© 2015 Pengfei Qiao

SURFACE-EMITTING LASERS FOR COMMUNICATIONS: NOVEL METAL-CAVITY MICROLASERS AND HIGH-CONTRAST-GRATING TUNABLE VCSELS

BY

PENGFEI QIAO

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Shun Lien Chuang, Director of Research
Professor Weng Cho Chew, Chair
Associate Professor John Dallesasse
Assistant Professor Daniel Wasserman
Professor Brian Cunningham

ABSTRACT

A comprehensive study of the theory and experiments of surface-emitting semiconductor lasers is presented. The design of novel micro and nanolasers using metal cavities for optical confinement is discussed. Theoretical modeling of quantum-well and quantum-dot emission properties, as well as experimental characterization of their coupling with optical cavities, are presented. Lasing behavior of our designed and fabricated devices is demonstrated at room temperature under continuous-wave and pulsed electrical injection with 3-μm and 1-μm cavity diameters, respectively. This work provides the research path toward dense-integrable power-efficient on-chip light sources. Surface-emitting tunable lasers for high-speed, long-haul communication are investigated. Novel laser designs using micro-electro-mechanical system controlled high-contrast gratings as tunable mirrors are presented. Rigorous, accurate, and efficient electromagnetic models for high-contrast gratings are developed. Our model enables us to design high-contrast gratings as one-dimensional or two-dimensional metastructures integrable on surface-emitting lasers. A wide range of optical functionalities such as broadband reflection, high-Q resonance, filtering, beam-steering, focusing, beam-conversion, and generation of photon orbital angular momentum are achieved. Our optical model is integrated with our laser cavity model and the rate-equation model to predict the temperature-dependent voltage tunable light output intensity and spectra. Future design and experimental strategies for heterogeneously integrated tunable surface-emitting lasers are discussed.

To my parents, for their love and support.

ACKNOWLEDGMENTS

I deeply thank my adviser Professor Shun Lien Chuang for the opportunity to work with him since August 2010. This work would not have been possible without his guidance and support. When I was about to finish my master's degree in May 2012, Professor Chuang was diagnosed with the stage IV lung cancer. When I was feeling hopeless and lost, he was the one who encouraged me not to give up, right after he was released from intensive care at Stanford. Despite of his physical conditions, his passion for scientific research and his care for his students never diminished. He continued to remotely meet with us and care for our research progress, while bravely fighting against his decease for almost two years. It was lucky for me that he came back to the University of Illinois and attended my preliminary exam in October 2013, though it was too soon and sudden for him to leave us in March 2014. Yet I feel thankful to have had Professor Chuang not only as my research adviser, but also my teacher and mentor. His perseverance, dignity, intelligence, and sense of humor will continue to guide me and enlighten me. I also thank Professor Chuang's wife, Lolita, for her care and encouragement.

I would like to thank Professor Weng Cho Chew, who was my undergraduate adviser, and is currently my committee chair. He was the one who stimulated my strong interest in electromagnetics. His intelligence, kindness, and desire to keep learning and exploring new areas, have always motivated me. I also thank Professor John Dallesasse for the help with my experiments. And I appreciate Professor Daniel Wasserman and Professor Brian Cunningham for serving as my final defense committee.

I also thank my collaborators Professor Connie Chang-Hasnain and Professor Ming Wu at University of California at Berkeley. They supported my work on the E-PHI project as well as my five-month visit at Berkeley. A significant amount of work in this dissertation was produced with Professor Chang-Hasnain's guidance. I also thank Professor Dieter Bimberg at Techni-

cal University of Berlin. He kindly provided us with high-quality quantum-well and quantum-dot samples, from which we were able to produce many significant experimental results.

I thank my previous group members Dr. Chien-Yao Lu, Dr. Akira Matsudaira, and Dr. Adrian Chi-Yu Ni. Chien-Yao guided me remotely on the nanolaser fabrication and characterization with great patience, even after he graduated. Akira trained me on the InP-based nanolaser fabrication, and I had much discussion with Adrian on laser modeling. I appreciate all the insightful discussion with my fellow group members: Daniel Zuo, Benjamin Kesler, Thomas O'Brien, and Guan-Lin Su.

Last but most importantly, I am deeply grateful to my parents who have raised me and always been supportive. They always encourage and trust me whatever difficulty I encounter. It would not have been possible for me to come to this stage of my life without their measureless love and unconditional support.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Motivation and Research Goals

The demand for data transmission in our daily lives is growing at an ever-increasing rate. The global data center traffic is expected to have a three-fold increase from 2013 to 2018 [1], as shown in Fig. 1.1. Such demand is a driving force for the development of technologies that enable faster and faster data communication. Semiconductor lasers are key components that enable the long-haul communication for the world-wide Internet connections. Endless research efforts have been devoted to searching novel data modulation and multiplexing schemes, as well as improving the modulation speed for individual devices.

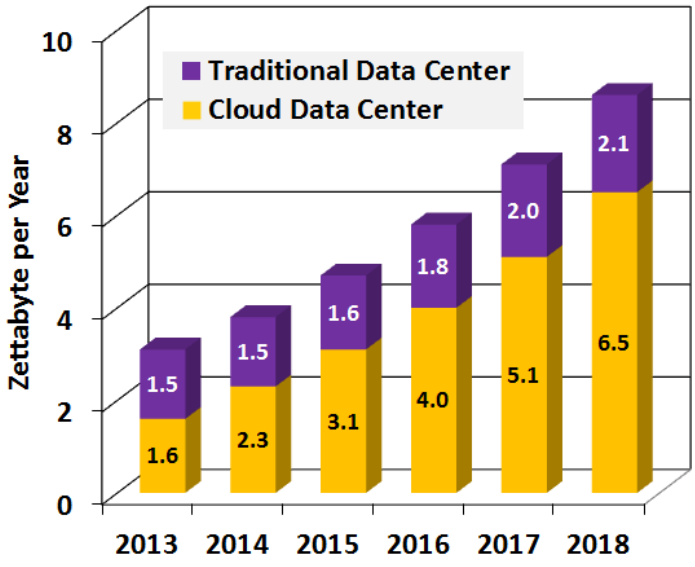


Figure 1.1: Cisco's prediction on total data center traffic growth adapted from [1].

Accompanying the heavy data traffic, the increase of data processing ca-

pacity is another demand. As a result, the integration density of electronic devices, such as transistors, is doubling every eighteen months, which is commonly regarded as Moore's law. Figure 1.2 shows Intel's effort to keep up with Moore's law, and developing new generations of technology becomes a crucial matter. One example is the on-chip copper transmission lines, a bottle-neck for high-speed data transmission, which are being gradually replaced by hybrid opto-electronic structures [2].

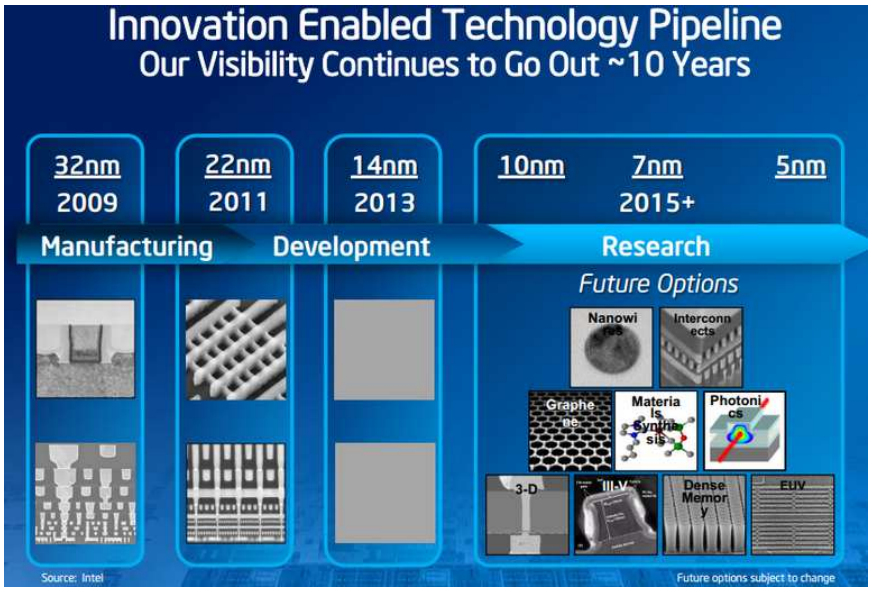


Figure 1.2: Intel's effort to keep up with Moore's law and develop new generations of technology. [2]

Semiconductor lasers, being efficient and high-performance coherent light sources, are key components that enable high-speed data transmission. Ever since the inventions of the first laser by Maiman in 1960 [3] and the semiconductor lasers by Hall and three other groups [4, 5, 6, 7], the miniaturization of lasers has been going on continuously. The invention of vertical-cavity surface-emitting lasers (VCSELs) by Iga in 1979 [8, 9] was a milestone, which enabled coherent light generation with small device footprint, low threshold, and high beam quality. The first room-temperature continuous-wave (CW) operation of a VCSEL was reported in 1988 [10]. In the 1990s and 2000s, researchers engineered novel laser cavities to confine the optical modes down to micrometer and sub-micrometer scales. Notable breakthroughs include the whispering-gallery mode microdisk lasers in 1992 [11] and the photonic crystal lasers in 2004 [12]. In 2007, the first metal-cavity nanolaser under

CW injection was demonstrated at 77 K [13], where the optical field was confined using the metal-cavity and mode-cutoff.

In this dissertation, we will focus on surface-emitting semiconductor lasers for optical communications. On one hand, our research goal is the miniaturization and integration of micro and nanolasers using novel cavity designs and optical gain materials. This will enable the dense integration of low-cost, low-footprint, low-crosstalk, and power-efficient coherent light sources for on-chip communication. On the other hand, our research goal is toward the high-speed, high-power, widely-tunable, low-cost, densely-integrable light sources, which enable long-haul optical communication using various modern multiplexing schemes.

1.2 Challenges and Issues with Current Technologies

In the current technology, copper still plays a major role in interconnects especially for chip-to-chip and board-to-board communication. As the data rate approaches tens and twenties of gigabits per second, copper is already approaching its limits with its significantly large attenuation and energy loss. Numerous researchers have proposed optical interconnects and the integration of photonic devices for high bandwidth, low loss, low power consumption and low noise transmission. One of the major challenges for integrated photonics is the on-chip light source.

As electrical signals are confined and guided in structures with a physical scale on the order of tens of nanometers, shrinking the dimensions of photonic devices to such nano scales is extremely challenging. As we approach such a size limit, known as the diffraction limit $(\lambda_0/2n_\mathrm{r})^3$, the optical diffraction loss drastically increases. The optical field becomes poorly confined and increasing crosstalk between neighboring devices forbids increasingly dense integration. The electrical contacts, isolation, as well as heat management all become more critical for devices at the sub-micron and nanometer scales.

Another challenge is the heterogeneous integration between the electronic and photonic systems. Most electronic devices are based on the silicon platform, while photonic devices, especially efficient semiconductor light sources, are based on III-V compound semiconductor materials. Research efforts are directed toward either establishing efficient group IV light sources, or finding

stable, high-performance integration methods.

Multiplexing technologies are crucial for modern data transmission, and wavelength-division multiplexing (WDM) has become the most widely used method for fiber-optic communication systems. Multiple channels with optical carriers at different wavelengths are packed into the C-Band of silica fibers around 1550 nm. Coarse WDM (CWDM) utilizes 8 channels and dense WDM (DWDM) utilizes up to 80 channels with 50 GHz spacing in the same window. However, further increasing the channel density is very challenging, because very strict requirements would be placed upon the laser linewidth and the fiber dispersion. Researchers have been seeking other independent multiplexing techniques besides the wavelength-division, time-division, and space-division multiplexing. Polarization is another degree of freedom for optical signals. However, photons can only have two different spin angular momenta, and thus each optical mode can only have two independent polarizations. The orbital angular momentum (OAM) of light, on the other hand, can theoretically be any integer multiple of $\hbar$. This unique feature has attracted extensive research interest in the last decade. The challenge lies in the efficient generation and stable guidance of these beams containing high orders of OAM, which are vortex beams.

1.3 State of the Art

In 2009, Hill et al., collaborating with Ning's group at Arizona State University (ASU), demonstrated metal-cavity nanolasers operating at 298 K under pulsed electrical injection [14]. Ding et al. from the ASU group further demonstrated CW operation of the sub-wavelength metal-cavity laser with a cavity volume of $0.67\lambda_0^3$ ($\lambda = 1591\,\text{nm}$) in 2013 [15, 16]. Optically pumped nanolasers were achieved at 4 K by Kwon et al. using the nanopan design [17], as well as at 77 K by Wu's group at the University of California at Berkeley (UCB) using the nanopatch design [18], both in 2010. Fainman's group at the University of California at San Diego (UCSD) demonstrated optically pumped subwavelength lasers at room temperature in 2010 [19]. Lee et al. from the UCSD group further demonstrated in 2011 the electrically pumped nanolaser at 140 K using the coaxial structure [20].

Here in our group at the University of Illinois at Urbana-Champaign

(UIUC), Lu et al. have designed and realized the metal-cavity surface-emitting 2-μm-diameter laser integrated on silicon, operating at room temperature under CW injection, in 2010 [21], and directions for further size reduction were subsequently proposed [22]. We have further replaced the conventional strained quantum wells (QWs) by the submonolayer quantum dots (QDs) as the gain medium, and demonstrated room temperature operation under electrical injection [23, 24, 25, 26, 27, 28].

Tunable lasers are important components for the WDM systems, and tunable VCSELs with micro-electro-mechanical systems (MEMS) prove to be favorable due to their continuous tuning capability, low cost, and low power consumption [29]. Gierl et al. from Amann's group demonstrated 102 nm tuning in 1550-nm VCSELs in 2011 [30]. Potsaid et al. demonstrated 110 nm wavelength tuning in a 1310-nm VCSEL in 2012 [31]. Yet due to the fabrication complexity, low-cost tunable VCSELs are still not widely available. Huang et al. from Chang-Hasnain's group at UCB demonstrated the first tunable VCSEL at 850 nm using the low-cost, compact, high-contrast grating (HCG) in 2007 [32]. The UCB group, in collaboration with our group here at UIUC, further extended the work to the 1550-nm VCSELs since 2010 [33, 34, 35, 36, 37]. In 2014, the first 1550-nm VCSEL heterogeneously integrated on silicon-on-insulator (SOI) using the silicon HCG was demonstrated [38, 39].

Besides improving the performance of individual devices, researchers are also seeking novel multiplexing schemes for optical data transmission on top of the existing WDM systems. Recently, Willner's group at the University of Southern California (USC) demonstrated terabit free-space data transmission using OAM multiplexing [40]. However, developing efficient integrated light sources for generating optical beams with OAM remains challenging. Our group here at UIUC and Chang-Hasnain's group at UCB are collaborating on the design and experimental demonstration of phase manipulation of surface-emitting lasers with two-dimensional high-contrast metastructures [41].

1.4 Dissertation Overview

This dissertation presents a comprehensive study on the theory, design, fabrication, and characterization of novel surface-emitting semiconductor lasers

for optical communication. These power-efficient, high-speed, low-cost, and low-footprint coherent light sources are key components in the upcoming electronic-photonic heterogeneously dense-integrated systems.

Chapter 2 presents the theory and design for novel surface-emitting micro and nanolasers. The light emission properties of QDs as the laser gain medium are predicted using our theoretical formulation with effective-mass approximation. The resonance behavior of micro and nanocavities is captured by our electromagnetic model. The coupling between the QD spontaneous emission and the cavity modes, and its effects on laser behavior, are investigated thoroughly. Our theory is used to explain the electron-photon interaction in laser cavities, as well as the light emission power and spectrum for a given injection current.

Chapter 3 presents the fabrication and characterization of our micro and nanolasers. The epitaxial wafers are grown by our collaborators and the devices are fabricated in the cleanrooms in the Micro and Nanotechnology Laboratory (MNTL) at UIUC. Detailed fabrication steps are included. The electrical and optical characterizations of our fabricated devices are performed in Professor Shun Lien Chuang's laboratory in MNTL. Light emission intensity, spectra, and field profiles are collected for devices with various sizes under various injection conditions. Our tested devices use metallic or dielectric materials to form the optical cavity, as well as QDs or QWs as the gain medium. We demonstrate lasing behavior of our devices at room-temperature under electrical injection. The experimental observation confirms our prediction on the sub-threshold and above-threshold emission properties.

Chapter 4 focuses on the theory and design of a novel optical component, known as the high-contrast grating (HCG). We develop electromagnetic models for HCGs with one-dimensional periodicity based on the mode-matching method and the surface-integral-equation method with periodic Green's functions. The results are compared with simulations using commercial packages. For HCGs with two-dimensional (2D) periodicity, we develop a rigorous coupled-wave analysis (RCWA) package, which is efficient for design purposes and accurate for understanding the physics of HCGs. We propose a design procedure for 2D HCGs to realize various optical functionalities, such as beam-steering, beam-focusing, and beam-conversion. We further design 2D phase plates which can modulate the orbital angular momentum of light. Our high-performance designs are confirmed by full-wave simulations.

Chapter 5 discusses the application of HCG as high-performance mirrors in tunable vertical-cavity surface-emitting lasers (VCSELs). We demonstrate a comprehensive model for the MEMS-tunable HCG VCSELs. The model calculates the temperature-dependent material gain and spontaneous emission spectra of the strained QW active region. The optical properties of the HCG reflector are obtained from the previous chapter and the HCG VCSEL cavity is modeled with the transfer matrix method, which produces important parameters for device-level simulation. The rate-equation model takes into account the thermal effects and our calculated temperature-dependent L-I curves show excellent agreement with experiment. Our MEMS model further correlates the tuning voltage with the resonance wavelength, threshold current, and peak power. The measurements can be accurately explained by our model.

Chapter 6 summarizes the major accomplishments of this dissertation and provides promising future research directions.

CHAPTER 2

THEORY AND DESIGN OF NOVEL SURFACE-EMITTING MICRO AND NANOLASERS

2.1 Introduction

Metal-cavity microlasers have attracted extensive research interest in the past few years. Since the first demonstration of lasers in 1961 and the first vertical-cavity surface-emitting laser (VCSEL) in 1979 [8], metals have been used by researchers for the miniaturization of semiconductor lasers. With the help of metals, micrometer- to nanometer-scale light sources are now promising candidates for low foot print high-speed optical communication systems. Novel structures based on metal cavities, photonic crystals, or surface plasmons have been employed to confine the optical mode close to or even below diffraction limit. Recent work on metal cavity nanolasers has demonstrated operation not only under optical pumping [18, 19], but also at room temperature under electrical injection [15, 16, 20, 22].

Low dimensional materials such as quantum dots (QDs) have been employed as the active material in lasers showing enormous advantages as compared to conventional quantum-well lasers, such as high differential gain, low threshold current, high temperature stability, and wide modulation bandwidth due to their discrete density-of-states (DOS) [42]. Submonolayer quantum dots (SML QDs) as an alternative to the Stranski-Krastanow grown quantum dots (S-K QDs), have been demonstrated with much higher modal gain, better uniformity, less inhomogeneous broadening, and sharper emission spectra [43]. The smaller size and shape deviations of SML QDs allow a much higher saturated gain. Since SML QDs do not have wetting layers (WLs), the carrier population in WL bound states and the carrier scattering from WL states into QDs are avoided [44]. Hence, both maximum gain and modulation bandwidth are expected to be larger.

One major challenge of achieving room-temperature electrical-injected mi-

crolasers is to have enough gain to balance the loss. Both the radiation and material loss increase rapidly as cavity size shrinks. Heat accumulation from increasing series resistance is another limiting factor for electrical-injection microlasers. Thermal effects result in unstable threshold current and early output power roll-over. SML QDs have shown improved thermal stability in threshold current and differential efficiency in high-power high-speed VCSELs [44] and are also promising for micro-cavity lasers.

In this chapter, we develop a theoretical model to investigate the size-dependent device performance of metal-cavity SML QD microlasers. Devices are demonstrated to lase at room temperature under electrical injection with device radius down to $2\,\mu$m for continuous wave (CW) and to $0.5\,\mu$m for pulsed mode operations. Using a quantum disk model for S-K grown QDs, we have successfully explained experimental results such as optical gain and linewidth enhancement factor [45]. In this work [26, 27, 28], we extend the model for multi-stack SML QDs with strong vertical correlation, and consider them as effective quantum disks. Strain effects on the heterojunctions are included in the Hamiltonian for calculating the electronic states in SML QDs. The QD material gain and the spontaneous emission rate are obtained with Fermi's golden rule and both homogeneous and inhomogeneous broadening effects are considered. The characteristics of the laser optical cavity are solved using Maxwell's equations semi-analytically, with the effective index method and the transfer matrix method. We then use the rate-equation model to study the interaction between the injected carriers and the generated photons. The calculated QD gain and cavity properties are used as inputs and the light output power versus current (L-I) behavior is predicted. Our theory agrees with experimental data for device radii from $5\,\mu$m down to $0.5\,\mu$m.

The coupling of spontaneous emission (SE) into the cavity modes becomes much more significant as the cavity size and the effective mode volume reduce [46, 47]. In conventional semiconductor lasers, such as VCSELs using quantum wells (QWs), the free-space SE bandwidth is much wider than the cavity linewidth, i.e. $\Gamma_{\mathrm{sp,\,free}} \gg \Gamma_{\mathrm{cav}}$. Then the SE coupled into the lasing mode can be approximated by the total SE over the full spectrum multiplied by a coupling factor β_{sp}, which is set as a constant [48]. On the other hand, researchers study how the radiation environment affects the emission property, i.e. the Purcell effect [46, 47, 49, 50, 51], and the formulation often

assumes a two-level system, such as an artificial atom, as the emitter. This is the case for the other extreme where $\Gamma_{\mathrm{sp,\,free}} \ll \Gamma_{\mathrm{cav}}$. However, we are interested in the intermediate regime $\Gamma_{\mathrm{sp,\,free}} \approx \Gamma_{\mathrm{cav}}$, which is the case for semiconductor QDs emitting in micro and nanocavities.

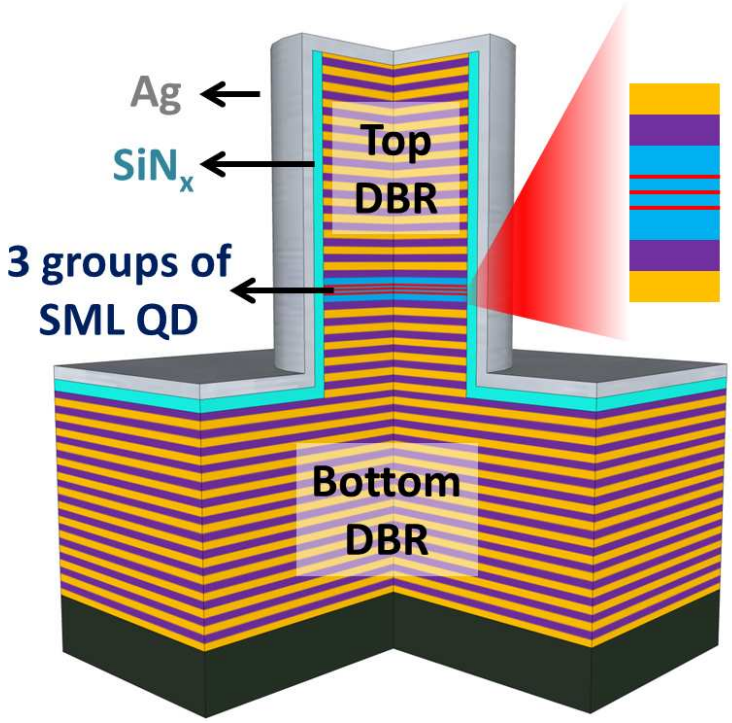
In this work [27, 28], we derive a rigorous expression for the coupling factor β_{sp} which accounts for both the emission properties of the QDs and the radiation environment modified by the cavity. The sub-threshold L-I behavior we observe is successfully explained by the increasing amount of carrier density-dependent spontaneous emission coupling into the cavity mode. The β_{sp} factor at threshold increases drastically as we reduce the device size due to the more sparse mode distribution within the gain spectrum.

Figure 2.1(a) shows the schematic of the metal-cavity microlasers. The active region contains three groups of SML QDs. The device sidewall is passivated by silicon nitride (SiN_x) for both electrical isolation and optical buffering to reduce metallic loss. The whole device is covered by silver to form the metal cavity. The top/bottom mirrors of the cylindrical $3\lambda/2n_{\mathrm{r}}$ microcavity are formed by 19/32 pairs of p-doped/n-doped AlGaAs/GaAs distributed Bragg reflectors (DBR), respectively.

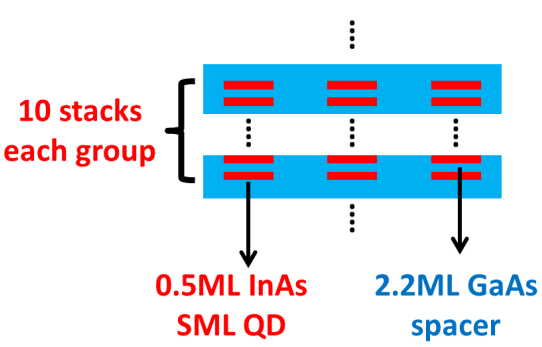
2.2 Optical Gain of Submonolayer Quantum Dots

The active region of the microlasers consists of three groups of SML QDs [44, 52], each being 8 nm thick and separated by 13 nm GaAs spacers. Each group of the SML QDs consists of ten stacks of 0.5-monolayer InAs QD layers, separated by 2.2-monolayer GaAs spacers. The nominal structure for each group of SML QDs is shown as Fig. 2.1(b) [52]. Vertically-correlated SML QDs in each group are modeled as effective quantum disks. Extensive work [53] has been done to analyze the effect of size, shape, and piezoelectricity on QD optical properties. In our model, since the SML QD layers are thin, we assume the effective quantum disks to have no variance in the growth direction. Such cylindrical high-symmetry structure leads to a negligible shear strain, thus we consider the QD strain as biaxial, namely, $\epsilon_{xx} = \epsilon_{yy} \neq \epsilon_{zz}$, but $\epsilon_{xy} = \epsilon_{yz} = \epsilon_{zx} = 0$.

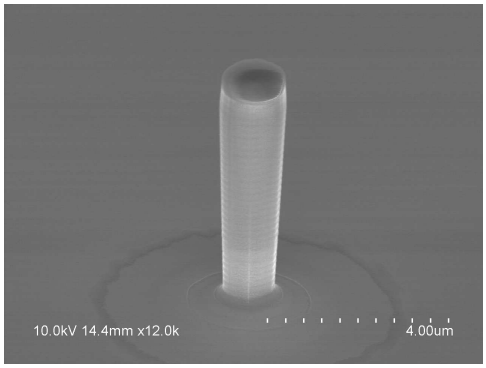
Since the shear strain is assumed negligible, there are two effects that also become negligible: the piezoelectric effect [53, 54] and valence-band mixing.



(a)



(b)



(c)

Figure 2.1: (a) Schematic of a submonolayer (SML) quantum-dot (QD) metal-cavity surface-emitting laser. The active region contains 3 groups of SML QDs. (b) Schematic of each group of SML QDs, consisting of 10 stacks of 0.5-monolayer InAs QDs separated by 2.2-monolayer thick GaAs spacers.[44] (c) A scanning electron micrograph of a 0.5-μm-radius microlaser before the SiN_x sidewall-passivation and metal-coating.

For materials with zincblende crystal structure, the first-order piezoelectric polarization $\mathbf{P}_1$ is only dependent on the shear strain because the only non-zero elements in the strain tensor $[\bar{\mathbf{e}}]_{3\times6}$ are $e_{14} = e_{25} = e_{36}$. The second-order piezoelectric polarization $\mathbf{P}_2$ also vanishes when no shear strain exists

[53, 54]. Due to the discrete energy levels of quantum dots, the band mixing effect is only caused by strain. That is, the off-diagonal terms in the coupled Hamiltonian [55, 56, 57] are only dependent on the shear strain (ϵ_{xy}, ϵ_{yz}, and ϵ_{zx}), where $\epsilon_{xx} - \epsilon_{yy}$ is also negligible due to symmetry. Therefore, we can treat the conduction, heavy-hole, light-hole, and spin-orbit split-off bands as fully decoupled, and we can solve the single-band Hamiltonian for each band individually using the corresponding Γ-point effective mass. The biaxial strain effect is included through the extra strain terms from the Pikus-Bir Hamiltonian [57]. The material parameters related to the band gap, strain and effective mass are taken from the experimental data summarized in [58]. The eigen-problem formed by the Hamiltonian is solved numerically using the 3D finite-difference method. The Dirichlet boundary condition is used for an isolated QD when the dot density is low or the 2D fill factor is small. Periodic boundary conditions are used at high dot density and 2D fill factor to include the lateral coupling among QDs.

Figure 2.2(a) shows the isosurfaces of the 3D wavefunctions for the first two conduction band states (CB1 and CB2) and heavy-hole states (HH1 and HH2). The QDs are assumed to be isolated laterally. In this figure, the effective diameter of the vertically-stacked SML QDs is assumed as 20 nm. We can see that the half maximum isosurface of the ground state and the first excited state wavefunctions are still well contained in the stacked SML QDs. Figure 2.2(b) shows the case when the lateral coupling is not negligible (2D fill factor being 62.8%) and a periodic boundary condition is used. In this case, the squared magnitude of the wavefunction at the mid-point between two unit cell centers is 19.8% of that at the unit cell center. When the 2D fill factor is 31.4%, the squared magnitude of the wavefunction at the mid-point between two unit cell centers becomes 0.72% of that at the unit cell center, which indicates the lateral coupling is negligible in this case.

The calculated ground state transition (C1 to HH1) energy as a function of the temperature and effective dot diameter is shown in Fig. 2.3(a). The temperature-dependent material band gap is used from the empirical Varshni equation [58, 59]. The measured ground state transition energy [44] from photoluminescence is shown for comparison. We conclude that the effective QD size is close to the size observed in [52].

To account for inhomogeneous broadening [42, 55] due to QD size variations, the electron and hole carrier densities (n and p) are related to the

quasi-Fermi levels (F_c and F_v) as

$$\begin{aligned} n &= 2\frac{N_{\text{dot}}^{\text{2D}}}{L_z}\sum_i \int dE \left[\frac{1}{\sqrt{2\pi}\sigma_c} e^{-(E-E_c^i)^2/2\sigma_c^2}\right] f_c(E,F_c), \\ p &= 2\frac{N_{\text{dot}}^{\text{2D}}}{L_z}\sum_j \int dE \left[\frac{1}{\sqrt{2\pi}\sigma_v} e^{-(E-E_v^j)^2/2\sigma_v^2}\right] f_v(E,F_v) \end{aligned} \tag{2.1}$$

respectively. The subscripts c and v indicate conduction and valence bands, respectively. The square brackets are the linewidth broadening functions, with linewidths being σ_c and σ_v. Here, $N_{\text{dot}}^{\text{2D}}$ is the 2D dot density, f_c and f_v are the Fermi occupation probabilities, and E_c^i and E_v^j are the energies for the i-th conduction band and j-th valence band, respectively. And L_z is the thickness of each QD group.

Figure 2.3(b) shows the carrier-dependent quasi-Fermi level (F_c) for the CB obtained from Eq. (2.1), together with the first fifty CB states, among which three are bound states (one ground state and two degenerate excited states). Similarly, the carrier-dependent quasi-Fermi level (F_v) for the VB is shown in Fig. 2.3(c). The first fifty HH states shown are all bound states, while only the first LH state out of the twenty states shown is a bound state.

Once the quasi-Fermi levels are obtained, we can calculate the carrier-dependent material gain (cm^{-1}) and the free-space spontaneous emission rate ($\text{cm}^{-3}\text{s}^{-1}\text{eV}^{-1}$) as [42, 45, 55]

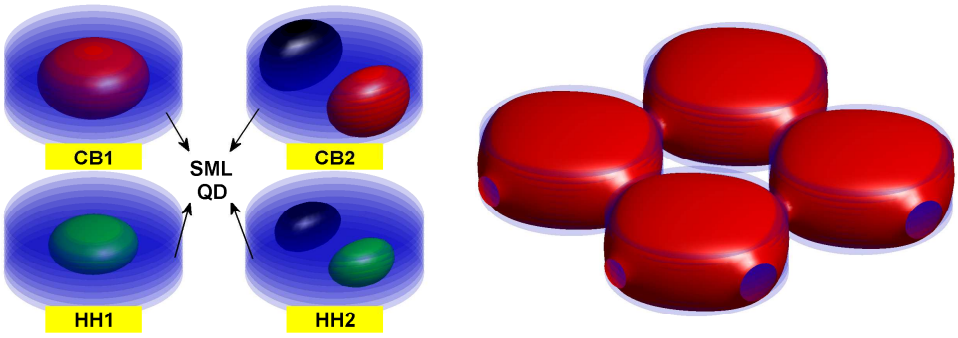
$$\begin{aligned} g(\hbar\omega) &= \frac{2N_{\text{dot}}^{\text{2D}}}{L_z} C_0 \sum_{i,j} \int dE |M_{\text{env}}^{ij}|^2 |\hat{e}\cdot\mathbf{p}_{cv}|^2 D(E,E_{cv}^{ij}) L(E,\hbar\omega)(f_{c,i}-f_{v,j}), \\ r_{\text{spon}}(\hbar\omega) &= \frac{2N_{\text{dot}}^{\text{2D}}}{L_z} B_0 C_0 \sum_{i,j} \int dE |M_{\text{env}}^{ij}|^2 |\hat{e}\cdot\mathbf{p}_{cv}|^2 D(E,E_{cv}^{ij}) L(E,\hbar\omega) f_{c,i}(1-f_{v,j}) \end{aligned} \tag{2.2}$$

where M_{env}^{ij} is the overlap integral between the envelop functions of i-th CB and j-th VB states, and E_{cv}^{ij} is the transition energy between the two states. Here, $\hat{e}\cdot\mathbf{p}_{cv}$ is the bulk momentum matrix element. The linewidth functions due to inhomogeneous broadening $D()$ and homogeneous broadening $L()$ are

expressed as

$$D(E, E_{cv}^{ij}) = \frac{1}{\sqrt{2\pi(\sigma_c^2 + \sigma_v^2)}} \exp\left[-(E - E_{cv}^{ij})^2/2(\sigma_c^2 + \sigma_v^2)\right],$$
$$L(E, \hbar\omega) = \frac{\Gamma_{cv}}{\pi} \frac{1}{\Gamma_{cv}^2 + (E - \hbar\omega)^2}, \qquad B_0 = \frac{n_a^2 \omega^2}{\pi^2 \hbar c^2}, \qquad C_0 = \frac{\pi e^2}{n_a c \epsilon_0 m_0^2 \omega}, \quad (2.3)$$
$$f_{c,i} = \frac{1}{1 + \exp\left(\frac{E_{c,i} - F_c}{kT}\right)}, \qquad f_{v,j} = \frac{1}{1 + \exp\left(\frac{E_{v,j} - F_v}{kT}\right)}$$

where the linewidth Γ_{cv} accounts for carrier scattering processes, and n_a is the refractive index for the active region. Figure 2.4(a) and 2.4(b) show the carrier-dependent TE-polarized (electric field normal to growth direction) material gain and spontaneous emission rate, respectively, at $T = 300\,\mathrm{K}$.

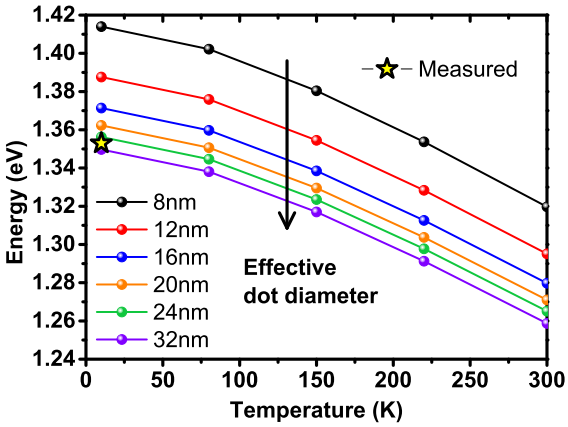


(a) (b)

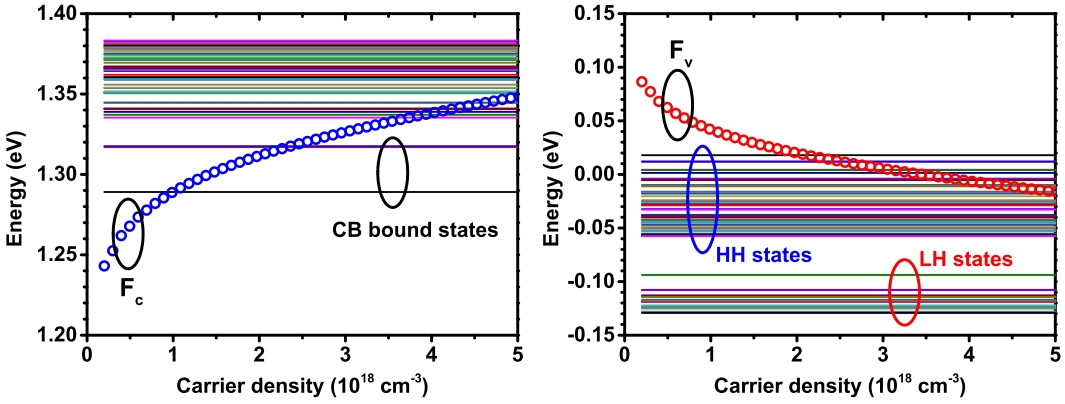
Figure 2.2: (a) Wavefunctions of the first two conduction band states (CB1 and CB2) and the first two heavy-hole states (HH1 and HH2). The wavefunctions are shown at the isosurface of $|\Psi|^2 = 0.5|\Psi|^2_{\max}$. The blue disks are the ten-fold vertically-correlated submonolayer quantum dots, assuming no lateral coupling. (b) Wavefunction of the conduction band ground state (CB1), considering lateral coupling. The wavefunction is shown at the isosurface of $|\Psi|^2 = 0.16|\Psi|^2_{\max}$. In this example, the quantum-dot 2D fill factor is 62.8% and the 2D dot density is 2×10^{11} cm^{-2}.

2.3 Spontaneous Emission Coupling in Micro and Nanocavities

Since the spontaneous emission is affected by the vacuum-field fluctuation and the interaction between the emitter and the optical modes, by modi-

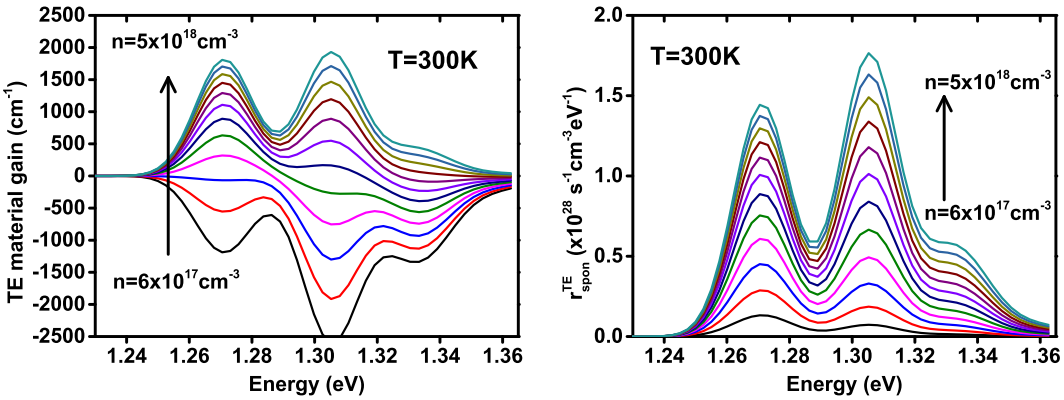


(a)



(b) (c)

Figure 2.3: (a) Ground state transition energies as functions of the temperature for different effective dot sizes. The measured photoluminescence peak [44] is shown as the star. (b) The quasi-Fermi level F_c for conduction band (blue circles) as a function of injected carrier density. Horizontal lines show fifty conduction band states with bound states circled. (c) The quasi-Fermi level F_v for valence band (red circles) as a function of injected carrier density. Horizontal lines show fifty heavy-hole and twenty light-hole states.



(a) (b)

Figure 2.4: (a) TE-polarized material gain and (b) TE-polarized spontaneous emission rate calculated for the submonolayer quantum dots at $T = 300\,\mathrm{K}$, with carrier densities from $n = 6 \times 10^{17}$ cm^{-3} to $n = 5.4 \times 10^{18}$ cm^{-3}.

fying the radiation environment we can potentially control the spontaneous emission rate. It was discovered by E. Purcell [49] that spontaneous emission is not only an inherent property of the emitter, but also dependent on the density of optical modes. Such enhancement of the spontaneous emission rate is characterized by the Purcell factor

$$F_p = \frac{3Q}{4\pi^2 V_{\mathrm{eff}}} \left(\frac{\lambda}{n_{\mathrm{r}}}\right)^3 \tag{2.4}$$

where Q is the quality factor of the optical cavity, V_{eff} is the effective mode volume, and n_{r} is the refractive index. The assumption to arrive at the Purcell factor given in Eq. (2.4) is that the emitter is a two-level system. Therefore, the Purcell factor contains information about the optical properties (radiation environment) but lacks information on the electronic density-of-states of the emitter.

The spontaneous emission coupling factor (β_{sp} factor) is defined as the ratio between the spontaneous emission coupled into the m-th mode $R_{\mathrm{sp},m}$ and the spontaneous emission coupled into all modes R_{sp} [46, 48, 55]. Gérard et al. [51] showed the relationship between the β_{sp} factor and the Purcell factor as

$$\beta_{\mathrm{sp}} = \frac{F_p/3}{gF_p/3 + 1} \tag{2.5}$$

where g is the mode degeneracy ($g = 2$ for circular pillars and disks [51]). Yamamoto et al. [50] gave an empirical formula as

$$\beta_{\mathrm{sp}} = \frac{\lambda^4}{4\pi^2 V \Delta\lambda \epsilon^{3/2}} \tag{2.6}$$

which is equivalent to the expression derived by Baba et al. [46]. We can see that the above empirical formula does not contain the detailed emission property of the emitter. Due to various linewidth broadening mechanisms, the quantum dot transition linewidth can be comparable to the cavity linewidth, and the detailed electronic density-of-states should be considered. Hence, we start with the discrete-mode spontaneous emission rate that couples into the m-th cavity mode [27, 28, 60], and take into account both the transition linewidth Γ_{cv} and the cavity mode linewidth Γ_m,

$$R_{\mathrm{sp},m} = \frac{2N_{\mathrm{dot}}^{\mathrm{2D}}}{L_z}\frac{2\pi}{\hbar}\int dE \sum_{i,j}\left\{\left|\left\langle \psi_c^i \left| e\mathbf{d}\cdot\frac{\mathbf{E}_m}{2}\right| \psi_v^j\right\rangle\right|^2 \cdot D(E, E_{cv}^{ij})\frac{\Gamma_{cv}+\Gamma_m}{\pi}\frac{f_{c,i}(1-f_{v,j})}{(E-\hbar\omega_m)^2+(\Gamma_{cv}+\Gamma_m)^2}\right\} \tag{2.7}$$

where Γ_m is the half-width half-maximum (HWHM) of the optical density-of-states, determined by the cavity quality factor ($\hbar\omega_m/2\Gamma_m = Q$). Here, $e\mathbf{d}$ is the dipole moment, and $N_{\mathrm{dot}}^{\mathrm{2D}}$ is the 2D density of the QDs. The number "2" before $N_{\mathrm{dot}}^{\mathrm{2D}}$ accounts for the spin degeneracy, and the number "2" below $\mathbf{E}_m$ comes from the phasor representation $\mathbf{E}_m(\mathbf{r},t) = \frac{\mathbf{E}_m}{2}e^{i\omega_m t} + \frac{\mathbf{E}_m^*}{2}e^{-i\omega_m t}$. For simplicity, we omit the inhomogeneous broadening integral $\int dE D(E, E_{cv}^{ij})$ for now (and we can re-include it anytime)

$$\begin{aligned} R_{\mathrm{sp},m} &= \frac{2N_{\mathrm{dot}}^{\mathrm{2D}}}{L_z}\frac{2\pi}{\hbar}\sum_{i,j}\left|\left\langle \psi_c^i \left| e\mathbf{d}\cdot\frac{\mathbf{E}_m}{2}\right| \psi_v^j\right\rangle\right|^2 \frac{\Gamma_{cv}+\Gamma_m}{\pi}\frac{f_{c,i}(1-f_{v,j})}{(E_{cv}^{ij}-\hbar\omega_m)^2+(\Gamma_{cv}+\Gamma_m)^2} \\ &\approx \frac{2N_{\mathrm{dot}}^{\mathrm{2D}}}{L_z}\frac{2\pi}{\hbar}\left[\int_{V_a}\frac{d^3\mathbf{r}}{V_a}\frac{|\mathbf{E}_m(\mathbf{r})|^2}{4}\right]\sum_{i,j}\left\{\left|\left\langle \psi_c^i \left| e\mathbf{d}\cdot\hat{e}\right| \psi_v^j\right\rangle\right|^2 \frac{\Gamma_{cv}+\Gamma_m}{\pi}\cdot \right. \\ &\left. \frac{f_{c,i}(1-f_{v,j})}{(E_{cv}^{ij}-\hbar\omega_m)^2+(\Gamma_{cv}+\Gamma_m)^2}\right\} \end{aligned} \tag{2.8}$$

where we have used the slow-varying approximation of cavity mode $\mathbf{E}_m$ com-

pared to crystal unit cells. The energy confinement factor is rigorously defined [61] as

$$\frac{V_a}{V_{\text{eff}}} = \Gamma_E = \frac{\int_{V_a} d^3\mathbf{r}\epsilon_0 n_a^2 |\mathbf{E}_m(\mathbf{r})|^2/2}{\int_V d^3\mathbf{r}\epsilon_0 n^2(\mathbf{r})|\mathbf{E}_m(\mathbf{r})|^2/2} = \frac{\int_{V_a} d^3\mathbf{r}\epsilon_0 n_a^2 |\mathbf{E}_m(\mathbf{r})|^2/2}{\hbar\omega_m} \tag{2.9}$$

where V_a is the active region volume and V is the total volume. The refractive indices n_a (active region) and n have to be replaced by

$$n^2 \to \frac{1}{2}\left[\left.\frac{\partial\left[\omega' n^2(\omega')\right]}{\partial\omega'}\right|_{\omega'=\omega} + n^2(\omega)\right] \tag{2.10}$$

in order to account for material dispersion and the negative permittivities of many noble metals at optical frequencies. The interaction matrix in Eq. (2.8) can be rewritten as

$$\left|\left\langle \psi_c^i \left| e\mathbf{d}\cdot\hat{e}\right| \psi_v^j \right\rangle\right|^2 = |M_{\text{env}}^{ij}|^2 \left|e\mathbf{d}_{cv}\cdot\hat{e}\right|^2 = |M_{\text{env}}^{ij}|^2 \frac{e^2}{m_0^2\omega_{ij}^2}\left|\hat{e}\cdot\mathbf{p}_{cv}^{ij}\right|^2 \tag{2.11}$$

The discrete-mode spontaneous emission rate becomes

$$R_{\text{sp},m} = \frac{2N_{\text{dot}}^{\text{2D}}}{L_z}\frac{2\pi}{\hbar}\left(\frac{\Gamma_E\hbar\omega_m}{2\epsilon_0 n_a^2 V_a}\right)\sum_{i,j}\left\{\frac{e^2}{m_0^2\omega_{ij}^2}|M_{\text{env}}^{ij}|^2\left|\hat{e}\cdot\mathbf{p}_{cv}\right|^2\frac{\Gamma_{cv}+\Gamma_m}{\pi}\cdot \frac{f_{c,i}(1-f_{v,j})}{(E_{cv}^{ij}-\hbar\omega_m)^2+(\Gamma_{cv}+\Gamma_m)^2}\right\} \tag{2.12}$$

Considering $\omega_m \approx \omega_{ij}$ due to the narrow linewidth for the discrete mode, and splitting the Lorentzian function as the convolution of two Lorentzian functions

$$\begin{aligned} R_{\text{sp},m} &= \frac{2N_{\text{dot}}^{\text{2D}}}{L_z}\frac{2\pi}{\hbar}\sum_{i,j}\left\{\frac{\hbar e^2}{2\epsilon_0 n_a^2 V_{\text{eff}} m_0^2\omega_m}|M_{\text{env}}^{ij}|^2\left|\hat{e}\cdot\mathbf{p}_{cv}\right|^2 f_{c,i}(1-f_{v,j})\cdot \right. \\ &\qquad \left. \int d(\hbar\omega)\frac{\Gamma_{cv}/\pi}{(E_{cv}^{ij}-\hbar\omega)^2+\Gamma_{cv}^2}\frac{\Gamma_m/\pi}{(\hbar\omega-\hbar\omega_m)^2+\Gamma_m^2}\right\} \\ &= \int d(\hbar\omega)\left\{\left[\frac{2N_{\text{dot}}^{\text{2D}}}{L_z}\frac{cC_0}{V_{\text{eff}}n_a}\sum_{i,j}|M_{\text{env}}^{ij}|^2\left|\hat{e}\cdot\mathbf{p}_{cv}\right|^2 f_{c,i}(1-f_{v,j})\frac{\Gamma_{cv}/\pi}{(E_{cv}^{ij}-\hbar\omega)^2+\Gamma_{cv}^2}\right]\cdot \right. \\ &\qquad \left. \frac{\Gamma_m/\pi}{(\hbar\omega-\hbar\omega_m)^2+\Gamma_m^2}\right\} \end{aligned} \tag{2.13}$$

where we have used the expression of C_0 in Eq. (2.3).

We notice that the term in the square bracket in Eq. (2.13) is proportional to the free-space spontaneous emission in Eq. (2.2). Therefore, we can write

$$\begin{aligned} R_{\mathrm{sp},m} &= \int d(\hbar\omega) \frac{c/(V_{\mathrm{eff}} n_a)}{B_0} r_{\mathrm{spon}}(\hbar\omega) \frac{\Gamma_m/\pi}{\Gamma_m^2 + (\hbar\omega - \hbar\omega_m)^2} \\ &\approx D_{\mathrm{cav}}(\hbar\omega_m) \int d(\hbar\omega) r_{\mathrm{spon}}(\hbar\omega) \frac{\Gamma_m/\pi}{\Gamma_m^2 + (\hbar\omega - \hbar\omega_m)^2} \end{aligned} \tag{2.14}$$

From Eq. (2.14) we see that the discrete-mode spontaneous emission rate ($\mathrm{s}^{-1}\mathrm{cm}^{-3}$) reduces to an overlap integral between the free-space spontaneous emission spectrum ($\mathrm{s}^{-1}\mathrm{cm}^{-3}\mathrm{eV}^{-1}$) and the photon density spectrum (eV^{-1}) for the m-th mode, expressed by a Lorentzian. The prefactor D_{cav} keeps the unit consistent and is dependent on the resonance wavelength for the m-th mode. By the comparison in Eq. (2.14), we see the prefactor D_{cav} is closely related to the Purcell factor as

$$\begin{aligned} D_{\mathrm{cav}}(\hbar\omega_m) &= \frac{c/(V_{\mathrm{eff}} n_a)}{B_0} = \frac{\pi^2 \hbar c^3}{V_{\mathrm{eff}} n_a^3 \omega_m^2} = \frac{\hbar\omega_m}{8\pi V_{\mathrm{eff}}} \left(\frac{\lambda_m}{n_a}\right)^3 = \frac{2\Gamma_m Q}{8\pi V_{\mathrm{eff}}} \left(\frac{\lambda_m}{n_a}\right)^3 \\ &= \pi\Gamma_m \left[\frac{Q}{4\pi^2 V_{\mathrm{eff}}} \left(\frac{\lambda_m}{n_a}\right)^3\right] = \pi\Gamma_m \frac{F_p}{3} \end{aligned} \tag{2.15}$$

where the square bracket is exactly the same as the Purcell factor in Eq. (2.4). Therefore, by definition we can write the β_{sp} factor for the m-th cavity mode as

$$\begin{aligned} \beta_{\mathrm{sp},m} &= \frac{R_{\mathrm{sp},m}}{R_{\mathrm{sp}}} = \frac{R_{\mathrm{sp},m}}{g R_{\mathrm{sp},m} + R_{\mathrm{sp,cont}}} \\ &= \frac{D_{cav}(\hbar\omega_m) \int d(\hbar\omega) r_{\mathrm{spon}}(\hbar\omega) \frac{\Gamma_m/\pi}{\Gamma_m^2 + (\hbar\omega - \hbar\omega_m)^2}}{g D_{cav}(\hbar\omega_m) \int d(\hbar\omega) r_{\mathrm{spon}}(\hbar\omega) \frac{\Gamma_m/\pi}{\Gamma_m^2 + (\hbar\omega - \hbar\omega_m)^2} + R_{\mathrm{spon,cont}}} \end{aligned} \tag{2.16}$$

where $R_{\mathrm{sp,cont}}$ accounts for the coupling into the continuum modes of the cavity, and g accounts for the degeneracy of the m-th mode. If the cavity supports other discrete modes, those can be lumped into $R_{\mathrm{sp,cont}}$ as well. Combining Eq. (2.15) and Eq. (2.16), we obtain the simplified form of β_{sp}

factor as

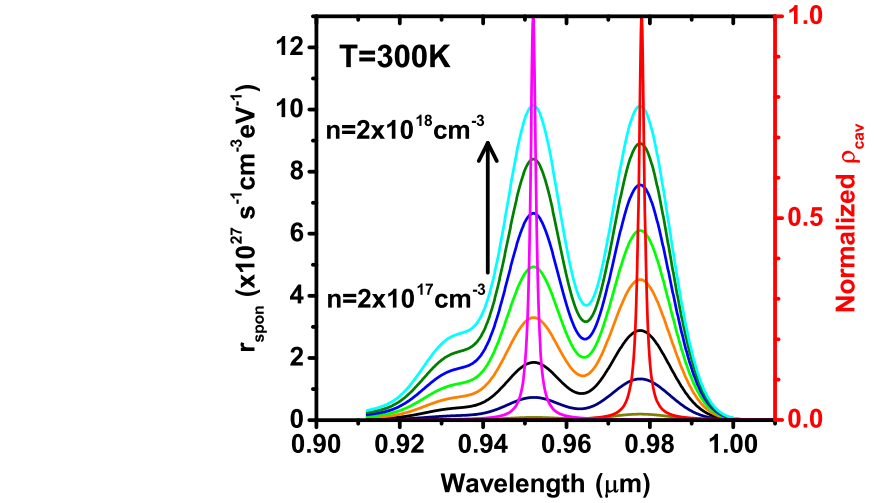
$$\beta_{\mathrm{sp},m} = \frac{\gamma F_p/3}{\gamma g F_p/3 + 1} \tag{2.17}$$

We can see the expression is similar to the empirical formula in Eq. (2.5). But now the β_{sp} factor considers not only the radiation environment (photon density-of-states), but also the radiation property of the emitter (electronic density-of-states), which is embedded in the unitless parameter γ as

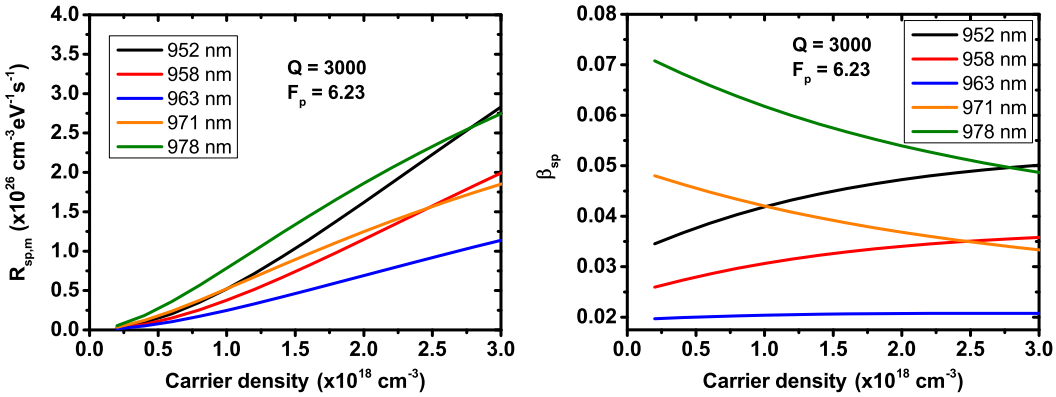
$$\begin{aligned} \gamma &= \frac{1}{R_{\mathrm{sp,cont}}} \int d(\hbar\omega) r_{\mathrm{spon}}(\hbar\omega) \frac{\Gamma_m^2}{\Gamma_m^2 + (\hbar\omega - \hbar\omega_m)^2} \\ &= \tau_{\mathrm{sp,cont}} V_a \int d(\hbar\omega) r_{\mathrm{spon}}(\hbar\omega) \frac{\Gamma_m^2}{\Gamma_m^2 + (\hbar\omega - \hbar\omega_m)^2} \end{aligned} \tag{2.18}$$

To evaluate the spontaneous emission coupled into the continuum modes is nontrivial not only because the continuum electric field mode has to be used in the optical matrix element, but also because the continuum photon density-of-states $\rho_{\mathrm{cont}}(\hbar\omega)$ is needed. Furthermore, the amount of coupling depends on the injected carrier density. In this case, we can replace $R_{\mathrm{sp,cont}}$ by a phenomenological lifetime $\tau_{\mathrm{sp,cont}}$ for an active region volume of V_a.

Figure 2.5 shows the theoretical free-space SE of the SML QDs at different injection carrier densities, as well as illustrations of the photon density-of-state for cavity modes aligned with the QD ground-sate and excited-state transitions, shown as the red ($\lambda_{\mathrm{peak}} = 978\,\mathrm{nm}$) and magenta ($\lambda_{\mathrm{peak}} = 952\,\mathrm{nm}$) curves, respectively. Figure 2.5(b) shows the carrier-dependent SE coupled to a specific cavity mode ($R_{\mathrm{sp},m}(\mathrm{n})$) and Fig. 2.5(c) shows the carrier-dependent $\beta_{\mathrm{sp},m}$ factor for a cavity mode when such a mode is resonating at different wavelengths. We can see that the alignment between the cavity resonance and the QD emission peak has major effects on both the $R_{\mathrm{sp},m}$ and $\beta_{\mathrm{sp},m}$. The $R_{\mathrm{sp},m}$ at different cavity resonances always increases with carrier density n, but the increasing rate is different. The ground-state SE at 978 nm first increases with n then saturates, while the excited-state SE at 952 nm appears later but takes over. As a result, the β_{sp} factor for the 978-nm cavity mode decreases with n because of the decreasing amount of the portion of the total SE that is coupled to this mode. This tells us that the positioning the cavity resonance is very critical for QD VCSELs.



(a)



(b) (c)

Figure 2.5: (a) Theoretical submonolayer (SML) quantum dot (QD) spontaneous emission (SE) at different injection carrier densities at $T = 300\,\mathrm{K}$. The red curve is the normalized photon density-of-state (DOS) for a $\lambda = 978\,\mathrm{nm}$ cavity mode aligned with the QD ground-state transition. The magenta curve is the normalized photon DOS for a $\lambda = 952\,\mathrm{nm}$ cavity mode aligned with the QD excited-state transition. (b) The SE coupled to the cavity mode and (c) the β_{sp} factor at different cavity resonances.

2.4 Size-Dependent Cavity Properties and Lasing Behavior

Solving the cavity properties of the microlaser usually requires 3D numerical methods such as finite-element method (FEM) and finite-difference time-domain method (FDTD). However, due to the metallic dispersion and the thin DBR layers, the computational cost for 3D methods is expensive. Alternatively, we can solve the 2D transverse waveguide modes and the effective indices n_{eff} from the equations derivable from Maxwell's equations,

$$\left[\nabla_t^2 + n^2(\boldsymbol{\rho})k_0^2\right] \begin{bmatrix} E_z \\ H_z \end{bmatrix} = k_z^2 \begin{bmatrix} E_z \\ H_z \end{bmatrix} = n_{\text{eff}}^2 k_0^2 \begin{bmatrix} E_z \\ H_z \end{bmatrix} \tag{2.19}$$

where the $k_0 = \omega\sqrt{\epsilon_0\mu_0}$ is the free-space wave number, and k_z is the wave number in the propagation direction. Here, $n(\boldsymbol{\rho})$ is the transverse profile for the refractive index. The effective index n_{eff} solved from Eq. (2.19) includes both the material dispersion and the modal dispersion due to size dependence. We use n_{eff} for each layer in the 1D transfer matrix method, which calculates the longitudinal field distribution, as well as the reflection spectra from the top and bottom mirrors. The cavity resonance condition (round-trip phase matching condition) is obtained from the 1D Fabry-Pérot model, which has been shown to be in excellent agreement with the full-structure FDTD simulation and the experimental data [25].

Figure 2.6(a) shows the longitudinal profile of the electric field across the device layer structure at resonance for a 2-μm-diameter device. The magnified view across the cavity region in Fig. 2.6(b) helps us verify the design of the layer structure in terms of the optical field overlap with the active region, which strongly affects the confinement factor Γ and the threshold gain.

The key size-dependent parameters to be obtained from the cavity model are the fundamental mode (HE_{11}) lasing wavelength λ_r, quality factor Q, photon lifetime τ_p, mirror loss α_m, and confinement factor Γ_E. Size-dependent resonance wavelength, photon lifetime and the quality factor are calculated for the HE_{11} mode and shown in Fig. 2.7(a) and Fig. 2.7(b). The effective mode volume in terms of $(\lambda_r/n_r)^3$ is calculated based on the confinement factor obtained from the Fabry-Pérot model.

To study the light output power versus current behavior of the metal-cavity

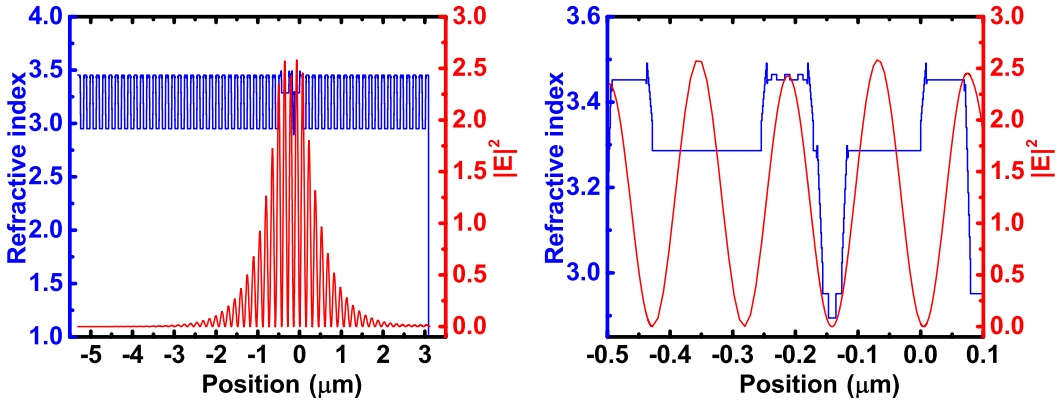


Figure 2.6: (a) Longitudinal profiles of the electric-field squared-magnitude (red) and the material indices across the device layer structure. (b) Magnified view of the device cavity showing the overlap between the optical field and the quantum-dot active region.

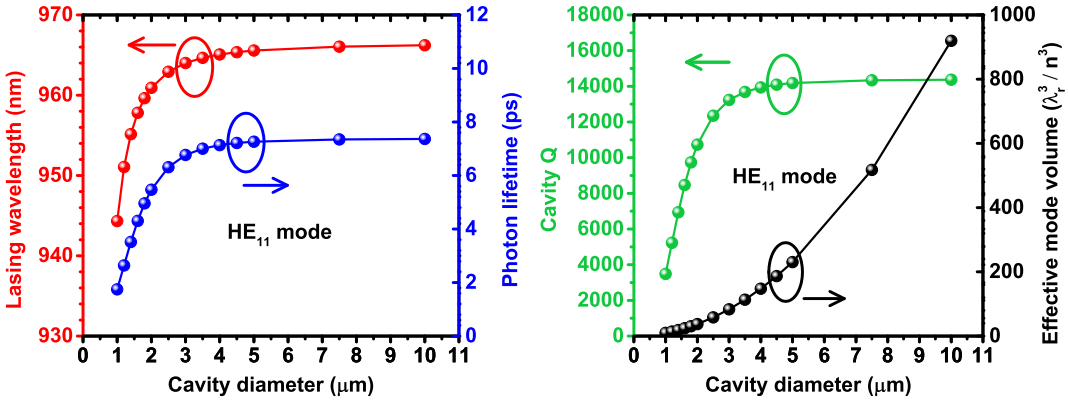


Figure 2.7: (a) Calculated lasing wavelength and the photon lifetime as a function of the cavity diameter for the fundamental HE_{11} transverse mode. (b) Calculated cavity quality factor and effective mode volume as a function of the cavity diameter for the HE_{11} transverse mode.

microlasers, we use the rate equations [48, 55, 62] of the carrier density n and the photon density S

$$\begin{aligned} \frac{dn}{dt} &= \eta_i \frac{I - I_l(n)}{qV_a} - (An + Cn^3) - R_{\text{sp}}(n) - v_g g(n) S, \\ \frac{dS}{dt} &= \Gamma_E v_g g(n) S - \frac{S}{\tau_p} + \Gamma_E \beta_{\text{sp},m}(n) R_{\text{sp}}(n) \end{aligned} \tag{2.20}$$

where the material gain $g(n)$ and spontaneous emission rate $R_{\text{sp}}(n)$ are obtained from our QD gain model which accounts for the inhomogeneous broadening effect. The photon lifetime τ_p and confinement factor Γ_E are obtained from our cavity transfer matrix model with effective indices. The rigorous definition for the energy confinement factor Γ_E is given in Eq. (2.9). Here, η_i is the intrinsic quantum efficiency, and v_g is the group velocity. And A is the surface recombination coefficient and C is the Auger coefficient. The A coefficient is dependent on the surface-to-volume ratio (A_a/V_a) of the active region

$$A = \frac{A_a}{V_a} v_s = \frac{4}{D} v_s \quad \text{(cylindrical)} \tag{2.21}$$

where v_s is the surface recombination velocity and D is the cavity diameter. Before the laser threshold is reached, as more carriers are injected, the quasi-Fermi level gets closer to the quantum dot barrier and more carriers leak out of the quantum dots without radiative recombination. Therefore such carrier leakage has an important effect on the threshold carrier density and threshold current. We consider such leakage current [63] as

$$I_l(n) = I_{l0} \cdot \exp\left(\frac{E_{g,\text{barrier}} - [F_c(n) - F_v(n)]}{kT}\right) \tag{2.22}$$

where $E_{g,\text{barrier}}$ is the band gap of the quantum dot barrier.

The spontaneous emission coupling factor $\beta_{\text{sp},m}(n)$ into the lasing mode in Eq. (2.20) is obtained from the previous section, and is dependent on the carrier density. The free-space spontaneous emission r_{spon} increases with carrier density as shown in Fig. 2.4, thus the γ parameter in Eq. (2.18) is also carrier-dependent. Although the continuum-mode spontaneous emission lifetime $\tau_{\text{sp,cont}}$ decreases with carrier density, the increase of the integral in Eq. (2.18) is faster if the free-space emission peak aligns with the cavity peak. Then, both the γ parameter in Eq. (2.18) and the $\beta_{\text{sp},m}$ factor in Eq. (2.17) increase with the carrier density.

After the rate equations are solved for each given injection current, we can obtain the output light power as

$$P = \beta_{c1} \hbar\omega \frac{V_a}{\Gamma_E} v_g \alpha_m S + \beta_{c2} \hbar\omega R_{\text{sp}} V_a \tag{2.23}$$

The first term is the contribution from stimulated emission, which depends

on the photon density S and the photon escape rate $v_g\alpha_m$, and V_a/Γ_E is the effective mode volume. The second term is the contribution from spontaneous emission. β_{c1} and β_{c2} are the coupling efficiencies to the detector for the stimulated emission and spontaneous emission, respectively, which account for the loss of light power from the experimental setup.

Figure 2.8 shows a comparison between the theoretical and experimental light output power versus current (L-I) curves for metal-cavity SML QD surface-emitting microlasers with different device diameters at T = 300 K. The L-I curves are measured under pulsed mode to eliminate laser self-heating for the ease of analyzing the size-dependent lasing characteristics. We can see that, as the device size reduces, the L-I curves exhibit a more obvious turn-on behavior below threshold. Such turn-on behavior was also observed for metal-cavity quantum-well microlasers in [62]. This "upward-bending" L-I behavior below threshold can be explained by the increasing amount of spontaneous emission coupling into the cavity mode as the injection current increases.

In the rate-equation model, the surface recombination velocity v_s is set to 6×10^5 cm/s and the Auger coefficient C is set to 1×10^{-29} cm^6/s. However, we do not see significant change of the L-I curves when we vary v_s in the 10^5 cm/s range or vary C in the $10^{-30} \sim 10^{-29}$ cm^6/s range. Due to the 1% pulsed mode operation, the thermal effect is negligible, and most carrier-dependent laser characteristics are pinned upon threshold. Table 2.1 summarizes the carrier density, material gain, threshold current density, and leakage current density at threshold. The threshold material gain $g_{\rm th}$ increases as size reduces because of both larger radiation loss and larger material loss from metal. Figure 2.9(a) shows the threshold carrier density and the threshold current density as functions of the cavity diameter. The threshold carrier density $n_{\rm th}$ is dependent on $g_{\rm th}$ as well as the lasing wavelength. Since the quantum dot emission bandwidth is narrow, the alignment between the cavity peak and gain peak is important. Furthermore, because $n_{\rm th}$ is larger for smaller devices, the quasi-Fermi levels are closer to the quantum dot barrier, and the leakage current density at threshold $J_{l,\rm th}$ is also larger. As a result of larger $n_{\rm th}$ and larger $J_{l,\rm th}$, we can see that the threshold current density also increases when we reduce the device size. Figure 2.9(b) shows the Purcell factor calculated with Eq. (2.4), as well as the $\beta_{\rm sp}$ factor at threshold extracted from the fitting of the L-I curves with the rate-equation model. We

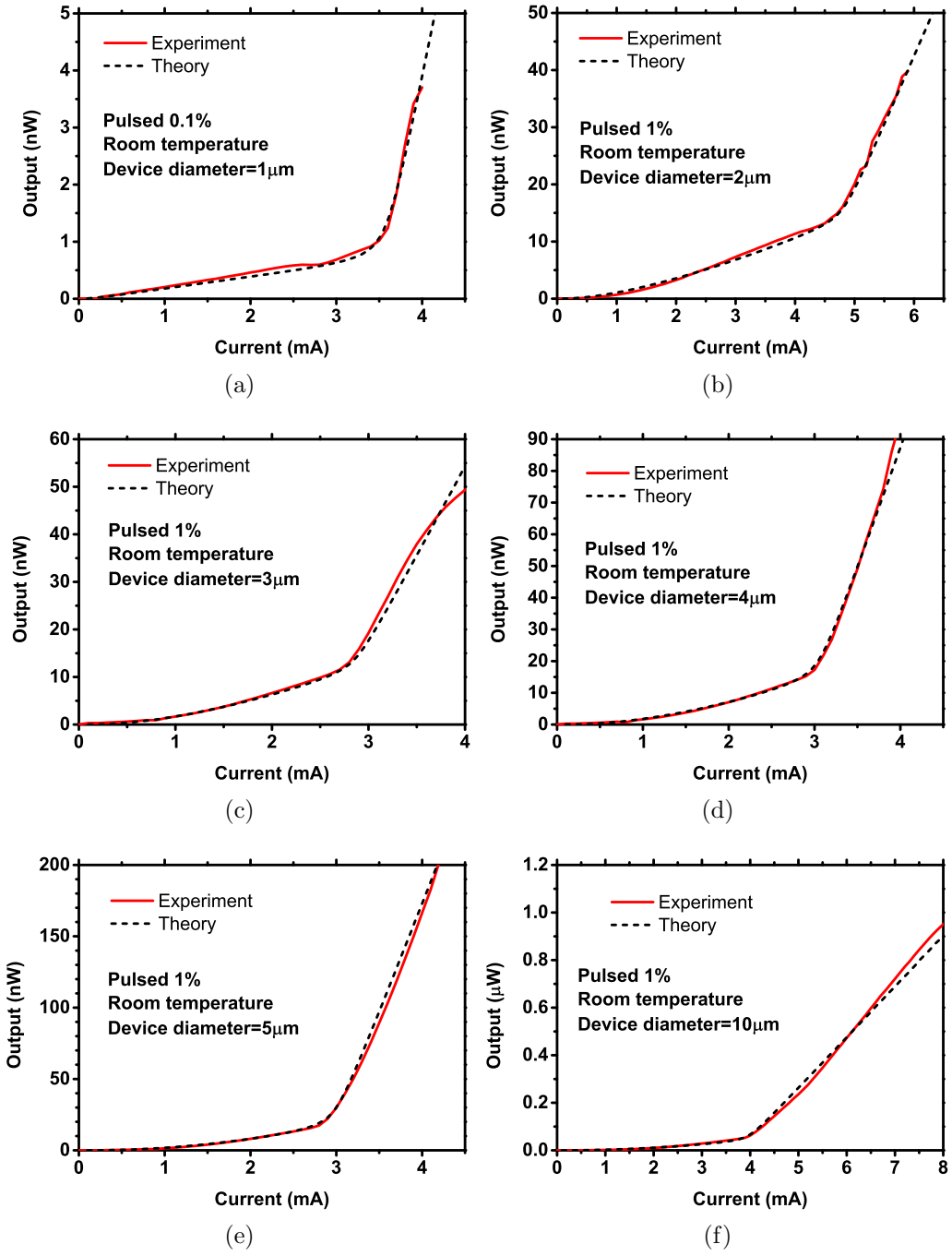
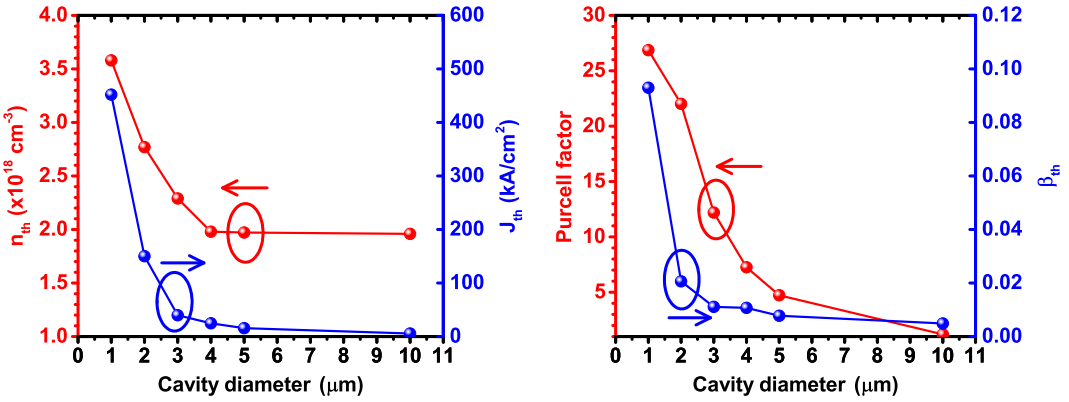


Figure 2.8: Theoretical and measured light output power versus current (L-I) curves for submonolayer quantum-dot metal-cavity surface-emitting microlasers with different device diameters at $T = 300\,\mathrm{K}$. The turn-on behavior below lasing threshold is explained by the increasing β_{sp} factor with carrier injection, i.e., increasing amount of spontaneous emission coupled into the cavity mode.



(a) (b)

Figure 2.9: (a) Threshold carrier density and the threshold current density solved from the rate equations as functions of the cavity diameter. (b) Purcell factor and the β_{sp} factor at threshold as functions of the cavity diameter.

Table 2.1: Size-dependent laser characteristics extracted from the rate-equation model.

Device diameter (μm)	Threshold carrier density n_{th} ($\times 10^{18}$ cm^{-3})	Threshold material gain g_{th} (cm^{-1})	Leakage current density at threshold $J_{l,\mathrm{th}}$ (kA/cm^2)	Threshold current density J_{th} (kA/cm^2)
10	1.96	160.72	1.488	5.730
5	1.97	163.01	6.305	15.83
4	1.98	165.96	14.21	24.89
3	2.29	175.02	25.96	39.82
2	2.77	216.83	128.7	150.2
1	3.58	685.68	411.4	452.1

can see that the portion of spontaneous emission coupled to the lasing mode increases drastically when the cavity size reduces. The parameters used in our model are summarized in Table 2.2.

2.5 Summary

Metal-cavity submonolayer quantum-dot surface-emitting lasers are demonstrated under electrical injection at room temperature with device radius down to 0.5 μm. We have developed a comprehensive model for analyzing

Table 2.2: Parameters used in our theoretical model.

Name and symbol	Value	Name and symbol	Value
Surface recombination velocity v_s	6×10^5 cm/s [48]	Injection efficiency η_i	0.4∼0.7
Auger coefficient C	1×10^{-29} cm^6/s [48]	Quantum dot diameter	20 μm
Coupling efficiency β_{c1}	0.05∼0.15	Quantum dot 2D density	1×10^{11} cm^{-2}
Coupling efficiency β_{c2}	0.005∼0.05	Series leakage current parameter I_{l0}	100∼300 mA
Homogeneous broadening linewidth Γ_{cv}	15 meV [55]	Inhomogeneous broadening linewidth for conduction band σ_c	15 meV [55]
Inhomogeneous broadening linewidth for valence band σ_v	2 meV [55]		

the size-dependent device performance. Our model yields the material gain and the spontaneous emission spectra of submonolayer quantum dots. We derive a rigorous expression for the spontaneous emission coupling factor. The laser cavity properties are solved with the effective index method and the transfer matrix method. For future work, we can extend our cavity model for cases where multiple transverse modes are present. This can be done using the analytical vector mode-matching method, and the mode conversion and mismatch can be included. The resonance wavelength and photon lifetime for each individual mode can be obtained. This then can be followed by a multi-mode rate-equation model to produce the mode-resolved L-I curves. In this work, a single-mode model has shown good prediction of the cavity resonance for our application. With the information on the quantum-dot emission and the optical cavity, a single-mode rate-equation model is used to investigate the laser characteristics. Our theory shows excellent agreement with the experiments and directs our future work toward the miniaturization of metal-cavity lasers.

CHAPTER 3

FABRICATION AND CHARACTERIZATION OF METAL-CAVITY SURFACE-EMITTING MICROLASERS

3.1 Introduction

In this chapter, we will discuss the fabrication steps for surface-emitting micro and nanolasers in detail. The device processing in this work is challenging due to its small size. Both the optical and electrical performance can be significantly affected by the fabrication imperfections. For example, many of the processing steps can cause the device sidewall to be rough and not perfectly vertical, reducing the cavity quality factor. The device top surface can also become non-flat and rough, causing problems for electrical contact as well as optical loss. Also, many processing steps involve high-vacuum, high-temperature, or high-power plasma, which all require robustness of the laser structure. Nonetheless, we demonstrate lasing behavior of our quantum-well (QW) and quantum-dot (QD) surface-emitting lasers at room temperature under electrical injection.

For our working devices, we characterize the light output versus current (L-I) and the current versus voltage (I-V) behavior. We study the threshold condition for various device sizes. From the current-dependent emission spectra, we study the thermal effects on the cavity resonance and the emission intensity. We verify the emission beam pattern from the near-field images. We also measure the sub-threshold emission spectra to extract the cavity information. Using our theory developed in Chapter 2, we are able to explain size-dependent spontaneous emission coupling with the cavity modes.

3.2 Fabrication of Surface-Emitting Micro and Nanolasers

The micro and nanolasers in this work are fabricated from top down using plasma etch (dry etch), where a SiN_x hard mask is needed to protect the patterned region from the plasma etch. Figures 3.1(a) to 3.1(d) show the processing steps for patterning the hard mask. We first deposit a uniform SiN_x layer by plasma-enhanced chemical vapor deposition (PECVD). A layer of photoresist is spin-coated and patterned by ultraviolet (UV) photolithography (365 nm) using the Karl Suss MJB3 Contact Mask Aligner. Then a thin layer (around 10 nm) of nickel is vacuum-evaporated using the electron beam, followed by a metal lift-off step. Thus the pattern remains with the nickel mask and can be transfered onto the SiN_x by Freon reactive-ion etch (RIE) using the combination of CHF_3 and O_2. Figure 3.1(e) shows the 1-μm-diameter SiN_x hard mask formed on top of the wafer to be processed.

Figures 3.2(a) to 3.2(e) show the processing steps for forming the laser cavity by plasma etch as follows. We first remove the remaining nickel on top of the SiN_x mask by the TFG nickel etchant. Then we use the inductively-coupled plasma (ICP) RIE with $SiCl_4$/Ar to form the cavity. The result after ICP RIE is shown in Fig. 3.2(c) and we can identify the distributed Bragg reflector (DBR) layers and the cavity region. Afterwards, we remove the remaining SiN_x mask and a 1-μm-diameter laser cavity is shown in Fig. 3.2(e). We etch past the cavity region and stop at around the seventh pair of the bottom DBR. Since the cavity quality is critical to the laser performance, we need precise control of the ICP etch rate and the etch stop. Besides calibrating the etch time, we also have an in-situ reflectometer to monitor the etch depth, as shown in Fig. 3.3(a). The collimated laser beam is incident through the monitor window of the etch chamber onto the sample. The reflected beam is collected by a photodetector. As the DBR lasers are removed in the etch chamber, the reflected power shows an oscillation, which allows us to monitor the etched thickness as time elapses, as shown in Fig. 3.3(b). From the reflected power we can further identify whether we have etched past the cavity region and reached our desired depth.

In order to have electrical injection through the laser top surface and the p-doped region, we need to provide sidewall electrical isolation, for which SiN_x is a good candidate. The refractive index of SiN_x is lower than GaAs,

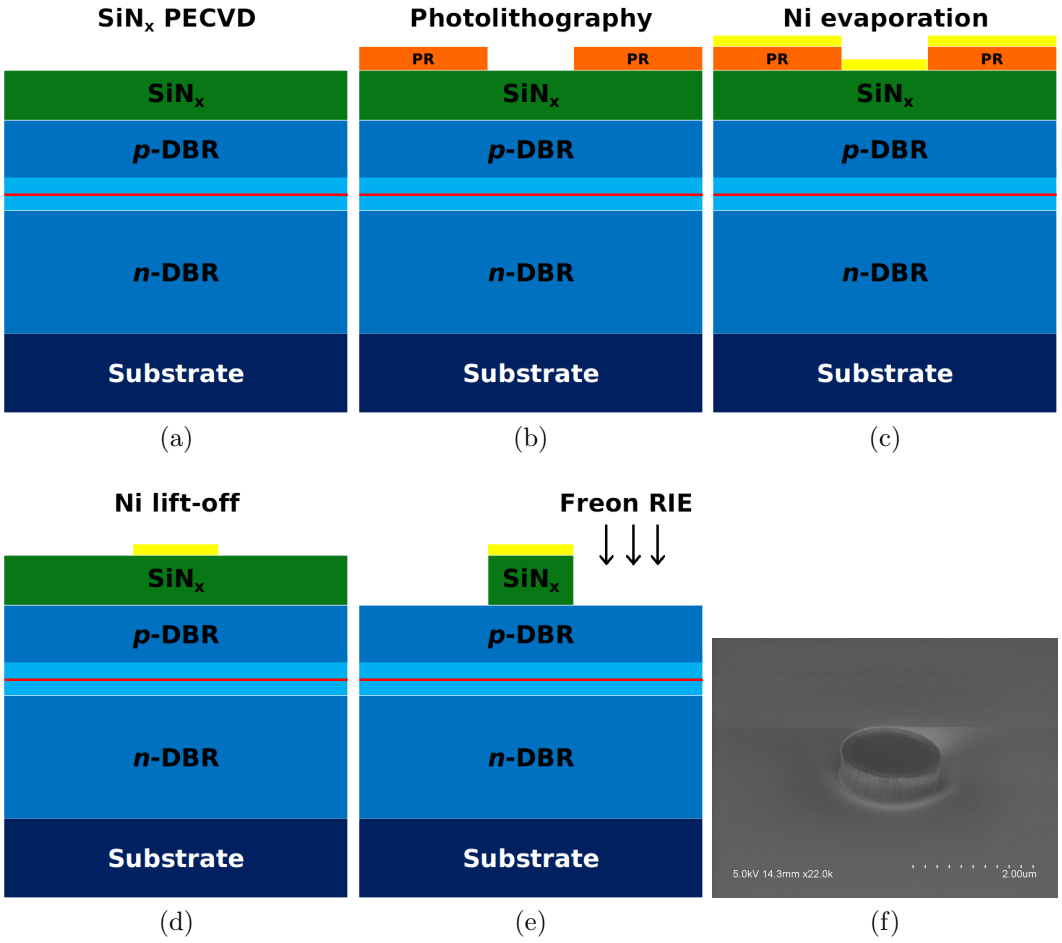


Figure 3.1: Processing steps for patterning the SiN_x hard mask for plasma etch. (a) SiN_x chemical vapor deposition; (b) Photolithography for patterning photoresist (PR); (c) Nickel evaporation; (d) Metal lift-off for nickel mask; (e) Freon reactive-ion etch (RIE); (f) Scanning electron micrograph of the SiN_x hard mask.

which provides optical mode confinement and reduces the field leakage into the sidewall metal and thus the material loss. Figures 3.4(a) to 3.4(c) show the processing steps. We first coat conformally our laser cavity with SiN_x using PECVD. Then we cover the sample with PR and planarize step by step using the UV exposure and PR development until the device top is exposed. Afterwards, we use the Freon RIE with CHF_3 to remove the SiN_x on the device top while the sidewall is protected by the PR. Figure 3.4(d) shows the scanning electron micrograph of a device with sidewall coated by SiN_x while the top surface is exposed for electrical injection and optical emission.

Since the bottom contact is on the *n*-doped GaAs substrate, careful steps

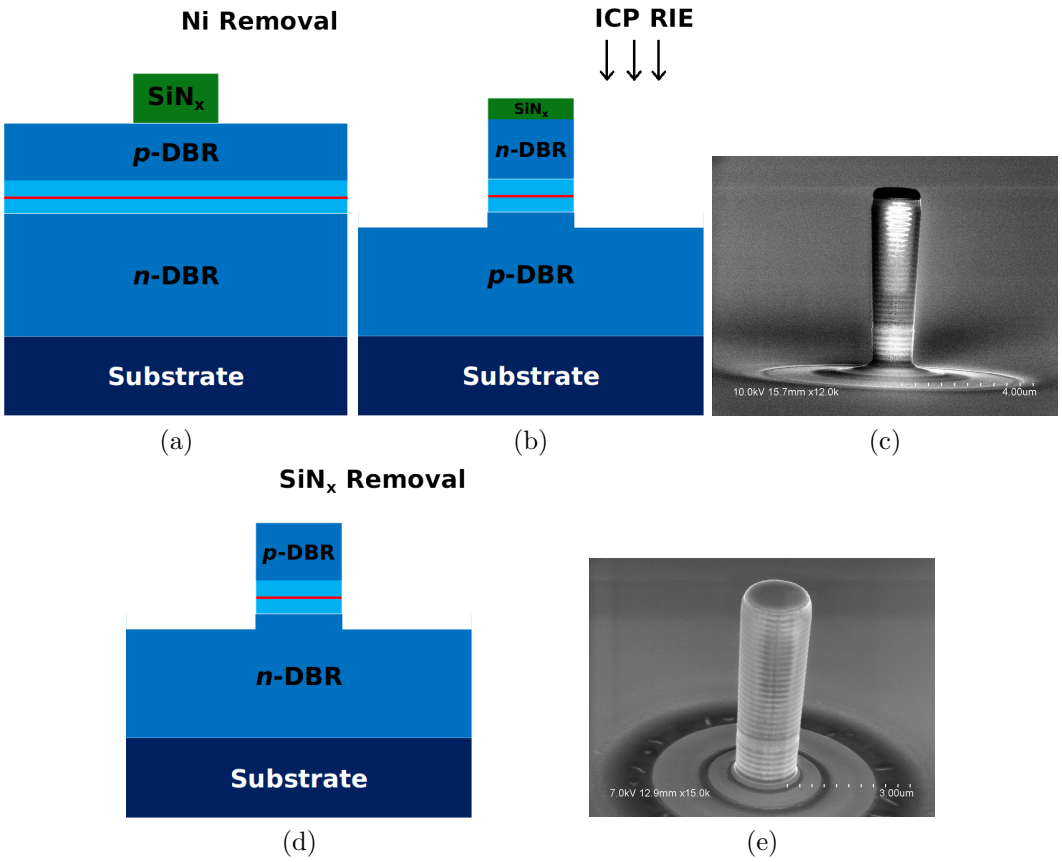
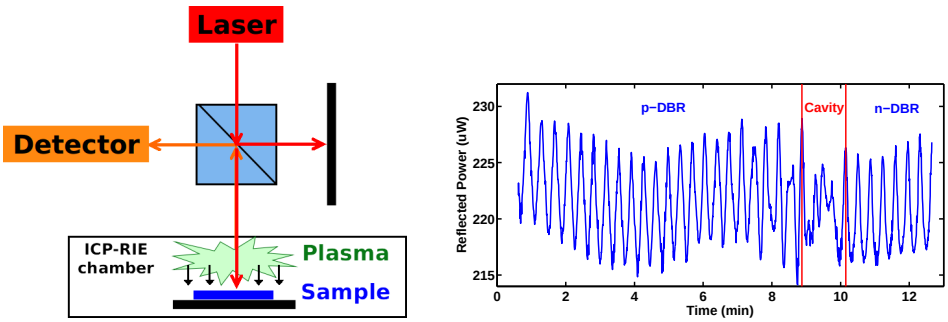


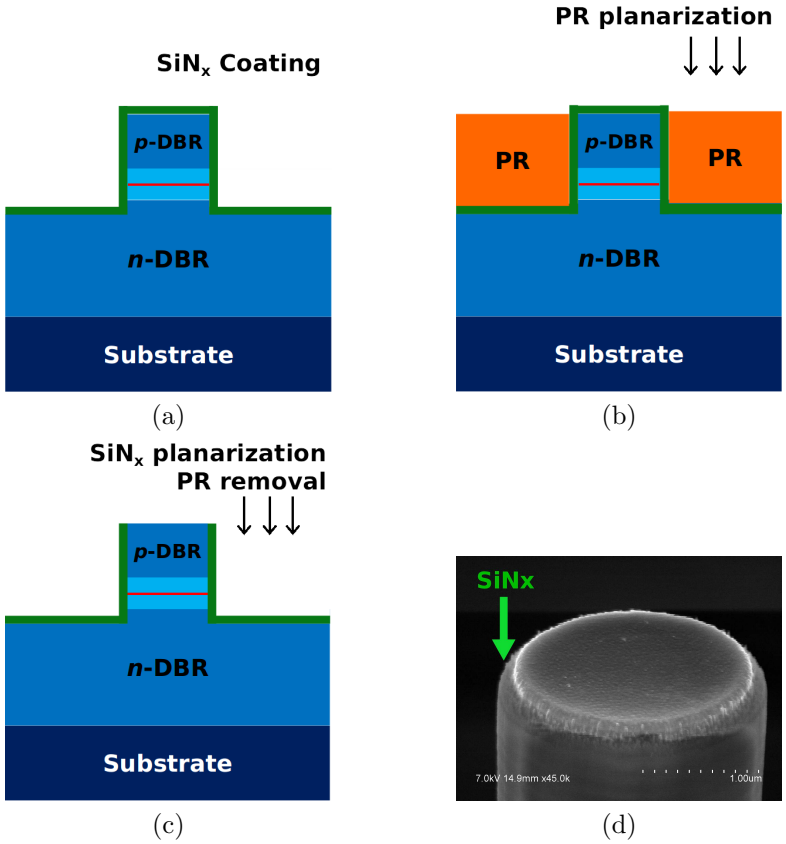
Figure 3.2: Processing steps for forming the laser cavity using plasma etch. (a) Nickel mask removal; (b) Inductively-coupled plasma (ICP) reactive-ion etch (RIE); (c) Scanning electron micrograph of the laser cavity after ICP RIE; (d) SiN_x mask removal; (e) Laser cavity of 1-μm-diameter formed.

are needed to form the ohmic contact. We vacuum-evaporate around 200 nm of AuGe/Ni/Au and anneal at the eutectic temperature around 370 °C, as shown in Fig. 3.5(a). Then we do photolithography to define the pattern for the top contacts. We vacuum-evaporate 5 nm of silver/gold as the top p-contact and do metal lift-off so that the device p-contacts are separated, as shown in Fig. 3.5(b) and Fig. 3.5(c), respectively. However, at this step the n-regions for all devices are still connected. Thus we spin-coat the PR and pattern with the UV lithography for device separation, as shown in Fig. 3.5(d). We then plasma-etch the SiN_x layer using Freon RIE with CF_4/O_2, and chemical-etch the n-DBR using 1:5:5:40 citric acid:H_2O:H_2O_2:H_2SO_4, where the citric acid is mixed from 1 mg:1 mL $C_6H_8O_7$ monohydrate to H_2O. After cleaning the device, we can proceed to testing, as shown in Fig. 3.5(e)



(a) (b)

Figure 3.3: (a) Schematic diagram of the in-situ reflectometer for monitoring the sample etching depth in the plasma chamber. (b) Reflected power collected by the detector as a function of the etching time. We identify the time when the etch depth reaches the cavity region, the bottom distributed Bragg reflector, and our desired etch-stop.



(a) (b)

(c) (d)

Figure 3.4: Processing steps for the sidewall SiN_x coating. (a) Conformal SiN_x coating; (b) Photoresist planarization; (c) Plasma etch of the top SiN_x and cleaning; (d) A sidewall-coated device with top emission window opened.

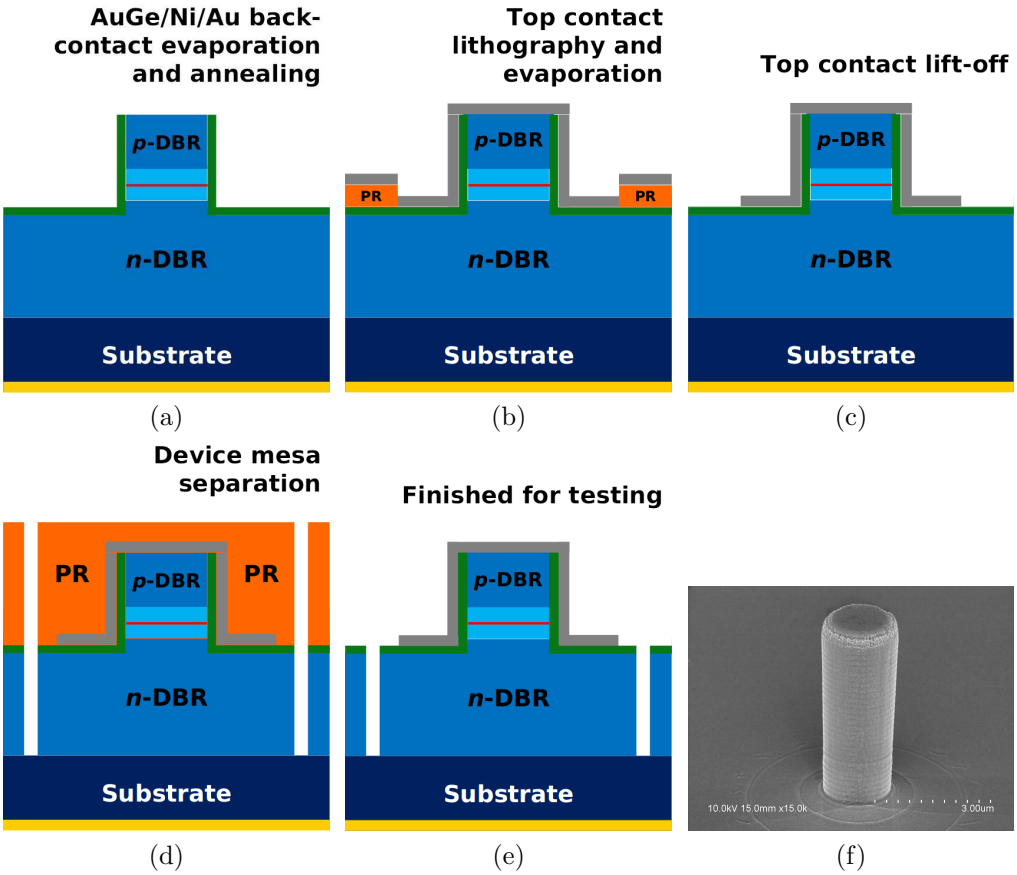


Figure 3.5: Processing steps of forming laser contacts. (a) Evaporation and annealing of the bottom n-contact; (b) Lithography for top p-contacts; (c) Metal lift-off for top contacts; (d) Mesa etch for device separation; (e) Schematic diagram and (f) scanning electron micrograph of a processed 1-μm-diameter metal-cavity surface-emitting laser.

and Fig. 3.5(f).

As a comparison, we also fabricate dielectric-cavity surface-emitting lasers with the same wafer. In this work, we use Benzocyclobutene (BCB) as the cavity cladding. Before forming the dielectric cavity, we first processed the bottom n-contact and the separation of devices, as shown in Fig. 3.6(a). Then we spin-coat the BCB and cure at 250°C, after which BCB becomes solid and stable. Afterwards, we use Freon RIE with CF_4/O_2 to planarization the BCB until the device top surface is exposed. Figure 3.6(c) shows the scanning electron micrograph after BCB planarization. We normally planarize slightly deeper than the device top surface for the better surface quality and contact adhesion. Figure 3.6(d) shows the patterning and the

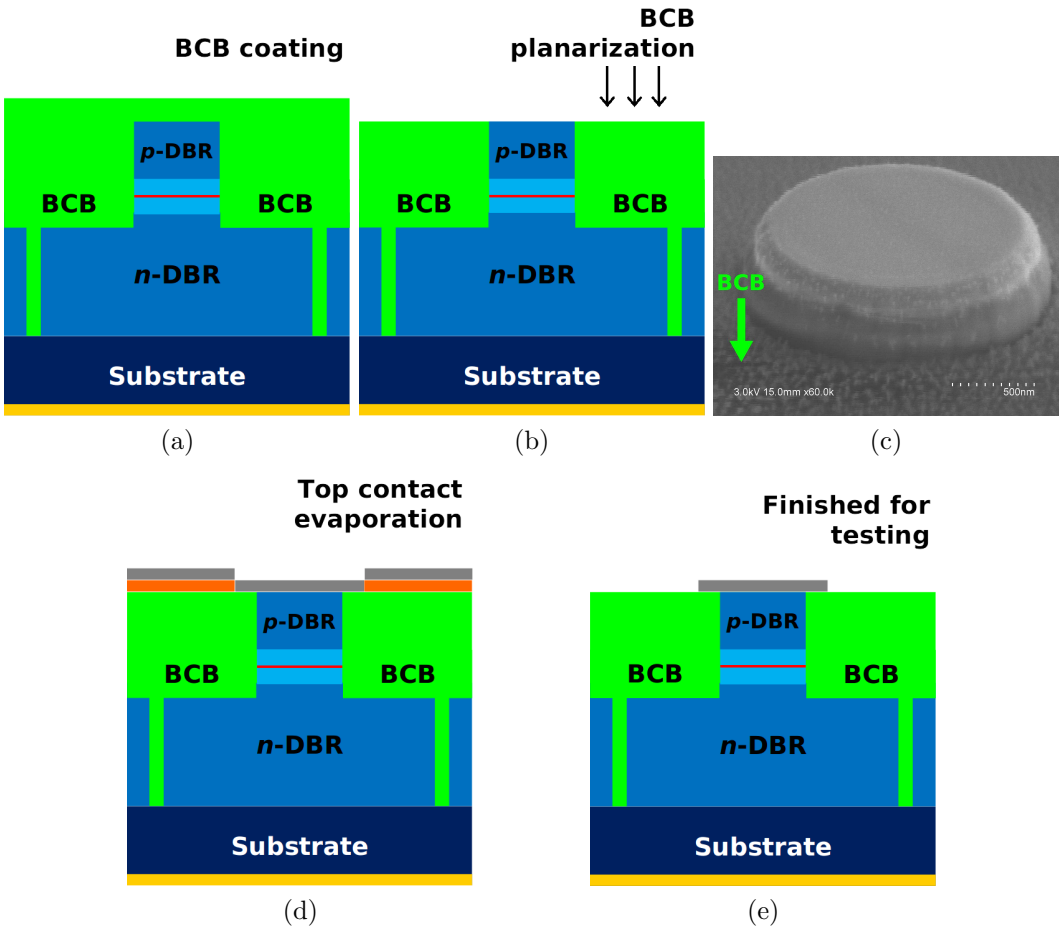


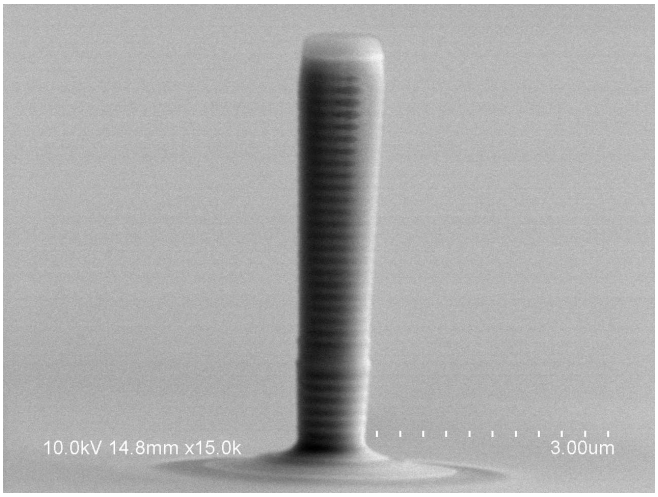
Figure 3.6: Processing steps for dielectric-cavity surface-emitting lasers. (a) Spin-coating and curing of Benzocyclobutene (BCB); (b) Planarization of BCB using Freon reactive-ion etch; (c) Device top surface exposed after planarization; (d) Top contact lithography; (e) Process finished and ready for testing.

evaporation of the top contacts. Figure 3.6(e) shows the schematic of the dielectric-cavity surface-emitting laser ready for testing.

3.3 Characterization of Quantum-Well Surface-Emitting Microlasers

The wafers for our QW surface-emitting lasers are grown on the GaAs substrate using the metal-organic chemical vapor deposition (MOCVD). The active region consists of three 6-nm-thick compressively strained $In_{0.17}Ga_{0.83}As$

QWs separated by 4-nm-thick $GaAs_{0.92}P_{0.08}$ barriers with tensile strain. The photoluminescence of the QWs has a peak wavelength of 970 nm. The $3\lambda/2n_r$ cavity is sandwiched between 23 pairs p-doped and 36 pairs of n-doped DBRs. The DBR consists of alternating $Al_{0.12}Ga_{0.88}As$ and $Al_{0.9}Ga_{0.1}As$ layers. Figure 3.7 shows the scanning electron micrograph of a QW laser cavity formed after the ICP RIE plasma etch, which passes the cavity and down to the seventh pair of the bottom DBR.



(a)

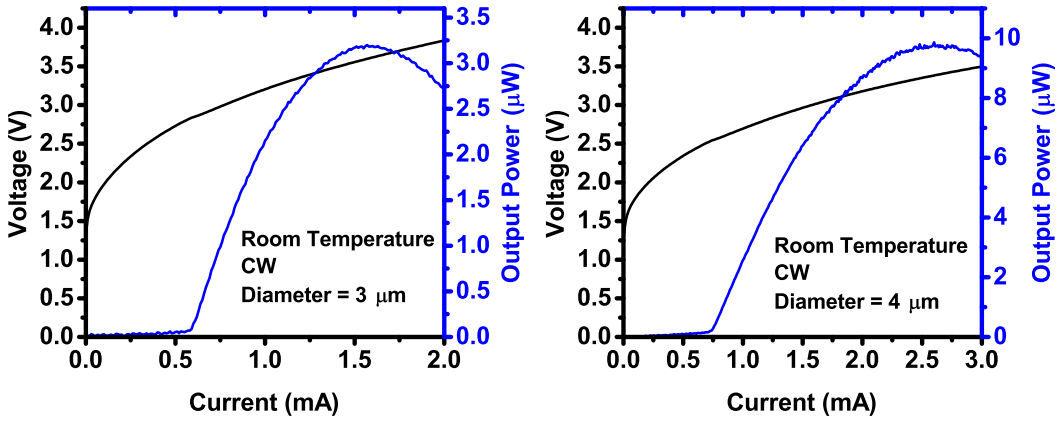
Figure 3.7: Scanning electron micrograph of a 1-μm-diameter quantum-well surface-emitting laser, viewed at a right angle.

Our metal-cavity QW surface-emitting lasers can operate at room temperature under continuous-wave (CW) electrical injection with device diameter down to 3 μm and threshold current being 0.58 mA (8.2 kA/cm^2 threshold current density). Figure 3.8 shows the light output versus current (L-I) and current versus voltage (I-V) characteristics of our devices. The light output is collected by the Lightwave Probe with a ball-lensed pigtail multimode (62.5-μm core) fiber, coupling to the calibrated Newport 818-SL silicon photodetector. The L-I and I-V characteristics are measured by the Hewlett-Packard 4145B Semiconductor Parameter Analyzer. The devices are tested using the Cascade MPS150 probe system. The 1.5 V turn-on voltages in Fig. 3.8 indicate proper electrical performance for the devices as p-i-n diodes. The roll-over behavior of the light output is observed due to the heating of the device active regions under CW operation. Figure 3.8(e) shows the extracted threshold currents I_{th} and threshold current densities J_{th} as functions

of the device diameter. The I_{th} reduces as the device size shrinks, yet the J_{th} increases. One reason for this is the larger detuning between the lasing wavelength and the high-reflection window of the DBR for smaller devices because of the change of the effective index. Thus the quality factor reduces and the radiation loss increases, and the devices have larger threshold gain and threshold carrier density. Thus, even with a smaller active region volume, a small device can have a larger threshold current density. Another reason is that for smaller devices, the optical field has more penetration into the metallic sidewall and more material loss is introduced, resulting in larger thresholds.

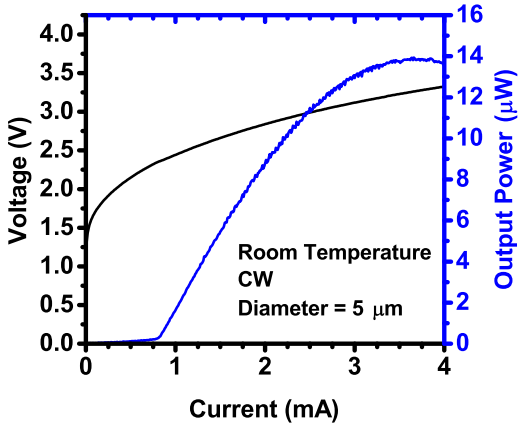
Figure 3.9 shows the near-field images of the emission from a 10-μm-diameter metal-cavity QW laser under various injection currents. The images are taken by the Industrial Vision Source model IV-CCAM2 digital color CCD camera, which covers the visible to infrared spectrum. The light emission is collected through a 50x objective lens, which focuses on the laser top surface. The threshold current for this device is 3.2 mA.

Figure 3.10 shows the emission spectra of the devices under CW injection at room temperature for various device sizes. The light emission is collected using the Lightwave Probe with a ball-lensed pigtail multi-mode (62.5-μm core) fiber. The emission peaks red-shift as more current is injected because of the heat-induced cavity expansion and the refractive index change. We also observe larger mode splitting with smaller sizes. The side-mode suppression is more than 15 dB for the 5-μm-diameter devices, and more than 40 dB for the 4-μm-diameter devices. Devices with diameters smaller than 4 μm can be well considered as single-mode lasers. Figure 3.10(e) shows the current-dependent peak wavelengths of the fundamental cavity mode for various device sizes. At very low current ($I < 1\,\mathrm{mA}$), we observe the mode dispersion due to geometry, since a smaller size supports a smaller peak resonance wavelength. As the current increases, the red-shift of the peak wavelength for a smaller device is much faster. One reason is the larger series resistance for smaller devices, which can be seen from the I-V curves in Fig. 3.8. Thus more heat is generated for smaller devices at the same current. Another reason is that the heat dissipation through the sidewall metal is not efficient enough. Smaller devices have larger surface-to-volume ratio and should be better in terms of sidewall heat dissipation. Yet because of the low thermal conductivity of the SiN_x sidewall buffer layer, there is still a large amount of heat flowing

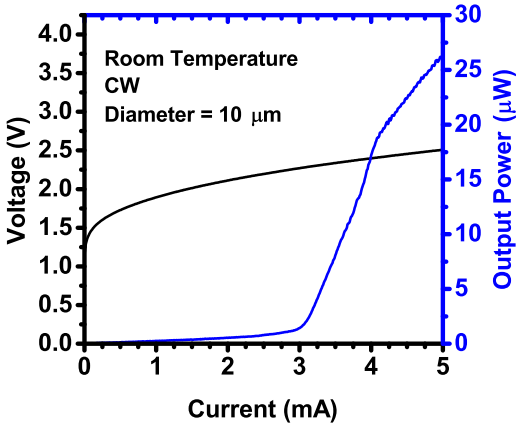


(a)

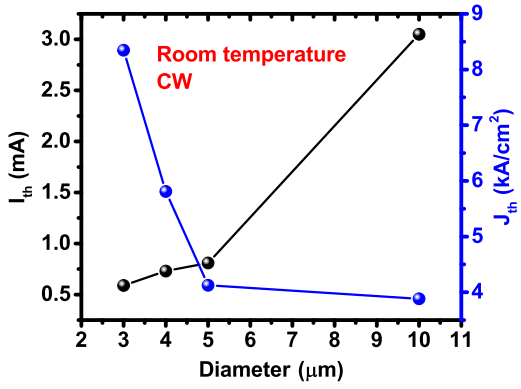
(b)



(c)



(d)



(e)

Figure 3.8: The light output versus current (L-I) and the current versus voltage (I-V) curves for metal-cavity quantum-well surface-emitting lasers operating at room temperature under continuous-wave electrical injection. The cavity diameters are (a) 3 μm, (b) 4 μm, (c) 5 μm, and (d) 10 μm, respectively. (e) Size-dependent laser threshold currents (black) and threshold current densities (blue).

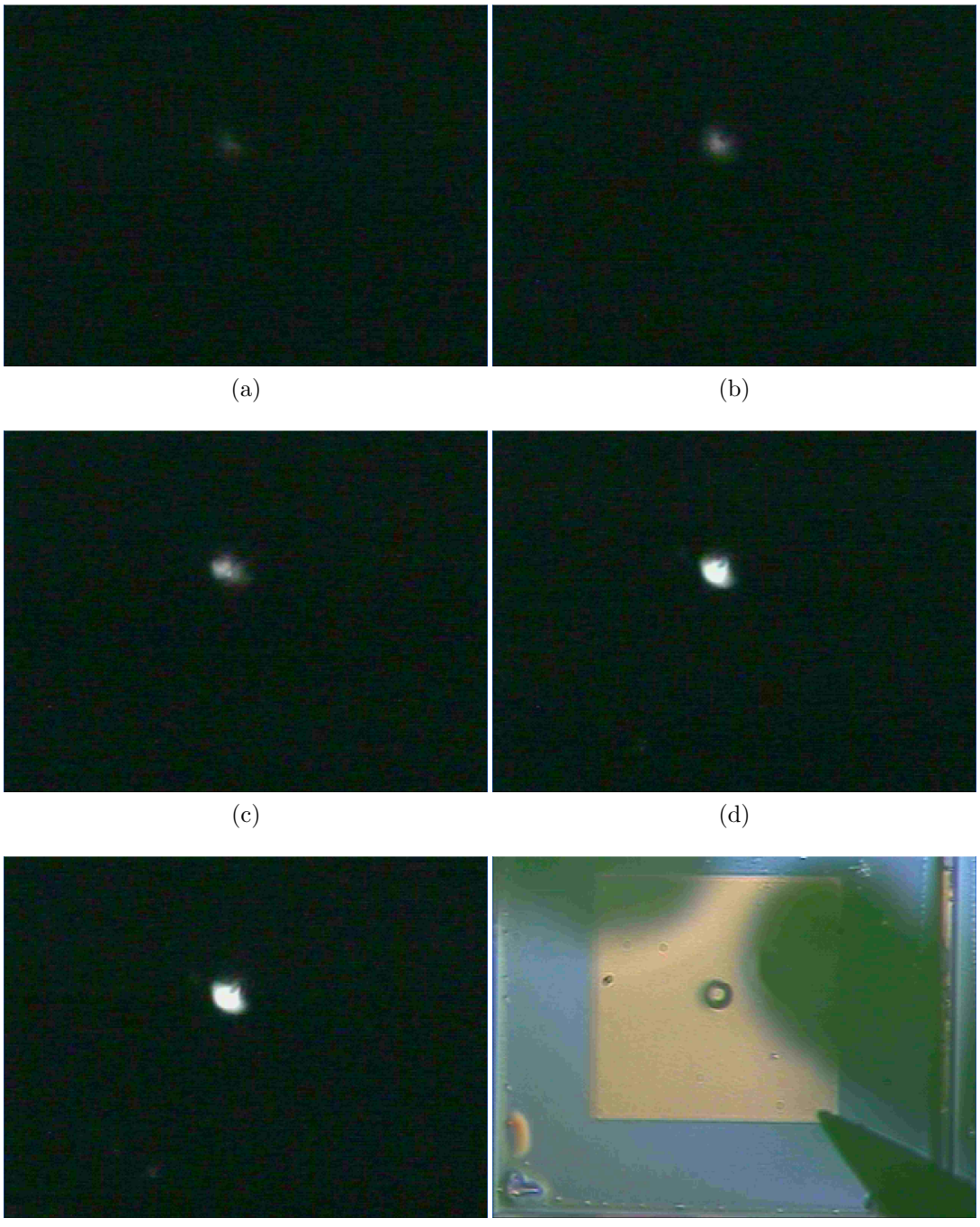


Figure 3.9: Near-field images of the top emission from a 10-μm-diameter metal-cavity quantum-well laser under (a) 1 mA, (b) 3 mA, (c) 5 mA, (d) 6 mA, and (e) 7 mA current injection. Images are collected through a 50x objective lens and taken by an infrared CCD camera. (f) Optical micrograph of the device being probed.

vertically through the bottom DBR to the substrate. And the vertical heat flow is less efficient for smaller devices.

We can also look at the contour plot of the current-dependent emission spectra. By tracing the peak location as current increases, we can map the relation between the peak wavelength and injection current, i.e. $\lambda_{\text{peak}}(I)$. At low current injection, we can often assume a linear relation between the temperature change ΔT and the dissipated power ΔP, i.e. a constant thermal resistance such that $\Delta T = R_{th}\Delta P$, where $\Delta P = IV - P_{\text{out}}$. Furthermore, the change of the peak wavelength $\Delta\lambda_{\text{peak}}$ is also nearly linear to the change of temperature ΔT. As a result, the peak wavelength should follow $\lambda_{\text{peak}}(I) - \lambda_{\text{peak}}(I = 0) \propto I\ln(I)$, where the natural logarithmic dependence comes from the diode equation. This gives rise to a dependence on I which is stronger than linear but weaker than quadratic, i.e. I^x where $1 < x < 2$. This argument is confirmed by Fig. 3.11.

As a comparison, we also fabricate simultaneously quantum-well lasers with dielectric cavities using the same wafer. The laser cavity is covered by BCB and the top emission window is exposed by the planarization of the BCB through plasma etch. Figure 3.12 shows the SEM of a processed BCB laser after contact metallization. The top surface of the laser is only covered by a thin layer (around 4 nm) of silver/gold for electrical contact. The devices are tested with the Cascade MPS150 probe station. The emission is collected by directly covering the devices with the Newport 818-SL silicon detector. Figure 3.13(a) shows our smallest BCB laser of 2-μm diameter at room temperature under pulsed injection. Figure 3.13(b) shows our smallest BCB laser of 4-μm diameter under CW operation at room temperature. We have later improved this work in [64], which have demonstrated CW lasing down to 2-μm diameter at room temperature, with the lasing threshold as small as 0.374 mA.

3.4 Characterization of Quantum-Dot Surface-Emitting Microlasers

Figure 3.14 shows the schematic diagram of the fabricated QD metal-cavity surface-emitting lasers. The layer structure is grown by molecular beam epitaxy (MBE). The gain medium consists of three groups of submonolayer

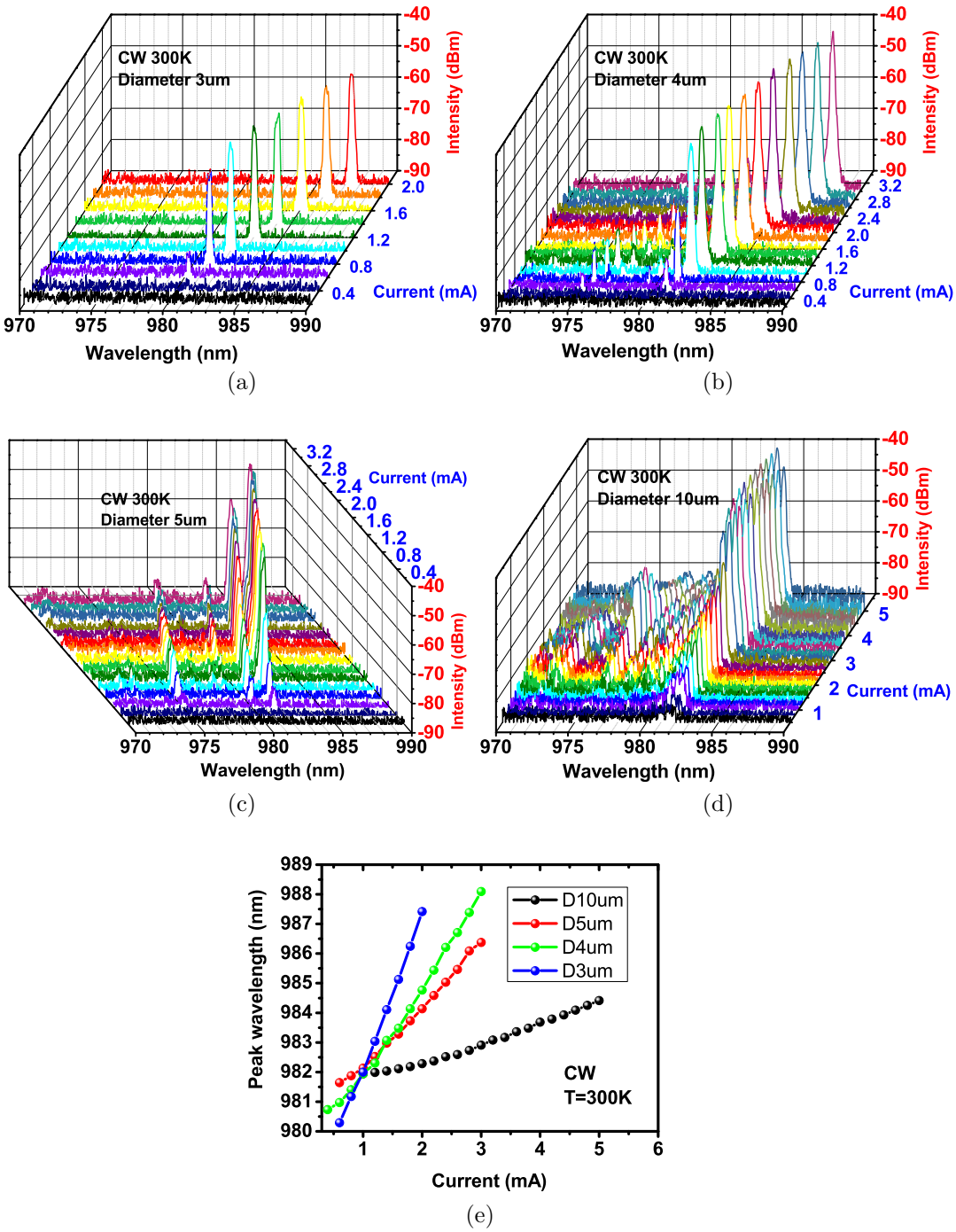


Figure 3.10: Current-dependent emission spectra of the metal-cavity quantum-well lasers with cavity diameters of (a) 3 μm, (b) 4 μm, (c) 5 μm, and (d) 10 μm. The lasers operate under continuous-wave electrical injection at room temperature. (e) Current-dependent peak wavelengths for the fundamental cavity mode with various device sizes.

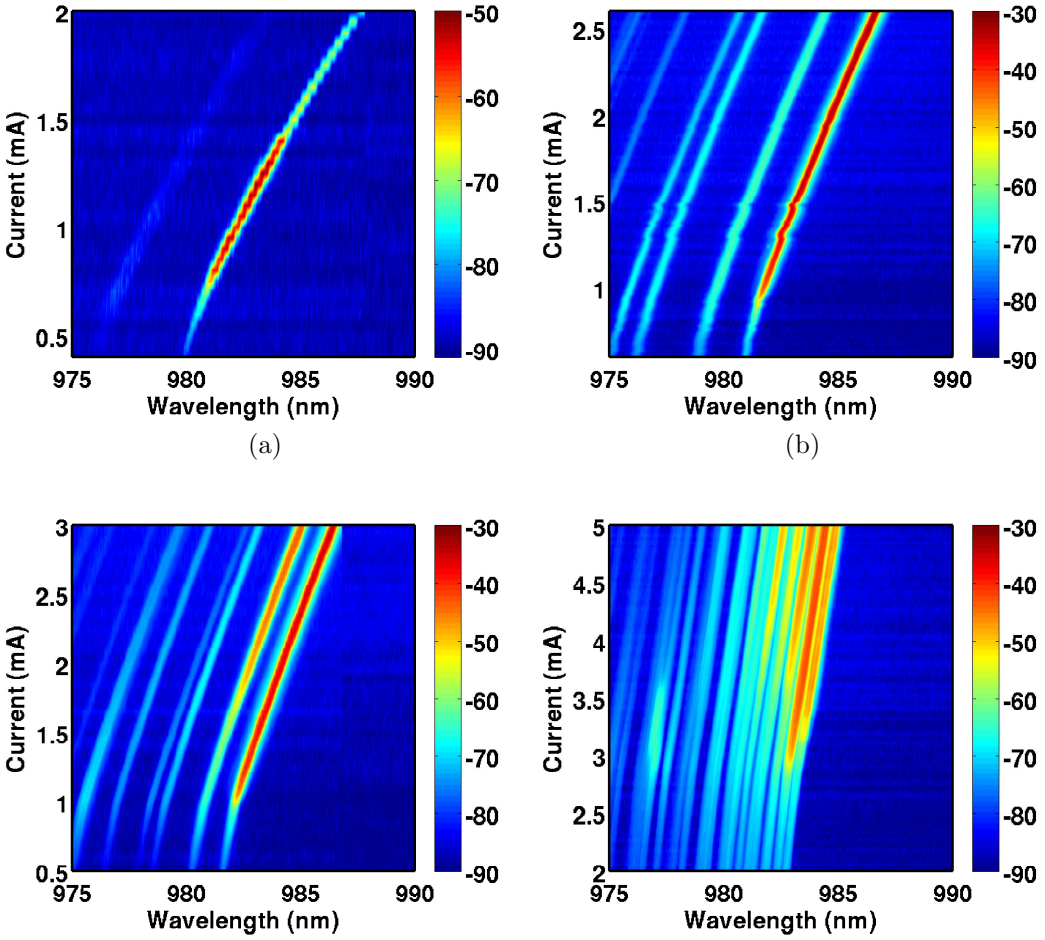
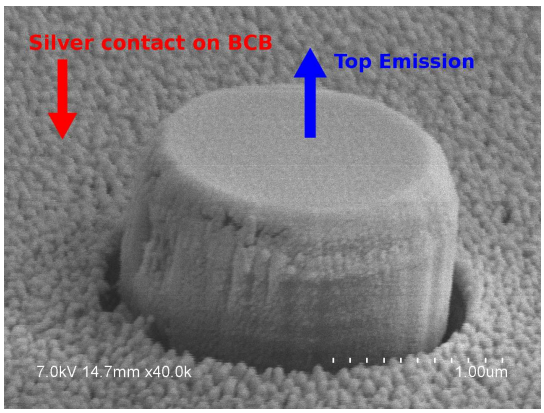
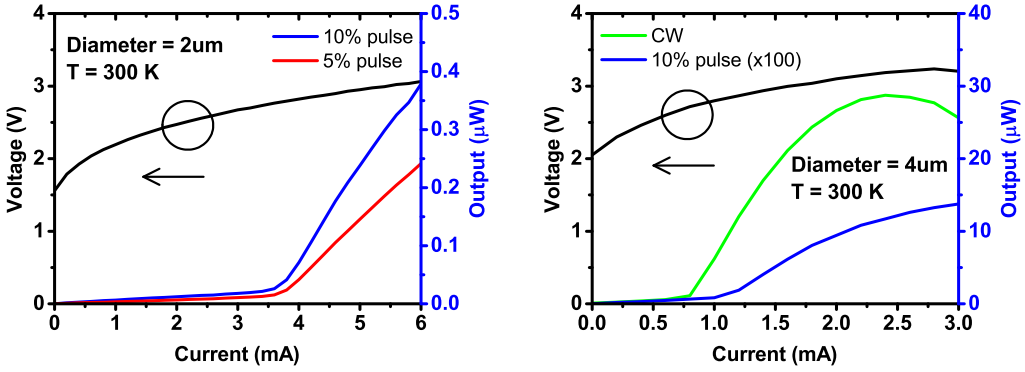


Figure 3.11: Contour plots of the current-dependent emission spectra of the metal-cavity quantum-well lasers with cavity diameters of (a) $3\,\mu$m, (b) $4\,\mu$m, (c) $5\,\mu$m, and (d) $10\,\mu$m.



(a)

Figure 3.12: Scanning electron micrograph of a 1-μm-diameter dielectric-cavity surface-emitting quantum-well laser after contact metallization.

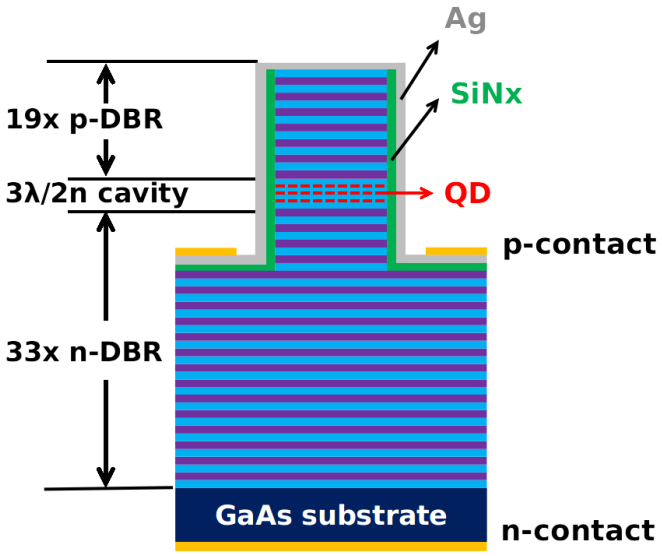


(a) (b)

Figure 3.13: Light output versus current (L-I) curves and current versus voltage (I-V) curves for our smallest dielectric-cavity quantum-well lasers at room temperature under (a) pulsed and (b) continuous-wave electrical injection.

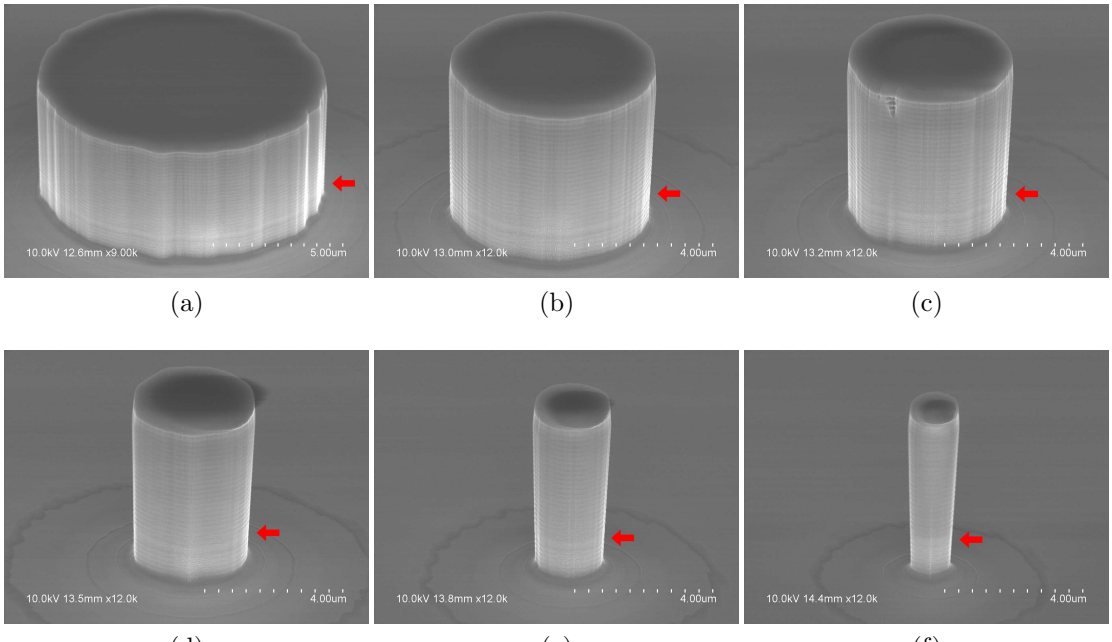
(SML) QDs, which are positioned within the $3\lambda/2n_{\mathrm{r}}$ cavity. The top and bottom mirrors are 19 pairs of p-doped DBR and 33 pairs of n-doped DBR, respectively. The DBR consists of alternating AlGaAs/GaAs layers. The lasing wavelength is designed to be around 975 nm.

In this section, we also carry out the size-dependent study of the QD laser properties. Devices with cavity diameters ranging from 10 μm down to 1 μm are fabricated simultaneously on the same chip. Figures 3.15(a)-(f) show the laser cavities with various diameters formed after the plasma etch and before



(a)

Figure 3.14: Schematic diagram of the metal-cavity quantum-dot surface-emitting microlaser.



(a) (b) (c)

(d) (e) (f)

Figure 3.15: Scanning electron micrographs showing the quantum-dot surface-emitting lasers with cavity diameters being (a) 10 μm, (b) 5 μm, (c) 4 μm, (d) 3 μm, (e) 2 μm, and (f) 1 μm. The red arrows indicate the quantum-dot active region.

the SiN_x sidewall coating.

The laser L-I curves are measured using the Newport 818-SL silicon photodetector calibrated with the Newport model 835 optical powermeter, and the ILX Lightwave LDP-3811 current source. The devices are mounted on the thermoelectric cooler, which is set at the room temperature (293 K) during measurements. Figure 3.16(a) shows the experimental and theoretical L-I curves for a 2-μm-diameter device. Good agreement is shown when the carrier-dependent β_{sp}-factor is used, as discussed in Chapter 2. Figure 3.16(b) shows the emission spectrum for the same device operating at room temperature under 5 mA pulsed electrical injection with 200 kHz repetition rate and 10% duty cycle. The spectrum indicates that the device is lasing at 5 mA pulsed injection. Figure 3.16(c) shows the light output versus current density (L-J) curves for devices with various diameters operating under 1%-duty cycle electrical injection. Our theory explains very well the measured laser behavior.

In order to experimentally determine the parameters of the cavity modes, we need to measure the emission spectrum before the laser reaches the threshold [25]. The emission is collected using an objective lens (50x, 0.60 NA). The emission spectrum is obtained through the SPEX 1250M monochromator with a liquid-nitrogen-cooled germanium detector. Since the sub-threshold emission is weak and we need a narrow slit width in the monochromator for high spectral resolution, it is necessary to use the lock-in technique to improve the signal-to-noise ratio. The Stanford Research Systems lock-in amplifier model SR530 is used in this experiment. Figure 3.17 shows the sub-threshold emission spectra for devices with diameters ranging from 10 μm down to 2 μm. The black arrows indicate the fundamental cavity modes.

Below the lasing threshold, the spontaneous emission can still be amplified by the cavity effect. Since the thermal effect is negligible under pulsed operation below threshold, we can estimate the cavity resonance from the amplified spontaneous emission, and the peak wavelength (λ_c) should have very little shift with current injection below threshold. Furthermore, from the full-width-half-maximum (FWHM) of the emission peaks, we can extract the quality factor (Q) for each cavity mode. Ideally, what we need is the cold-cavity Q, where all the loss comes from the cavity, including material loss (ohmic loss) and diffraction loss. This means that the gain medium should have neither amplification nor absorption, and should be electrically biased

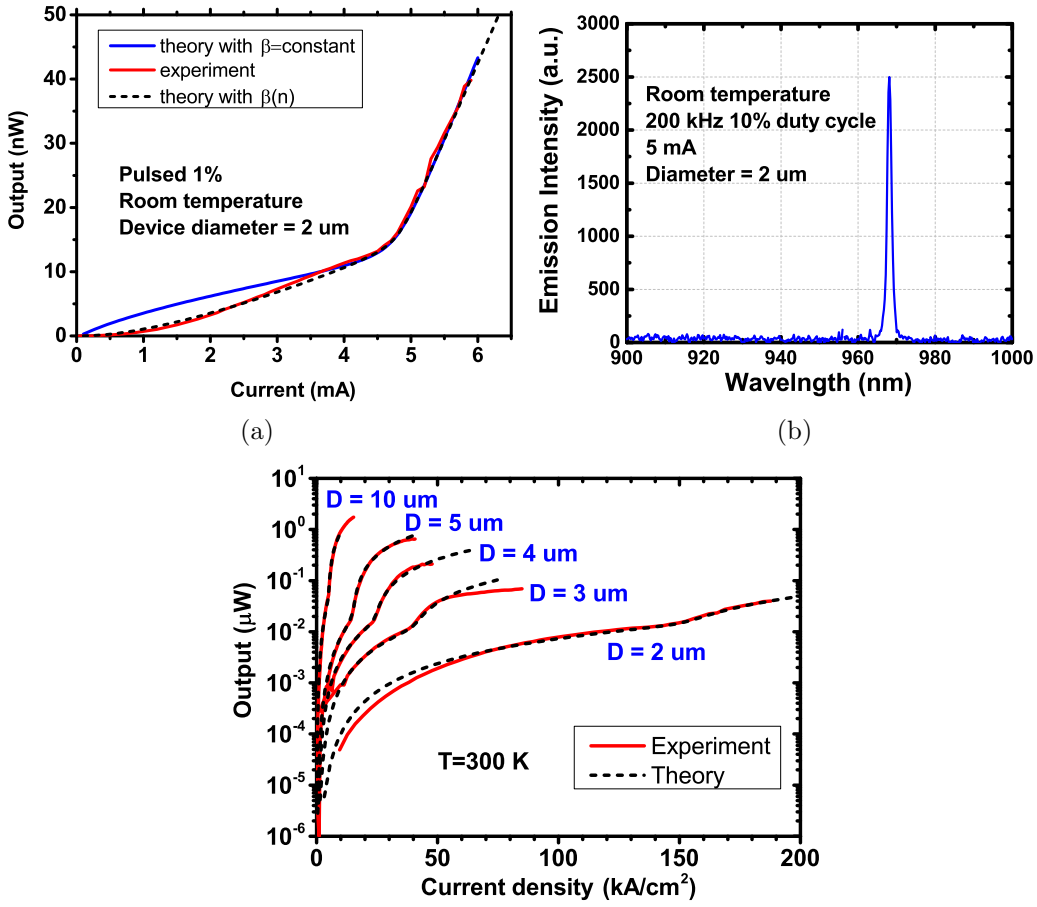
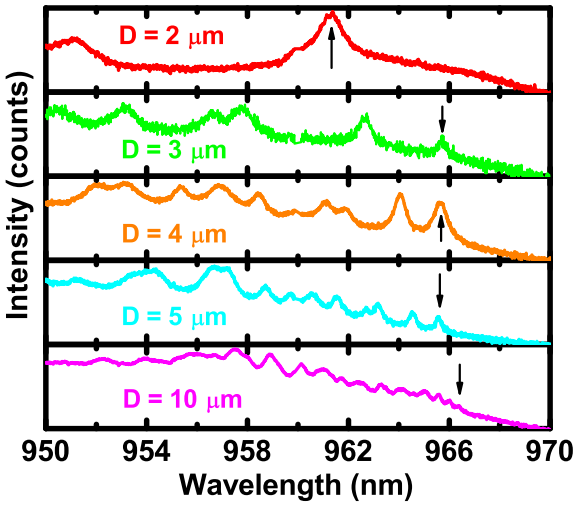


Figure 3.16: (a) Experimental light output versus current (L-I) curve (red) for a 2-μm-diameter surface-emitting quantum-dot laser operating at room temperature. Theoretical L-I curves with (dashed) and without (blue) using a carrier-dependent $\beta_{\rm sp}$-factor. (b) Experimental emission spectrum for the same laser with 5 mA pulsed electrical injection at room temperature, showing the lasing behavior. (c) Light output versus current density (L-J) curves for devices with different cavity diameters. Comparison between theory (dashed) and experiment (red) is shown.

at the transparency condition. The transparency carrier density for the gain region to have zero gain or loss is difficult to determine experimentally for VCSELs, yet we can estimate theoretically from the gain spectra as in Chapter 2 for QDs and Chapter 5 for QWs. In the case of our SML QDs, the transparency carrier density $n_{\rm tr}$ is around $1 \times 10^{18}\,{\rm cm}^{-3}$, which is around 1/3 to 1/2 of the threshold carrier density $n_{\rm th}$, depending on the size of the device. Figure 3.18 shows the magnified view of the sub-threshold emission



(a)

Figure 3.17: Emission spectra below threshold for metal-cavity surface-emitting quantum-dot lasers with various cavity diameters. The black arrows indicate the fundamental cavity modes.

Table 3.1: Extracted size-dependent cavity-mode parameters.

Device diameter (μm)	Q(HE_{11})	Q(TE_{01})	λ_c(HE_{11}) (nm)	λ_c(TE_{01}) (nm)
2	682	320	961.07	950.51
3	870	762	965.40	962.60
4	2702	2258	965.63	964.05
5	2997	2264	965.57	964.54
10	2.77	216.83	966.32	965.97

spectra of devices with various sizes. Lorentzian fits are shown for the emission peaks corresponding to the two fundamental cavity modes (HE_{11} and TE_{01}). From the extracted FWHMs, we can determine the Q for each mode. The values of λ_c and Q are shown in Table 3.1.

With the extracted λ_c and Q, we can theoretically estimate the spontaneous emission coupled to a specific mode based on the formulation in Chapter 2. Figure 3.19 shows the size-dependent experimental Q and the calculated discrete-mode spontaneous emission rate ($R_{\mathrm{sp},m}$). The Q decreases with the reduction of the device size due to the increase of cavity diffraction loss. But at the same time, the effective mode volume (V_{eff}) decreases as the device size shrinks down, resulting in a larger Purcell factor, which

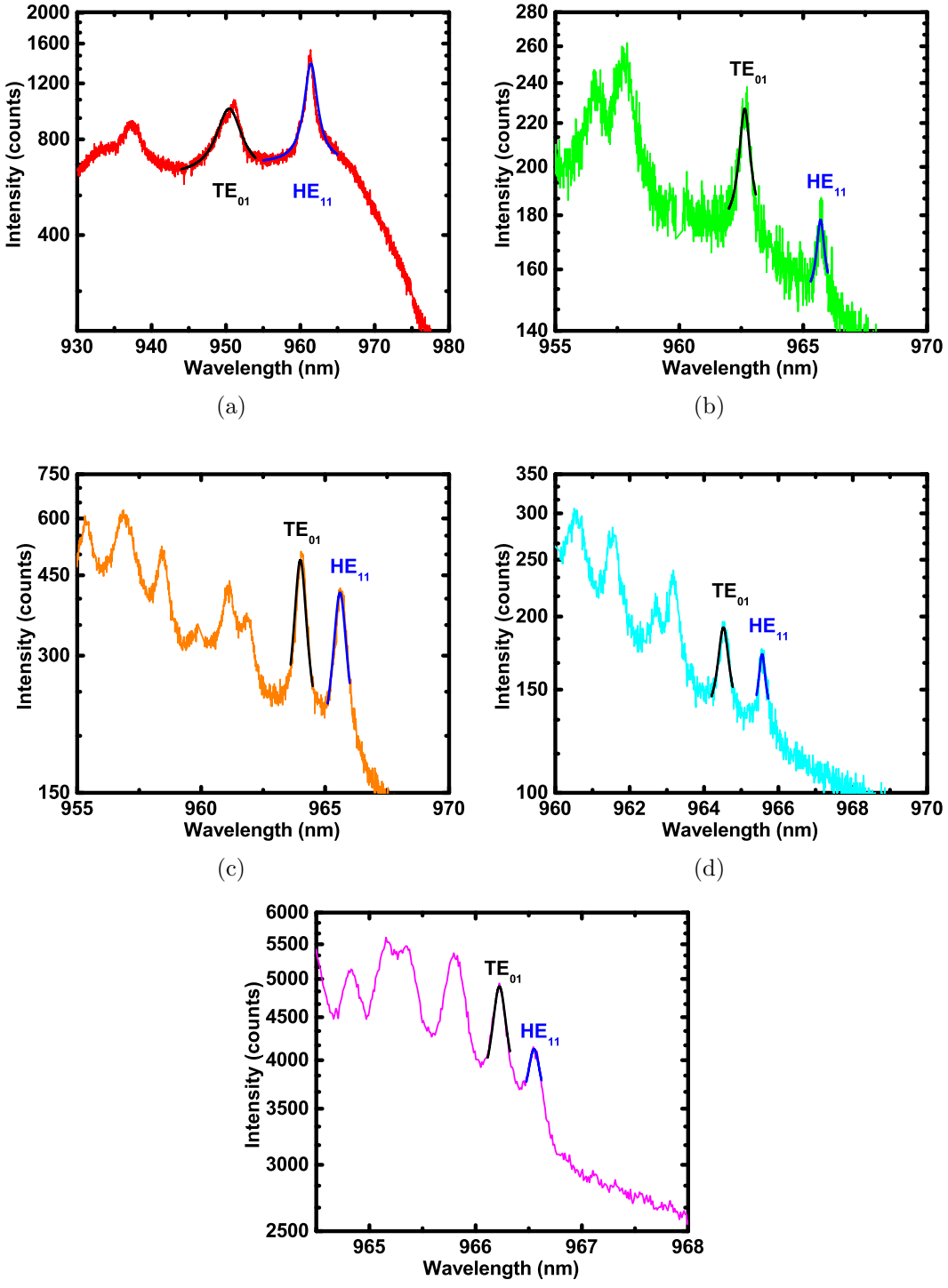


Figure 3.18: Lorentzian fits of the emission spectra for parameter extraction of the HE_{11} (blue) and TE_{01} (black) cavity modes. The device diameters are (a) 2 μm, (b) 3 μm, (c) 4 μm, (d) 5 μm, and (e) 10 μm, respectively.

is proportional to the ratio between Q and V_{eff}. Therefore, we can see an increasing amount of spontaneous emission coupled into the cavity mode as the size shrinks down in Fig. 3.19.

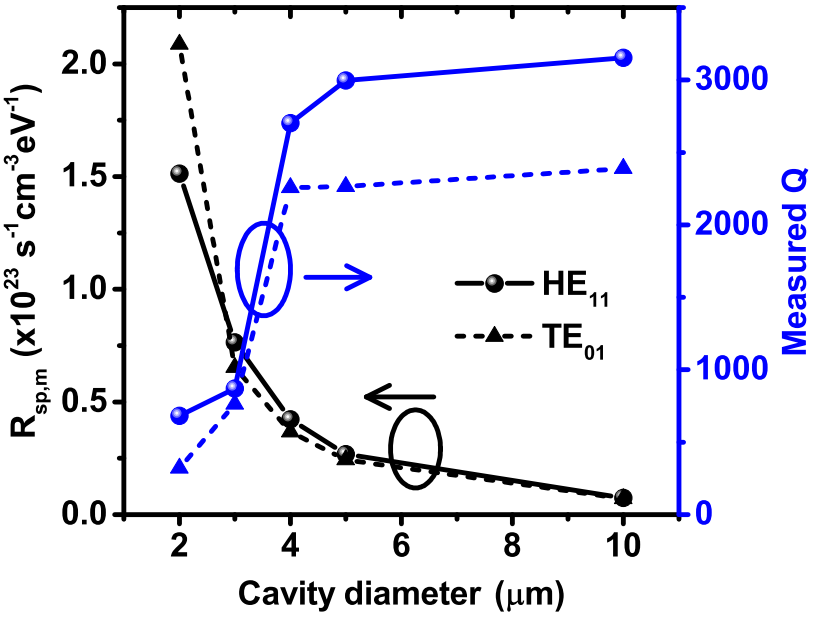


Figure 3.19: Size-dependent experimental cavity quality factor (blue) and the calculated discrete-mode spontaneous emission (black) for the HE_{11} mode (round dots) and the TE_{01} mode (triangles).

3.5 Summary

In summary, we show the detailed fabrication procedures for both the metal-cavity and the dielectric-cavity surface-emitting micro and nanolasers. We demonstrate the room-temperature operation under electrical injection of both quantum-well (QW) and quantum-dot (QD) lasers. Continuous-wave lasing is observed for 3-μm-diameter QW lasers with 0.5-mA threshold. Current-dependent emission spectra are collected and the size-dependent thermal properties are investigated. Dielectric-cavity lasers are characterized for comparison. We observe room-temperature lasing for QD devices with only 2-μm cavity diameters. We further investigate the sub-threshold emission spectra for QD lasers and extract the cavity information for our analysis of the spontaneous emission coupling with cavity modes. Our theory successfully

explains our experimental observation.

Different from many other groups' work on nanolasers, our work has a fixed constraint of electrical injection at room temperature. And our strategy is to start with a larger device size and move toward the device miniaturization without relaxing the operation criteria. As a result, our devices always meet the practical working conditions, where high-power optical pumps and cryogenic temperature are not feasible.

CHAPTER 4

MODELING AND DESIGN OF HIGH-CONTRAST GRATINGS

4.1 Introduction

Diffraction gratings have been crucial components in the broad area of optics for centuries [65]. In the recent decade, researches have shown extensive interest in an operation regime where the grating period Λ is comparable or slightly smaller than the wavelength λ [32, 35, 66, 67, 68], where strong field interaction among the near-wavelength structures can provide extraordinary optical behaviors. One interesting type of grating, known as the high-contrast grating (HCG), is made of periodic high-index materials surrounded by materials with much smaller indices [35]. HCGs have been used as ultrabroadband reflectors [69] and high-Q resonators [70]. They can be integrated as compact, high-performance tunable mirrors in optoelectronic devices, such as vertical-cavity surface-emitting lasers (VCSELs) [32, 33, 38].

In Sections 4.2 and 4.3 the optical properties of one-dimensional (1D) gratings are investigated using the in-house developed mode-matching method and the surface-integral-equation (SIE) method, respectively. The results are compared with the finite-element method (FEM) using COMSOL Multiphysics. Excellent agreement is demonstrated. The developed methods allow the efficient top-down design of 1D gratings for various optical applications. To use 1D HCGs as broadband high-reflection mirrors in vertical-cavity surface-emitting lasers (VCSELs), the key property is the complex reflection coefficient for the fundamental reflection order.

Recently, much research interest has been devoted to the design of 2D arrays of nanostructures [66, 71]. In Section 4.4 the properties of 2D HCGs under various incidence conditions and with different structural parameters are studied using an in-house developed rigorous coupled-wave analysis (RCWA) package. In Section 4.5 a top-down design procedure for 2D HCGs is pro-

posed with the RCWA package as an efficient design tool.

One important application for 2D HCGs is the 2D phase plate, which allows novel designs of optical components, such as deflectors, lenses [72], axicons for generating Bessel beams [73, 74], vortex phase plates for generating orbital angular momentum of light [75, 76], or even holographic plates. 2D phase plates normally require non-periodic designs. However, the high-index material can largely confine the field around the local position, where effective-medium approximation is applicable [76, 77, 78]. Based on our optimized design for periodic 2D HCGs, we can apply the grating parameters as the local parameters for the non-periodic grating, according to the design requirement at the local position. Our designed non-periodic phase plate is then verified through the finite-difference time-domain (FDTD) simulation. The results show excellent agreement with our designs.

4.2 Mode-Matching Method for 1D Periodic Structures

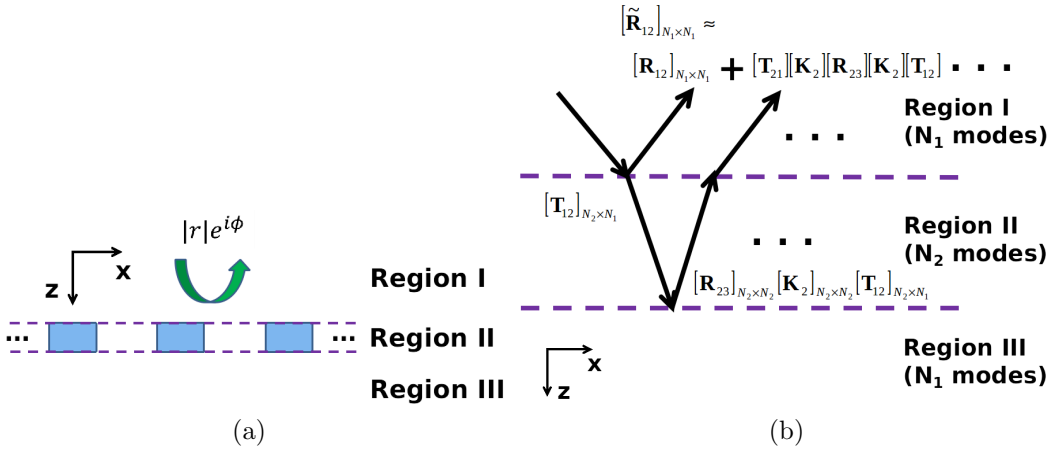
HCGs are referred as one-dimensional (1D) if the periodicity is only in one direction (e.g. periodic bars, stripes, or grooves). In this case the problem is reduced to a 2D problem, and the fields can be decoupled into the transverse-electric (TE) and transverse-magnetic (TM) polarizations, where the electric field is parallel and perpendicular to the grating bars, respectively. An incident wave with one polarization will not excite an HCG mode with an orthogonal polarization.

4.2.1 Theoretical Formulation

The complex reflection coefficient of the HCG is an important parameter for the design and modeling of HCG VCSELs. For normal incidence on a subwavelength HCG, i.e. the wavelength is smaller than the grating period Λ, only the zeroth-order Floquet mode (normal reflection) is propagating, while all higher-order Floquet modes are evanescent. In this case, a complex reflection coefficient, instead of a reflection matrix, can be defined. In this subsection, we compare the reflection spectra calculated using three different methods: analytical mode-matching [79], mode-matching [80] using numerically solved eigenmodes, and parameter extraction from the finite-element

simulation. As shown in Fig. 4.1(a), we define Region I and Region III to be the incidence and transmission air regions, respectively, and Region II to be the HCG layer.

For the analytical mode-matching method, the modes inside Region II are expanded in terms of sinusoidal functions in both the transverse and longitudinal directions. In Regions I and III, the field is expanded into a summation of discrete propagating waves, the wavenumbers and propagation directions of which are governed by the Floquet theorem. The boundary conditions for the electric and magnetic fields are matched at the air-HCG interfaces, and we can obtain a reflection matrix, out of which the zeroth-order reflection coefficient can be extracted. For the numerical mode-matching method, we first solved the eigenmodes in the HCG layer using the 1D finite-difference frequency-domain (FDFD) method, and using Floquet modes in the air regions.



(a) (b)

Figure 4.1: (a) Schematic of a high-contrast-grating (HCG) reflector with a normal incident plane wave. The complex reflection coefficient $|r|e^{i\phi}$ for the zeroth-order reflected wave is to be determined. (b) Calculation of the HCG reflection by mode-matching method. Fields in each region are expanded by eigenmodes, and the reflection and transmission matrices $\bar{\mathbf{R}}$ and $\bar{\mathbf{T}}$ are determined at each interface. The generalized reflection matrix $\widetilde{\mathbf{R}}$ is obtained.

For the 2D TE (E_y) case, the electric field (invariant in $\hat{y}$) in the grating region (Region II) can be written as

$$\mathbf{E}^{\mathrm{II}}(x,z) = \hat{y}E_y^{\mathrm{II}}(x)e^{ik_z z} \tag{4.1}$$

Substituting the above into the Helmholtz equation $\nabla^2\mathbf{E} + k^2\mathbf{E} = 0$, we get

$$\frac{\partial^2}{\partial x^2}E_y^{\mathrm{II}}(x) + k_0^2 n_{\mathrm{r}}^2(x)E_y^{\mathrm{II}}(x) = k_z^2 E_y^{\mathrm{II}}(x) \tag{4.2}$$

from which we see that the eigen-problem in the grating is now a 1D problem. According to the Bloch's theorem, the periodicity in x-direction requires the field to take the form of $E_y^{\mathrm{II}}(x) = U_{k_{x,\mathrm{Bloch}}}(x)e^{ik_{x,\mathrm{Bloch}}x}$, where $U_{k_{x,\mathrm{Bloch}}}(x + \Lambda) = U_{k_{x,\mathrm{Bloch}}}(x)$ and Λ is the period. If the field repeats for N periods, i.e. $E_y^{\mathrm{II}}(x + N\Lambda) = E_y^{\mathrm{II}}(x)$, then the Bloch wave number $k_{x,\mathrm{Bloch}}$ takes discrete values in the first Brillouin zone,

$$k_{x,\mathrm{Bloch}} = \frac{2\pi n}{N\Lambda}, \text{ where } n = \begin{cases} 0, \pm1, \pm2, \ldots, \pm\frac{N}{2}, & N \text{ is even} \\ 0, \pm1, \pm2, \ldots, \pm\frac{N-1}{2}, & N \text{ is odd} \end{cases} \tag{4.3}$$

Then the eigenvalue problem is solved in one period using the FDFD method with the periodic boundary condition $E_y^{\mathrm{II}}(x + \Lambda) = e^{ik_{x,\mathrm{Bloch}}\Lambda}E_y^{\mathrm{II}}(x)$.

Figure 4.2 shows the first four eigenmodes corresponding to $k_{x,\mathrm{Bloch}} = 0$ (Γ-point) for 10 periods. Figure 4.3 shows the contour plots for these four eigenmodes as they propagate along the $\hat{z}$-direction. We can see starting from the fourth mode to higher orders, the fields become evanescent.

Because of the periodic structure, the scattered waves in Regions I and III have discrete wave vectors, and we can write the fields as

$$\mathbf{E}^{\mathrm{I,III}}(x, z) = \hat{y}E_{y,n}^{\mathrm{I,III}}(x)e^{ik_{z,n}z} \tag{4.4}$$

for the n-th mode, where the transverse profile takes the form $E_{y,n}^{\mathrm{I,III}}(x) = \sqrt{\frac{1}{N\Lambda}}e^{ik_{x,n}x}$, and $k_{x,n} = k_{ix} + \frac{2n\pi}{\Lambda},\ n = 0, \pm1, \pm2, \ldots$.

Now we can expand the transverse electric field in terms of eigenmodes in each region,

$$\begin{aligned} \mathbf{E}_t =& \hat{y}\sum_n^N \psi_n(x)e^{ik_{z,n}z}E_n \\ =& \hat{y}\left[\psi_1(x), \ldots, \psi_n(x)\right] \begin{bmatrix} \exp(ik_{z,1}t_g) & 0 & \cdots \\ 0 & \exp(ik_{z,2}t_g) & \\ \vdots & & \ddots \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{bmatrix} \end{aligned} \tag{4.5}$$

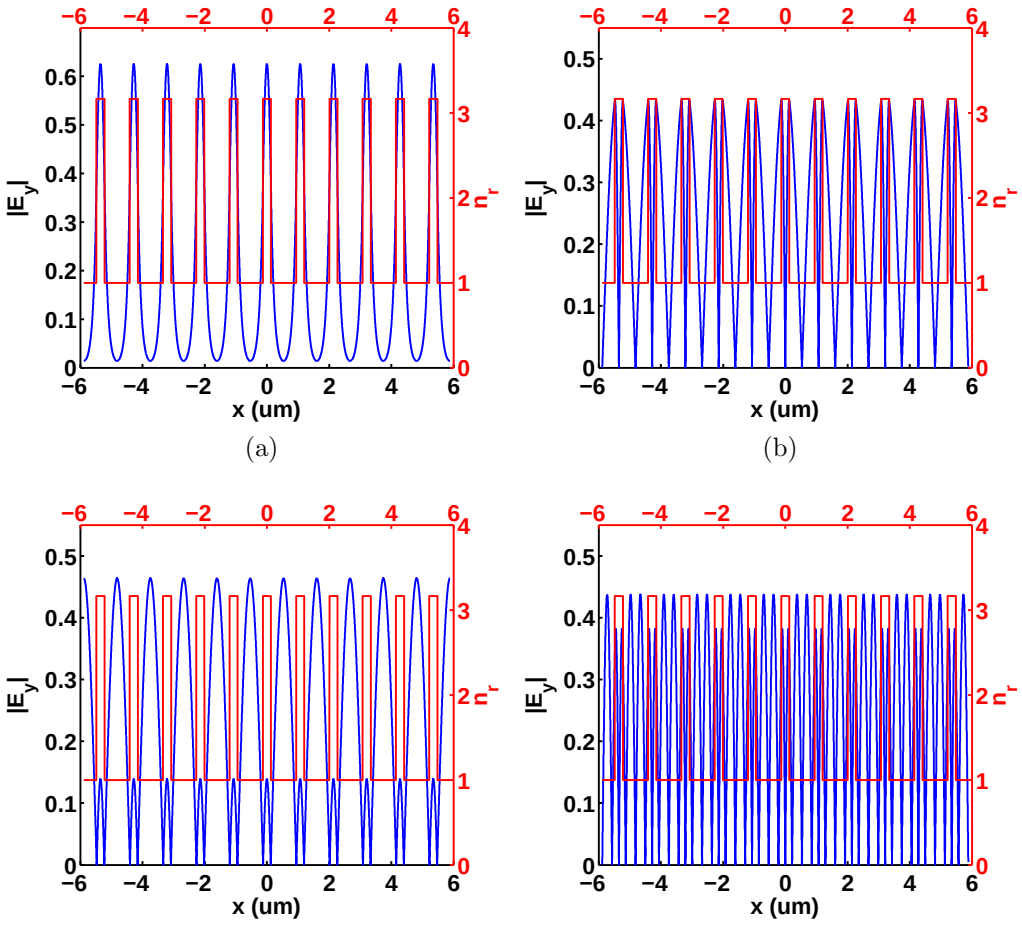


Figure 4.2: $|E_y|$ for the first four TE-polarized eigenmodes in the grating region for 10 periods. Parameters: period $\Lambda = 1070\,\mathrm{nm}$, grating width $D = 260\,\mathrm{nm}$, wavelength $\lambda = 1.55\,\mu\mathrm{m}$, and index $n_\mathrm{r} = 3.164$.

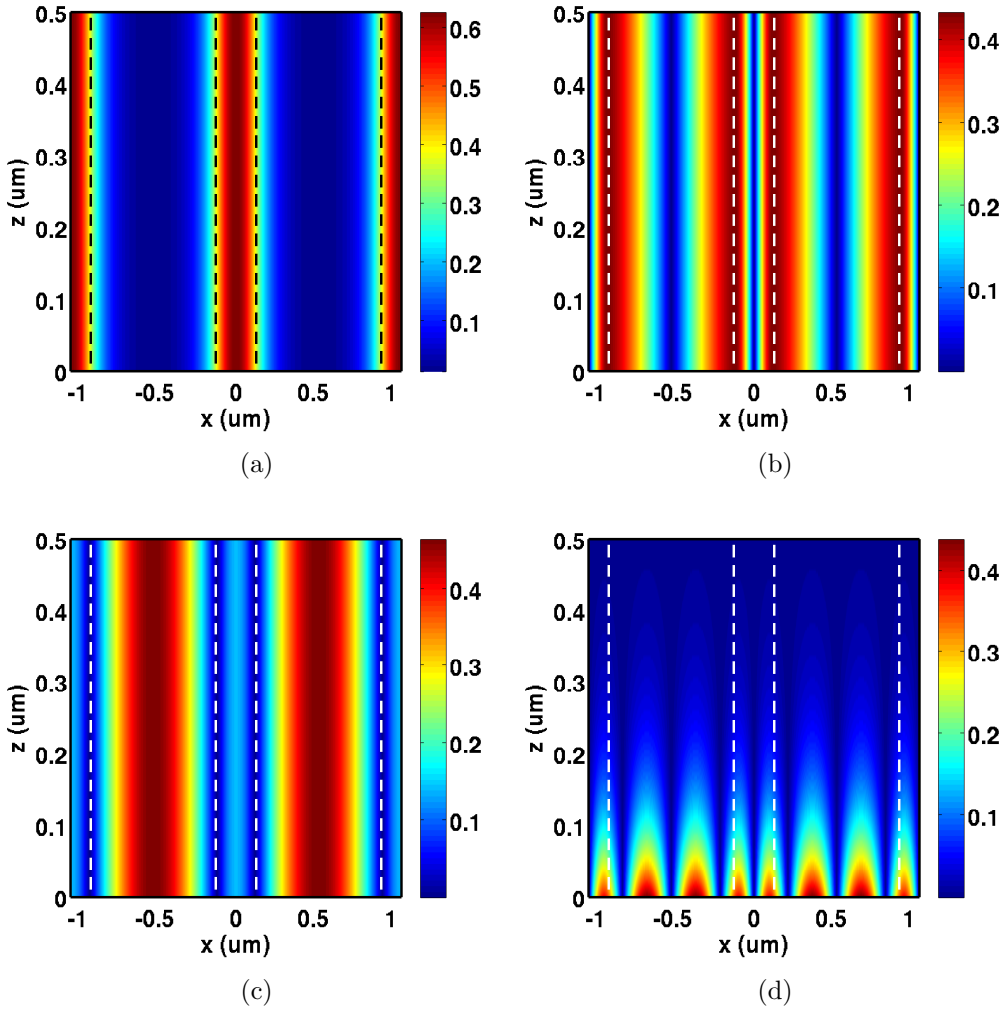


Figure 4.3: Contour plots of $|E_y|$ for the first four TE-polarized eigenmodes in one unit cell as they propagate along the $+\hat{z}$-direction in the grating region. Starting from the fourth mode the fields become evanescent in the $+\hat{z}$-direction.

Similarly, the transverse magnetic field can be written as

$$\begin{aligned}
-\hat{z}\times\mathbf{H}_t =& \hat{y}\frac{1}{i\omega\mu}\frac{\partial}{\partial z}E_y = \hat{y}\sum_n^N \frac{k_{z,n}}{\omega\mu}\psi_n(x)e^{ik_{z,n}z}E_n \\
=& \hat{y}\left[\psi_1(x),\ldots,\psi_n(x)\right]\begin{bmatrix} k_{z,1}/\omega\mu & 0 & \ldots \\ 0 & k_{z,2}/\omega\mu & \\ \vdots & & \ddots \end{bmatrix} \\
& \begin{bmatrix} \exp(ik_{z,1}z) & 0 & \ldots \\ 0 & \exp(ik_{z,2}z) & \\ \vdots & & \ddots \end{bmatrix}\begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{bmatrix}
\end{aligned} \tag{4.6}$$

In terms of the matrix form, we can write

$$\begin{aligned}
E_y =& \bar{\mathbf{\Psi}}^\dagger(x)\cdot e^{i\bar{\mathbf{K}}_z z}\cdot\mathbf{e} \\
\left[-\hat{z}\times\mathbf{H}_t\right]_y =& \bar{\mathbf{\Psi}}^\dagger(x)\cdot\bar{\mathbf{G}}\cdot e^{i\bar{\mathbf{K}}_z z}\cdot\mathbf{e}
\end{aligned} \tag{4.7}$$

where $\bar{\mathbf{\Psi}}^\dagger(x) = [\psi_1(x),\ldots,\psi_n(x)]$, $\bar{\mathbf{G}}$ and $e^{i\bar{\mathbf{K}}_z z}$ are diagonal matrices with elements being $k_{z,n}/\omega\mu$ and $\exp(ik_{z,n}z)$, respectively. To account for both forward and backward propagation as the wave transmits from Region I to Region II, the transverse fields become

$$\begin{aligned}
E_y^{\mathrm{I}} =& \bar{\mathbf{\Psi}}_1^\dagger(x)\cdot\left(e^{i\bar{\mathbf{K}}_{1z}z}\cdot\mathbf{e}_1 + e^{-i\bar{\mathbf{K}}_{1z}z}\cdot\bar{\mathbf{R}}_{12}\cdot\mathbf{e}_1\right) \\
\left[-\hat{z}\times\mathbf{H}_t^{\mathrm{I}}\right]_y =& \bar{\mathbf{\Psi}}_1^\dagger(x)\cdot\bar{\mathbf{G}}_1\cdot\left(e^{i\bar{\mathbf{K}}_{1z}z}\cdot\mathbf{e}_1 - e^{-i\bar{\mathbf{K}}_{1z}z}\cdot\bar{\mathbf{R}}_{12}\cdot\mathbf{e}_1\right)
\end{aligned} \tag{4.8}$$

in Region I and

$$\begin{aligned}
E_y^{\mathrm{II}} =& \bar{\mathbf{\Psi}}_2^\dagger(x)\cdot e^{i\bar{\mathbf{K}}_{2z}z}\cdot\mathbf{e}_2 = \bar{\mathbf{\Psi}}_2^\dagger(x)\cdot e^{i\bar{\mathbf{K}}_{2z}z}\cdot\bar{\mathbf{T}}_{12}\cdot\mathbf{e}_1 \\
\left[-\hat{z}\times\mathbf{H}_t^{\mathrm{II}}\right]_y =& \bar{\mathbf{\Psi}}_2^\dagger(x)\cdot\bar{\mathbf{G}}_2\cdot e^{i\bar{\mathbf{K}}_{2z}z}\cdot\mathbf{e}_2 = \bar{\mathbf{\Psi}}_2^\dagger(x)\cdot\bar{\mathbf{G}}_2\cdot e^{i\bar{\mathbf{K}}_{2z}z}\cdot\bar{\mathbf{T}}_{12}\cdot\mathbf{e}_1
\end{aligned} \tag{4.9}$$

in Region II, where $\bar{\mathbf{R}}_{12}$ is a $N_1\times N_1$ reflection matrix from Region I to Region II, and $\bar{\mathbf{T}}_{12}$ is a $N_2\times N_1$ transmission matrix from Region I to Region II. N_1 and N_2 are the numbers of modes kept in Regions I and II, respectively.

The continuity of the transverse fields from Region I to Region II gives

$$\begin{aligned}
&\bar{\mathbf{\Psi}}_1^{\dagger}(x)\cdot\left(e^{i\bar{\mathbf{K}}_{1z}z}\cdot\mathbf{e}_1+e^{-i\bar{\mathbf{K}}_{1z}z}\cdot\bar{\mathbf{R}}_{12}\cdot\mathbf{e}_1\right)=\bar{\mathbf{\Psi}}_2^{\dagger}(x)\cdot e^{i\bar{\mathbf{K}}_{2z}z}\cdot\bar{\mathbf{T}}_{12}\cdot\mathbf{e}_1\\
&\bar{\mathbf{\Psi}}_1^{\dagger}(x)\cdot\bar{\mathbf{G}}_1\cdot\left(e^{i\bar{\mathbf{K}}_{1z}z}\cdot\mathbf{e}_1-e^{-i\bar{\mathbf{K}}_{1z}z}\cdot\bar{\mathbf{R}}_{12}\cdot\mathbf{e}_1\right)=\bar{\mathbf{\Psi}}_2^{\dagger}(x)\cdot\bar{\mathbf{G}}_2\cdot e^{i\bar{\mathbf{K}}_{2z}z}\cdot\bar{\mathbf{T}}_{12}\cdot\mathbf{e}_1
\end{aligned}\tag{4.10}$$

We can set the interface location as $z=0$ and do an inner-product with $\bar{\mathbf{\Psi}}_1$. Then the reflection and transmission matrices can be obtained at the I-II interface as

$$\begin{aligned}
\bar{\mathbf{T}}_{12}&=2\left(\bar{\mathbf{D}}_1^{-1}\bar{\mathbf{L}}_{12}+\bar{\mathbf{G}}_1^{-1}\bar{\mathbf{D}}_1^{-1}\bar{\mathbf{L}}_{12}\bar{\mathbf{G}}_2\right)^{-1}\\
\bar{\mathbf{R}}_{12}&=\left(\bar{\mathbf{D}}_1^{-1}\bar{\mathbf{L}}_{12}-\bar{\mathbf{G}}_1^{-1}\bar{\mathbf{D}}_1^{-1}\bar{\mathbf{L}}_{12}\bar{\mathbf{G}}_2\right)^{-1}\left(\bar{\mathbf{D}}_1^{-1}\bar{\mathbf{L}}_{12}+\bar{\mathbf{G}}_1^{-1}\bar{\mathbf{D}}_1^{-1}\bar{\mathbf{L}}_{12}\bar{\mathbf{G}}_{12}\right)
\end{aligned}\tag{4.11}$$

where $\bar{\mathbf{D}}_i=\langle\bar{\mathbf{\Psi}}_i,\bar{\mathbf{\Psi}}_i^{\dagger}\rangle$ and $\bar{\mathbf{L}}_{ij}=\langle\bar{\mathbf{\Psi}}_i,\bar{\mathbf{\Psi}}_j^{\dagger}\rangle$.

Similarly, we can find the reflection and transmission matrices at the II-III interface. Then the generalized reflection matrix $[\widetilde{\mathbf{R}}_{12}]_{N_1\times N_1}$ can be obtained as [80]

$$\widetilde{\mathbf{R}}_{12}=\bar{\mathbf{R}}_{12}+\bar{\mathbf{T}}_{21}(\bar{\mathbf{I}}-e^{i\bar{\mathbf{K}}_{2z}t_g}\bar{\mathbf{R}}_{23}e^{i\bar{\mathbf{K}}_{2z}t_g}\bar{\mathbf{R}}_{21})^{-1}e^{i\bar{\mathbf{K}}_{2z}t_g}\bar{\mathbf{R}}e^{i\bar{\mathbf{K}}_{2z}t_g}\bar{\mathbf{T}}_{12}\tag{4.12}$$

where $[\bar{\mathbf{R}}_{12}]_{N_1\times N_1}$, $[\bar{\mathbf{R}}_{23}]_{N_2\times N_2}$, $[\bar{\mathbf{R}}_{21}]_{N_2\times N_2}$, $[\bar{\mathbf{T}}_{12}]_{N_2\times N_1}$, and $[\bar{\mathbf{T}}_{21}]_{N_1\times N_2}$ are the reflection and transmission matrices at the interfaces. $[\bar{\mathbf{I}}]_{N_2\times N_2}$ is the identity matrix. The subscript ij indicates wave incidence from Region i to Region j. N_1 and N_2 are the number of modes in Region I and Region II, respectively. The propagation matrix $[e^{i\bar{\mathbf{K}}_{2z}t_g}]_{N_2\times N_2}$ can be written as

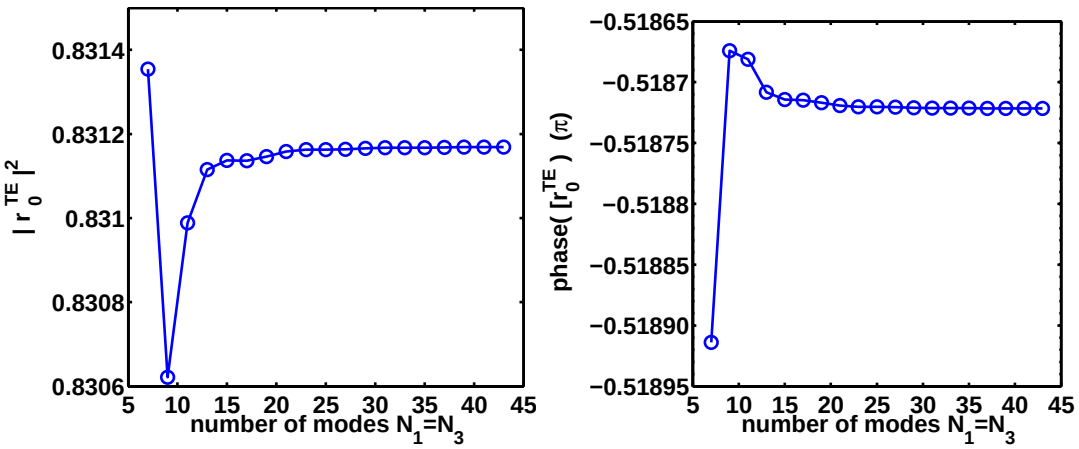
$$e^{i\bar{\mathbf{K}}_{2z}t_g}=\begin{pmatrix}\exp(ik_{z,1}t_g) & 0 & \cdots\\ 0 & \exp(ik_{z,2}t_g) & \\ \vdots & & \ddots\end{pmatrix}\tag{4.13}$$

where t_g is the thickness of the grating, and $k_{z,i}$ is the propagation constant for the i-th eigenmode in Region II. The complex reflection coefficient can then be obtained from the $\widetilde{\mathbf{R}}_{12}$ matrix. As shown in Fig. 4.1(b), the idea of generalized reflection for layered medium is still applicable when there are multiple modes in each region, except that the reflection and transmission at each interface are characterized by matrices solved from mode-matching. The dimensions of the matrices also match with the number of modes in each

region.

4.2.2 Convergence Tests and Result Verification

The convergence of the generalized reflection matrix is tested with different numbers of modes N_1, N_2, and N_3, in Regions I, II, and III, respectively. We choose the matrix element corresponding to the zeroth order TE reflection coefficient (r_0^{TE}) for analysis. Figure 4.4 shows the convergence of the squared magnitude and phase of r_0^{TE} as N_1 and N_3 ($N_1 = N_3$) increase, with N_2 fixed. Figure 4.5 shows the convergence of the squared magnitude and phase of r_0^{TE} as N_2 increases, with N_1 and N_3 fixed.

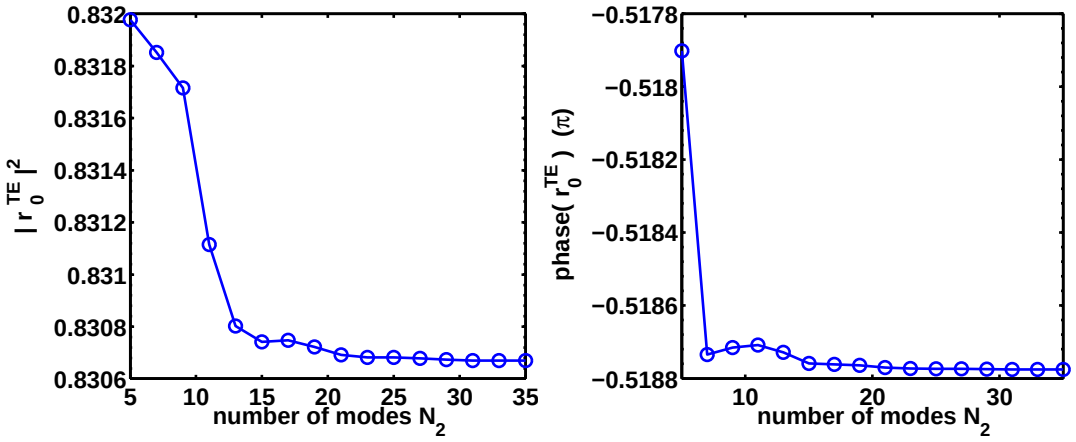


(a) (b)

Figure 4.4: Convergence of the (a) squared magnitude and (b) phase of the zeroth-order generalized reflection coefficient upon the number of modes in Regions I and III ($N_1 = N_3$). Parameters: $N_2 = 11$, $\lambda = 1.4\,\mu$m, $\Lambda = 1.07\,\mu$m, $D = 0.26\,\mu$m, $t_g = 0.195\,\mu$m, and $n_{\mathrm{r}} = 3.164$.

We further simulate the total field distribution using the finite-element method (COMSOL Multiphysics). By fitting the field distribution using the complex reflection coefficient as the fitting parameter, we are able to extract the reflection spectrum.

Figure 4.6(a) shows the total field distribution when a normal incident wave ($\lambda = 1550$ nm) is reflected by a TE-HCG (electric field parallel to HCG bars), where the cross-sections of HCG bars are indicated by the white boxes. Figure 4.6(b) and Figure 4.6(c) show excellent agreement among the three methods for both the magnitude and the phase of the complex reflection



(a) (b)

Figure 4.5: Convergence of the (a) squared magnitude and (b) phase of the zeroth-order generalized reflection coefficient upon the number of modes in Region II N_2. Parameters: $N_1 = N_3 = 13$, $\lambda = 1.4\,\mu$m, $\Lambda = 1.07\,\mu$m, $D = 0.26\,\mu$m, $t_g = 0.195\,\mu$m, and $n_{\mathrm{r}} = 3.164$.

coefficient. The green dashed line indicates $\lambda = \Lambda$. When $\lambda > \Lambda$, we no longer have a single reflected mode since higher order Floquet modes become propagating. The power is not conserved for the zeroth-order mode, and incident power will be carried away by higher order Floquet modes.

4.3 Surface Integral Equations and Periodic Green's Function for 1D Periodic Structures

The scattering problem of a periodic structure can be solved by reducing the problem to only one period. According to the Bloch theorem, the fields satisfy certain periodic properties and the surface integral equations (SIEs) can be reduced to only one period. The summation of all Bloch phase-shifted terms gives rise to the periodic Green's function (PGF).

The method of moments (MOM) is used to solve the surface integral equations. When the surface elements are close to each other, the periodic Green's function and its derivatives converge very slowly. For the self-interaction terms (or the diagonal terms in the impedance matrix), the PGF is singular if summed over the spatial phased-shifted terms, and not converging at all if summed over the spectral modes. A technique to accelerate the convergence

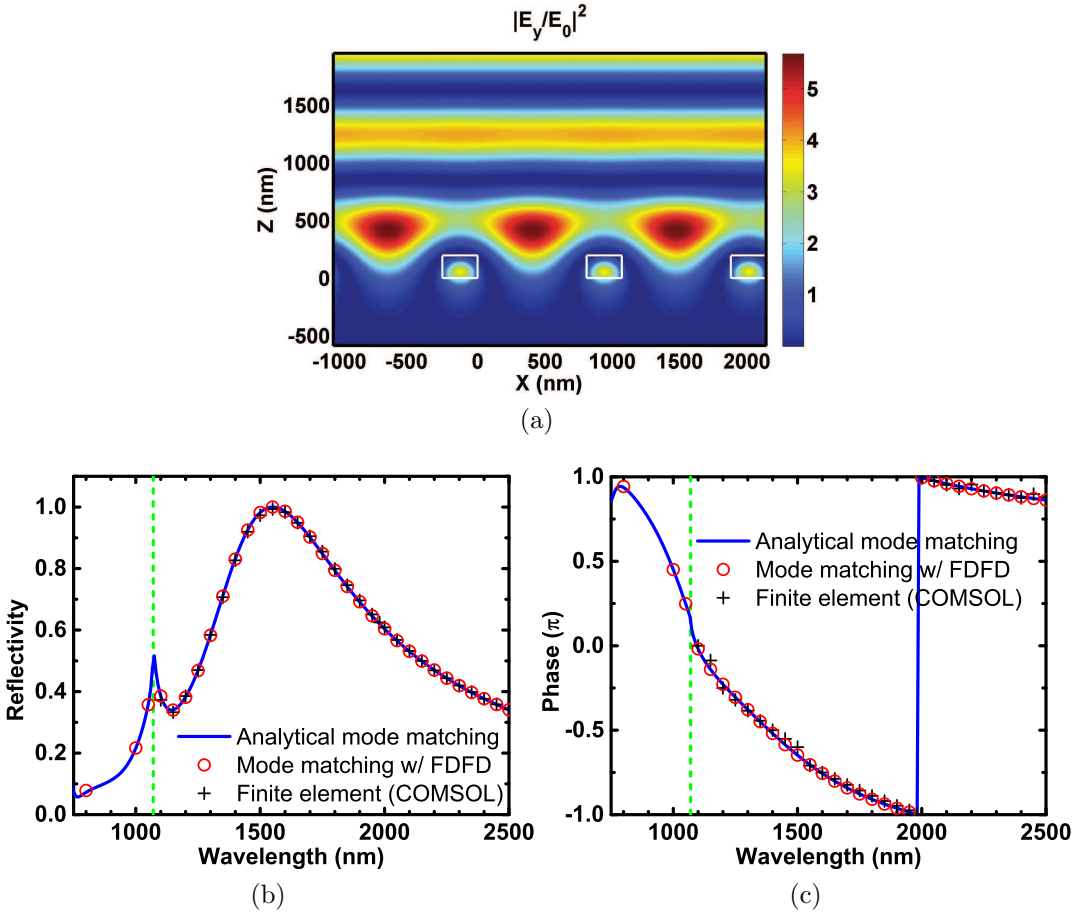


Figure 4.6: (a) Total electric field distribution calculated with the mode-matching method for a 1550 nm normal incident plane wave reflected by a TE-HCG. The HCG parameters are: $n_{\mathrm{r}} = 3.164$, grating period $\Lambda = 1070\,\mathrm{nm}$, thickness $t_g = 195\,\mathrm{nm}$, width $w = 260$ nm. (b) The reflectivity and (c) the phase of the complex reflection coefficient of a TE-HCG calculated using analytical mode-matching method [79] (solid), numerical mode-matching method (circle), and finite-element method with COMSOL Multiphysics (cross). The green dashed line indicates $\lambda = \Lambda$.

of PGFs based on the Poisson summation formula is used and compared to the direct summation over the spectral Floquet modes. PGFs converge with less than 20 terms with the acceleration technique. The derivatives of PGFs converge with less than 40 terms.

4.3.1 Convergence of the Periodic Green's Function and Its Derivatives

The periodic Green's function for 1D gratings can be derived as [81]

$$\begin{aligned} g_p(\mathbf{r},\mathbf{r}') &= \frac{i}{4}\sum_{n=-\infty}^{\infty} e^{ik_{ix}n\Lambda} H_0^{(1)}\left(k\left|\mathbf{r}-\mathbf{r}'-\hat{x}n\Lambda\right|\right) \\ &= \frac{i}{4}\sum_{n=-\infty}^{\infty} e^{ik_{ix}n\Lambda}\frac{1}{\pi}\int_{-\infty}^{\infty} dk_x \frac{1}{k_z} e^{ik_x(x-x'-n\Lambda)+ik_z|z-z'|} \\ &= \frac{i}{4\pi}\int_{-\infty}^{\infty} dk_x \frac{1}{k_z} e^{ik_x(x-x')+ik_z|z-z'|}\sum_{n=-\infty}^{\infty} e^{i(k_{ix}-k_x)n\Lambda} \\ &= \frac{i}{4\pi}\int_{-\infty}^{\infty} dk_x \frac{1}{k_z} e^{ik_x(x-x')+ik_z|z-z'|}\sum_{n=-\infty}^{\infty} \frac{2\pi}{\Lambda}\delta(k_x-k_{ix}-\frac{2n\pi}{\Lambda}) \\ &= \frac{i}{2\Lambda}\sum_n \frac{1}{k_{zn}} e^{ik_{xn}(x-x')+ik_{zn}|z-z'|} \end{aligned} \tag{4.14}$$

where $k_{xn} = k_{ix} + \frac{2n\pi}{\Lambda}$, $k_{zn} = \sqrt{k^2 - k_{xn}^2}$, ($\Im m[k_{zn}] > 0$ for proper waves), and we represent an infinite summation of complex exponentials by a periodic train of Dirac delta functions using the Fourier expansion.

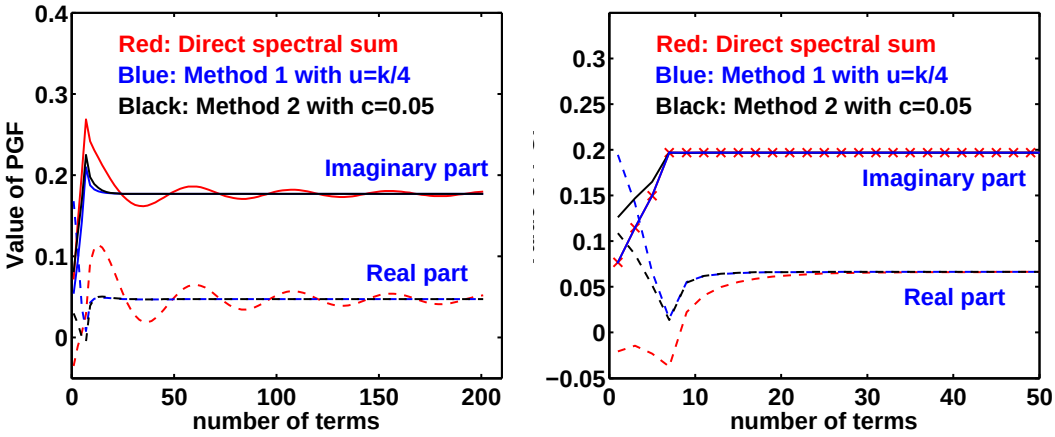
Two acceleration techniques for the summation are used. The first method [82] is to sum in the spectral domain for the fast converging part, but convert the slow converging part into spatial domain using the Poisson summation formula. The slow converging part then becomes modified Bessel functions. The second method [83] is to sum in the spatial domain (all image terms of the Green's function), but the asymptotic behavior is subtracted out by introducing auxiliary terms. These auxiliary terms are then converted back to spectral domain and converge rapidly. Figure 4.7(a) is the convergence of $g_p(x-x', z-z')$ when $x-x' = \Delta s$ and $z-z' = 0$ with Δs being the element length in MOM. Figure 4.7(b) is the convergence of $g_p(x-x', z-z')$ when $x-x' = 0$ and $z-z' = \Delta s$. The first (blue) and the second (black) acceleration

methods both converge much faster than direct spectral summation (red).

Note that for the self-interaction terms ($x - x' = z - z' = 0$) in MOM, the acceleration method will have singularity and the direct summation is never converging. But we can directly integrate the element [84] and the result is

$$P.V. \int_{-\Delta s/2}^{\Delta s/2} ds' g_p|_{x=x',z=z',n=0} = \Delta s \frac{i}{4} \left[1 + \frac{2i}{\pi} \ln \left(\frac{k\Delta s}{4e} \gamma \right) \right] \tag{4.15}$$

where $\gamma = 1.781$.



(a) (b)

Figure 4.7: (a) Convergence of the real and imaginary parts of the periodic Green's function with direct spectral sum (red), acceleration method 1 (blue), acceleration method 2 (black). Here, $x - x' = \Delta s$ and $z - z' = 0$. (b) Convergence of the real and imaginary parts of the periodic Green's function with direct sum (red), acceleration method 1 (blue), acceleration method 2 (black). Here, $x - x' = 0$ and $z - z' = \Delta s$.

The derivatives of the periodic Green's function can be derived as

$$\begin{aligned} \frac{\partial}{\partial x'} g_p(\mathbf{r}, \mathbf{r}') = \frac{i}{4} \sum_{n=-\infty}^{\infty} \Bigg[& e^{ik_{ix} n\Lambda} H_1^{(1)} \left(k\sqrt{(x - x' - n\Lambda)^2 + (z - z')^2} \right) \\ & \cdot \frac{k(x - x' - n\Lambda)}{\sqrt{(x - x' - n\Lambda)^2 + (z - z')^2}} \Bigg] \\ = \frac{1}{2\Lambda} \sum_n & \frac{k_{xn}}{k_{zn}} e^{ik_{xn}(x-x') + ik_{zn}|z-z'|} \end{aligned} \tag{4.16}$$

and

$$\begin{aligned}\frac{\partial}{\partial z'}g_p(\mathbf{r},\mathbf{r}') &= \frac{i}{4}\sum_{n=-\infty}^{\infty}\left[e^{ik_{ix}n\Lambda}H_1^{(1)}\left(k\sqrt{(x-x'-n\Lambda)^2+(z-z')^2}\right)\right. \\ &\left.\cdot\frac{k(z-z')}{\sqrt{(x-x'-n\Lambda)^2+(z-z')^2}}\right] \\ &= \frac{1}{2\Lambda}\sum_n e^{ik_{xn}(x-x')+ik_{zn}|z-z'|}\mathrm{sgn}(z-z')\end{aligned} \tag{4.17}$$

where sgn() is the sign function. The derivatives of the periodic Green's function converge more slowly, and the acceleration method that subtracts the asymptotic behavior of the Hankel function is used as in [83]. The cross-terms in MOM can be derived as

$$\begin{aligned}\frac{\partial}{\partial x'}g_p(\mathbf{r},\mathbf{r}') &= \frac{i}{4}\sum_{n=-\infty}^{\infty}\left[e^{ik_{ix}n\Lambda}H_1^{(1)}\left(k\sqrt{(x-x'-n\Lambda)^2+(z-z')^2}\right)\right. \\ &\left.\cdot\frac{k(x-x'-n\Lambda)}{\sqrt{(x-x'-n\Lambda)^2+(z-z')^2}}\right] \\ &-\frac{i}{4}\sum_{n=-\infty}^{\infty}\left[e^{ik_{ix}n\Lambda}H_1^{(1)}\left(k\sqrt{(x-x'-n\Lambda)^2+(|z-z'|+c\Lambda)^2}\right)\right. \\ &\left.\cdot\frac{k(x-x'-n\Lambda)}{\sqrt{(x-x'-n\Lambda)^2+(|z-z'|+c\Lambda)^2}}\right] \\ &+\frac{1}{2\Lambda}\sum_n\frac{k_{xn}}{k_{zn}}e^{ik_{xn}(x-x')+ik_{zn}(|z-z'|+c\Lambda)}\end{aligned} \tag{4.18}$$

$$\begin{aligned}\frac{\partial}{\partial z'}g_p(\mathbf{r},\mathbf{r}') = \frac{i}{4}\sum_{n=-\infty}^{\infty}\Bigg[e^{ik_{ix}n\Lambda}H_1^{(1)}\left(k\sqrt{(x-x'-n\Lambda)^2+(z-z')^2}\right)\\ \cdot\frac{k(z-z')}{\sqrt{(x-x'-n\Lambda)^2+(z-z')^2}}\Bigg]\\ -\frac{i}{4}\sum_{n=-\infty}^{\infty}\Bigg[e^{ik_{ix}n\Lambda}H_1^{(1)}\left(k\sqrt{(x-x'-n\Lambda)^2+(|z-z'|+c\Lambda)^2}\right)\\ \cdot\frac{k(|z-z'|+c\Lambda)\mathrm{sgn}(z-z')}{\sqrt{(x-x'-n\Lambda)^2+(|z-z'|+c\Lambda)^2}}\Bigg]\\ +\frac{1}{2\Lambda}\sum_n e^{ik_{xn}(x-x')+ik_{zn}(|z-z'|+c\Lambda)}\mathrm{sgn}(z-z')\end{aligned} \tag{4.19}$$

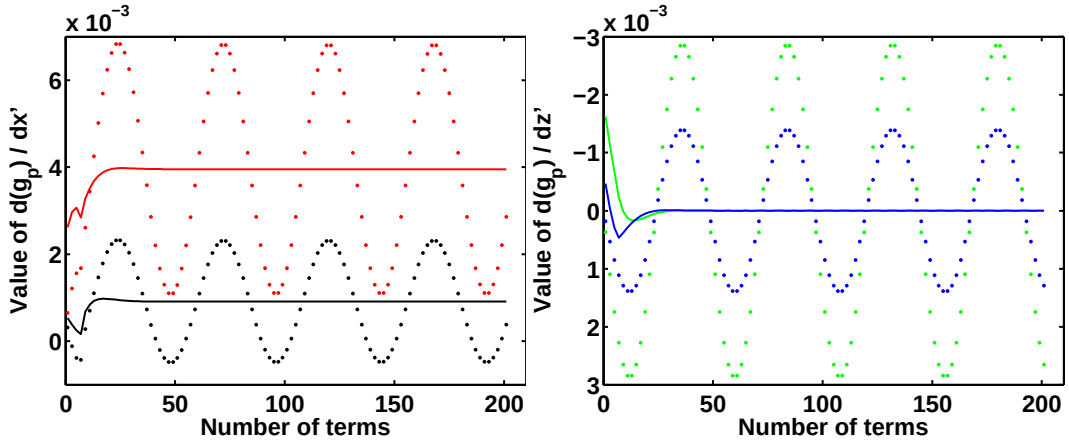
where c is a dimensionless control parameter chosen as 0.05. Figures 4.8(a) and 4.8(b) show the convergence of $\partial g_p/\partial x'$ and $\partial g_p/\partial z'$, respectively, when $x-x'=\Delta s$ and $z-z'=0$. Figures 4.9(a) and 4.9(b) show the convergence of $\partial g_p/\partial x'$ and $\partial g_p/\partial z'$, respectively, when $x-x'=0$ and $z-z'=\Delta s$.

For the self-interaction terms ($x-x'=z-z'=0$) in MOM, we use the expression in Eq. (4.16) and Eq. (4.17) in the spatial domain and directly do the principal value integration within the element. For the $n=0$ term of the derivatives, the integration is zero because of the odd-symmetry,

$$P.V.\int_{-\Delta s/2}^{\Delta s/2}ds'\frac{\partial}{\partial x'}g_p|_{x=x',z=z',n=0} = P.V.\int_{-\Delta s/2}^{\Delta s/2}ds'\frac{\partial}{\partial z'}g_p|_{x=x',z=z',n=0} = 0 \tag{4.20}$$

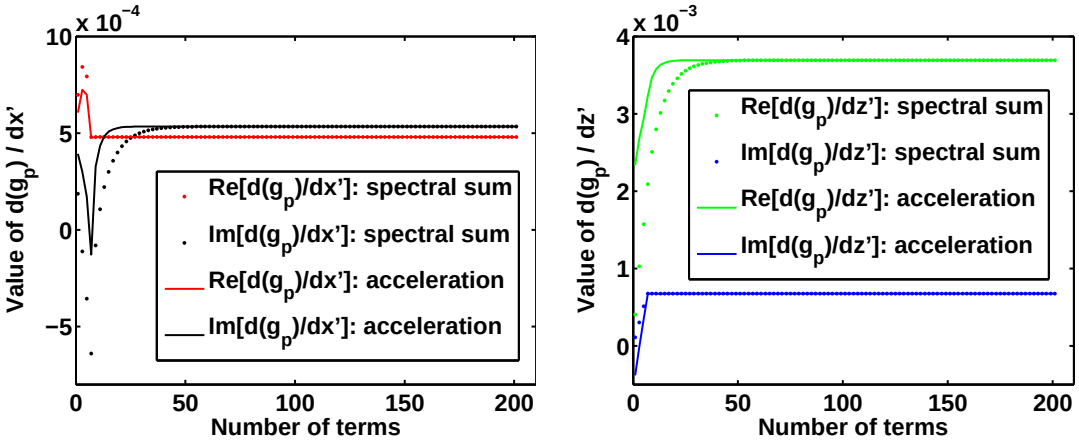
For $n\neq 0$ terms,

$$\begin{aligned}\frac{\partial}{\partial x'}g_p|_{x=x',z=z',n=0} &= -\frac{i}{4}e^{ik_{ix}n\Lambda}H_1^{(1)}\left(k|n\Lambda|\right)\frac{kn}{|n|}\\ \frac{\partial}{\partial z'}g_p|_{x=x',z=z',n=0} &= 0\end{aligned} \tag{4.21}$$



(a) (b)

Figure 4.8: (a) Convergence of the real (red) and imaginary (black) parts of $\partial g_p/\partial x'$ with direct spectral sum (dotted), acceleration (solid). Here, $x - x' = \Delta s$ and $z - z' = 0$. (b) Convergence of the real (green) and imaginary (blue) parts of $\partial g_p/\partial z'$ with direct sum (dotted), acceleration (solid). Here, $x - x' = \Delta s$ and $z - z' = 0$.



(a) (b)

Figure 4.9: (a) Convergence of the real (red) and imaginary (black) parts of $\partial g_p/\partial x'$ with direct spectral sum (dotted), acceleration (solid). Here, $x - x' = 0$ and $z - z' = \Delta s$. (b) Convergence of the real (green) and imaginary (blue) parts of $\partial g_p/\partial z'$ with direct sum (dotted), acceleration (solid). Here, $x - x' = 0$ and $z - z' = \Delta s$.

4.3.2 TE (s-Polarized) Scattering by 1D Periodic Dielectric Structures

Two sets of SIEs are required to solve the surface fields since both tangential electric field and magnetic field are present. Based on the formulation of SIEs and the PGFs in Appendix A, the scattering problem becomes a matrix equation with given incidence conditions. Once we have solved the MOM expansion coefficients, we can find the tangential electric field and its normal derivative. Further we can find the complex mode expansion coefficients r_n and t_n for each diffraction order n in Region I and Region II, respectively. The total field can then be calculated as

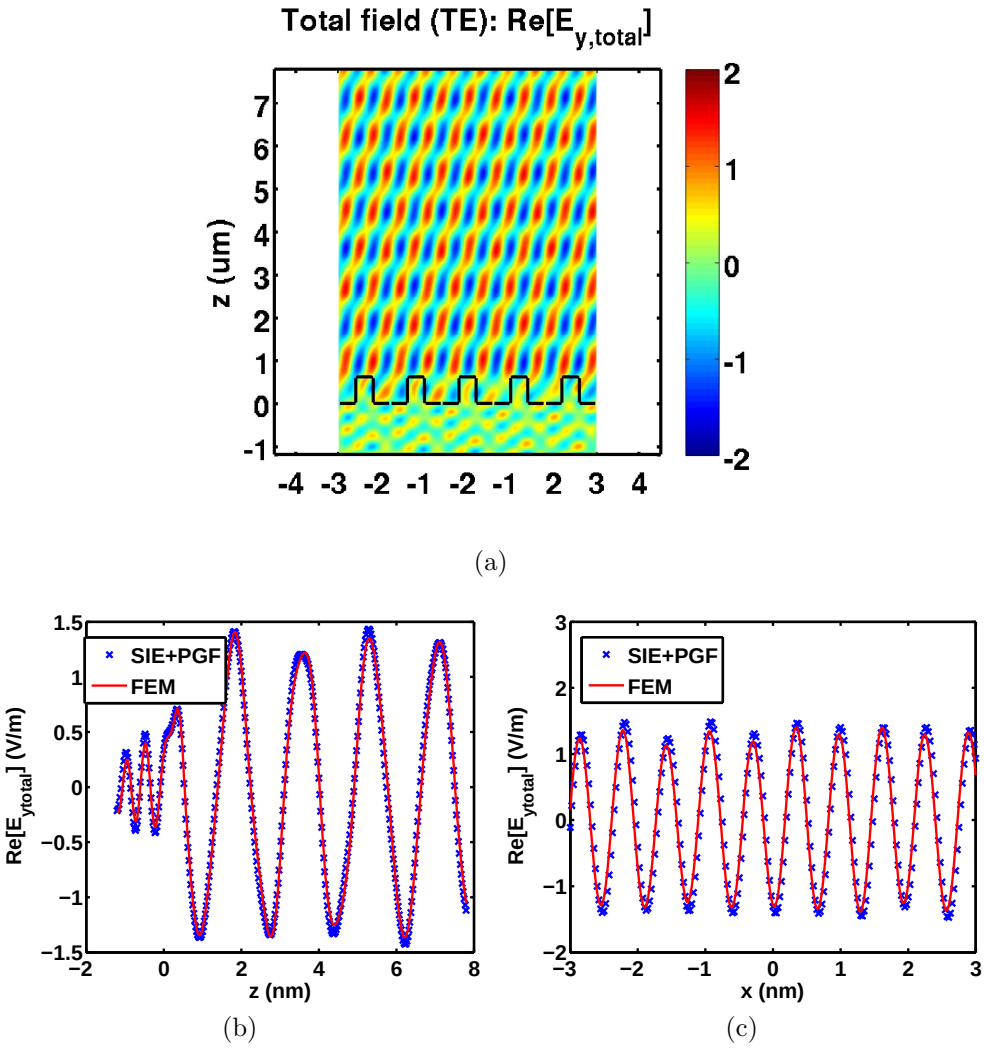
$$\begin{cases} E_y(\mathbf{r}) = E_{iy}(\mathbf{r}) + \sum_n r_n e^{i\mathbf{k}_{1n}^{+}\cdot\mathbf{r}} & \mathbf{r} \in V_1 \\ E_y(\mathbf{r}) = \sum_n t_n e^{i\mathbf{k}_{2n}^{-}\cdot\mathbf{r}} & \mathbf{r} \in V_2 \end{cases} \tag{4.22}$$

where the expressions for the coefficients r_n and t_n are given in Appendix A, and $\mathbf{k}_n^{\pm} = \hat{x}k_{xn} \pm \hat{z}k_{zn}$ with k_{zn} being real and positive for propagating modes.

Figure 4.10 (a) shows the total electric field for a plane wave incident upon a 1D grating calculated using SIEs with PGFs. The grating thickness t_g is 600 nm, width D is 400 nm, period Λ is 1200 nm, the incident polar angle θ is 70 degrees, the incident azimuthal angle ϕ is 0, $\varepsilon_{\mathrm{r}1} = 1$, $\varepsilon_{\mathrm{r}2} = 2$, $\mu_{\mathrm{r}1} = \mu_{\mathrm{r}2} = 1$, and the free-space wavelength λ_0 is 600 nm. We also compare the results with the those from finite-element methods (FEM) using COMSOL Multiphysics on the $x = 0$ and $z = 1\,\mu$m lines. As shown in Fig. 4.10(b) and Fig. 4.10(c), the two methods show good agreement. The numerical error in solving the surface fields comes from the imperfect expansion as MOM basis functions. The numerical error in the total field also comes from using a finite number of diffraction orders in Eq. (4.22). When the observation point is close to the grating, the contribution from evanescent waves can be strong. In the example we are using four propagating modes and six evanescent modes.

4.3.3 TM (p-Polarized) Scattering by 1D Periodic Dielectric Structures

Following the same procedure as the TE case, we can obtain the complex expansion coefficient of each diffraction order for the TM case. And we can



(a)

(b)

(c)

Figure 4.10: (a) Real part of the total electric field ($\Re e[E_y]$) for a TE scattering problem with an oblique plane-wave incidence and a 1D dielectric grating, calculated using surface integral equations (SIEs) with periodic Green's function. (b) Comparison between $\Re e[E_y]$ along the $x = 0$ line calculated by SIEs and finite-element method (FEM). (c) Comparison between $\Re e[E_y]$ along the $z = 1\,\mu$m line calculated by SIEs and FEM.

use the same summation as in Eq. (4.22) to find the total field, except the field component is changed to H_y and the expansion coefficients r_n and t_n for the TM case are included in Appendix A.

Figure 4.11 (a) shows the total magnetic field for a plane wave incident upon a 1D grating calculated using SIEs with PGFs. The grating parameters and the incidence conditions are the same as the previous TE case. We also compare the results with the those from finite-element methods (FEM) using

COMSOL Multiphysics on the $x = 0$ and $z = 1\,\mu$m lines. As shown in Fig. 4.11(b) and Fig. 4.11(c), the two methods again show good agreement.

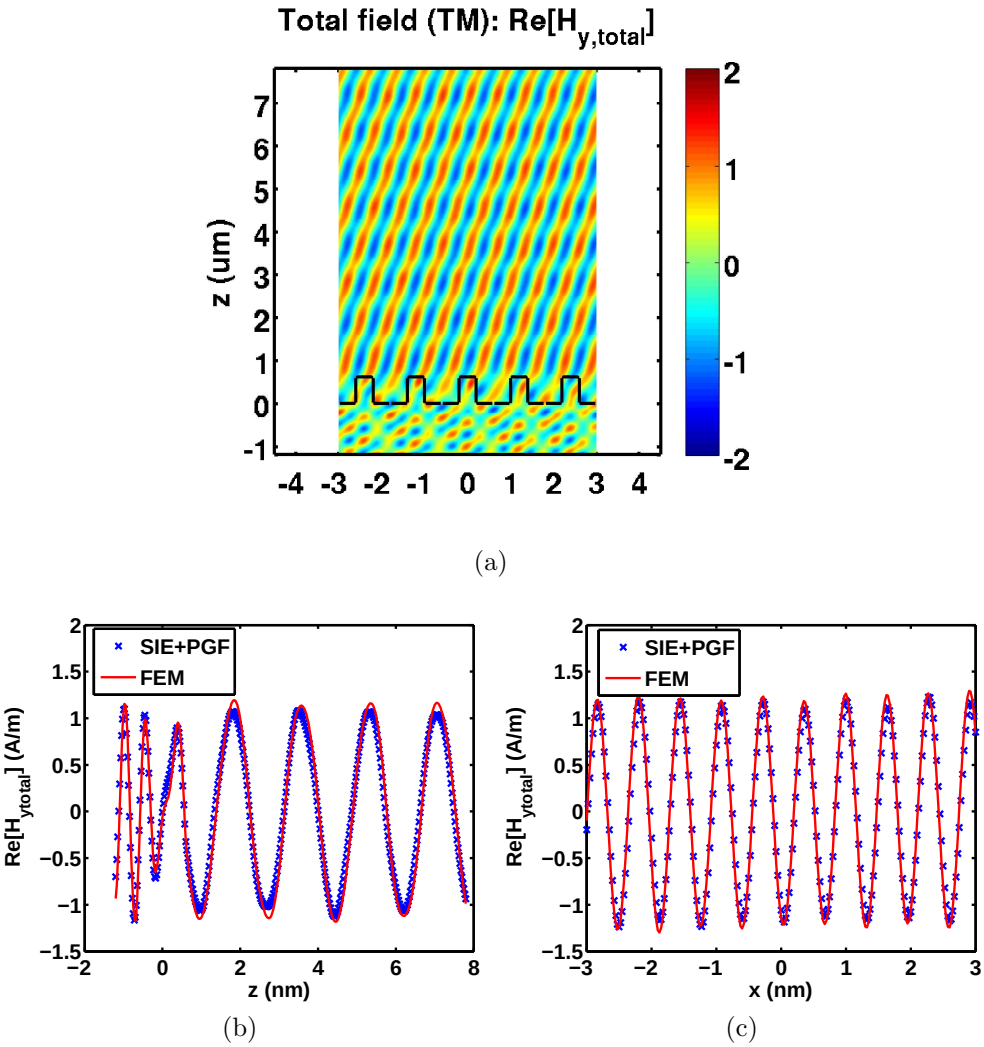


Figure 4.11: (a) Real part of the total magnetic field ($\Re e[H_y]$) for a TM scattering problem with an oblique plane-wave incidence and a 1D dielectric grating, calculated using surface integral equations (SIEs) with periodic Green's function. (b) Comparison between $\Re e[H_y]$ along the $x = 0$ line calculated by SIEs and finite-element method (FEM). (c) Comparison between $\Re e[H_y]$ along the $z = 1\,\mu$m line calculated by SIEs and FEM.

4.4 Optical Properties of 2D High-Contrast Gratings

For 2D HCGs, structures are periodic in two directions and the modes supported by the grating have hybrid polarizations, that is, they are not separable into TE and TM. The modes in 2D HCGs are also more closely-packed spectrally. In other words, with a given incident wave, many more modes can be excited in a 2D HCG than in a 1D HCG. The diffracted modes, also known as the Floquet modes, have one extra degree of freedom for the propagation direction, which makes it very difficult to predict the optical properties of a 2D HCG. Moreover, there are many more structural parameters in 2D HCGs, which largely complicate the design and optimization procedure. Up to now, majority of the gratings in use are still one-dimensional.

However, there are applications which require beam-steering or focusing in both directions, or generating optical vortices [72, 76]. In these cases, 2D HCGs are more advantageous. Extensive experimental work has been done in demonstrating the functionalities of 2D gratings [66, 72]. In this work, we examine the physics of 2D HCGs and the mechanisms which allow high-performance operation. The dual-mode analysis we propose can largely simplify the searching of the initial design and the subsequent optimization. A top-down design procedure is provided. To model 2D HCGs, the analytical mode-matching method [35, 79], which was previously used in 1D HCGs, is rather inefficient and requires complex root-searching for coupled nonlinear equations. An in-house developed rigorous coupled-wave analysis (RCWA) [85, 86] program is shown to be very convenient for design purposes, and both the convergence and accuracy are verified.

We first study the underlying physics of 2D high-contrast gratings. Figure 4.12(a) shows a general 2D grating on a substrate (optional for suspended grating membranes). The incident plane wave is characterized by the polar angle θ and the azimuthal angle ϕ of its wave vector,

$$\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz}) = (k_0 \sin\theta\cos\phi, k_0 \sin\theta\sin\phi, k_0\cos\theta) \tag{4.23}$$

The polarization of the incident electric field should also be characterized by two angles,

$$\mathbf{E}_i = (E_0 \sin\alpha_1\cos\alpha_2, E_0 \sin\alpha_1\sin\alpha_2, E_0\cos\alpha_1) \tag{4.24}$$

Yet the electric field is transverse to the wave vector for propagating waves, thus one angle α is sufficient to describe the polarization. For oblique incidence, it is convenient to decompose the fields into s- and p-polarizations, which will be discussed later.

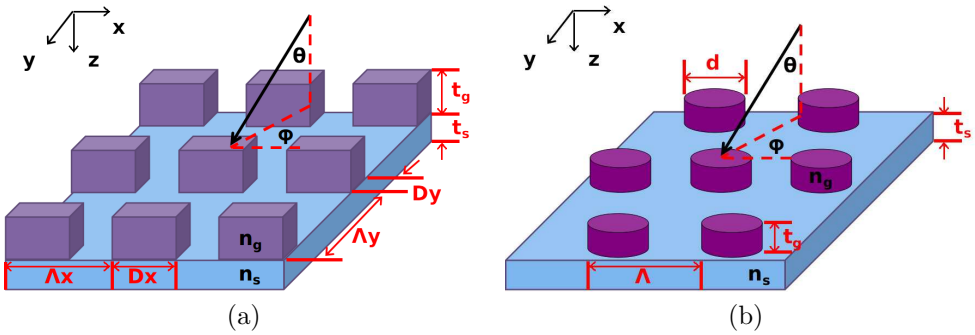


Figure 4.12: (a) Schematic diagram of a 2D high-contrast grating on a rectangular lattice with a substrate. The plane-wave incident polar and azimuthal angles are θ and ϕ, respectively. Parameters: Λ_x and Λ_y (periods in $\hat{x}$ and $\hat{y}$), D_x and D_y (box widths in $\hat{x}$ and $\hat{y}$), n_g and n_s (grating and substrate indices), and t_g and t_s (grating and substrate thicknesses). (b) Schematic diagram of a 2D high-contrast grating on a hexagonal lattice with a substrate. Parameters: period Λ, rod diameter d, grating and substrate indices n_g and n_s, grating and substrate thicknesses t_g and t_s.

Design parameters include grating periods and widths in both $\hat{x}$ and $\hat{y}$ directions (Λ_x, Λ_y, D_x, and D_y), as well as thicknesses and refractive indices of both the high-index material and the substrate (n_g, n_s, t_g, and t_s), as shown in Fig. 4.12(a). We define duty cycles for the high-index material as $\eta_x = D_x/\Lambda_x$ and $\eta_y = D_y/\Lambda_y$.

For applications where optical fields express rotational symmetry, 2D circular gratings on a hexagonal lattice, as shown in Fig. 4.12(b), are more advantageous. Design parameters are simplified to grating period (Λ), rod diameter (d), thicknesses and refractive indices of both the grating and substrate (t_g, t_s, n_g and n_s). The duty cycle is defined as $\eta = d/\Lambda$. Gratings on a hexagonal lattice, being more closely packed, require $\Lambda < 2\lambda_0/\sqrt{3}$ in order to have only the main harmonic being propagating, as compared to $\Lambda_{x,y} < \lambda_0$ for gratings on a rectangular lattice.

We first look at the incident and transmitted regions, which are typically homogeneous. According to the Floquet theorem, the scattered field consists of Floquet modes with different orders of tangential wave vectors, as a result

of the phase matching between the homogeneous region and the 2D grating with the translational symmetry. In rectangular lattice, the (m,n)-th order Floquet mode is a plane wave with a wave vector being

$$\begin{aligned}
&\mathbf{k}_{(m,n)} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z, \quad k_x = k_{ix} + G_{x,m}, \quad k_y = k_{iy} + G_{y,n}, \\
&G_{x,m} = \frac{m\pi}{\Lambda_x}, \quad m = 0, \pm 1, \pm 2, ..., \pm N_x, \\
&G_{y,n} = \frac{n\pi}{\Lambda_y}, \quad n = 0, \pm 1, \pm 2, ..., \pm N_y, \\
&k_z = \pm\sqrt{\frac{4\pi^2 n_{\mathrm{r}}^2}{\lambda^2} - k_x^2 - k_y^2}
\end{aligned} \tag{4.25}$$

where n_{r} is the index of the incident or transmitted region. Thus the total number of modes is $N = (2N_x + 1)(2N_y + 1)$. Using the RCWA formulation, we can further expand the transverse fields of the i-th eigenmode in the grating layer in terms of the same Floquet basis,

$$\begin{aligned}
\mathbf{E}_t^{(i)} &= e^{iK_z^{(i)}z} \left(\sum_{m,n} (\hat{x}\widetilde{E}_{x,(m,n)}^{(i)} + \hat{y}\widetilde{E}_{y,(m,n)}^{(i)}) e^{iG_{x,m}x + iG_{y,n}y} \right) e^{ik_{ix}x + ik_{iy}y} \\
&= e^{iK_z^{(i)}z} \sum_{\mathbf{G}} \widetilde{\mathbf{E}}_{t,\mathbf{G}}^{(i)} \cdot e^{i\mathbf{G}\cdot\mathbf{r}} e^{ik_{ix}x + ik_{iy}y}
\end{aligned} \tag{4.26}$$

where $\mathbf{G} = (G_{x,m}, G_{y,n}, 0)$, $\widetilde{\mathbf{E}}_{t,\mathbf{G}}^{(i)} = (\widetilde{E}_{x,(m,n)}^{(i)}, \widetilde{E}_{y,(m,n)}^{(i)})$, and the tilde symbols indicate the spectral coefficients. Here, $K_z^{(i)}$ is the propagation constant for the i-th eigenmode. Since the displacement field in spatial domain can be related to the electric field as

$$\mathbf{D}^{(i)}(\mathbf{r}) = \int \varepsilon(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')\mathbf{E}^{(i)}(\mathbf{r}')d\mathbf{r}' \tag{4.27}$$

then in spectral domain we can write

$$\widetilde{\mathbf{D}}_{\mathbf{G}}^{(i)} = \sum_{\mathbf{G}'} \varepsilon_{\mathbf{G}\mathbf{G}'} \widetilde{\mathbf{E}}_{\mathbf{G}'}^{(i)} \tag{4.28}$$

where

$$\varepsilon_{\mathbf{G}\mathbf{G}'} = \frac{1}{S} \int \varepsilon(\mathbf{r}) \exp[i((\mathbf{G}' - \mathbf{G})\mathbf{r})] dxdy \tag{4.29}$$

The above integration is performed over the unit cell with an area S. For a

rectangular unit cell with a rectangular core medium ε_2 and a surrounding medium ε_1, we have an analytical expression [86],

$$\varepsilon_{\mathbf{GG'}} = \varepsilon_1 \delta_{\mathbf{GG'}} + (\varepsilon_2 - \varepsilon_1) \frac{\sin\left[\pi(m-m')\frac{D_x}{\Lambda_x}\right] \sin\left[\pi(n-n')\frac{D_y}{\Lambda_y}\right]}{\pi^2 (m-m')(n-n')} \tag{4.30}$$

For 2D circular materials ε_2 with a radius r and a surrounding material ε_1 [87],

$$\varepsilon_{\mathbf{GG'}} = \begin{cases} \varepsilon_1 + (\varepsilon_2 - \varepsilon_1)\frac{\pi r^2}{S} & \text{for } \mathbf{G} = \mathbf{G'} \\ 2(\varepsilon_2 - \varepsilon_1)\frac{J_1(|\mathbf{G}-\mathbf{G'}|r)}{|\mathbf{G}-\mathbf{G'}|r}\frac{\pi r^2}{S} & \text{for } \mathbf{G} \neq \mathbf{G'} \end{cases} \tag{4.31}$$

where J_1 is the Bessel function of the first order. If the 2D grating is on a hexagonal lattice with a period Λ, we can still define a rectangular unit cell with a fixed aspect ratio $\sqrt{3} : 1$. Then there will be equivalently two phase-shifted core structures in a cell. If the core structure is circular, we can write

$$\varepsilon_{\mathbf{GG'}} = \begin{cases} \varepsilon_1 + (\varepsilon_2 - \varepsilon_1)\frac{2\pi r^2}{S} & \text{for } \mathbf{G} = \mathbf{G'} \\ 2(\varepsilon_2 - \varepsilon_1)\frac{J_1(|\mathbf{G}-\mathbf{G'}|r)}{|\mathbf{G}-\mathbf{G'}|r}\frac{\pi r^2}{S}\left(1 + e^{[i\pi(m-m')+i\pi(n-n')]}\right) & \text{for } \mathbf{G} \neq \mathbf{G'} \end{cases} \tag{4.32}$$

Substituting the Fourier expansion Eq. (4.26) and Eq. (4.28) into the vector wave equation $\nabla \times \nabla \times (\mathbf{E}_t^{(i)} + \hat{z}E_z^{(i)}) = \frac{\omega^2}{c^2}\mathbf{D}^{(i)}$, and eliminating $E_z^{(i)}$ using the divergence theorem, we obtain the eigen-equation in the spectral domain [86],

$$\bar{\mathbf{M}} \cdot \widetilde{\mathbf{E}}_t^{(i)} = (K_z^{(i)})^2 \bar{\mathbf{N}} \cdot \widetilde{\mathbf{E}}_t^{(i)} \tag{4.33}$$

where $\bar{\mathbf{M}}$ and $\bar{\mathbf{N}}$ are $2N \times 2N$ matrices dependent on the frequency ω and refractive index $n_{\mathrm{r}}(x, y)$, $(K_z^{(i)})^2$ is the eigenvalue, and $\widetilde{\mathbf{E}}_t^{(i)}$ is a $2N$-dimensional vector consisting of spectral coefficients,

$$\widetilde{\mathbf{E}}_t^{(i)} = \left(\widetilde{E}_x^{\mathbf{G}_1}, ..., \widetilde{E}_x^{\mathbf{G}_N}, \widetilde{E}_y^{\mathbf{G}_1}, ..., \widetilde{E}_y^{\mathbf{G}_N}\right)^T \tag{4.34}$$

The $\bar{\mathbf{M}}$ and $\bar{\mathbf{N}}$ can be written in terms of block matrices as [86]

$$\bar{\mathbf{M}} = \begin{bmatrix} \frac{\omega^2}{c^2}\varepsilon_{\mathbf{GG'}} - \delta_{\mathbf{GG'}}(k_{iy}+G_y)^2 & (k_{ix}+G_x)(k_{iy}+G_y)\delta_{\mathbf{GG'}} \\ (k_{ix}+G_x)(k_{iy}+G_y)\delta_{\mathbf{GG'}} & \frac{\omega^2}{c^2}\varepsilon_{\mathbf{GG'}} - \delta_{\mathbf{GG'}}(k_{ix}+G_x)^2 \end{bmatrix} \tag{4.35}$$

$$\bar{\mathbf{N}} = \begin{bmatrix} \delta_{\mathbf{GG'}} - (k_{ix}+G_x)Z^{-1}_{\mathbf{GG'}}(k_{ix}+G'_x) & -(k_{ix}+G_x)Z^{-1}_{\mathbf{GG'}}(k_{iy}+G'_y) \\ -(k_{iy}+G_x)Z^{-1}_{\mathbf{GG'}}(k_{ix}+G'_x) & \delta_{\mathbf{GG'}} - (k_{iy}+G_y)Z^{-1}_{\mathbf{GG'}}(k_{iy}+G'_y) \end{bmatrix} \tag{4.36}$$

where

$$Z_{\mathbf{GG'}} = -\frac{\omega^2}{c^2}\varepsilon_{\mathbf{GG'}} + \delta_{\mathbf{GG'}}\left[(k_{ix}+G_x)^2 + (k_{ix}+G_x)^2\right] \tag{4.37}$$

In general, the $\bar{\mathbf{M}}$ and $\bar{\mathbf{N}}$ are not positive definite and Hermitian, and we can have eigenvalues $(K_z^{(i)})^2$ not being real, that is, $K_z^{(i)}$ being neither purely real or imaginary but complex, which corresponds to leaky waves in periodic waveguides.

The total transverse field in the grating layer can be expressed in terms of eigenmodes as

$$\widetilde{\mathbf{E}}_t = \sum_i A_i \widetilde{\mathbf{E}}_t^{(i)} = \bar{\boldsymbol{\mathcal{E}}}_t \cdot \mathbf{A} \tag{4.38}$$

where $\bar{\boldsymbol{\mathcal{E}}}_t$ is a $2N \times 2N$ matrix with columns being eigenvectors $\widetilde{\mathbf{E}}_t^{(i)}$, and $\mathbf{A}$ contains $2N$-dimensional eigenmode expansion coefficients. From the eigenequations in Eq. (4.33) one can obtain forward and backward propagation constants for each eigenmode $(\pm K_z^{(i)})$.

Similar to Eq. (4.38) we can expand the transverse magnetic field in terms of the magnetic components $\widetilde{\mathbf{H}}_t^{(i)}$ of the eigenmodes ($\widetilde{\mathbf{H}}_t^{(i)}$ obtained from $\widetilde{\mathbf{E}}_t^{(i)}$).

$$\widetilde{\mathbf{H}}_t = \frac{c}{\omega}(\bar{\mathbf{C}}\bar{\boldsymbol{\mathcal{E}}}_t\bar{\mathbf{K}}) \cdot \mathbf{A} = \bar{\boldsymbol{\mathcal{H}}}_t \cdot \mathbf{A} \tag{4.39}$$

where

$$\bar{\bar{\mathbf{C}}} = \begin{bmatrix} (k_{iy}+G_x)Z^{-1}_{\mathbf{GG}'}(k_{ix}+G'_x) & -\delta_{\mathbf{GG}'}+(k_{iy}+G_y)Z^{-1}_{\mathbf{GG}'}(k_{iy}+G'_y) \\ \delta_{\mathbf{GG}'}-(k_{ix}+G_x)Z^{-1}_{\mathbf{GG}'}(k_{ix}+G'_x) & -(k_{ix}+G_x)Z^{-1}_{\mathbf{GG}'}(k_{iy}+G'_y) \end{bmatrix}$$

$$\bar{\mathbf{K}} = \begin{pmatrix} K_z^{(1)} & 0 & 0 & \dots & 0 \\ 0 & K_z^{(2)} & 0 & \dots & 0 \\ \dots & & & & \\ 0 & \dots & & & K_z^{(2N)} \end{pmatrix} \tag{4.40}$$

We can then express the z-dependent transverse electric and magnetic fields including forward and backward propagation as

$$\begin{aligned} \begin{bmatrix} \widetilde{\mathbf{E}}_t(z) \\ \widetilde{\mathbf{H}}_t(z) \end{bmatrix} &= \begin{bmatrix} \bar{\boldsymbol{\mathcal{E}}}_t & \bar{\boldsymbol{\mathcal{E}}}_t \\ \bar{\boldsymbol{\mathcal{H}}}_t & -\bar{\boldsymbol{\mathcal{H}}}_t \end{bmatrix} \begin{bmatrix} \mathbf{A}^+(z) \\ \mathbf{A}^-(z) \end{bmatrix} \\ &= \begin{bmatrix} \bar{\boldsymbol{\mathcal{E}}}_t & \bar{\boldsymbol{\mathcal{E}}}_t \\ \bar{\boldsymbol{\mathcal{H}}}_t & -\bar{\boldsymbol{\mathcal{H}}}_t \end{bmatrix} \begin{bmatrix} e^{i\bar{\mathbf{K}}_z z} & 0 \\ 0 & e^{-i\bar{\mathbf{K}}_z z} \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{A}^- \end{bmatrix} \end{aligned} \tag{4.41}$$

where $\mathbf{A}^{\pm} = [A_1^{\pm}, ..., A_{2N}^{\pm}]$. Here, $e^{\pm i\bar{\mathbf{K}}_z z}$ are diagonal matrices with elements being $\exp(\pm iK_z^{(i)}z)$. Matching the boundary conditions at the interface between two regions (e.g. I for air and II for grating), we have

$$\begin{bmatrix} \mathbf{A}^+_{\mathrm{II}} \\ \mathbf{A}^-_{\mathrm{II}} \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\mathcal{E}}}^{\mathrm{II}}_t & \bar{\boldsymbol{\mathcal{E}}}^{\mathrm{II}}_t \\ \bar{\boldsymbol{\mathcal{H}}}^{\mathrm{II}}_t & -\bar{\boldsymbol{\mathcal{H}}}^{\mathrm{II}}_t \end{bmatrix}^{-1} \begin{bmatrix} \bar{\boldsymbol{\mathcal{E}}}^{\mathrm{I}}_t & \bar{\boldsymbol{\mathcal{E}}}^{\mathrm{I}}_t \\ \bar{\boldsymbol{\mathcal{H}}}^{\mathrm{I}}_t & -\bar{\boldsymbol{\mathcal{H}}}^{\mathrm{I}}_t \end{bmatrix} \begin{bmatrix} \mathbf{A}^+_{\mathrm{I}} \\ \mathbf{A}^-_{\mathrm{I}} \end{bmatrix} = \left[\bar{\boldsymbol{\mathcal{T}}}^{\mathrm{I-II}}\right]_{4N\times 4N} \begin{bmatrix} \mathbf{A}^+_{\mathrm{I}} \\ \mathbf{A}^-_{\mathrm{I}} \end{bmatrix} \tag{4.42}$$

where $\bar{\boldsymbol{\mathcal{T}}}^{\mathrm{I-II}}$ is the transfer matrix at the I-II interface. For a given incident $\mathbf{A}_{\mathrm{inc}}$, we can solve for the reflected $\mathbf{A}_{\mathrm{ref}}$ and transmitted $\mathbf{A}_{\mathrm{trans}}$, once we find the total transfer matrix or scattering matrix, which satisfy

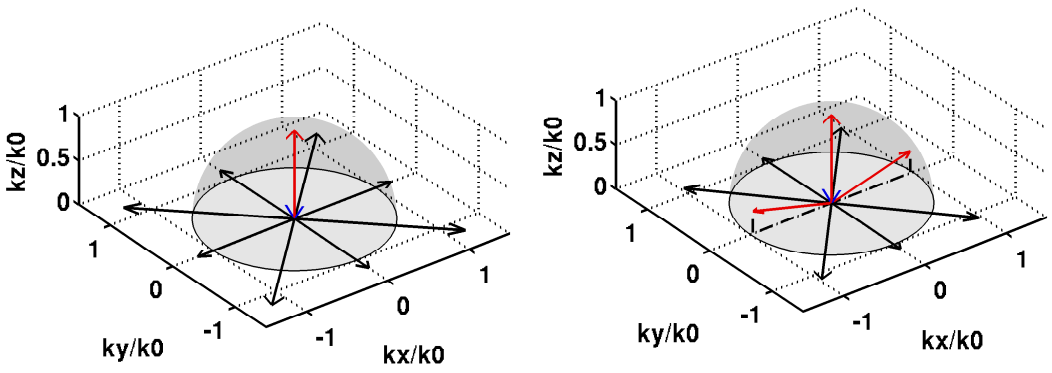
$$\begin{bmatrix} \mathbf{A}_{\mathrm{trans}} \\ 0 \end{bmatrix} = \left[\bar{\boldsymbol{\mathcal{T}}}^{\mathrm{total}}\right]_{4N\times 4N} \begin{bmatrix} \mathbf{A}_{\mathrm{inc}} \\ \mathbf{A}_{\mathrm{ref}} \end{bmatrix} \tag{4.43}$$

and

$$\begin{bmatrix} \mathbf{A}_{\text{trans}} \\ \mathbf{A}_{\text{ref}} \end{bmatrix} = \left[\bar{\boldsymbol{\mathcal{S}}}^{\text{total}}\right]_{4N\times 4N} \begin{bmatrix} \mathbf{A}_{\text{inc}} \\ 0 \end{bmatrix} \tag{4.44}$$

respectively. We can find the coefficients $\widetilde{\mathbf{E}}_t$ and $\widetilde{\mathbf{H}}_t$ for each spectral order using Eq. (4.38), and the spatial representations $\mathbf{E}_t$ and $\mathbf{H}_t$ using the Fourier series as in Eq. (4.26).

The phase matching diagram for a 2D grating is a matching sphere, as shown in Fig. 4.13. We can see that, when the wavelength is larger than both Λ_x and Λ_y, only the (0,0) diffracted mode is propagating. Otherwise there will be higher diffraction orders.



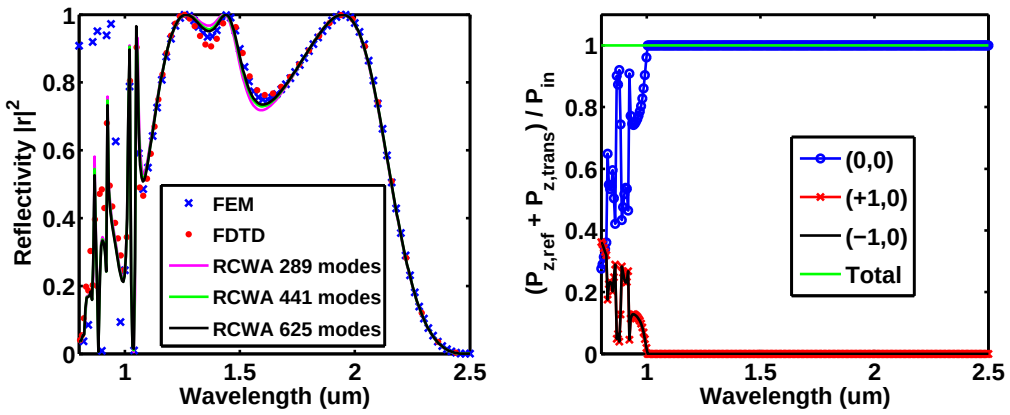
(a) (b)

Figure 4.13: (a) Phase matching diagram for $\Lambda_x = \Lambda_y = 1.3\,\mu$m and $\lambda = 1.55\,\mu$m. (b) Phase matching diagram for $\Lambda_x = 1.4\,\mu$m, $\Lambda_y = 1.4\,\mu$m, and $\lambda = 1.55\,\mu$m.

Figure 4.14(a) shows the (0,0)-th order reflection spectra for a 2D HCG under normal incidence. Results are calculated with the finite-element method (FEM) using COMSOL Multiphysics, the finite-difference time-domain method (FDTD) using Lumerical Solutions, as well as an in-house developed 2D rigorous coupled-wave analysis (RCWA) program. Good agreement is shown among three methods and good convergence of our RCWA program upon total number of 2D spectral orders (N) is observed. Moreover, calculating each frequency point using RCWA with $N = 625$ is only 90 seconds on a personal computer with a Intel Core i7-3520M processor, whereas the computation

cost is much heavier for FDTD and FEM due to the full discretization of the 3D solution domain. When $\lambda < \max\{\Lambda_x, \Lambda_y\}$, higher spectral orders become propagating, and extracting each spectral order accurately from the total field solution in FDTD or FEM becomes more challenging.

We also verify that our RCWA simulation is energy-conserving. We calculate the scattered power flux (reflection and transmission) normalized by the incident power flux for the (0,0), (+1,0), (−1,0) modes, as shown in Fig. 4.14(b). The power scattered into the (0,0) mode drops from unity when $\lambda < 1\,\mu$m, indicating power transfering into higher diffraction orders. Nonetheless, the total scattered power remains unity, satisfying the conservation of energy.



(a) (b)

Figure 4.14: (a) Comparison among the (0,0)-th order reflectivity spectra of a 2D high-contrast grating (HCG) under normal incidence calculated using finite-element method (blue star), finite-difference time-domain (red dots), and rigorous coupled-wave analysis (RCWA) using $N = 289$ (magenta), $N = 441$ (green), and $N = 625$ (blue). HCG parameters are $\Lambda_x = 1\,\mu$m, $\Lambda_y = 0.5\,\mu$m, $D_x = 0.6\,\mu$m, $D_y = 0.3\,\mu$m, and $t_g = 0.5\,\mu$m. (b) Spectra of normalized scattered power in the $\hat{z}$-direction for (0,0), (+1,0), and (−1,0) spectral orders (blue, red, and black, respectively), and the sum of the three spectra (green).

We can also study the 2D HCG properties under oblique incidence. The field polarization is defined relative to the incidence plane, which is formed by the surface normal and the incidence wave vector. We use s- and p-polarization for electric field perpendicular and parallel to the incidence plane, respectively. Figure 4.15(a) shows the (0,0)-th order reflectivity and

transmissivity as functions of the incidence polar angle θ with wavelength $\lambda = 2\,\mu$m and incidence azimuthal angle $\phi = 0$. In this case, all higher spectral orders remain evanescent for the whole 90-degree range of θ. Thus the sum of reflectivity and transmissivity remains unity. Results from our 2D RCWA program and FEM software agree very well.

As the wavelength decreases, the higher spectral orders may be propagating. The critical condition for the (m,n)-th spectral order to transition between propagating and evanescent is

$$\left(k_0 \sin\theta\cos\phi + \frac{m\pi}{\Lambda_x}\right)^2 + \left(k_0 \sin\theta \sin\phi + \frac{n\pi}{\Lambda_y}\right)^2 = k_0^2 \tag{4.45}$$

where $0 < \theta < 90°$. If we consider the incidence plane being the xz-plane (i.e. $\phi = 0$), wavelength being fixed, and $\lambda > \Lambda_y$, then the cutoff angles from the above critical condition become

$$\begin{aligned} \theta_c &= \sin^{-1}\left(1 - \frac{|m|\lambda}{\Lambda_x}\right) \quad \text{for } (+|m|, 0) \text{ to disappear} \\ \theta_c &= \sin^{-1}\left(\frac{|m|\lambda}{\Lambda_x} - 1\right) \quad \text{for } (-|m|, 0) \text{ to appear} \end{aligned} \tag{4.46}$$

Figures 4.15(b) and 4.15(c) show the normalized scattered power in the $\hat{z}$-direction for different spectral orders under s- and p-polarized oblique plane wave incidence, respectively. The cutoff angles for the (−2,0), (+1,0), and (−3,0) modes are calculated from Eq. (4.46) to be 11.54°, 23.58°, and 53.13°, which agree with the results indicated by the arrows in Fig. 4.15(b) and Fig. 4.15(c).

We also observe resonance behavior around the cutoff angles for p-polarized but not the s-polarized incidence. Such resonance is not a numerical artifact since it remains as the numerical accuracy is varied (N in RCWA). It is indeed the Wood's anomaly which has been studied extensively since 1902 [88, 89, 90, 91, 92, 93]. Hessel and Oliner [92] classified the anomaly into two types: a Rayleigh type where a diffraction order appears or disappears, and a resonance type which is related to the coupling between the diffracted wave and the leaky wave inside the grating. The two types may occur very closely or separately. The second type is dependent on the incident field polarization, which requires the surface reactance of the grating to be capacitive or inductive for resonance to occur. The surface reactance is affected by grating

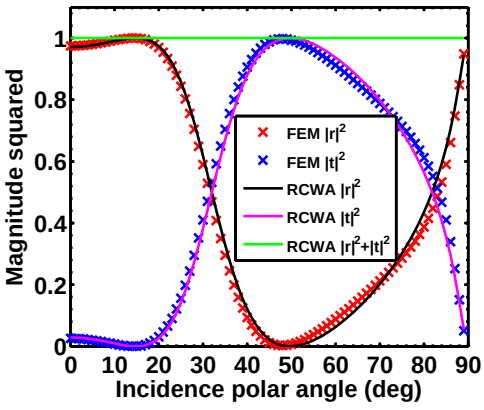
thickness, which explains why for a range of thicknesses, the anomaly only happens to one polarization. Previous investigation on Wood's anomaly in 1D gratings is well applicable to 2D gratings, though rigorous theoretical analysis could be rather involved. In this case, our 2D RCWA program is a convenient tool to study this effect. One thing to note is that, many studies [92, 93] on 1D grating defined s- and p-polarizations relative to the grooves of the grating, while in this work on 2D grating, polarizations are defined relative to the incidence plane. Thus the analysis on "s-anomalies" in [92] is applicable to the p-polarized case in Fig. 4.15(c).

Regardless of the anomalies, the energy conservation is still satisfied, as confirmed by our RCWA results. The total scattered power flux in the $\hat{z}$-direction remains a smooth function and has a cosine dependence on the incident polar angle when normalized by the total incident power flux, as shown by the magenta lines in Fig. 4.15(b) and Fig. 4.15(c).

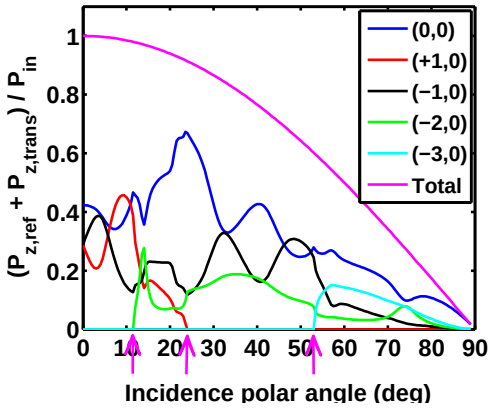
Even using only the fundamental order of reflection or transmission with normal plane wave incidence, the 2D grating can be designed as various optical components, such as polarizers and waveplates. Figure 4.16(a) shows the magnitude of the reflected, transmitted, and total scattered wave under normal plane wave incidence in the subwavelength regime ($\lambda > \max\{\Lambda_x, \Lambda_y\}$). The scattered wave is decomposed into two polarizations, one being parallel and the other being perpendicular to the incidence. As the incident polarization rotates relative to the grating about the $\hat{z}$-axis, the magnitude of the parallel component varies relative to the perpendicular component. This provides the possibility of using 2D gratings as polarizers. Moreover, the phase difference between the two components is also dependent on the azimuthal angle of the incident electric field, as shown in Fig. 4.16(b). This enables us to design quarter-wave plates. In this example, the parallel and perpendicular reflected waves have a 90° phase difference when the incident polarization angle is 27.7°. If such angle is 52.7°, the two transmitted components have a $-90°$ phase difference.

4.5 Design Rules for 2D High-Contrast Gratings

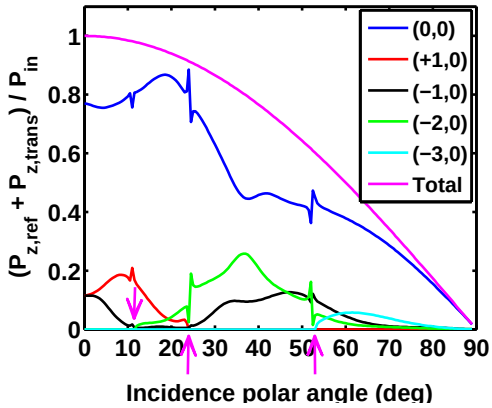
We have verified our 2D RCWA program as a design tool and studied the optical properties of 2D HCGs, yet the design of 2D HCGs remains a chal-



(a)

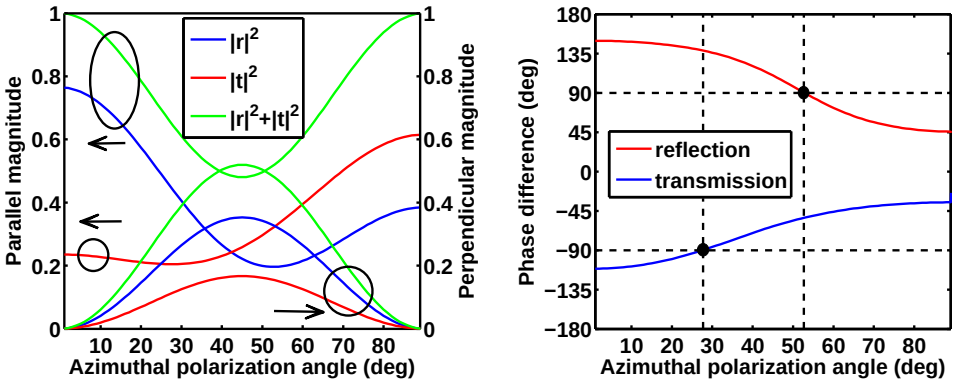


(b)



(c)

Figure 4.15: (a) Incident polar angle dependent reflectivity (black), transmissivity (magenta), and their sum (green) for the (0,0) fundamental order under p-polarized incidence with $\lambda = 2\,\mu$m calculated using rigorous coupled-wave analysis, and comparison with finite-element method (red and blue crosses). The same grating structure is solved under (b) s-polarized and (c) p-polarized incidence with wavelength being $\lambda = 0.6\,\mu$m. Normalized scattered power fluxes in the $\hat{z}$-direction for the (0,0)-th, (+1,0)-th, (−1,0)-th, (−2,0)-th, and (−3,0)-th spectral orders are shown as the blue, red, black, green, and cyan lines, respectively. The magenta line indicates the total normalized scattered power flux in the $\hat{z}$-direction. Arrows indicate the cutoff angles for spectral orders to appear or disappear.



(a) (b)

Figure 4.16: (a) Reflectivity (blue), transmissivity (red), and the normalized scattered power (green) of a normal incident plane wave with $\lambda = 1.55\,\mu$m as functions of the azimuthal polarization angle. The arrows indicate the components parallel and perpendicular to the incident wave. (b) Phase differences between the parallel and perpendicular components for reflection (red) and transmission (blue) as functions of the azimuthal polarization angle. The $\pm 90°$ phase differences correspond to polarization angles of 27.7° and 52.7° for reflection and transmission types, respectively.

lenge. The structural parameters include $n_{\rm r}$, t_g, Λ_x, Λ_y, η_x, and η_y. The incidence condition includes wavelength (λ), propagation direction (ϕ, θ), and polarization (α). Here we present a design procedure which largely simplifies the searching of initial parameters and the optimization process. As shown in Fig. 4.17, we start from the eigenmodes in the 2D HCG, which is considered as an infinitely long waveguide along the $\hat{z}$-direction and periodic in the xy-plane. Based on the dispersion relations of the eigenmodes, we can obtain initial designs for the grating structure in the xy-plane. Then using the dual-mode analysis we can find the resonance condition, and thus the values of grating thickness to provide high and possibly perfect reflection or transmission. With the initial structural parameters, we can then optimize the design for various applications. For reflectors, filters and resonators, we are concerned with locations of the passband, stopband, and transition band. By checking the frequency spectra and adjusting the grating thickness, we can shift these bands to our desired frequency range. For polarizers and waveplates, we want to obtain desired amplitude selectivity and relatively phase shift between two orthogonal polarizations. Both can be achieved by rotating the incident polarization relative to the 2D HCG, as shown in Fig.

4.16. For 2D phase plate, we need to obtain a full 2π phase tuning range by varying the HCG transverse structure (e.g. Λ or η). At the same time, we need to maintain high reflection or transmission for designs to be practical. More details about the phase plate are in the next section.

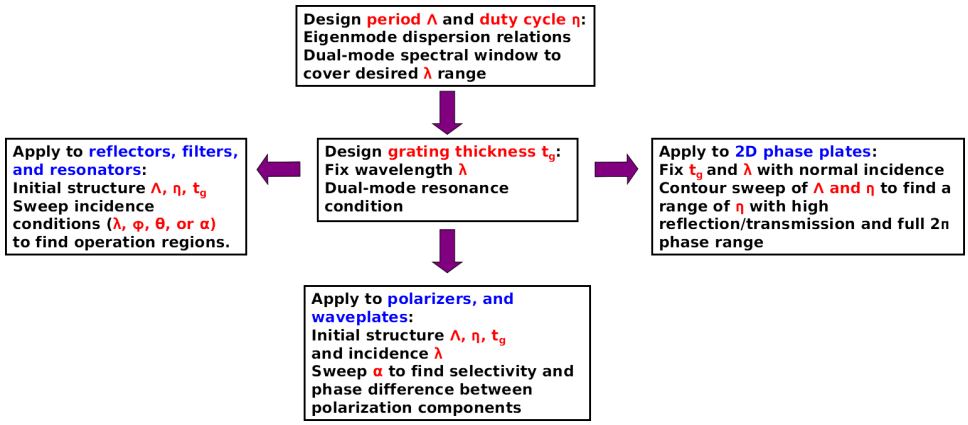


Figure 4.17: Design procedures of 2D high-contrast gratings.

4.5.1 Eigenmodes in 2D High-Contrast Gratings

Eigenmodes in 2D dielectric gratings are hybrid modes, which cannot be separated into transverse electric (TE) and transverse magnetic (TM), as in the case of 1D gratings. Furthermore, eigenmodes are more closely spaced in the frequency domain due to one extra degree of freedom (one more mode number), which largely complicates our analysis and prediction. However, the majority of the modes with strong confinement have dominant transverse field components, namely E_x/H_y-dominant (EH-like) and E_y/H_x-dominant (HE-like). These modes in periodic gratings resemble the EH and HE modes in the rectangle dielectric waveguides, using Marcatili's approximation [94]. We can analyze the two separately under the approximation that grating modes with one dominant polarization only couple strongly with the free-space Floquet modes with the same polarization. Thus we can reduce the number of modes to analyze by almost half. Moreover, the field distribution of the incident plane wave possesses even symmetry in the xy-plane across the unit cell except for particular incident angles. Therefore eigenmodes with odd symmetry will not be excited for most incidence conditions and we can further narrow down our analysis. Figure 4.18(a) shows the dispersion

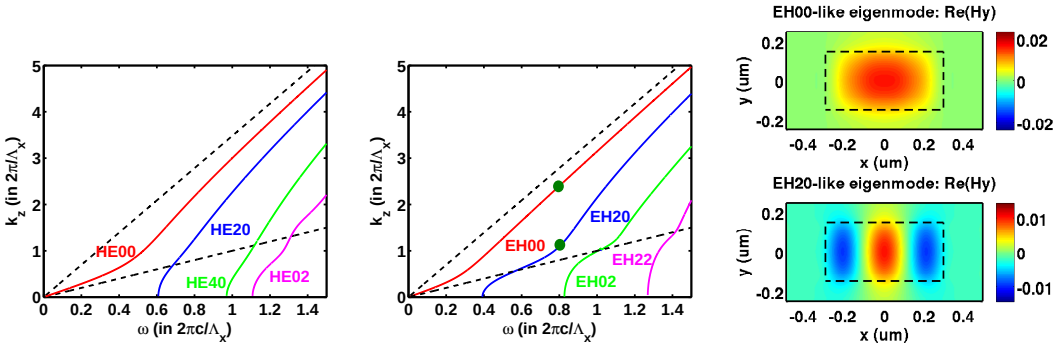
curves of the eigenmodes with even symmetry and H_x-component dominant over H_y-component. In fact, the four modes shown in Fig. 4.18(a) are very much like the HE modes in rectangular dielectric waveguides when they are well-confined between the light lines ($k_z = \omega/c$ and $k_z = n_g\omega/c$). Similarly, we can separate out E_x/H_y-dominant eigenmodes with even symmetry, as shown in Fig. 4.18(b). The H_y field profiles for EH_{00}-like and EH_{20}-like modes at $\omega = 0.8\frac{2\pi c}{\Lambda_x}$ indeed are very close to those in dielectric waveguides, as shown in Fig. 4.18(c).

Now we are able to focus on the modes which have major contribution to the HCG behavior. For instance, between the cutoff frequencies ω_{c20} and ω_{c02} (for the EH_{20}-like and EH_{02}-like modes, respectively), we have a spectral region where only two key modes exist. This means within this region we can engineer the interference between the two modes to produce desired functionalities.

In other cases, 2D gratings on the hexagonal lattice are preferred. Our RCWA program can be easily modified for this case. As shown in Fig. 4.19(a), the unit cell can be chosen as the yellow rectangle with a fixed aspect ratio of $\sqrt{3} : 1$. We can still separate eigenmodes with H_x-dominant and H_y-dominant polarizations. The dispersion curves of the first few H_x- and H_y-dominant eigenmodes are shown in Fig. 4.19(b) and Fig. 4.19(c), respectively. We further identify the eigenmodes which are symmetric with the translation along $\frac{1}{2}\hat{x} + \frac{\sqrt{3}}{2}\hat{y}$ from those which are anti-symmetric, as indicated by the red and blue lines in the dispersion curves. The H_y components of the first symmetric, first anti-symmetric, and second symmetric H_y-dominant modes are shown in Fig. 4.19(d), Fig. 4.19(e), and Fig. 4.19(f), respectively.

4.5.2 Dual-Mode Analysis for Perfect Reflection and Transmission

In many cases, the incident wave only strongly couples to a few eigenmodes due to mismatch of the field polarization or symmetry. We are particularly interested in the case when only two eigenmodes are strongly excited. Then we can engineer the structure to produce nearly perfect interference between them. In such case we can simplify our study to a dual-mode analysis. The



(a) (b) (c)

Figure 4.18: (a) HE-like even eigenmodes in a 2D rectangular grating on a rectangular lattice. The two dashed lines indicate the $k_z = \omega/c$ and $k_z = n_g\omega/c$ light lines. (b) EH-like even eigenmodes in a 2D rectangular grating on a rectangular lattice. (c) $\Re e[H_y]$ for the EH_{00}-like eigenmode and EH_{20}-like eigenmode at $\omega = 0.8\frac{2\pi c}{\Lambda_x}$, as indicated by the green dots in (b).

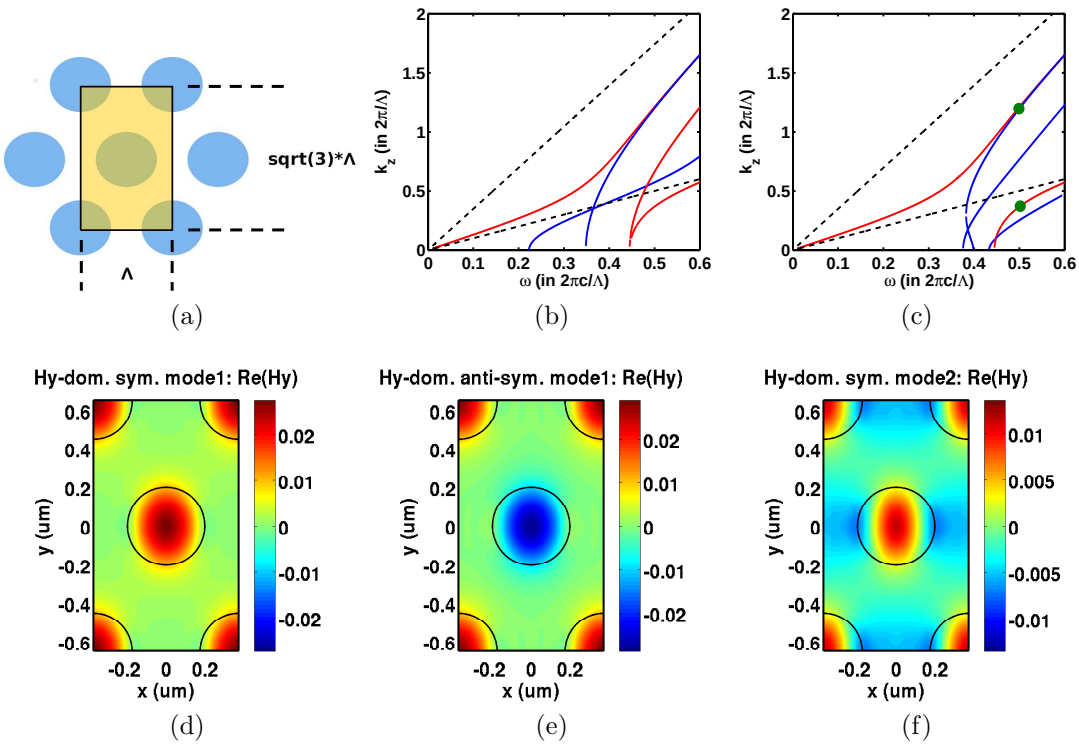
transverse field in Eq. (4.38) is rewritten as

$$\widetilde{\mathbf{E}}_t \approx \bar{\boldsymbol{\mathcal{E}}}_t^{\mathrm{DM}} \cdot \mathbf{A_{DM}} \tag{4.47}$$

where the "DM" indicates dual-mode. Here, $\bar{\boldsymbol{\mathcal{E}}}_t^{\mathrm{DM}}$ is a $2N \times 2$ matrix containing only two eigen-vectors. And $\mathbf{A_{DM}} = [A_p, A_q]^T$ contains two elements which are the expansion coefficients of the two modes. The superposition of eigenmodes in the HCG layer (Region II) forms a supermode. At certain wavelength λ and HCG thickness t_g, such a supermode can satisfy the Fabry-Pérot resonance condition in the $\hat{z}$-direction. Assume the field is scaled by a complex number $\Omega = |\Omega|e^{i\varphi}$ after a round-trip, we have

$$\begin{aligned}|\Omega|e^{i\varphi}\mathbf{A}_{\mathrm{II}}^{+}(0) &= \bar{\mathbf{R}}^{\mathrm{II-I}}\mathbf{A}_{\mathrm{II}}^{-}(0) = \bar{\mathbf{R}}^{\mathrm{II-I}}e^{-i\bar{\mathbf{K}}_z\cdot(-t_g)}\mathbf{A}_{\mathrm{II}}^{-}(t_g) = \bar{\mathbf{R}}^{\mathrm{II-I}}e^{i\bar{\mathbf{K}}_z t_g}\bar{\mathbf{R}}^{\mathrm{II-III}}\mathbf{A}_{\mathrm{II}}^{+}(t_g)\\ &= \bar{\mathbf{R}}^{\mathrm{II-I}}e^{i\bar{\mathbf{K}}_z t_g}\bar{\mathbf{R}}^{\mathrm{II-III}}e^{i\bar{\mathbf{K}}_z t_g}\mathbf{A}_{\mathrm{II}}^{+}(0) = \bar{\mathbf{M}}(\lambda, t_g)\mathbf{A}_{\mathrm{II}}^{+}(0)\end{aligned} \tag{4.48}$$

where $\bar{\mathbf{R}}^{\mathrm{II-I}}$ and $\bar{\mathbf{R}}^{\mathrm{II-III}}$ are the reflection matrices for the Region II eigenmodes bounced back by Region I (incident region) and Region III (transmitted region), respectively. Both reflection matrices can be obtained from Eq. (4.42). Equation (4.48) is an eigen-equation for the round-trip propagation matrix $\bar{\mathbf{M}}(\lambda, t_g)$. When the phase φ of the eigenvalue Ω is even or



(a) (b) (c)

(d) (e) (f)

Figure 4.19: (a) Circular high-contrast grating on a hexagonal lattice. The yellow rectangle indicates a choice of the unit cell. Parameters: $\Lambda = 1\,\mu$m, $\eta = 0.6$. Dispersion curves of (b) H_x-dominant and (c) H_y-dominant eigenmodes possessing symmetry (red) and anti-symmetry (blue). $\Re e[H_y]$ for (d) the first symmetric, (e) the first anti-symmetric, and (f) the second symmetric modes at frequency $\omega = 0.5\frac{2\pi c}{\Lambda}$, as indicated by the green dots in (c).

odd multiples of π, we have perfect constructive or destructive interference, respectively.

Using our dual-mode analysis, the dimensions of matrices and vectors in Eq. (4.48) are reduced to contain only two modes,

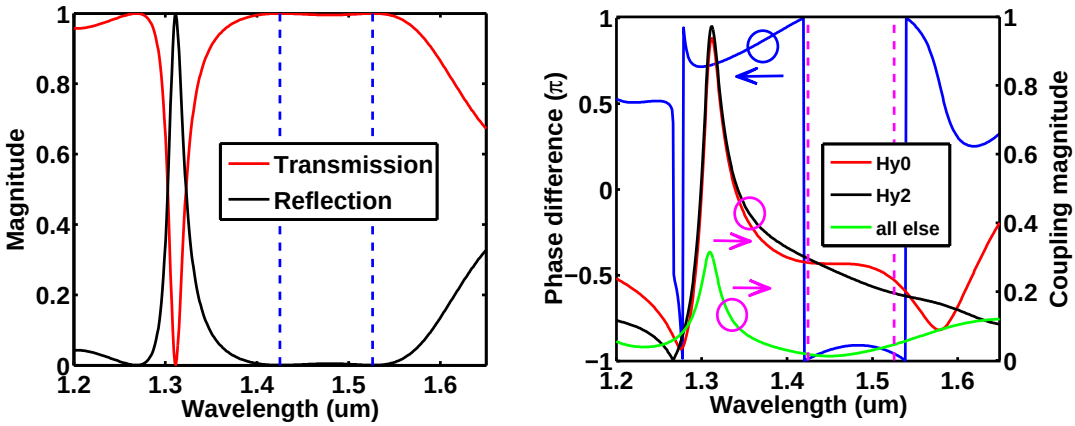
$$\left[\bar{\mathbf{M}}_{\mathrm{DM}}\right]_{2\times 2}(\lambda, t_g)\begin{bmatrix} A_p \\ A_q \end{bmatrix} = |\Omega_{\mathrm{DM}}|e^{i\varphi_{\mathrm{DM}}}\begin{bmatrix} A_p \\ A_q \end{bmatrix} \tag{4.49}$$

where the phase φ_{DM} can be adjusted by the wavelength and the grating structure.

In Fig. 4.20(a) we see the reflection and transmission spectra of a 2D circular HCG on a hexagonal lattice under E_x/H_y-polarized normal incidence at $\lambda = 1.55\,\mu$m. We have perfect transmission (zero reflection) at two wavelengths and ultra-high transmission in between. Using our dual-mode analysis, we can easily extract the contribution to the reflected (0,0) mode from the back-coupling of each of the two strongly excited eigenmodes, as well as the contribution from all other modes, as shown in Fig. 4.20(b). Furthermore, we can calculate the phase difference of the contribution from the two modes. We can see that at zero reflection, indicated by the dashed lines, the coupling magnitudes of the two modes are almost the same, and the phase difference is almost $\pm\pi$. This indicates the interference of the two modes is very close to but not yet perfectly destructive. We still see some coupling from other eigenmodes. This is because the eigenmodes in 2D HCG are hybrid modes and we have no perfect selectivity of polarizations. The result is a slight difference between the wavelengths for zero-reflection and the dual-mode perfect interference. Nonetheless, this analysis provides a very good estimation of the zero-reflection condition. The zero-transmission case can be analyzed similarly when we look at the forward-coupling to the transmitted region.

4.5.3 Resonance Conditions in 2D High-Contrast Gratings

For a given desired wavelength, we can also adjust the HCG thickness t_g to study the resonance condition. We first look at a rectangular HCG with the same structure as in Fig. 4.14 under E_x/H_y-polarized normal incidence. The dual-mode eigen-equation in Eq. (4.49) yields two eigen-solutions which we

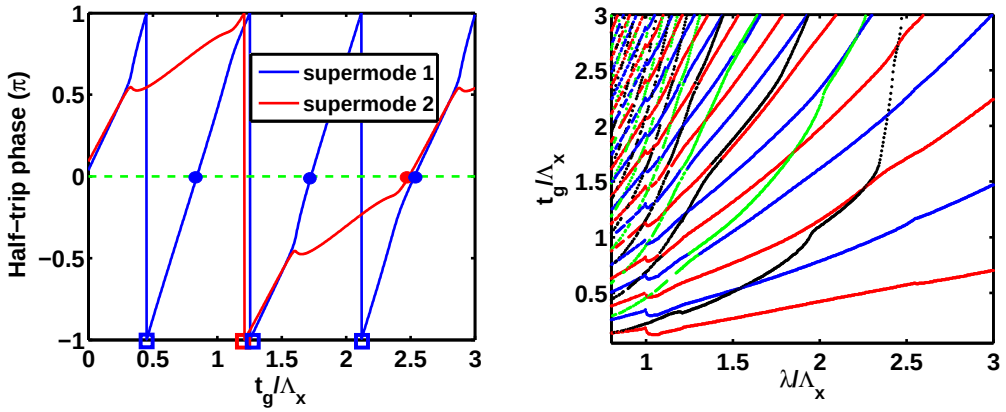


(a) (b)

Figure 4.20: (a) Reflectivity (red) and transmissivity (blue) of a 2D circular high-contrast grating on a hexagonal lattice with $\Lambda = 750\,\text{nm}$, $\lambda = 1.55\,\mu\text{m}$, and $t_g = 713\,\text{nm}$. Blue dashed lines indicate the wavelengths for perfect transmission. (b) Back-coupling magnitude from the first (red), second (black) symmetric eigenmodes and all other eigenmodes (green) in the grating to the (0,0) mode in the incident region. The dashed lines in (b) are at the same wavelengths as in (a). Blue solid line is the phase difference between the reflected waves back-coupled from the two eigenmodes.

will name as supermode 1 and supermode 2. The phases of the eigenvalues ($\varphi_{1,\text{DM}}$ and $\varphi_{2,\text{DM}}$), which are the round-trip phases for the two supermodes, are functions of the grating thickness. Figure 4.21(a) shows the half-trip phases ($\varphi^{\text{half}}_{1,\text{DM}}$ and $\varphi^{\text{half}}_{2,\text{DM}}$), and we record the HCG thicknesses at which the half-trip phases are even and odd multiples of π, indicated by the empty boxes and solid circles, respectively. At each given wavelength, we can solve this dual-mode eigen-problem and obtain the thicknesses for resonance to occur. Figure 4.21(b) shows the contour plot for the resonance conditions. Red and black lines indicate the half-trip phases $\varphi^{\text{half}}_{1,\text{DM}}$ and $\varphi^{\text{half}}_{2,\text{DM}}$ being odd multiples of π, respectively. Blue and green lines indicate the half-trip phases $\varphi^{\text{half}}_{1,\text{DM}}$ and $\varphi^{\text{half}}_{2,\text{DM}}$ being even multiples of π, respectively.

Figure 4.22(a) shows the reflectivity contour plot as a function of the grating thickness and wavelength. From Fig. 4.18(b) we can find the cutoff frequencies for the EH_{20} and EH_{02} to be $\omega_c^{\text{EH}_{20}} = 0.392\frac{2\pi c}{\Lambda_x}$ ($\lambda_c^{\text{EH}_{20}} = 2.55\Lambda_x$) and $\omega_c^{\text{EH}_{02}} = 0.827\frac{2\pi c}{\Lambda_x}$ ($\lambda_c^{\text{EH}_{02}} = 1.21\Lambda_x$). Between these two frequencies we see a dual-mode spectral window from the dispersion curves. These two frequencies are also indicated by the black lines in Fig. 4.22. We can see the

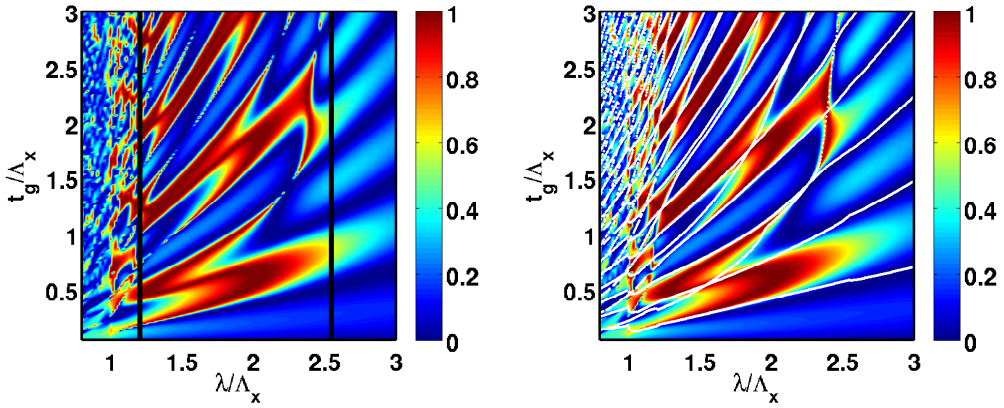


(a) (b)

Figure 4.21: (a) Half-trip phases of the supermodes 1 (blue) and 2 (red) in the same grating as in Fig. 4.14 under $\lambda = 2.1\,\mu$m normal incidence. Empty boxes and solid circles indicate half-trip phases being odd and even multiples of π, respectively. (b) Resonance conditions for the grating thicknesses at given wavelengths. Red and blue lines indicate the half-trip phase of supermode 1 being odd and even multiples of π, respectively. Black and green lines indicate the half-trip phase of supermode 1 being odd and even multiples of π, respectively.

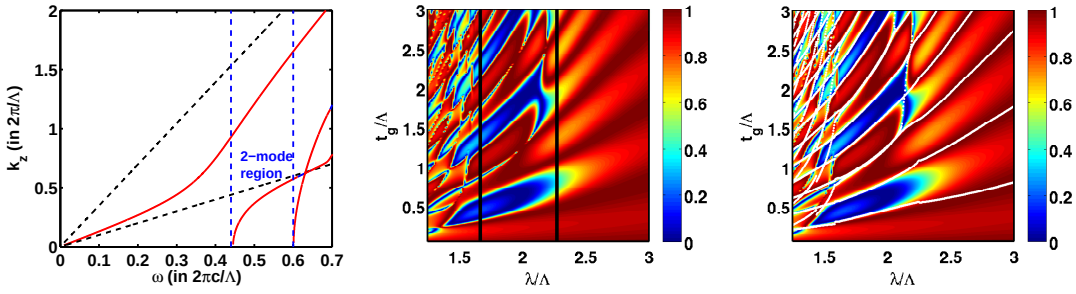
ultra-high-reflection regions are mostly within the dual-mode window. Figure 4.22(b) shows the overlap between the resonance lines in Fig. 4.21(b) and the reflectivity contour in Fig. 4.22(a). Excellent agreement is shown which indicates that our dual-mode analysis can successfully predict the resonance and high-reflection conditions in a 2D HCG.

For a circular HCG on a hexagonal lattice, we can also identify a dual-mode spectral window where only two symmetric modes with the same dominant polarization are strongly excited, as shown in Fig. 4.23(a). From the transmissivity contour plot of our designed transmission-type HCG in Fig. 4.23(b), we see wide high-transmission regions inside the dual-mode window. These high-transmission regions are mostly bounded by or along the resonance lines, as shown in Fig. 4.23(c). We again see an excellent overlap between the dual-mode resonance lines and the contour plot. From the resonance lines, we can easily identify the design regions for broadband or sharp-transition applications, such as reflectors, filters, and resonators.



(a) (b)

Figure 4.22: (a) Reflectivity contour plot of a rectangular high-contrast grating (same as in Fig. 4.14) as a function of the thickness t_g and wavelength λ. The solid lines indicate $\lambda = 2.55\Lambda_x$ and $\lambda = 1.21\Lambda_x$, corresponding to the cutoff frequencies $\omega_c^{\mathrm{EH}_{20}} = 0.392\frac{2\pi c}{\Lambda_x}$ and $\omega_c^{\mathrm{EH}_{02}} = 0.827\frac{2\pi c}{\Lambda_x}$ for the EH_{20} and EH_{02}-like modes, as shown in Fig. 4.18(b). (b) Overlap between the resonance contour plot (white) and the reflectivity contour plot.

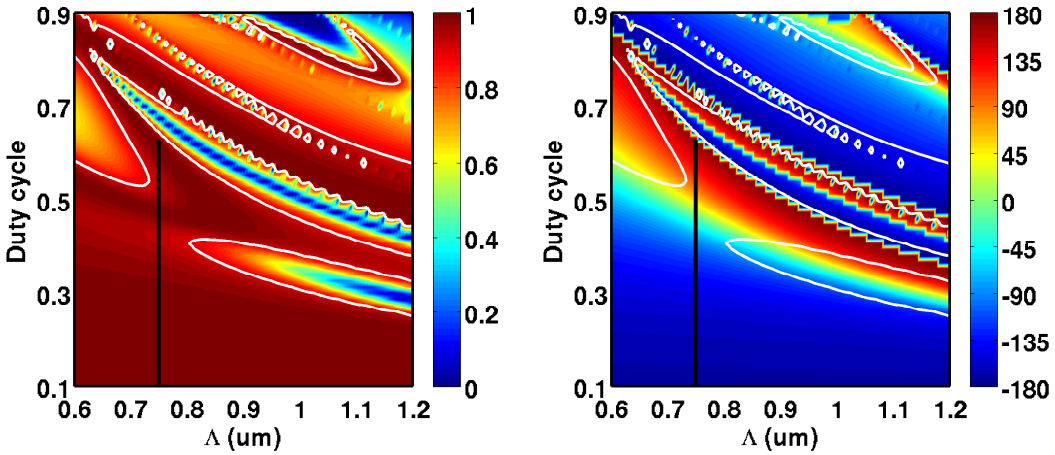


(a) (b) (c)

Figure 4.23: (a) Dual-mode dispersion curves in a hexagonal-lattice grating and the dual-mode window (same as in Fig. 4.19) is indicated by the blue dashed lines at $\omega = 0.44\frac{2\pi c}{\Lambda}$ and $\omega = 0.6\frac{2\pi c}{\Lambda}$. (b) Transmission contour plot as a function of the grating thickness t_g and wavelength λ. The dual-mode window indicated by the two black lines. (c) Overlap between the resonance lines (white) and the transmission contour plot.

4.6 Engineering of 2D Phased Arrays Using High-Contrast Gratings

With the knowledge of the resonance conditions and the mechanisms for ultra-high reflection or transmission, we can also apply the initial design to 2D phase plates using HCGs. Now we fix the wavelength and grating thickness, but vary the transverse structure to find a full 2π reflection (or transmission) phase tuning range while maintaining the power efficiency. Figure 4.24 shows an example of a transmission-type hexagonal-lattice HCG operating at $\lambda = 1.55\,\mu$m with a thickness of $0.8\,\mu$m. We are able to find an optimum design range by fixing $\Lambda = 0.75\,\mu$m, and tuning η from 10% to 63%, as indicated by the black lines. Thus we have a mapping between the transmission phase and the transverse geometry. The design of the phase plate is essentially aimed at generating a transverse position-dependent phase alternation profile, that is, $\Delta\Phi(x, y)$. This can be translated into the design of the $\eta(x, y)$ profile. In many applications, we would need substrates for the



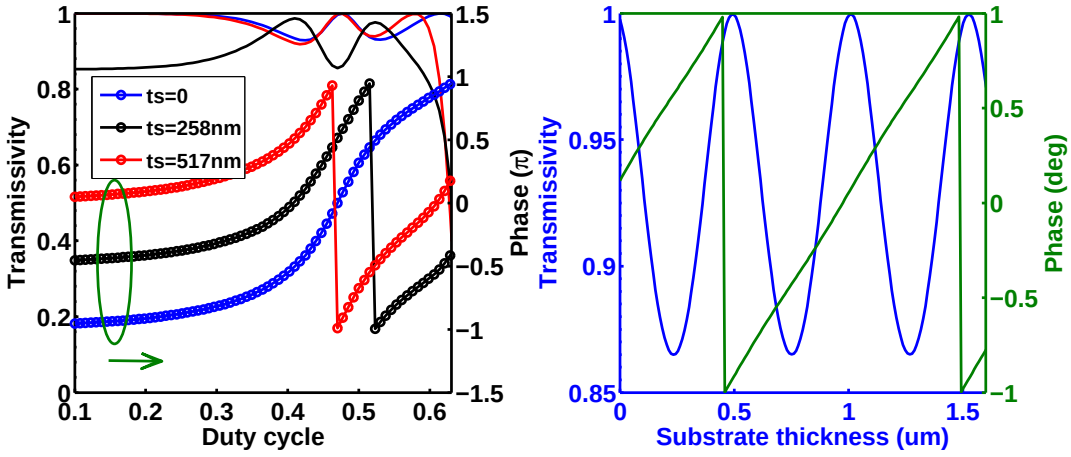
(a) (b)

Figure 4.24: Contour plots as functions of the grating period Λ and duty cycle η for the (a) magnitude and (b) phase of the transmission through a hexagonal-lattice grating under $\lambda = 1.55\,\mu$m normal incidence. The black lines indicates a 2π phase range at $\Lambda = 750\,$nm. The white lines indicate the contour for 90% transmission.

HCGs. Figure 4.25(a) shows the magnitude and phase tuning by the HCG duty cycle with different substrate thicknesses ($t_g = 0$, $\lambda/4$, and $\lambda/2$). We can see the full 2π phase tuning range is still obtainable except the phase

discontinuity occurs at different locations. The magnitude remains almost unchanged for thickness being multiples of half-wavelength, but is strongly perturbed at quarter-wavelength. In the situation when the grating transmission is not close enough to unity, we can use the substrate to further improve the transmission, and find the optimum value of t_g, as shown in Fig. 4.25(b).

By introducing a phase alternation $\Delta\Phi$ based on the transverse spatial location (x, y), we can have 2D HCGs behaving like many conventional optical components. For example, having $\Delta\Phi(x, y) = \frac{2\pi}{\lambda}(x\cos\phi\sin\theta + y\sin\phi\sin\theta)$, we can steer the normal-incident plane wave into an oblique direction of $\hat{x}\cos\phi\sin\theta + \hat{y}\sin\phi\sin\theta + \hat{z}\cos\theta$, where θ and ϕ are the polar and azimuthal angles of the transmitted wave vector. For a lens with a focusing length f, we need $\Delta\Phi(x, y) = \frac{2\pi}{\lambda}\left(f - \sqrt{f^2 + x^2 + y^2}\right)$. For an axicon [73] with an opening angle α_{axi}, an index n_{r} and a radius R, we have $\Delta\Phi(x, y) = (R - \sqrt{x^2 + y^2})(n_{\text{r}} - 1)\tan\alpha_{\text{axi}}$. For a focusing lens that generates orbital angular momentum (OAM) of $m\hbar$ ($m = 0, \pm1, \pm2...$), we have $\Delta\Phi(x, y) = \frac{2\pi}{\lambda}\left(f - \sqrt{f^2 + x^2 + y^2} + m\phi\right)$. Using Fig. 4.25 we can design the local duty cycle $\eta(x, y)$ for the 2D HCG based on the desired phase alternation $\Delta\Phi(x, y)$.

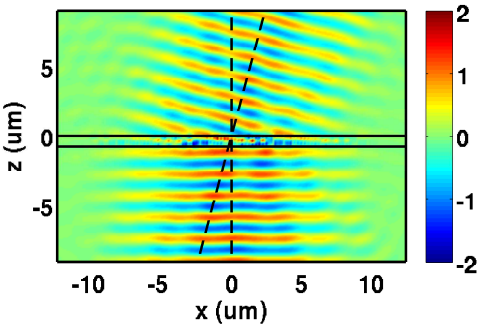


(a) (b)

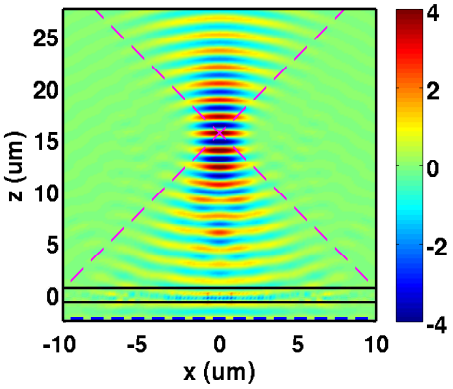
Figure 4.25: (a) Tuning of the transmission magnitude (solid) and phase (dots) by grating duty cycle with thickness being 0 (blue), 258 nm (black), and 517 nm (red). Same structure as in Fig. 4.19 and $\lambda = 1.55\,\mu$m. (b) Transmission magnitude (blue) and phase (green) as functions of the substrate thickness with 48% duty cycle.

Once we have the design for 2D HCGs, we use the 3D FDTD method to verify the performance. Figure 4.26(a) shows a normal-incident x-polarized plane wave toward the $+\hat{z}$-direction at $\lambda = 1.55\,\mu$m passing through our designed HCG beam steering plate with a 15° steering angle. The two dashed lines indicate the normal and 15° from normal directions. The transmitted wave is indeed traveling along the desired direction with high power efficiency. Figure 4.26(b) shows a normal-incident x-polarized Gaussian beam with a beam waist $w_0 = 7\,\mu$m transmitting through a 10-μm-radius HCG phase plate, which is designed to have a focal length $f = 15\,\mu$m. The magenta dashed lines indicate the focusing behavior of an equivalent lens with a numerical aperture of 0.5547. Figure 4.26(c) shows that the same Gaussian beam passes through a HCG phase plate, which acts as an axicon [74], and becomes a Bessel beam . The magenta dashed lines indicate an Bessel beam opening angle of $2\theta_{\mathrm{BS}} = 0.4\,\mathrm{rad}$, which is the result of an equivalent axicon with a radius $R = 10\,\mu$m, an index $n_{\mathrm{r}} = 1.5$, and an axicon opening angle $\alpha_{\mathrm{axi}} = \theta_{\mathrm{BS}}/(n_{\mathrm{r}} - 1) = 0.4\,\mathrm{rad}$. Such a Bessel beam has a maximum propagation distance $z_{\max} \approx w_0/\theta_{\mathrm{BS}} = 35\,\mu$m, within which the beam will have no diffraction. This maximum distance can be increased by reducing the axicon opening angle α_{axi}.

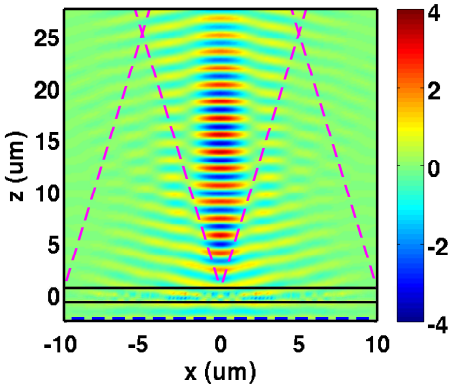
We can allow our 2D HCG phase plate to focus and simultaneously generate orbital angular momentum as a beam passes through. Such modification can be easily done by adding a spatial-dependent, or more specifically, an azimuthal angle-dependent phase delay. Using the phase information from Fig. 4.25(a) we arrive at a 2D HCG design in Fig. 4.27(a) on top of a glass substrate, with an operation wavelength at $\lambda = 1.55\,\mu$m. The transmitted field intensity profiles at $z = f - 4\lambda$, $z = f$, and $z = f + 4\lambda$ are shown in Fig. 4.27(b), 4.27(c), and 4.27(d), respectively, where the focal length is designed to be $20\,\mu$m. We can see a clear intensity null at the center, which is necessary for nonzero OAM. The incident beam has a Gaussian distribution with a beam waist of $7\,\mu$m and zero OAM, as shown in Fig. 4.27(e). The focusing behavior is observed by comparing the beam waists and the peak intensities 4λ below, 4λ above, and exactly at the focal plane, as well as those of the source. Our design provides high transmissivity with a power efficiency above 90%. Figure 4.27(f), 4.27(g), and 4.27(h) show the phase distribution profiles at $z = f - \lambda/3$, $z = f$, and $z = f + \lambda/3$. The spatial-dependent phase distribution already indicates nonzero OAM. The phase of the transmitted



(a)



(b)



(c)

Figure 4.26: $\Re e[E_x]$ in the xz-plane for an x-polarized normal-incident wave toward $+\hat{z}$ at $\lambda = 1.55\,\mu$m transmitting through HCG phase plates designed for Gaussian beams to: (a) be deflected by 15°; (b) focus at a 15-μm distance; (c) be converted to Bessel beams. Black solid lines indicate the grating layer. Blue dashed lines indicate the source locations. Other dashed lines are for visual aids. All field intensities are normalized by the incident field intensity.

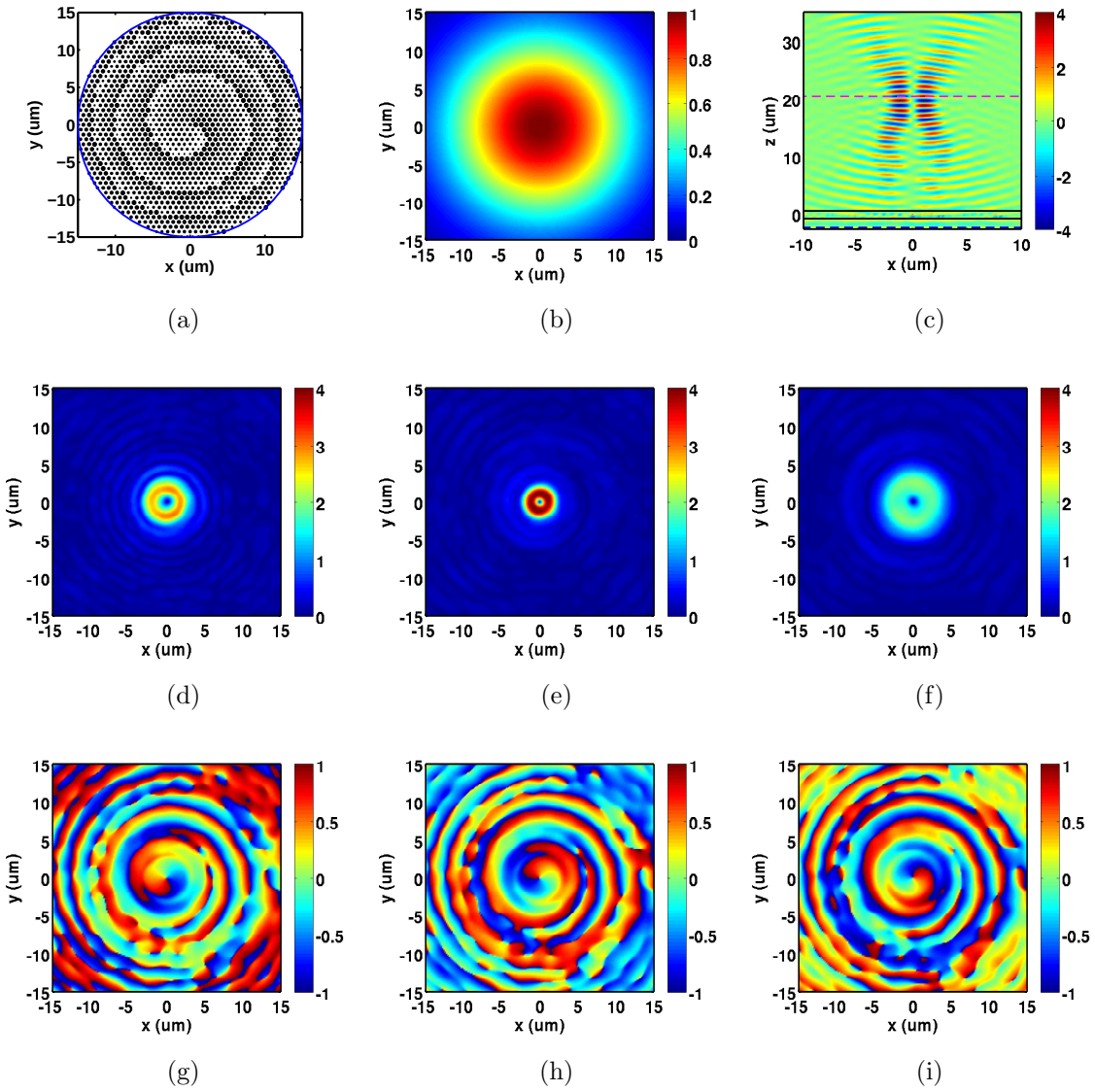


Figure 4.27: (a) Structure of the 2D high-contrast grating phase plate for focusing at $f = 20\,\mu$m and generating $+1\hbar$ orbital angular momentum. (b) Gaussian source intensity with a beam waist of $12\,\mu$m. (c) Cross-section view of $\Re e[E_x]$ in the xz-plane, where the red dashed line indicates the focal plane, the blue dashed line indicates the source, and the black lines indicate the phase plate. Field intensities at $z = f - 4\lambda$, $z = f$, and $z = f + 4\lambda$ are shown in (d), (e), and (f), respectively. Phase distributions (in π) of the transmitted wave at $z = f - \lambda/3$, $z = f$, and $z = f + \lambda/3$ are shown in (g), (h), and (i), respectively.

wave increases from $-\pi$ to π as the azimuthal angle increases in a 2π range, indicating the OAM is $+1\hbar$ per photon. Alternatively, we see the phase profile rotates clockwise by $2\pi/3$ as the beam propagates by $\Delta z = \lambda/3$ in the $+\hat{z}$-direction, showing a left-hand helical phase front. Therefore the OAM is indeed $+1\hbar$ per photon.

4.7 Summary

In this chapter we study the optical properties of high-contrast gratings (HCGs), and the design of HCGs as novel high-performance, low cost, compact optical components. We start with modeling the HCGs with one-dimensional periodicity using the mode-matching method and the method of moments with periodic Green's functions. Our results show good agreement with the finite-element method. Our model is efficient and accurate for the design and optimization of the HCGs.

We have also investigated the physics of 2D HCGs. Our in-house developed 2D rigorous coupled-wave analysis (RCWA) program is shown to be an efficient and accurate tool for understanding the optical behavior of 2D HCGs with various structural and incidence parameters. We further demonstrate the design rules for various optical applications using 2D HCGs, based on the HCG mode properties we obtained from RCWA. Using our dual-mode analysis, the design process is largely simplified. Once our high-performance initial design is obtained, it is optimized according to the desired functionalities, such as broadband reflection, high-Q resonance, filtering, polarizing, etc. At last, we discuss the design of phase plates using 2D HCGs. Our designed HCG phase plates can function as conventional optical components, such as lenses, deflectors, axicons, spiral phase shifters, with excellent agreement.

There is much recent research progress on vortex beam generation using plasmonic metastructures. The localized resonance behavior resulting from the plasmonic effect also enables a wide range of beam engineering. However, materials supporting plasmonic resonance also introduce much loss. Compared to these designs, our designs using low-loss high-index dielectric materials demonstrate excellent performance in terms of power efficiency.

CHAPTER 5

HIGH-CONTRAST-GRATING MEMS-TUNABLE VCSELS

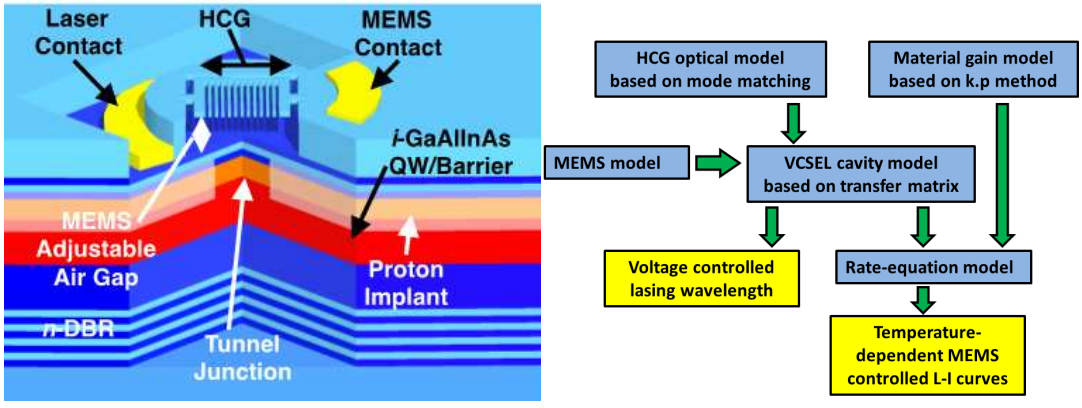
5.1 Introduction

Long wavelength vertical-cavity surface-emitting lasers (VCSELs) emitting at around 1550 nm have drawn extensive research interest in the past few decades due to their fast-growing applications in fiber communication, optical interconnects, and laser spectroscopy [9, 95, 96, 97]. VCSELs with tunable wavelengths are particularly important for wavelength-division-multiplexing and light ranging and detection [29, 98, 99]. The high contrast grating (HCG) controlled by a micro-electro-mechanical system (MEMS) has become a promising candidate for tunable VCSELs due to its high reflectivity, good mode selectivity, low power consumption, and low cost [32, 33, 34, 38, 39, 99].

The InP-based HCG tunable VCSELs have been demonstrated with single mode operation, a wide wavelength tuning range, and fast MEMS tuning speed [33]. However, a theoretical model has yet to be developed for analyzing the laser performance. In this chapter, we demonstrate a comprehensive model which covers the theories of quantum-well (QW) material gain, the reflectivity of the top and bottom mirrors composed of both the HCG and distributed Bragg reflectors (DBRs). The optical modeling of the resonant cavity is connected to the electrostatic modeling of the MEMS, and the results provide a deeper understanding of the tunable device properties and accurately predict the lasing wavelength and threshold current. This chapter also establishes a rate-equation model that correlates the injected carrier density and output photon density, accurately predicting the laser light output versus current (L-I) behavior. Thermal effects are especially important to consider for short cavity lasers such as VCSELs [100]. In this work we consider the temperature change of the active region as more current is injected. The degradation of the material gain as the active region temperature

increases is considered, as well as the red-shift of both the gain peak and the cavity resonance due to thermal effects. The current that leaks through the quantum wells without undergoing recombination is also included. The unpinning effect of the carrier density above threshold is investigated. The temperature-dependent L-I curves calculated by our model agree very well with the experiment. This comprehensive model can be an important tool for designing high-speed, low-chirp, low-threshold tunable lasers with wide tuning ranges.

Figure 5.1(a) shows the schematic diagram of the HCG tunable VCSEL to be modeled in this work [36, 37]. The active region consists of InGaAlAs multiple quantum wells. The top reflector consists of the HCG and DBR. The air gap between the HCG and top DBR is tunable by the MEMS control voltage. The device is electrically injected and the proton implantation serves to form the current confining aperture. Figure 5.1(b) shows the flow of the modeling procedure. The HCG optical properties are calculated by the mode matching method, and the air-gap thickness is correlated with the tuning voltage by the MEMS model. The VCSEL structure is modeled by the transfer matrix method, which can predict the MEMS-controlled lasing wavelength and provide the cavity parameters. The rate-equation model takes in the cavity parameters and the QW material gain calculated by the $\mathbf{k}\cdot\mathbf{p}$ method. Finally, the temperature-dependent MEMS-controlled L-I curves are calculated.



(a) (b)

Figure 5.1: (a) Schematic diagram of the high contrast grating tunable VCSEL. (b) Block diagram of the theoretical modeling procedure.

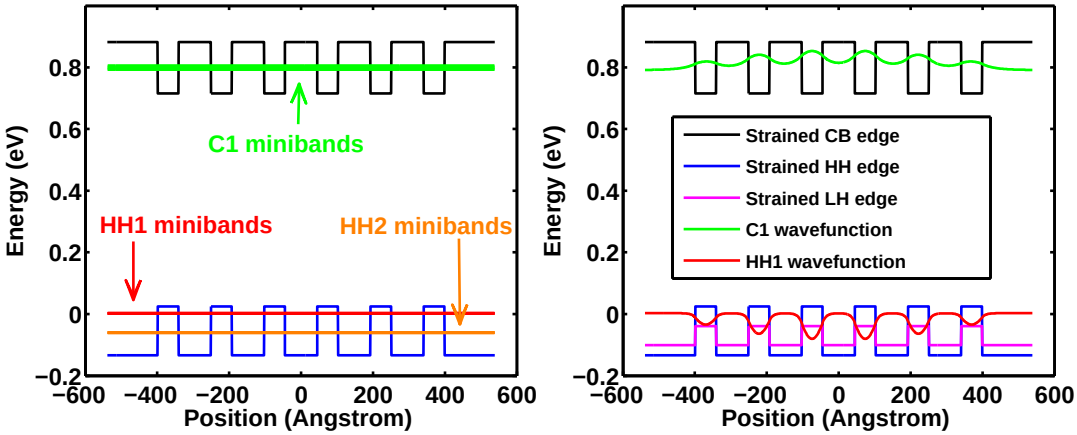
5.2 Material Gain of Strained Quantum Wells

The active region of the long-wavelength HCG VCSEL consists of InGaAlAs multiple quantum wells (MQWs). The modeling of the optoelectronic properties of the strained MQWs is based on our rigorous band structure model, and the detailed formulation is shown in Appendix B. The valence band structure and eigenstates are solved with the 4-band $\mathbf{k} \cdot \mathbf{p}$ method using the Luttinger-Kohn model [55, 101, 102, 103], which includes the heavy-hole (HH) and light-hole (LH) mixing. The conduction band (CB) structure and eigenstates are solved with the single-band effective-mass approximation. The strain effect is included through the Pikus-Bir strain terms in the Hamiltonian [57].

The QW region is a 57.5Å $InGa_{0.274}Al_{0.046}As$ layer with compressive strain on the InP substrate. The barriers between the QWs are 90Å $InGa_{0.366}Al_{0.174}As$ layers with tensile strain on the InP substrate. The binary material parameters as well as the bowing parameters for ternary and quaternary materials can be found in [58, 104]. Figure 5.2(a) shows the energy levels solved from the 4-band Hamiltonian for six coupled QWs. Due to the coupling among QWs, each state becomes a miniband with six coupled states. The energy splitting among the states depends on the wavefunction leakage from one QW to others, and thus is affected by the effective mass and the barrier height. This is the reason why the CB miniband has a broader splitting. Figure 5.2(b) shows the envelop functions $\psi(z)$ and $g^{(1)}(z)$ for the lowest CB miniband state and the highest HH1 miniband state, respectively.

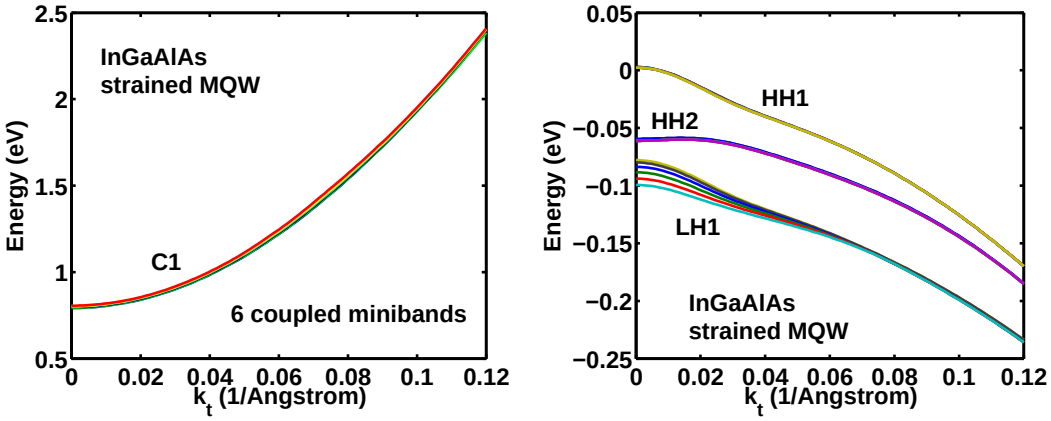
The coupled Hamiltonian is solved at each given transverse wave number k_t, and we can obtain the energy dispersion relations for each subband, as shown in Fig. 5.3(a) and 5.3(b). The CB dispersion is very close to parabolic due to the effective mass approximation and the fact that the barrier effective mass has little contribution for bound states. However, the VB dispersion is non-parabolic due to the band-mixing effect.

The radiative transition rate can be obtained from Fermi's golden rule, which accounts for the Fermi-Dirac occupation of the conduction and valence subbands. Therefore we can write the material gain and spontaneous



(a) (b)

Figure 5.2: (a) Energy levels for the minibands in the first conduction band (CB) (green), the first heavy-hole (HH) band (red), and the second HH band (orange). Black and blue solid lines are the strained CB and HH band edges, respectively. (b) Green and red lines are the wave functions for the lowest and highest states among the C1 and HH1 minibands, respectively. The magenta line is the light-hole band edges.



(a) (b)

Figure 5.3: Energy dispersion relations for (a) the first conduction band (C1), (b) the first heavy-hole band, the second heavy-hole band, and the first light-hole band. Each subband is split into six minibands due to the six coupled quantum wells.

emission rate for the quantum wells as [105]

$$
\begin{aligned}
g(\hbar\omega) =& \frac{\pi e^2}{n_{\mathrm{r}} c\epsilon_0 m_0^2 \omega} \sum_{\sigma_2} \sum_{n,m} \int_0^\infty \frac{k_t dk_t}{\pi L_z} |M_{nm}^{\sigma_2}(k_t)|^2 [f_c^n(k_t) - f_v^{\sigma_2,m}(k_t)] L(k_t, \hbar\omega) \\
r_{\mathrm{spon}}(\hbar\omega) =& \frac{n_{\mathrm{r}} \omega e^2}{\pi\hbar c^3 \epsilon_0 m_0^2} \sum_{\sigma_2} \sum_{n,m} \int_0^\infty \frac{k_t dk_t}{\pi L_z} |M_{nm}^{\sigma_2}(k_t)|^2 f_c^n(k_t) [1 - f_v^{\sigma_2,m}(k_t)] L(k_t, \hbar\omega)
\end{aligned}
\tag{5.1}
$$

where σ_2 accounts for the valence band spin degeneracy, where the conduction band spin degeneracy is already accounted for. The lineshape function $L(k_t, \hbar\omega)$ accounts for the finite transition linewidth due to various scattering mechanisms. Here, $M_{nm}^{\sigma_2}$ is the momentum matrix element, and f_c and f_v are the Fermi distribution functions for electrons in the n-th conduction subband and m-th valence subband, respectively.

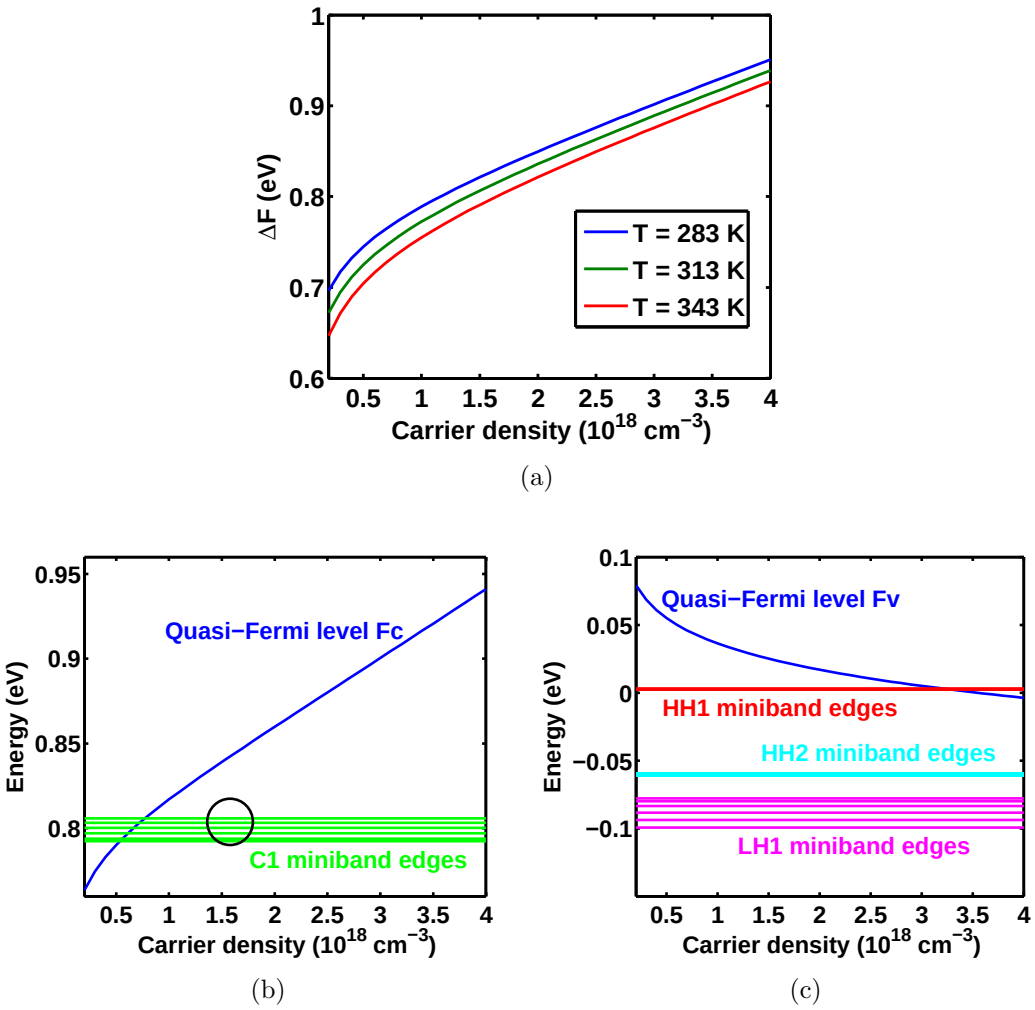
$$
\begin{aligned}
f_c^n(k_t) &= \frac{1}{1 + \exp(\frac{E_{c,n}(k_t) - F_c}{kT})} \\
f_v^m(k_t) &= \frac{1}{1 + \exp(\frac{E_{v,m}(k_t) - F_v}{kT})}
\end{aligned}
\tag{5.2}
$$

The quasi-Fermi levels (QFLs) F_c and F_v in Eq. (5.2) are related to the injected carrier density by

$$
\begin{aligned}
n &= \sum_{\sigma_1}^{\uparrow,\downarrow} \sum_n \frac{1}{2\pi L_T} \int_0^\infty k_t dk_t f_c^{\sigma_1,n}(k_t) \\
p &= \sum_{\sigma_2}^{U,L} \sum_m \frac{1}{2\pi L_T} \int_0^\infty k_t dk_t \left[1 - f_v^{\sigma_2,m}(k_t)\right]
\end{aligned}
\tag{5.3}
$$

Figure 5.4 shows how the quasi-Fermi levels F_c, F_v, and their separation ΔF depend on the injection carrier density. From Fig. 5.4(a) we can determine at different temperatures the values for injection carrier density such that the Fermi-level separation is equal to the effective band gap, i.e. the Bernard-Duraffourg inversion condition. The corresponding carrier density is known as the transparency carrier concentration n_{tr} [48, 55]. Also, the difference between ΔF and the QW barrier height determines the rate for carriers to tunnel into or leak out of the QW region [37, 63]. Therefore this should be considered in our design, which will be discussed in later sections.

Figures 5.4(b) and 5.4(c) show the F_c and F_v calculated from Eq. (5.3) and the miniband energies at the zone center ($k_t = 0$) are shown for reference. Note that at high injection levels, the CB QFL can get very close to the QW barrier energy level and the calculation in Eq. (5.3) should include the continuum states; otherwise, the QFL value will be overestimated.



(a)

(b)

(c)

Figure 5.4: (a) Quasi-Fermi-level (QFL) separation ΔF as a function of the injection carrier density at temperature $T = 283\,\mathrm{K}$, $T = 313\,\mathrm{K}$, and $T = 343\,\mathrm{K}$. (b) QFL for conduction band and (c) QFL for valence band as functions of the injection carrier density. The zone-center ($k_t = 0$) energies for the minibands are shown for reference.

Once we understand how the injected carriers populate in the energy bands, we need to calculate the strength of the transition, which is determined by the optical transition matrix, or the momentum matrix. The optical

transition matrix is calculated from the wavefunction overlap between subbands, as shown in Appendix B. We consider the transition separately when the interacting light polarization is TE (parallel to QW) and TM (perpendicular to QW). Figure 5.5 shows the squared magnitude of the momentum matrix (normalized by $m_0 E_p/6$) for different light polarizations and different types of interband transitions. We can see the C1-HH1 transition favors the TE polarization and allows little TM polarization unless going far away from zone center (large k_t). This is due to the lack of the $|Z\rangle$ p-state in the HH basis. On the other hand, the C1-LH1 transition favors the TM polarization.

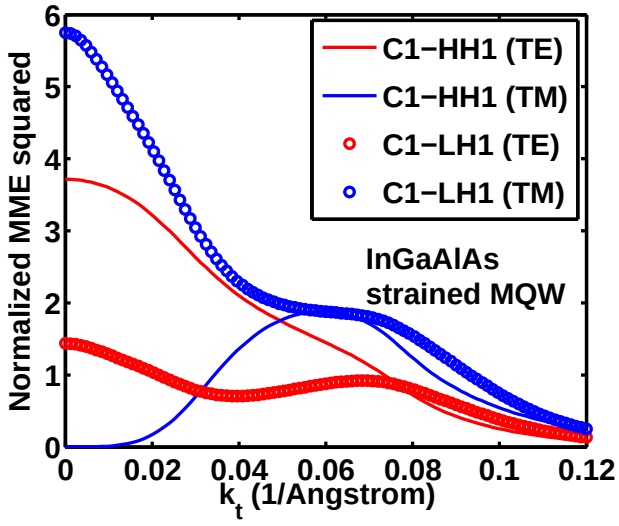


Figure 5.5: Normalized squared magnitude of the momentum matrix for C1-HH1 (solid) and C1-LH1 (circle) transitions with the TE (red) and TM (blue) light polarizations.

The large densities of electrons and holes in the laser active region bring in the many-body effects, which cause the band gap renormalization. Thus, we need to account for the red-shift of the band edge with the increasing injection level. The band gap shrinkage is modeled with a cubic-root dependence on the carrier density [48] as

$$\Delta E_g = -\Delta E_{\mathrm{BR}} (n_{\mathrm{2D}})^{1/3} \tag{5.4}$$

where ΔE_{BR} is the band gap renormalization constant for quantum wells, and n_{2D} is the surface carrier density in each quantum well normalized by

10^{12} cm^{-2}. Furthermore, we include the temperature dependence of the material band gap [58] as

$$E_g(T) = E_g(T=0) - \frac{\alpha T^2}{T+\beta} \tag{5.5}$$

where α and β are the Varshni parameters [59].

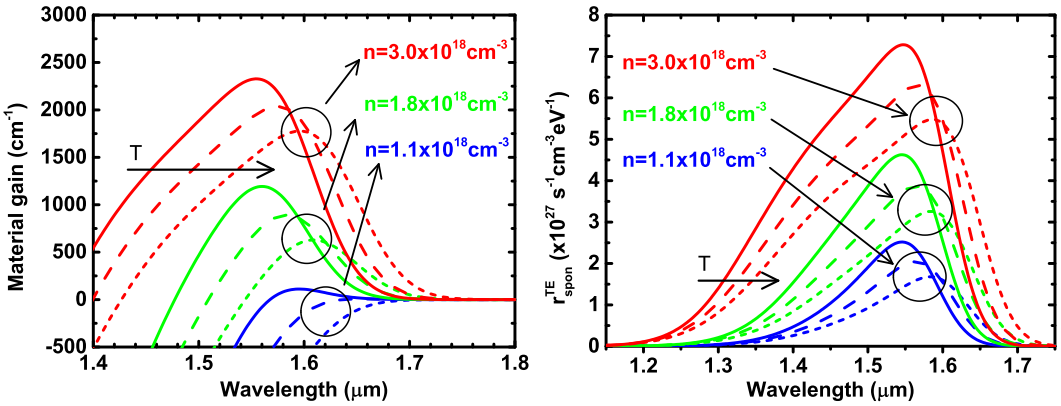
Our InGaAlAs gain model has been verified by experimental data [104]. Figures 5.6(a) and 5.6(b) show the transverse electric (TE) polarized (electric field parallel to QWs) material gain and the TE spontaneous emission rate per unit volume per unit energy interval ($s^{-1}cm^{-3}eV^{-1}$), respectively, for the InGaAlAs QWs at different temperatures and different carrier densities. Increasing temperature results in the red-shift of the gain and spontaneous emission spectra. Increasing carrier density results in the blue-shift of both spectra under low injection due to band-filling, but red-shift under high injection due to band gap renormalization. The total spontaneous emission rate per unit volume ($s^{-1}cm^{-3}$) is the integration over the emission spectrum averaged among the TE and TM polarizations [55],

$$R_{\text{sp}} = \int_0^\infty \frac{1}{3}[2 \times r_{\text{spon}}^{\text{TE}}(\hbar\omega) + r_{\text{spon}}^{\text{TM}}(\hbar\omega)]d(\hbar\omega) \tag{5.6}$$

We see both the peak gain and the total spontaneous emission rate decrease with temperature.

5.3 Optical Modeling of High-Contrast-Grating Tunable VCSELs

The top mirror of the tunable HCG VCSEL consists of 2-4 pairs of p-doped DBR and a TE-HCG (electric field parallel to HCG bars) with an air gap in between the two regions. The bottom mirror consists of 40-55 pairs of n-doped DBR composed of alternating InGaAlAs and InP layers. The air-gap thickness and, consequently, the lasing wavelength can be tuned by the MEMS control voltage. Since the device diameter is large (between 10-25 μm) compared to the emission wavelength, the fundamental transverse mode profile approaches a plane wave, and the effective index approaches the material refractive index. In this case, the transfer matrix method [55] can reduce



(a) (b)

Figure 5.6: (a) TE-polarized material gain and (b) TE-polarized spontaneous emission rate calculated for InGaAlAs quantum wells at $T = 283$ K (solid), $T = 313$ K (dashed), and $T = 343$ K (dotted), with carrier densities $n = 1.0 \times 10^{18}$ cm^{-3} (blue), $n = 1.8 \times 10^{18}$ cm^{-3} (green), and $n = 3.0 \times 10^{18}$ cm^{-3} (red).

the 3D problem to a 1D problem in the longitudinal direction, and provide an accurate prediction of the top and bottom mirror reflectivity, cavity resonance wavelength, confinement factor, quality factor, and threshold material gain.

In Chapter 4 we have obtained the optical properties of HCGs, particularly the fundamental-mode reflection coefficient. In this section, the complex reflection coefficient of the TE-HCG is used as the boundary condition for the transfer matrix method, as illustrated in Fig. 5.7(a). The transfer matrix method uses a forward and a backward propagating plane wave for each layer, characterized by two complex coefficients A_n and B_n. The boundary conditions at the last layer determine the relationship $B_{n+1} = r_{\mathrm{HCG}} A_{n+1}$, and the complex coefficient r_{HCG} for the fundamental-mode HCG reflection is obtained in Chapter 4.

Figure 5.7(b) shows the reflectivity of the top mirror (including the DBR, air gap and HCG), bottom mirror, and HCG alone. The top DBR, though only four pairs, increases the reflection bandwidth of the HCG alone. The bottom DBR bandwidth is narrower because of the large number of pairs, and the round-trip high reflection window of the Fabry-Pérot cavity is determined by the bottom DBR.

Figure 5.8(a) shows the longitudinal profile of the electric field magnitude

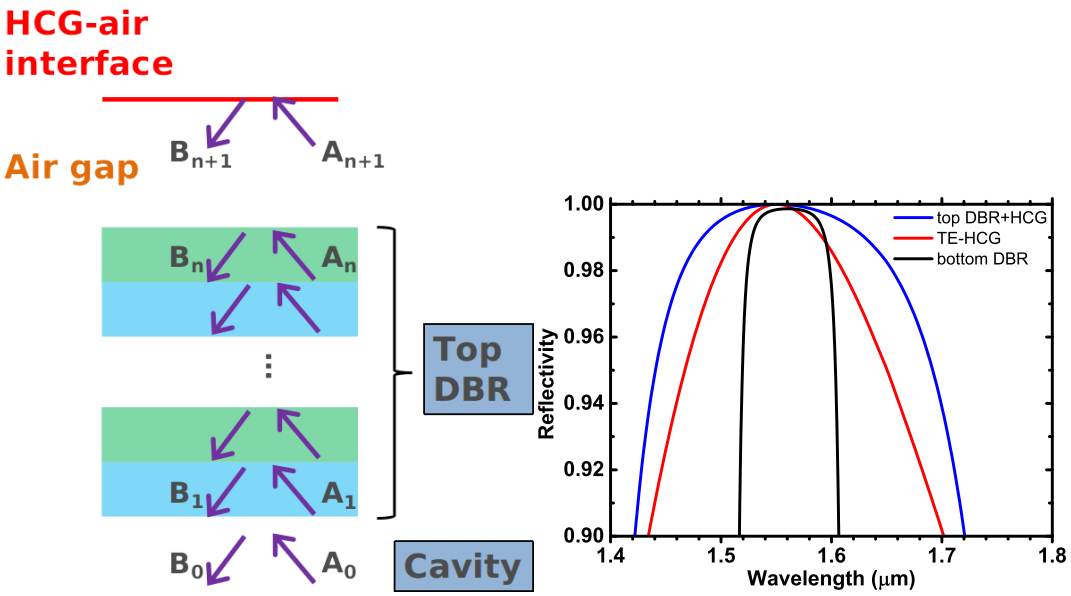


Figure 5.7: (a) Transfer matrix method for calculating the reflection magnitude and phase of the top mirror. The high-contrast grating (HCG) reflection properties are used as the boundary conditions at the HCG-air interface. (b) The reflectivity of the top (blue) and bottom (black) mirrors calculated by the transfer matrix method, plotted with the reflectivity of HCG alone (red).

obtained from transfer matrix method and the refractive index across the device layer structure at wavelength $\lambda = 1.55\,\mu$m. We observe that at this desired wavelength, the laser cavity is indeed at the Fabry-Pérot resonance, and the air gap between the HCG and the VCSEL top surface supports full periods of standing waves. Figure 5.8(b) shows the magnified profiles in the laser cavity. We see that the peak of the electric field overlaps well with the laser active region made of the MQWs. From here we can also obtain the cavity quality factor Q, photon lifetime τ_p for the resonant mode, and the optical confinement factor Γ, which are important design parameters for a laser cavity and will be discussed later.

By adjusting the air-gap thickness, we can investigate the tunability of both the magnitude and phase of the top mirror reflection. The peak reflectivity of the top mirror changes little with the air-gap thickness. However, there is significant variation in the shape and asymmetry of the reflection spectrum, as shown in Fig. 5.9(a). The air-gap thickness largely perturbs the phase of the wave reflected by HCG, which ultimately determines the

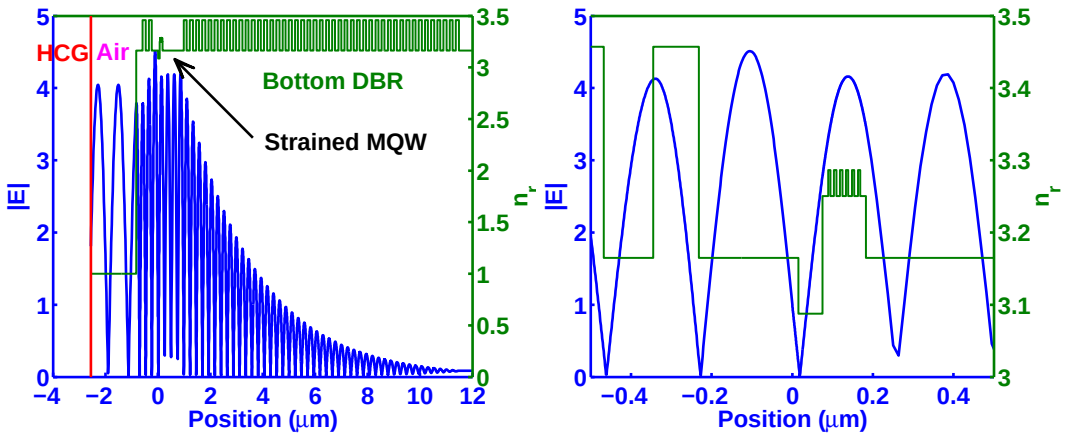


Figure 5.8: (a) Longitudinal standing wave pattern of the electric field magnitude (blue) and the refractive index profile (green) across the device layer structure. The red line indicates the interface between the air gap and the high-contrast grating. (b) Magnified view of (a) showing the overlap between the electric field and the multiple quantum wells as the gain region.

resonance wavelength. Similar to the Fabry-Pérot model, we define the total round-trip phase as

$$\phi_{\text{total}}(\lambda) = \phi_{\text{top}}(\lambda) + \phi_{\text{cavity+bottom}}(\lambda) \tag{5.7}$$

where ϕ_{top} is the top reflection phase and $\phi_{\text{cavity+bottom}}$ is the bottom reflection phase that includes the cavity region. The resonance condition is determined by

$$\phi_{\text{total}}(\lambda_{\text{r}}) = 2m\pi, \quad m \in \text{integer} \tag{5.8}$$

Figure 5.9(b) shows the round-trip phase spectra for different air-gap thicknesses, indicating the tunability of the cavity, where zero-crossing points correspond to cavity resonances. Due to the change of the resonance wavelength, the reflectivity at resonance also changes largely with the air-gap thickness, as indicated by the circles in Fig. 5.9(a), though the peak reflectivity remains nearly the same. The transfer matrix method can take into

account complex effective indices in the layered medium as

$$n = n' + in'' = \begin{cases} n' + i(\alpha_{\mathrm{i}} - g)\dfrac{\lambda}{4\pi}, & \text{in QWs} \\ n' + i\alpha_{\mathrm{i}}\dfrac{\lambda}{4\pi}, & \text{elsewhere} \end{cases} \tag{5.9}$$

where g is the QW material gain and α_{i} is the material intrinsic loss. The effective index real part n' is assumed constant since the change induced by the gain in QWs is negligible. Further, the small change in the thin QWs has little effect on optical modes. In order to calculate the threshold material gain g_{th} and mirror loss α_{m}, we define the round-trip gain at resonance to be

$$G_{\mathrm{total}}(g) = -\ln\left(\frac{1}{|r_{\mathrm{top}}(\lambda)|^2 |r_{\mathrm{cavity+bottom}}(\lambda, g)|^2}\right)\Bigg|_{\lambda=\lambda_{\mathrm{r}}} \tag{5.10}$$

where r_{top} and $r_{\mathrm{cavity+bottom}}$ are the complex reflection coefficients of the top mirror, and the bottom region (including cavity region and bottom DBR), respectively. Then the threshold material gain can be found by setting the round-trip gain to be zero

$$G_{\mathrm{total}}(g_{\mathrm{th}}) = 0 \tag{5.11}$$

Since the mirror loss is equal to the threshold modal gain G_{th} when the intrinsic loss is zero, we can find the mirror loss as

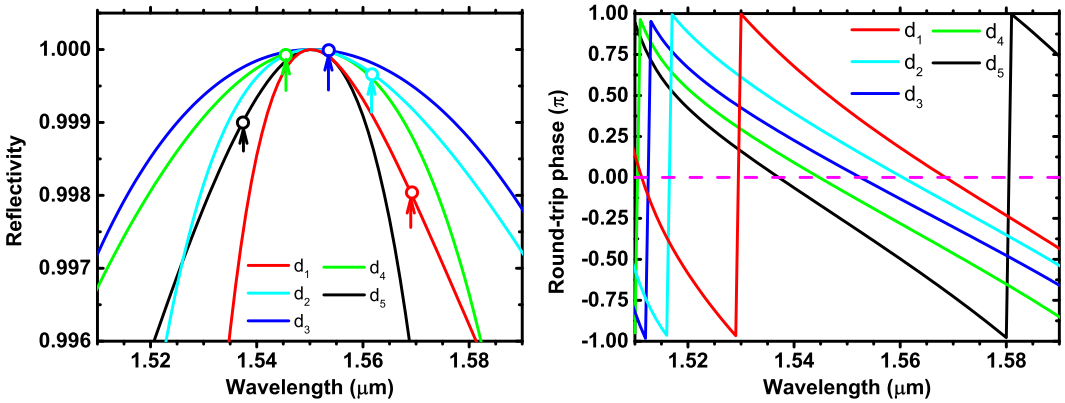
$$\alpha_{\mathrm{m}} = G_{\mathrm{th}}|_{\alpha_{\mathrm{i}}=0} = \Gamma g_{\mathrm{th}}|_{\alpha_{\mathrm{i}}=0} \tag{5.12}$$

where Γ is the confinement factor calculated from the transfer matrix method at a given air-gap thickness. The photon lifetime can be found as

$$\frac{1}{\tau_p} = v_g(\alpha_{\mathrm{m}} + \alpha_{\mathrm{i}} + \alpha_{\mathrm{d}}) = \frac{\omega}{Q_{\mathrm{rad}}} + \frac{\omega}{Q_{\mathrm{mat}}} + \frac{\omega}{Q_{\mathrm{d}}} \tag{5.13}$$

where α_{d} accounts for the diffraction loss due to the finite-size effect, and the tilting and bending of the HCG caused by the MEMS tuning [106]. Q_{rad}, Q_{mat}, and Q_{d} refer to the quality factors associated with the radiation loss, material loss, and diffraction loss, respectively.

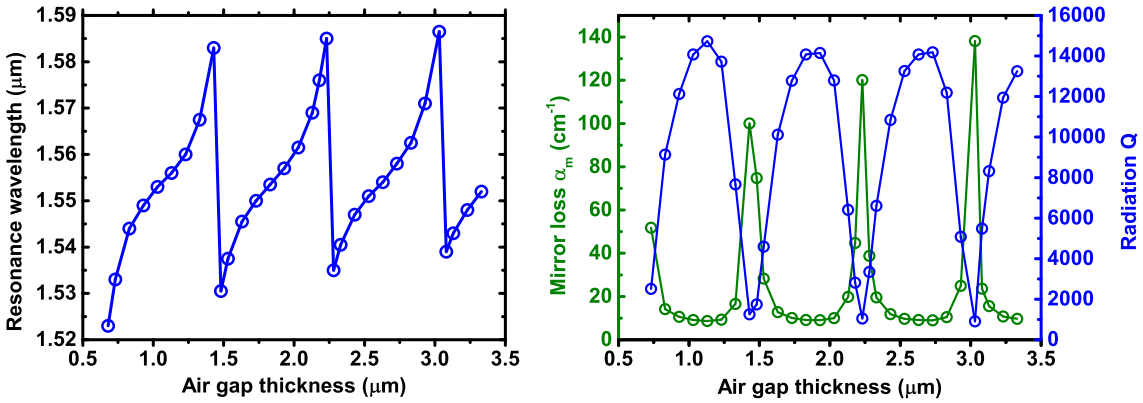
Figure 5.10(a) shows the tuning of the cavity resonance wavelength by controlling the air-gap thickness through the MEMS. The linear tuning range can be as wide as 20 nm, with a tuning efficiency around 0.04 nm/nm. Figure



(a) (b)

Figure 5.9: (a) The reflectivity of the top mirror with different air-gap thicknesses: $d_1 = 2.13$ μm, $d_2 = 2.03$ μm, $d_3 = 1.83$ μm, $d_4 = 1.63$ μm, and $d_5 = 1.53$ μm. The circles indicate the corresponding resonance wavelengths at different air-gap thicknesses. (b) Total round-trip phase spectra in the Fabry-Pérot model with different air-gap thicknesses. The zero-crossing points of the total round-trip phase determine the resonance wavelengths.

5.10(b) shows the cavity mirror loss and the radiation Q at different air-gap thicknesses. We can see that when air-gap thickness is 1.83 μm, the reflection spectrum in Fig. 5.9(a) is most symmetric. It also corresponds to the center of the tuning range in Fig. 5.10(a), and the lowest mirror loss and the highest radiation Q in Fig. 5.10(b).



(a) (b)

Figure 5.10: (a) Cavity resonance wavelengths of the HCG VCSEL at different air-gap thicknesses controlled by the MEMS. (b) Cavity mirror loss α_{m} and radiation quality factor Q_{rad} as functions of the air-gap thickness.

Moving away from the linear tuning range, the lasing wavelength begins to change very quickly with the air-gap thickness, which is mainly caused by the bottom DBR. The phase delay from the air gap changes linearly with the gap thickness, while the HCG phase dispersion is also nearly linear, yet the phase dispersion from the bottom DBR is highly nonlinear for off-resonance. When the tuning goes beyond a certain point, there is a sudden jump in the resonance wavelength, indicated by the discontinuities in Fig. 5.10(a). The discontinuity is due to the switching between two longitudinal cavity modes. Within the linear tuning range, the mirror loss and radiation Q change very little, while outside of this range, the mirror loss sharply increases and the radiation Q sharply decreases. This is due to the significant decrease in reflectivity from the bottom DBR as the resonance wavelength shifts away from the center of the reflection bandwidth spectrum.

5.4 Rate Equations for Tunable VCSELs

After obtaining the gain $g(\lambda, n, T)$ and spontaneous emission rate $R_{\rm sp}(n, T)$ from the $\mathbf{k} \cdot \mathbf{p}$ method, and the photon lifetime τ_p, the confinement factor Γ and the mirror loss $\alpha_{\rm m}$ from the transfer matrix method, the output power of the HCG tunable VCSELs is modeled using the rate equations [48, 55, 62] for the carrier density n and the photon density S

$$\begin{aligned} \frac{dn}{dt} &= \eta_i \frac{I - I_l(n, T_a) - I_{\rm sh}(I)}{qV_a} - R_{\rm nr}(n) - R_{\rm sp}(n, T_a) - R_{\rm st}(n, T_a)S \\ \frac{dS}{dt} &= \Gamma R_{\rm st}(n, T_a)S - \frac{S}{\tau_p} + \Gamma \beta_{\rm sp} R_{\rm sp}(n, T_a) \end{aligned} \tag{5.14}$$

where $\beta_{\rm sp}$ is the spontaneous emission coupling factor, and η_i is the current injection efficiency. The active region temperature T_a can be obtained from the substrate temperature $T_{\rm sub}$, input electric power (VI), and output light power P as

$$T_a = T_{\rm sub} + R_{\rm th}(VI - P) \tag{5.15}$$

where $R_{\rm th}$ is the thermal resistance in K/mW. The cavity resonance wavelength also has a red-shift with increasing temperature due to the change of the material refractive index and the thermal expansion of the cavity. The

change of lasing wavelength due to thermal effects is

$$\Delta\lambda = \frac{d\lambda}{dT}\Delta T \tag{5.16}$$

where $d\lambda/dT$, obtained from experiments, is around 0.102 nm/K, and ΔT is known once the active region temperature is obtained in Eq. (5.15). The non-radiative recombination rate and the stimulated emission rate can be calculated as

$$R_{\mathrm{nr}}(n) = v_s \frac{A_a}{V_a} n + Cn^3 \tag{5.17}$$

$$R_{\mathrm{st}}(n, T_a) = v_g g(\lambda, n, T_a) \tag{5.18}$$

where v_s is the surface recombination velocity, C is the Auger recombination coefficient, A_a and V_a are the surface area and volume of the active region, respectively, and v_g is the group velocity in the active region.

In order to account for injected carriers that pass through the quantum wells without undergoing recombination, we consider the series leakage current as

$$I_l(n, T_a) = I_{l0} \cdot \exp\left(\frac{(F_c - F_v) - E_{g,\mathrm{barrier}}}{kT_a}\right) \tag{5.19}$$

where F_c and F_v are the quasi-Fermi levels in conduction band and valence band, respectively, $E_{g,\mathrm{barrier}}$ is the band gap of the QW barrier, and I_{l0} is a leakage current parameter. As the quasi-Fermi level separation $(F_c - F_v)$ becomes closer to the QW barrier band gap, the leakage current significantly increases, which indicates large leakage currents at high injection levels.

Due to the incomplete electrical isolation of the proton implantation, we include $I_{\mathrm{sh}}(I)$ in the rate equations as the shunt leakage current. The shunt leakage is dependent on the injection current rather than the carrier density, and the carrier pinning effect does not clamp the shunt leakage. The shunt leakage path can be considered as a leakage diode in parallel with the laser diode. When the laser diode has a small turn-on voltage compared to the shunt diode, the laser diode path behaves like a small resistance, and the voltage is almost linear with the total current. The shunt diode current depends on the voltage exponentially. Thus, in this case, it is a good approximation to model the shunt leakage current as an exponential function of the total current.

If the shunt diode turns on earlier than the laser diode, the shunt leakage

current increases with total current linearly at first, and the laser diode is nearly an open circuit. As the voltage increases, the laser diode turns on, and the circuit becomes two parallel diodes. Since the current through each diode depends on the voltage exponentially, the two currents are polynomial functions of each other. Thus, in this case, we can relate the total current to the shunt leakage current as a linear function at first and a polynomial function after the laser diode turns on.

The output light power can be obtained as

$$P = \beta_{c1}\hbar\omega S\frac{V_a}{\Gamma}v_g\alpha_{\rm m} + \beta_{c2}\hbar\omega R_{\rm sp}V_a \tag{5.20}$$

where β_{c1} and β_{c2} account for the coupling efficiencies for the stimulated emission and spontaneous emission power.

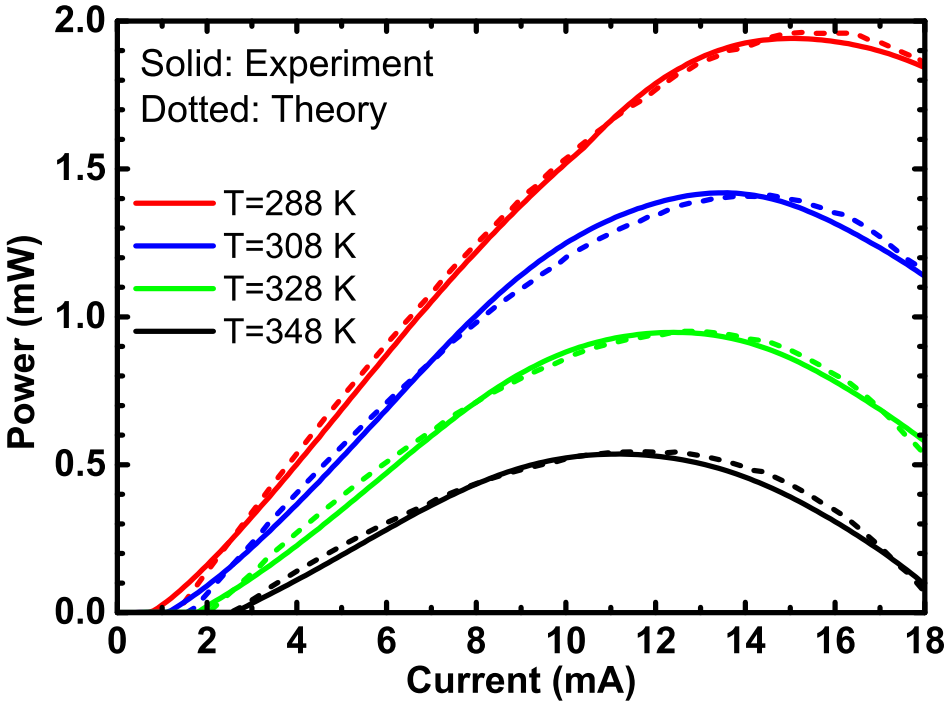


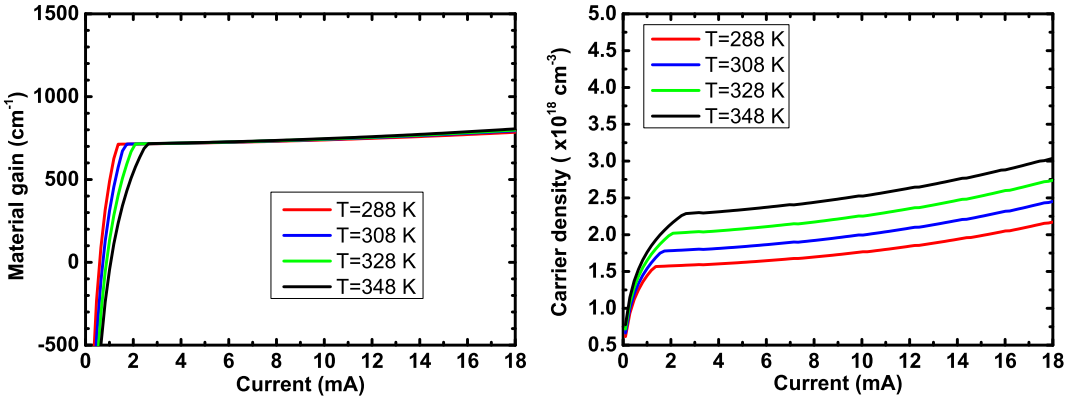
Figure 5.11: Comparison between the theoretical and experimental L-I curves for a fixed-gap TE-HCG VCSEL at different temperatures.

Figure 5.11 shows the theoretical temperature-dependent L-I curves of a fixed-gap (fixed-wavelength) TE-HCG VCSEL and the excellent agreement with experimental data. Our theory can accurately explain the temperature dependence of the threshold current and the rollover of the output power

due to thermal effects. The threshold current increases with temperature because the material gain is reduced at higher temperature for a given carrier density. Therefore more carriers are required to increase gain to a high enough level to overcome the loss. Thus the threshold current is larger at higher temperature. The rollover can be caused by a combination of many mechanisms. Firstly, as larger current is injected, the active region temperature increases, which causes the thermal expansion of the cavity and the change of the material refractive indices. Thus, the lasing wavelength has a red-shift with temperature. Meanwhile, the material gain also has a red-shift with increasing temperature, as shown in Fig. 5.6. However, the lasing wavelength red-shift is slower than that of the material gain, causing the detuning of the gain peak and the lasing wavelength and the reduction of the stimulated emission rate and output power. Secondly, the material gain itself decreases with temperature even without considering the detuning, as also shown in Fig. 5.6. Thirdly, the series leakage current increases at higher carrier densities and high temperatures, as indicated in Eq. (5.19). Higher carrier density also gives rise to a larger non-radiative recombination current and larger spontaneous emission current. Therefore, the current contributing to the lasing mode is reduced, resulting in less lasing power. Furthermore, the shunt leakage increases with the injection current, and directly contributes to the rollover.

Figure 5.12(a) and Fig. 5.12(b) show the gain and carrier density solved from the rate equations as functions of the injection current. The kinks in the curves correspond to the lasing threshold, with higher substrate temperatures resulting in larger threshold currents, as described above. Both the gain and the carrier density should be pinned at their threshold values if thermal effects are not considered. However, the unpinning effect is observed in our theoretical results. The red-shift of the lasing wavelength at higher injection currents causes the decrease of the HCG and DBR reflectivity and the increase of mirror loss. Therefore, the gain is pinned at slightly higher values to overcome the loss when current increases, as shown in Fig. 5.12(a).

The carrier unpinning [63, 107] shown in Fig. 5.12(b) is mainly caused by the degradation of the material gain at elevated temperatures as injection current increases, and the detuning between the gain peak and the cavity resonance. To compensate the reduction of gain at cavity resonance, more carriers are required as current increases. At the same injection current above



(a) (b)

Figure 5.12: (a) Material gain and (b) carrier density solved from the rate equations as functions of the injection current at different substrate temperatures.

threshold, higher substrate temperature also requires a larger carrier density to maintain enough material gain, thus larger non-radiative, spontaneous emission, and leakage currents. This results in a smaller portion of the injected current contributing to the lasing mode at a higher temperature, and both the output power and the wall-plug efficiency are reduced, as shown in Fig. 5.11.

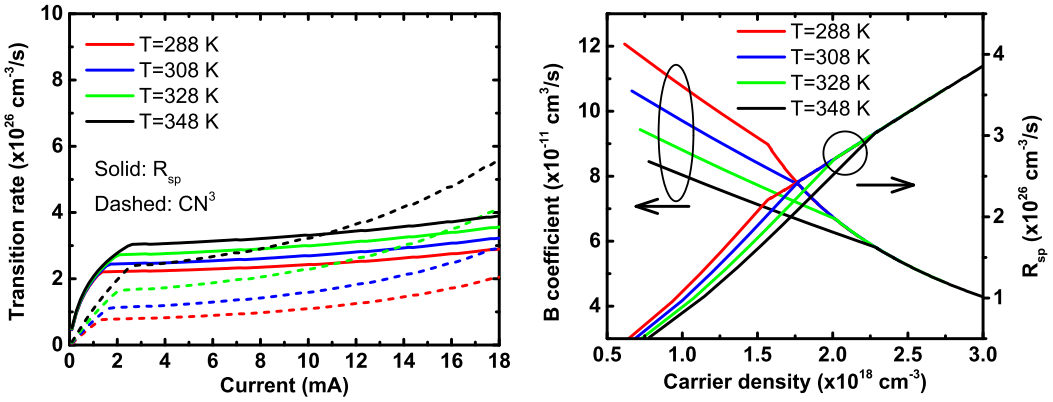
Figure 5.13(a) shows the spontaneous emission rate $R_{\rm sp}$ and the Auger recombination rate $R_{\rm Auger}$ calculated as functions of the injection current. Due to the unpinning of the carrier density, both $R_{\rm sp}$ and $R_{\rm Auger}$ keep increasing above threshold, and they are both larger for higher substrate temperatures. However, $R_{\rm sp}$ is less temperature-sensitive than $R_{\rm Auger}$. Even though the carrier density n is larger with higher substrate temperature at a given injection current, as shown in Fig. 5.12(b), the increase of temperature also causes $R_{\rm sp}(n, T_a)$ to drop, as shown in Fig. 5.6(b). Therefore, compared to $R_{\rm Auger}$, $R_{\rm sp}$ increases with substrate temperature much slower at a fixed injection current.

To see the temperature-dependent spontaneous emission, we obtain the B coefficient as

$$B(n,T) = \frac{R_{\rm sp}(n,T_a)}{n^2} \tag{5.21}$$

where $R_{\rm sp}$, n, and T_a are all solved from the rate equations at a given injection current. The relationship among $R_{\rm sp}$, n, and T_a at different substrate

temperatures are shown in Fig. 5.13(b). The B coefficient decreases with carrier density due to the increasing active region temperature. At the same carrier density, the B coefficient with lower substrate temperature is indeed larger. The four curves are pinned to the same curve due to stimulated emission, where the kinks indicate the thresholds. At the same carrier density, $R_{\rm sp}$ is also larger with lower substrate temperature. Below threshold, $R_{\rm sp}$ increases with n almost quadratically, yet the curvature is reduced by the increase of the active region temperature.



(a) (b)

Figure 5.13: (a) Calculated spontaneous emission rate (solid) and Auger recombination rate (dashed) as functions of the injection current at different substrate temperatures as labeled. (b) Relationship among the B coefficient, spontaneous emission rate, and the carrier density, all of which are solved from the rate equations at different substrate temperatures as labeled.

Figure 5.14 shows the five current mechanisms that comprise the injection current, including the current contributing to stimulated emission, spontaneous emission, non-radiative recombination, the current leaking through quantum wells, and the shunt leakage current, at four different substrate temperatures. We can see indeed smaller percentage of the injection current goes into the lasing mode when the substrate temperature is higher.

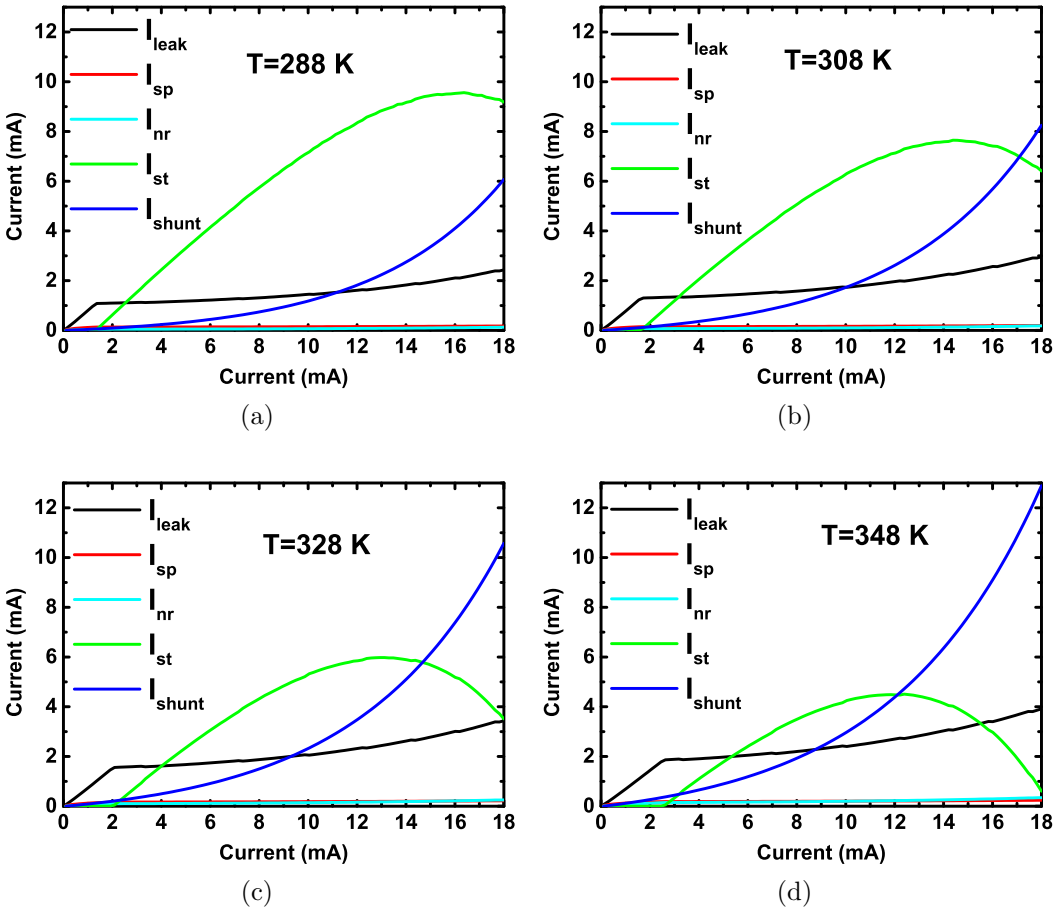


Figure 5.14: The stimulated emission current (green), spontaneous emission current (red), non-radiative recombination current (cyan), series leakage current (black), and shunt leakage current (blue) solved as functions of the injection current with substrate temperatures at (a) $T = 288$ K, (b) $T = 308$ K, (c) $T = 328$ K, and (d) $T = 348$ K.

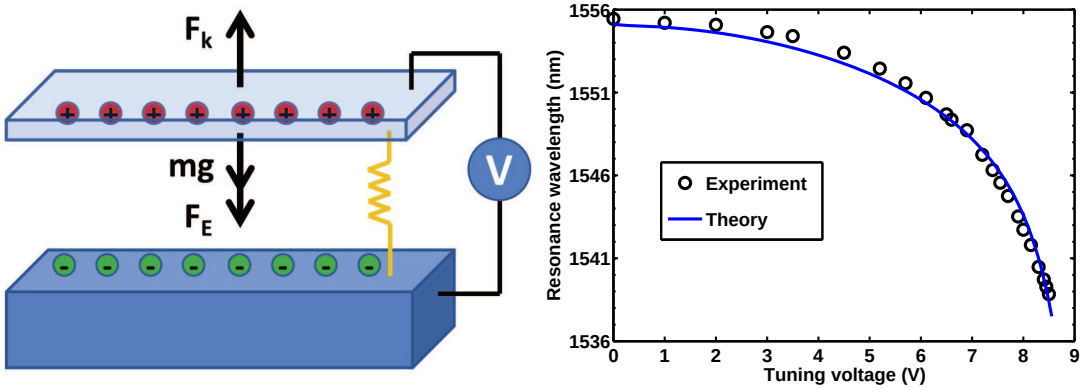
5.5 Electrostatic Model for MEMS and Tunable Resonance

In order to correlate the lasing wavelength and the L-I curve with the MEMS tuning voltage, we use an electrostatic model for the MEMS that controls the air-gap thickness, as shown in Fig. 5.15(a). The gravity of the MEMS top plate and the attractive force due to the opposite charges on the two MEMS plates are balanced by the MEMS elastic force. The force equations are

$$\begin{aligned} k(h_0 - x_0) &= mg, && \text{for } V = 0 \\ k(h_0 - x) &= mg + F_E = mg + \frac{\epsilon A V^2}{2x^2}, && \text{for } V \neq 0 \end{aligned} \tag{5.22}$$

where k is the spring constant for the elastic force F_k, F_E is the electrostatic force, h_0 is the air-gap thickness when no charge is on the plate and gravity is not considered, i.e., the MEMS has no elastic deformation. x_0 is the air-gap thickness when the control voltage is zero (no charge), and x is the air-gap thickness when the control voltage is V. From Eq. (5.22) we can obtain the mapping between the control voltage V and the air-gap thickness x as

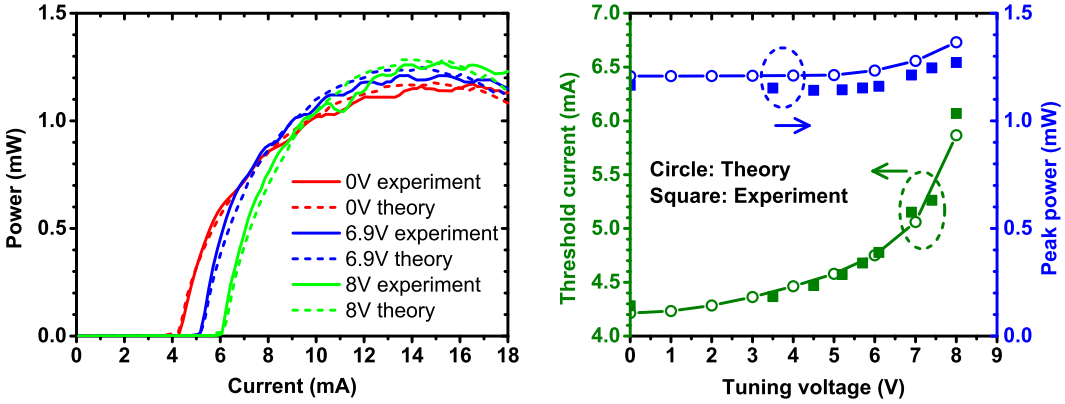
$$x^2(x_0 - x) = \frac{\epsilon A V^2}{2k} \tag{5.23}$$



(a) (b)

Figure 5.15: (a) Schematic diagram of the electrostatic model for MEMS controlling the air-gap thickness. (b) Theoretical and experimental resonance wavelengths versus tuning voltage.

Since the mapping between the air-gap thickness and resonance wavelength is obtained from Fig. 5.10(a), we can correlate the tuning voltage and resonance wavelength using Eq. (5.23). By taking the spring constant k as the only fitting parameter, our theoretical results match very well with the experimental data, as shown in Fig. 5.15(b). The fitted spring constant k is 0.16 N/m.



(a) (b)

Figure 5.16: (a) Theoretical (dashed) and experimental (solid) L-I curves of a MEMS-controlled TE-HCG VCSEL with different tuning voltages. (b) Theoretical (circle) and experimental (square) peak output powers and threshold currents as functions of the MEMS tuning voltage.

At a given MEMS control voltage, we can calculate the air-gap thickness, which is used as the input to the transfer matrix model, outputting the mirror loss α_{m}, confinement factor Γ, and the quality factor Q. Our rate-equation model further produces the L-I curves at different tuning voltages, as shown in Fig. 5.16(a). We can also see that as we increase the tuning voltage, the air-gap thickness is tuned away from the center of the linear tuning range in Fig. 5.10(a), and the threshold current increases due to the increase of the mirror loss. As shown in Fig. 5.16(b), the change of the threshold current and the peak power is small below $V = 4$ V because the shift of the lasing wavelength is small, as indicated in Fig. 5.15(b). Yet above $V = 4$ V we see a fast increase of the threshold current. Besides of the increase of the mirror loss, the increase of diffraction loss also has a contribution to the large increase of the threshold current. The increase of diffraction loss can be caused by the bending of the HCG reflector due to MEMS tuning. From our model, we estimate the additional diffraction loss $\Delta\alpha_{\mathrm{d}}$ (relative to 0 V)

at 5 V, 7 V, and 8 V to be 12, 20, and 26 cm^{-1}, respectively, which equates to a 0.1%, 0.17%, and 0.23% reduction in the reflectivity, respectively. Both the peak power and the slope of the L-I curve increase slightly with tuning voltage due to the increase of mirror loss. The parameters used in our theoretical model are listed in Table. 5.1.

Table 5.1: Parameters used in our theoretical model.

Name and symbol	Value	Name and symbol	Value
Surface recombination velocity v_s	800∼1200 m/s	Varshni parameter α	0.42 meV/K
Auger coefficient C	2∼9×10^{-29} cm^6/s	Varshni parameter β	271 K
Band gap renormalization constant ΔE_{BR}	25 meV	Thermal resistance R_{th} (measured)	∼1.5 mW/K
Series leakage current parameter I_{l0}	80∼120 mA	Intrinsic loss α_{i}	5∼15 cm^{-1}
Cavity resonance shift $d\lambda/dT$ (measured)	∼0.102 nm/K	Injection efficiency η_i	0.7∼0.9
Diffraction loss α_{d} (Tuning V=0∼8 V)	34∼60 cm^{-1}		

5.6 Thermal Management in High-Contrast-Grating VCSELs

Thermal management is a crucial matter in the VCSEL design. Unlike edge-emitting lasers, VCSELs have much wider cavity mode spacing and the cavity resonance is rather sensitive to heat due to the temperature-dependent material indices and the cavity length. Furthermore, the active region temperature strongly affects the optical gain not only on the peak wavelength but also the maximum gain achievable at a given carrier injection level. The thermal effect comes in not only through the substrate temperature but also through electrical injection due to the finite device thermal resistance. As a result, the substrate temperature affects the threshold current, and the injection increases, so we can often observe the rollover behavior in the output power, as discussed in the previous section.

Figure 5.17 illustrates the thermal detuning between the cavity resonance and the material gain peak as the injection current increases with a fixed substrate temperature at 75°C. The cavity resonance indicated by the vertical lines red-shifts as current increases and cavity temperature rises, due to the increase of material indices and the expansion of cavity length. The material gain also red-shifts with temperature and it is much faster than the red-shift of the cavity resonance. From Fig. 5.6, we know that at the same carrier density, higher temperature gives lower peak gain. From Fig. 5.17 we can tell that in order to maintain the gain at resonance to overcome cavity loss, indeed more carriers are injected because the peak gain is higher at $6I_{th}$. The carrier density at a given injection current is solved from the rate equations in the previous section.

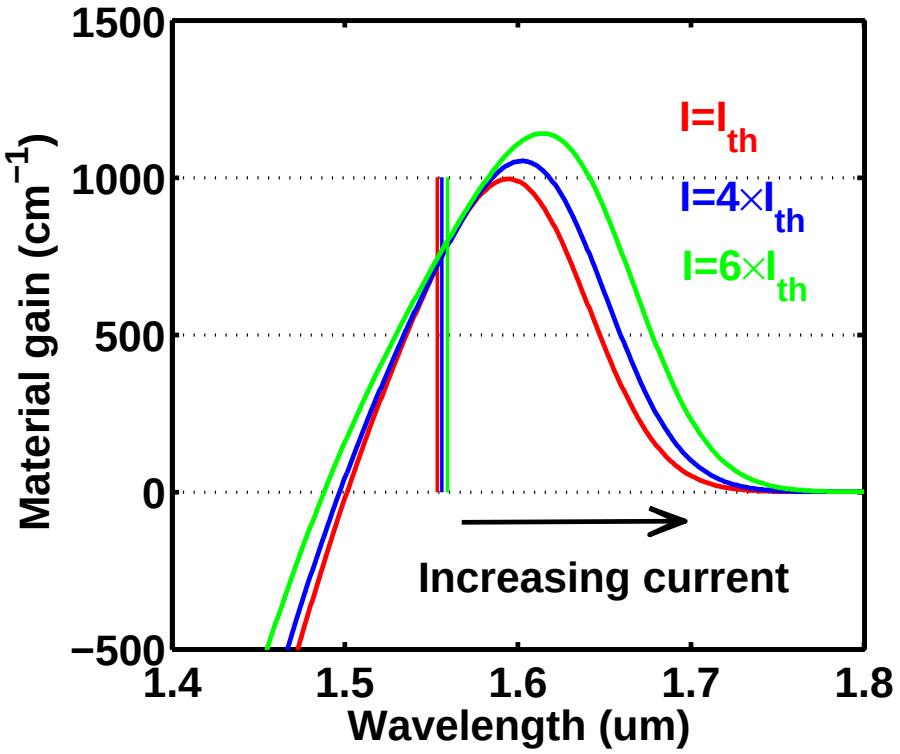
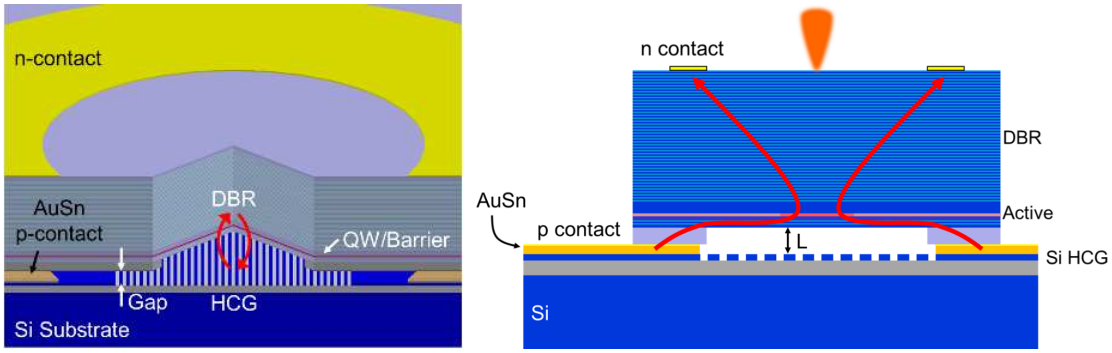


Figure 5.17: Thermal detuning between the cavity resonance (vertical lines) and the material gain peak when the injection current is I_{th} (red), $4I_{th}$ (blue), and $6I_{th}$ (green). The substrate temperature is set at 75°C and I_{th} is the threshold current at this substrate temperature.

Using COMSOL Multiphysics, a finite-element-method simulation tool, the heat distribution of the HCG VCSEL structure, as shown in Fig. 5.18,



(a) (b)

Figure 5.18: Schematic diagram of a VCSEL with silicon HCG as bottom mirror [38]. (a) Tilted-view of the VCSEL cross-section with circulating red arrows indicating optical cavity, drawn to scale. The two material systems are heterogeneously integrated via AuSn thin film, with a hermetically sealed air gap of length L within the cavity. (b) The VCSEL employs a proton implant-defined aperture for current confinement, indicated by red curved lines between contacts.

was modeled by an 8 μm aperture with 25, 45, and 65 mW heat sources in the active region. The boundary conditions are set as thermal insulation at the semiconductor-air interface to provide a worse-case scenario for heat dissipation. The heat transfer in the vertical and radial directions can be very different in a multilayer system such as a DBR [108], where the thermal resistance of each layer adds up in series and in parallel, respectively. Moreover, the heat flow is strongly affected by the alloy impurities and layer interfaces, due to the restriction of the phonon mean free path [109]. In conventional standalone VCSELs, the heat dissipation relies on the bottom DBR in between the heat source and heat sink. In flip-chip-bonded VCSELs, the heat can be carried away through the AuSn bonding layer.

In Fig. 5.19, we show a comparison between a flip-chip bonded VCSEL on SOI using AuSn, and a standalone III-V VCSEL structure, with the DBR thermal conductivity modeled after the approach in [110]. The maximum temperature generated is 125°C for the VCSEL on SOI, whereas the standalone III-V structure reached 183°C, shown in Fig. 5.19(a) and 5.19(b), respectively. Although the flip-chip VCSEL is resting on insulator (SOI), the graph in Fig. 5.19(c) shows that, at a given input thermal power, the average temperature in the active region has a stronger dependence on the vertical thermal conductivity of the DBR for standalone VCSELs (red) than

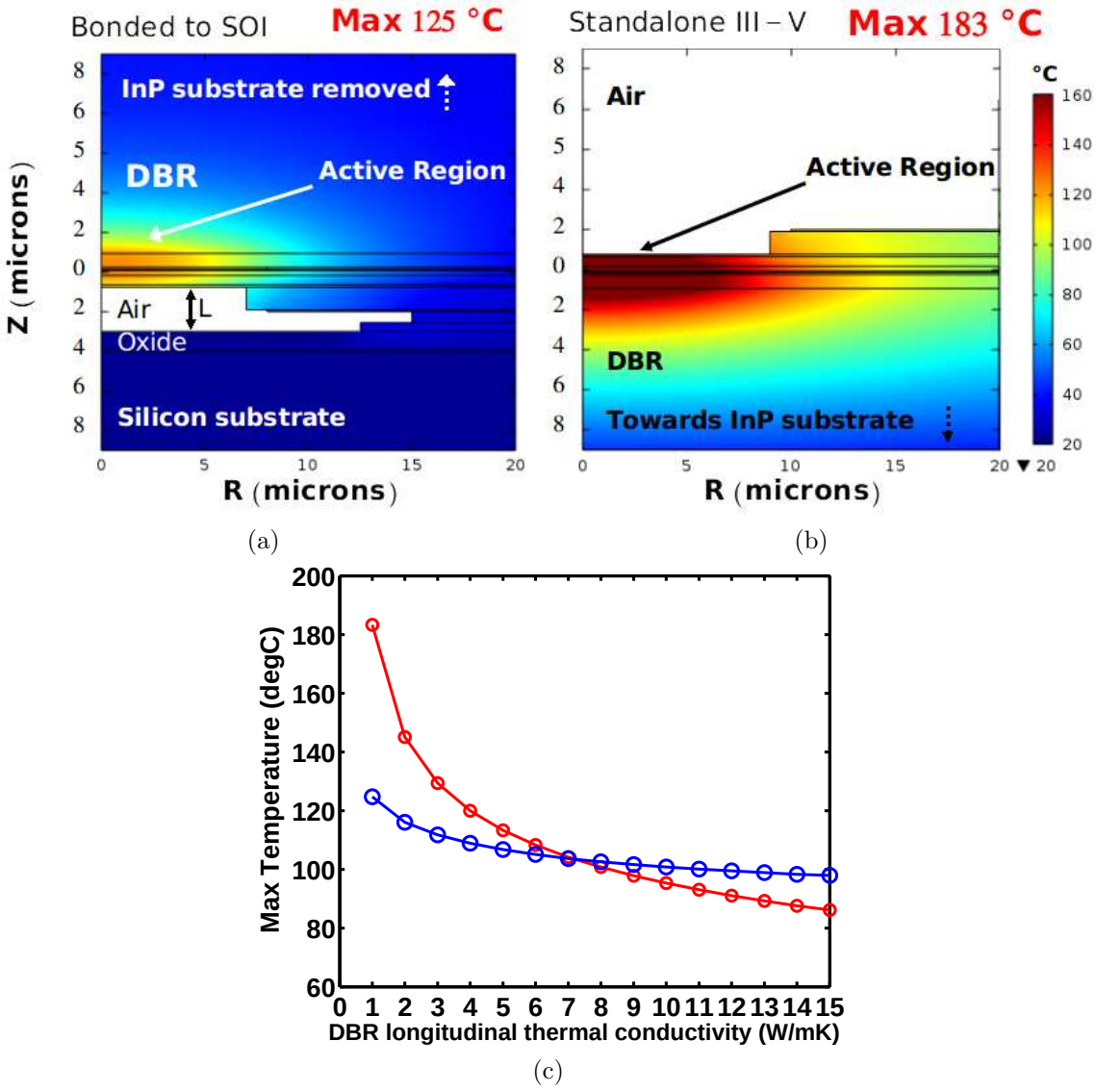


Figure 5.19: Radially-symmetric COMSOL finite-element simulation modeling temperature distribution in VCSEL using 45 mW heat source in active region. (a) Bonded Si-HCG VCSEL on SOI substrate; active region $z = 0$, DBR extends from $0.8\,\mu\text{m} \leq z \leq 11.3\,\mu\text{m}$, air $z \geq 11.3\,\mu\text{m}$ (since InP substrate is removed), thermal oxide $-4\,\mu\text{m} \leq z \leq -4\,\mu\text{m}$, silicon substrate $z \leq 4\,\mu\text{m}$ and $350\,\mu\text{m}$ thick. (b) Standalone III-V VCSEL with InP substrate; structure has air above active region from $z \geq 0.8\,\mu\text{m}$, DBR extends from $-11.3\,\mu\text{m} \leq z \leq -0.8\,\mu\text{m}$, InP substrate $z \leq 11.3\,\mu\text{m}$ and $350\,\mu\text{m}$ thick. (c) Maximum temperature versus DBR quaternary alloy thermal conductivity. The standalone III-V structure reaches higher temperature for DBR thermal conductivities below 7.2 W/mK.

flip-chip-bonded VCSELs (blue). This means the flip-chip-bonded VCSELs can overcome the restriction of heat flow caused by the DBR alloy impurities and interfaces. The thermal conductivity of AuSn is 57 W/mK, roughly an order of magnitude better than for a standard DBR quaternary alloy, i.e. 5 W/mK for AlGaInAs [111]. For alloy compositions with thermal conductivities below 7.2 W/mK, the bonded device exhibits superior thermal performance.

5.7 Summary

We have successfully demonstrated a comprehensive model for the MEMS-tunable HCG VCSELs. The model calculates the temperature-dependent material gain and spontaneous emission spectra of the quantum-well active region. The optical properties of the HCG reflector are modeled with both analytical and numerical methods, showing good convergence. The HCG VCSEL cavity is modeled with the transfer matrix method, which produces important parameters for device-level simulation. The rate-equation model takes into account the thermal effects and our calculated temperature-dependent L-I curves show excellent agreement with experiment. Our MEMS model further correlates the tuning voltage with the resonance wavelength, threshold current, and peak power. The measurements can be accurately explained by our model.

CHAPTER 6

CONCLUSION

6.1 Summary of Key Results

In this dissertation, we have investigated the theory and experiment of novel surface-emitting semiconductor lasers, which enable a wide range of applications in optical communications. On one hand, we focus on metal-cavity micro and nanolasers as promising light sources for dense photonic integrated circuits and power-efficient on-chip optical links. On the other hand, we examine the high-power tunable surface-emitting lasers, which are important components in the wavelength division multiplexing (WDB) systems for high-speed fiber optical communication. Fast and widely tunable light sources also enable applications such as optical coherent tomography (OCT), and light detection and ranging (LIDAR).

We first present the theory and design for surface-emitting micro and nanolasers using submonolayer quantum-dots (QD) as the optical gain medium. Our theoretical formulation of the QD gain allows us to predict the available gain at the desired wavelength with the given level of carrier injection. Our size-dependent cavity model provides the information of cavity loss and resonance. Our rate-equation model allows us to predict the laser light output at a given injection current. This comprehensive theory enables us to design cavity structures where the total loss can be balanced and the lasing threshold is achievable. We further investigate the spontaneous emission from QDs coupled to the optical modes in micro and nanocavities. Our theory successfully explains the sub-threshold and above-threshold QD laser behavior.

The fabrication steps for surface-emitting micro and nanolasers are discussed in details. We demonstrate high-quality plasma etch for forming laser cavities. We develop recipes for the coating of sidewall SiN_x and the opening for the laser top emission window. We present lasing behavior of our QD

and quantum-well (QW) lasers with both metallic and dielectric cavities, at room temperature under electrical injection. Our QW lasers work with continuous-wave (CW) injection with device diameter down to $3\,\mu$m and threshold current as low as 0.58 mA. Our QD lasers work under CW and pulsed injection with device diameters down to $4\,\mu$m and $1\,\mu$m, respectively. Our current-dependent laser emission spectra allow us to study thermal effects on cavity properties and QW/QD light emission. We further investigate the sub-threshold light emission spectra of our QD devices. The extracted spontaneous emission coupling in micro and nanocavities agrees with our theoretical prediction.

We then study the optical properties of the high-contrast grating (HCG), which is a key component in tunable lasers based on the micro-electro-mechanical systems (MEMS). We develop theoretical models for gratings with one-dimensional (1D) periodicity using both the mode-matching method and the surface-integral-equation method based on the periodic Green's function. Excellent agreement is shown between our developed methods and the commercial packages. To understand the physics of HCGs with two-dimensional (2D) periodicity, we develop the rigorous coupled-wave analysis (RCWA) package, which allow us to efficiently predict the optical performance of 2D HCGs with various combinations of structural parameters and wave incidence conditions. Furthermore, we present a top-down design procedure that enables us to efficiently engineer the 2D HCGs for a wide range of applications, such as reflectors, filters, resonators, waveplates, and 2D phase plates. The simulation results of our designed 2D phase plates show ultra-high power efficiency, and excellent agreement with our predicted functionalities, such as beam steering, focusing, beam conversion, and generation of orbital angular momentum of light.

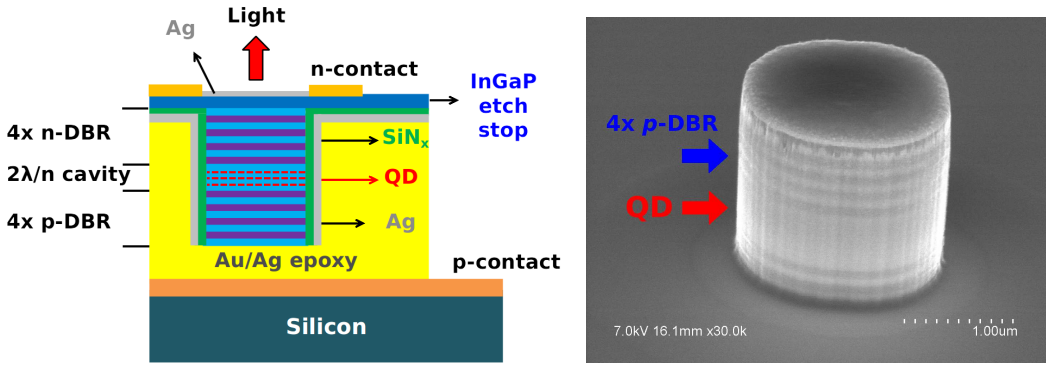
Finally, the HCGs that we design are integrated onto the vertical-cavity surface-emitting lasers (VCSELs), and can be made tunable using MEMS. We present a comprehensive theoretical model for the HCG VCSELs, and it covers from the scale of quantum structures, lasing devices, up to the scale of photonic integration. Our band structure model based on the $\mathbf{k}\cdot\mathbf{p}$ method calculates the optical gain and spontaneous emission of the strained QW active region. The grating reflection properties and the cavity resonance condition are investigated through our optical model. Correlating the results with the electrostatic model for the MEMS, we accurately predict

the measurements on the voltage-controlled lasing wavelength. Furthermore, our calculated temperature-dependent wavelength-tunable light output versus current behavior shows excellent agreement with experiment. Thermal effects on HCG VCSELs are discussed, and directions for heat management are provided.

6.2 Future Prospects

We have demonstrated standalone metal-cavity QD lasers working at room temperature. The future directions include further size reduction and the integration on silicon platform. We propose the new design by reducing the number of distributed Bragg reflector (DBR) pairs from 19/32 to only 4 pairs, as shown in Fig. 6.1. Using fewer pairs of DBR, the high reflection window can be widened, and we can maintain the reflectivity by adding a silver layer to form the hybrid mirror. Furthermore, the series resistance can be largely reduced and the heat dissipation also becomes more efficient. As the device size becomes small, the carrier leakage and temperature stability are important issues to consider. The multi-stack submonolayer quantum dots can provide good carrier confinement and are less temperature sensitive as compared to quantum wells. Thus the thermal rollover of the device L-I curves is expected to be pushed much farther. In order the form the backside (n-side) hybrid mirror, the device is flip-chip-bonded to silicon, and a substrate removal is performed, followed by the deposition of the silver reflector. An InGaP layer is used as the etch-stop layer during substrate removal. The substrate-removed devices now have their sizes largely reduced and are compatible for integration on silicon platform. Flip-chip-bonded devices also are expected to have largely improved thermal properties. As shown in Fig. 6.1, the new design has a total volume of $2.16\,\lambda_0^3$ for a 1-μm-diameter structure, and a total volume of λ_0^3 for a 0.65-μm-diameter structure. The new design is nearly three times smaller than the current design in terms of the size of the metal cavity. In terms of threshold gain, the new design has values of $1302\,\mathrm{cm}^{-1}$ and $1471\,\mathrm{cm}^{-1}$ for the 1-μm-diameter and 0.65-μm-diameter structures, respectively. With the use of submonolayer quantum dots, the threshold condition is achievable at room temperature.

For the HCG VCSELs, we have demonstrated the heterogeneous integra-



(a) (b)

Figure 6.1: (a) Schematic diagram of a metal-cavity quantum-dot surface-emitting nanolaser integrated on silicon. (b) Scanning electron micrograph of a nanocavity with quantum-dot active region.

tion of InP-based long-wavelength devices on the silicon-on-insulator (SOI) substrate, using silicon as the HCG. However, after flip-chip bonding the bottom DBR is still kept. The future direction is to replace the bottom DBR by the HCG mirror, and thus VCSEL will be free of DBR and the cavity is formed by two HCG layers. The device height can be shrunk down to micrometer or sub-micrometer scale without sacrificing the cavity quality factor. The thermal properties are expected to improve even more, and the device structure would be more favorable for dense photonic integration.

We have also provided the designs for 2D phase plates using HCGs, which enables us to engineer the beam properties. The future direction is to integrate the HCG phase plates onto VCSELs. Then the VCSEL output can be steered, focused, or converted to Bessel beams with minimal divergence. The phase plates can also introduce orbital angular momentum (OAM) to the output beam, which provides another degree of freedom for carrying information. The VCSEL arrays integrated for OAM multiplexing are expected to greatly improve the optical communication bandwidth on top of the existing technologies, such as the wavelength division multiplexing.

APPENDIX A

SURFACE INTEGRAL EQUATIONS FOR PERIODIC DIELECTRIC STRUCTURES

We consider a dielectric structure which is periodic in the $\hat{x}$-direction, and invariant along the $\hat{y}$-direction. The scattering problem of an incident plane wave, of which the wave vector is within the xz-plane, can then be separated into a TE (s-polarized) case and a TM (p-polarized) case. We first consider the TE case, where E_y is the component for our formulation. The integral equations based on the Green's function in Region I (incident region) and Region II (transmitted region) can be written as [81]

$$E_{iy}(\mathbf{r}) + \int_{\Lambda} dS' \left\{ E_{1y}(\mathbf{r}')\hat{n}\cdot\nabla' g_{1p}(\mathbf{r},\mathbf{r}') - g_{1p}(\mathbf{r},\mathbf{r}')\hat{n}\cdot\nabla' E_{1y}(\mathbf{r}') \right\} = \begin{cases} E_{1y}(\mathbf{r}), & \mathbf{r} \in V_1 \\ 0, & \mathbf{r} \in V_2 \end{cases} \tag{A.1}$$

$$-\int_{\Lambda} dS' \left\{ E_{2y}(\mathbf{r}')\hat{n}\cdot\nabla' g_{2p}(\mathbf{r},\mathbf{r}') - g_{2p}(\mathbf{r},\mathbf{r}')\hat{n}\cdot\nabla' E_{2y}(\mathbf{r}') \right\} = \begin{cases} 0, & \mathbf{r} \in V_1 \\ E_{2y}(\mathbf{r}), & \mathbf{r} \in V_2 \end{cases} \tag{A.2}$$

Combining the integral equations, we obtain the two SIEs on the surface S',

$$\begin{aligned} -\int_{\Lambda} dS' \left\{ E_{1y}(\mathbf{r}')\hat{n}\cdot\nabla' g_{1p}(\mathbf{r},\mathbf{r}') - g_{1p}(\mathbf{r},\mathbf{r}')\hat{n}\cdot\nabla' E_{1y}(\mathbf{r}') \right\} + \frac{1}{2}E_{1y}(\mathbf{r}) &= E_{iy}(\mathrm{r}) \\ \int_{\Lambda} dS' \left\{ E_{2y}(\mathbf{r}')\hat{n}\cdot\nabla' g_{2p}(\mathbf{r},\mathbf{r}') - g_{2p}(\mathbf{r},\mathbf{r}')\hat{n}\cdot\nabla' E_{2y}(\mathbf{r}') \right\} + \frac{1}{2}E_{2y}(\mathbf{r}) &= 0 \end{aligned} \tag{A.3}$$

where $\mathbf{r} \in S'$. According to the boundary conditions,

$$\begin{cases} E_{1y}(\mathbf{r}') = E_{2y}(\mathbf{r}') = E_y(\mathbf{r}') \\ \frac{\hat{n}\cdot\nabla' E_{1y}(\mathbf{r}')}{\mu_{\mathrm{r}1}} = -i\omega H_{1t}(\mathbf{r}') = -i\omega H_{2t}(\mathbf{r}') = \frac{\hat{n}\cdot\nabla' E_{2y}(\mathbf{r}')}{\mu_{\mathrm{r}2}} \end{cases} \tag{A.4}$$

Then we can let

$$\begin{cases} E_{1y}(\mathbf{r}') = E_{2y}(\mathbf{r}') = E_y(\mathbf{r}') = \sum_l d_l v_l(\mathbf{r}') \\ \hat{n} \cdot \nabla' E_{1y}(\mathbf{r}') = \mu_{\mathrm{r}1} \sum_j c_j v_j(\mathbf{r}') \\ \hat{n} \cdot \nabla' E_{2y}(\mathbf{r}') = \mu_{\mathrm{r}2} \sum_j c_j v_j(\mathbf{r}') \end{cases} \tag{A.5}$$

where $v(\mathrm{r})$ is the basis function for MOM. The matrix equation becomes

$$\begin{bmatrix} \bar{\bar{\mathbf{B}}}_{m\times l} & \bar{\bar{\mathbf{A}}}_{m\times j} \\ \bar{\bar{\mathbf{D}}}_{m\times l} & \bar{\bar{\mathbf{C}}}_{m\times j} \end{bmatrix} \begin{bmatrix} \bar{d}_{l\times 1} \\ \bar{c}_{j\times 1} \end{bmatrix} = \begin{bmatrix} \bar{b}_{m\times 1} \\ \bar{0} \end{bmatrix} \tag{A.6}$$

where

$$\begin{aligned} \bar{\bar{\mathbf{B}}}_{ml} &= -P.V. \int_\Lambda dS' \hat{n} \cdot \nabla' g_{1p}(\mathbf{r}_m, \mathbf{r}') v_l(\mathbf{r}') + \frac{1}{2}\delta_{ml} \\ \bar{\bar{\mathbf{A}}}_{mj} &= \mu_{\mathrm{r}1} P.V. \int_\Lambda dS' g_{1p}(\mathbf{r}_m, \mathbf{r}') v_j(\mathbf{r}') \\ \bar{b}_m &= E_{iy}(\mathbf{r}_m) \\ \bar{\bar{\mathbf{D}}}_{ml} &= P.V. \int_\Lambda dS' \hat{n} \cdot \nabla' g_{2p}(\mathbf{r}_m, \mathbf{r}') v_l(\mathbf{r}') + \frac{1}{2}\delta_{ml} \\ \bar{\bar{\mathbf{C}}}_{mj} &= -\mu_{\mathrm{r}2} P.V. \int_\Lambda dS' g_{2p}(\mathbf{r}_m, \mathbf{r}') v_j(\mathbf{r}') \end{aligned} \tag{A.7}$$

From the matrix equation we obtain the expansion coefficients for the fields on the surface in terms of MOM basis functions. Then we can find the total fields in Regions I and II as the summation of all propagating diffraction orders,

$$\begin{cases} E_y(\mathbf{r}) = E_{iy}(\mathbf{r}) + \sum_n r_n e^{i\mathbf{k}_{1n}^+ \cdot \mathbf{r}} & \mathbf{r} \in V_1 \\ E_y(\mathbf{r}) = \sum_n t_n e^{i\mathbf{k}_{2n}^- \cdot \mathbf{r}} & \mathbf{r} \in V_2 \end{cases} \tag{A.8}$$

where $\mathbf{k}_{1n}^+ = \hat{x}k_{xn} + \hat{z}\sqrt{\omega^2\mu_1\varepsilon_1 - k_{xn}^2}$ and $\mathbf{k}_{2n}^- = \hat{x}k_{xn} + \hat{z}\sqrt{\omega^2\mu_2\varepsilon_2 - k_{xn}^2}$. The mode expansion coefficients can be found as [81]

$$\begin{aligned} r_n &= \frac{i}{2\Lambda} \int_\Lambda dS' \left\{ E_y(\mathbf{r}') \hat{n} \cdot \nabla' \frac{e^{-i\mathbf{k}_{1n}^+ \cdot \mathbf{r}'}}{k_{1zn}} - \frac{e^{-i\mathbf{k}_{1n}^+ \cdot \mathbf{r}'}}{k_{1zn}} \hat{n} \cdot \nabla' E_y(\mathbf{r}') \right\} \\ t_n &= -\frac{i}{2\Lambda} \int_\Lambda dS' \left\{ E_y(\mathbf{r}') \hat{n} \cdot \nabla' \frac{e^{-i\mathbf{k}_{2n}^- \cdot \mathbf{r}'}}{k_{2zn}} - \frac{e^{-i\mathbf{k}_{2n}^- \cdot \mathbf{r}'}}{k_{2zn}} \hat{n} \cdot \nabla' E_y(\mathbf{r}') \right\} \end{aligned} \tag{A.9}$$

We then consider the TM case, where H_y is the component for our formu-

lation. The integral equations in Region I, Region II, and on surface S' are the same as Eq. (A.1), (A.2), and (A.3), except that we change E_{iy}, E_{1y}, and E_{2y} to H_{iy}, H_{1y}, and H_{2y}, respectively. The key difference comes from the boundary conditions, where we need to rewrite Eq. (A.4) as

$$\begin{cases} H_{1y}(\mathbf{r}') = H_{2y}(\mathbf{r}') = H_y(\mathbf{r}') \\ \frac{\hat{n}\cdot\nabla' H_{1y}(\mathbf{r}')}{\varepsilon_{\mathrm{r}1}} = i\omega E_{1t}(\mathbf{r}') = i\omega E_{2t}(\mathbf{r}') = \frac{\hat{n}\cdot\nabla' H_{2y}(\mathbf{r}')}{\varepsilon_{\mathrm{r}2}} \end{cases} \tag{A.10}$$

Then we can let

$$\begin{cases} H_{1y}(\mathbf{r}') = H_{2y}(\mathbf{r}') = \sum_l d_l v_l(\mathbf{r}') \\ \hat{n}\cdot\nabla' H_{1y}(\mathbf{r}') = \varepsilon_{\mathrm{r}1}\sum_j c_j v_j(\mathbf{r}') \\ \hat{n}\cdot\nabla' H_{2y}(\mathbf{r}') = \varepsilon_{\mathrm{r}2}\sum_j c_j v_j(\mathbf{r}') \end{cases} \tag{A.11}$$

The matrix equation is the same as Eq. (A.6) except that

$$\begin{aligned} \bar{\bar{\mathbf{B}}}_{ml} &= -P.V.\int_\Lambda dS'\hat{n}\cdot\nabla' g_{1p}(\mathbf{r}_m,\mathbf{r}')v_l(\mathbf{r}') + \frac{1}{2}\delta_{ml} \\ \bar{\bar{\mathbf{A}}}_{mj} &= \varepsilon_{\mathrm{r}1}P.V.\int_\Lambda dS' g_{1p}(\mathbf{r}_m,\mathbf{r}')v_j(\mathbf{r}') \\ \bar{b}_m &= H_{iy}(\mathbf{r}_m) \\ \bar{\bar{\mathbf{D}}}_{ml} &= P.V.\int_\Lambda dS'\hat{n}\cdot\nabla' g_{2p}(\mathbf{r}_m,\mathbf{r}')v_l(\mathbf{r}') + \frac{1}{2}\delta_{ml} \\ \bar{\bar{\mathbf{C}}}_{mj} &= -\varepsilon_{\mathrm{r}2}P.V.\int_\Lambda dS' g_{2p}(\mathbf{r}_m,\mathbf{r}')v_j(\mathbf{r}') \end{aligned} \tag{A.12}$$

Finding the total field and the mode expansion coefficient for each diffraction order is the same as Eq. (A.8) and Eq. (A.9), except that all the fields are changed to H_y and the surface field expansion uses Eq. (A.11).

APPENDIX B

MODELING OF STRAINED MULTIPLE QUANTUM WELLS WITH THE $\mathbf{k}\cdot\mathbf{p}$ METHOD

The Bloch wave functions at the zone center can be taken from the s-state and p-state wave functions of a hydrogen atom model, which are the spherical harmonics [101], that is, $|S\uparrow\rangle$, $|S\downarrow\rangle$ for conduction band, and $|X\uparrow\rangle$, $|X\downarrow\rangle$, $|Y\uparrow\rangle$, $|Y\downarrow\rangle$, $|Z\uparrow\rangle$, and $|Z\downarrow\rangle$ for valence band.

For most semiconductor photonic applications in the near-infrared and visible range, the effective band gap is sufficiently large such that the mixing among conduction band (CB) and valence band states is negligible. We can solve the CB states from a single-band Hamiltonian with the effective mass approximation,

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left[\frac{1}{m^*(z)}\frac{\partial}{\partial z}\psi(z)\right]+\frac{\hbar^2 k_t^2}{2m^*(z)}\psi(z)+V(z)\psi(z)=E_n(k_t)\psi(z) \quad \text{(B.1)}$$

with $E_n(k_t)$ being the n-th subband energy with a given transverse wave number k_t. In order to include the strain effect on the band edge, we have $V = E_c + A_\epsilon$, where E_c is the unstrained CB edge, and A_ϵ is the strain contribution to be expressed later.

Using the $\mathbf{k}\cdot\mathbf{p}$ method for degenerate bands based on the Luttinger-Kohn (LK) model [102], the valence band state can be expanded into the following basis functions to include the coupling among degenerate bands,

$$\begin{aligned}
|1\rangle &= -\frac{1}{\sqrt{2}}|(X+iY)\uparrow\rangle, & |4\rangle &= \frac{1}{\sqrt{2}}|(X-iY)\downarrow\rangle \\
|2\rangle &= \frac{1}{\sqrt{6}}|-(X+iY)\downarrow+2Z\uparrow\rangle, & |5\rangle &= \frac{1}{\sqrt{6}}|(X-iY)\uparrow+2Z\downarrow\rangle \\
|3\rangle &= \frac{1}{\sqrt{3}}|(X+iY)\downarrow+Z\uparrow\rangle, & |6\rangle &= \frac{1}{\sqrt{3}}|(X-iY)\uparrow-Z\downarrow\rangle
\end{aligned} \quad \text{(B.2)}$$

where $|1\rangle = |\frac{3}{2},\frac{3}{2}\rangle$ and $|4\rangle = |\frac{3}{2},-\frac{3}{2}\rangle$ are known as the heavy-hole (HH) states, $|2\rangle = |\frac{3}{2},\frac{1}{2}\rangle$ and $|5\rangle = |\frac{3}{2},-\frac{1}{2}\rangle$ are known as the light-hole (LH) states, and

$|3\rangle = |\frac{1}{2}, \frac{1}{2}\rangle$ and $|6\rangle = |\frac{1}{2}, -\frac{1}{2}\rangle$ are known as the spin-orbit split-off (SOSO) states. Note that here, we assume the valence band has weak coupling with the conduction band. But for narrow-gap materials we have to use an eight-fold basis [105].

We can further block-diagonalize the 6-band LK Hamiltonian using the basis transformation as

$$\begin{aligned}
|u_1\rangle &= \frac{1}{\sqrt{2}}[|1\rangle e^{-i3\phi/2} - i|4\rangle e^{i3\phi/2}], \quad |u_4\rangle = \frac{1}{\sqrt{2}}[|1\rangle e^{-i3\phi/2} + i|4\rangle e^{i3\phi/2}] \\
|u_2\rangle &= \frac{1}{\sqrt{2}}[i|2\rangle e^{-i\phi/2} - |5\rangle e^{i\phi/2}], \quad |u_5\rangle = \frac{1}{\sqrt{2}}[-i|2\rangle e^{-i\phi/2} - |5\rangle e^{i\phi/2}] \\
|u_3\rangle &= \frac{1}{\sqrt{2}}[-i|3\rangle e^{-i\phi/2} - |6\rangle e^{i\phi/2}], \quad |u_6\rangle = \frac{1}{\sqrt{2}}[i|3\rangle e^{-i\phi/2} - |6\rangle e^{i\phi/2}]
\end{aligned} \tag{B.3}$$

The resultant Hamiltonian will become

$$H^{LK}_{6\times6}(k_t) = \begin{bmatrix} H^{U}_{3\times3}(k_t) & 0 \\ 0 & H^{L}_{3\times3}(k_t) \end{bmatrix} \tag{B.4}$$

where

$$\begin{aligned}
H^{U}_{3\times3}(k_t) &= \begin{bmatrix} E_v - P - Q & R_\rho + iS_\rho & \sqrt{2}R_\rho - i\frac{1}{\sqrt{2}}S_\rho \\ R_\rho - iS_\rho & E_v - P + Q & -\sqrt{2}Q - i\sqrt{\frac{3}{2}}S_\rho \\ \sqrt{2}R_\rho + i\frac{1}{\sqrt{2}}S_\rho & \sqrt{2}Q + i\sqrt{\frac{3}{2}}S_\rho & E_v - P - \Delta \end{bmatrix} \\
H^{L}_{3\times3}(k_t) &= \begin{bmatrix} E_v - P - Q & R_\rho - iS_\rho & \sqrt{2}R_\rho + i\frac{1}{\sqrt{2}}S_\rho \\ R_\rho + iS_\rho & E_v - P + Q & -\sqrt{2}Q + i\sqrt{\frac{3}{2}}S_\rho \\ \sqrt{2}R_\rho + i\frac{1}{\sqrt{2}}S_\rho & \sqrt{2}Q - i\sqrt{\frac{3}{2}}S_\rho & E_v - P - \Delta \end{bmatrix}
\end{aligned} \tag{B.5}$$

and

$$
\begin{aligned}
A &= \frac{\hbar^2}{2m_c'}(k_t^2 + k_z^2) + A_\epsilon \\
P &= \frac{\hbar^2}{2m_0}\gamma_1(k_t^2 + k_z^2) + P_\epsilon \\
Q &= \frac{\hbar^2}{2m_0}\gamma_2(k_t^2 - 2k_z^2) + Q_\epsilon \\
R_\rho &= -\frac{\hbar^2}{2m_0}\sqrt{3}(\frac{\gamma_2 + \gamma_3}{2})k_t^2 = -\frac{\hbar^2}{2m_0}\sqrt{3}\bar{\gamma}k_t^2 \\
S_\rho &= \frac{\hbar^2}{2m_0}2\sqrt{3}\gamma_3 k_t^2
\end{aligned}
\tag{B.6}
$$

For QWs, the k_z term in the Hamiltonian is replaced by the operator $-i\partial/\partial z$. E_v is the unstrained valence band edge. The strain is assumed to be biaxial with the substrate and epilayer lattice constants being a_0 and a, respectively. The strain effect is introduced into the Hamiltonian through the Pikus-Bir deformation potentials a_c, a_v, and b [57], and the strain tensor elements C_{ij},

$$
\begin{aligned}
A_\epsilon &= a_c(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}), \\
P_\epsilon &= -a_v(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}), \\
Q_\epsilon &= -\frac{b}{2}(\epsilon_{xx} + \epsilon_{yy} - 2\epsilon_{zz}) \\
\epsilon_{xx} &= \epsilon_{yy} = \frac{a_0 - a}{a},\ \epsilon_{zz} = -\frac{2C_{12}}{C_{11}}\epsilon_{xx}
\end{aligned}
\tag{B.7}
$$

Then the m-th upper-eigenstate and m'-th lower-eigenstate in the VB can be written as

$$
\begin{aligned}
\Psi_m^U(\mathbf{k}_t, \mathbf{r}) &= \frac{e^{i\mathbf{k}_t\cdot\mathbf{r}_t}}{\sqrt{A}}\sum_{i=1}^{3} g_m^{(i)}(k_t, z))|u_i\rangle \\
\Psi_{m'}^L(\mathbf{k}_t, \mathbf{r}) &= \frac{e^{i\mathbf{k}_t\cdot\mathbf{r}_t}}{\sqrt{A}}\sum_{i=4}^{6} g_{m'}^{(i)}(k_t, z))|u_i\rangle
\end{aligned}
\tag{B.8}
$$

A state is named as an HH, LH, or SOSO state if the dominant envelop function is $g^{(1)}$, $g^{(2)}$, or $g^{(3)}$, respectively. Once the envelop functions are solved we can find the squared magnitude of the momentum matrix elements

$|M_{nm}^{\sigma_1\sigma_2}(k_t)|^2 = |\langle\Psi_c^{\sigma_1,n}(\mathbf{k}_t,z)|\hat{e}\cdot\mathbf{p}|\Psi_v^{\sigma_2,m}(\mathbf{k}_t,z)\rangle|^2$ as

$$\begin{aligned}|M_{\mathrm{TE}}|^2 &= |\hat{x}\cdot\langle\Psi_c^{\sigma_1,n}|\mathbf{p}|\Psi_v^{\sigma_2,m}\rangle|^2 = |\hat{y}\cdot\langle\Psi_c^{\sigma_1,n}|\mathbf{p}|\Psi_v^{\sigma_2,m}\rangle|^2 \\ &= \frac{1}{4}\left[3\left|\langle\psi_n|g_m^{(1)}\rangle\right|^2 + \left|\langle\psi_n|g_m^{(2)}\rangle + \sqrt{2}\langle\psi_n|g_m^{(3)}\rangle\right|^2\right] M_b^2 \quad \text{for } \sigma_2 = U \\ &= \frac{1}{4}\left[3\left|\langle\psi_n|g_m^{(4)}\rangle\right|^2 + \left|\langle\psi_n|g_m^{(5)}\rangle + \sqrt{2}\langle\psi_n|g_m^{(6)}\rangle\right|^2\right] M_b^2 \quad \text{for } \sigma_2 = L\end{aligned} \tag{B.9}$$

for TE polarization, and

$$\begin{aligned}|M_{\mathrm{TM}}|^2 &= |\hat{z}\cdot\langle\Psi_c^{\sigma_1,n}|\mathbf{p}|\Psi_v^{\sigma_2,m}\rangle|^2 \\ &= \left|\langle\psi_n|g_m^{(2)}\rangle - \frac{1}{\sqrt{2}}\langle\psi_n|g_m^{(3)}\rangle\right|^2 M_b^2 \quad \text{for } \sigma_2 = U \\ &= \left|\langle\psi_n|g_m^{(5)}\rangle - \frac{1}{\sqrt{2}}\langle\psi_n|g_m^{(6)}\rangle\right|^2 M_b^2 \quad \text{for } \sigma_2 = L\end{aligned} \tag{B.10}$$

for TM polarization. The matrix element M_b^2 can be obtained from the optical matrix energy parameter E_p in many data books [55] as $M_b^2 = |\langle S|p_x|X\rangle|^2/3 = m_0E_p/6$.

In many applications, the SOSO band can also be far from the VB edge and has little contribution. Then we further reduce the formulation to four bands. In this case we will disregard $|3\rangle$ and $|6\rangle$ in Eq. (B.2), and $|u_3\rangle$ and $|u_6\rangle$ in Eq. (B.3). Deleting the third rows and columns in the matrices in Eq. (B.5) gives the 2×2 upper and lower Hamiltonian for coupled HH and LH bands. We will also remove all the momentum matrix terms containing $g^{(3)}$ and $g^{(6)}$ in Eq. (B.8), (B.9), and (B.10).

REFERENCES

[1] *Cisco Global Cloud Index: Forecast and Methodology 2013–2018 White Paper*, Cisco, 2014.

[2] M. Bohr, "Silicon technology leadership for the mobility era," in *Intel Developer Forum*, 2012.

[3] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, no. 4736, pp. 493–494, 1960.

[4] R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. O. Carlson, "Coherent light emission from GaAs junctions," *Physical Review Letters*, vol. 9, no. 9, pp. 366–368, 1962.

[5] M. I. Nathan, W. P. Dumke, G. Burns, F. H. Dill, and G. Lasher, "Stimulated emission of radiation from GaAs *p-n* junctions," *Applied Physics Letters*, vol. 1, no. 3, pp. 62–64, 1962.

[6] N. Holonyak and S. F. Bevacqua, "Coherent (visible) light emission from $Ga(As_{1-x}P_x)$ junctions," *Applied Physics Letters*, vol. 1, no. 4, pp. 82–83, 1962.

[7] T. M. Quist, R. H. Rediker, R. J. Keyes, W. E. Krag, B. Lax, A. L. McWhorter, and H. J. Zeigler, "Semiconductor maser of GaAs," *Applied Physics Letters*, vol. 1, no. 4, pp. 91–92, 1962.

[8] H. Soda, K. Iga, C. Kitahara, and Y. Suematsu, "GaInAsP/InP surface emitting injection lasers," *Japanese Journal of Applied Physics*, vol. 18, no. 12, p. 2329, 1979.

[9] K. Iga, "Surface-emitting laser—its birth and generation of new optoelectronics field," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 6, pp. 1201–1215, 2000.

[10] F. Koyama, S. Kinoshita, and K. Iga, "Room temperature CW operation of GaAs vertical cavity surface emitting laser," *IEICE Transactions*, vol. 71, no. 11, pp. 1089–1090, 1988.

[11] S. L. McCall, A. F. J. Levi, R. E. Slusher, S. J. Pearton, and R. A. Logan, "Whispering-gallery mode microdisk lasers," *Applied Physics Letters*, vol. 60, no. 3, pp. 289–291, 1992.

[12] H.-G. Park, S.-H. Kim, S.-H. Kwon, Y.-G. Ju, J.-K. Yang, J.-H. Baek, S.-B. Kim, and Y.-H. Lee, "Electrically driven single-cell photonic crystal laser," *Science*, vol. 305, no. 5689, pp. 1444–1447, 2004.

[13] M. T. Hill, Y.-S. Oei, B. Smalbrugge, Y. Zhu, T. de Vries, P. J. van Veldhoven, F. W. M. van Otten, T. J. Eijkemans, J. P. Turkiewicz, H. de Waardt, E. J. Geluk, S.-H. Kwon, Y.-H. Lee, R. Notzel, and M. K. Smit, "Lasing in metallic-coated nanocavities," *Nature Photonics*, vol. 1, no. 10, pp. 589–594, 2007.

[14] M. T. Hill, M. Marell, E. S. P. Leong, B. Smalbrugge, Y. Zhu, M. Sun, P. J. van Veldhoven, E. J. Geluk, F. Karouta, Y.-S. Oei, R. Nötzel, C.-Z. Ning, and M. K. Smit, "Lasing in metal-insulator-metal subwavelength plasmonic waveguides," *Optics Express*, vol. 17, no. 13, pp. 11 107–11 112, 2009.

[15] K. Ding, Z. Liu, L. Yin, H. Wang, R. Liu, M. T. Hill, M. J. H. Marell, P. J. van Veldhoven, R. Nötzel, and C. Z. Ning, "Electrical injection, continuous wave operation of subwavelength-metallic-cavity lasers at 260 K," *Applied Physics Letters*, vol. 98, no. 23, p. 231108, 2011.

[16] K. Ding, M. T. Hill, Z. C. Liu, L. J. Yin, P. J. van Veldhoven, and C. Z. Ning, "Record performance of electrical injection sub-wavelength metallic-cavity semiconductor lasers at room temperature," *Optics Express*, vol. 21, no. 4, pp. 4728–4733, 2013.

[17] S.-H. Kwon, J.-H. Kang, C. Seassal, S.-K. Kim, P. Regreny, Y.-H. Lee, C. M. Lieber, and H.-G. Park, "Subwavelength plasmonic lasing from a semiconductor nanodisk with silver nanopan cavity," *Nano Letters*, vol. 10, no. 9, pp. 3679–3683, 2010.

[18] K. Yu, A. Lakhani, and M. C. Wu, "Subwavelength metal-optic semiconductor nanopatch lasers," *Optics Express*, vol. 18, no. 9, pp. 8790–8799, 2010.

[19] M. P. Nezhad, A. Simic, O. Bondarenko, B. Slutsky, A. Mizrahi, L. Feng, V. Lomakin, and Y. Fainman, "Room-temperature subwavelength metallo-dielectric lasers," *Nature Photonics*, vol. 4, no. 6, pp. 395–399, 2010.

[20] Q. Gu, J. S. T. Smalley, M. P. Nezhad, A. Simic, J. H. Lee, M. Katz, O. Bondarenko, B. Slutsky, A. Mizrahi, V. Lomakin, and Y. Fainman, "Subwavelength semiconductor lasers for dense chip-scale integration," *Advances in Optics and Photonics*, vol. 6, no. 1, pp. 1–56, 2014.

[21] C.-Y. Lu, S.-W. Chang, S. L. Chuang, T. D. Germann, and D. Bimberg, "Metal-cavity surface-emitting microlaser at room temperature," *Applied Physics Letters*, vol. 96, no. 25, p. 251101, 2010.

[22] C.-Y. Lu, C.-Y. Ni, M. Zhang, S. L. Chuang, and D. Bimberg, "Metal-cavity surface-emitting microlasers with size reduction: Theory and experiment," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 5, pp. 1–9, 2013.

[23] C.-Y. Lu, S. L. Chuang, D. Bimberg, and F. Hopfer, "Metal-cavity quantum-dot surface-emitting microlaser," in *Conference on Lasers and Electro-Optics (CLEO) 2012*, 2012, p. CW3N.1.

[24] C.-Y. Lu, M. Zhang, S. L. Chuang, E. Stock, and D. Bimberg, "Size dependence of quantum-dot metal-cavity surface-emitting microlaser," in *2012 23rd IEEE International Semiconductor Laser Conference (ISLC)*, 2012, pp. 118–119.

[25] C.-Y. Lu, S. L. Chuang, and D. Bimberg, "Metal-cavity surface-emitting nanolasers," *IEEE Journal of Quantum Electronics*, vol. 49, no. 1, pp. 114–121, 2013.

[26] C.-Y. Lu, P. Qiao, A. Matsudaira, S. L. Chuang, and D. Bimberg, "Metal-cavity quantum-dot surface-emitting microlaser: theory and experiment," in *Photonics Society Summer Topical Meeting Series, 2013 IEEE*, July 2013, pp. 3–4.

[27] P. Qiao, C.-Y. Lu, D. Bimberg, and S. L. Chuang, "Theory and experiment of submonolayer quantum-dot metal-cavity surface-emitting microlasers," *Optics Express*, vol. 21, no. 25, pp. 30 336–30 349, 2013.

[28] P. Qiao, C.-Y. Lu, D. Bimberg, and S. L. Chuang, "Metal-cavity submonolayer quantum-dot surface-emitting microlasers," in *Proceedings of the SPIE*, vol. 8980, 2014, pp. 89 800E–89 800E–9.

[29] C. Chang-Hasnain, "Tunable VCSEL," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 6, pp. 978–987, 2000.

[30] C. Gierl, T. Gruendl, P. Debernardi, K. Zogal, C. Grasse, H. A. Davani, G. Böhm, S. Jatta, F. Küppers, P. Meißner, and M.-C. Amann, "Surface micromachined tunable 1.55 μm-VCSEL with 102 nm continuous single-mode tuning," *Optics Express*, vol. 19, no. 18, pp. 17 336–17 343, 2011.

[31] B. Potsaid, V. Jayaraman, J. G. Fujimoto, J. Jiang, P. J. S. Heim, and A. E. Cable, "MEMS tunable VCSEL light source for ultrahigh speed 60 kHz-1 MHz axial scan rate and long range centimeter class OCT imaging," in *Proceedings of the SPIE*, vol. 8213, 2012, pp. 82 130M–82 130M–8.

[32] M. C. Huang, Y. Zhou, and C. J. Chang-Hasnain, "A surface-emitting laser incorporating a high-index-contrast subwavelength grating," *Nature Photonics*, vol. 1, no. 2, pp. 119–122, 2007.

[33] C. Chase, Y. Rao, W. Hofmann, and C. J. Chang-Hasnain, "1550 nm high contrast grating VCSEL," *Optics Express*, vol. 18, no. 15, pp. 15 461–15 466, 2010.

[34] Y. Rao, W. Yang, C. Chase, M. Huang, D. Worland, S. Khaleghi, M. Chitgarha, M. Ziyadi, A. Willner, and C. Chang-Hasnain, "Long-wavelength VCSEL using high-contrast grating," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 4, p. 1701311, 2013.

[35] C. J. Chang-Hasnain and W. Yang, "High-contrast gratings for integrated optoelectronics," *Advances in Optics and Photonics*, vol. 4, no. 3, pp. 379–440, 2012.

[36] P. Qiao, G.-L. Su, Y. Rao, C. Chang-Hasnain, and S. L. Chuang, "Modeling of long-wavelength high contrast grating VCSELs and comparison with experiment," in *Conference on Lasers and Electro-Optics (CLEO) 2013*, 2013, p. CF2F.5.

[37] P. Qiao, G.-L. Su, Y. Rao, M. C. Wu, C. J. Chang-Hasnain, and S. L. Chuang, "Comprehensive model of 1550 nm MEMS-tunable high-contrast-grating VCSELs," *Optics Express*, vol. 22, no. 7, pp. 8541–8555, 2014.

[38] J. Ferrara, W. Yang, L. Zhu, P. Qiao, and C. J. Chang-Hasnain, "Heterogeneously integrated long-wavelength VCSEL using silicon high contrast grating on an soi substrate," *Optics Express*, vol. 23, no. 3, pp. 2512–2523, 2015.

[39] J. Ferrara, L. Zhu, W. Yang, P. Qiao, and C. J. Chang-Hasnain, "Heterogeneously-integrated VCSEL using high-contrast grating on silicon," in *Proceedings of the SPIE*, vol. 9372, 2015, pp. 937 208–937 208–12.

[40] J. Wang, J.-Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, and A. E. Willner, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nature Photonics*, vol. 6, no. 4, pp. 488–496, 2012.

[41] P. Qiao, L. Zhu, and C. Chang-Hasnain, "Design rule of 2D high contrast gratings and engineering of orbital angular momentum of light," in *Conference on Lasers and Electro-Optics (CLEO) 2015*, 2015, p. SW1I.5.

[42] D. Bimberg, "Quantum dot based nanophotonics and nanoelectronics," *Electronics Letters*, vol. 44, no. 3, pp. 168–171, 2008.

[43] S. S. Mikhrin, A. E. Zhukov, A. R. Kovsh, N. A. Maleev, V. M. Ustinov, Y. M. Shernyakov, I. P. Soshnikov, D. A. Livshits, I. S. Tarasov, D. A. Bedarev, B. V. Volovik, M. V. Maximov, A. F. Tsatsul'nikov, N. N. Ledentsov, P. S. Kop'ev, D. Bimberg, and Z. I. Alferov, "0.94 μm diode lasers based on Stranski-Krastanow and sub-monolayer quantum dots," *Semiconductor Science and Technology*, vol. 15, no. 11, p. 1061, 2000.

[44] F. Hopfer, A. Mutig, G. Fiol, M. Kuntz, V. Shchukin, V. Haisler, T. Warming, E. Stock, S. Mikhrin, I. Krestnikov, D. Livshits, A. Kovsh, C. Bornholdt, A. Lenz, H. Eisele, M. Dahne, N. Ledentsov, and D. Bimberg, "20 Gb/s 85°C error-free operation of VCSELs based on sub-monolayer deposition of quantum dots," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 5, pp. 1302–1308, 2007.

[45] J. Kim and S. L. Chuang, "Theoretical and experimental study of optical gain, refractive index change, and linewidth enhancement factor of p-doped quantum-dot lasers," *IEEE Journal of Quantum Electronics*, vol. 42, no. 9, pp. 942–952, 2006.

[46] T. Baba, T. Hamano, F. Koyama, and K. Iga, "Spontaneous emission factor of a microcavity DBR surface-emitting laser," *IEEE Journal of Quantum Electronics*, vol. 27, no. 6, pp. 1347–1358, 1991.

[47] J. M. Gérard, B. Sermage, B. Gayral, B. Legrand, E. Costard, and V. Thierry-Mieg, "Enhanced spontaneous emission by quantum boxes in a monolithic optical microcavity," *Physical Review Letters*, vol. 81, no. 5, pp. 1110–1113, 1998.

[48] L. Condren and S. Corzine, *Diode Lasers and Photonic Integrated Circuits*. New York: Wiley, 1995.

[49] E. M. Purcell, "Spontaneous emission probabilities at radio frequencies," *Physical Review*, vol. 69, p. 681, 1946.

[50] Y. Yamamoto, S. Machida, and G. Björk, "Microcavity semiconductor laser with enhanced spontaneous emission," *Physical Review A*, vol. 44, no. 1, pp. 657–668, 1991.

[51] J. Gérard and B. Gayral, "InAs quantum dots: artificial atoms for solid-state cavity-quantum electrodynamics," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 9, no. 1, pp. 131–139, 2001.

[52] F. Hopfer, A. Mutig, M. Kuntz, G. Fiol, D. Bimberg, N. N. Ledentsov, V. A. Shchukin, S. S. Mikhrin, D. L. Livshits, I. L. Krestnikov, A. R. Kovsh, N. D. Zakharov, and P. Werner, "Single-mode submonolayer quantum-dot vertical-cavity surface-emitting lasers with high modulation bandwidth," *Applied Physics Letters*, vol. 89, no. 14, p. 141106, 2006.

[53] A. Schliwa, M. Winkelnkemper, and D. Bimberg, "Impact of size, shape, and composition on piezoelectric effects and electronic properties of In(Ga)As/GaAs quantum dots," *Physical Review B*, vol. 76, no. 20, p. 205324, 2007.

[54] G. Bester, X. Wu, D. Vanderbilt, and A. Zunger, "Importance of second-order piezoelectric effects in zinc-blende semiconductors," *Physical Review Letters*, vol. 96, no. 18, p. 187602, 2006.

[55] S. L. Chuang, *Physics of Photonic Devices*, 2nd ed. New York: Wiley, 2009.

[56] C. Pryor, "Eight-band calculations of strained InAs/GaAs quantum dots compared with one-, four-, and six-band approximations," *Physical Review B*, vol. 57, no. 12, pp. 7190–7195, 1998.

[57] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors*. New York: Wiley, 1974.

[58] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III–V compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

[59] Y. P. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149–154, 1967.

[60] S.-W. Chang, C.-Y. A. Ni, and S.-L. Chuang, "Theory for bowtie plasmonic nanolasers," *Optics Express*, vol. 16, no. 14, pp. 10 580–10 595, 2008.

[61] S.-W. Chang and S. L. Chuang, "Fundamental formulation for plasmonic nanolasers," *IEEE Journal of Quantum Electronics*, vol. 45, no. 8, pp. 1014–1023, 2009.

[62] S.-W. Chang, C.-Y. Lu, S. L. Chuang, T. D. Germann, U. W. Pohl, and D. Bimberg, "Theory of metal-cavity surface-emitting microlasers and comparison with experiment," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 6, pp. 1681–1692, 2011.

[63] P. Mena, J. Morikuni, S. Kang, A. Harton, and K. Wyatt, "A comprehensive circuit-level model of vertical-cavity surface-emitting lasers," *Journal of Lightwave Technology*, vol. 17, no. 12, pp. 2612–2632, 1999.

[64] G.-L. Su, P. Qiao, C.-Y. Lu, D. Bimberg, and S. L. Chuang, "Low-threshold dielectric-cavity microlasers," in *Conference on Lasers and Electro-Optics (CLEO) 2014*, 2014, p. SW3G.7.

[65] M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge University Press, 1999.

[66] N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nature Materials*, vol. 13, no. 2, pp. 139–150, 2014.

[67] H. Yang, D. Zhao, S. Chuwongin, J.-H. Seo, W. Yang, Y. Shuai, J. Berggren, M. Hammar, Z. Ma, and W. Zhou, "Transfer-printed stacked nanomembrane lasers on silicon," *Nature Photonics*, vol. 6, no. 9, pp. 615–620, 2012.

[68] B. Zhang, Z. Wang, S. Brodbeck, C. Schneider, M. Kamp, S. Hofling, and H. Deng, "Zero-dimensional polariton laser in a subwavelength grating-based vertical microcavity," *Light: Science and Applications*, vol. 3, p. e135, 2014.

[69] C. F. R. Mateus, M. C. Y. Huang, Y. Deng, A. R. Neureuther, and C. Chang-Hasnain, "Ultrabroadband mirror using low-index cladded subwavelength grating," *IEEE Photonics Technology Letters*, vol. 16, no. 2, pp. 518–520, 2004.

[70] Y. Zhou, M. Moewe, J. Kern, M. C. Huang, and C. J. Chang-Hasnain, "Surface-normal emission of a high-Q resonator using a subwavelength high-contrast grating," *Optics Express*, vol. 16, no. 22, pp. 17 282–17 287, 2008.

[71] I. R. Hooper, B. Tremain, J. A. Dockrey, and A. P. Hibbins, "Massively sub-wavelength guiding of electromagnetic waves," *Scientific Reports*, vol. 4, p. 7495, 2014.

[72] D. Fattal, J. Li, Z. Peng, M. Fiorentino, and R. G. Beausoleil, "Flat dielectric grating reflectors with focusing abilities," *Nature Photonics*, vol. 4, no. 7, pp. 466–470, 2010.

[73] J. H. Mcleod, "The axicon: a new type of optical element," *Journal of the Optical Society of America A*, vol. 44, no. 8, pp. 592–592, 1954.

[74] G. Indebetouw, "Nondiffracting optical fields: some remarks on their analysis and synthesis," *Journal of the Optical Society of America A*, vol. 6, no. 1, pp. 150–152, 1989.

[75] F. Bouchard, I. De Leon, S. A. Schulz, J. Upham, E. Karimi, and R. W. Boyd, "Optical spin-to-orbital angular momentum conversion in ultra-thin metasurfaces with arbitrary topological charges," *Applied Physics Letters*, vol. 105, no. 10, p. 101905, 2014.

[76] S. Vo, D. Fattal, W. Sorin, Z. Peng, T. Tran, M. Fiorentino, and R. Beausoleil, "Sub-wavelength grating lenses with a twist," *IEEE Photonics Technology Letters*, vol. 26, no. 13, pp. 1375–1378, 2014.

[77] J. N. Mait, D. W. Prather, and M. S. Mirotznik, "Design of binary subwavelength diffractive lenses by use of zeroth-order effective-medium theory," *Journal of the Optical Society of America* A, vol. 16, no. 5, pp. 1157–1167, 1999.

[78] M. A. Golub and A. A. Friesem, "Effective grating theory for resonance domain surface-relief diffraction gratings," *Journal of the Optical Society of America A*, vol. 22, no. 6, pp. 1115–1126, 2005.

[79] V. Karagodsky, F. G. Sedgwick, and C. J. Chang-Hasnain, "Theoretical analysis of subwavelength high contrast grating reflectors," *Optics Express*, vol. 18, no. 16, pp. 16 973–16 988, 2010.

[80] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: Wiley-IEEE, 1995.

[81] J. A. Kong, *Electromagnetic Wave Theory*. New York: Wiley, 2000.

[82] G. Valerio, P. Baccarelli, P. Burghignoli, and A. Galli, "Comparative analysis of acceleration techniques for 2-D and 3-D Green's functions in periodic structures along one and two directions," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 6, pp. 1630–1643, 2007.

[83] R. Jorgenson and R. Mittra, "Efficient calculation of the free-space periodic Green's function," *IEEE Transactions on Antennas and Propagation*, vol. 38, no. 5, pp. 633–642, 1990.

[84] R. F. Harrington, *Field Computation by Moment Methods*. New York: Wiley-IEEE press, 1993.

[85] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *Journal of the Optical Society of America A*, vol. 71, no. 7, pp. 811–818, 1981.

[86] S. G. Tikhodeev, A. L. Yablonskii, E. A. Muljarov, N. A. Gippius, and T. Ishihara, "Quasiguided modes and optical properties of photonic crystal slabs," *Physical Review B*, vol. 66, no. 4, p. 045102, 2002.

[87] K. Sakoda, *Optical Properties of Photonic Crystals*, 2nd ed. Springer-Verlag Berlin Heidelberg, 2005.

[88] R. W. Wood, "On a remarkable case of uneven distribution of light in a diffraction grating spectrum," *Proceedings of the Physical Society of London*, vol. 18, no. 1, pp. 269–275, 1902.

[89] L. Rayleigh, "III. Note on the remarkable case of diffraction spectra described by prof. wood," *Philosophical Magazine Series 6*, vol. 14, no. 79, pp. 60–65, 1907.

[90] U. Fano, "The theory of anomalous diffraction gratings and of quasi-stationary waves on metallic surfaces (Sommerfeld's waves)," *Journal of the Optical Society of America*, vol. 31, no. 3, pp. 213–222, 1941.

[91] J. C. Harvey Palmer, "Parallel diffraction grating anomalies," *Journal of the Optical Society of America*, vol. 42, no. 4, pp. 269–273, 1952.

[92] A. Hessel and A. A. Oliner, "A new theory of Wood's anomalies on optical gratings," *Applied Optics*, vol. 4, no. 10, pp. 1275–1297, 1965.

[93] A. Maurel, S. Félix, J.-F. Mercier, A. Ourir, and Z. E. Djeffal, "Wood's anomalies for arrays of dielectric scatterers," *Journal of the European Optical Society Rapid Publications*, vol. 9, p. 14001, 2014.

[94] E. A. J. Marcatili, "Dielectric rectangular waveguide and directional coupler for integrated optics," *Bell System Technical Journal*, vol. 48, no. 7, pp. 2071–2102, 1969.

[95] M.-C. Amann and W. Hofmann, "InP-based long-wavelength VCSEL and VCSEL arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 15, no. 3, pp. 861–868, 2009.

[96] P. A. Martin, "Near-infrared diode laser spectroscopy in chemical process and environmental air monitoring," *Chemical Society Reviews*, vol. 31, no. 4, pp. 201–210, 2002.

[97] D. Babic, K. Streubel, R. Mirin, N. Margalit, J. Bowers, E. Hu, D. Mars, L. Yang, and K. Carey, "Room-temperature continuous-wave operation of 1.54-μm vertical-cavity lasers," *IEEE Photonics Technology Letters*, vol. 7, no. 11, pp. 1225–1227, 1995.

[98] N. Satyan, A. Vasilyev, G. Rakuljic, V. Leyva, and A. Yariv, "Precise control of broadband frequency chirps using optoelectronic feedback," *Optics Express*, vol. 17, no. 18, pp. 15 991–15 999, 2009.

[99] M. Li, W. Yuen, G. Li, and C. Chang-Hasnain, "Top-emitting micromechanical VCSEL with a 31.6-nm tuning range," *IEEE Photonics Technology Letters*, vol. 10, no. 1, pp. 18–20, 1998.

[100] D. Young, J. Scott, F. Peters, M. Peters, M. Majewski, B. Thibeault, S. Corzine, and L. Coldren, "Enhanced performance of offset-gain high-barrier vertical-cavity surface-emitting lasers," *IEEE Journal of Quantum Electronics*, vol. 29, no. 6, pp. 2013–2022, 1993.

[101] E. Kane, "The $\mathbf{k} \cdot \mathbf{p}$ method," in *Physics of III-V Compounds,* Vol. 1 of *Semiconductors and Semimetals*, R. Willardson and A. Beer, Eds. New York: Academic Press, 1966, pp. 75–100.

[102] J. M. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," *Physical Review*, vol. 97, pp. 869–883, 1955.

[103] C.-S. Chang and S. L. Chuang, "Modeling of strained quantum-well lasers with spin-orbit coupling," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 1, no. 2, pp. 218–229, 1995.

[104] J. Minch, S.-H. Park, T. Keating, and S. Chuang, "Theory and experiment of $In_{1-x}Ga_xAs_yP_{1-y}$ and $In_{1-x-y}Ga_xAl_yAs$ long-wavelength strained quantum-well lasers," *IEEE Journal of Quantum Electronics*, vol. 35, no. 5, pp. 771–782, 1999.

[105] P.-F. Qiao, S. Mou, and S. L. Chuang, "Electronic band structures and optical properties of type-II superlattice photodetectors with interfacial effect," *Optics Express*, vol. 20, no. 3, pp. 2319–2334, 2012.

[106] M. Y. Li and C. J. Chang-Hasnain, "Tilt loss in wavelength tunable micromechanical vertical cavity lasers," in *Conference on Lasers and Electro-Optics (CLEO) 1999*, 1999.

[107] J. Scott, D. Young, B. Thibeault, M. Peters, and L. Coldren, "Design of index-guided vertical-cavity lasers for low temperature-sensitivity, sub-milliamp thresholds, and single-mode operation," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 1, no. 2, pp. 638–648, 1995.

[108] J. Piprek and S. Yoo, "Thermal comparison of long-wavelength vertical-cavity surface-emitting laser diodes," *Electronics Letters*, vol. 30, no. 11, pp. 866–868, 1994.

[109] T. Yao, "Thermal properties of AlAs/GaAs superlattices," *Applied Physics Letters*, vol. 51, no. 22, pp. 1798–1800, 1987.

[110] M. Osinski and W. Nakwaski, "Effective thermal conductivity analysis of 1.55 μm InGaAsP/InP vertical-cavity top-surface-emitting microlasers," *Electronics Letters*, vol. 29, no. 11, pp. 1015–1016, 1993.

[111] W. Nakwaski, "Thermal conductivity of binary, ternary, and quaternary III-V compounds," *Journal of Applied Physics*, vol. 64, no. 1, pp. 159–166, 1988.