. 2-7 and Fig. 2-8: very high levels of injection can be achieved at a very low doping level. At the same time, injected carrier density can be maintained at a constant level over a very large density range, meaning that the exact doping level is not important in a core-shell structure. Therefore, NW with unintentional doping can be used for electrical injection device, removing the difficult requirement of doping control in nano-devices.

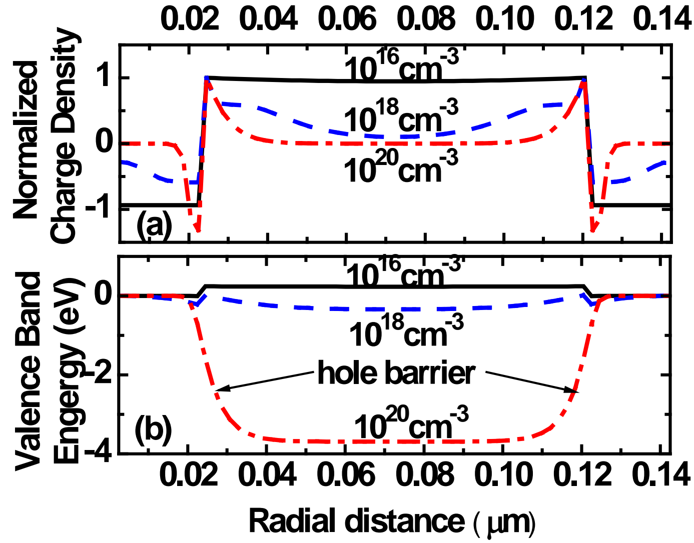

Fig. 2-8. (a) Normalized charge density and (b) valence band profile along the diameter of InAs/GaAs core-shell structure with different doping densities at zero bias.

2.2.4. Injection efficiency verses band edge alignment

Both the p-i-n and the core-shell NWs studied above have type-I band alignment, i. e., the active region (i region for p-i-n and core region for core-shell) has lower conduction and higher valence band edge compared to its surrounding materials, as shown in Fig. 2-4. This band configuration is required for traditional p-i-n diode, since both electrons and holes can be confined within the active region by the potential well. It is, however, not necessary for core-shell NWs, because type-II band alignment is also able to confine sufficient carriers in the core region under proper bias. To show this, we consider an InP/GaAs core-shell NW as an example, where an electron well and hole barrier exist in the core

region under zero bias. The dependence of the carrier density in the core region on the bias is shown in Fig. 2-9, compared with the results for a homo-GaAs n-core/p-shell NW. Both the electron and hole densities in the InP/GaAs core-shell NW are large enough after about 2 V; on the other hand, the hole density in the homo-GaAs core-shell NW is much smaller than its electron density even at 5 V bias. In order to clarify the reason, the valence band lineups of these two structures along the diameter are plotted in Fig. 2-10 at 3 V. It can be seen that there is a shallow hole potential well in the core region of the heterostructure, but not in the core region of the homostructure. As a result, more holes can be confined in the core region of the former structure than of the latter one. Of course for large enough bias, there will also be a hole well in homostructure core region, but the injection efficiency is rather low compared to its heterostructure counterpart. The band edge behavior in the core region is caused by the ring contact around the shell. When applying positive voltage on the shell contact, the direction of the electric field in the core points to the end of the NW because the potential at that end is zero. If considering a cross section of the core region, the valence band in this section is raised due to the electric field perpendicular to this section. This remarkable property of the core-shell structure will lead to a huge design advantage, since high density injection can be achieved regardless of the band lineup for the core-shell material combination.

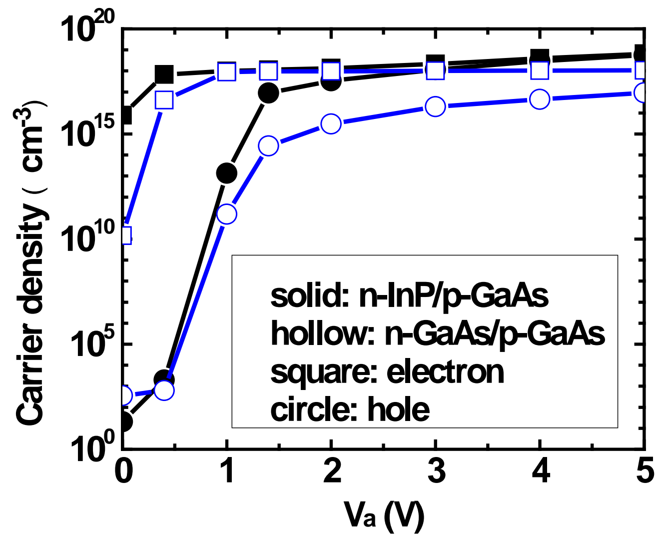

Fig. 2-9. Electrons (squares) and holes (circles) density at the center of core region of InP/GaAs and GaAs/GaAs core-shell structure as a function of bias. Doping density is 1×10^{18} cm^{-3}.

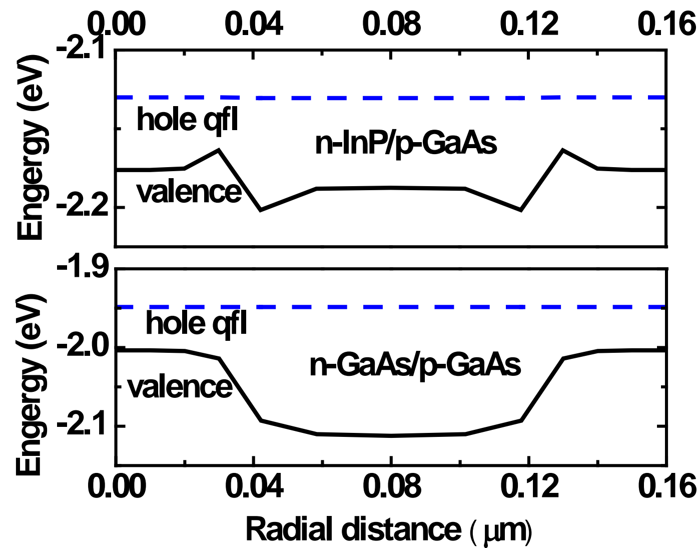

Fig. 2-10. Comparison of valence band lineups along the diameter between InP/GaAs and GaAs/GaAs core-shell structures at 3 V bias, together with the corresponding hole QFLs. Doping density is 1×10^{18} cm^{-3}.

It should be noted that the hole barrier will become larger when doping density is increased, the same as in the type-I structure (Fig. 2-8(b)). Therefore, hole injection efficiency will be very low if the core doping level is extremely high in an n-core/p-shell structure, no matter which kind of band lineup it is. On the other hand, the situation becomes quite complicated when the core becomes barriers for both electrons and holes. The details of carrier injection depend on the barrier heights and the applied bias, and further investigation is needed.

Fig. 2-11 shows the carrier density at the center of core region of InAs/GaAs and InP/GaAs core-shell structure as a function of core region radius. The shell thickness is fixed at 20 nm. In the InAs/GaAs type-I structure, carrier density decreases monotonically as core radius becomes smaller due to the reduction of the length of the depletion region in the core. However in the InP/GaAs type-II structure, as core radius increases, electron density firstly decreases while hole density increases quickly, then both electron and hole concentrations increase slowly with core radius, and the density finally decreases as the case in the type-I structure.

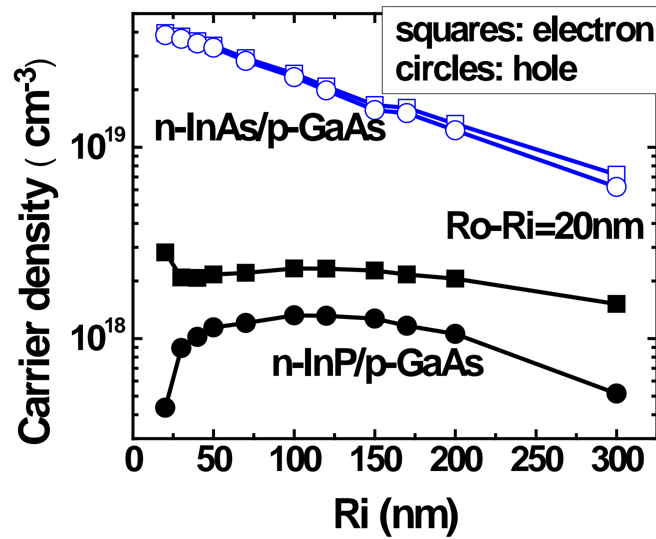

Fig. 2-11. Electrons (squares) and holes (circles) density at the center of core region of InAs/GaAs and InP/GaAs core-shell structure as a function of core region radius. The shell thickness is fixed at 20 nm. The bias voltage is 3 V and doping density is 1×10^{18} cm^{-3}.

Fig. 2-12 shows the carrier density at the center of core region of InAs/GaAs and InP/GaAs core-shell structure as a function of shell region thickness. The core radius is fixed at 50 nm. In the InAs/GaAs type-I structure, carrier density keeps almost a constant value as shell thickness varies. However in the InP/GaAs type-II structure, as shell thickness increases, carrier density firstly decreases slowly, and then increases. Electron and hole density becomes closer and closer as shell becomes thicker.

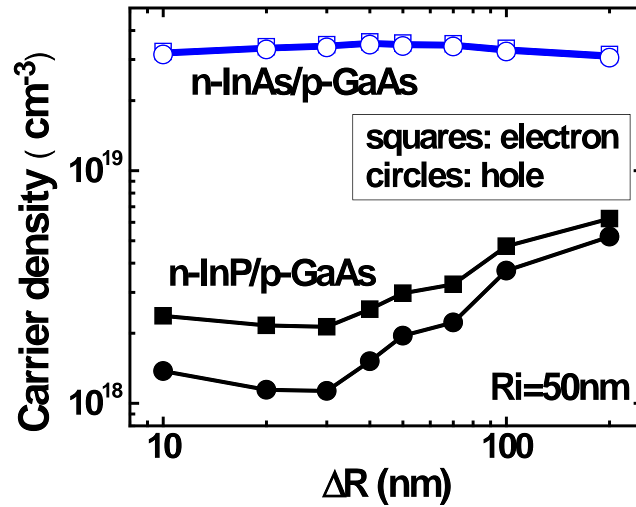

Fig. 2-12. Electrons (squares) and holes (circles) density at the center of core region of InAs/GaAs and InP/GaAs core-shell structure as a function of shell thickness. The core region radius is fixed at 50 nm. The bias voltage is 3 V and doping density is 1×10^{18} cm^{-3}.

2.3. Conclusion

We have performed a 3D numerical simulation on longitudinal p-i-n and coaxial p-n core-shell NWs. The influence of the bias, doping density, and band alignment on the electrical injection characteristics is systematically investigated. It is shown from the simulation results that the injection efficiency in the core-shell NW can be higher than that in the p-i-n NW, that is, the carrier density in the core region is higher than that in the i region under the same bias. Furthermore, the injected carrier density in core-shell NW can be maintained at a high level over a wide doping range due to the fully or nearly fully depleted core

region. Finally, type-II band alignment is also allowable for the core-shell NW to obtain the high level injection, while only type-I band profile is feasible in the longitudinal p-i-n NW.

3. Electrically driven single nanowire laser

In chapter 2, we proposed a core-shell NW structure where the core wire is coated by a heavily doped thin shell. It is, however, not easy to grow and fabricate such structures because doping the thin shell is difficult to realize and control. The metal contacts are hard to make to such structures. To reduce the complexity of the device fabrication while still taking the advantage of the high injection efficiency of the core-shell configuration, we investigated another variation of a core-shell structure for achieving single NW lasing. This design takes into account several issues of device physics: electrical injection efficiency, optical confinement of modes, and the feasibility of fabrication. We in particular studied the following issues of importance.

3.1. A proposal of device structure

The proposed design of the laser structure is shown in Fig. 3-1. The NW without intentional doping is grown in an anodic aluminium oxide (AAO) template [75]. A portion of the wire at each end is coated by p and n-type material, respectively, forming a local core-shell structure between the p and n contact region and the NW. The whole device is actually a general longitudinal p-i-n structure using core-shell configuration at p and n contact region to improve injection efficiency.

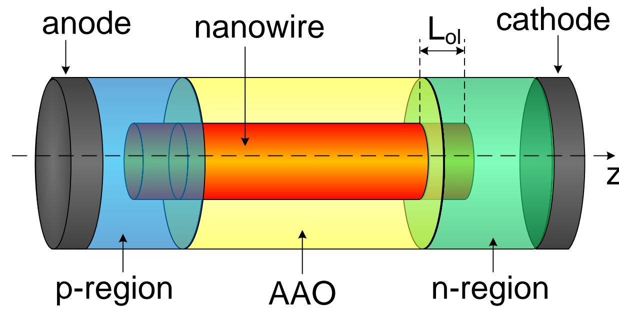

Fig. 3-1. A proposed device structure of an electrically-driven nanowire laser with core-shell structure at p and n contact region.

To show the improvement in electrical injection efficiency and possible lasing by electrical pumping in this structure, we use CdS and CdSe NWs as the core components. Since broad wavelength-tunable lasing in continuous composition grading CdSSe ternary alloy NWs by optical pumping has been reported [10], it is also interested to see whether electrically pumped lasing in CdSSe can be achieved using this structure. we use ZnO and GaN as the n and p contact material, respectively. ZnO is a well-studied typical n-type material. It can be doped with Al, Ga or In with doping density as high as 10^{19} cm^{-3} [76]. Common metal alloy such as Ti/Au can form ohmic contact with n-ZnO [77]. GaN can be doped with Mg to form p-type material with a doping concentration higher than 10^{18} cm^{-3} [78]. Ohmic contact can be made to p-GaN with Ni/Pt/Au alloy [79]. The bandgap, carrier mobility and effective mass of all semiconductors are taken from [80], and are listed in table 3-1. we assume the Auger

recombination coefficient is 10^{-30} cm^3s^{-1} and the carrier lifetime due to Shockley-Read-Hall recombination is 100 ns for all semiconductors. The radius of NW is fixed at 200 nm in the following study.

Table 3-1. Material Parameters at Room Temperature. [80]

	bandgap (eV)	electron affinity (eV)	electron mobility ($cm^2/V \cdot s$)	hole mobility ($cm^2/V \cdot s$)	electron effective mass (m_0)[a]	hole effective mass (m_0)
CdS	2.48	4.5	160	15	0.2	0.8
CdSe	1.74	4.95	720	40	0.12	0.6
ZnO	3.37	4.57	167	30	0.3	0.59
GaN	3.44	4.1	760	350	0.2	0.67

[a] m_0 is the electron rest mass

3.2. Electrical injection efficiency

We carried out a 3D device simulation to investigate the electrical injection properties of the structure shown in Fig. 3-1 by using Silvaco ATLAS. The doping density for ZnO and GaN is 10^{19} cm^{-3} and 10^{18} cm^{-3}, respectively in the simulation. Fig. 3-2 shows the carrier concentration along the length (z axis) of CdS and CdSe NWs for different overlapping length of the contact regions and the NW, L_{ol}, at 4V and 8V bias. The length of the wire is 10 μm, and the p region is on the left side and the n region is on the right side. The peaks in each curve are the boundaries of the contact regions and the AAO region. We can see in each panel of Fig. 3-2 that the carrier density increases as the overlapping length

increases. This is the evidence that the electrical injection efficiency is improved by using core-shell configuration. The more the p and n shell overlaps with the NW core, the longer the electrons and holes diffuse in the wire. The carrier concentration thus becomes higher when the core-shell overlapping length increases. Such improvement in carrier density by core-shell structure is more significant in the area coated by AAO between two contact regions. We can also see from the top two curves of each panel that the curve is tilted from left down to the right, showing a non-uniform carrier density distribution along z axis. This is because the hole mobility is much smaller than the electron mobility in both CdS and CdSe, the hole thus has smaller diffusion length than electron. To maintain charge neutrality, the electron concentration on the n contact side (the right side in Fig. 3-2) of the NW is clamped by the low hole concentration, resulting in lower carrier density than that on the p contact side (the left side in Fig. 3-2).

One method to improve the uniformity is to shrink the length of the NW, so that holes can reach closer to the right end of the NW. Fig. 3-3 shows the carrier concentration along the length of CdS and CdSe NWs for different NW length at 4V and 8V bias. The overlapping length is 1 μm at each contact region. We can see in each panel of Fig. 3-3 that the carrier density becomes higher and more uniform as NW becomes shorter. If the length of the NW is shorter than the diffusion length of electron, the carrier density on the p contact side will not

change a lot, as the cases of L_{NW}=3 and 5 μm at 4V bias in Fig. 3-3(a) and 3-3(c) and those of L_{NW}=3, 5 and 7 μm at 8V bias in Fig. 3-3(b) and 3-3(d). In these cases, the uniformity is improved mainly in the middle and on the right side of the NW. It seems that increasing bias from 4V to 8V contributes little improvement in the uniformity. Take L_{NW}=3 μm as the example, the highest concentration is 2.7 and 4.1 times the lowest concentration for CdS and CdSe at 4V, respectively; these values change to 2 and 3.4 at 8V bias, which is less than 50% improvement in uniformity while bias voltage increases by 100%. Note that the carrier density of the NW in any cross section plane vertical to z axis is always uniform.

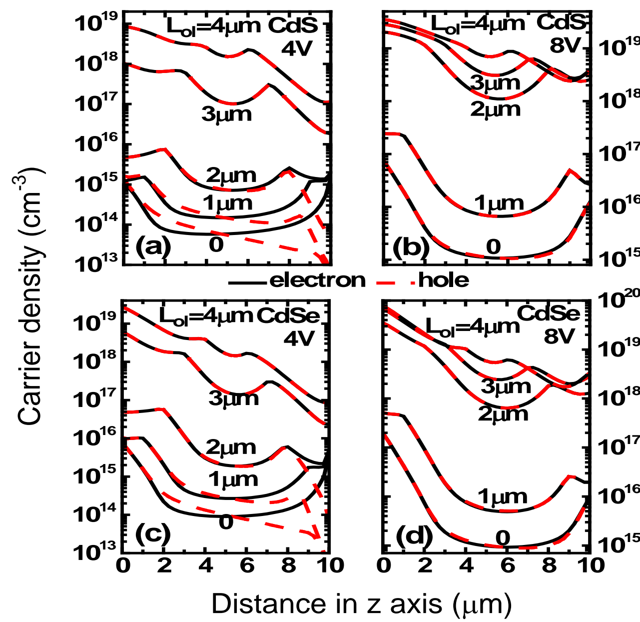

Fig. 3-2. Electron and hole concentration along z axis (the length direction of the NW) in (a) and (b) CdS, and (c) and (d) CdSe at 4V and 8V bias for different NW overlapping length between the contact regions and the NW. The length of the wire is 10 μm.

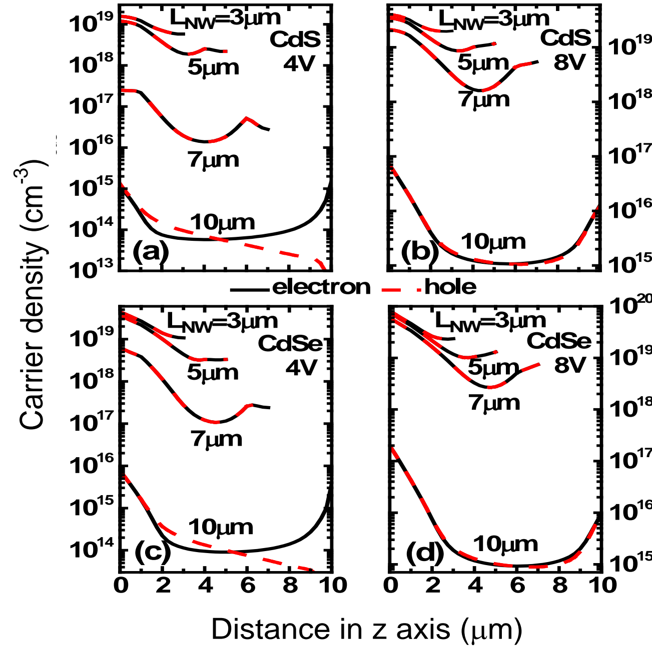

Fig. 3-3. Electron and hole concentration along z axis (the length direction of the NW) in (a) and (b) CdS, and (c) and (d) CdSe at 4V and 8V bias for different NW length. The overlapping length between the contact regions and the NW is 1 μm.

3.3. Mode analysis and laser design

To see the feasibility of electrically driven lasing in this structure, we performed a device design in the following. There are two contradictory design issues in electronics and optics for the core-shell configuration. First, to get better optical confinement of cavity mode in the NW, the refractive index contrast between core and shell should be as large as possible. This requires the AAO region whose refractive index is around 1.79 covers the entire NW side surface,

because the indexes of ZnO and GaN are both larger than 2. However, longer AAO region means shorter overlap between contact material and NW, which will lead to low electrical injection efficiency in the wire core. Second, long wire is desired to reduce the lasing threshold since the mirror (or reflection) loss is inversely proportional to the length of the laser cavity. Using long wire, on the other hand, degrades the uniformity of the carrier concentration along the wire, as shown in Fig. 3-3. Therefore, trade-offs should be made for these two issues. Since the operating wavelength in the case (from 500 nm to 710 nm) is larger than the NW diameter studied here (400 nm), the light is easier to spread out from the side of the NW compared to the cases where the operating wavelength is smaller than the NW diameter. To get better optical confinement, we limited the overlapping length up to 1 μm in the device design. The sacrifice in the injection efficiency can be somewhat compensated by increasing bias and reducing the NW length. Because the carrier density uniformity is disastrous in the long NW, e.g., 7 and 10 μm, as shown in Fig. 3-3, we restricted the NW length within 5 μm in the design.

In a NW waveguide with a few hundred nanometers in diameter, there are a lot of cylindrical waveguide modes such as HE_{11}, TE_{01}, TM_{01}, etc [81]. Among all these modes, HE_{11} mode has the best field overlap with the active core region since its field maximum is right at the center of the NW while other modes have

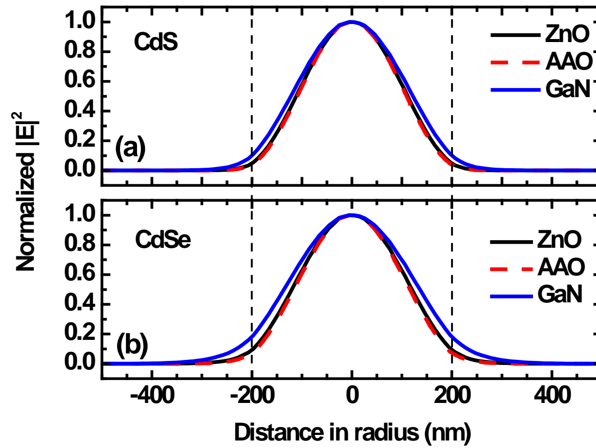

Fig. 3-4. Normalized $|E|^2$ profile of a HE11p mode along the radius of a CdS and a CdSe NW across ZnO, AAO and GaN region, respectively. The area between two dashed lines corresponds to the NW core while the rest of the area is shell region (ZnO, AAO or GaN). The length of the NWs is 3 μm, and the overlap between the contact region and NW is 0.5 μm.

nodes in their field at the center of the NW. The HE_{11} mode also has the longest cut-off wavelength, allowing a wide range of wavelength tuning. Therefore, we focus the design on the HE_{11p} (p=1, 2, 3,…) modes in the cavity. Here we performed the modal analysis by using COMSOL [82]. Fig. 3-4 shows the normalized $|E|^2$ profile of a HE_{11p} mode along the radius of a CdS and a CdSe NW across ZnO, AAO and GaN region, respectively. The length of the NWs is 3 μm, and the overlap between the contact region and NW is 0.5 μm. The mode profiles ($|E|^2$) across ZnO and AAO region almost overlap due to their similar refractive index. The profile across GaN is a little broader than the other two profiles, but

over 80% of the energy is still confined in the NW core even in the p-contact region (NW core/GaN shell). The structure thus provides a good optical confinement.

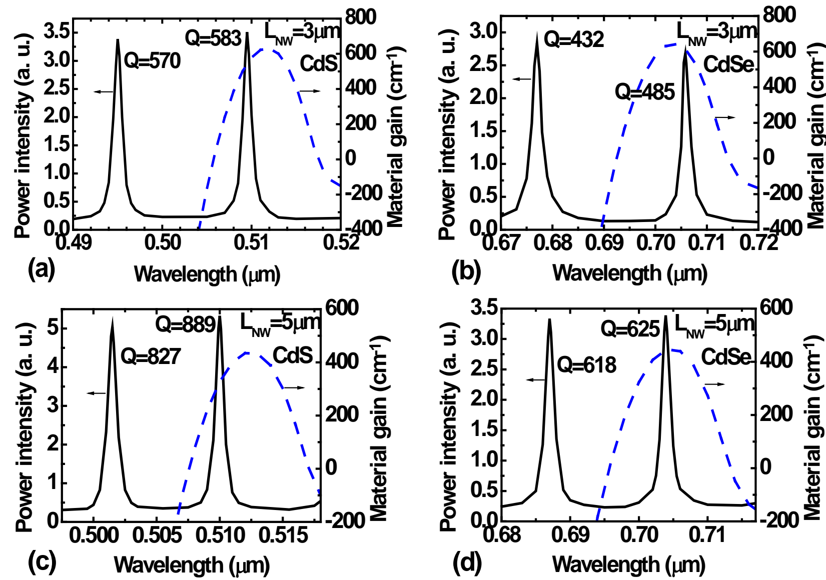

Fig. 3-5. Spectrum of HE11p (p=1, 2, 3,…) modes in (a) and (c) CdS, and (b) and (d) CdSe NWs for 3 and 5 μm length. The quality factor of each mode is marked in the figure. The material gain spectrum at lasing threshold of the very right mode in each panel is shown by the dashed lines as well.

Fig. 3-5 shows the mode spectrum (solid lines) in CdS and CdSe NW for 3 and 5 μm length. The overlapping length between the contact region and NW is 0.5 μm. Each peak of the solid lines represents a cavity mode whose quality factor is marked near it. By increasing the pumping, the mode above and the closest to the CdS or CdSe bandgap will reach its lasing threshold first, because it has the largest quality factor thus lowest lasing threshold. Such mode is the right-side

mode in each panel of Fig. 3-5. The material gain spectrum at lasing threshold of such mode, calculated from the free-electron model [83], is plotted in the figure by the dashed lines. We can see the right-side mode has the best overlap with the material gain spectrum, ensuring the single mode lasing operation. The peak material gain is up to 630 cm^{-1}, which is achievable in CdS and CdSe.

The semiconductor material gain is monolithically dependent on the carrier density according to the free electron model. Hence, once the threshold material gain is found, the corresponding threshold carrier density is determined as well. Since the carrier density is not uniform in NWs, the threshold carrier density is required to be the minimum value of the density profile throughout the NWs to achieve lasing. Fig. 3-6 shows the carrier concentration along the length of CdS and CdSe NWs in the same structure shown in Fig. 3-5 for different biases. The radiative recombination due to spontaneous emission has been included in the simulation. The calculation of the spontaneous emission rate will be shown in the next section. Note that the electron and hole concentrations are the same for all biases. To achieve the threshold carrier density, the bias voltage for CdS and CdSe with 3 μm in length is 5.1V and 3V, respectively; the bias voltage for CdS and CdSe with 5 μm in length is 8V and 4.8V, respectively. The minimum bias voltage applied to achieve lasing condition can be defined as the threshold voltage. We can see that even though the structure with longer wire has smaller threshold gain,

thus smaller threshold carrier density, than the structure with shorter wire, its threshold voltage is even higher, especially for CdS NW. This is the result of the worse uniformity of carrier density in longer wire. Therefore, good trade-off between threshold gain and threshold voltage is critical in the design of such laser structure.

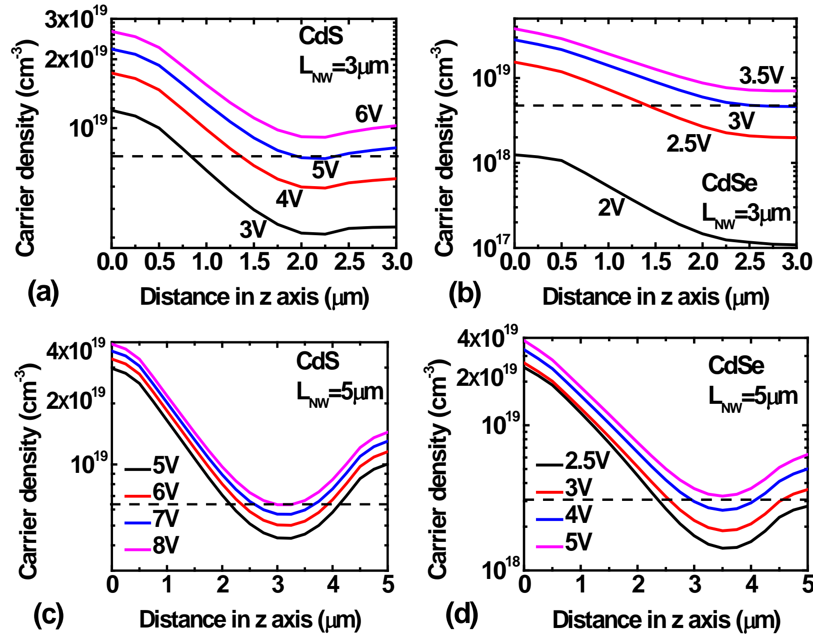

Fig. 3-6. Carrier concentration along the length of (a) and (c) CdS, and (b) and (d) CdSe NWs in the same structure shown in Fig. 4 for different biases. The threshold carrier density is indicated by the horizontal dashed line in each panel.

Fig. 3-7 shows the threshold voltage ((a) and (b)) and threshold gain ((c) and (d)) of CdS and CdSe NWs with various overlapping length and wire length. We can see from Fig. 3-7(a) and 3-7(b) that threshold voltage can reach a minimum

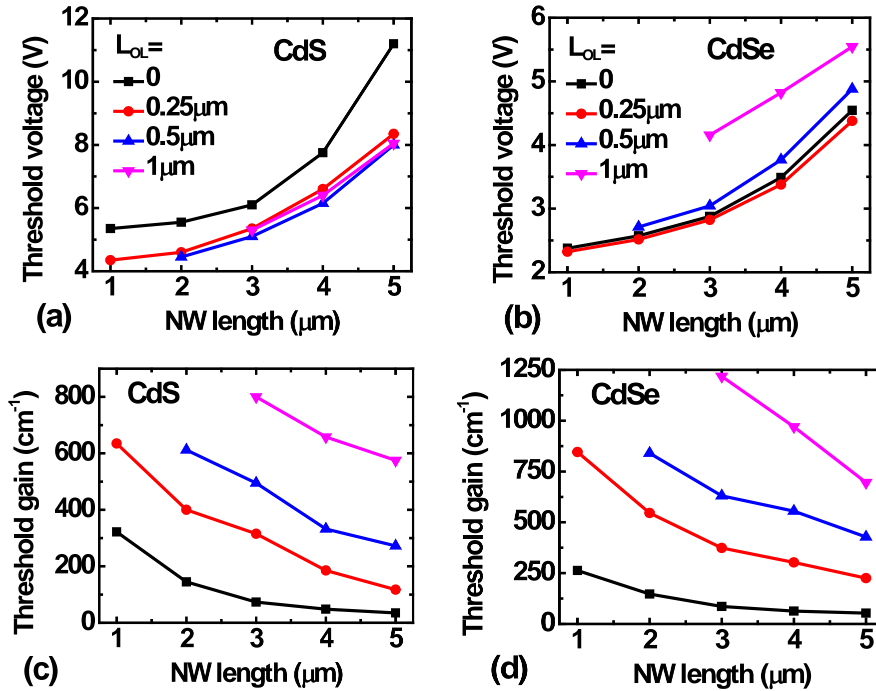

Fig. 3-7. Threshold voltage in (a) CdS and (b) CdSe, and threshold gain in (c) CdS and (d) CdSe with various overlapping length and NW length.

value for a certain overlapping length at any NW length. This overlapping length is around 0.5 μm in CdS and 0.25 μm in CdSe. The threshold voltage increases rapidly as NW length increases, for example, it is doubled when NW length increases from 1 μm to 5 μm in CdS. The threshold gain shown in Fig. 3-7(c) and 3-7(d) increases as overlap increases due to the weaker mode confinement in the NW, and decreases as NW length increases due to the reduction of mirror loss, as predicted at the beginning of this section. The design and optimization of the laser structure can be carried out according to this figure. For example, if it is required that the threshold gain not exceed 400 cm^{-1} while keeping the threshold voltage as low as possible, then the structure with 3-μm-long wire and

0.25-μm-long overlap meets the design requirement after applying the limit in Fig. 3-7.

3.4. Laser performance

To investigate the performance of the laser, we solved the following rate equations [83] to get the relation between output power and injection current (L vs. I) for the laser.

$$\begin{aligned} \frac{dN_s}{dt} &= \eta \frac{I}{qV_a} - R_{nr} - R_{sp,cont} - \sum_i R_{sp,i} - R_{st} N_{ph} \\ \frac{dN_{ph}}{dt} &= -\frac{N_{ph}}{\tau_p} + \Gamma R_{sp,i} + \Gamma R_{st} N_{ph} \end{aligned} \quad (3\text{-}1)$$

where N_s and N_{ph} is the carrier and photon density, respectively, I is the injection current, η is the injection efficiency, q is the charge of electron, V_a is the volume of active region, R_{nr} is the non-radiative recombination rate, $R_{sp,cont}$ and $R_{sp,i}$ are the spontaneous emission rate into continuum modes and cavity mode i, respectively, R_{st} is the stimulated emission rate, τ_p is the cavity photon lifetime, and Γ is the confinement factor [56]. The expressions for R_{nr}, $R_{sp,cont}$, $R_{sp,i}$, and R_{st} can be found in [84]. The output power is given by [83]

$$P_o = \hbar \omega N_{ph} V_a v_g \alpha_{rad} \quad (3\text{-}2)$$

where v_g is the group velocity and α_{rad} is the radiation loss. Fig. 3-8 shows the L-I curve of CdS and CdSe lasers with 3 μm NW length and 0.25 μm overlapping length. We can see from the double logarithmic scale a “S” shape with three

different slopes in each L-I curve. The lasers are turned on in the middle segment of the “S” shaped curve with the largest slope. The lasing threshold current hence is ranged from 30 to 80 μA for CdS and from 10 to 50μA for CdSe. This value is even smaller than the threshold current of 200 μA measured in the single CdS NW laser at 8 K [4] but comparable with that of >39 μA estimated for a single wire in the ZnO NW array laser at room temperature [17]. The external differential quantum efficiency, defined as $\eta_e = \frac{dP_o / \hbar\omega}{dI / e}$, above threshold is 2% for CdS and 1.8% for CdSe NW laser. When the injected current is equal to 1 mA, the output power for CdS NW laser is around 50 μW while that for CdSe NW laser is about 30 μW.

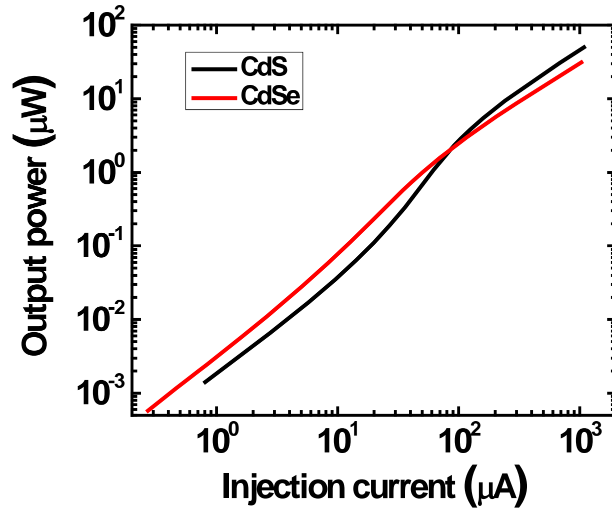

Fig. 3-8. L-I curve of CdS and CdSe lasers with 3 μm NW length and 0.25 μm overlapping length.

3.5. Conclusion

We proposed a new device structure as an electrically pumped single nanowire laser. This structure takes the advantage of core-shell configuration at p and n contact regions to improve electrical injection efficiency into the NW. Taking CdS and CdSe NWs as examples, we investigated the injection efficiency and carrier concentration uniformity in the wires for various NW lengths and core-shell overlapping lengths by carrying out a 3D device simulation. We then performed the design and optimization of the electrically driven laser based on the HE_{11p} (p=1, 2, 3, …) modes. We showed a design methodology to obtain a good trade-off between the threshold gain and the threshold voltage by tuning the NW length and overlapping length. The laser performance was further studied by solving rate equations. It is shown in an optimized structure that the threshold current is less than 80 μA at room temperature and the output power can exceed 30 μW if the injected current is greater than 1 mA for both CdS and CdSe NW lasers.

4. Metal-semiconductor surface plasmonic waveguides

In the previous two chapters, electrical injection characteristics of NWs and NW lasers have been studied. NW photonic devices usually have a length over 1 μm and a diameter over 100 nm, and are subjected to the diffraction limit in all three dimensions. Plasmonic structures are alternative device structures used to further reduce the size of photonic devices even below the diffraction limit. We then focused the study on plasmonic structures in the next three chapters.

4.1. Fundamental physics of surface plasmon polaritons

The exponential decay of optical wave near a metallic surface can serve as a mechanism of confining optical wave to a much smaller spatial region than otherwise possible. A proper design of metal-dielectric (semiconductor) structures can therefore lead to nanophotonic devices smaller than the wavelength. At the nanometer scale, metal penetration depth becomes comparable to the feature size of the devices. Therefore, metal cannot be treated as a perfect metal anymore, rather, active excitations of plasmon modes in metals are important. Furthermore, manipulation and engineering of such plasmon modes become interesting and sometimes very beneficial. Before probing further, some basic physics of plasmonics are introduced first [85].

Over a wide frequency range, the optical properties of metals can be

explained by a plasma model, where a gas of free electrons of number density N moves against a fixed background of positive ion cores. The motion for an electron of the plasma sea subjected to an external electric field $\mathbf{E}$ can be described by the following equation:

$$m\ddot{\mathbf{u}} + m\gamma\dot{\mathbf{u}} = -e\mathbf{E} \tag{4-1}$$

where the damping factor $\gamma = 1/\tau$, and τ is known as the relaxation time of the free electron gas. For a time-harmonic field, using macroscopic polarization $\mathbf{P} = -eN\mathbf{u}$, the electric displacement is in the form:

$$\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P} = \varepsilon_0\left(1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}\right)\mathbf{E} \tag{4-2}$$

where $\omega_p = \left(\frac{Ne^2}{\varepsilon_0 m}\right)^{1/2}$ is the plasmon frequency of the free electron gas. Therefore, the dielectric function of the free electron gas is

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega} \tag{4-3}$$

For large frequency where $\omega\tau >> 1$, $\varepsilon(\omega)$ becomes dominantly real:

$$\varepsilon(\omega) \approx 1 - \frac{\omega_p^2}{\omega^2} \tag{4-4}$$

The AC conductivity $\sigma(\omega)$ of metals is given by the classical Drude model:

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau} \tag{4-5}$$

where $\sigma_0 = \frac{Ne^2\tau}{m}$ is the DC conductivity. Comparing σ_0 with ω_p, the

dielectric function can be written as

$$\varepsilon(\omega) = 1 + \frac{i\sigma(\omega)}{\varepsilon_0 \omega} \quad (4\text{-}6)$$

This result can be obtained directly from Maxwell equations.

All the description above has assumed an ideal metal, but real metals are more important for applications. For noble metal (e.g. Au, Ag, Cu), the filled d band close to the Fermi surface causes a highly polarized environment. This residual polarization due to the positive background of the ion cores can be described using an additional polarization term $\mathbf{P}_\infty$. This effect is therefore described by a dielectric constant ε_∞ (usually $1 \leqslant \varepsilon_\infty \leqslant 10$), and the dielectric function is now:

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\gamma\omega} \quad (4\text{-}7)$$

Usually for frequency higher than infrared region, transitions between electronic bands occur in the noble metal, and the free electron gas model breaks down. For example, Fig. 4-1 shows the real and the imaginary parts of the dielectric function for silver [86] and Drude model fits to the data. Clearly, this model is not adequate for describing either real or imaginary parts at high frequencies due to the interband transition. Therefore, numerical fitting to the dielectric function is required for the real metal.

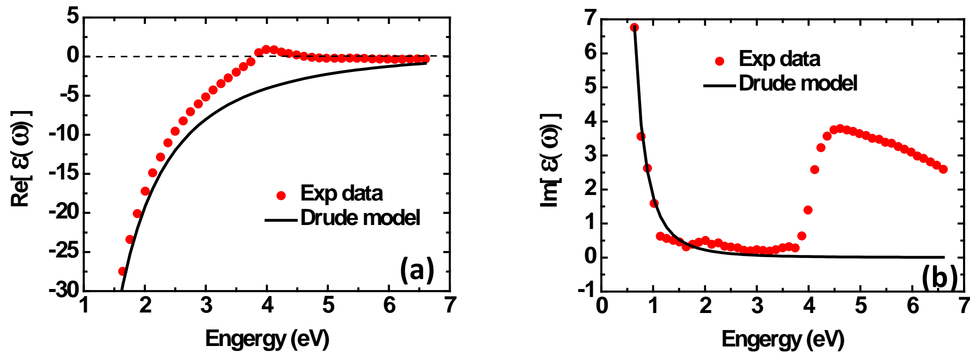

Fig. 4-1. (a) Real and (b) imaginary parts of the dielectric function of silver [86]. Drude model fits to the data.

Surface plasmon polaritons are electromagnetic excitations propagating at the interface between a dielectric and a conductor, evanescently confined in the perpendicular direction. These electromagnetic surface waves arise via the coupling of the electromagnetic fields to plasma oscillations of the conductors. The simplest geometry sustaining SPPs is that of a single, flat interface (Fig. 4-2) between a dielectric, non-absorbing half space ($x > 0$) with positive real dielectric constant ε_1 and an adjacent conducting half space ($x < 0$) described via a dielectric function $\varepsilon_2(\omega)$ whose real part is negative. In the case of Fig. 4-2, these two layers are a semi-infinite large semiconductor and a semi-infinite large metal respectively. Only TM polarized SPPs exist for this single interface configuration, and the following equation should be satisfied according to the boundary condition:

$$k_1\varepsilon_2 + k_2\varepsilon_1 = 0 \tag{4-8}$$

and one also has for two regions:

$$k_1^2 = k_z^2 - \frac{\omega^2}{c^2}\varepsilon_1 \tag{4-9a}$$

$$k_2^2 = k_z^2 - \frac{\omega^2}{c^2}\varepsilon_2 \tag{4-9b}$$

The dispersion relation of SPPs propagating at the interface between the two half spaces thus is given by

$$k_z = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}} = k_0\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}} \tag{4-10}$$

where $k_0 = \omega / c$. The characteristic surface plasmon frequency is the frequency at which k_z goes to infinity. It can be obtained using Eq. (4-4), and has the value

$$\omega_{sp} = \frac{\omega_p}{\sqrt{1+\varepsilon_1}} \tag{4-11}$$

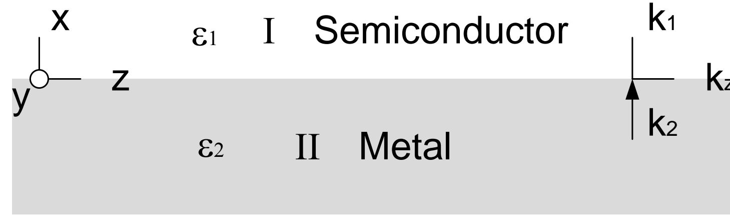

Fig. 4-2. Geometry for SPP propagation at a single interface between a metal and a semiconductor.

Two SPPs can be coupled with each other by putting two metal/dielectric interfaces close together. Fig. 4-3 shows a possible system consisting of a thin mid-layer semiconductor with two thick enough (larger than the SPPs decaying

length) metal claddings. This structure serves as a planar waveguide for surface-plasmonic wave. For non-oscillatory TM modes, the field components in each layer can be written as

$x \le -d/2$	$\lvert x \rvert \le d/2$	$x \ge d/2$
$H_y = Ae^{k_2 x + ik_z z}$	$H_y = B(e^{-k_1 x} + e^{k_1 x})e^{ik_z z}$	$H_y = Ae^{-k_2 x + ik_z z}$
$E_x = A\dfrac{k_z}{\omega\varepsilon_2\varepsilon_0}e^{k_2 x + ik_z z}$	$E_x = B\dfrac{k_z}{\omega\varepsilon_1\varepsilon_0}(e^{-k_1 x} + e^{k_1 x})e^{ik_z z}$	$E_x = A\dfrac{k_z}{\omega\varepsilon_2\varepsilon_0}e^{-k_2 x + ik_z z}$
$E_z = -A\dfrac{ik_2}{\omega\varepsilon_2\varepsilon_0}e^{k_2 x + ik_z z}$	$E_z = B\dfrac{ik_1}{\omega\varepsilon_1\varepsilon_0}(e^{-k_1 x} - e^{k_1 x})e^{ik_z z}$	$E_z = A\dfrac{ik_2}{\omega\varepsilon_2\varepsilon_0}e^{-k_2 x + ik_z z}$

where $A/B = e^{-(k_1-k_2)d/2} + e^{(k_1+k_2)d/2}$. The dispersion relation is described by two transcendental equations:

$$k_1\varepsilon_2 \tanh(\frac{k_1 d}{2}) + k_2\varepsilon_1 = 0 \tag{4-12a}$$

$$k_1\varepsilon_2 \coth(\frac{k_1 d}{2}) + k_2\varepsilon_1 = 0 \tag{4-12b}$$

for symmetric (even) and anti-symmetric (odd) modes, respectively. The propagation wavevector is defined by

$$k_z^2 = \varepsilon_{1,2}k_0^2 + k_{1,2}^2. \tag{4-13}$$

It is easy to show that the eigenvalue Eq. (4-12) for the three layer structure becomes the well-known expression for the SPP mode along a bi-layer interface in the limit of large $\mathrm{Re}[k_1 d]$, i.e., Eq. (4-10). For a finite *d*, eigenvalue Eq. (4-12) has to be solved numerically.

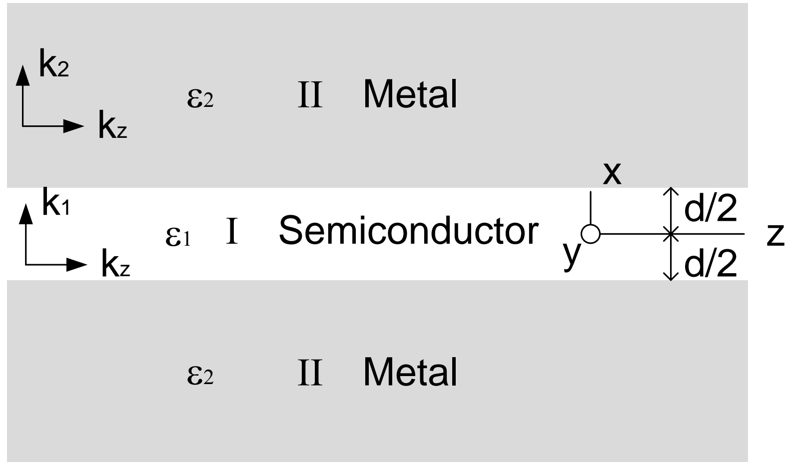

Fig. 4-3. Geometry of a three-layer structure consisting of a thin mid-layer semiconductor sandwiched between two infinite (or large enough) metal layers.

4.2. Planar surface plasmon waveguides with optical gain

In the following work, a study on finding the net gain near the SPP resonance was carried out on the structure shown in Fig. 4-3, and the results show that a giant modal gain in a MSM waveguide near the SPP resonance is able to be realized.

4.2.1. Modal gain in active plasmonic waveguide

The real part of the semiconductor dielectric constant is 12 and the metal is silver whose dielectric function is taken from Ref. [86] and curve fitted for our numerical solution. All energy-loss channels from semiconductor to metal are contained in the experimentally determined dielectric function $\varepsilon_2(\omega)$ for silver. The material gain in the semiconductor can be modeled by adding a small

negative imaginary part ε_1'' to the dielectric constant, and the material gain is given by $G_0 = -\varepsilon_1''\omega / (nc)$, where n is the real part of refractive index. As we will show in the later, such a frequency independent ε_1'' can be justified by using a more realistic gain model. By solving Eq. (4-12a), we can get the dispersion relation for a symmetric TM mode in the MSM waveguide as an example.

The calculation was started from the frequency far below SPP resonance, and the imaginary part of k_z as a function of photon energy with d=100 nm for 5 different values of ε_1'', corresponding to different levels of material gain in semiconductors was plotted in Fig. 4-4. The modes in the MSM waveguide in the frequency range far below SPP resonance are normal optical modes (e.g., the left side of the figure), since they have linear dispersion relation. As material gain increases, Im[k_z] decreases from positive to negative, meaning the occurrence of a small modal gain. When frequency approaches the SPP resonance frequency, the small modal gain could disappear due to the increased metal loss as for $\varepsilon_1'' = -0.3$, or get amplified if material gain overcomes the metal loss as for $\varepsilon_1'' = -0.4$ and –0.5. The small gain at low frequency becomes a huge gain or loss near SPP resonance depending on the delicate balance of the semiconductor material gain and metal loss, together with the respective modal profiles in each of the media. The sign choice of k_z from Eq. (4-12a) has been an issue of discussion in the literature. The choice is based on the gain or loss far below SPP resonance as

shown in Fig. 4-4, where the sign of the imaginary part of k_z is determined by the net gain or loss we have in our system. Once the sign is known far below SPP resonance, continuous tracking of a solution over the frequency domain of our interest determines the sign uniquely.

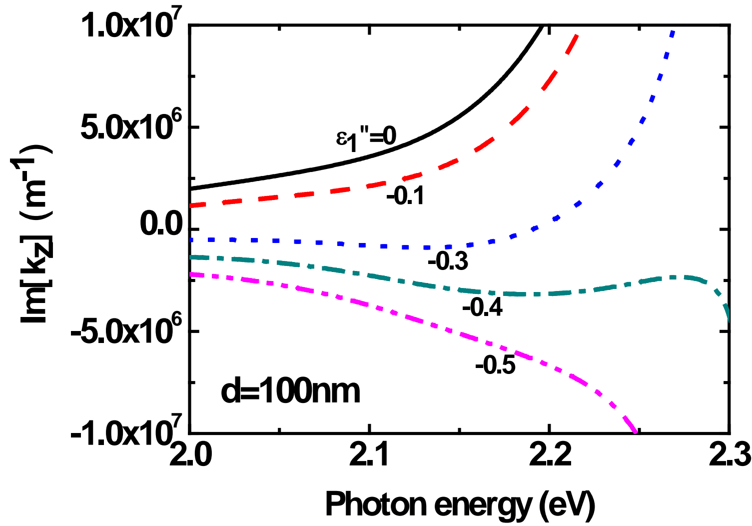

Fig. 4-4. The imaginary part of k_z as a function of photon energy with d=100 nm for 5 different values of the imaginary part of semiconductor dielectric constant ε_1''.

For the situation around SPP resonance, Fig. 4-5(a) and (b) show respectively the real and imaginary parts of k_z as a function of photon energy with d=100 nm and 200 nm for 4 different values of ε_1'', compared with the dispersion relations in a MS bi-layer structure. It is seen that all structures have the same behavior around SPP resonance (within a range of ±25 meV around the peak), independent of the number of interfaces or the thickness of the middle layer. This

means the guided modes in the MSM waveguide are decoupled into two independent SPP modes at the two MS interfaces. We observe in Fig. 4-5(a) that material gain in semiconductor layer can significantly modify the resonance behavior, as can be easily seen by comparing different panels of Fig. 4-5(a). There exists an optimum level of semiconductor gain (around $\varepsilon_1'' = -0.4$), where SPP resonance has the strongest response (with highest and narrowest resonance peak). For application of active optical devices, it is more important to study the imaginary part of k_z as it describes the optical loss or gain. The *modal gain* is given by, G_m ($=-2\mathrm{Im}[k_z]$), which describes the overall gain that a given mode experiences in a waveguide containing active region. For example, the intensity of a field mode traveling along the z-axis of the waveguide can be written as $I = I_0 e^{G_m z}$, where I_0 is the initial intensity. The *material gain*, G_0 defined earlier, is a pure material property of the semiconductor and describes how a plane wave is amplified in an infinitely large medium without a waveguide. In a typical dielectric waveguide with gain medium, the modal gain is typically much smaller than material gain. Note that the modal gain in a MSM structure already takes into account the metal loss and is therefore a net gain for a mode. In Fig. 4-5(b) we see that, for small material gain (e.g., when $\varepsilon_1'' > -0.3$), $\mathrm{Im}[k_z]$ is positive or modal gain is negative near SPP resonance due to metal loss. However, $\mathrm{Im}[k_z]$ becomes extremely negative or modal gain becomes a giant positive value near the SPP

resonance when the semiconductor material gain G_0 is large enough (or ε_1'' is enough negative). For $\varepsilon_1'' = -0.4$, which corresponds to a material gain of about 1.35×10^4 cm^{-1}, the modal gain at the SPP resonance is around 2×10^7 cm^{-1} (see Fig. 4-5(b)), more than 1000 times larger than the material gain. Such a large modal gain has not been seen in any other systems or situations. It may sound counter-intuitive that a guided mode can experience more gain (modal gain) per unit propagation length than the gain provided by the gain medium (material gain). The ratio of the two is known as *confinement factor* in conventional waveguide with a gain material and much smaller than 1 typically. For instance, a GaAs/AlGaAs waveguide with a single quantum well has a confinement factor of 0.02-0.03. The concept of the confinement factor was indeed an issue of debate when a larger-than-unity confinement factor was first shown in the case of a semiconductor NW [87]. In that case, the physical origin for the larger modal gain is the group velocity (defined as $v_g = \partial\omega / \partial(\mathrm{Re}[k_z])$) slowing down dramatically in a strongly guided situation, leading to a larger optical gain per unit length than the material gain.

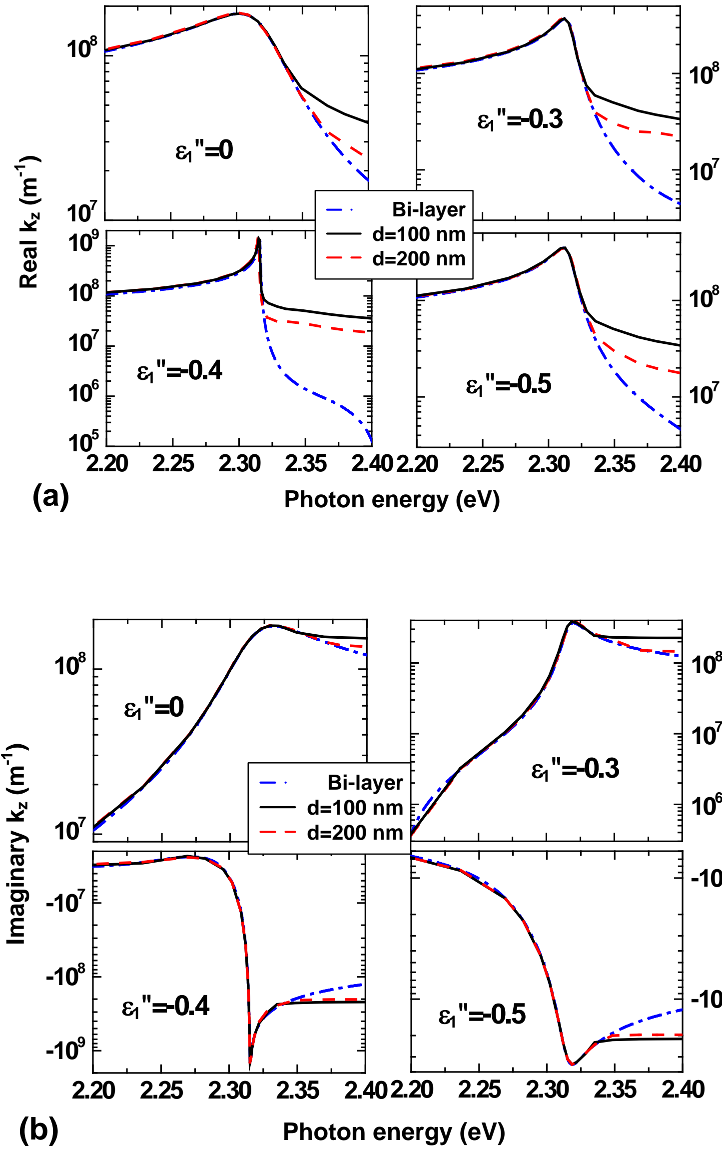

Fig. 4-5. Real (a) and imaginary (b) parts of wavevector k_z versus photon energy for d=100 nm (solid) and 200 nm (dash) for the MSM structure or for the bi-layer MS structure (dash dot).

4.2.2. Explanation of giant modal gain

To understand the physical origin of the giant modal gain and to see if the huge enhancement of modal gain is related to the slowing down of group velocity, we plot in Fig. 4-6(a) the group velocity for $\varepsilon_1''=-0.4, -0.5$, compared with the corresponding modal gain in Fig. 4-6(b). As expected, the group velocity in a plasmonic waveguide can take both positive and negative values around the SPP resonance. However, the maximum of modal gain does not correspond to any special features of the group velocity. Recall that, in a dispersive medium with loss or gain, a physically more fundamental velocity is not group velocity, but energy velocity v_E, or the velocity of energy propagation for a particular mode. Energy velocity [88] is defined as the ratio of energy flux density S over stored energy density w, i.e., $v_E=S/w$. Since energy propagates in opposite directions in semiconductors and metals, an average energy velocity $\bar{v}_E$ is introduced and is given by:

$$\bar{v}_E = \frac{\int w^s v_E^s d\mathbf{v} + \int w^m v_E^m d\mathbf{v}}{\int w^s d\mathbf{v} + \int w^m d\mathbf{v}} = \frac{\int S^s d\mathbf{v} + \int S^m d\mathbf{v}}{\int w^s d\mathbf{v} + \int w^m d\mathbf{v}} \tag{4-14}$$

where superscripts "*s*" and "*m*" represent quantities in semiconductor and metal respectively. The *z*-component of the energy flux density S_z in this case can be written as

$$S_z = \frac{1}{2}\mathrm{Re}[E_x H_y^*] \tag{4-15}$$

and the energy density in semiconductor is defined in the usual way as

$$w^s = \frac{1}{4}\mathrm{Re}[\mathbf{E}^s \cdot \mathbf{D}^{s*} + \mathbf{B}^s \cdot \mathbf{H}^{s*}] \tag{4-16}$$

while energy density in metal is derived from Poynting's theorem in linear, lossy, and dispersive media [88]:

$$w^m = \frac{1}{4}\mathrm{Re}[\frac{d(\omega\varepsilon_2)}{d\omega}\mathbf{E}^m \cdot \mathbf{E}^{m*} + \mathbf{B}^m \cdot \mathbf{H}^{m*}] \tag{4-17}$$

Figure 4-6(c) shows the z-component of the average energy velocity as a function of photon energy for $\varepsilon_1'' = -0.4, -0.5$. It is seen that $\overline{v}_E$ has its minimum value whenever G_{m} is the largest. This relationship is true for all the values of the material gain. This means that the giant modal gain comes from the slowing down of the average energy propagation. A slowed energy transport allows more energy exchange events between waves and gain/loss media, or a slow wave will experience more absorption or emission events than fast wave when traveling through the same distance. If the total system is lossy, this enhanced exchange will lead to large modal loss near resonance; similarly in the opposite situation, it will lead to a huge modal gain. At the SPP resonance, the slowest energy velocity ($\varepsilon_1'' = -0.4$) is about 200 m/s. This is a factor of 1.5 million slowing down compared to the speed of light in vacuum. Such a dramatic slowing down allows much longer interaction time per unit propagation length, so that any small model gain increases to a giant value.

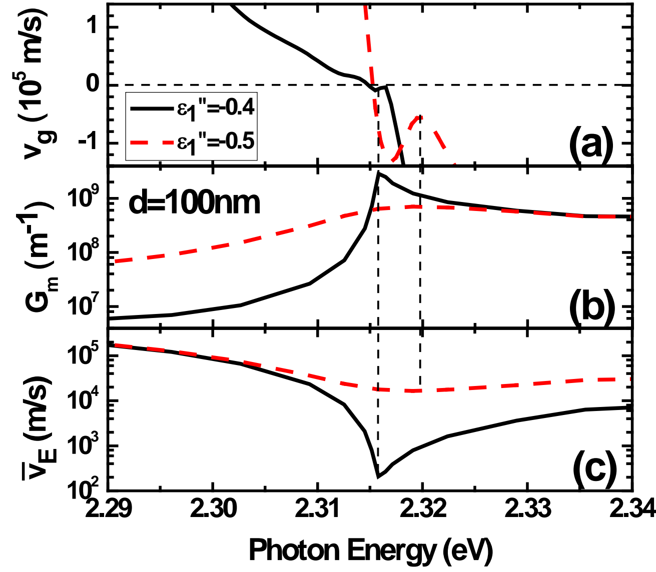

Fig. 4-6. Group velocity (a), modal gain (b), and average energy velocity (c) versus photon energy for ε_1'' =0.4 (solid) and 0.5 (dash) with d=100 nm.

Fig. 4-7(a) shows the three-dimensional distribution of two electric field components E_x and E_z in the MSM waveguide with d=100 nm and $\varepsilon_1'' = -0.4$ for three different photon energies, as marked in the dispersion curve in Fig. 4-7(b). We can see from Fig. 4-7(a) more intuitively how the electric fields are amplified when traveling along z-axis in the waveguide, especially for the case near SPP resonance.

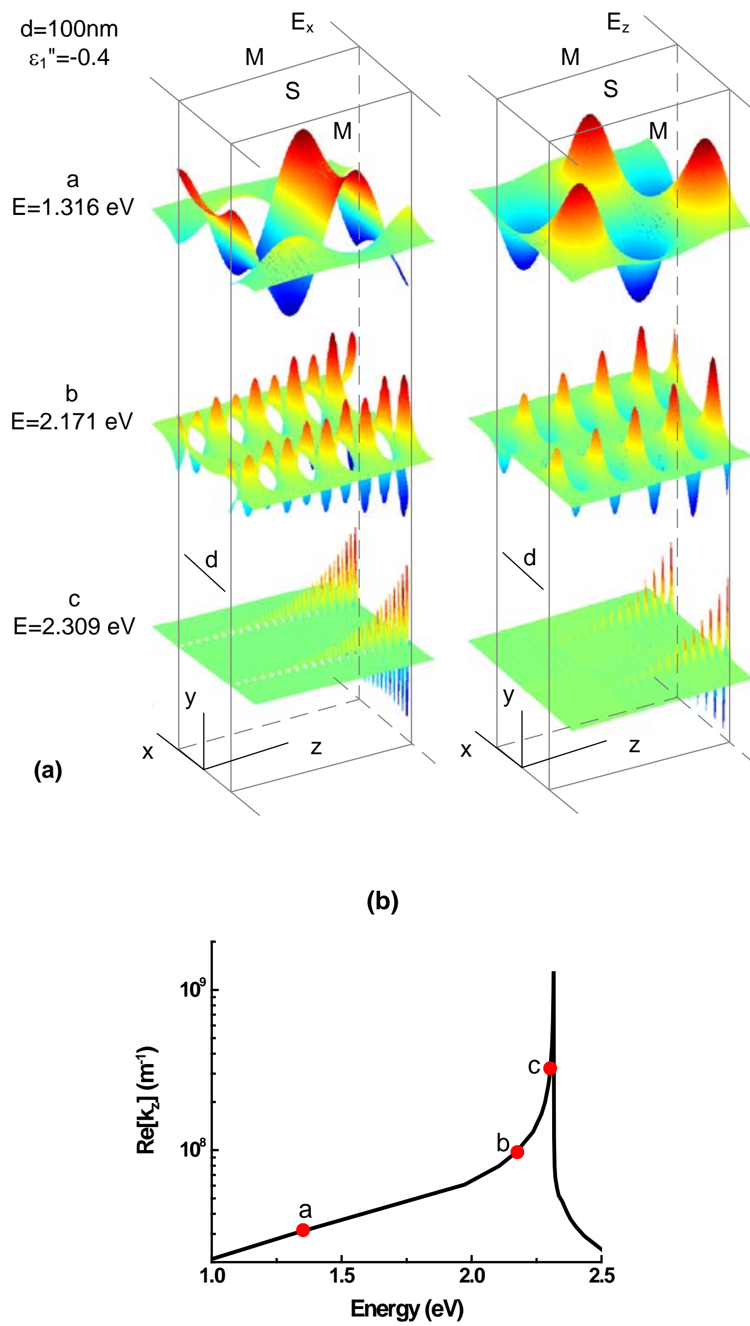

Fig. 4-7. Distribution of two electric field components E_x and E_z in the MSM waveguide with d=100 nm and $\varepsilon_1'' = -0.4$ at 3 different photon energies (a) as indicated in the dispersion curve (b).

The giant modal gain also exists in the anti-symmetric SPP mode which can be calculated from Eq. (4-12b). Fig. 4-8 shows modal gain and average energy velocity of an anti-symmetric SPP mode for ε_1''=0.4 and 0.5 with d=100 nm as a function of photon energy. The value and the position of the modal gain are identical to those of the symmetric SPP mode, and the cause of the giant modal gain can be explained by the same theory shown before. Therefore, the modal properties of anti-symmetric SPP mode are similar to those of symmetric SPP mode near SPP resonance.

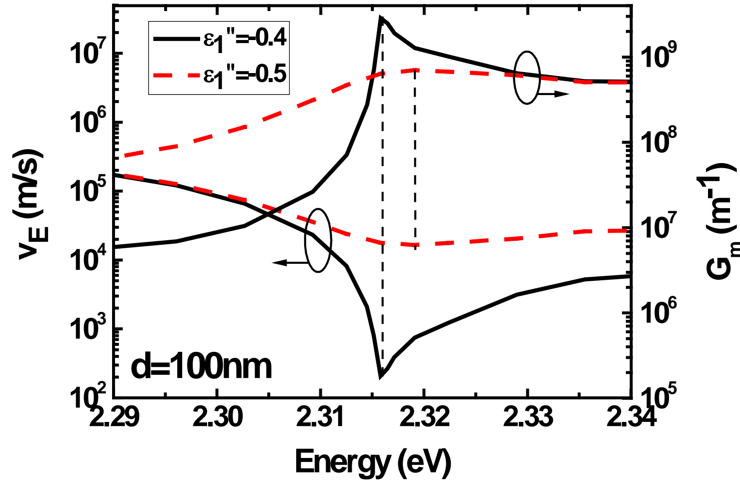

Fig. 4-8. Average energy velocity and modal gain versus photon energy of an anti-symmetric SPP mode for ε_1''=0.4 (solid) and 0.5 (dash) with d=100 nm.

4.2.3. Feasibility of giant modal gain

While the enhancement of modal gain is obviously of great significance, this only occurs when the material gain is sufficiently large to balance out the metal

loss and to achieve such a huge modal gain. It is thus important to assess the feasibility of such material gain. For example, in order to obtain the giant modal gain in the MSM waveguide with 100 nm-thick middle layer, the required minimum material gain ($G_{\min}$) is about 1.34×10^{4} cm^{-1} (corresponding to ε_1'' ≈ -0.396). While material gain of this magnitude or as large as 2×10^{4} cm^{-1} has been shown possible in wide-gap semiconductor quantum wells such as nitrides or II-VI semiconductors [89], it is more desirable to have this gain enhancement occur at a lower material gain level. A possible initial experiment can be performed at low temperature. As is well known, optical loss in metals such as silver is significantly reduced at low temperature, while gain in semiconductor is significantly increased. It was shown experimentally that the imaginary part of silver dielectric constant is reduced as temperature decreases [90]. The reduction rate at 3 eV is about 5×10^{-4} K^{-1}. Groeneveld *et al.* [91] theoretically estimated the temperature coefficients of both real and imaginary parts of silver dielectric constant at 2 eV to be 8.5×10^{-4} K^{-1} and 1.5×10^{-3} K^{-1} respectively. Using the experimental results of Ref. [90], we estimate that the required imaginary part of the semiconductor dielectric constant can be as low as -0.271 at 77 K to achieve the giant modal gain, corresponding to a material gain of 9×10^{3} cm^{-1}.

To find the proper materials to test the feasibility of giant modal gain, we need to consider the factors of both bandgap and refractive index. It is required

that the position of SPP resonance, which is determined by the material refractive index, must be above the material bandgap, hence one can draw a upper limit for such requirement, as shown in the solid line in Fig. 4-9. The position of SPP resonance should not be too high so that it is still located within the range of material gain spectrum. The width of material gain spectrum in common semiconductors is usually smaller than 200 meV [89], one thus can draw the lower limit 200 meV below the upper limit, as shown by the solid line in Fig. 4-9. It seems only II-VI group materials may fit within the allowable range. The bandgap of several II-VI materials versus their refractive index in both cubic (cub) and hexagonal (hex) structures is shown in Fig. 4-9 at 77 K and 300 K. We can see only CdS, ZnTe and ZnSe fall in the allowable region. However, some alloy between those II-VI materials can also fall in the region, such as ZnCdSe. As a specific example, we considered a MSM waveguide with $Zn_{0.8}Cd_{0.2}Se$ as the middle layer material at 77 K. We calculated the material gain of $Zn_{0.8}Cd_{0.2}Se$ via a free-carrier model [89], using material parameters of ZnSe and CdSe in Ref. [92]. The material parameters of $Zn_{0.8}Cd_{0.2}Se$ were obtained by linear interpolation. The calculated gain G_0 spectrum is shown in Fig. 4-10 (dashed line) with carrier density 1.1×10^{19} cm^{-3}. Such a gain level at 77 K is reasonable in $Zn_{0.8}Cd_{0.2}Se$ [93]. The actual spectral dependence of imaginary part of the dielectric function obtained using $\varepsilon_1'' = -ncG_0/\omega$ is then used to solve for the

propagation wavevector, k_z, which is shown for the fundamental TM mode in Fig. 4-10 (solid lines). As we see, a huge negative imaginary k_z (red curve) occurs near SPP resonance, meaning the giant modal gain occurs in this more realistic case. In addition, this specific example also justifies our earlier treatment of imaginary part of the dielectric function of the semiconductor as a frequency independent value, since semiconductor gain is much broader than the imaginary part of the wavevector near SPP resonance.

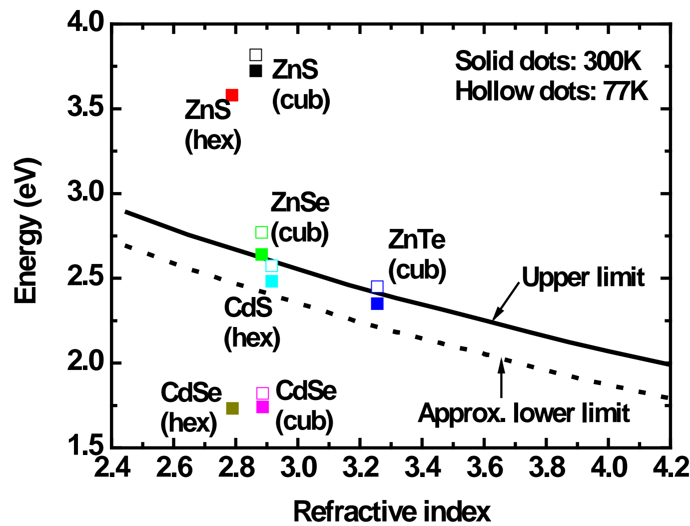

Fig. 4-9. Bandgap energy versus refractive index of some II-VI materials in both cubic (cub) and hexagonal (hex) structures at 77 K and 300 K. The upper limit and approximate lower limit of allowable material region are shown by the solid line and dashed line, respectively.

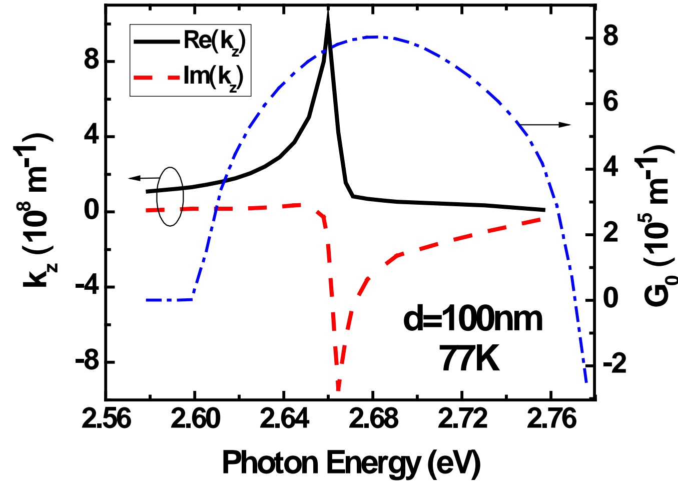

Fig. 4-10. Real (solid) and imaginary (dash) part of k_z for the fundamental TM mode of silver/$Zn_{0.8}Cd_{0.2}Se$/silver waveguide with d=100 nm at 77 K. G0 (dash dot) in the $Zn_{0.8}Cd_{0.2}Se$ layer was calculated at carrier density 1.1×10^{19} cm^{-3}.

It is possible to search such giant gain by experiments. As a first experiment, we suggest a MS structure with grating coupling for the incident light to excite SPP. In a MS structure, pumping can be incident from the semiconductor side and the SPP can be excited through spontaneous emission. In reality with long propagation length, nonlinear saturation could be also important both inside semiconductor and inside metal. All these issues need to be examined.

4.3. Confinement factor in the planar surface plasmon waveguides

As mentioned in the last section, the ratio of the modal gain to the material

gain is often called confinement factor[1]. In a typical semiconductor waveguide such as GaAs/AlGaAs, the confinement factor is a few percent per quantum well. Due to small difference between the indices of refraction of the core and clad layers, the confinement factor in such "weakly" guided structure can be approximated by a "energy ratio": the energy in the active core material to that in the entire wave. Even though this usage of confinement factor has been popular and sometimes assumed as generally valid, it has been realized recently that the error of energy-ratio definition can be very significant. Nanophotonic waveguides provide many such examples where the power-ratio definition fails. The reason is that nanophotonic devices necessarily require strongly confined waveguides, due to small size of waveguides which requires often large difference of core-clad indices of refraction. A remarkable example of such devices is provided by semiconductor NW lasers where confinement factor larger than unity was demonstrated [87], in direct contradiction to the power-ratio definition. Another example is metal-dielectric or metal-semiconductor waveguide which becomes favored choice for many nanophotonic devices. The giant modal gain presented in section 4.2 indicates an unusual property of confinement factor, but this issue has not been investigated in detail so far. Given the widespread use of plasmonic waveguide and other nanophotonic waveguide and the existence of large modal

[1] This ratio is also called gain coefficient in some literatures.

gain, it is necessary to discuss the issue of confinement factor in plasmonic waveguide in detail. Although many previous authors have discussed the confinement factor [94-97], the features of the confinement factor in a metal-semiconductor multi-layer structure have not been examined in detail, especially near SPP resonance. In this section we consider confinement factor in a MSM waveguide and express it in a way that it is more transparent to show its physical relation to the power transport in the structure.

4.3.1. Derivation of confinement factors in a plasmonic waveguide

In the linear regime, the intensity, I, of a given waveguide mode propagates along the z direction according to the following equation:

$$\frac{dI}{dz} = G_{net} I \tag{4-18}$$

where G_{net} is the net modal gain which is the difference between the modal gain G_{mod} and modal loss α_{mod}, or $G_{net} = G_{\text{mod}} - \alpha_{\text{mod}}$. The concept of "modal gain" used in section 4.2 is actually the net modal gain here. Similar to the concept of modal gain, modal loss is the absorptive loss experienced by a propagating mode. The modal gain and loss are in general different from the so-called material gain G_0 and material loss α_0, due to the waveguiding effects. The material gain and loss can be defined in general as:

$$G_0 = -\frac{\omega \varepsilon_s''}{c n_s}, \qquad \alpha_0 = \frac{\omega \varepsilon_m''}{c n_m} \tag{4-19}$$

where c is the speed of light in vacuum, ε_s'' and n_s are the imaginary part of dielectric constant and real part of refractive index in the gain medium (such as semiconductor), while ε_m'' and n_m are those in the lossy medium (such as metal). These modal quantities are typically related to the corresponding material quantities through the introduction of the so-called confinement factors as follows:

$$G_{\mathrm{mod}} = \Gamma_s G_0, \qquad \alpha_{\mathrm{mod}} = \Gamma_m \alpha_0 \tag{4-20}$$

Γ_s is the confinement factor define in the first paragraph and has been widely used. Γ_m is the loss confinement factor and analogous to gain confinement factor, and it takes into account possible different spatial distribution of material loss, such as metal loss in a MSM waveguide.

We first consider an arbitrary waveguide which is uniform along the propagation direction z. The active region has a uniform complex dielectric constant $\varepsilon = \varepsilon' + i\varepsilon''$. The real field variables are decomposed as follows:

$$\mathbf{E}(x,y,z) = A(z)\overline{\mathbf{E}}(x,y)e^{ik_z' z - i\omega t} + c.c., \quad \mathbf{H}(x,y,z) = A(z)\overline{\mathbf{H}}(x,y)e^{ik_z' z - i\omega t} + c.c. \tag{4-21}$$

where k_z' is the real part of the propagation wave vector, and c.c. stands for complex conjugate. $A(z)$ describes amplitude that changes along z direction; $\overline{\mathbf{E}}(x,y)$ and $\overline{\mathbf{H}}(x,y)$ are complex filed components depending only on the

transverse coordinates x and y. Following the procedure of Ref. [87], (also given in appendix A) one can obtain the gain (index, s) and loss (index, m) confinement factor as

$$\Gamma_{s,m} = \frac{n_b c\varepsilon_0 \iint_{s,m} |A(z)|^2 |\bar{\mathbf{E}}(x,y)|^2 \, dxdy}{\iint_{\infty} |A(z)|^2 \, \mathrm{Re}[\bar{\mathbf{E}}(x,y) \times \bar{\mathbf{H}}(x,y)^*] \bullet \hat{\mathbf{z}} dxdy} \tag{4-22}$$

where n_b is the real part of the background refractive index of the active material. This expression is a direct result from the definition of confinement factor in Eq. (4-20) and can be obtained by applying Poynting's theorem [87, 96]. Eq. (4-22) is generally valid and equivalent to those given in Ref. [94-97] and has been also reproduced recently [98-100]. For Γ_s, the integration in the numerator runs over the gain region, while for Γ_m it runs over the lossy region, e.g., the metal region in a MSM waveguide.

We perform integration in Eq. (4-22) in metal and semiconductor regions separately. The total power flux along the propagation direction, I, is divided into

$$I = \iint_m S_z^m dxdy + \iint_s S_z^s dxdy = \iint_m w^m v_{E,z}^m dxdy + \iint_s w^s v_{E,z}^s dxdy \tag{4-23}$$

where S_z, w and $v_{E,z}$ have already been given in the last section. Recall that the average energy velocity $\bar{v}_{E,z}$ was introduced by:

$$\bar{v}_{E,z} = \frac{\iint_m w^m v_{E,z}^m dxdy + \iint_s w^s v_{E,z}^s dxdy}{\iint_m w^m dxdy + \iint_s w^s dxdy} = \frac{I}{W} \tag{4-24}$$

where

$$W = \iint_m w^m dxdy + \iint_s w^s dxdy \tag{4-25}$$

is the total energy per unit length in z direction. Dividing the numerator and denominator of Eq. (4-22) by $4W$ simultaneously and using Eq. (4-24) (note that the denominator of Eq. (4-22) is equal to $2I$), one can rewrite equation (4-22) as

$$\Gamma_{s,m} = \frac{c}{\overline{v}_{E,z}} \frac{n_b \iint_{s,m} \frac{1}{4}\varepsilon_0 |A(z)|^2 |\overline{\mathbf{E}}(x,y)|^2 \, dxdy}{W/2} \tag{4-26}$$

The numerator of this expression contains information on total electric energy. Specifying Eq. (4-26) for the cases of semiconductor and metal respectively and using Eqs. (4-16) and (4-17), we obtain confinement factors for metal and for semiconductor as follows:

$$\Gamma_m = \frac{c/n_m}{\overline{v}_{E,z}} \frac{W_e^m}{W/2} \frac{n_m^2}{\frac{d(\omega\varepsilon_m')}{d\omega}} \tag{4-27a}$$

$$\Gamma_s = \frac{c/n_s}{\overline{v}_{E,z}} \frac{W_e^s}{W/2} \tag{4-27b}$$

where W_e^m and W_e^s are the total electric energy per unit length in metal and semiconductor respectively. Here we use the approximation: $\varepsilon_s' \approx n_s^2$, since $|\varepsilon_s''| \ll \varepsilon_s'$ is always true in the active region consisting of common semiconductors. The confinement factor for metal region contains an extra factor (the third factor in Eq. (4-27a)). This factor stems from the dispersiveness of metals and the large imaginary part of metal dielectric function. This factor reduces to 1 when both are negligible. In that case, Eqs. (4-27a) and (4-27b) are completely analogous. Note

that Eq. (4-27) gives general expressions for waveguides whose cross section contained metals, semiconductors, or multiple layers of other materials. In a lossless dielectric waveguide, the average energy velocity is the same as the group velocity. As a result, the first factor in Eq. (4-27) becomes the ratio n_g/n_b, or the ratio of group index to the phase (background) index and the second factor becomes the well-known energy ratio: the energy in the active region to the energy in the entire waveguide. The energy ratio (which is always smaller than 1) is often mistaken in the literature as a general expression for confinement factor. Eq. (4-27b) shows the energy ratio as a limiting case of the confinement factor when the energy velocity approaches the phase velocity.

4.3.2. Enhancement of confinement factors

In the following we consider a MSM sandwich waveguide with semiconductor in the middle. Fig. 4-11 shows the confinement factors in semiconductor (Γ_s) and metal layer (Γ_m) for a symmetric TM mode for two situations: no gain ($\varepsilon_s'' = 0$) and high gain ($\varepsilon_s'' = -0.4$). We can see from the figure that Γ_s and Γ_m increase with photon energy until reaching the SPP resonance and then decrease thereafter. Such enhancement near SPP resonance, especially when $\varepsilon_s'' = -0.4$, can be explained by the slow average energy propagation. Near the SPP resonance, the total energy flux of the entire waveguide is extremely

small because power travels in metal and semiconductor in opposite directions with almost the same magnitude. The total energy flux becomes vanishingly small if the material gain can exactly compensate the metal loss, as in the case where $\varepsilon_s'' \sim -0.4$. Such small total energy flux leads to a small average energy velocity from Eq. (4-27), resulting in huge confinement factors for both the semiconductor part and metal part. Since both Γ_s and Γ_m are inversely proportional to the average energy velocity (thus is G_{net}), it is possible to obtain a huge positive value of net modal gain near SPP resonance, as is shown in Fig. 4-12, where the modal gain from semiconductor, modal loss from metal, and the net modal gain are shown as a function of photon energy with $\varepsilon_s'' = -0.4$. We can see that the modal gain and loss have the same order of magnitude around SPP resonance. Due to the huge amplifying effect of slow energy velocity, any excess of gain over metal loss would lead to a giant net modal gain. This is the origin of giant model gain discussed in the last section.

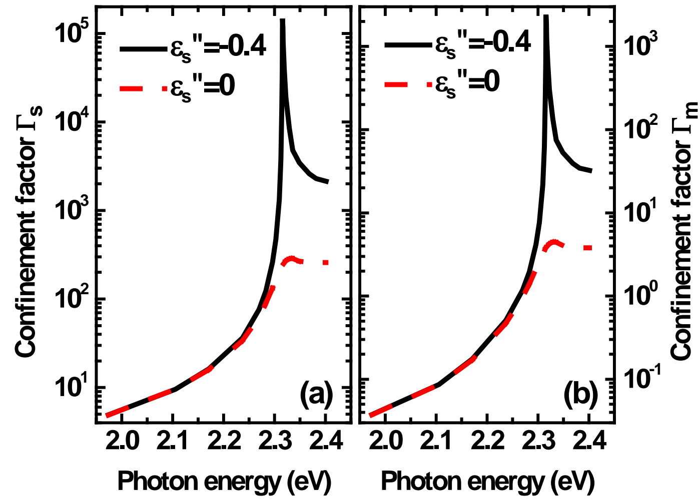

Fig. 4-11. Confinement factors in (a) semiconductor and (b) metal layer as a function of photon energy, for two different gain levels (different values of ε_s'') in a MSM waveguide. The thickness of semiconductor layer is 100 nm.

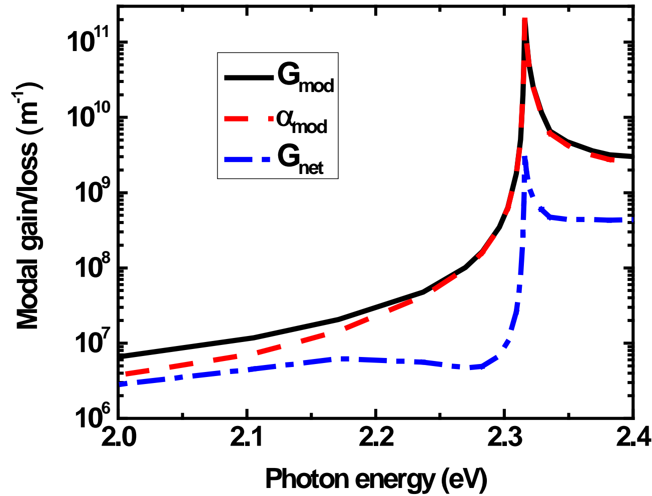

Fig. 4-12. Modal gain from semiconductor (solid line), modal loss from metal (dashed line) and net modal gain from Eq. (4-18) (dash-dot line) as a function of photon energy with $\varepsilon_s'' = -0.4$.

4.3.3. Gain dependence of confinement factors

It is interesting to note from Fig. 4-11(a) that Γ_s values at different ε_s'' agree well at low frequency (well-below SPP resonance), but deviate from each other near and above the SPP resonance. Similar behavior occurs for Γ_m as shown in Fig. 4-11(b). This means that the confinement factors depend on the levels of gain in the semiconductor near and above SPP resonance. In a conventional semiconductor laser, confinement factor as given by Eq. (4-26) is known as a pure geometric factor that only depends on the configuration and geometry of a waveguide, but not the material gain. To understand such a perplexing aspect of confinement factor near SPP resonance and the relation between confinement factor and ε_s''(material gain), we plot Γ_s as a function of ε_s'' at different photon energies in Fig. 4-13. Γ_s does not show ε_s''-dependence below 2.237 eV, consistent with the conventional understanding of the confinement factor. But a peak around $\varepsilon_s''=-0.4$ appears when frequency is increased further, reaching its maximum value at the SPP resonance frequency (top curve peaked at $\varepsilon_s''=-0.4$). Further above the SPP resonance, the peak decreases. Thus the confinement factor is strongly influenced by ε_s'' only near the SPP resonance. To understand this, we recast Eq. (4-26) for Γ_s into:

$$\Gamma_s = \frac{n_s}{k_0}\frac{|k_z|^2 Q_s}{k_z'\varepsilon_s' Q_x^s + k_z'\varepsilon_m' Q_x^m + k_z''\varepsilon_s'' Q_x^s + k_z''\varepsilon_m'' Q_x^m} \qquad \text{(4-28)}$$

where $Q_s = \int_S |\overline{\mathbf{E}}_s|^2 \, dx$, $Q_x^s = \int_S |\overline{E}_x^s|^2 \, dx$, $Q_x^m = \int_M |\overline{E}_x^m|^2 \, dx$; k_z'' is the imaginary part of the propagation wave vector k_z. Near SPP resonance ($\varepsilon_s' \approx -\varepsilon_m'$), the electric field intensity has almost the same profile in metal as in semiconductor, thus $Q_x^s \approx Q_x^m$. Therefore the sum of the first two terms in the denominator of Eq. (4-28) is extremely small so that the denominator is mainly determined by the last two terms. Since ε_m'' is positive, a small change in the negative value of ε_s'' will result in a large change of Γ_s. This explains the sensitive dependence of confinement factors near the SPP resonance. For the frequency far bellow SPP resonance, most of electric field is confined in the semiconductor, thus $Q_x^s \gg Q_x^m$. We always have $\varepsilon_s' \gg |\varepsilon_s''|$. As a result, the first term in the denominator is much larger than the other terms and dominates Eq. (4-28), Γ_s thus shows weak dependence on ε_s''. In the latter case, the metallic waveguide behaves like a dielectric waveguide where the sum of first two terms in the denominator of Eq. (4-28) is much larger than the last two terms. The analysis for Γ_m is the analogous since Γ_m can be written in a similar form to Eq. (4-28) with identical denominator. Note that the material gain is able to affect the confinement factor only in the structure where the real parts of the dielectric constants of two adjacent layers have almost the same magnitude but opposite sign. The metal-semiconductor waveguide near its SPP resonance is a perfect example.

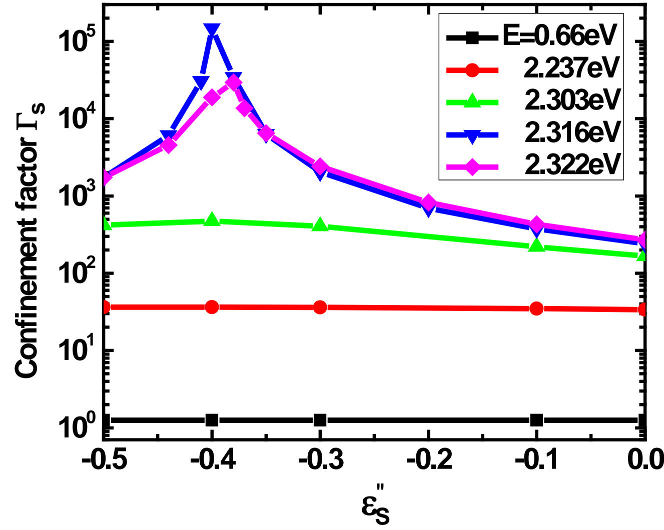

Fig. 4-13. Confinement factor Γ_s as a function of ε_s'' at different photon energies in the MSM waveguide (same as in Fig. 4-9).

4.4. Conclusion

We have investigated the interaction between photons and plasmons near the SPP resonance in a MSM plasmonic waveguide. We discovered a rather unusual phenomenon, i.e., the existence of a giant modal gain near the SPP resonance that is more than 1000 times the material gain of the semiconductor. We also identified the physical origin of this giant gain and attributed it to the dramatic slowing down of the energy velocity to as slow as 200 m/s. The concept of the confinement factor is re-examined and analyzed in detail in the case of a MSM waveguide. The familiar definition of the confinement factor is recast into an expression where its inverse-dependence on the average energy velocity is

explicit. This helps to explain the hugely enhanced confinement factor and thus the net modal gain by a slowing-down of average energy propagation near the SPP resonance. Such definition can be easily extended to a multilayer structure consisting of metal and semiconductor layers. We also find an unusual feature of the confinement factor: it depends on the material gain near the SPP resonance in a MSM waveguide, contrary to the pure geometric character in a dielectric waveguide. We show that such a unique feature originates from the SPP resonance in a plasmonic waveguide.

5. Metal-semiconductor-metal nano-laser

The gain and loss mechanisms in a MSM waveguide with the size limit only in the thickness direction have been studied in chapter 4. If such waveguide has finite width and length, it can work eventually as a laser cavity. In order to realize electrical injection SPP lasing in the MSM cavity at room temperature, we carried out a systematic study of various loss mechanisms and focused on the design and optimization of the structure of a plasmonic nanolaser with the possible smallest size.

5.1. Metal-semiconductor-metal nano-cavity

In this section, we studied a prototype of a MSM cavity as shown in Fig. 5-1(a) through design and simulation by using COMSOL Multiphysics. The cavity has length L in z direction, width W in y direction, and thickness h_s of semiconductor core layer and h_m of metal (silver) cladding layer in x direction. The semiconductor has a dielectric constant ε_s with a real part of 12, and the silver dielectric function $\varepsilon_m(\omega)$ is taken from [86] and curve fitted for our numerical solution. Such configuration can be easily extended to a manufacturable structure with electrical injection and be scaled down in three dimensions separately for optimization. In a realistic electrical injection device, very often an insulator layer is sandwiched between the metal and semiconductor

layers, which can be approximated by the MSM structure for the purpose of modal property study. Since we only study the possibility of SPP mode lasing in the cavity, the modes are always decaying in the x direction from the metal/semiconductor interface. By exciting the cavity with a transverse magnetic (TM)-polarized incident wave, one can obtain the intensity spectrum of the SPP modes. The cavity modes are denoted by TM_{0nm}. The first index, "0", means that there is no half-cycle variation of field in x direction (to avoid dielectric or optical modes), the second and the third indices stand for the number of half-cycle variations in y and z direction, respectively. Fig. 5-1(b) shows the intensity spectrum of a MSM cavity with W=L=50 nm and h_s=h_m=10 nm. The inset of Fig. 5-1(b) is the field pattern of TM_{012} mode in x-z plane.

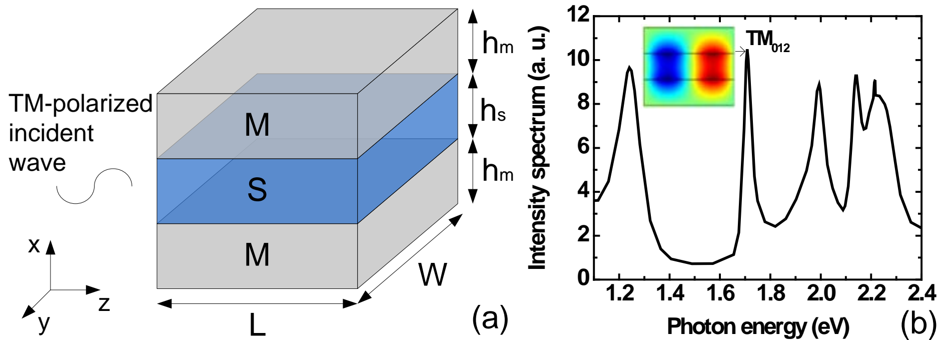

Fig. 5-1 (a) Schematic of a MSM cavity. (b) The intensity spectrum of a MSM cavity with W=L=50 nm and h_s=h_m=10 nm. The color inset is the field pattern of TM_{012} mode in x-z plane.

As the cavity being pumped, the semiconductor layer will provide optical

gain or material gain to the cavity and become the active region. The material gain spectrum in a semiconductor can be calculated by a microscopic or free-carrier theory and can be added to the imaginary part ε_s'' of the dielectric function, and is given by the following formula at frequency ω:

$$G_0 = -\frac{\varepsilon_s''\omega}{n_s c} \tag{5-1}$$

where n_s is the refractive index of semiconductor and c is the speed of light in vacuum. In the linear regime, the material gain will increase with pumping. When the modal gain (material gain × confinement factor) is equal to the overall loss of the cavity seen by a mode, which is the sum of modal (internal) loss and facet (transmission or radiation) loss, this mode reaches its lasing threshold, and the corresponding material gain is called threshold gain. In the linear theory of lasers, spectral width of a mode is proportional to the net gain or loss. Thus the spectral linewidth approaches zero at threshold. This fact can be used to obtain threshold numerically by following the spectral width with increasing gain in the semiconductor. The theory of this method is given in appendix B. Fig. 5-2 shows the linewidth and peak height of the intensity spectrum in a MSM cavity with W=L=30 nm and h_s=h_m=10 nm as a function of material gain. As material gain increases, the spectral linewidth decreases linearly which is predicted by the expression of γ in appendix B. When the material gain is large enough so that the

modal gain can compensate the total loss exactly, the linewidth goes to zero. The corresponding material gain is the threshold material gain which is around 12000 cm^{-1} in Fig. 5-2. If the material gain keeps increasing, the linewidth increases again since the simulation has been done in the linear region. In a real lasing system, the linewidth will keep nearly a constant at threshold due to the nonlinear saturation of material gain. The peak height of the intensity increases as material gain increases, and reaches its maximum value at the threshold. The peak height is thus another measure to determine the threshold gain.

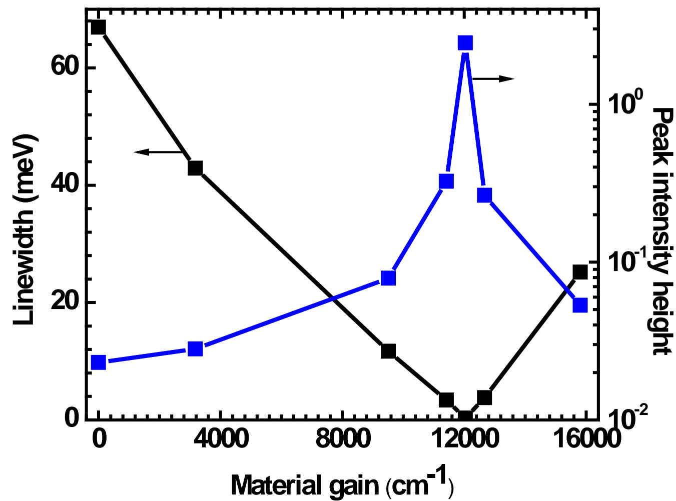

Fig. 5-2 Linewidth and peak height of the intensity spectrum in a MSM cavity with W=L=30 nm and h_s=h_m=10 nm as a function of material gain.

5.2. Parametric dependence of loss and gain

The study is focused on the loss mechanisms and their parametric dependence on nanolaser geomtries. Such study will allow people to determine

the smallest size of the laser cavity and the corresponding modes with threshold gain that is practically achievable. In a nanolaser cavity of a finite size, there are several loss processes: 1) The cavity transmission or radiation to outside; 2) The SPP modal loss; and 3) The so-called lossy surface waves as a result of pair-excitations in electron gas [101, 102]. The pair excitation loss leads to strong quenching of dipole emission in a semi-infinite medium near the surface of a semi-infinite metal. But since the pair excitation wave-vector band is around 10^9-10^{10} m^{-1} in Au or Ag, such quenching layer is extremely thin and can be easily avoided by adding a thin layer on the order of 1 nm between gain medium and metal. Since the SPP wave-vector is at least one order of magnitude smaller, strong interaction can still be maintained between metal and gain medium through the SPP modes. This argument has been demonstrated [101] in a similar situation where dipoles in a medium sandwiched by two metal layers dissipate over 99% of their power into the waveguide mode instead of lossy surface wave even if the middle medium layer is only 3 nm. Thus we will focus on the first two loss mechanisms, while assuming that the last loss mechanism can be avoided by adding a very thin layer. Later, we will use a quantum-well heterostructures design to include the thin wide gap layer in the design.

5.2.1. Mode analysis

Fig. 5-3 shows the intensity spectrum of MSM cavities with various W and L and h_s=h_m=10 nm. The position of a series of TM_{01m} modes is marked by the color arrows. We can see the intensity drops quickly when the photon energy is larger than 2.3 eV and diminishes after 2.4 eV. To get better understanding of the mode behaviour, we plot the real part of the propagation wavevector k_z of some high order TM_{01m} modes in the MSM cavities with various W and L in Fig. 5-4 by the square dots. The real k_z of the SPP mode in a MSM waveguide with infinite width and length and 10-nm-thick middle layer is also shown by the solid curve for comparison. The discrete k_z in the MSM cavity agrees well with the continuous k_z in the MSM waveguide, showing the unifying feature of SPP modes in a cavity with finite length and width and in a waveguide with infinite length and width. The peak in the dispersion curve of the waveguide corresponds to the position of SPP resonance which is about 2.33 eV. All modes in the MSM cavity are below the SPP resonance, no cavity mode exists at and above the SPP resonance, which explains the diminishing of excitation intensity in Fig. 5-3 when the photon energy is larger than 2.4 eV. Fig. 5-5 shows the intensity spectrum of MSM cavities with various h_s at W=L=50 nm and h_m=10 nm. The intensity tends to diminish when the photon energy exceeds the SPP resonance, similar to the phenomena in Fig. 5-3.

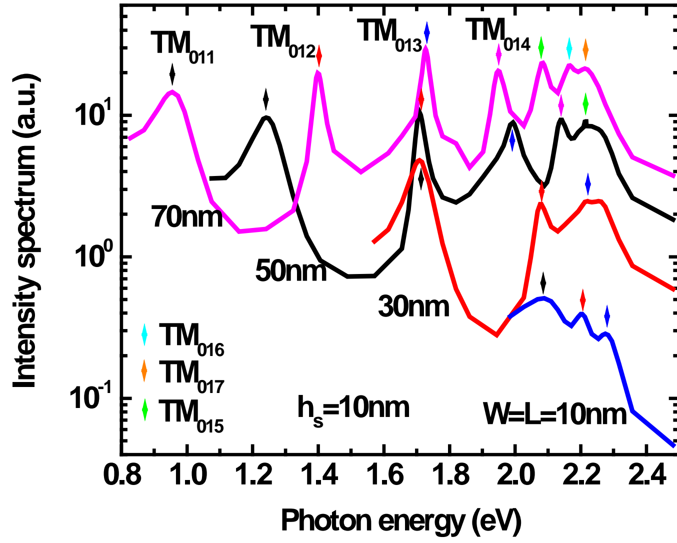

Fig. 5-3 Intensity spectrum of MSM cavities with various W and L and h_s=h_m=10 nm. The position of a series of TM_{01m} modes is marked by the color arrows.

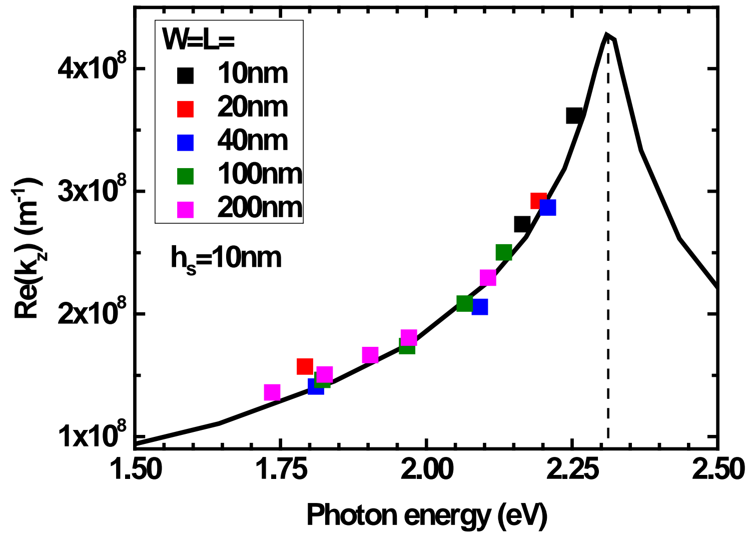

Fig. 5-4 Real part of the propagation wavevector k_z of some high order modes in the MSM cavities with various W and L and h_s=h_m=10 nm. The real k_z of the SPP mode in a MSM waveguide is shown by the solid curve as well.

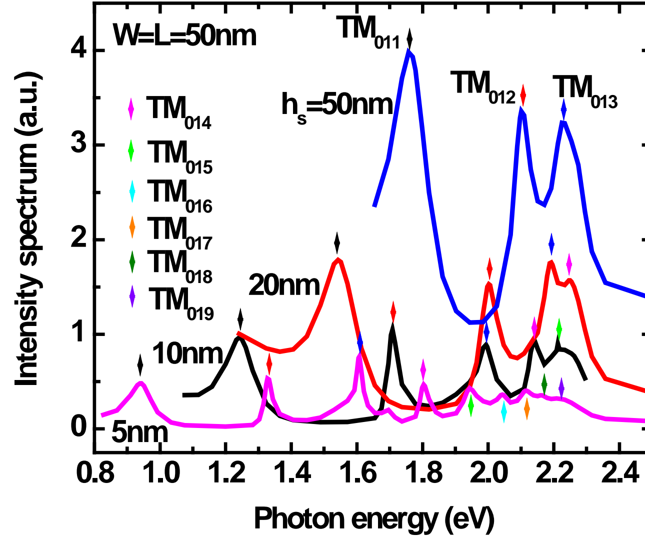

Fig. 5-5 Intensity spectrum of MSM cavities with various h_s at $W=L=50$ nm and $h_m=10$ nm. The position of a series of TM_{01m} modes is indicated by the color arrows.

5.2.2. Dependence on middle layer thickness

To study the effect of middle gain layer thickness, we first fix the in-plane size W and L both at 50 nm and the metal layer thickness at h_m=10nm, and vary the thickness of the core layer. The threshold gain for each cavity mode with different core layer thicknesses is shown as a function of photon energy ($=\hbar\omega$) in Fig. 5-6. First, we notice that in the cavity with h_s=10 nm TM_{02m} and TM_{03m} modes have higher threshold than TM_{01m} modes. It is true in general for various middle layer thickness that the threshold gain is higher for mode TM_{0nm} with larger index n. Therefore, we will discuss from now on TM_{01m} modes only.

Second, it is important to realize that for each mode group TM_{01m}, there is a mode index m with a given photon energy, for which the thresholds gain is minimum. That is because the threshold gain for lower order modes (smaller m) is dominated by the large facet loss while the threshold gain for higher order modes is determined by the large SPP modal loss. As photon energy increases, the facet loss decreases since higher order mode is more tightly confined inside the cavity while the SPP modal loss increases monotonically as the photon energy approaches the SPP resonance [55]. The threshold gain, proportional to the sum of facet loss and SPP modal loss, thus has a minimum value at an intermediate photon energy. Thirdly, we can see that the minimum threshold gain decreases with the core layer thickness which is mainly due to the reduction of facet loss, and the detail will be shown later. The red-shift of the modes with minimum threshold with decreasing core layer thickness further reduces the threshold gain since material gain is proportional to the frequency. The existence of a mode with the minimum threshold gain for a given core thickness and the decrease of such minimum gain with core thickness leads to the overall reduction of threshold gain to a level of about 4000 cm^{-1}, achievable in a typical semiconductor. We emphasize that the delicate interplay between SPP modal loss and facet transmission loss leads to the existence of a minimum-threshold mode. Note that for very small thickness, the pair-excitation loss will have to be included.

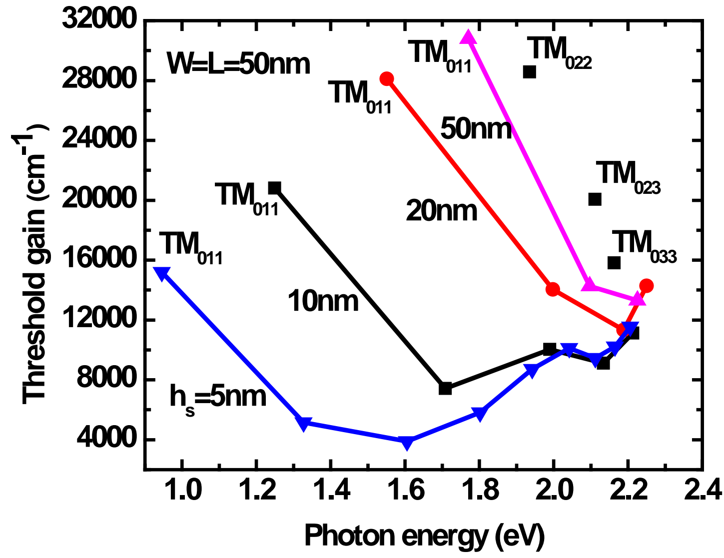

Fig. 5-6 Threshold gain of consecutive TM_{01m} (m=1, 2, 3, …) modes (symbols connected by lines) in the cavity with different core layer thickness, h_s. The threshold gain of all TM_{02m} and TM_{03m} modes in the cavity with h_s=10 nm is shown by the un-connected squares.

5.2.3. Dependence on in-plane sizes

We next fix the thickness of the core and metal layer both at 10 nm, and vary the in-plane size W and L by assuming W=L. Fig. 5-7 shows the threshold gain of consecutive modes with different in-plane sizes as a function of photon energy. The overall behaviour is similar to that in Fig. 5-6: There is a mode with minimum threshold gain for each given in-plane size. The photon energy with the minimum threshold is red-shifted as in-plane size increases and the minimum threshold gain decreases accordingly. Note that for the case with W=L=200 nm,

the TM_{015} mode (with a photon energy of 1.2 eV) has threshold gain around 2000 cm^{-1}, a low enough value achievable by many common semiconductors at room temperature.

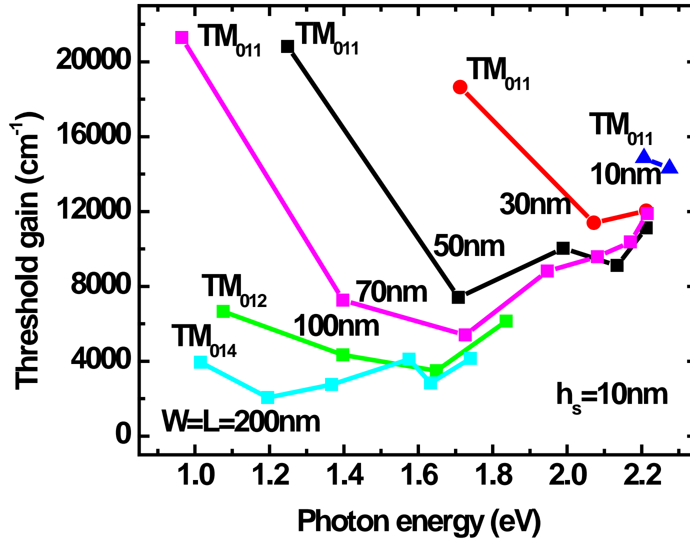

Fig. 5-7 Threshold gain of consecutive TM_{01m} (m=1, 2, 3, …) modes for different in-plane sizes, *W* and *L*.

By reorganizing the results from Fig. 5-6 and Fig. 5-7, we plot the minimum threshold gain as a function of (a) core layer thickness and (b) in-plane size in Fig. 5-8. We can see from Fig. 5-8(a) that the minimum threshold gain drops quickly if the core layer thickness is smaller than 30 nm and keeps decreasing even if the core size is as thin as 5 nm. The reason for the feature has been explained by the reduction of facet loss in the previous section. The facet loss, defined as the threshold modal gain subtracting the SPP modal loss, is shown in Fig. 5-8(a) for

the modes with minimum threshold at different core layer thicknesses. We can see the facet loss coincides with the threshold gain very well except the case with h_s=10 nm. In the cavity with h_s=10 nm, TM_{012} mode has the minimum threshold, while in other cavities with larger h_s TM_{013} mode has the minimum threshold. Although TM_{013} modes in some cavities with h_s>10 nm may have less facet loss than the TM_{012} mode at h_s=20 nm, their threshold is still higher due to their high mode frequency. Therefore, there is no low limit on core layer thickness to achieve minimum threshold gain for core thickness down to 5 nm. Note that for very small thickness, the pair-excitation loss will have to be included. The dashed curve in Fig. 5-8(a) shows the minimum threshold gain as a function of core layer thickness by including the influence of the dipole quenching near metal interface. We can see that the minimum threshold gain is larger than that without dipole quenching. The difference is less than 5% when core layer is thicker than 30 nm, but is over 70% when the core is 5nm thick. The minimum threshold gain reaches a smallest value at core thickness about 6 nm. Therefore the core thickness of the MSM structure has a lower limit due to the existence of the pair-excitation loss. In Fig. 5-8(b), the minimum threshold becomes practically constant for in-plane size larger than 200 nm. Although in-plane size over 200 nm×200 nm is on the same order of magnitude as reported [49], the advantage of scaling down of the core layer thickness enables significant reduction of the overall size of the cavity. The

corresponding cavity volume (in λ_0^3) for the cases in Fig. 5-8(b) is also shown in Fig. 5-8(b). Taking the cavity with 200 nm in in-plane size and 10 nm in core layer thickness as an example, the cavity volume is only 3.6×10^{-4} λ_0^3 and the required threshold gain is as low as 2000 cm^{-1}, achievable with many semiconductors.

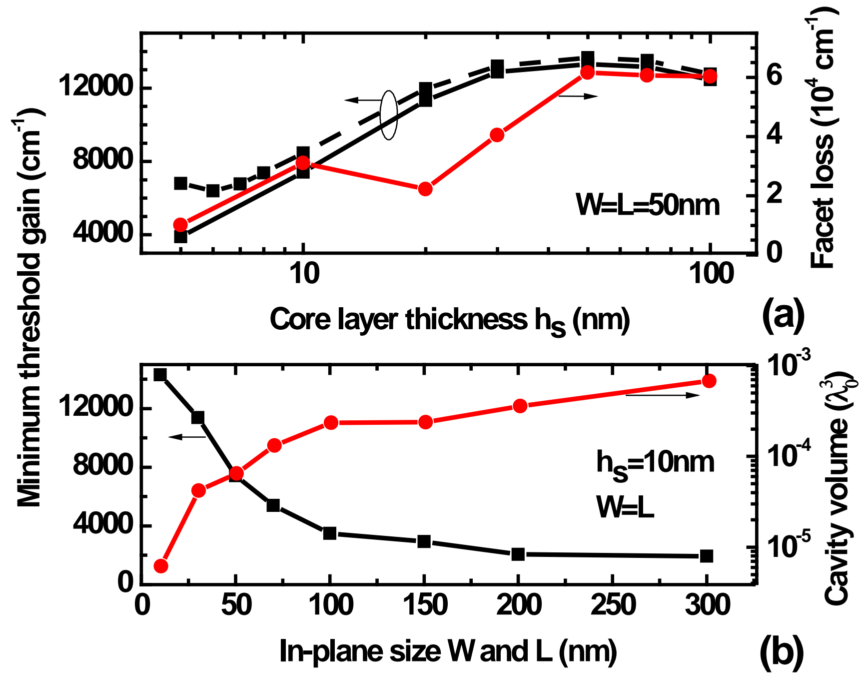

Fig. 5-8 The minimum threshold gain of the cavity with different (a) core layer thicknesses in the cases with (dashed line) and without (solid line) dipole quenching and (b) in-plane sizes. The facet loss for the cases in Fig. 5-8(a) is shown in Fig. 5-8(a), and the cavity volume (in λ_0^3) for the cases in Fig. 5-8(b) is shown in Fig. 5-8(b) as well.

It is also interesting to see how the loss and gain will change if the thickness and in-plane sizes of the core region is reduced at the same time. Here we assume the core region of the cavity is a cubic box with $W=L=h_s$ and the metal layer has a

fixed thickness of 10 nm. Fig. 5-9(a) shows the intensity spectra for cavities with various sizes of core region. The three lowest order modes: TM_{011}, TM_{012} and TM_{013} mode are marked in the figure. All modes are blue-shifted as core size decreases, and the frequency for the highest order mode in each curve is close to 2.33eV, which is the SPP resonance frequency in a MSM waveguide with the same thickness. The threshold gain of each mode is shown as a function of core size in Fig. 5-9(b). The threshold gain for the fundamental TM_{011} mode (black points) increases dramatically as core size increases while the threshold gain for TM_{012} (red points) and TM_{013} (blue points) mode does not change significantly as core size is varied. The reason is that the fundamental modes with low frequency are loosely confined inside the cavity, thus suffer large facet loss which dominates the threshold gain; for the fundamental modes with high frequency and higher order modes, the threshold gain is dominated by the SPP modal loss and the facet loss can be ignored. In the case of Fig. 5-9(a), the facet loss at low frequency is much larger than the SPP modal loss at high frequency, hence the threshold gain for the fundamental modes with low frequency is much larger than that of other modes. We can see that the threshold gain for all modes is larger than 12000 cm^{-1} which is almost unachievable in semiconductors. Therefore, a cubic core design is not feasible for semiconductor nanolaser.

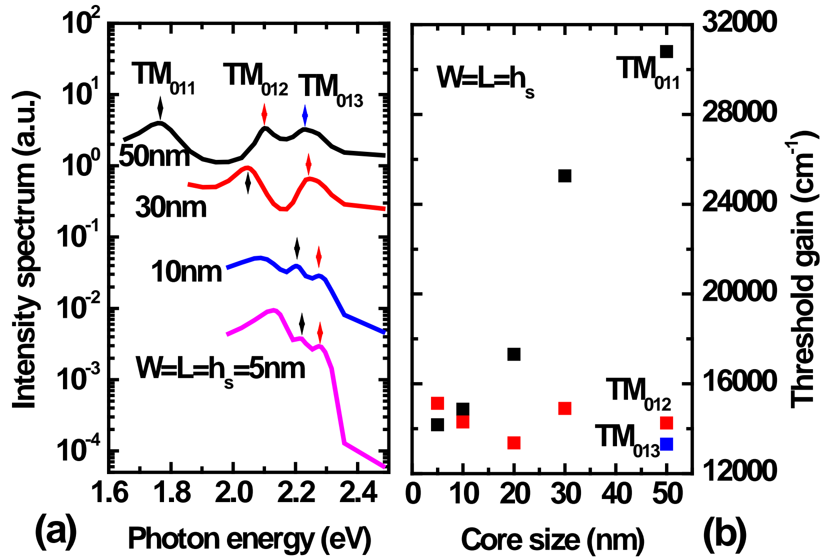

Fig. 5-9 (a) Intensity spectra for cubic core cavities with various size of core region. The three lowest order modes are marked by color arrows. (b) Threshold gain of each mode in cavities with different core size.

5.2.4. Dependence on metal layer thickness

The metal cladding layers are usually designed to be thick enough to prevent mode leakage from the core. However, as we will show later, using thin metal layer is crucial in reducing lasing threshold gain. Therefore it is necessary to study the influence of the thickness of metal layer. Fig. 5-10(a) shows energy profile in the x-z plane for three different values of metal layer thickness. We can see that, as metal layer thickness increases from 5 to 20 nm, less energy leaks through the metal claddings and most of the light comes out from the side of the cavity in z direction. The percentage of the energy leaked through the metal cladding as a

function of metal layer thickness is plotted in Fig. 5-10(b). We can see that the leakage decreases exponentially as metal layer thickness increases linearly, and less than 0.1% of the energy leaks through the metal cladding if the thickness of metal layer is 15 nm and thicker. Fig. 5-10(c) shows the relation between the minimum threshold gain and metal layer thickness. The minimum threshold gain increases rapidly from h_m=5 nm to 15 nm and keeps nearly constant as h_m>20 nm. The reduction of the minimum threshold gain with decreased metal thickness is partly due to the decreased metal loss but mainly due to the red shift of the mode with minimum threshold (TM_{012} mode). The red-shift of the mode is probably caused by the coupling between the mode inside the core and that outside the metal. Fig. 5-11 shows the influence of metal layer thickness in the cavity with W=L=100 nm. All features in the figure are similar to those in Fig. 5-10. We should note that the threshold gain can be reduced to around 2500 cm^{-1}, as shown in Fig. 5-11(a), which is realizable in common semiconductors.

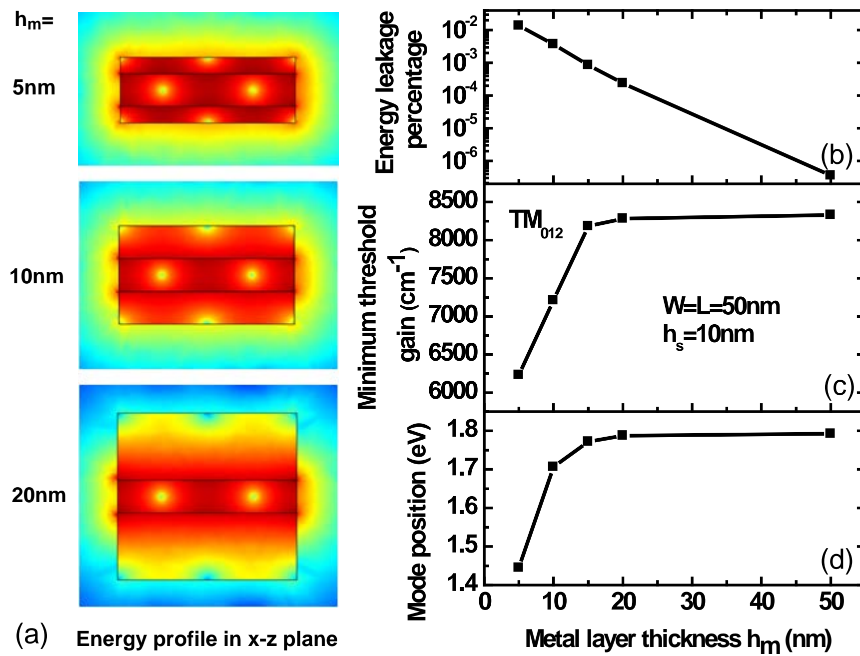

Fig. 5-10 (a) Energy profiles in the x-z plane with metal layer thicknesses: 5 nm, 10 nm, and 20 nm, for W=L=50 nm and h_s=10 nm. (b) The percentage of the energy leaked through the metal cladding layers out of the cavity, (c) the minimum threshold gain, and (d) the position of the mode with minimum threshold gain as a function of metal layer thickness.

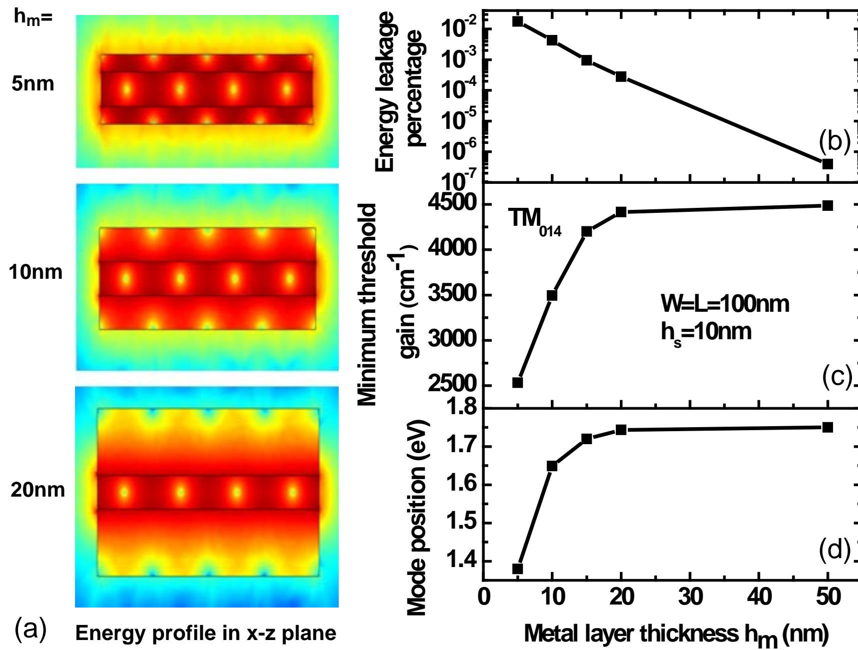

Fig. 5-11 (a) Energy profiles in the x-z plane with metal layer thicknesses: 5 nm, 10 nm, and 20 nm, for W=L=100 nm and h_s=10 nm. (b) The percentage of the energy leaked through the metal cladding layers out of the cavity, (c) the minimum threshold gain, and (d) the position of the mode with minimum threshold gain as a function of metal layer thickness.

5.2.5. Effects of insulating layer

In an actual metal-semiconductor nanolaser design, an insulating layer is usually inserted between metal cladding and semiconductor active region (thus forming the MISIM structure) to prevent charge transfer from active region to metal layer. Such insulating layer can be SiO_2 or Si_3N_4 [45, 46, 49] or wide-gap semiconductors [50, 103]. It was reported that [49, 104] there is an optimal thickness of the insulating layer, with which the laser has the smallest threshold

gain. The conclusion is correct only for TE-like modes which have very weak electric field component vertical to the insulating layer or metal layer, the low-index insulating layer can push the field towards the high-index core and away from the metal cladding, thus reducing the metal confinement factor [43, 56] and the metal modal loss. However, for TM-like modes such as SPP mode, there is a strong electric field component vertical to the insulating layer or metal layer, so that a slot mode [105] will be formed with a strong field intensity inside the insulating layer due to the low index of the insulating layer and the continuity of the displacement vector. The slot mode will mix with SPP mode, resulting in the strongest mode intensity located at the metal-insulating layer interface, as shown in the inset of Fig. 5-12(a), the field thus cannot be pushed away from the metal cladding.

If the dielectric constant of the insulating layer, ε_d , is too small compared to ε_s , as is the case for many insulators, the electric field in the insulating layer will be much larger than that in the core, resulting in small mode volume in the active core region. The weak field intensity inside the active layer will lead to small confinement factor and increased threshold material gain. Therefore, it is preferable to use materials with dielectric constant close to that of the semiconductor core to enhance the modal gain. An ideal choice is a quantum well structure with a narrow gap sandwiched between two layers of wide gap

semiconductor. Very often the index difference between well and barrier materials is relatively small and the quantum well structure provide both the required index profile and mechanism for electrical injection and confinement. Assume $\varepsilon_d = 10$, Fig. 5-12(a) shows the minimum threshold gain as a function of insulating layer thickness with a fixed total thickness h_s+h_d=20 nm. We can see the thinner the insulating layer is, the smaller the minimum threshold gain will be. Therefore the insulating layer should be made as thin as possible for SPP mode lasing. Once the thickness of the insulating layer is determined, the thickness of the semiconductor core is the only parameter that needs to be optimized. Remember we showed in Fig. 5-8(a) that the minimum threshold gain decreases as core thickness is reduced if the total thickness of the layers between metal layers is smaller than 30 nm. On the other hand, with fixed thickness of the insulating layer, the decrease of the active core thickness as the total size between metals decreases will lead to the increase of the threshold gain. Therefore, there will be an optimal value for the thickness of the active core at which the threshold gain is minimized. Fig. 5-12(b) shows the minimum threshold gain as a function of core layer thickness with h_d=3 and 5 nm, respectively. As expected, an optimal value of the core thickness (8 and 13 nm, respectively) exists for each case, depending on the thickness of insulating layer.

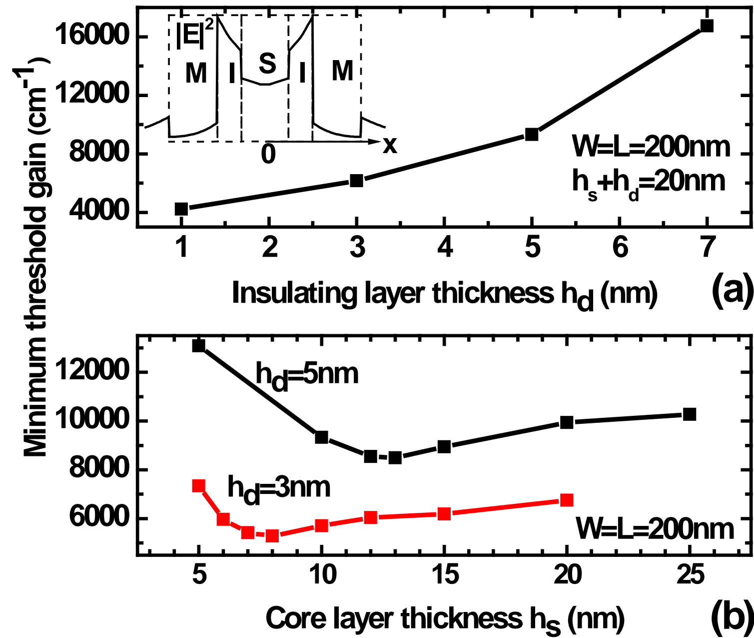

Fig. 5-12 (a) Minimum threshold gain as a function of insulating layer thickness with a fixed total thickness of insulating and core layers of 20 nm. The $|E|^2$ profile of a slot mode mixed with SPP mode along x axis in a MISIM cavity is shown in the inset. (b) Minimum threshold gain as a function of core layer thickness with two insulating layer thicknesses: 3 and 5 nm.

5.3. Toward the smallest semiconductor laser

Using the design and optimization procedures for the MSM cavity discussed above, we now consider a specific example: an $Al_{0.2}Ga_{0.8}As$/GaAs/$Al_{0.2}Ga_{0.8}As$ single quantum well structure sandwiched between two Ag layers, as shown in Fig. 5-13(a). The thickness of $Al_{0.2}Ga_{0.8}As$, GaAs and Ag layers is optimized to be 3 nm, 10 nm and 15 nm, respectively, and the length and width of the cavity is 200 nm. For electrical injection operation, the whole cavity is usually coated by silicon nitride (SiN_x) in device fabrication for electric isolation and mechanical

support [45, 46, 54]. We use the dielectric constant of silicon nitride in near infrared range (~4) as the ambient dielectric constant. The dielectric constant of $Al_xGa_{1-x}As$ and GaAs can be found in Ref. [106]. Since the material gain is available within a limited frequency range above the bandgap of active semiconductor, we only study the modes within the gain bandwidth of AlGaAs/GaAs single quantum well (from 1.4 to 1.65 eV according to Ref. [107]). The simulation results showed that two modes: TM_{015} and TM_{016} exist in this range, as shown by the intensity spectrum in Fig. 5-13(b). The peak wavelength of TM_{015} mode is λ_0=846 nm (1.467 eV) and the lasing threshold gain is 3646 cm^{-1}, achievable at room temperature [107]. Since the plasmonic modes have two strong electric field components in x and z direction, respectively, material gain due to the in-plane (TE) and off-plane (TM) dipole transitions in the quantum well contributes to the mode. The quality factor for this mode without material gain is about 56. TM_{016} mode is at λ_0=769 nm (1.613 eV) whose quality factor and lasing threshold gain is 72 and 3427 cm^{-1}. Although the threshold gain is slightly smaller than that of TM_{015} mode, it can hardly be achieved since the TM material gain provided to TM_{016} mode is from the transition between the first excited state in conduction band and the first excited state in light-hole band, which is too small to satisfy the requirement of the threshold gain [107]. The cavity hence works as a single-mode operation laser. The physical volume of

AlGaAs/GaAs/AlGaAs core is only 6.4×10^{-4} μm^3 (~0.001 λ_0^3), and the total volume of the structure, including metal layers, is 1.84×10^{-3} μm^3 (~0.03 λ_0^3).

Fig. 5-13(c) shows the near field energy density pattern of TM_{015} mode at lasing threshold in the x-z plane. We can see most energy is confined in the cavity, as shown by the deep red color. The far-field pattern, shown in Fig. 5-13(d), is an important characteristic of any laser. As can be seen, the full-width at half-maximum of the far-field angle is about 88°. If the sides of the structure in Fig. 5-13(a) are left with air, instead of coated by SiN, the facet loss will be reduced due to the larger index contrast. According to the simulation, the TM_{015} mode is then blue shifted to 839 nm with a threshold gain of 3721 cm^{-1}, even larger than that shown previously, since material gain is also proportional to modal frequency according to Eq. (5-1).

The size of the nanolaser can be potentially shrunk further. Assume reducing the size of the structure in Fig. 5-13(a) by half in all three dimensions and keeping the thickness of metal layers unchanged to provide good cladding, then the thickness of each layer (from top to bottom) will be 15 nm, 1.5 nm, 5 nm, 1.5 nm, and 15 nm with a lateral size W×L=100 nm×100 nm. Using $Al_{0.3}Ga_{0.7}As$ as the insulating layer instead of $Al_{0.2}Ga_{0.8}As$, the laser structure is shown in Fig. 5-14(a). The simulation showed that the TM_{014} mode exists within the bandwidth of the quantum well with a wavelength of 813 nm and a threshold gain of 3859 cm^{-1}, as

shown in Fig. 5-14(b). Such threshold gain is achievable in the quantum well at room temperature [108]. The physical volume of the cavity is only $1.5\times10^{-4}\,\lambda_0^3$, which is the smallest nano-laser reported so far. The near field energy pattern and far field radiation pattern is shown in Fig. 5-14(c) and 5-14(d), respectively. The full-width at half-maximum of the far-field angle is around 93°, similar to that of the larger structure in Fig. 5-13.

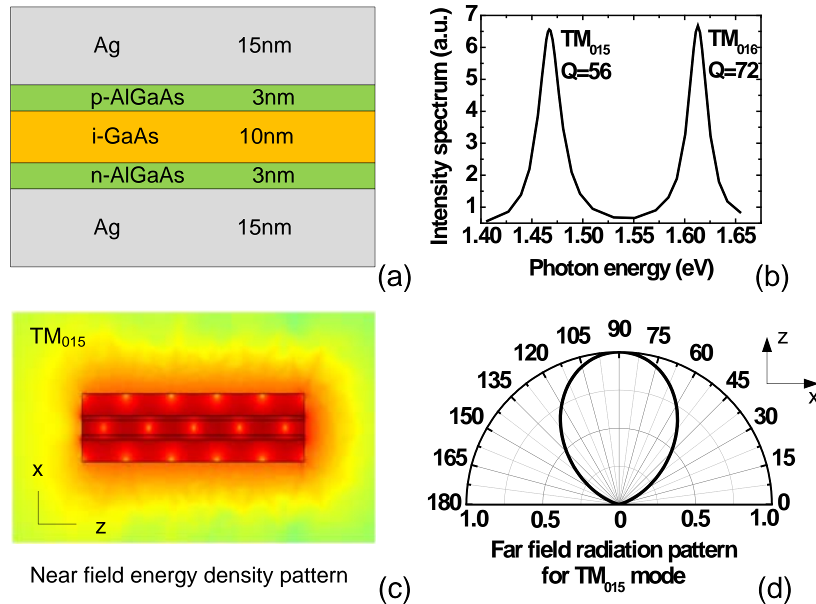

Fig. 5-13 (a) Schematic structure of an optimized MISIM nanolaser. (b) Intensity spectrum of the nanolaser within the gain bandwidth of the AlGaAs/GaAs/AlGaAs quantum well, showing two possible modes. (c) Near field energy density pattern of the TM_{015} at the lasing threshold in the *x-z* plane. (d) Angular dependence of the far field $|E|^2$ radiation pattern of the TM_{015} mode at the lasing threshold in the *z-x* plane.

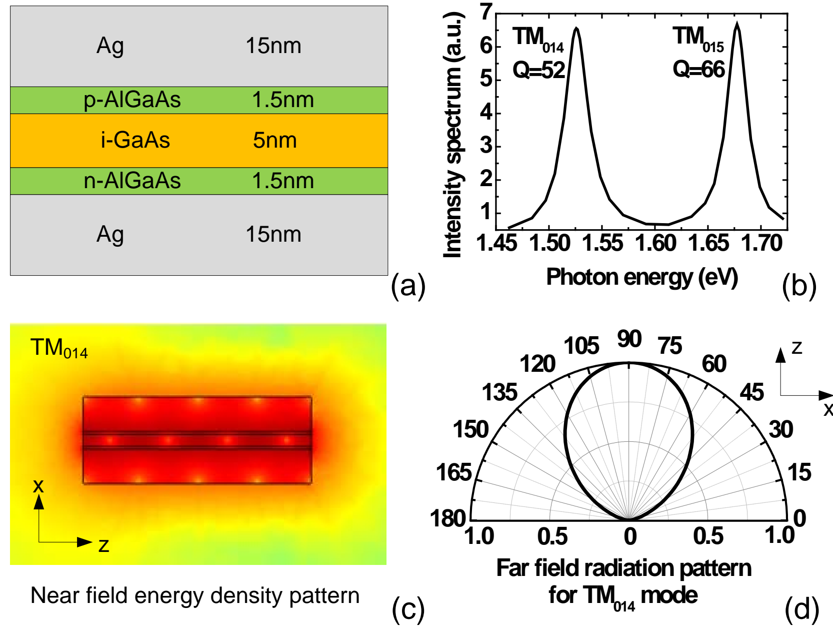

Fig. 5-14 (a) Schematic structure of an optimized MISIM nanolaser. (b) Intensity spectrum of the nanolaser within the gain bandwidth of the AlGaAs/GaAs/AlGaAs quantum well, showing two possible modes. (c) Near field energy density pattern of the TM_{014} at the lasing threshold in the *x*-*z* plane. (d) Angular dependence of the far field $|E|^2$ radiation pattern of the TM_{014} mode at the lasing threshold in the *z*-*x* plane.

Note that many of the quantitative results shown above depend on the values of metal loss used for Ag. The values in [86] which were used for the simulation are known to have overestimated the metal loss, most likely because they were obtained with metals deposited with technologies decades ago. According to the more recent dielectric function data [109] for Ag, the threshold gain could be smaller since the Ag material loss (given by Eq. (5-1) with positive sign) is 5% to 10% smaller than that calculated by using Ag dielectric function data from Ref. [86] within the wavelength range studied here. Therefore, the threshold gain

required here is likely larger and thus the design and feasibility study represent a more conservative upper limit. We believe that with the improvement of deposition techniques of metals, metal loss can be further reduced, increasing the likelihood of observing the SPP mode lasing studied here.

5.4. Conclusion

We studied a prototype of a SPP nanolaser with a MSM rectangular cavity, where the cavity modes are pure SPP modes. To find the smallest cavity with practically achievable lasing threshold gain, we performed a systematic study of the dependence of the threshold gain on structure parameters. We found that the threshold gain decreases continuously when shrinking the thickness of the semiconductor core from 30 nm to 5 nm and maintains a minimum value if the length and width is over 200 nm. The actual smallest core thickness is limited by the dipole quenching occurring in the gain material within 2 nm from the metal interface due to pair-excitations. The threshold gain can be reduced by reducing the thickness of the metal cladding layer from 20 nm to 10 nm, still maintaining negligible leakage through the metal layers. We also showed that the existence of an insulating layer between the semiconductor core and the metal layer always raises the threshold due to the characteristics of the mixing of the slot mode and the SPP mode, but the threshold gain can be minimized by adjusting the core layer

thickness if the insulating layer thickness is fixed. We finally proposed a more realistic cavity consisting of an $Al_{0.3}Ga_{0.7}As/GaAs/Al_{0.3}Ga_{0.7}As$ single quantum well sandwiched by Ag layers with optimized parameters. Single mode lasing in a SPP mode can be achieved in the cavity at room temperature, and the physical volume of the cavity, including insulating and semiconductor core layers, is only $1.5\times10^{-4}\ \lambda_0^3$.

6. All-semiconductor plasmonics

We investigated the metallic plasmonic structures in the previous two chapters. As discussed in the introduction, metals have their own drawbacks such as large intrinsic loss and poor layer quality in fabrication. The concept of all-semiconductor plasmonics thus becomes important in utilizing plasmonic devices, especially for the on-chip integration purpose. In this chapter, we demonstrated an all-semiconductor plasmonics by using InAs heterostructures which offer the shortest plasmon wavelengths among common semiconductors with the wide tunability. The plasmonic properties of InAs-based heterostructures are superior to those of metals with a much reduced optical loss and a much improved modal confinement or field enhancement. We also proposed and studied a prototype of a monolithic all-semiconductor active plasmonic system including the source, active waveguide, and detector, all realizable on a chip in a single epitaxial growth process [101].

6.1. Interband and intraband transition in a semiconductor

For a semiconductor with high electron density, it is important to include both interband and intraband transitions, especially for narrow gap semiconductors. The relationship between electric displacement vector **D** and electric field **E** can be written as:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}_b + \mathbf{P}_e \tag{6-1}$$

where ε_0 is the permittivity in vacuum, $\mathbf{P}_b$ and $\mathbf{P}_e$ are the polarizations for the interband transitions and intraband transitions (or free electrons), respectively. The free electron polarization is related to the electric field E through the linear susceptibility, χ_e, according to $\mathbf{P}_e = \varepsilon_0 \chi_e \mathbf{E}$. χ_e can be approximated by the Drude model:

$$\chi_e(\omega) = -\frac{Ne^2}{\varepsilon_0 m(\omega^2 + i\omega/\tau)} = -\frac{\varepsilon_\infty \omega_p^2}{\omega^2 + i\omega/\tau} \tag{6-2}$$

where N is the free electron concentration, e is the electron charge unit, m is the effective mass of electrons in semiconductor and τ is the free electron relaxation time. The plasmon frequency ω_p is given by

$$\omega_p = \sqrt{\frac{Ne^2}{\varepsilon_\infty \varepsilon_0 m}} \tag{6-3}$$

where ε_∞ is the semiconductor dielectric constant for $\omega >> \omega_p$. The dielectric function corresponding to the interband transitions is given by $\varepsilon_b = 1 + \chi_b$, which can be calculated using a semi-empirical expression for undoped or lightly doped semiconductors [111]. This model used a multi-transition best fit to the experimentally measured dielectric function to get ε_b. Semiconductors with small effective mass in conduction band, such as InSb or InAs, are easy to become degenerate at high electron density. In that case the absorption cannot happen near band edge since the bottom of conduction band is fully or nearly

fully occupied by electrons, as was first noticed by Burstein [112]. The absorption coefficient, α, can be given by taking into account of the occupation probabilities of electrons and holes, f_e and f_h, respectively:

$$\alpha = \alpha_0[(1-f_e)(1-f_h) - f_e f_h] = \alpha_0(1-f_e-f_h) \tag{6-4}$$

where α_0 is the absorption coefficient in undoped or lightly doped semiconductor. The absorption coefficient is related to the imaginary part of semiconductor dielectric constant by $\varepsilon_b'' = n_b c\alpha / \omega$, where n_b is the background refractive index and c is the speed of light in vacuum. Finally the total dielectric constant is given as $\varepsilon_r = \varepsilon_b + \chi_e$.

6.2. Plasmonic features of InAs heterostructures

Carrier densities that can be introduced in common semiconductors are orders of magnitude lower than in metals, so that plasmon resonances in doped semiconductors are typically in wavelength range longer than 10 microns. One of the key tasks for semiconductor plasmonics is to identify semiconductors and related heterostructures that can be highly doped or injected, such that the plasmon resonances can be in the MIR wavelength range. To compare the relative merits of common semiconductors, we plot in Fig. 6-1(a) the corresponding plasmon wavelengths ($\lambda_p = 2\pi c / \omega_p$) as a function of electron density. As we can see, InAs has the shortest plasmon wavelength for a given electron density. Even

though all semiconductors can in principle be doped for plasmonic applications, the required doping levels are typically too high to be practical, especially for short wavelength applications. From Eq. (6-3), it is clear that InAs is the best candidate to provide the largest plasmon frequency, since InAs is among few semiconductors with the smallest electron mass. Figs. 6-1(b) and 6-1(c) show respectively the real and imaginary part of the dielectric constant ε_r of InAs from 1 to 10 μm for different electron concentrations. Heavily doped InAs becomes metallic, as indicated by the negative real part of ε_r in the entire wavelength range shown. Higher electron density makes the metallic property more pronounced. Another important advantage of InAs is that it can be grown lattice matched with AlSb and GaSb [113] to form several interesting heterostructures, as shown in Fig. 6-1(d), where a "broken-gap" InAs/GaSb structure and a deep AlSb/InAs/AlSb quantum well are shown. Both of these structures allow high electron density to be introduced in the InAs layer at equilibrium without intentional doping [114, 115]. The broken-gap lineup is a spatial "semimetal" structure with conduction band edge (in InAs) lower than the valence band edge (in GaSb), similar to the common semimetals in k-space. Electrons and holes thus are separated spatially, avoiding major recombination in InAs.

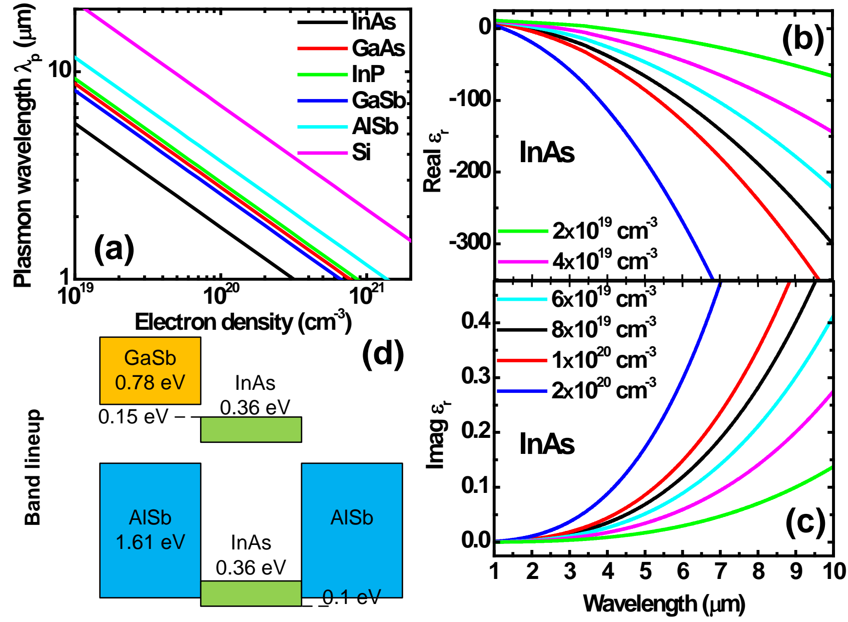

Fig. 6-1. Plasmonic features of InAs. (a) Plasmon wavelength vs. electron density for several common semiconductors. (b) Real and (c) imaginary part of InAs dielectric constant vs. wavelength for 6 different electron concentrations. (d) Band lineup of a GaSb/InAs (top) and a AlSb/InAs/AlSb (bottom) heterostructure.

To achieve high electron density, electrical bias has to be applied. Here we consider a p-GaSb/n-InAs/p-GaSb three layer structure as an example, as shown in the inset of Fig. 6-2. We performed a two-dimensional (2D) simulation of this structure using ATLAS assuming InAs layer is uniformly doped at 6×10^{17} cm^{-3} and GaSb layers are doped at 5×10^{18} cm^{-3}. Forward bias is applied on both GaSb layers. The electron density at the edge and in the middle of InAs layer as a function of bias is shown in Fig. 6-2 by black and red curves, respectively. Without bias, the electron density at the edge of InAs layer is 8.5×10^{17} cm^{-3} and

that in the middle of InAs layer is 7.7×10^{17} cm^{-3}. Even though higher doping densities can be introduced in each of these layers, these density levels are still too low. The electron density throughout the InAs layer is uniform but not large enough for plasmonic applications. As bias increases, the density at the edge increases faster than that in the middle of InAs layer, resulting in non-uniform distribution in InAs layer. However, since the strongest intensity of a surface mode is usually located at the InAs/GaSb interface, the electron density at the edge of InAs layer is more important for plasmonic application. We can see that the electron density at the edge reaches 1×10^{20} cm^{-3} at a bias of 4 V, a large enough value for our plasmonic application.

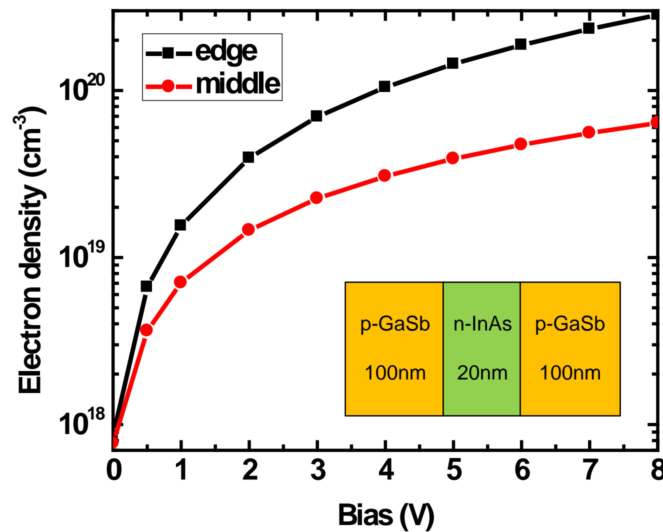

Fig. 6-2. Electron density as a function of bias in a GaSb/InAs/GaSb structure (inset). The black and red curves show respectively the electron density at the edge and in the middle of InAs layer. The thickness of each layer in the GaSb/InAs/GaSb structure is shown in the inset.

It is interesting to see what will happen if only the doping density of InAs layer is increased. Fig. 6-3(a) shows the electron density profile vertical to all layers in the GaSb/InAs/GaSb structure shown in Fig. 6-2 with different InAs doping densities. Doping density in GaSb layers is 8×10^{18} cm^{-3} and the bias voltage is 5V. We can see that the improvement of electron concentration is just in the middle of InAs layer, and is not significant even the doping density is as high as 8×10^{18} cm^{-3}. The electron density at the edge of InAs layer is independent on the InAs doping levels. On the other hand, the overall electron density in the InAs will increase as bias increases, as shown in Fig. 6-3(b) where InAs and GaSb layers are doped at 6×10^{17} cm^{-3} and 8×10^{18} cm^{-3}, respectively. Therefore, bias is more critical to achieve high electron density.

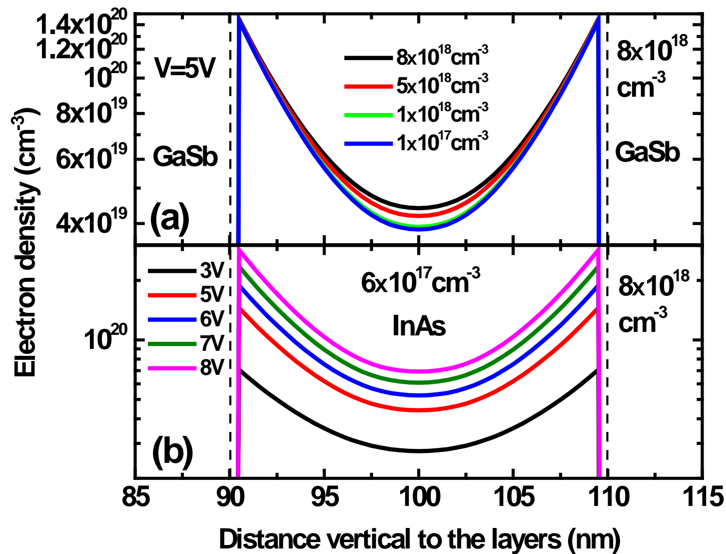

Fig. 6-3. Electron density profile vertical to all layers in a GaSb/InAs/GaSb structure with different (a) InAs doping densities and (b) anode biases.

To demonstrate the advantages of plasmonic features of InAs structures, we consider a prototype SPP structure: a bi-layer structure with a heavily doped InAs layer (dielectric constant ε_m) as a metallic layer interfaced with a GaSb layer with a dielectric constant ε_s. Assume the SPP wave propagates along the InAs/GaSb interface in the z direction, the propagation wavevector k_z is given by

$$k_z = \frac{\omega}{c}\sqrt{\frac{\varepsilon_s \varepsilon_m}{\varepsilon_s + \varepsilon_m}} \qquad (6\text{-}5)$$

The real and imaginary k_z for different electron concentrations in InAs as a function of wavelength are plotted in Fig. 6-4. The peaks of real k_z in Fig. 6-4(a) stand for the SPP resonance, coinciding with the peaks of imaginary k_z in Fig. 6-4(b). Since the loss that a waveguide mode experiences is equal to twice imaginary k_z, large value of imaginary k_z means large loss from the waveguide. We can see from Fig. 6-4(b) that the loss in the waveguide with low InAs electron density is larger than that in the waveguide with high InAs electron density at the wavelength larger than the SPP resonance wavelength. Therefore, large electron concentration in InAs is helpful to reduce the waveguide loss.

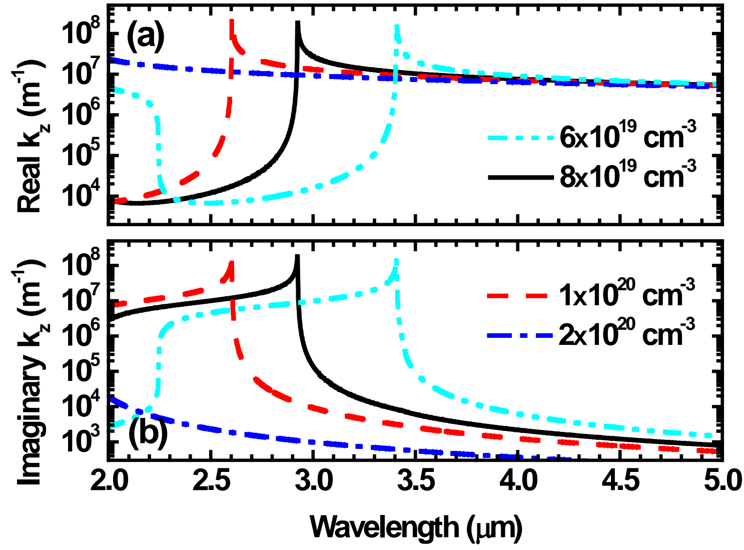

Fig. 6-4. (a) Real and (b) imaginary part of propagation wavevector k_z vs. wavelength of a SPP mode in an InAs/GaSb bi-layer structure for different electron concentrations of InAs layer.

An important factor for comparing the quality between waveguides is the quality factor (Q factor). For a planar waveguide, the Q factor is obtained via the definition in the following formula [117]:

$$Q = \omega \frac{W_{stored}}{P_{dissipated}} = \frac{\omega}{2k_z'' \overline{v}_{E,z}} \quad (6\text{-}6)$$

where k_z'' is the imaginary part of the propagation wavevector k_z and $\overline{v}_{E,z}$ is the average energy velocity of the mode. Note that W_{stored} and $P_{dissipated}$ in Eq. (6-6) have the units of J·m^{-2} and J·m^{-2}·s^{-1}, respectively. Figure 6-5 shows the Q factors as a function of wavelength of a SPP mode in an InAs/GaSb structure with different electron concentrations, in an Ag/GaSb structure, and in an Au/GaSb

structure. The dip (minimum) in each curve corresponds to the SPP resonance. We can see that the Q-factor of InAs/GaSb structure increases with electron density on the longer wavelength side of plasmon resonance. The situation becomes more complicated below plasmon resonance wavelength. One important result is that the Q factor of InAs/GaSb structure is several times larger than that of Ag/GaSb and Au/GaSb structures. Therefore, highly doped InAs structure is an excellent plasmonic waveguide and is superior to typical metal-based structures.

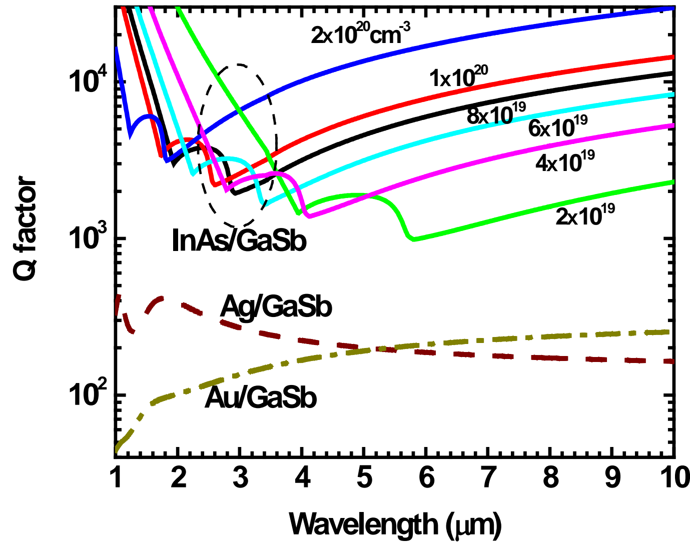

Fig. 6-5. Quality factors vs. wavelength of a SPP mode in an InAs/GaSb structure with different electron concentrations (cm^{-3}), in an Ag/GaSb structure, and in an Au/GaSb structure.

It is interesting to see how tightly a mode can be bound at the InAs/GaSb interface, since tight confinement of a mode is important for nanophotonic applications. Fig. 6-6 shows normalized energy density and power flux (absolute

value) profile across the interface of the InAs/GaSb structure for different electron concentrations in InAs at 3 and 5 μm, respectively. We can see from Figs. 6-6(a) and 6-6(b) that the energy and power densities for InAs/GaSb at electron density 6×10^{19} cm^{-3} are almost flat in both InAs and GaSb layers. This is because the working wavelength is shorter than the SPP resonance wavelength so that SPP mode is very loosely bound at the interface. The situation is the same for the case with density 2×10^{19} cm^{-3} in Figs. 6-6(c) and 6-6(d). The operating wavelength thus has to be longer than the SPP wavelength for better SPP localization. In order to quantitatively describe the mode confinement at the interface, we introduce effective widths (W_t) across the interface as the sum of the width in InAs layer, W_m, and that in GaSb layer, W_s (see Figs. 6-6(a) and 6-6(b)). They are defined as the distance from the interface through which the energy or power decays to e^{-2} of their values at the interface. The three widths measured at the working wavelength of 3 μm are listed in table 6-1. The effective widths for InAs/GaSb waveguide at density 6×10^{19} cm^{-3} are extremely long, corresponding to the flat curve in Figs. 6-6(a) and 6-6(b). For larger electron density, the effective widths are small. The total effective width at density 8×10^{19} cm^{-3} is the smallest one, because the working wavelength is very close to but still longer than the SPP resonance wavelength where the SPP mode has the best confinement. The total effective width at 73 nm is one fortieth of the working wavelength in this case, representing

a huge compression of the effective wavelength. There are similar conclusions for the case with working wavelength at 5 μm, with the required density somewhat smaller. Since the Q factor at the SPP resonance is the smallest (see Fig. 6-5), the electron density in InAs needs to be well controlled to meet the requirement of both quality and confinement in a waveguide design.

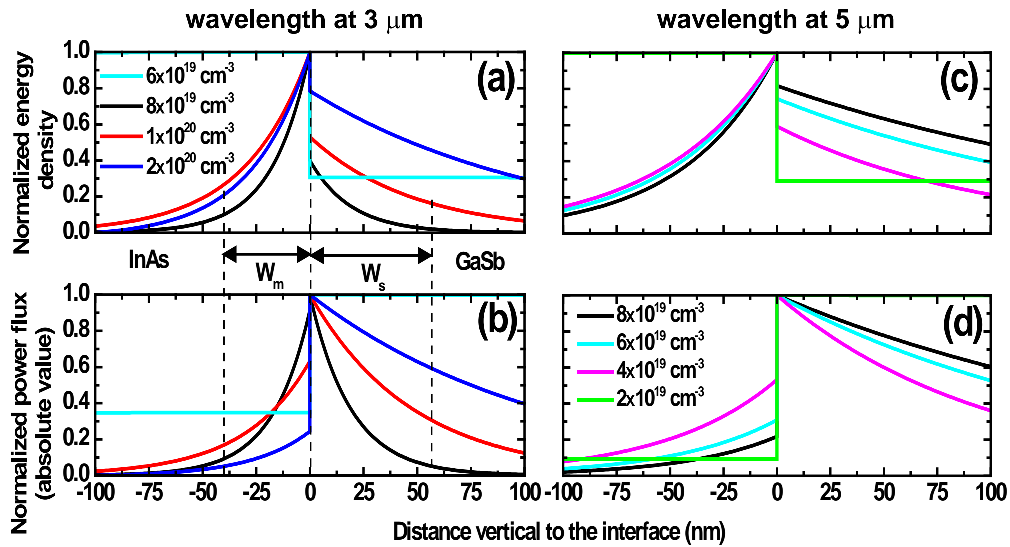

Fig. 6-6. (a) Normalized energy density and (b) power flux (absolute value) profile across the interface of an InAs/GaSb structure at 3 μm. The effective widths in InAs layer, W_m, and in GaSb layer, W_s, are schematically shown as well. (c) Normalized energy density and (d) power flux (absolute value) profile at 5 μm.

Table 6-1. Effective Widths in InAs/GaSb Structure at 3 μm.

Electron density (cm^{-3})	6×10^{19}	8×10^{19}	1×10^{20}	2×10^{20}
W_s (nm)	$>10^4$	38	95	218
W_m (nm)	$>10^4$	35	60	54
W_t (nm)	$>2\times10^4$	73	155	272

6.3. All-semiconductor active plasmonic system

The most interesting aspect of an all-semiconductor plasmonic structure is the possibility of monolithically growing an entire plasmonic system with various components on a chip. To illustrate this exciting aspect, we proposed and conducted a design study of an all-semiconductor plasmonic system consisting of a SPP source, a waveguide with integrated amplifiers, and a detector, as shown in Fig. 6-7(a). All three components can be grown in a single epitaxial growth process and each component can be then defined lithographically. The high electron density (on the order of 10^{20} cm^{-3}) in InAs can be achieved by electrical bias of the doped structures. An active structure is needed to provide gain in the source and amplifier. In addition to InAs and GaSb, AlSb can be used as a barrier material. The layer structures of all three components are identical and shown in Fig. 6-7(b) where the material composition, doping types, thickness and polarity of electrodes of each layer are also shown. The active region consists of p-AlSb/i-InAs/n-AlSb triple layers, with the intrinsic InAs layer serving as the

gain layer. This structure has been shown to be able to produce strong interband gain despite being a type-II structure for the undoped system [116]. The plasmonic structure consists of p-GaSb/n-InAs/p-GaSb triple layers. SPP mode is formed and propagating at the two n-InAs/p-GaSb interfaces, as we studied above in section 6.2. The overall working principle of this system is illustrated in Fig. 6-7(c).

Source: The bias V_1 is positive in the source, electrons and holes are thus injected into the i-InAs layer. Optical dipoles formed by the electron-hole pairs in this layer then recombine to excite and transfer energy preferentially to the SPP mode at the two p-GaSb/n-InAs interfaces [101]. The SPPs are thus generated in the source. If a time-varying signal voltage $V_S(t)$ is added to V_1, the amplitude of the SPPs will be modulated according to the strength of the signal.

Amplifier: The amplifier functions similar to the source. The bias V_1 has the same positive constant value V_0 as that in the source. When passing through the amplifier, the SPPs from source will acquire energy emitted from the electron-hole pairs in the i-InAs layer and get amplified.

Detector: The bias V_1 in the detector is negative, so the i-InAs layer is depleted. The incoming SPPs are then absorbed by InAs layers of the detector. The portion of the energy absorbed by the i-InAs initiates the generation of the electrons and holes in this layer which can be collected through the electrodes of the reversely

biased p-i-n structure. Since the density of generated carriers is proportional to the amplitude of the arrived SPPs, the signal from the source is demodulated and detected by the detector. Note that the gain spectrum provided by the active region has a peak frequency larger than InAs bandgap, as shown earlier [116]. Therefore, most of the energy of the SPP mode can be absorbed by the InAs layer in the detector.

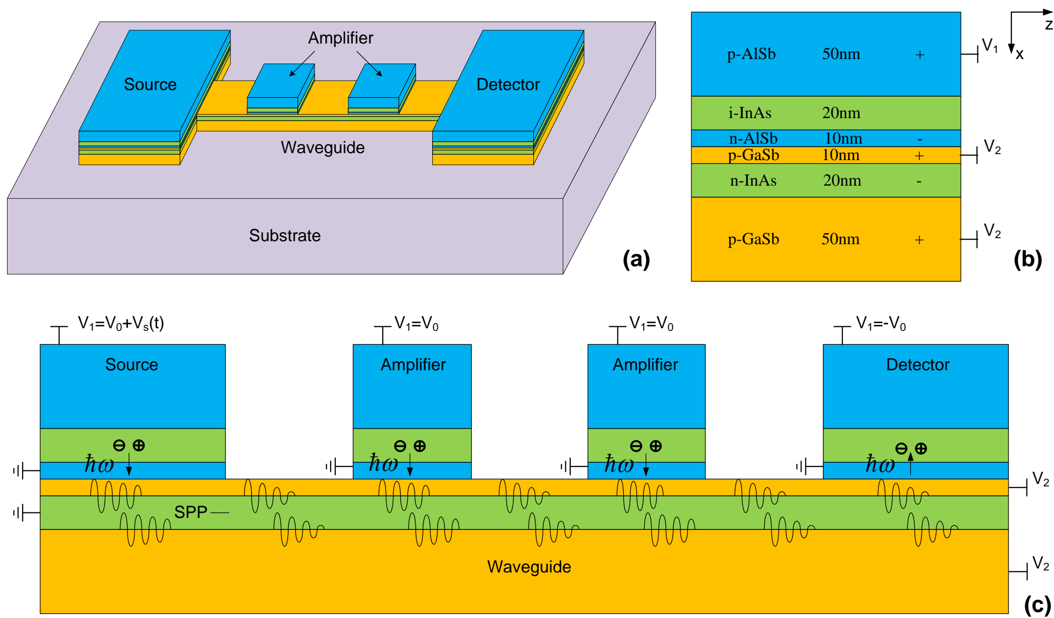

Fig. 6-7. All-semiconductor active plasmonic system. (a) Schematic of the proposed plasmonic system consisting of a SPP source, a waveguide with amplifiers and a detector. (b) The layer structures of all three components of this system. The material composition, doping type, thickness and polarity of electrodes of each layer are also shown. (c) Illustration of the working principle of the system.

We performed a 2D simulation of this system using ATLAS assuming the

layers in the active structure are uniformly doped at 2.5×10^{18} cm^{-3} and the layers in the plasmonic structure are doped at 5×10^{18} cm^{-3}. The electron and hole densities were calculated at V_1=0.8 V (source and amplifier) and V_2=5 V, and their profiles are shown in Fig. 6-8. The metallic InAs layer is located from x=90 to 110 nm. At the edge of this layer, the maximum electron density is about 1.5×10^{20} cm^{-3} while the hole density is lower than 1×10^{14} cm^{-3}. This situation is exactly what we need for the InAs application as a metal. In the middle of the metallic InAs layer (x=100 nm), the minimum electron density is about 4.2×10^{19} cm^{-3}, ensuring the maximum real part of the dielectric constant of InAs is still negative in the layer. Therefore this InAs layer can be treated as a pure metallic layer. Note that the hole density at the edge of GaSb layer is higher than the electron density at the edge of InAs layer, the metallic property of GaSb due to hole absorption thus needs to be considered. The collective motion of high-density holes can also be approximated by Drude model and the treatment of dielectric function of GaSb is similar to that shown in section 6.1 for InAs. The InAs gain layer is located from x=50 to 70 nm. The carrier density on the right half of this layer is not uniform because holes are preferentially swept from the middle AlSb layer into the GaSb layer on the right instead into the InAs layer on the left, leading to low hole density on the right half of InAs gain layer. The carrier density is, however, high and uniform enough on the left half of this InAs

layer, which can still provide high material gain to the whole system.

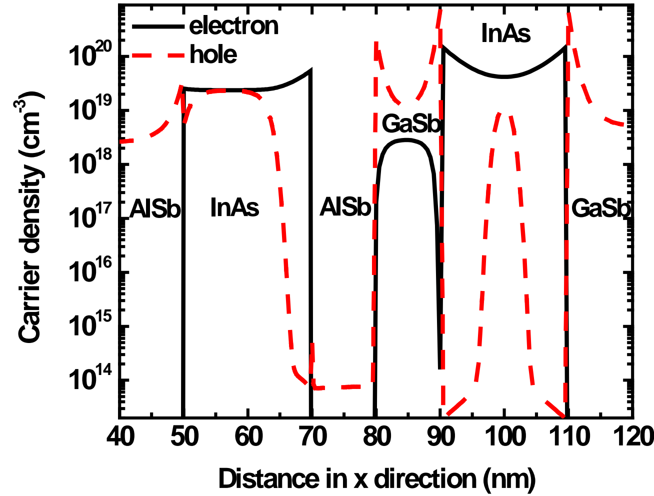

Fig. 6-8. Electron and hole density profiles along *x* axis at V_1=0.8 V and V_2=5 V in the plasmonic system.

The non-uniform dielectric constant profile in the metallic InAs layer and its adjacent GaSb layers was calculated by using the actual electron and hole density profiles from the simulation. The results then were used in a 2D electromagnetic mode analysis by COMSOL to find the SPP modes in the structure with a length 500 nm in *z* direction. Note that only the SPP mode is analyzed in this work since its highly-localized feature ensures the design of a system whose transverse size (the direction vertical to the wave propagation) can be on nanoscale. Other modes like TE mode, however, will spread out everywhere in the system consisting of layers with thickness of only a few tens of nanometers. The material gain in the InAs gain layer can be estimated from [116]. The power flux pattern of a SPP mode at an operating wavelength of 3.3 μm is shown in Fig. 6-9(a). The mode

propagates from left to right and increases exponentially due to the gain provided by the active structure. The normalized power flux (absolute value) profile taken at the location of the dashed line in Fig. 6-9(a) is shown in Fig. 6-9(b). We can see that most of the power of the structure travels outside the metallic InAs layer but is tightly bound at the InAs/GaSb interfaces. The effective width across the left-hand-side layers of the metallic InAs layer is about 26 nm. The propagation length ($=1/(2k_z'')$) of the SPP mode in the waveguide without amplifier is below 10 μm. However, since the amplifier provides the SPPs with a net modal gain ($=-2k_z''$) of 26.3 cm^{-1}, the SPPs can propagate over hundreds of microns if multiple-amplifier configuration is used. This structure is thus suitable for the nanoscale SPP waveguiding at 3.3 μm. The propagation length and the layout of the amplifiers required depend on the operating wavelength. To provide a gain at longer wavelengths, InAsSb alloy can be also used so that the operating wavelength can be somewhat further above the plasmon resonance wavelength of the InAs plasmonic structure.

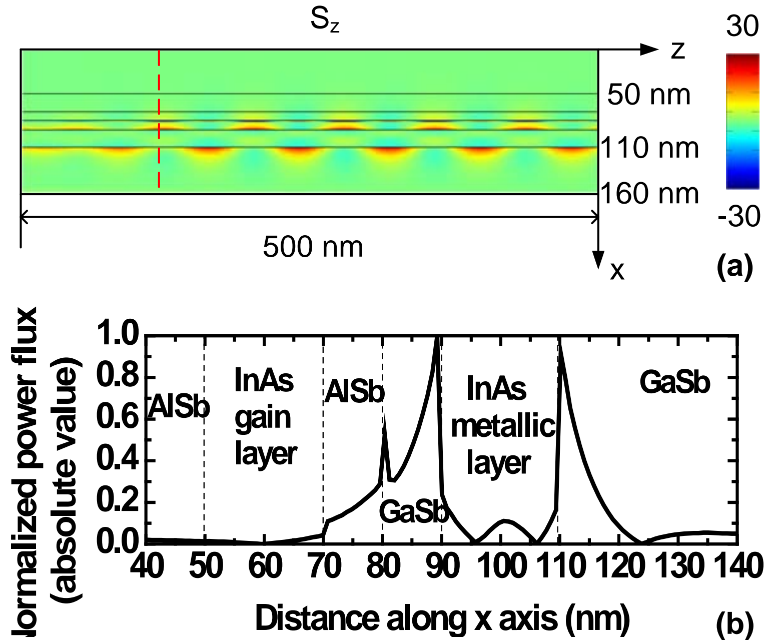

Fig. 6-9. Modal pattern of a SPP mode in the plasmonic system. (a) Power flux pattern of a SPP mode at 3.3 μm in the plasmonic waveguide with a length 500 nm in z direction. (b) Normalized power flux (absolute value) profile taken from the dashed line in Fig. 6-8(a).

6.4. Conclusion

We have investigated the plasmonic properties of InAs-heterostructures by using the Drude model. The result showed that InAs-heterostructures are superior to other common semiconductors and metals for applications in plasmonic structures and metamaterials at MIR wavelengths. The SPP modal properties have been studied in detail using an InAs/GaSb bi-layer structure for different electron concentrations in the InAs layer. Finally, we proposed and studied a prototype system of a monolithic all-semiconductor integrated plasmonic system on a chip, realizable in a single epitaxial growth process.

7. Summary

A theoretical study and numerical modelling and simulation on nano-scale semiconductor optoelectronics has been performed by programming in Matlab or using commercial software: Silvaco ATLAS and COMSOL Multiphysics. Silvaco ATLAS has been used in the study of electrical properties in semiconductor devices while COMSOL Multiphysics has been used in the study of optical and electromagnetic issues. Design and optimization of nanophotonic devices whose size (at least in one dimension) is smaller than the diffraction limit of the operating wavelength have also been carried out. The research is focused on electrically driven NW lasers, active plasmonic waveguides, plasmonic nano-lasers, and all-semiconductor plasmonic systems. The key results and potential impacts on the society of technology are summarized as follows.

1) The longitudinal p-i-n and coaxial p-n core-shell NWs were comparatively studied. The influence of the bias, doping density, and band alignment on the electrical injection characteristics was systematically investigated. It was shown that the injection efficiency in the core-shell NW can be higher than that in the p-i-n NW, that is, the carrier density in the core region is higher than that in the i region under the same bias. Furthermore, the injected carrier density in the core-shell NW can be maintained at a high level over a wide doping range due to

the fully or nearly fully depleted core region. Finally, type-II band alignment is also allowable for the core-shell NW to obtain a high level injection, while only type-I band profile is feasible in the longitudinal p-i-n NW. Such fundamental understanding is important for developing electrically pumped NW lasers which is still a challenging task for the community.

2) A new device structure used as an electrically pumped single nanowire laser was proposed. This structure takes the advantage of the core-shell configuration at p and n contact regions to improve electrical injection efficiency into the NW. Taking CdS and CdSe NWs as examples, we investigated the injection efficiency and carrier concentration uniformity in the wires for various NW lengths and core-shell overlapping lengths by carrying out a 3D device simulation. We then performed the design and optimization on the electrically driven laser based on a series of HE_{11p} (p=1, 2, 3, …) modes. We showed a design methodology to obtain a good trade-off between the threshold gain and the threshold voltage by tuning the NW length and overlapping length. The laser performance was further studied by solving rate equations. It is shown in an optimized structure that the threshold current is less than 80 μA at room temperature and the output power can exceed 30 μW if the injected current is greater than 1 mA for both CdS and CdSe NW lasers. The design also shows the promising application as an electrically pumped broad wavelength-tunable laser

utilizing in continuous composition grading CdSSe ternary alloy NWs. Such device structure is easy to implement any NW laser under the electrical pumping.

3) The propagation of SPPs near the SPP resonance in a MSM plasmonic waveguide which was pumped with an optical gain was investigated. We discovered a rather unusual phenomenon, i.e., the existence of a giant modal gain near the SPP resonance that is more than 1000 times the material gain of the semiconductor. Since the real part of the propagation wavevector is maximum (or, the effective wavelength is minimum) near the SPP resonance, devices operating near the SPP resonance can be made the smallest in size. The discovery of this giant gain will make it possible to achieve both a positive gain and the smallest size simultaneously, necessary features of nanophotonic devices. We also identified the physical origin of this giant gain and attributed it to the dramatic slowing down of the energy velocity to as slow as 200 m/s. This realization will contribute significantly to our rapidly increasing understanding of SPP phenomena in active plasmonics and will likely to stimulate experiments to search for such giant gain. As a first experiment, we suggest a MS waveguide with gratings coupling for the incident light to excite SPPs. In a MS structure, pumping light can be incident from the semiconductor side and the SPPs can be excited through spontaneous emission. In the reality with a long propagation length, the nonlinear saturation could be also important both inside the semiconductor and

inside the metal. All these issues need to be examined. In addition, amplified or long propagating SPPs will enable many experiments and will greatly impact many fields of physics such as condensed matter, classical and quantum optics, and metamaterials. Finally, such unprecedented modal gain is also expected to have a great impact to many plasmonic devices like plasmonic waveguides for future on-chip interconnects, subwavelength sources, and many other truly nanophotonic devices.

4) The concept of the confinement factor in the case of a MSM waveguide was re-examined and analyzed in detail. The familiar definition of the confinement factor was recast into an expression where its inverse-dependence on the average energy velocity is explicit. This helps to explain the hugely enhanced confinement factors and thus the net modal gain by a slowing-down of average energy propagation near the SPP resonance. Such definition can be easily extended to a multilayer structure consisting of metal and semiconductor layers. We also found an unusual feature of the confinement factor: it depends on the material gain near the SPP resonance in a MSM waveguide, contrary to the pure geometric character in a dielectric waveguide. We showed that such a unique feature originates from the SPP resonance in a plasmonic waveguide. The understanding of this unique feature of plasmonic waveguides will help people understand and develop nanophotonic devices based on plasmonic effects.

5) A prototype of a SPP nanolaser with a MSM rectangular cavity was studied, where the cavity modes are pure SPP modes. To find the smallest cavity with practically achievable lasing threshold gain, we performed a systematic study of the dependence of the threshold gain on structure parameters. We found that there is an interesting interplay between facet loss and SPP modal loss and that such interplay leads to the existence of a minimum-threshold mode. Either shrinking the thickness of the semiconductor core or increasing the length and width of the core leads to the reduction of threshold gain, which makes the SPP lasing threshold achievable for many semiconductors at room temperature. However, the actual smallest core thickness is limited by the dipole quenching occurring in the gain material within 2 nm from the metal interface due to pair-excitations in the metal. In addition, we found that the threshold can be further reduced by using thinner metal cladding without much exacerbating the mode leakage. This is partly due to the decreased metal loss but mainly due to the red shift of the minimum-threshold mode. We also showed that the existence of an insulating layer between the semiconductor core and the metal layer always increases the threshold due to the characteristics of the mixing between the slot mode and the SPP mode, but the threshold gain can be minimized by adjusting the thickness of the core layer if the thickness of the insulating layer is fixed. We finally proposed a more realistic cavity consisting of an AlGaAs/GaAs/AlGaAs

single quantum well sandwiched by Ag layers with optimized parameters. Single mode lasing in a SPP mode can be achieved in the cavity at room temperature, and the physical volume of the cavity, including insulating and semiconductor core layers, is only about 0.001 λ_0^3. We also pointed out that such laser can be potentially shrunk by half in all dimensions, leading to a physical volume of only $1.5\times10^{-4}\ \lambda_0^3$, which is the smallest nano-laser reported so far.

6) The plasmonic properties of InAs-heterostructures by using the Drude model were investigated. It was shown that InAs-heterostructures are superior to other common semiconductors and metals for applications in plasmonic structures and metamaterials at MIR wavelengths. The SPP modal properties were studied in detail using an InAs/GaSb bi-layer structure for different electron concentrations in the InAs layer. Finally, we proposed and studied a prototype system of a monolithic all-semiconductor integrated plasmonic system on a chip, realizable in a single epitaxial growth process. The significance and impact of the proposed plasmonic heterostructures and integrated system can be appreciated in several ways. First the InAs-based structures fill an important bandgap window between 2 and 8 microns for plasmonic applications in MIR range. While common metals can be used at MIR wavelengths as good mirrors, they cannot be used as plasmonic structures, since there is almost no localization near the metal surfaces at these long wavelengths. The available semiconductors such as GaSb and AlSb

that are lattice matched to InAs can work as barrier materials or allow easy design of the active source, amplifying waveguide, and detector, all based on the same substrate. Thus an entire plasmonic system can be grown in a single epi-process. This is a significant advantage compared to plasmonic or metamaterial structures that use a combination of metals and semiconductors, since high quality semiconductors have to be grown epitaxially, while metals are typically deposited in a thermal process with much poorer crystal quality. The all-epitaxially grown integrated plasmonic system in a single process can offer unprecedented quality and reliability. The low loss (or high quality factor) of InAs is another important advantage. Such metal-free, all-epi-growth approaches are especially important for metamaterials where more complicated structures on deep-subwavelength scales with higher fabrication precision are required. All these remarkable features will make the proposed all-semiconductor active plasmonic system the most promising one for tunable applications in MIR range.

Through a completely and systematically theoretical investigation on semiconductor nanophotonics, new phenomena and new physics on new and existing nanophotonic devices have been discovered. The proposals and designs of new devices have been demonstrated of great importance in realizing truly nano-scale semiconductor optoelectronic devices. It is believed that the understanding of new physics and phenomena and the methodology of modelling

and simulation in the research will potentially impact the current and next generation of nano-optoelectronic devices in the following decade. For example, the proposed NW laser and MSM laser would be very useful in the electrically driven nano-scale light source since they both can work at room temperature or a little higher temperature. Such devices will have significant advantages such as low power and potentially high speed, particularly for all-optical modulation. They could serve as a basis for high speed, low-power, high-complexity, all-optical signal processing systems that outperforms electronics in high speed applications. The proposed all-semiconductor plasmonic system can work as an on-chip communication system based on the light and plasmonics, which operates inside or between silicon chips. The easy tunability of plasmonic wavelength, i.e. the carrier wave wavelength, by simply changing the bias voltage is an important advantage of such communication system. This allows a fabricated communication system to be able to tune to different modulation wavelengths "on the fly", a significant important feature for adaptable communication systems. In addition, active gain materials can be integrated to provide a longer signal transportation length.

Acknowledgement

The research work in this thesis was supported by the Defense Advanced Research Project Agency (DARPA) program Nanoscale Architectures of Coherent Hyper-Optical Sources (NACHOS) and Science Foundation of Arizona.

References

1. F. Léonard and J. Tersoff, Phys. Rev. Lett., 83, 5174, (1999).

2. J. Hu, Y. Liu, C. Z. Ning, R. Dutton, and S. M. Kang, Appl. Phys. Lett., 92, 083503, (2008).

3. F. Léonard and A. A. Talin, Phys. Rev. Lett., 97, 026804, (2006).

4. X. F. Duan, Y. Huang, R. Agarwal, and C. M. Lieber, Nature, 421, 241, (2003).

5. M. S. Gudiksen, L. J. Lauhon, J. F. Wang, D. C. Smith, and C. M. Lieber, Nature, 415, 617, (2002).

6. O. Hayden, A. B. Greytack, and D. C. Bell, Adv. Mater., 17, 701, (2005).

7. M. Huang, S. Mao, H. Feick, H. Yan, Y. Wu, et al, Science, 292, 1897, (2001).

8. J. C. Johnson, H. J. Choi, K. P. Knutsen, R. D. Schaller, P. Yang, and R. J. Saykally, Nat. Mater., 1, 102, (2002).

9. A. H. Chin, S. Vaddiraju, A. V. Maslov, C. Z. Ning, M. K. Sunkara, and M. Meyyappan, Appl. Phys. Lett., 88, 163115, (2006).

10. A. Pan, W. Zhou, E. S. P. Leong, R. Liu, A. H. Chin, et al, Nano Lett., 9, 784, (2009).

11. Y. Xiao, C. Meng, P. Wang, Y. Ye, H. Yu, et al, Nano Lett., 11, 1122, (2011).

12. M. A Zimmler, F. Capasso, S. Müller and C. Ronning, Semicond. Sci. Technol., 25, 024001, (2010).

13. X. Y. Ma, J. Pan, P. Chen, D. Li, H. Zhang, Y. Yang, and D. Yang, Opt. Expr., 17, 14426, (2009).

14. H. Zhu, C. X. Shan, J. Y. Zhang, Z. Zhang, B. H. Li, et al. Adv. Mater. 22, 1877, (2010).

15. S. Chu, M. Olmedo, J. Y. Kong, Z. Yang, and J. L. Liu, Appl. Phys. Lett. 93, 181106, (2008).

16. C. Y. Liu, H. Y. Xu, J. G. Ma, X. H. Li, X. T. Zhang, et al, Appl. Phys. Lett., 99, 063115, (2011).

17. S. Chu, G. Wang, W. Zhou, Y. Lin, L. Chernyak, et al, Nat. Nanotech., 6, 506, (2011).

18. J. Hu, Y. Liu, A. Maslov, C. Z. Ning, R. Dutton, and S. M. Kang, Proc. of SPIE 6468, 64681E, (2007).

19. E. N. Economou, Phys. Rev., 182, 539, (1969).

20. R. B. Pettit, J. Silcox, and R. Vincent, Phys. Rev. B, 11, 3116, (1975).

21. J. J. Burke and G. I. Stegeman, Phys. Rev. B, 33, 5186, (1986).

22. P. Berini, Phys. Rev. B, 61, 10484, (2000).

23. J. A. Dionne, L. A. Sweatlock, and H. A. Atwater, Phys. Rev. B, 72, 075405, (2005).

24. F. Kusunoki, T. Yotsuya, J. Takahara, and T. Kobayashi, Appl. Phys. Lett., 86, 211101, (2005).

25. S. A. Maiera and H. A. Atwater, J. Appl. Phys., 98, 011101, (2005).

26. J. A. Dionne, L. A. Sweatlock, and H. A. Atwater, Phys. Rev. B, 73, 035407, (2006).

27. H. T. Miyazaki and Y. Kurokawa, Phys. Rev. Lett., 97, 097401, (2006).

28. J. Chen, G. A. Smolyakov, S. R. J. Brueck, and K. J. Malloy, Opt. Expr., 16, 14902, (2008).

29. N. M. Lawandy, Appl. Phys. Lett., 85, 5040, (2004).

30. M. Ambati, S. H. Nam, E. U.-Avila, D. A. Genov, G. Bartal, and X. Zhang, Nano Lett., 8, 3998, (2008).

31. T. Okamoto, F. H'Dhili, and S. Kawata. Appl. Phys. Lett., 85, 3968, (2004).

32. G. Winter, S. Wedge, and W. L. Barnes, New J. Phys., 8, 125, (2006).

33. I. Avrutsky, Phys. Rev. B, 70, 155416, (2004).

34. M. P. Nezhad, K. Tetz, and Y. Fainman, Opt. Expr., 12, 4072, (2004).

35. J. Seidel, S. Grafstrom, and L. Eng, Phys. Rev. Lett., 94, 177401, (2005).

36. M. A. Noginov, G. Zhu, M. Bahoura, J. Adegoke, C. E. Small, et al, Opt. Lett., 31, 3022, (2006).

37. M. A. Noginov, V. A. Podolskiy, G. Zhu, M. Mayy, M. Bahoura, et al, Opt. Expr., 16, 1385, (2008).

38. M. A. Noginov, G. Zhu, M. Mayy, B. A. Ritzo, N. Noginova, and V. A. Podolskiy, Phys. Rev. Lett., 101, 226806, (2008).

39. S. A. Maier, Opt. Comm., 258, 295, (2006).

40. Z. Yu, G. Veronis, S.Fan, and M. L. Brongersma, Appl. Phys. Lett., 92, 041117, (2008).

41. X. Chen, B. Bhola, Y. Huang, and S. T. Ho, Opt. Expr., 18, 17220, (2010).

42. A. V. Maslov and C. Z. Ning, Proc. of SPIE 6468, 646801, (2007).

43. C.Z. Ning, Phys. Stat. Sol. (b), 247, 774, (2010).

44. M. T. Hill, J. Opt. Soc. Am. B, 27, B36, (2010).

45. M. T. Hill, Y. S. Oei, B. Smalbrugge, Y. Zhu, T. de Vries, et al, Nat. Photon., 1, 589, (2007).

46. M. T. Hill, M. Marell, E. S. P. Leong, B. Smalbrugge, Y. Zhu, et al, Opt. Expr., 17, 11107, (2009).

47. M. A. Noginov, G. Zhu, A. M. Belgrave, R. Bakker, V. M. Shalaev, et al, Nature, 460, 1110, (2009).

48. R. F. Oulton, V. J. Sorger, T. Zentgraf, R.-M. Ma, C. Gladden, et al, Nature, 461, 629, (2009).

49. M. P. Nezhad, A. Simic, O. Bondarenko, B. Slutsky, A. Mizrahi, et al, Nat. Photon., 4, 395, (2010).

50. S. Kwon, J. Kang, C. Seassal, S. Kim, P. Regreny, et al, Nano Lett., 10, 3679, (2010).

51. K. Yu, A. Lakhani, and M. Wu, Opt. Expr., 18, 8790, (2010).

52. Ren-Min Ma, R. F. Oulton, V. J. Sorger, G. Bartal and X. Zhang, Nat. Mater., 10, 110, (2011).

53. C. Y. Lu and S. L. Chuang, Opt. Expr., 19, 13225, (2011).

54. K. Ding, Z. Liu, L. Yin, H. Wang, R. Liu, et al, Appl. Phys. Lett., 98, 231108, (2011).

55. D.B. Li and C.Z. Ning, Phys. Rev. B, 80, 153304, (2009).

56. D.B. Li and C.Z. Ning, Appl. Phys. Lett., 96, 181109, (2010).

57. R. Soref, R. E. Peale, and W. Buchwald, Opt. Expr., 16, 6507, (2008).

58. R. Soref, Nat. Photon., 4, 495, (2010).

59. S.-Y. Cho, and R. A. Soref, Opt. Lett., 34, 1759, (2009).

60. G. V. Naik and A. Boltasseva, Phys. Stat. Sol., RRL 4, 295, (2010).

61. A. J. Hoffman, L. Alekseyev, S. S. Howard, K.J. Franz, et al, Nat. Mater., 6, 946, (2007).

62. A. Boltasseva and H. A. Atwater, Science, 331, 290, (2011).

63. R. Teissier, D. Barate, A. Vicet, C. Alibert, A. N. Baranov, et al, Appl. Phys. Lett., 85, 167, (2004).

64. K. Ohtani, K. Fujita, and H. Ohno, Appl. Phys. Lett., 87, 211113, (2005).

65. J. Devenson, D. Barate, O. Cathabard, R. Teissier, and A. N. Baranov, Appl. Phys. Lett., 89, 191115, (2006).

66. Z. Tian, R. Q. Yang, T. D. Mishima, M.B. Santos, and M.B. Johnson, IEEE Photon. Tech. Lett., 21, 1588, (2009).

67. E. Feigenbaum, K. Diest, and H. A. Atwater, Nano Lett., 10, 2111, (2010).

68. D. B. Li and C. Z. Ning. Nano Lett., 8, 4234, (2008).

69. www.silvaco.com.

70. W. Shockley and W.T. Read, Phys. Rev., 87, 835, (1952).

71. Dziewior J. and W. Schmid, Appl. Phys. Lett., 31, 346, (1977).

72. S. C. Jane, J. M. McGregor, and D. J. Roulson, J. Appl. Phys., 68, 3747, (1990).

73. S. M. Sze. Physics of semiconductor devices, 2nd ed., (Wiley, Hoboken, 1981).

74. M. Levinstein, S. Rumyantsev, and M. Shur, Handbook Series on Semiconductor Parameters, Vol. 2, (World Scientific, London, 1999).

75. M. H. Lee, N. Lim, D. J. Ruebusch, A. Jamshidi, R. Kapadia, et al, Nano Lett., 11, 3425, (2011).

76. H. Katoa, M. Sanoa, K. Miyamotoa, and T. Yaob, J. Crystal Growth, 237, 538, (2002).

77. L. J. Brillson and Y. Lu, J. Appl. Phys., 109, 121301, (2011).

78. S. Nakamura, M. Senoh, T. Mukai, Jpn. J. Appl. Phys., 30, L1708, (1991).

79. J.S. Jang, K.H. Park, H.K. Jang, H.G. Kim, and S.J. Park, J. Vac. Sci. Technol. B, 16, 3105, (1998).

80. Otfried Madelung, Semiconductors: Data Handbook, 3rd ed., (Springer, Berlin, 2004).

81. A.V. Maslov and C.Z. Ning, Appl. Phys. Lett., 83, 1237, (2003).

82. www.comsol.com

83. S. L. Chuang, Physics of photonic devices, 2nd ed., (Wiley, Hoboken, 2008).

84. S.W. Chang, T.R. Lin, and S. L. Chuang, Opt. Expr., 18, 15039, (2010).

85. H. Raether, Surface Plasmons on Smooth and Rough Surfaces and on Gratings, (Springer, New York, 1988); S. A. Maier, Plasmonics: Fundamentals and Applications, (Springer, New York, 2006).

86. P. B. Johnson and R. W. Christy, Phys. Rev. B, 6, 4370, (1972).

87. A. V. Maslov and C. Z. Ning, IEEE J. Quantum Electron., 40, 1389, (2004).

88. L. D. Landau and E. M. Lifshitz, Electrodynamics of Continuous Media, 2nd ed., (Elsevier, Oxford, 1984).

89. W. W. Chow and S. W. Koch, Semiconductor-Laser Fundamentals, (Springer-Verlag, Berlin, 1999).

90. P. Winsemius, F. F. van Kampen, H. P. Lengkeek, and C. G. van Went, J. Phys. F: Met. Phys., 6, 1583, (1976).

91. R. H. M. Groeneveld, R. Sprik, and A. Lagendijk, Phys. Rev. Lett., 64, 784, (1990).

92. E. D. Palik, Handbook of Optical Constants of Solids, (Academic, New York, 1985); M. C. Tamargo, Optoelectronic Properties of Semiconductors and Superlattices, edited by M. O. Manasreh, Vol. 12, (CRC, New York, 2002).

93. D. Ahn, T.-K. Yoo, and H. Y. Lee, Appl. Phys. Lett., 59, 2669, (1991); Y.-H. Wu, IEEE J. Quan. Electron., 30, 1562, (1994).

94. Y. Z. Huang, Z. Pan, and R. Wu, J. Appl. Phys., 79, 3827, (1996).

95. Y. Z. Huang, IEE Proc.: Optoelectron., 148, 131, (2001).

96. T. D. Visser, H. Blok, B. Demeulenaere, and D. Lenstra, IEEE J. Quan. Electron., 33, 1763, (1997).

97. A. V. Maslov and C. Z. Ning, in Nitride Semiconductor Devices: Principles and Simulation, edited by J. Piprek, (Wiley, Hoboken, 2007).

98. J. T. Robinson, K. Preston, O. Painter, and M. Lipson, Opt. Expr., 16, 16659, (2008).

99. S. W. Chang and S. L. Chuang, IEEE J. Quan. Electron., 45, 1014, (2009).

100. A. V. Maslov and M. Miyawaki, J. Appl. Phys., 108, 083105, (2010).

101. G. W. Ford, and W. H. Webber, Phys. Rep., 113, 195, (1984).

102. H. Haug and S.W. Koch, Quantum theory of the optical and electronic properties of semiconductors, 5th ed., (World Scientific, Hoboken, 2009).

103. C. Manolatou and F. Rana, IEEE J. Quan. Electron., 44, 435, (2008).

104. A. Mizrahi, V. Lomakin, B. Slutsky, M. Nezhad, L. Feng, and Y. Fainman, Opt. Lett., 33, 1261, (2008).

105. V. Almeida, Q. Xu, C. Barrios, and M. Lipson, Opt. Lett., 29, 1209, (2004).

106. S. Adachi, J. Appl. Phys., 58, R1, (1985).

107. E. Avrutin, I. Chebunina, I. Eliachevitch, S. Gurevich, M. Portnoi, and G. Shtengel, Semicond. Sci. Technol., 8, 80, (1993).

108. W. Batty, U. Ekenberg, A. Ghitit, and E. P. O'Reilly, Semicond. Sci. Technol., 4, 904, (1989).

109. D. W. Lynch and W. R. Hunter, in Handbook of Optical Constants of Solids, edited by E. D. Palik, (Academic, New York, 1998).

110. D. B. Li and C. Z. Ning, Opt. Expr., 19, 14594, (2011).

111. S. Adachi, Phys. Rev. B, 35, 7454, (1987).

112. E. Burstein, Phys. Rev., 93, 632, (1954).

113. H. Kroemer, Phys. E, 20, 196, (2004).

114. D. C. Larrabee, G. A. Khodaparast, J. Kono, K. Ueda, Y. Nakajima, et al, Appl. Phys. Lett., 83, 3936, (2003).

115. J. Li, K. I. Kolokolov, C. Z. Ning, D. C. Larrabee, G. A. Khodaparast, et al, Progress in Semiconductors II: Electronic and Optoelectronic Applications, MRS Proceedings, 744, 571, (2003).

116. K. Kolokolov and C. Z. Ning, Appl. Phys. Lett., 83, 1581, (2003).

117. J. D. Jackson, Classical Electrodynamics, 3rd ed., (Wiley, New York, 1999).

118. S. Adachi, Handbook on Physical Properties of Semiconductors, Vol. 1, 2, (Springer, Norwell, 2004).

APPENDIX A

DERIVATION OF CONFINEMENT FACTOR

In a source-free linear medium, the Poynting theorem for a monochromatic field with angular frequency, ω, is given by [117]

$$\oint_S (\mathbf{E}\times\mathbf{H}^*)\bullet\mathbf{n}da + i\omega\int_V (\mathbf{E}\bullet\mathbf{D}^* - \mathbf{B}\bullet\mathbf{H}^*)dv = 0 \quad \text{(A-1)}$$

Consider now a waveguide with waves propagating along the positive z direction. We assume that the dielectric constant in the active region is uniform and is given as a complex value $\varepsilon = \varepsilon' + i\varepsilon''$. Eq. (A-1) thus can be written as

$$\iint\left[(\mathbf{E}\times\mathbf{H}^*)_{z+\Delta z} - (\mathbf{E}\times\mathbf{H}^*)_z\right]dxdy + i\omega\iint(\varepsilon^*\varepsilon_0|\mathbf{E}|^2 - \mu_0|\mathbf{H}|^2)dxdy\Delta z = 0 \quad \text{(A-2)}$$

Taking the real part of Eq. (A-2) yields

$$\iint(S_{z+\Delta z} - S_z)dxdy = -\frac{1}{2}\omega\iint\varepsilon''\varepsilon_0|\mathbf{E}|^2\,dxdy\Delta z \quad \text{(A-3)}$$

where $S_z = \frac{1}{2}\mathrm{Re}(\mathbf{E}\times\mathbf{H}^*)_z$, is the time-average Poynting vector in z direction. By using the field expressions in Eq. (4-21), the total power through the entire waveguide at location z is thus written as

$$I(z) = \iint S_z dxdy = \frac{1}{2}|A(z)|^2\iint\mathrm{Re}[\overline{\mathbf{E}}(x,y)\times\overline{\mathbf{H}}(x,y)^*]\bullet\hat{\mathbf{z}}dxdy \quad \text{(A-4)}$$

and Eq. (A-3) can be rewritten, when $\Delta z \to 0$, as

$$\frac{dI(z)}{dz} = -\frac{1}{2}\omega\varepsilon''\varepsilon_0|A(z)|^2\iint|\overline{\mathbf{E}}(x,y)|^2\,dxdy \quad \text{(A-5)}$$

Since $dI/dz = \Gamma G_0 I$ according to Eq. (4-18), where

$$G_0 = -\frac{\varepsilon''\omega}{n_b c} \quad \text{(A-6)}$$

Divide Eq. (A-5) by Eqs. (A-4) and (A-6), we obtain the confinement factor Γ as

$$\Gamma = \frac{n_b c\varepsilon_0\iint|\overline{\mathbf{E}}(x,y)|^2\,dxdy}{\iint\mathrm{Re}[\overline{\mathbf{E}}(x,y)\times\overline{\mathbf{H}}(x,y)^*]\bullet\hat{\mathbf{z}}dxdy} \quad \text{(A-7)}$$

where n_b is the real part of the background refractive index of the active material and c is the speed of light in vacuum.

APPENDIX B

DETERMINATION OF LASING THRESHOLD

In time domain, the field intensity of a mode, I, in an active cavity is described by

$$\frac{dI}{dt} = (\Gamma G_0 - \alpha_m) v_g I \tag{B-1}$$

where Γ is the confinement factor of the cavity for the mode, thus ΓG_0 is the corresponding modal gain per unit length, v_g is the group velocity and α_m is the overall loss of the cavity seen by the mode. The solution of Eq. (B-1) has the form of

$$I = I_0 e^{-\gamma t} \tag{B-2}$$

where $\gamma = -(\Gamma G_0 - \alpha_m) v_g$. Assume the center frequency of the mode is ω_0, the Fourier transform of Eqn. (A2) yields

$$L(\omega) = \int_0^\infty I_0 e^{-\gamma t} e^{-i(\omega - \omega_0)t} dt = \frac{I_0}{\gamma - i(\omega - \omega_0)} \tag{B-3}$$

The real part of Eq. (B-3) gives the field intensity in frequency domain or the intensity spectrum as

$$\text{Re}[L(\omega)] = \frac{I_0 \gamma}{\gamma^2 + (\omega - \omega_0)^2} \tag{B-4}$$

Now $|\gamma|$ is the linewidth of the intensity spectrum. At the lasing threshold, the modal gain is equal to the overall loss of the cavity seen by the mode, i. e., $\Gamma G_{0,th} = \alpha_m$. Therefore the linewidth of the intensity spectrum is zero. We then are able to determine the threshold gain by tuning material gain to the situation where the linewidth of the intensity spectrum is equal to or very close to zero, depending on the numeral accuracy.