Title
High Contrast Gratings for Integrated Optoelectronics

Author
Yang, Weijian

Publication Date
2013

Peer reviewed|Thesis/dissertation

High Contrast Gratings for Integrated Optoelectronics

by

Weijian Yang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Engineering-Electrical Engineering and Computer Sciences
and the Designated Emphasis
in
Nanoscale Science and Engineering
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Constance J. Chang-Hasnain, Chair
Professor Ming C. Wu
Professor Xiang Zhang
Professor Eli Yablonovitch

Fall 2013

High Contrast Gratings for Integrated Optoelectronics

by Weijian Yang

Abstract

High Contrast Gratings for Integrated Optoelectronics

by

Weijian Yang

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences

and the Designated Emphasis in Nanoscale Science and Engineering

University of California, Berkeley

Professor Constance J. Chang-Hasnain, Chair

Integrated optoelectronics has seen its rapid development in the past decade. From its original primary application in long-haul optical communications and access network, integrated optoelectronics has expanded itself to data center, consumer electronics, energy harness, environmental sensing, biological and medical imaging, industry manufacture control etc. This revolutionary progress benefits from the advancement in light generation, manipulation, detection and its interaction with other systems. Device innovation is the key in this advancement. Together they build up the component library for integrated optoelectronics, which facilities the system integration.

High contrast grating (HCG) is an emerging element in integrated optoelectronics. Compared to the other elements, HCG has very rich properties and design flexibility. Some of them are fascinating and extraordinary, such as broadband high reflectivity, and high quality factor resonance – all it needs is a single thin-layer of HCG. Furthermore, it can be a microelectromechanical structure. These rich properties are readily to be harnessed and turned into novel devices.

This dissertation is devoted to investigate the physical origins of the extraordinary features of HCG, and explore its applications in novel devices for integrated optoelectronics. An intuitive picture will be presented to explain the HCG physics. The essence of HCG lies in its superb manipulation of light, which can be coupled to applications in light generation and detection. Various device innovations, such as low-loss hollow-core waveguide, fast optical phased array, tunable VCSEL and detector are demonstrated with the HCG as a key element. This breadth of functionality of HCG suggests that HCG has reached beyond a single element in integrated optoelectronics; it has enabled a new platform for integrated optoelectronics.

To my parents

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I thank my advisor Professor Connie Chang-Hasnain for her continuous guidance, inspiration and encouragement throughout my graduate school career. From her, I have not only learnt the philosophy and methods for conducting research, but also the attributes to be a great researcher, i.e. passion, dedication, and optimism. I am particularly grateful for the opportunities that she has brought to me. Without all her support, I would not reached this far.

I also thank Professor Ming Wu, Professor Eli Yablonovitch, and Professor Xiang Zhang for their intellectual guidance in my graduate work, and the review on this dissertation and serving on my qualifying exam committee.

My journey in graduate school has been accompanied by numerous supports from a group of remarkable researchers. I thank Dr. Chris Chase, Dr. Roger Chen, Dr. Jeffrey Chou and Dr. Forrest Sedgwick for their wisdom and philosophy shared with me in research and life. I am particularly grateful to Dr. Frank Yi Rao, Billy Kar Wei Ng, Thai Tran, James Ferrara, and Wai Son Ko. As we work together, we share a lot, not only the joy of our accomplishment, but also the stress and frustration in the difficult times. Without their help and encouragement, I would never have today's achievement. I also thank Dr. Devang Parekh, Dr. Wendy Xiaoxue Zhao and Dr. Erwin Lau for their guidance and training. Meanwhile, I appreciate the collaboration works in the group with Linda Kun Li, Tianbo Sun, Andy Li Zhu, Adair Gerke and Fanglu Lv. I would like to thank Vadim Karagodsky for his development of the HCG theory, which paved the road for my research work. Finally, I thank all the optoelectronics students who have created such a nice research environment.

I am very grateful to have the multiple collaboration work with Professor Ming Wu's group and Professor Alan Willner's group at University of Southern California. I would like to thank Karen Krutter, Anthony Yeh, and Dr. Byung-Wook Yoo from Professor Ming Wu's group; and Dr. Yang Yue from Professor Alan Willner's group. Also my special thanks and appreciation go to Professor David Horsley and his research fellows Dr. Mischa Megens and Dr. Trevor Chan at UC Davis, for their collaboration work on the optical phased array. I also like to express my gratitude and respect to Professor Zhangyuan Chen, Professor Weiwei Hu and Dr. Peng Guo in Peking University, Professor Fumio Koyama at Tokyo Institute of Technology and Dr. Weimin Zhou at US Army Research Laboratory for the fruitful collaborations.

I thank the Berkeley Marvell Nanolab staffs for their help and support in the device fabrication and characterization.

I am also thankful to the financial supports for my graduate studies and research. This includes the Maxine Pao Memorial Fellowship, DARPA iPhod project, DARPA SWEEPER project, and NSF.

Last but not least, to my parents, I am forever indebted for their sacrifice, love, patience, encouragement and support. I could not make this happen without them.

Chapter 1
Introduction

1.1 Introduction to high contrast grating

Optical gratings are among the most fundamental building blocks in optics. They are well understood in two regimes: the diffraction regime, where the grating period (Λ) is greater than the wavelength (λ) [1, 2], and the deep-subwavelength regime, where the grating period is much less than the wavelength [3]. Between these two well-known regimes lies a third, relatively unexplored regime: the near-wavelength regime, where the grating period is between the wavelength of the grating material and that of its surrounding media. At this regime, the grating behaves radically differently and exhibits many distinct features that are not commonly attributed to gratings. These features become more pronounced when there is a large refractive index contrast between the grating bars and their surrounding area. With an intuitive top-down design guidelines, broadband ultra-high reflectivity (>98.5% over a wavelength range of $\Delta\lambda/\lambda$>35%), broadband high transmission windows, 100% reflection and 100% transmission, as well as high quality-factor resonance (quality-factor Q > 10^7) can be obtained [4-9]. This is thus a new class of grating, and is referred as high contrast grating (HCG), schematically shown in Figure 1.1. The high index grating bars are fully surrounded by low index materials -- a key differentiator from other near-wavelength gratings which are etched on a high-index substrate without the additional index contrast at the exiting plane [10-16]. With many extraordinary properties, HCG establishes a new platform for planar optics and integrated optics.

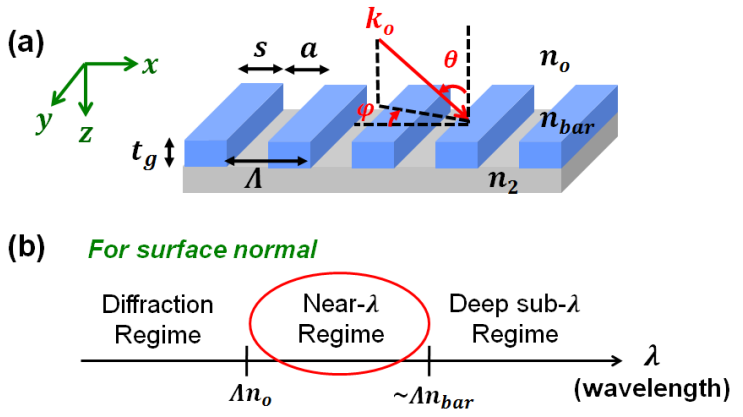


Figure 1.1 (a) Generic HCG structure. The grating comprises of simple dielectric bars with high refractive index n_{bar}, surrounded by a low index medium n_o. A second low-index material n_2 is beneath the bars. (b) The three operation region for gratings. High contrast grating operates at the near-wavelength regime.

A single-layer ultra-thin HCG with broadband high reflectivity for surface-normal incidence was first proposed with numerical simulation [6], followed by experimental demonstration of HCG having reflectivity >98.5% over a wavelength span of $\Delta\lambda/\lambda$>35% [7]. The HCG is subsequently incorporated as the top mirror in vertical cavity surface

emitting lasers (VCSELs) emitting at 850 nm, 980 nm, 1330 nm and 1550 nm wavelength regimes, replacing traditional multi-layer distributed Bragg reflectors (DBR) [17-30]. This HCG top mirror is naturally a microelectromechanical structure (MEMS), and can be electrostatically actuated. Monolithic, continuously tunable HCG-VCSELs have been demonstrated at 850 nm, 1060 nm and 1550 nm [20, 22, 23, 28, 29]. With the same epitaxial structure, the tunable HCG-VCSEL can be operated as a tunable detector [31]. By removing the quantum wells and turning the device into totally passive, all-pass-filters can be constructed as phase tuners [32]. These tunable devices can be further fabricated into arrays on a wafer scale, facilitating large scale photonic integration. By replacing the DBR with HCG, not only the manufacturing cost can be reduced due to a much thinner epitaxial layer, but the MEMS footprint can also be reduced, leading to a more compact structure and higher integration capability. Recently, lasers using two HCG reflectors are realized in a VCSEL structure [30] and in-plane-emitting structure with vertically standing HCGs [33].

Besides replacing DBRs as a highly reflective mirror, HCG has many other special properties and enabled various applications. A single-layer HCG can form a mirrorless high-Q cavity with a Q value as high as 10^7 and emits in surface-normal direction, orthogonal to the periodicity [34, 35]. A single HCG itself has been demonstrated as a surface-emitting laser [36]. In addition, HCG has a rich phase response in reflection and transmission, and this promises applications in phase-engineered optics. For example, a multi-wavelength VCSEL array can be realized by changing the reflection phase of HCG with its dimensions [37]. Furthermore, arbitrary wave-fronts can be obtained by spatially chirping the grating dimensions. This leads to planar, single-layer lens and focusing reflectors with high focusing power [38-40], or light splitter and router [41]. This has created a new field termed as “flat optics”. Other interesting applications of HCG include the transverse mode control in VCSELs, utilizing the angular dependence on HCG reflection [42]; extremely high reflection mirror for high-precision metrology, where the “coating thermal noise” is greatly suppressed [43, 44].

While HCG has many distinct features with light at surface normal incidence, it can be designed to provide reflection and resonances for incident light at an oblique angle as well. One of the applications is hollow-core waveguide, where high reflection mirrors are required for light confinement. By placing two HCGs in parallel [45, 46], or four HCGs in a cage-like scheme, hollow-core waveguides can be made to provide two-dimensional light confinement [45, 47]. On this hollow-core waveguide platform, novel functionalities such as low-loss slow light [47, 48], ultra-compact optical coupler, splitter [49] and switch [50] can be realized with special HCG designs. In addition, an HCG photon cage is proposed for light confinement in the low-index material [51]. Finally, HCG vertical-to-in-plane coupler can be designed to direct light from surface-normal to in-plane index guided waveguide and vice versa [52].

A simple but elegant analytical treatment has been proposed to solve the HCG’s reflection and transmission property, as well as the electromagnetic field distribution in the HCG [9, 53, 54]. The analytical formulation provides an intuition to explain the various peculiar phenomena in HCG, and furthermore, a design guideline of HCG for various applications.

The essence of HCG lies in its superb manipulation of light, which can be coupled to applications in light generation and detection. The breadth of functionality of HCG has covered a whole range of aspects in integrated optoelectronics. This suggests that HCG has reached beyond a single element in integrated optoelectronics; it has enabled a new platform for integrated optoelectronics.

Since its invention in 2004, HCG has seen rapid advances in both experimental and theoretical aspects. It has drawn attentions from the researchers around the world, and gradually formed its own research field. A conference named high contrast metastructures has been established and is devoted to this field in SPIE photonic west since 2012. It is believed that HCG will continue to expand its application and play an important role in integrated optics and optoelectronics.

1.2 Dissertation Overview

Motivated by its potentials in integrated optics and optoelectronics, this dissertation devotes to the development of HCG, both theoretically and experimentally. There are mainly two tasks. The first one is to further develop the theory and physics of HCG, and the second one is to expand the applications of HCG, i.e. design and fabrication of novel devices based on HCG, and characterization and application of these devices in system. Chapter 2 is devoted to study the rich physics of HCG. The analytical formulation of HCG is first studied, for light incidence at an oblique angle, which can be subsequently simplified into the special case of surface normal incidence. This analytical treatment captures the physics of the HCG, and is much more elegant and intuitive than the conventional grating analysis like rigorous coupled wave analysis. This analytical formulation outlines a clear physical picture of HCG, which provides an understanding of the extraordinary features of the HCG. It will be shown that HCG can be easily designed using simple guidelines, rather than requiring heavy numerical simulations through the parameter space. An HCG band diagram is developed to differentiate HCG from photonic crystals (PhCs). The subsequent chapters explore various applications of HCG. Chapter 3 reports a novel hollow-core waveguide using HCG. The hollow-core waveguide using two parallel HCG layers is designed and fabricated. Two dimensional light guiding is experimentally demonstrated though the hollow-core waveguide only has a one-dimension confinement in geometry. Record low propagation loss of 0.37 dB/cm is demonstrated. Another two schemes of hollow-core waveguide would also be discussed. One of them uses DBR as side walls for lateral confinement, while the other uses four HCG layers to construct the waveguide, which is configured as a cage structure. These experimental demonstrations set up a hollow-core waveguide platform for integrated optics, where various functionalities such as slow light, optical switch, gas sensing, etc can be built up. In particular, an optical switch with short switching length is proposed and discussed. Chapter 4 shows an optical phased array using HCG all-pass-filter. The all-pass-filter provides an efficient mechanism for phase tuning. Fast optical beam steering using this phased array is experimentally demonstrated. Chapter 5 reports a 1550 nm tunable HCG detector using a tunable HCG VCSEL. The tunable detector is shown to select a wavelength channel over a large wavelength span while effectively reject the other channels. The device is essentially bi-functional, which can operate as a

VCSEL or detector. A high speed link between two such devices is demonstrated in a single fiber, where one device is functioned as a transmitter and the other as a receiver, or verse vice. Based on this tunable HCG detector, a novel configuration of the device as a tracking detector, and a novel chip-scale optical spectrometer would also be presented. Chapter 6 summaries the dissertation and discuss the outlook and future work for the further development of HCG.

Chapter 2
Physics of High Contrast Grating

To understand near-wavelength gratings, the electromagnetic field profile inside the grating can neither be approximated (as in the deep-subwavelength regime) nor ignored (as in the diffraction regime). Fully rigorous electromagnetic solutions exist for gratings [55-57]; however, they tend to involve heavy mathematical formulation. Recently, our group published a simple analytic formulation to explain the broadband reflection and resonance at surface normal incidence [53, 54, 58]. Independently, Lalanne *et. al.* published a quasi-analytical method using coupled Bloch modes [59] with rather similar formulation. In this chapter, we generalize the formulation to oblique incident angle, and provide a complete and in-depth discussion of the rich physics of HCG.

Figure 1.1 shows the schematic of a generic HCG, with air as the low index medium on top and between the grating bars, a second low-index material beneath the bars and an incident plane wave at an oblique angle. The HCG is polarization sensitive by its nature of 1D periodicity. Incident beam with E-field polarization along and perpendicular to the grating bars are referred to as transverse electric (TE) and transverse magnetic (TM) polarizations, respectively. The two incident characteristic angles, θ and φ, are the incident direction between the incident beam and the *y*-*z* and *x*-*z* plane, respectively. Two special conditions are $\varphi = 0$ and $\varphi = 90^o$, where the light propagation direction is perpendicular and parallel to the grating bars respectively. We term the former case as transverse HCG [Figure 2.1 (a)], and the latter, longitudinal HCG [Figure 2.1 (b)]. The HCG input plane is defined as the plane $z = 0$, whereas the HCG exit plane is defined as the plane $z = t_g$.

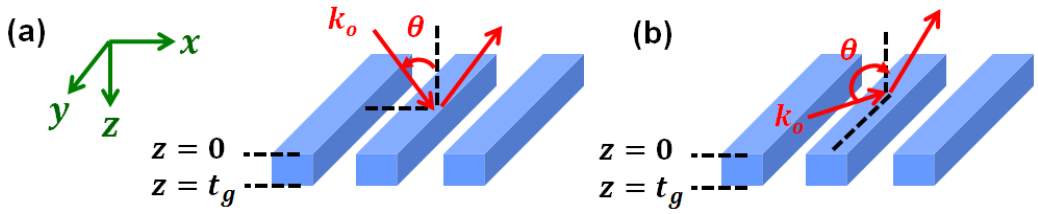


Figure 2.1 (a) Transverse HCG where φ=0. (b) Longitudinal HCG where φ=90°.

To understand the basic principle, we simplify the structure by considering the second low-index material and the substrate as air. We first focus our discussion on the special operation condition of $\varphi = 0$ and $\varphi = 90^o$. Later we generalize the analytical solution to multiple layer / HCG structures with a transmission matrix method. The main design parameters include the refraction index of the grating bars n_{bar}, grating period Λ, grating thickness t_g, grating bar width s, the incident angle θ, and the operation wavelength λ. The grating duty cycle η is defined as the ratio between the width of the high index material and period, i.e. $\eta = s/\Lambda$.

In section 2.1, we first outline the underlying physics and present the distinct features of HCG. The full analytical formulation is presented in section 2.2. In section 2.3,

the HCG supermodes and their interference are discussed, which can lead to a physical insight of the extraordinary properties of HCG. An HCG band diagram is developed in section 2.4; clear differentiation between HCG and 1D PhC is seen. In section 2.5, we briefly discuss the analytical treatment for longitudinal HCG. Section 2.6 provides a summary and closing remark for this chapter.

2.1 Overview of the Underlying Principles

The grating bars can be considered as merely a periodic array of waveguides with wave being guided along z direction. Upon plane wave incidence, depending on wavelength and grating dimensions, only a few waveguide-array modes are excited. Due to a large index contrast and near-wavelength dimensions, there exists a wide wavelength range where only two waveguide-array modes have real propagation constants in the z direction and, hence, carry energy. This is the regime of interests, referred as the dual-mode regime.

The two waveguide-array modes then depart from the grating input plane ($z = 0$) and propagate downward ($+z$ direction) to the grating exit plane ($z = t_g$), and then reflect back up. The higher order modes are typically below cutoff and in the form of evanescent surface-bound waves.

After propagating through the HCG thickness, each propagating waveguide-array mode accumulates a different phase. At the exit plane, due to a strong mismatch to the exit plane wave, the waveguide-array modes not only reflect back to themselves but also couple into each other. Similar mode coupling occurs at the input plane when the modes make one round trip and return to the input plane. Following the waveguide-array modes through one round trip, the reflectivity solution can be attained. At both input and exit planes, the waveguide-array modes also transmit out to the low index media, or air in this case. Due to HCG's near-wavelength period in air, only the 0^{th} diffraction order carries energy in reflection and transmission, which are plane waves. This is the most critical factor contributing to the extraordinary HCG properties.

The HCG thickness determines the phase accumulated by the waveguide-array modes and controls their interference at the HCG input plane and exit plane, making HCG thickness one of the most important design parameter. To obtain high reflection, the HCG thickness should be chosen such that a destructive interference is obtained at the exit plane, which cancels the transmission. For full transmission, on the other hand, the thickness should be chosen such that the interference is well matched with the input plane wave at the input plane; or in analogy, the "impedance" of HCG matches with that of the input space. Finally, when a constructive interference is obtained at both input and exit planes, a high-Q resonator results.

Here, destructive interference does not mean that the fields are zero everywhere. Rather, it means that the spatial mode-overlap with the transmitted plane wave is 0, yielding a zero transmission coefficient. This prevents optical power from being launched into a transmitting propagating wave, and thus causes full reflection. A matched interference at the input plane means the spatial mode-overlap with the input plane wave

is 1, and thus vanishes any reflection. For a high-Q resonator, each HCG waveguide-array mode couples strongly to each other and becomes self-sustaining after each round trip. The resonator can thus reach high Q without conventional mirrors – another unique and distinct feature of HCG.

The three main phenomena observed in HCG are shown in Figure 2.2. The reflectivity spectrum of *resonant gratings* (in red) exhibits several high-Q resonances, characterized by very sharp transitions from 0 to ~100% reflectivity and vice versa, e.g. 1.682 μm and 1.773 μm in Figure 2.2 (a) marked by two arrows. Figure 2.2 (b) shows reflectivity spectrum of *broadband reflector* (in blue). Broadband transmission (in green) can be obtained with different set of HCG parameters, also shown in Figure 2.2 (b). For the broad-band high reflection case, the 99% reflection bandwidth is 578 nm (from 1.344 μm to 1.922 μm); this corresponds to $\Delta\lambda/\lambda$~35%. For the broad-band high transmission case, the transmission is larger than 99.8% over a broad spectrum, not limited from 1.3 μm to 2 μm shown in the figure. The HCG parameters for the three different cases are: high-Q resonances, Λ=0.716 μm, t_g=1.494 μm, η=0.70, TE polarization light; broad-band high reflection, Λ=0.77 μm, t_g=0.455 μm, η=0.76, TM polarization light; broad-band high transmission, Λ=0.8 μm, t_g=0.6 μm, η=0.1, TM polarization light. θ=0 and n_{bar}=3.48 for all three cases.

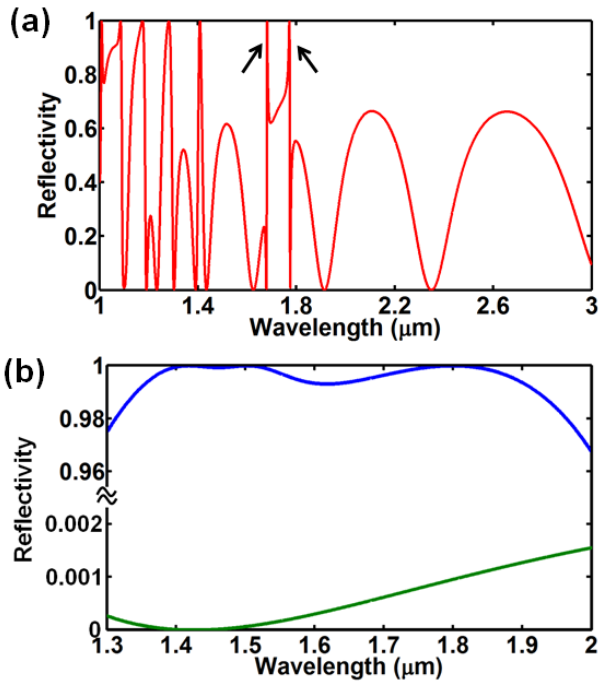


Figure 2.2 Examples of three types of extraordinary reflectivity / transmission features of HCG. (a) High-Q resonances (red). (b) Broad-band high reflection (blue), and broad-band high transmission (green).

From the above analysis, the HCG can be treated as a Fabry-Perot cavity, which is composed of the waveguide array with the HCG thickness as the cavity length. The input plane and the exit plane interface the HCG and the outside media, serving as two mirrors. Waveguide-array modes are supported by the cavity. These modes interact and interfere

with each other at the input and exit plane, and give rise to the rich characteristics of the HCG. To better understand the HCG as a cavity, we plot the reflectivity contour map versus normalized wavelength and grating thickness (both by HCG period Λ), for surface normal condition, transverse HCG with light incident angle of 50^o, and longitudinal HCG with light incident angle of 30^o, for both TE and TM polarization, shown in Figure 2.3. The HCG conditions are labeled in the figures. The HCG duty cycles are 0.45, 0.75, 0.45, 0.6, 0.5 and 0.75 for Figure 2.3 (a)~(f); $n_{bar} = 3.48$ for all the cases.

First of all, a fascinatingly well-behaved, highly ordered, checker-board pattern is observed for ***all*** six HCG cases. The similar and strong dependence on both wavelength and HCG thickness supports the interference effect discussed above. We further note that half of the "checker-board" have high reflectivity (deep red contour), while the other half have lower reflectivity (light red contour). This interesting phenomenon relates to the interference of the two waveguide-array modes at the HCG input/exit planes. This is also strong evidence that HCG effect is not merely a grating or photonic crystal effect since the longitudinal HCGs, with beam propagating in the direction orthogonal to the direction of periodicity, also exhibit similar behavior as the transverse ones.

The checker board pattern in Figure 2.3 is gridded by the resonance curves of the HCG supermodes. The HCG supermodes are a summation of the waveguide-array modes. They are the Eigen modes of the HCG. It should be note that though the waveguide-array modes are the Eigen modes of the waveguide array, they are not the Eigen modes of the HCG; they exchange energy and couple to the plane waves at HCG input and exit plane. The Eigen modes of the HCG can be expressed as a summation of these waveguide-array modes, termed as supermodes. The resonance curves in Figure 2.3 indicate where the round trip phase of these HCG supermodes reaches an integer number of 2π. Within one family of resonance curves, the different curves come from different integer multiple of 2π, rather analogous to longitudinal modes in a traditional Fabry-Perot cavity. The different families of resonance curves correspond to different orders of supermodes, analogous to transverse modes in the Fabry-Perot cavity. At longer wavelength, only the fundamental mode exists, and HCG operates in the deep-subwavelength regime, behaving like a quasi-uniform layer. The reflectivity contour is governed by a simple Fabry-Perot mechanism, which is recognizable by the (quasi) linear bands in Figure 2.3. The reflectivity in this regime, however, never gets to be very high.

The dual-mode regime locates between the cutoff wavelengths of the first-order waveguide-array mode and the second-order waveguide-array mode. This is where the most pronounced checker-board pattern appears. For wavelengths shorter than the cutoff wavelength of the second-order waveguide-array mode, more and more modes emerge, and the reflectivity contour shapes become less and less ordered. The diffraction region locates at $\lambda < (1 + sin\theta)\Lambda$, where higher diffraction orders emerge outside the grating and reflectivity is significantly reduced. The three waveguide regimes: deep-subwavelength, near-wavelength and diffraction are labeled in Figure 2.3. For the longitudinal HCG, the dual-mode regime is not well defined, as the cutoff wavelength for the first two high order modes are very close. In the section 2.2, we detail the analytical treatment of the HCG so as to better understand these extraordinary features of HCG.

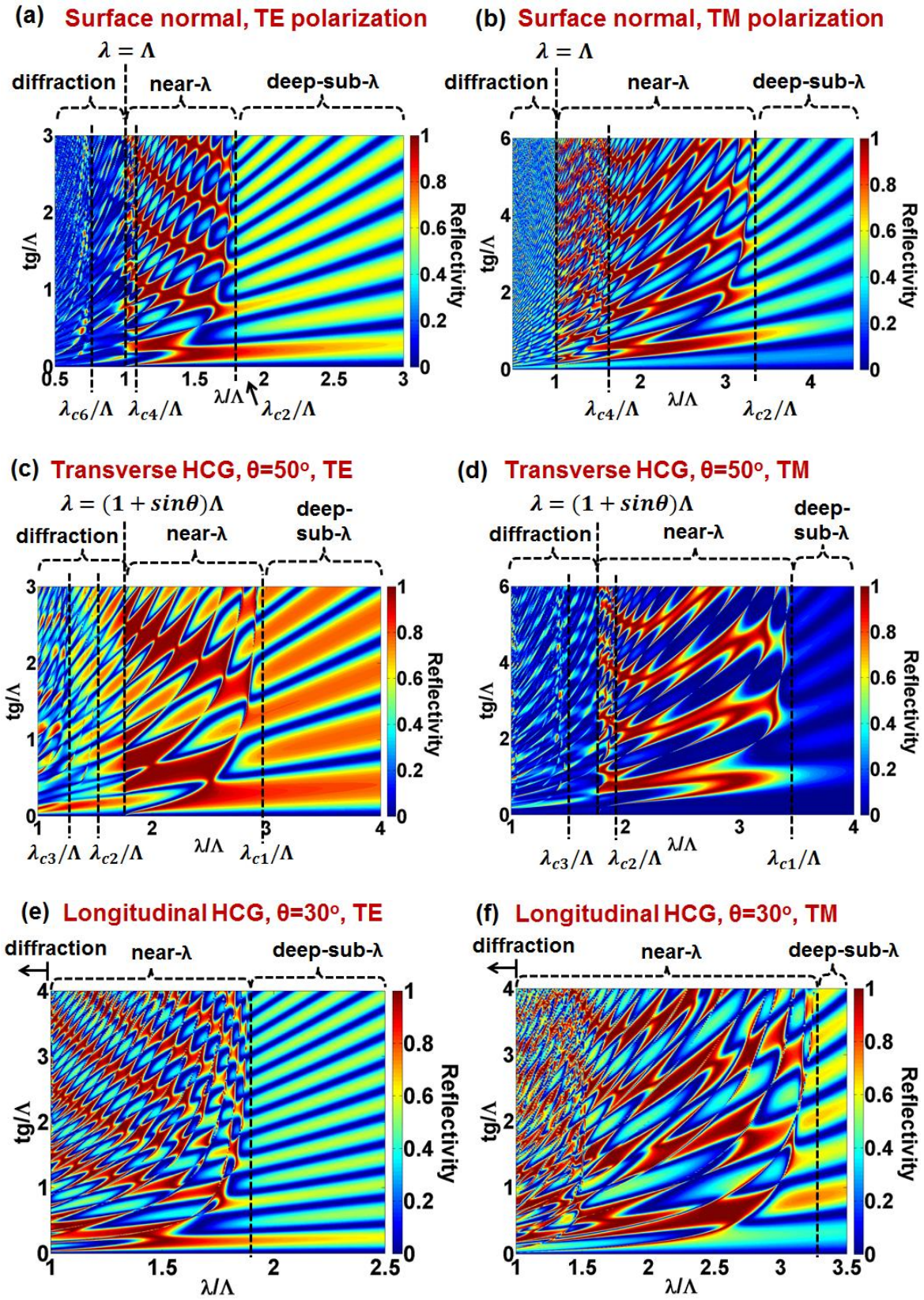

(a) Surface normal, TE polarization
(b) Surface normal, TM polarization
(c) Transverse HCG, θ=50°, TE
(d) Transverse HCG, θ=50°, TM
(e) Longitudinal HCG, θ=30°, TE
(f) Longitudinal HCG, θ=30°, TM
diffraction
near-λ
deep-sub-λ
deep-sub-λ
Reflectivity
t/g/λ
λ/Λ

Figure 2.3 Reflectivity contour of HCGs as a function of wavelength and grating thickness. A fascinatingly well-behaved, highly ordered, checker-board pattern reveals its strong property dependence on both wavelength and HCG thickness, which indicates an interference effect. Surface normal incidence, oblique incidence for transverse HCG and oblique incidence for longitudinal HCG are shown.

2.2 Analytical Formulation

In this section, we discuss the analytical treatment for HCG [9, 53, 54, 58]. While we focus on the transverse HCG, the methodology can be extended to longitudinal HCG. A transmission matrix method is also discussed to solve multiple layer HCG structures. The analytical solutions described here are in excellent agreement with numerical simulations using rigorous couple wave analysis (RCWA) and finite-difference time-domain method (FDTD). The analytical formulation not only facilitates a more intuitive understanding but also yields much faster computation speeds.

2.2.1 TM-polarized Incidence

We first focus on TM-polarized plane wave as incident light. The HCG is assumed to be infinite in y and infinitely periodic in x. The solution is thus two-dimensional ($\partial/\partial y = 0$). We consider three regions, separated by the HCG input plane $z = 0$ and exit plane $z = t_g$, as shown in Figure 2.4. In region I, $z < 0$, there are incoming plane wave and reflected waves. In region III, $z > t_g$, there exist only the transmitted waves. The incident angle of the incoming light is θ, corresponding to an x-wavenumber of k_{x0}. Due to the periodic nature of the grating, the reflection and transmission wave can be decomposed into plane waves with different diffraction order n; their x-wavenumber can be expressed as $k_{x0} + 2\pi n/\Lambda$, where $n = 0, \pm 1, \pm 2$ ... ... The z-wavenumber of the nth order mode is γ_n. In the near-wavelength and deep subwavelength regimes, only the 0th order diffraction is propagation wave (i.e. γ_0 is a real number), while others are evanescent waves (i.e. γ_n is an imaginary number). In region II, $0 < z < t_g$, the solutions are modes of a periodic array of 1D waveguides (i.e. 1D confinement in x direction and no confinement in y direction, also known as slab waveguides). k_a and k_s are the x-wavenumbers in the air-gaps and in the grating bars respectively for the waveguide-array mode. The z-wavenumber β is the same in both the air-gaps and the bars. For simplicity, all low-index media have refractive index of 1. We keep track of the field in the x, y direction, from which the field in the z direction can be easily derived with Maxwell equations.

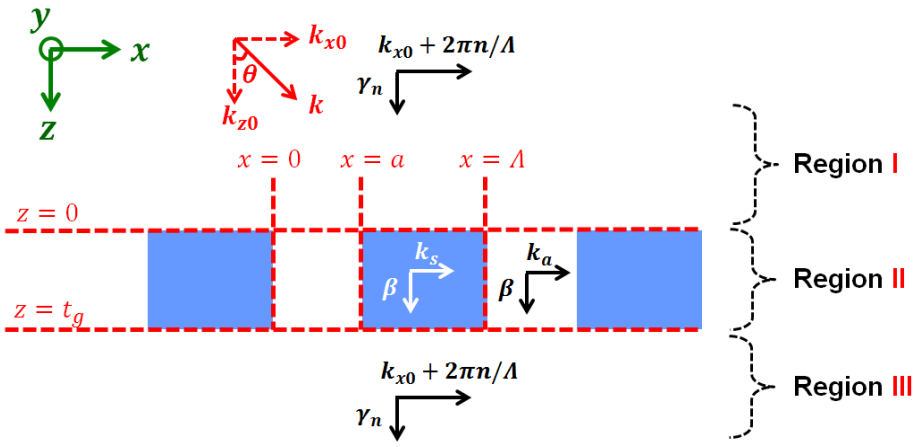


Figure 2.4 Nomenclature for Eqs. 2.1a-2.1d.

The magnetic and electric fields in Region I, $H_y^I(x,z)$ and $E_x^I(x,z)$, respectively, are expressed as the sum of the incident plane wave and multiple orders of reflected diffraction waves, whose magnetic field, electric field, reflection coefficient, and propagation constant are $\mathcal{H}_{y,n}^I(x)$, $\mathcal{E}_{x,n}^I(x)$, r_n, and γ_n, respectively, where $n = 0, \pm1, \pm2, \ldots$ is the reflected mode number.

$$H_y^I(x,z) = e^{-jk_0 cos\theta z - jk_0 sin\theta x} - \sum_{n=-\infty}^{+\infty} r_n \mathcal{H}_{y,n}^I(x) e^{+j\gamma_n z} \qquad z < 0$$

$$E_x^I(x,z) = \gamma_0/k_0 \sqrt{\frac{\mu_0}{\varepsilon_0}}\, e^{-jk_0 cos\theta z - jk_0 sin\theta x} + \sum_{n=-\infty}^{+\infty} r_n \mathcal{E}_{x,n}^I(x) e^{+j\gamma_n z} \qquad z < 0 \qquad (2.1a)$$

$$\mathcal{H}_{y,n}^I(x) = e^{-jk_{xn}x}$$

$$\mathcal{E}_{x,n}^I(x) = \gamma_n/k_0 \sqrt{\mu_0/\varepsilon_0}\, e^{-jk_{xn}x} \qquad (2.1b)$$

$$k_0 = \frac{2\pi}{\lambda}$$

$$k_{xn} = k_0 sin\theta + \frac{2n\pi}{\Lambda}$$

$$\gamma_n^2 = k_0^2 - {k_{x0}}^2 - \left(\frac{2n\pi}{\Lambda}\right)^2 \qquad (2.1c)$$

μ_0 and ε_0 are the vacuum permeability and vacuum permittivity respectively. As evident from Eq. 2.1c, when $\lambda > (1 + sin\theta)\Lambda$, only the 0^{th} diffraction order is propagating (γ_0 is real), while the higher orders are all evanescent (γ_1, γ_2 etc. are imaginary). This regime, $\lambda > (1 + sin\theta)\Lambda$, can be viewed as a general definition of subwavelength regime for the input (and output) region.

The transmitted mode profiles for Region III can be similarly expressed by Eq. 2.1d:

$$H_y^{III}(x,z) = \sum_{n=-\infty}^{+\infty} \tau_n \mathcal{H}_{y,n}^{III}(x) e^{-j\gamma_n (z-t_g)} \qquad z > t_g$$

$$E_x^{III}(x,z) = \sum_{n=-\infty}^{+\infty} \tau_n \mathcal{E}_{x,n}^{III}(x) e^{-j\gamma_n (z-t_g)} \qquad z > t_g \qquad (2.1d)$$

Here $\mathcal{H}_{y,n}^{III}(x) = \mathcal{H}_{y,n}^{I}(x)$ $\mathcal{E}_{x,n}^{III}(x) = \mathcal{E}_{x,n}^{I}(x)$

The reflection and transmission can be expressed in vector form, $\boldsymbol{r}$ and $\boldsymbol{\tau}$, respectively, in terms of all orders of diffractive waves, r_n and τ_n, $n \in (-\infty, \infty)$.

$$\boldsymbol{r} \equiv (\ldots\ r_{-2}\ \ r_{-1}\ \ r_0\ \ r_1\ \ r_2\ \ldots)^T \qquad (2.1e)$$

$$\boldsymbol{\tau} \equiv (\ldots\ \tau_{-2}\ \ \tau_{-1}\ \tau_0\ \ \tau_1\ \ \tau_2\ \ldots)^T \qquad (2.1f)$$

The incident wave in Region I can also be expressed as $(\ldots 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ldots)^T$, where only the 0th order is non-zero. The HCG reflectivity matrix **R** and transmission matrix **T** are thus found below.

$$\boldsymbol{r} \equiv \mathbf{R}(\ldots 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ldots)^T \qquad (2.1g)$$

$$\boldsymbol{\tau} \equiv \mathbf{T}(\ldots 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ldots)^T \qquad (2.1h)$$

The magnetic field $H_y^{II}(x,z)$ and electric field $E_x^{II}(x,z)$ in Region II comprises of waveguide-array modes (with mode index $m = 0, 1, 2, \ldots$) with lateral (x) magnetic and electric field of $\mathcal{H}_{y,m}^{II}(x)$ and $\mathcal{E}_{x,m}^{II}(x)$, respectively, and propagating constants of β_m in longitudinal direction (z). The forward ($+z$) and backward ($-z$) propagating components have coefficients of a_m and b_m respectively.

$$H_y^{II}(x,z) = \sum_{m=0}^{\infty} \mathcal{H}_{y,m}^{II}(x)\left[a_m e^{-j\beta_m (z-t_g)} - b_m e^{+j\beta_m (z-t_g)}\right] \qquad 0 < z < t_g$$

$$E_x^{II}(x,z) = \sum_{m=0}^{\infty} \mathcal{E}_{x,m}^{II}(x)\left[a_m e^{-j\beta_m (z-t_g)} + b_m e^{+j\beta_m (z-t_g)}\right] \qquad 0 < z < t_g \qquad (2.2a)$$

$$\mathcal{H}_{y,m}^{II}(x) = A_m cos\left[k_{a,m}(x - a/2)\right] + B_m sin\left[k_{a,m}(x - a/2)\right]; \qquad 0 < x < a$$

$$\mathcal{H}_{y,m}^{II}(x) = C_m cos\left\{k_{s,m}[x - (a+\Lambda)/2]\right\} + D_m sin\left\{k_{s,m}[x - (a+\Lambda)/2]\right\}; \quad a < x < \Lambda$$

$$\mathcal{E}_{x,m}^{II}(x) = (\beta_m / k_0)\sqrt{\mu_0/\varepsilon_0}\,\mathcal{H}_{y,m}^{II}(x) \qquad 0 < x < a$$

$$\mathcal{E}_{x,m}^{II}(x) = n_{bar}^{-2}(\beta_m / k_0)\sqrt{\mu_0/\varepsilon_0}\,\mathcal{H}_{y,m}^{II}(x) \qquad a < x < \Lambda \qquad (2.2b)$$

Here, k_a and k_s are the lateral wavenumbers inside air-gaps and grating bars, respectively. Eq. 2.2b is a general solution including both sine functions (odd modes) and cosine functions (even modes). At surface normal incidence where θ=0, there would be only even modes, since the odd modes cannot be excited by the surface normal incident plane waves in Region I.

By matching the boundary conditions at $x = a$ [$\mathcal{H}_{y,m}^{II}(x)$ and $\mathcal{E}_{z,m}^{II}(x)$ being continuous] and the following Bloch boundary conditions,

$$\mathcal{H}_{y,m}^{II}(x + l\Lambda) = \mathcal{H}_{y,m}^{II}(x)e^{-ik_{x0}l\Lambda}, \quad l \text{ is an integer}$$

$$\mathcal{E}_{x,m}^{II}(x + l\Lambda) = \mathcal{E}_{x,m}^{II}(x)e^{-ik_{x0}l\Lambda}, \quad l \text{ is an integer} \qquad (2.2c)$$

one can solve *A*, *B*, *C* and *D* into the following,

$$\begin{aligned} A_m = {} & 1/n_{bar}^2 k_{s,m} \sin(k_{a,m}a/2) + k_{a,m} \cos(k_{a,m}a/2) \sin(k_{s,m}s) e^{-ik_{x0}\Lambda} \\ & + 1/n_{bar}^2 k_{s,m} \sin(k_{a,m}a/2) \cos(k_{s,m}s) e^{-ik_{x0}\Lambda} \end{aligned}$$

$$\begin{aligned} B_m = {} & 1/n_{bar}^2 k_{s,m} \cos(k_{a,m}a/2) + k_{a,m} \sin(k_{a,m}a/2) \sin(k_{s,m}s) e^{-ik_{x0}\Lambda} \\ & - 1/n_{bar}^2 k_{s,m} \cos(k_{a,m}a/2) \cos(k_{s,m}s) e^{-ik_{x0}\Lambda} \end{aligned}$$

$$\begin{aligned} C_m = {} & k_{a,m} \sin(k_{s,m}s/2) \left[e^{-ik_{x0}\Lambda} + 2\cos^2(k_{a,m}a/2) - 1\right] \\ & + 2/n_{bar}^2 \, k_{s,m} \cos(k_{a,m}a/2) \sin(k_{a,m}a/2) \cos(k_{s,m}s/2) \end{aligned}$$

$$\begin{aligned} D_m = {} & -k_{a,m} \cos(k_{s,m}s/2) \left[e^{-ik_{x0}\Lambda} - 2\cos^2(k_{a,m}a/2) + 1\right] \\ & - 2/n_{bar}^2 \, k_{s,m} \cos(k_{a,m}a/2) \sin(k_{a,m}a/2) \sin(k_{s,m}s/2) \end{aligned} \qquad (2.2d)$$

The characteristic equation between the lateral wavenumbers $k_{a,m}$ and $k_{s,m}$ can be written as:

$$\begin{aligned} & n_{bar}^{-2} k_{a,m} k_{s,m} \left[1 + e^{-2ik_{x0}\Lambda} - 2\cos(k_{a,m}a) \cos(k_{s,m}s) e^{-ik_{x0}\Lambda}\right] \\ & + \sin(k_{a,m}a) \sin(k_{s,m}s) \left[k_{a,m}^2 + (n_{bar}^{-2})^2 k_{s,m}^2\right] e^{-ik_{x0}\Lambda} = 0 \end{aligned} \qquad (2.2e)$$

The longitudinal wave number β_m in Eq. 2.2a is given by:

$$\beta_m^2 = (2\pi/\lambda)^2 - k_{a,m}^2 = (2\pi n_{bar}/\lambda)^2 - k_{s,m}^2 \qquad (2.2f)$$

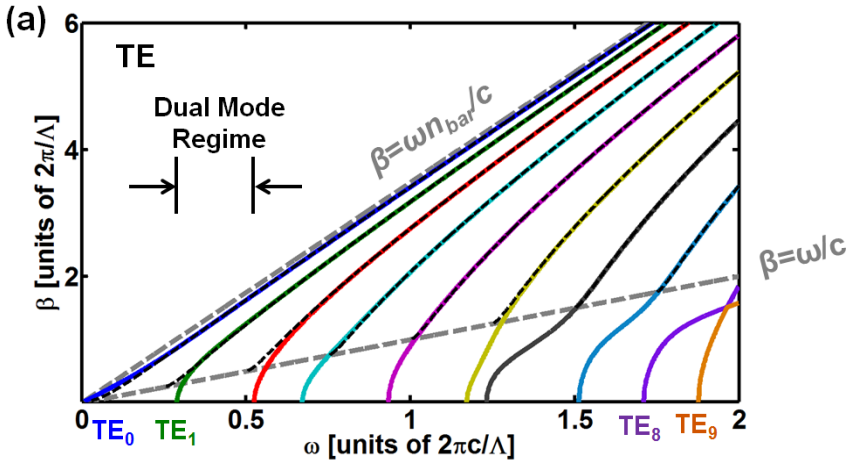
Thus

$$k_{s,m}^2 - k_{a,m}^2 = (2\pi/\lambda)^2 (n_{bar}^2 - 1) \qquad (2.2g)$$

The $\omega - \beta$ relations ($\omega = 2\pi c/\lambda$) resulting from Eq. 2.2e-g are essentially waveguide array $\omega - \beta$ diagrams, which are independent of HCG thickness t_g. They are shown in Figure 2.5 for both TE and TM polarization with a light incident angle of 50^o, compared with that of a single slab waveguide with the same waveguide width s. Here, we see several trends which are typical to waveguides: (1) as frequency increases, more and more modes exist; (2) except for the fundamental mode, there is a cut-off frequency (denoted by ω_{c1}, ω_{c2}, ω_{c3}, etc.) for each higher order mode (cutoff condition being $\beta = 0$); and (3) the higher the mode order, the higher the cut-off frequency (the shorter the cut-off wavelength). Similar to the notation in the single slab waveguide, we denote the waveguide array modes as TE_0, TE_1, TE_2, etc. for TE light; and TM_0, TM_1, TM_2, etc. for TM light. The even order modes have symmetric mode profiles whereas the odd order modes have asymmetric mode profiles. The cutoff frequencies in Figure 2.5 are important, since they determine the regime in which the grating operates. As will be shown

subsequently, the extraordinary features discussed in Section 2.1 are observed primarily in the regime where exactly two modes exist, hereinafter the "dual-mode" regime (e.g. TE_0 and TE_1, or TM_0 and TM_1). An example of the mode profile of the waveguide array mode for both TE $[E_y(x)]$ and TM $[H_y(x)]$ conditions are shown in Figure 2.6. The derivative of $H_y(x)$ over x is not continuous at $x = a \; and \; x = \Lambda$, simply due to the refractive index difference between the grating bars and air.

The dispersion curves of the waveguide array between the air light line, $\beta = \omega/c$, and the dielectric light line, $\beta = \omega n_{bar}/c$ resemble those of a single slab waveguide. They start to deviate from each other when ω approaches the cutoff frequency of the single slab waveguide, i.e. the air light line. This can be understood as the following: well above light line, the mode is tightly confined in each individual high index waveguides, and the coupling between adjacent waveguides is weak. Thus the ω-β diagrams of the waveguide array overlaps well with that of the single waveguide. This also indicates that the ω-β diagrams of the waveguide array are independent of the incident angle in this regime. Approaching the air light line, the mode is less confined in the high index waveguide, and coupling between adjacent waveguide becomes strong, leading to a deviation of its ω-β diagram from the single slab waveguide. Below the air light line lays the main difference between the waveguide arrays and single slab waveguide. Waveguide arrays have a discrete set of modes for $\beta<\omega/c$, in contrast to the continuum of radiation modes observed for a single slab waveguide. This discretization is a direct consequence of the grating periodicity.



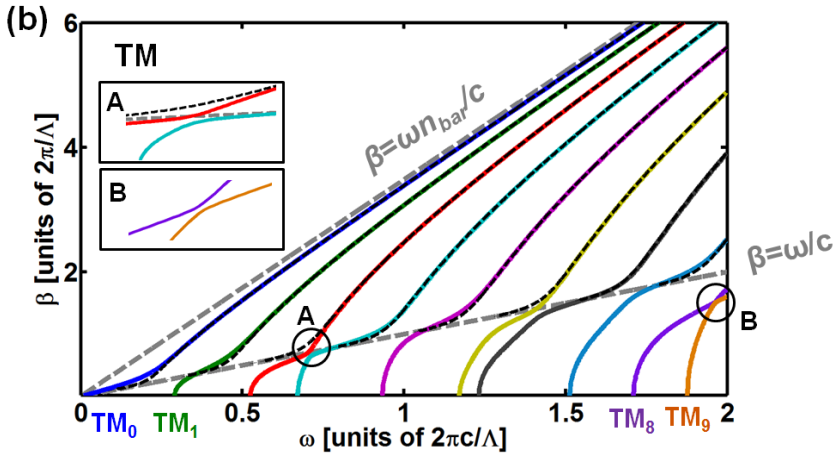


Figure 2.5 Dispersion curves (ω-β diagram) of a waveguide array (solid) and a single slab waveguide (dash), for the same bar width s and index n_{bar}, β being the z-wavenumbers. In this calculation, η=0.6, n_{bar}=3.48, θ=50°.

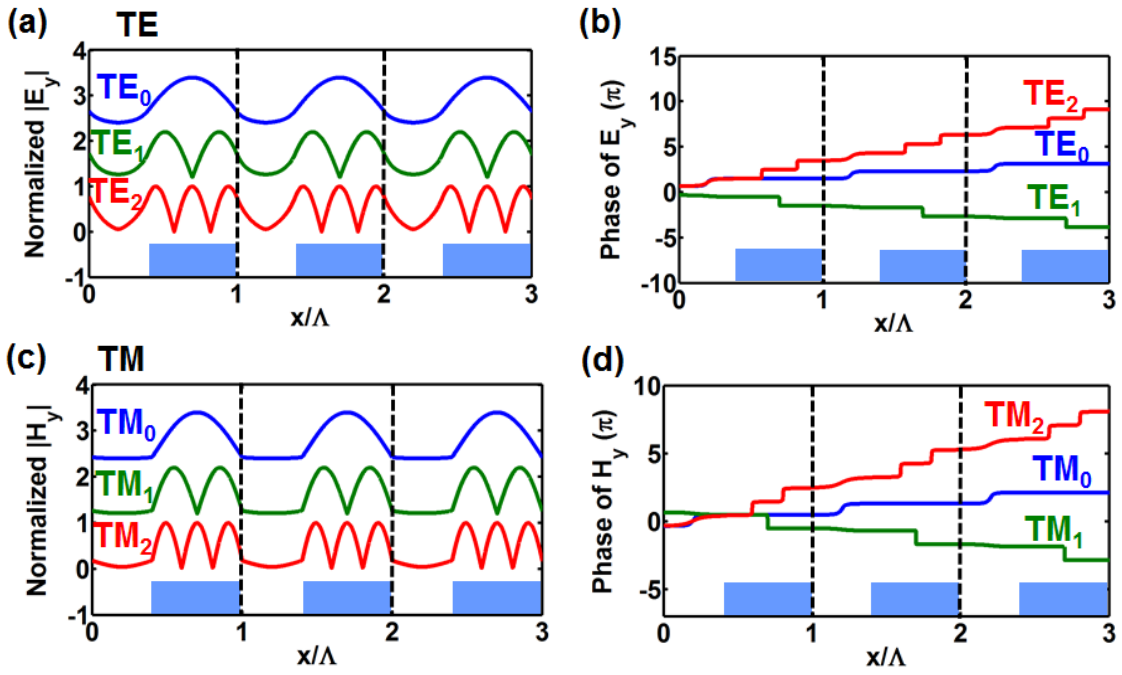


Figure 2.6 Mode profile of the TE waveguide array modes: (a) amplitude, (b) phase. (c) and (d) shows those for the TM waveguide array modes. The blue block indicates the HCG bars. In this calculation, λ=1.3Λ, η=0.45, n_{bar}=3.48, θ=50°.

One distinguished difference between TE and TM in ω-β diagrams locates at their crossing with the air light line. For TE light, the slopes of the ω-β curve are nearly constants, whereas for TM light, the slopes change significantly near the air light line crossings. This difference largely contributes to the different properties of TE and TM HCG, as will be discussed later.

It is interesting to see that there are anti-crossings between the ω-β curves of different modes. These anti-crossings happen when k_a is a real number, i.e. the region below the air light line in the ω-β diagrams. The physical intuition can be explained in the $\omega - k_s s$ diagram shown in Figure 2.7. At high optical frequency, the modes are well confined in the high index waveguide and their mode profiles are distinguishingly different from each other. As the frequency decreases, the $k_{s,m}s$ value of the mth order mode stays nearly constant at $(m+1)\pi$, until the optical frequency approaches $\omega_m = \beta_m c$, i.e. the air light line where $k_{a,m} = 0$. At this region, the modes lose their confinement in the high index waveguide, and the mode profiles spread more and more to the air gap as the optical frequency decreases. As an indicator of the mode confinement inside the high index waveguide, $k_{s,m}s$ value drops from $(m+1)\pi$ to $m\pi$ rapidly as well. Due to the different ω_m of different modes, there are regions where $k_{s,m+1}s$ of the m+1th order drops to $m\pi$, while $k_{s,m}s$ of the mth order mode is still $m\pi$. Thus the mth order mode looks similar as m+1th order; this provides a mode coupling mechanism, leading to an anti-crossing in the $\omega - \beta$ diagrams. This unique feature only appears when k_a is a real number, and is thus not seen in the conventional single slab waveguide.

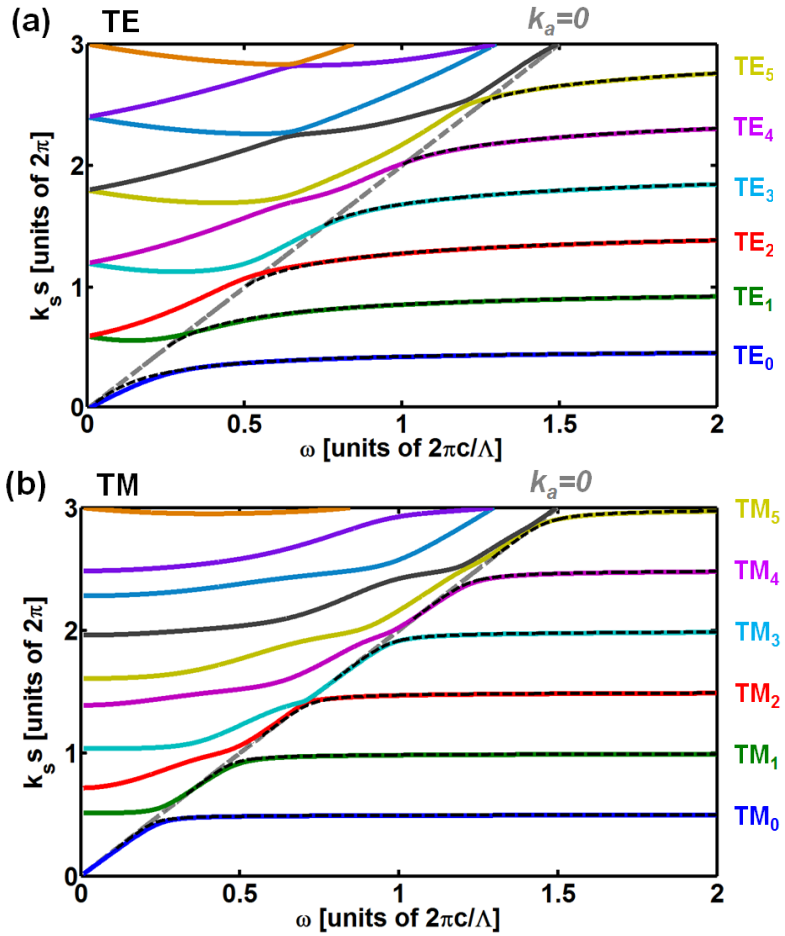


Figure 2.7 ω-$k_s s$ diagram of a waveguide array (solid) and a single slab waveguide (dash), for the same bar width s and index n_{bar}. This provides a physical intuition of the anti-crossing behaviors in the ω-β diagram. The k_a=0 line is essentially the same as the air light line $\beta=\omega/c$ in Figure 2.5. In this calculation, η=0.6, n_{bar}=3.48, θ=50^o.

Based on the mode profiles for Region I-III described in Eqs. 2.1-2.2, we can now calculate the HCG reflectivity by matching the boundary conditions of H_y and E_x at HCG input and output planes, $z = 0$ and $z = t_g$.

To gain physical insight, we first have an overview on the mode behavior in Region II. As described in Eq. 2.2a, once the waveguide-array modes are excited in Region II (inside the HCG), they would propagate along the HCG in $+z$ direction. When they reach the output plane $z = t_g$, they get reflected and coupled into each other. The same happens when the modes return to the input plane $z = 0$. We thus define two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$ to describe the waves propagating in the $+z$ and $-z$ direction, in terms of coefficients of waveguide-array modes. Note that $\boldsymbol{a}$ and $\boldsymbol{b}$ are independent of x and z.

$$\boldsymbol{a} \equiv (a_0 \;\; a_1 \;\; a_2 \ldots)^T; \qquad \boldsymbol{b} \equiv (b_0 \;\; b_1 \;\; b_2 \ldots)^T; \tag{2.3a}$$

The relationship between $\boldsymbol{a}$ and $\boldsymbol{b}$ can be described by a reflection matrix $\boldsymbol{\rho}$, such that

$$\boldsymbol{b} \equiv \boldsymbol{\rho a} \tag{2.3b}$$

Note the reflection matrix $\boldsymbol{\rho}$ in Eq. 2.3b is typically non-diagonal, which means that the HCG modes in Region II couple into each other during the reflection. This does not contradict the orthogonality of the modes in region II, since the reflection involves interaction with the external modes of region III, which are not orthogonal to the modes in region II.

We further define a propagation matrix $\boldsymbol{\varphi}$, to describe the accumulated phase when each mode propagates from $z = 0$ to $z = t_g$. $\boldsymbol{\varphi}$ matrix is a diagonal matrix, since the different modes are orthogonal to each other and thus would not couple to each other during propagation.

$$\boldsymbol{\varphi}_{n,m} = \begin{cases} e^{-j\beta_m t_g} & for\ n = m \\ 0 & for\ n \neq m \end{cases} \tag{2.3c}$$

By matching the boundary conditions of electromagnetic field at the HCG input and exit plane, $\boldsymbol{\rho}$ matrix can be calculated, and thus $\mathbf{R}$ and $\mathbf{T}$. As will be discussed later, it is $\boldsymbol{\varphi}$ (mode propagation inside HCG) and $\boldsymbol{\rho}$ (mode interaction at HCG input and exit plane) that give rise to the rich properties of HCG.

Next, we first match the boundary conditions that H_y and E_x are continuous at the HCG *exit* plane, $z=t_g$.

$$\sum_{n=-\infty}^{+\infty} \tau_n \mathcal{H}_{y,n}^{III}(x) = \sum_{m=0}^{\infty} \mathcal{H}_{y,m}^{II}(x)(a_m - b_m)$$

$$\sum_{n=-\infty}^{+\infty} \tau_n \mathcal{E}_{x,n}^{III}(x) = \sum_{m=0}^{\infty} \mathcal{E}_{x,m}^{II}(x)(a_m + b_m) \tag{2.4a}$$

By performing a Fourier overlap integral on both sides of Eq. 2.4a, we can express the transmitted coefficients τ_n in terms of the overlaps between the lateral (x) magnetic field profiles $\mathcal{H}_{y,m}^{II}(x)$ and $\mathcal{H}_{y,n}^{III}(x)$:

$$\tau_n \Lambda^{-1} \int_0^{\Lambda} \left|\mathcal{H}_{y,n}^{III}(x)\right|^2 dx = \sum_{m=0}^{\infty} (a_m - b_m)\Lambda^{-1} \int_0^{\Lambda} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)dx \qquad (2.4b)$$

$$\tau_n \Lambda^{-1} \int_0^{\Lambda} \left|\mathcal{E}_{x,n}^{III}(x)\right|^2 dx = \sum_{m=0}^{\infty} (a_m + b_m)\Lambda^{-1} \int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,n}^{III,*}(x)dx \qquad (2.4c)$$

$\mathcal{H}_{y,n}^{III,*}(x)$ and $\mathcal{E}_{x,n}^{III,*}(x)$ is the conjugate of $\mathcal{H}_{y,n}^{III}(x)$ and $\mathcal{E}_{x,n}^{III}(x)$ respectively.

Next we define the overlap matrixes **H** and **E** (both are unit-less), for the magnetic and electric field profiles respectively. We rewrite Eqs. 2.4b and 2.4c into the following matrix-vector format:

$$\mathbf{H}_{n,m} = \Lambda^{-1} \int_0^{\Lambda} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)\, dx$$

$$\mathbf{E}_{n,m} = \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}|\gamma_n|/k_0\right)^{-2} \int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,n}^{III,*}(x)\, dx$$

$$\boldsymbol{\tau} = \mathbf{H}(\boldsymbol{a} - \boldsymbol{b}) = \mathbf{E}(\boldsymbol{a} + \boldsymbol{b}) \qquad (2.5a)$$

The closed form of $\mathbf{H}_{n,m}$ and $\mathbf{E}_{n,m}$ can be further derived with Eq. 2.1b, 2.1d, 2.2b and 2.2d. They are presented in the Appendix A.

Eq. 2.5a can be rewritten as follows, using the definition of **ρ** matrix in Eq. 2.3b:

$$\boldsymbol{\tau} = \mathbf{H}(\mathbf{I} - \boldsymbol{\rho})\boldsymbol{a} = \mathbf{E}(\mathbf{I} + \boldsymbol{\rho})\boldsymbol{a} \qquad (2.5b)$$

Since Eq. 2.5b holds for every coefficient vector $\boldsymbol{a}$, we can now derive the reflection matrix $\boldsymbol{\rho}$ as a function of the overlap matrixes **E** and **H**:

$$\mathbf{H}(\mathbf{I} - \boldsymbol{\rho}) = \mathbf{E}(\mathbf{I} + \boldsymbol{\rho}) \;\Rightarrow\; \boldsymbol{\rho} = (\mathbf{H} + \mathbf{E})^{-1}(\mathbf{H} - \mathbf{E}) \qquad (2.6a)$$

Vector $\boldsymbol{\tau}$ defines the coefficient of different modes in Region III, while $\boldsymbol{a}$ defines that in Region II. From Eq. 2.5b, we can define a transmission matrix $\boldsymbol{t}$, which describe how the modes in Region II scatter to Region III.

$$\boldsymbol{t} = \mathbf{H}(\mathbf{I} - \boldsymbol{\rho}) = \mathbf{H}[\mathbf{I} - (\mathbf{H} + \mathbf{E})^{-1}(\mathbf{H} - \mathbf{E})] = 2\mathbf{H}(\mathbf{H} + \mathbf{E})^{-1}\mathbf{E} \qquad (2.6b)$$

Having matched the boundary conditions at the HCG output plane ($z = t_g$), we now repeat the steps in Eqs. 2.4 in order to match the boundary conditions at the HCG *input* plane ($z = 0$):

$$\sum_{n=0}^{\infty}(\delta_{n,0}-r_n)\mathcal{H}_{y,n}^{I}(x)=\sum_{m=0}^{\infty}\mathcal{H}_{y,m}^{II}(x)\left[a_m e^{+j\beta_m t_g}-b_m e^{-j\beta_m t_g}\right]$$

$$\sum_{n=0}^{\infty}(\delta_{n,0}+r_n)\mathcal{E}_{x,n}^{I}(x)=\sum_{m=0}^{\infty}\mathcal{E}_{x,m}^{II}(x)\left[a_m e^{+j\beta_m t_g}+b_m e^{-j\beta_m t_g}\right] \qquad (2.7a)$$

With similar derivation steps as Eqs. 2.4-2.6, we can derive the following equation from Eq. 2.7a:

$$(\mathbf{I}-\mathbf{R})^{-1}\mathbf{H}(\mathbf{I}-\boldsymbol{\varphi}\boldsymbol{\rho}\boldsymbol{\varphi})=(\mathbf{I}+\mathbf{R})^{-1}\mathbf{E}(\mathbf{I}+\boldsymbol{\varphi}\boldsymbol{\rho}\boldsymbol{\varphi}) \qquad (2.7b)$$

By rearranging Eq. 2.7b, we can implicitly express the HCG reflectivity matrix **R** in terms of the matrixes **E, H, ρ** and **φ**:

$$\mathbf{E}(\mathbf{I}+\boldsymbol{\varphi}\boldsymbol{\rho}\boldsymbol{\varphi})(\mathbf{I}-\boldsymbol{\varphi}\boldsymbol{\rho}\boldsymbol{\varphi})^{-1}\mathbf{H}^{-1}=(\mathbf{I}+\mathbf{R})(\mathbf{I}-\mathbf{R})^{-1}\triangleq\mathbf{Z}_{in} \qquad (2.7c)$$

This definition of input impedance matrix $\mathbf{Z}_{\mathbf{in}}$ resembles that used in transmission lines [60], with only difference of the latter being scalar. Using Eq. 2.7c, the HCG reflectivity matrix **R** can finally be calculated, resulting in an equation very common in transmission line theory:

$$\mathbf{R}=(\mathbf{Z}_{in}+\mathbf{I})^{-1}(\mathbf{Z}_{in}-\mathbf{I}) \qquad (2.8a)$$

Similarly, the HCG transmission matrix **T** can be shown to be:

$$\mathbf{T}=\mathbf{2E}(\mathbf{I}+\boldsymbol{\rho})\boldsymbol{\varphi}\left[\left(\mathbf{Z}_{\mathbf{in}}^{-\mathbf{1}}+\mathbf{I}\right)\mathbf{E}(\mathbf{I}+\boldsymbol{\varphi}\boldsymbol{\rho}\boldsymbol{\varphi})\right]^{-\mathbf{1}} \qquad (2.8b)$$

Finally, for subwavelength regime where $>(1+sin\theta)\Lambda$, the HCG reflectivity and transmission are simply $|\mathbf{R}_{00}|^2$and $|\mathbf{T}_{00}|^2$, where

$$|\mathbf{R}_{00}|^2+|\mathbf{T}_{00}|^2\equiv 1 \qquad (2.9)$$

Since subwavelength gratings do not have diffraction orders other than the zeroth order, all power that is not transmitted through the 0th diffraction order gets reflected back. This fact is essential for the design of high reflectivity gratings, since high reflectivity can be achieved by merely cancelling the 0th transmissive diffraction order (i.e. $\tau_0=0$).

It should be note that the above formulas are general, and work well for any wavelength range. In the diffraction regime, not only 0th order propagates and thus carries energy in the z direction, but also the diffraction orders n that satisfies the following condition,

$$\left|\frac{2\pi}{\lambda}sin\theta+\frac{2\pi}{\Lambda}n\right|<\frac{2\pi}{\lambda} \qquad (2.10)$$

The total power for the nth order reflection and transmission is described by $|R_{n0}|^2$ and $|T_{n0}|^2$ respectively, and can be expressed in the following,

$$R_{nth\ order} = |R_{n0}|^2 \frac{\gamma_n}{\gamma_0} \tag{2.11a}$$

$$T_{nth\ order} = |T_{n0}|^2 \frac{\gamma_n}{\gamma_0} \tag{2.11b}$$

The extra factor γ_n/γ_0 in Eq. 2.11a and b comes from the cross section difference between the input light and diffraction light.

With the energy conservation, the total energy in the reflection and transmission light equals that of the input light. We thus have the following equation,

$$\sum_n |R_{n0}|^2 \frac{\gamma_n}{\gamma_0} + \sum_n |T_{n0}|^2 \frac{\gamma_n}{\gamma_0} = 1 \tag{2.11c}$$

where n satisfies Eq. 2.10.

2.2.2 TE-polarized incidence

The analysis for TE polarized incidence follows the same steps as in section 2.2.1. The only differences are summarized in Table 2.1. The reason for the differences is simple: TM solution relies largely on Maxwell-Ampère's equation: $\nabla \times \vec{H} = j\omega\varepsilon_0\varepsilon_r\vec{E}$, whereas TE solution relies largely on Maxwell-Faraday's equation: $\nabla \times \vec{E} = -j\omega\mu\vec{H}$. The lack of ε_r in Faraday's equation explains why n_{bar}^{-2} is replaced by 1, and the minus sign in Faraday's equation explains the minus sign in the last two entries of Table 2.1. The inversion of the wavenumber-ratios at the last two entries of Table 2.1 is also a common difference between TM and TE modes in waveguides.

Table 2.1 Differences between TM and TE polarizations of incidence

Change from TM to TE: Replace (1) by (2)		Equations (Check-marks indicate the changes required for the equation. Equations not listed in this table require no changes.)							
(1)	(2)	2.1a	2.1b	2.1d	2.2a	2.2b	2.2c	2.2d	2.2e
H_y	H_x	√		√	√				
E_x	E_y	√		√	√				
$\mathcal{H}_y$	$\mathcal{H}_x$	√	√	√	√	√	√		
$\mathcal{E}_x$	$\mathcal{E}_y$	√	√	√	√	√	√		
n_{bar}^{-2}	1					√		√	√
β_m/k_0	$-k_0/\beta_m$					√			
γ_n/k_0	$-k_0/\gamma_n$	√	√						

Change from TM to TE: Replace (1) by (2)		Equations (Check-marks indicate the changes required for the equation. Equations not listed in this table require no changes.)							
(1)	(2)	2.4a	2.4b	2.4c	2.5a	2.7a	2.11a	2.11b	2.11c

H_y	H_x								
E_x	E_y								
$\mathcal{H}_y$	$\mathcal{H}_x$	√	√		√	√			
$\mathcal{E}_x$	$\mathcal{E}_y$	√		√	√	√			
n_{bar}^{-2}	1								
β_m/k_0	$-k_0/\beta_m$								
γ_n/k_0	$-k_0/\gamma_n$				√				
γ_n/γ_0	γ_0/γ_n						√	√	√

2.2.3 Comparison with Numerical Simulations

The analytical solutions are compared with the numerical simulation using RCWA and FDTD. Figure 2.8(a) presents the HCG reflectivity spectrum calculated by the analytical solution and RCWA. Figure 2.8 (b) shows the field intensity profile of an HCG reflector using analytical solution and FDTD. Excellent agreements are obtained. The HCG parameter in Figure 2.8(a) are t_g/Λ=0.9, η=0.65, n_{bar}=3.48, θ=40°, TM polarization; whereas in Figure 2.8(b) are λ/Λ =1.60, t_g/Λ =0.2, η =0.4, n_{bar}=3.48, θ=10°, TE polarization.

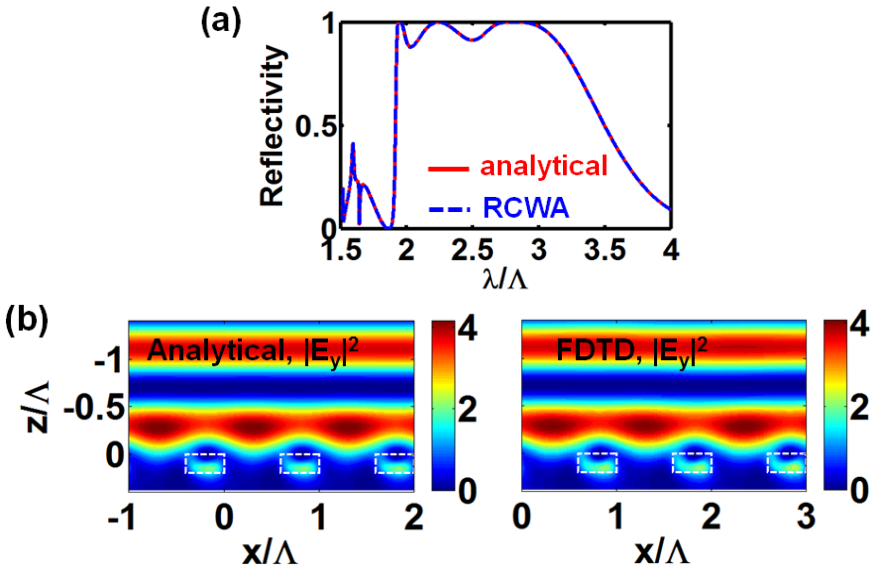


Figure 2.8 Excellent agreement between analytical solutions and commercial numerical simulation on HCG reflectivity and field intensity profile. (a) Comparison of HCG reflectivity spectrum calculated by analytical solution (red) and RCWA (blue). (b) Comparison of HCG field intensity profile ($|\boldsymbol{E}_y|^2$) calculated by analytical solution and FDTD. The white boxes indicate the HCG bars.

2.2.4 General Analytical Treatment – Transmission Matrix

The above section solves the HCG reflection and transmission coefficients, by first defining the modes inside HCG and outside HCG, and then matching the boundary

conditions at the HCG input and exit plane. In fact, the reflection and transmission can be solved in a more general way, by using transmission matrix (T-matrix) method. As described in Figure 2.4, there are three regions in the whole system. Take TM polarization as an example. The electric magnetic field in each region can be expressed with its own mode basis set, i.e. $\mathcal{E}_{x,n}^{I}(x)$ and $\mathcal{H}_{y,n}^{I}(x)$ in region I, $\mathcal{E}_{x,m}^{II}(x)$ and $\mathcal{H}_{y,m}^{II}(x)$ in region II, and $\mathcal{E}_{x,n}^{III}(x)$ and $\mathcal{H}_{y,n}^{III}(x)$ in region III. The electric field in Region I, II and III can then be expressed as,

$$E_x^I(x,z) = \sum_{n=-\infty}^{+\infty} [\boldsymbol{A_n^I}(z) + \boldsymbol{B_n^I}(z)]\mathcal{E}_{x,n}^{I}(x) \tag{2.12a}$$

$$E_x^{II}(x,z) = \sum_{m=0}^{\infty} [\boldsymbol{A_m^{II}}(z) + \boldsymbol{B_m^{II}}(z)]\mathcal{E}_{x,m}^{II}(x) \tag{2.12b}$$

$$E_x^{III}(x,z) = \sum_{m=0}^{\infty} [\boldsymbol{A_m^{III}}(z) + \boldsymbol{B_m^{III}}(z)]\mathcal{E}_{x,m}^{III}(x) \tag{2.12c}$$

Eq. 2.12 are consistent with Eq. 2.1a, 2.1d, and 2.2a, but in a more general form. Coefficients $\boldsymbol{A}$, and $\boldsymbol{B}$ are complex numbers, and denote the amplitude and phase of the corresponding mode, in forward propagation $(+z)$ and backward propagation $(-z)$ direction. They can be written as a column vector,

$$\boldsymbol{A^I} = [\ldots \quad \boldsymbol{A_{-2}^I} \quad \boldsymbol{A_{-1}^I} \quad \boldsymbol{A_0^I} \quad \boldsymbol{A_1^I} \quad \boldsymbol{A_2^I} \quad \ldots]^{\boldsymbol{T}} \tag{2.13a}$$

$$\boldsymbol{A^{II}} = [\boldsymbol{A_0^{II}} \quad \boldsymbol{A_1^{II}} \quad \boldsymbol{A_2^{II}} \quad \ldots]^{\boldsymbol{T}} \tag{2.13b}$$

$$\boldsymbol{A^{III}} = [\ldots \quad \boldsymbol{A_{-2}^{III}} \quad \boldsymbol{A_{-1}^{III}} \quad \boldsymbol{A_0^{III}} \quad \boldsymbol{A_1^{III}} \quad \boldsymbol{A_2^{III}} \quad \ldots]^{\boldsymbol{T}} \tag{2.13c}$$

The above concept is illustrated in Figure 2.9.

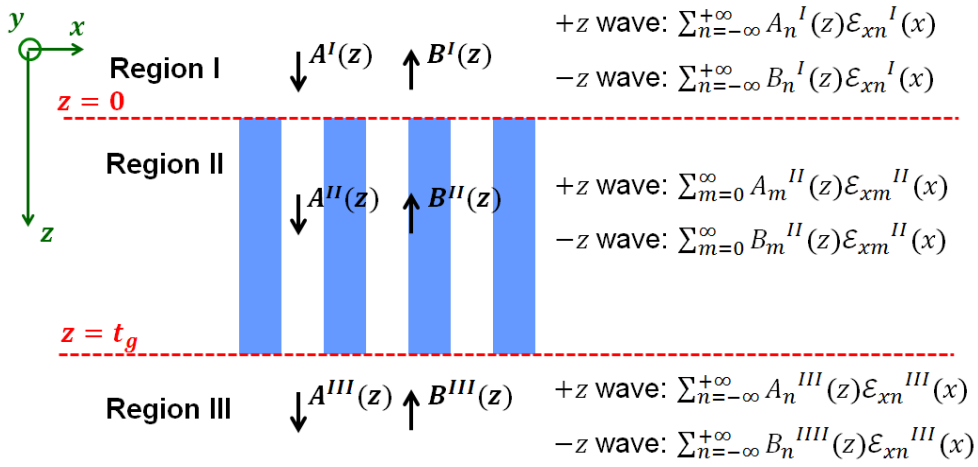


Figure 2.9 General formulation of the electric field in Region I, II and III of the HCG.

A T-matrix can be defined to relate $\boldsymbol{A}$, $\boldsymbol{B}$ vectors at the interface between Region I and II:

$$\begin{pmatrix} \boldsymbol{A^I}(z=0^-) \\ \boldsymbol{B^I}(z=0^-) \end{pmatrix} = \mathbf{T}_{\boldsymbol{I/II}} \begin{pmatrix} \boldsymbol{A^{II}}(z=0^+) \\ \boldsymbol{B^{II}}(z=0^+) \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{A^{II}}(z=0^+) \\ \boldsymbol{B^{II}}(z=0^+) \end{pmatrix} \tag{2.14}$$

By matching the boundary condition, we have

$$\begin{cases} \boldsymbol{B_I}(z=0^-) = \boldsymbol{\rho'} \boldsymbol{A_I}(z=0^-) + \boldsymbol{t} \boldsymbol{B_{II}}(z=0^+) \\ \boldsymbol{A_{II}}(z=0^+) = \boldsymbol{t'} \boldsymbol{A_I}(z=0^-) + \boldsymbol{\rho} \boldsymbol{B_{II}}(z=0^+) \end{cases} \tag{2.15}$$

Matrix $\boldsymbol{\rho}$ / $\boldsymbol{\rho'}$ describes how the modes in region II / I reflect to themselves and cross couple to each other at the interface; matrix $\boldsymbol{t}$ / $\boldsymbol{t'}$ describes how the modes in region II / I transmit to the modes in region I / II. $\boldsymbol{\rho}$ and $\boldsymbol{t}$ is already defined in Eq. 2.6; $\boldsymbol{\rho'}$ and $\boldsymbol{t'}$ can be solved in a similar way. We summarize their expression in Table 2.2. $\boldsymbol{\rho}$, $\boldsymbol{\rho'}$, $\boldsymbol{t}$ and $\boldsymbol{t'}$ describe the property of the interface, and solely depend on the matrix $\boldsymbol{H}$ and $\boldsymbol{E}$ in Eq. 2.5a, which are the overlap integral between the modes in region I and region II. More generally, matrix $\boldsymbol{H}$ and $\boldsymbol{E}$ can be written in the following,

$$H_{n,m} = \frac{\int_0^{\Lambda} \mathcal{H}_{y,m}^{II}(x) \left[\mathcal{H}_{y,n}^{I}(x)\right]^* dx}{\int_0^{\Lambda} \left|\mathcal{H}_{y,n}^{I}(x)\right|^2 dx} \tag{2.16a}$$

$$E_{n,m} = \frac{\int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x) \left[\mathcal{E}_{x,n}^{I}(x)\right]^* dx}{\int_0^{\Lambda} \left|\mathcal{E}_{x,n}^{I}(x)\right|^2 dx} \tag{2.16b}$$

Table 2.2. Matrix $\boldsymbol{\rho}$, $\boldsymbol{\rho'}$, $\boldsymbol{t}$, and $\boldsymbol{t'}$

$\boldsymbol{\rho} = (\boldsymbol{H}+\boldsymbol{E})^{-1}(\boldsymbol{H}-\boldsymbol{E})$
$\boldsymbol{\rho'} = -\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}(\boldsymbol{H}-\boldsymbol{E})\boldsymbol{H}^{-1}$
$\boldsymbol{t} = 2\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}\boldsymbol{E}$
$\boldsymbol{t'} = 2(\boldsymbol{H}+\boldsymbol{E})^{-1}$

Though the expressions of $\boldsymbol{H}$, $\boldsymbol{E}$, $\boldsymbol{\rho}$, $\boldsymbol{\rho'}$, $\boldsymbol{t}$ and $\boldsymbol{t'}$ are derived for HCG, they are general and can be applied to other systems as well. For example, if the two regions are homogeneous media, and there is only a plane wave mode in each region, then $\boldsymbol{\rho}$, $\boldsymbol{\rho'}$, $\boldsymbol{t}$ and $\boldsymbol{t'}$ are reduced to the plane wave reflection and transmission coefficient between two media, denoted by the Fresnel's Law. See Appendix B for this derivation.

With the matrix algebra, **T-matrix** for the HCG interface between region I and II can be expressed as,

$$\mathbf{T}_{\boldsymbol{I/II}} = \begin{pmatrix} \boldsymbol{t'^{-1}} & -\boldsymbol{t'^{-1}}\boldsymbol{\rho} \\ \boldsymbol{\rho'}\boldsymbol{t'^{-1}} & -\boldsymbol{\rho'}\boldsymbol{t'^{-1}}\boldsymbol{\rho} + \boldsymbol{t} \end{pmatrix} \tag{2.17}$$

T-matrix can also be defined to describe the mode propagation in each region. For example, T-matrix for HCG with thickness $\boldsymbol{t_g}$ can relate the $\boldsymbol{A}$, $\boldsymbol{B}$ vectors at $\boldsymbol{z = 0^+}$ and $\boldsymbol{z = t_g^-}$:

$$\begin{pmatrix} \boldsymbol{A^{II}}(z=0^+) \\ \boldsymbol{B^{II}}(z=0^+) \end{pmatrix} = \mathbf{T}_{\boldsymbol{HCG}} \begin{pmatrix} \boldsymbol{A^{II}}(z=\boldsymbol{t_g^-}) \\ \boldsymbol{B^{II}}(z=\boldsymbol{t_g^-}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi}^{-1} & 0 \\ 0 & \boldsymbol{\varphi} \end{pmatrix} \begin{pmatrix} \boldsymbol{A^{II}}(z=\boldsymbol{t_g^-}) \\ \boldsymbol{B^{II}}(z=\boldsymbol{t_g^-}) \end{pmatrix} \tag{2.18}$$

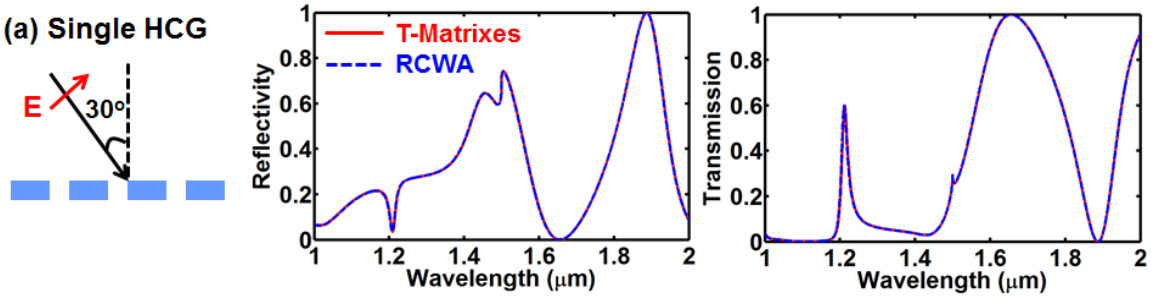
where $\boldsymbol{\varphi}$ matrix is defined at Eq. 2.3c.

With all these **T-matrix** defined, one can multiple them together to relate $\boldsymbol{A}$, $\boldsymbol{B}$ vectors at $z = 0^-$ and $z = t_g^+$:

$$\begin{pmatrix} \boldsymbol{A^I}(z=0^-) \\ \boldsymbol{B^I}(z=0^-) \end{pmatrix} = T_{I/II} T_{HCG} T_{II/III} \begin{pmatrix} \boldsymbol{A^{II}}(z=\boldsymbol{t_g^+}) \\ \boldsymbol{B^{II}}(z=\boldsymbol{t_g^+}) \end{pmatrix} \tag{2.19}$$

By setting $\boldsymbol{A^I}(z=0^-) = (\ldots\ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ldots)^T$ and $\boldsymbol{B^{II}}(z=\boldsymbol{t_g^+}) = (0\ \ 0\ \ 0\ \ldots)^T$, one can obtain $\boldsymbol{r} = \boldsymbol{B^I}(z=0^-)$ and $\boldsymbol{\tau} = \boldsymbol{A^{II}}(z=\boldsymbol{t_g^+})$.

The above T-matrix method essentially follows the same principle with that for multiple-layer structure where each layer is a homogenous material. Instead of there being only plane waves in each layer, there is a specific mode set in each "layer". The propagation constants of these modes define the T-matrix in the corresponding layer, and the mode overlap between different layers defines the property and thus the T-matrix of the interface. T-matrix can also be defined for the interface between two different HCGs. By cascading these T-matrixes together, one can solve any configuration of the HCG, such as HCG on silicon-on-insulator, two stacking HCGs, etc. This thus greatly expands the analytical treatment of HCG. Figure 2.10 shows the reflection and transmission spectrum of some exemplary HCG structures calculated by T-matrixes method. They are well matched with the results calculated by RCWA. Figure 2.10(a) shows a single TM HCG, with $\Lambda = 1\ \mu m$, $t_g = 0.3\ \mu m, \eta = 0.6$, grating bar index 3.48, incident angle 30^o; (b) shows a TE HCG on silicon-on-insulator (SOI) wafer, $\Lambda = 1\ \mu m$, $t_g = 0.35\ \mu m, \eta = 0.55$, grating bar index 3.48, incident angle 40^o, oxide thickness $2\ \mu m$, oxide index 1.44, and substrate index 3.48; (c) shows a TE HCG stack, with top HCG $t_g = 0.3\ \mu m$, grating bar index 3.48, bottom HCG $t_g = 0.5\ \mu m$, grating bar index 2.5, the thickness and index of the layer between the two HCGs $0.4\ \mu m$ and 1.44, both HCG $\Lambda = 1\ \mu m$, $\eta = 0.6$, incident angle 80^o; (d) shows a TM HCG stack, with all the parameters same with (c) but the thickness of the layer between the two HCGs $0.3\ \mu m$.



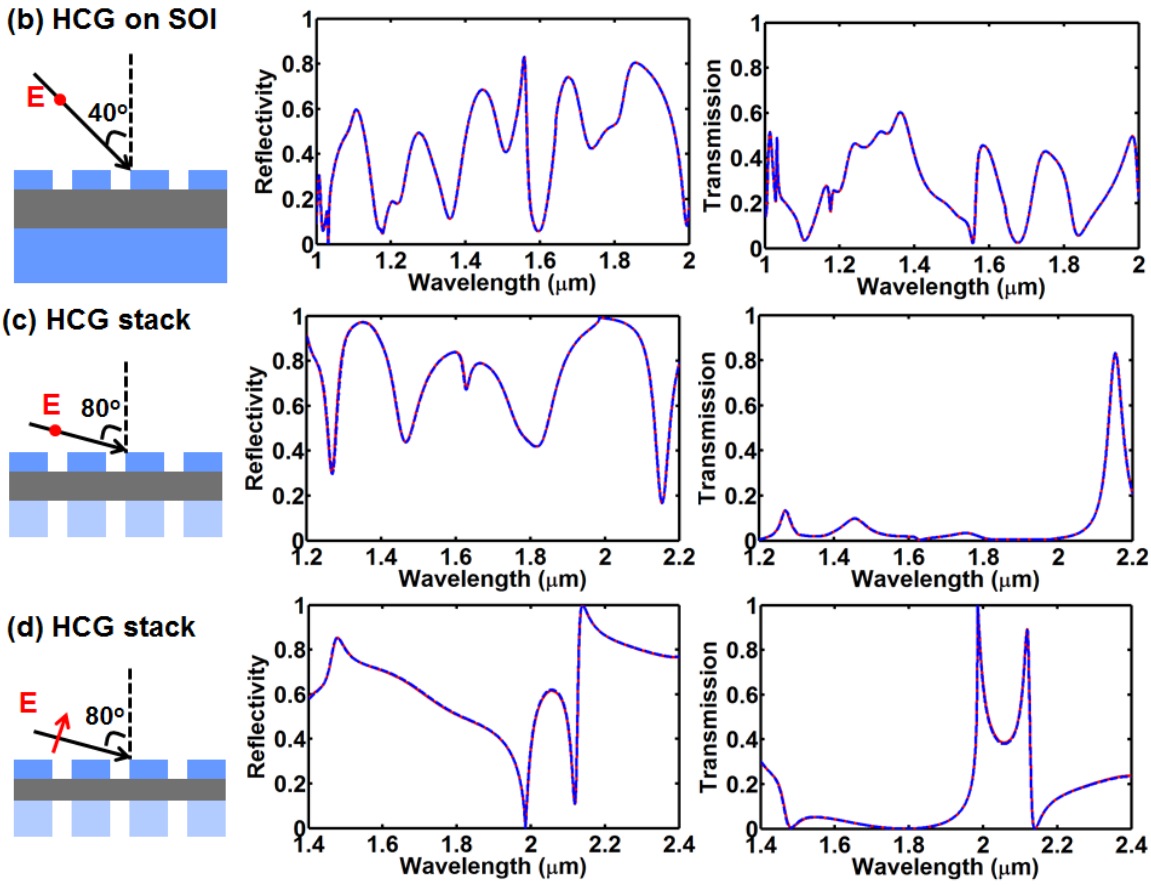


Figure 2.10 Reflection and transmission spectrum of exemplary HCG structures. The results solved by T-matrixes are indicated by red solid lines. They are well matched with the resulted calculated by RCWA, indicated by the blue dashed lines. (a) a single TM HCG, with Λ=1 μm, t_g=0.3 μm, η=0.6, grating bar index 3.48, incident angle 30°; (b) a TE HCG on SOI wafer, Λ=1 μm, t_g=0.35 μm, η=0.55, grating bar index 3.48, incident angle 40°, oxide thickness 2 μm, oxide index 1.44, and substrate index 3.48; (c) a TE HCG stack, with top HCG t_g=0.3 μm, grating bar index 3.48, bottom HCG t_g=0.5 μm, grating bar index 2.5, the thickness and index of the layer between the two HCGs 0.4 μm and 1.44, both HCG Λ=1 μm, η=0.6, incident angle 80°; (d) a TM HCG stack, with all the parameters same with (c) but the thickness of the layer between the two HCGs 0.3 μm.

2.2.5 Special case at surface normal incidence

The above HCG analytical formulation applies to any incident angle θ. A special case is the surface normal incidence, i.e. $\theta = 0$. At surface normal, the field in all the three regions in Figure 2.4 is symmetric along $x = a$ or $x = a + s/2$. In region I and III, $\mathcal{E}_{x,-n}^{I/III}(x)$ and $\mathcal{E}_{x,+n}^{I/III}(x)$ can be combined and expressed as cosinusoidal function. The same applies to $\mathcal{H}_{y,n}^{I/III}(x)$. In region II, there are only even modes TE_0/TM_0, TE_2/TM_2 etc, since the odd modes cannot be excited by the surface normal incident plane waves in Region I. By setting $\theta = 0$ in the above formulation, the analytical solution for surface normal

incidence can be obtained. Table 2.3 denotes these equations at normal incidence condition.

Table 2.3. Formulations at surface normal incidence

Equation	Expression
2.1a	$H_y^I(x,z) = e^{-jk_0 z} - \sum_{n=0}^{+\infty} r_n \mathcal{H}_{y,n}^I(x) e^{+j\gamma_n z} \qquad z < 0$ $E_x^I(x,z) = \sqrt{\frac{\mu_0}{\varepsilon_0}} e^{-jk_0 z} + \sum_{n=0}^{+\infty} r_n \mathcal{E}_{x,n}^I(x) e^{+j\gamma_n z} \qquad z < 0$
2.1b	$\mathcal{H}_{y,n}^I(x) = cos[(2n\pi/\Lambda)(x - a/2)]$ $\mathcal{E}_{x,n}^I(x) = (\gamma_n/k_0)\sqrt{\frac{\mu_0}{\varepsilon_0}} cos[(2n\pi/\Lambda)(x - a/2)]$
2.1c	$k_0 = 2\pi/\lambda$ $\gamma_n^2 = k_0^2 - (2n\pi/\Lambda)^2$
2.1d	$H_y^{III}(x,z) = \sum_{n=0}^{\infty} \tau_n \mathcal{H}_{y,n}^{III}(x) e^{-j\gamma_n (z - t_g)} \qquad z > t_g$ $E_x^{III}(x,z) = \sum_{n=0}^{\infty} \tau_n \mathcal{E}_{x,n}^{III}(x) e^{-j\gamma_n (z - t_g)} \qquad z > t_g$
2.1e	$\boldsymbol{r} \equiv (r_0 \;\; r_1 \;\; r_2 \;\; \ldots)^T$
2.1f	$\boldsymbol{\tau} \equiv (\tau_0 \;\; \tau_1 \;\; \tau_2 \;\; \ldots)^T$
2.1g	$\boldsymbol{r} \equiv \mathbf{R}(1 \;\; 0 \;\; 0 \ldots)^T$
2.1h	$\boldsymbol{\tau} \equiv \mathbf{T}(1 \;\; 0 \;\; 0 \ldots)^T$
2.2a, b	Same as the general formulation
2.2c	$\mathcal{H}_{y,m}^{II}(x) = \mathcal{H}_{y,m}^{II}(x + l\Lambda), \qquad l \; is \; an \; integer$ $\mathcal{E}_{x,m}^{II}(x) = \mathcal{E}_{x,m}^{II}(x + l\Lambda), \qquad l \; is \; an \; integer$
2.2d	$A = cos(k_{s,m} s/2)$ $C = cos(k_{a,m} a/2)$ $B = 0$ $D = 0$
2.2e	$n_{bar}^{-2} k_{s,m} tan(k_{s,m} s/2) = -k_{a,m} tan(k_{a,m} a/2)$
2.2f, g	Same as the general formulation
2.3-2.4	Same as the general formulation
2.5a	$\mathbf{H}_{n,m} = \Lambda^{-1}(2 - \delta_{n,0}) \int_0^{\Lambda} \mathcal{H}_{y,m}^{II}(x) \mathcal{H}_{y,n}^{III}(x)\, dx$ $\mathbf{E}_{n,m} = \Lambda^{-1}(2 - \delta_{n,0}) \left(\sqrt{\frac{\mu_0}{\varepsilon_0}} \gamma_n / k_0 \right)^{-2} \int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x) \mathcal{E}_{x,n}^{III}(x)\, dx$ $\boldsymbol{\tau} = \mathbf{H}(\boldsymbol{a} - \boldsymbol{b}) = \mathbf{E}(\boldsymbol{a} + \boldsymbol{b})$
2.5b, 2.6	Same as the general formulation

2.7-2.9	Same as the general formulation
2.10	$\left\|\frac{2\pi}{\Lambda}n\right\| < \frac{2\pi}{\lambda}$
2.11a	$R_{nth\ order} = R_{-nth\ order} = \frac{1}{4}\|R_{n0}\|^2 \frac{\gamma_n}{\gamma_0}$ $T_{nth\ order} = T_{-nth\ order} = \frac{1}{4}\|T_{n0}\|^2 \frac{\gamma_n}{\gamma_0}$

Table 2.1 applies to surface normal incidence.

2.3 HCG Supermodes and Their Interferences

The analytical formulation in Section 2.2 helps to understand the various HCG properties. We have briefly discussed the HCG supermodes in Section 2.1, and the resonance curves of these modes in Figure 2.3. Now we can have more insight on it.

From the HCG cavity point of view, $\boldsymbol{\varphi}$ and $\boldsymbol{\rho}$ are the most important matrices. $\boldsymbol{\varphi}$ characterizes the propagation of the waveguide-array mode inside HCG and $\boldsymbol{\rho}$ describes the mode interaction at HCG input and exit plane. The Eigen modes of this entire "system", i.e. a finite-thickness HCG with input/exit planes, are called supermodes. The supermodes can be obtained by the diagonalization of the $\boldsymbol{\rho\varphi}$ matrix. The diagonal $\boldsymbol{\rho\varphi}$ matrix can be written as $\boldsymbol{V}^{-1}(\boldsymbol{\rho\varphi})\boldsymbol{V}$, where $\boldsymbol{V}$ is the eigenvector matrix. The supermodes are essentially a linear combination of the waveguide array modes, linked by the eigenvector matrix $\boldsymbol{V}$. The phase of the mth diagonal elements in $\boldsymbol{V}^{-1}(\boldsymbol{\rho\varphi})\boldsymbol{V}$ describes the phase accumulated by the supermode *m* after one single-trip (or half-round-trip) through the grating in *z* direction, denoted as ψ_m:

$$\psi_m \equiv phase[\boldsymbol{V}^{-1}(\boldsymbol{\rho\varphi})\boldsymbol{V}]_{mm} \tag{2.20}$$

The Fabry-Perot resonance conditions for phases defined in Eq. 2.20 are simply:

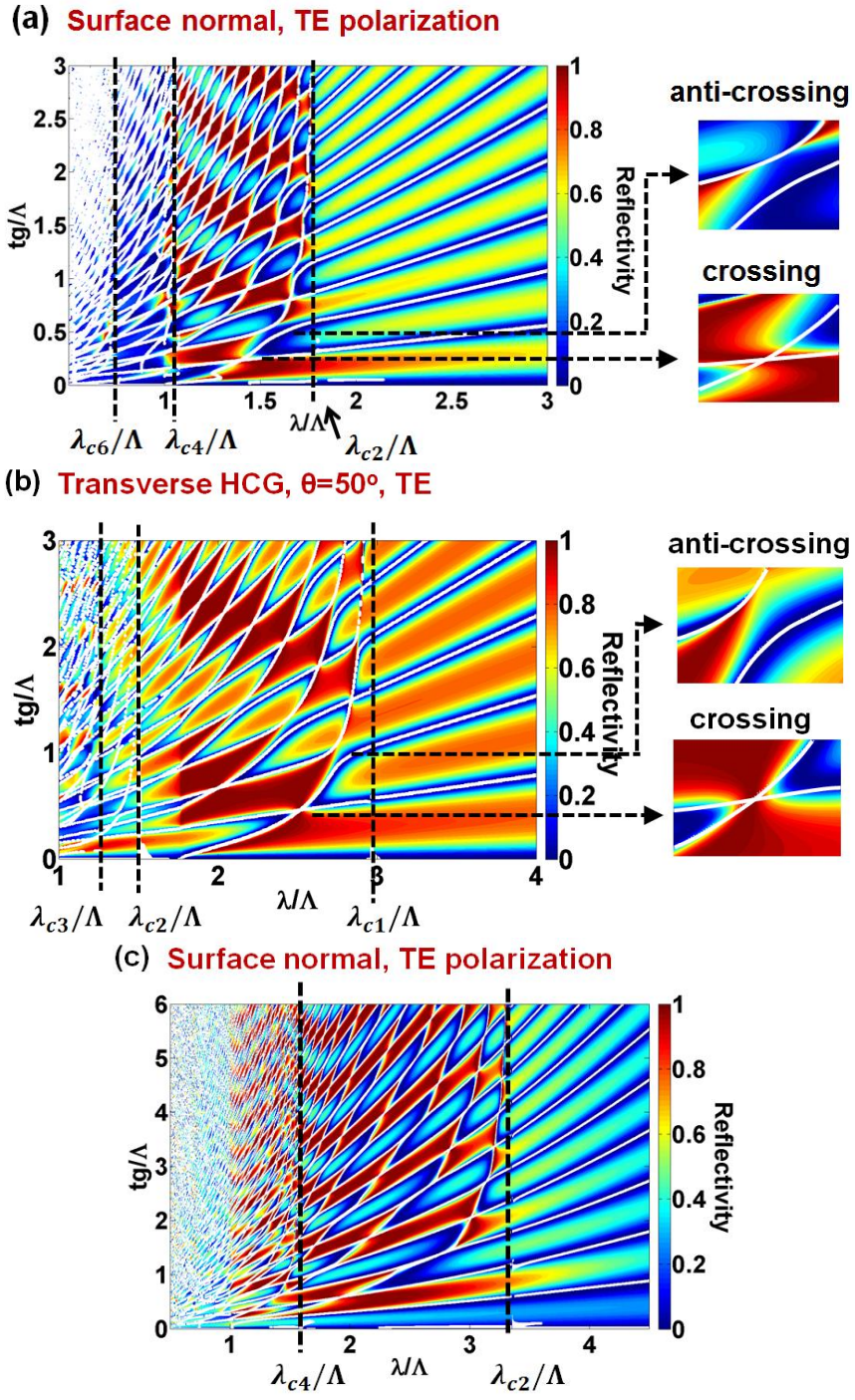
$$\psi_m = n\pi, \text{where } n = 0, 1, 2, \ldots \tag{2.21}$$

The analytical solutions of Eq. 2.21 are the resonance curves. Figure 2.11 shows these resonance curves for the transverse HCG both at surface normal incidence and oblique incidence [example in Figure 2.3 (a-d)]. The HCG duty cycles are 0.45, 0.45, 0.75 and 0.6 for Figure 2.11 (a)~(d); $n_{bar} = 3.48$ for all the cases. The resonance curves are plotted on top of the reflectivity contour plot for comparison. Several families of curves are seen, and they correspond to different order of supermodes. Resonance curves of different supermodes can intersect with each other. These intersection points corresponds to

$$|\Delta\psi| = |\psi_{\mathrm{m1}} - \psi_{\mathrm{m2}}| = l\pi, \text{where } l = 0, 1, 2, \ldots \tag{2.22}$$

The intersection of two resonance lines indicates that two supermodes reach their Fabry-Perot conditions *simultaneously*. For odd *l*, the two FP lines have an odd number

multiple of π-phase difference, and the resonance lines simply cross each other, herein referred to as "crossing". On the other hand, for even *l*, the intersecting FP lines are in-phase, and the two FP lines repel against each other, forming an "anti-crossing". The detailed views of these intersections are shown at the two right insets of Figure 2.11 (a) and (b). The crossing and anti-crossing points lead to the extraordinary phenomena of high-Q resonances, 100% reflectivity, or 100% transmission. In the surface normal incidence condition, there are only even modes excited, and thus only the resonance curves of the even modes show up.



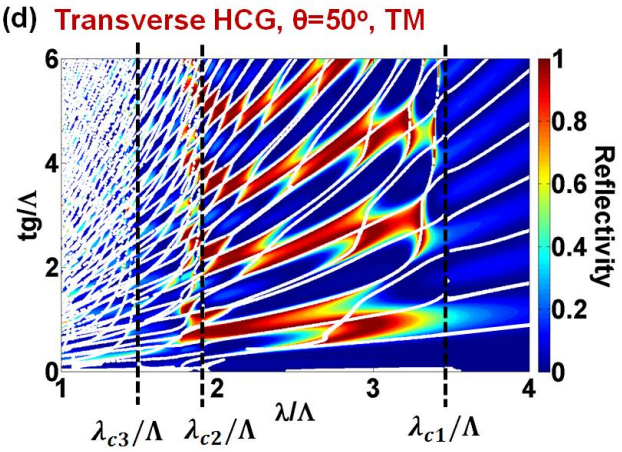


Figure 2.11 Analytical solutions of Fabry-Perot resonance conditions of the individual supermodes, shown by the white curves, superimposed on the reflectivity contour of an HCG as a function of wavelength and grating thickness. The inset in (a) and (b) shows the anti-crossing and crossing conditions.

2.3.1 The mechanism of high reflectivity and 100% reflectivity

The HCG can be easily designed to exhibit a wide dual-mode regime. Intuitively, if the two modes have complete destructive interference at the HCG output plane, the transmission will vanish, yielding a 100% reflectivity. Mathematically, this requires $T_{00} = 0$, and thus $\tau_0 = 0$. According to Eq. 2.5a, $\boldsymbol{\tau} = \mathbf{E}(\boldsymbol{a} + \boldsymbol{b})$, and therefore we can summarize the 100% reflectivity condition as follows:

$$|\mathbf{R}_{00}| = 100\% \quad \leftrightarrow \quad \tau_0 = [\mathbf{E}(\boldsymbol{a} + \boldsymbol{b})]_0 = \sum_m \mathbf{E}_{0m}(a_m + b_m) \tag{2.23}$$

This implies that (for TM light):

$$\sum_m (a_m + b_m)\Lambda^{-1} \int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,0}^{III,*}(x)\, dx = 0$$

$$\rightarrow \quad \sum_m (a_m + b_m)\, \Lambda^{-1} \int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x) e^{j(k_0 sin\theta)x}\, dx = 0 \tag{2.24}$$

The underlying concept in Eq. 2.24 is straight-forward. At the HCG exit plane, the overlap integral of the grating modes in Region II with the 0th order transmissive plane wave $\mathcal{E}_{x,0}^{III}(\boldsymbol{x})$ (a tilted wavefront with angle θ from surface normal) indicates the total energy transferred from HCG to the 0th order transmissive plane wave. More generally, if we represent the overall field profile (from all grating modes combined) at the output plane $\boldsymbol{z} = \boldsymbol{t_g}^-$ in terms of a Fourier series as a function of x (Λ being the Fourier series periodicity), the overlap integral in Eq. 2.24 has the mathematical meaning of 0th order Fourier coefficient (DC Fourier coefficient). In general, nth Fourier coefficient in such expansion would correspond directly to the transmission coefficient τ_n of nth

transmissive diffraction orders. However, since only the 0^{th} order Fourier component carries power along z due to the subwavelength nature of HCG, it is only the DC Fourier coefficient that we need to suppress in order to have full reflection.

By applying the two-mode approximation, we obtain the simplified condition for 100% reflectivity:

$$|\mathbf{R}_{00}| = 100\%$$

$$\leftrightarrow \underbrace{\underbrace{(a_0 + b_0)\Lambda^{-1}\int_0^{\Lambda} \mathcal{E}_{x,0}^{II}(x)e^{j(k_0 sin\theta)x}\,dx}_{overlap\ integral\ of\ the\ 0th\ order\ mode\ (forward+backward)\ with\ a\ tilted\ wavefront} + \underbrace{(a_1 + b_1)\Lambda^{-1}\int_0^{\Lambda} \mathcal{E}_{x,1}^{II}(x)e^{j(k_0 sin\theta)x}\,dx}_{overlap\ integral\ of\ the\ 1st\ order\ mode\ (forward+backward)\ with\ a\ tilted\ wavefront}}_{\text{"destructive interference" (cancellation) between the 0th order and the 1st order modes at z=}t_g^-}$$

$$= 0 \qquad (2.25)$$

We refer the above two mode cancellation condition as the "destructive interference" between the two modes. We use quote-marks due to the unusual *mode overlap integral* (or *lateral average*) interpretation of interference. Figure 2.12 illustrates the destructive interference concept for a TE HCG. The average lateral electric fields of the two modes are plotted as a function of wavelength, for both the amplitude and phase. HCG parameters are: $n_{bar} = 3.48$, $s/\Lambda = 0.45$, $t_g/\Lambda = 0.495$, $\theta = 50^o$ and TE polarization of incidence.

At the wavelengths of 100% reflectivity, the two DC Fourier components in Eq. 2.25 have equal magnitude and opposite phase ($\Delta\phi = \pi$), i.e. a perfect cancellation, shown in Figure 2.12 (a). If two such 100% reflectivity points are located at close spectral vicinity, a broad-band of high reflectivity is achieved. Figure 2.12 (b) illustrates the non-traditional interpretation for destructive interference. The lateral field profiles of 0^{th} and 1^{st} order modes, weighted by the tilted wavefront $e^{j(k_0 sin\theta)x}$, and their sum are plotted as a function of x. The summation field profile is zero only in terms of DC-component. Had the grating not been subwavelength, this cancellation would not be enough to yield 100% reflection, since higher diffraction orders would not be zero and thus can transmit energy.

To further testify this destructive interference concept, at the HCG exit plane, the two modes' DC-component "magnitude difference" as well as "phase difference" are plotted against normalized wavelength and grating thickness (both by Λ), compared with the reflectivity contour plot, in Figure 2.13. The light is at surface normal incidence, and Figure 2.13 (a) shows the TM HCG with $\eta = 0.75$, $n_{bar} = 3.48$, and (b) shows the TE HCG with $\eta = 0.45$, $n_{bar} = 3.48$.

It is clearly seen that the "checker-board" pattern aligns well with the two modes' phase difference contour. The amplitude difference, on the other hand, is relatively uniform. At the two modes' region, except at the resonance lines where the input plane wave is strongly coupled into one of the modes, the plane wave couples relatively uniformed to the two modes. Thus the two modes' phase difference dominantly determines the reflectivity contour. Furthermore, the HCG conditions which satisfies Eq. 2.25 is overlaid onto the reflectivity contour plot. Those are exactly the high reflectivity

regions. Excellent agreement is thus achieved between the simulation results and the physical insights given by Eq. 2.25. If one plots a series of contour plot with different duty cycle η, one can easily choose an η and a t_g to design a broadband high reflection mirror.

In the above discussion, the interference of the two modes determines the HCG reflection. As seen in Eq. 2.23-2.25, these two modes refer to the waveguide-array modes. Equivalently, one can also explain the reflectivity phenomena with the supermode concept. This is probably more suitable to explain the checker-board pattern on the reflectivity contour map. As seen in Figure 2.13, each compartment of the checker board (diamond shape) is bounded by two sets of quasi-parallel resonance curves of the two supermodes (Eq. 2.21). Inside the compartment, the phase accumulation of each supermode in a half-round-trip is either from 0 to π, or π to 2π. Thus the two supermodes are either destructive interference or constructive interference at the HCG exit plane. Since the phase value of the resonance curves alternates by π, the destructive interference or constructive interference alternates between adjacent compartments, yielding a checker-board pattern on the reflectivity.

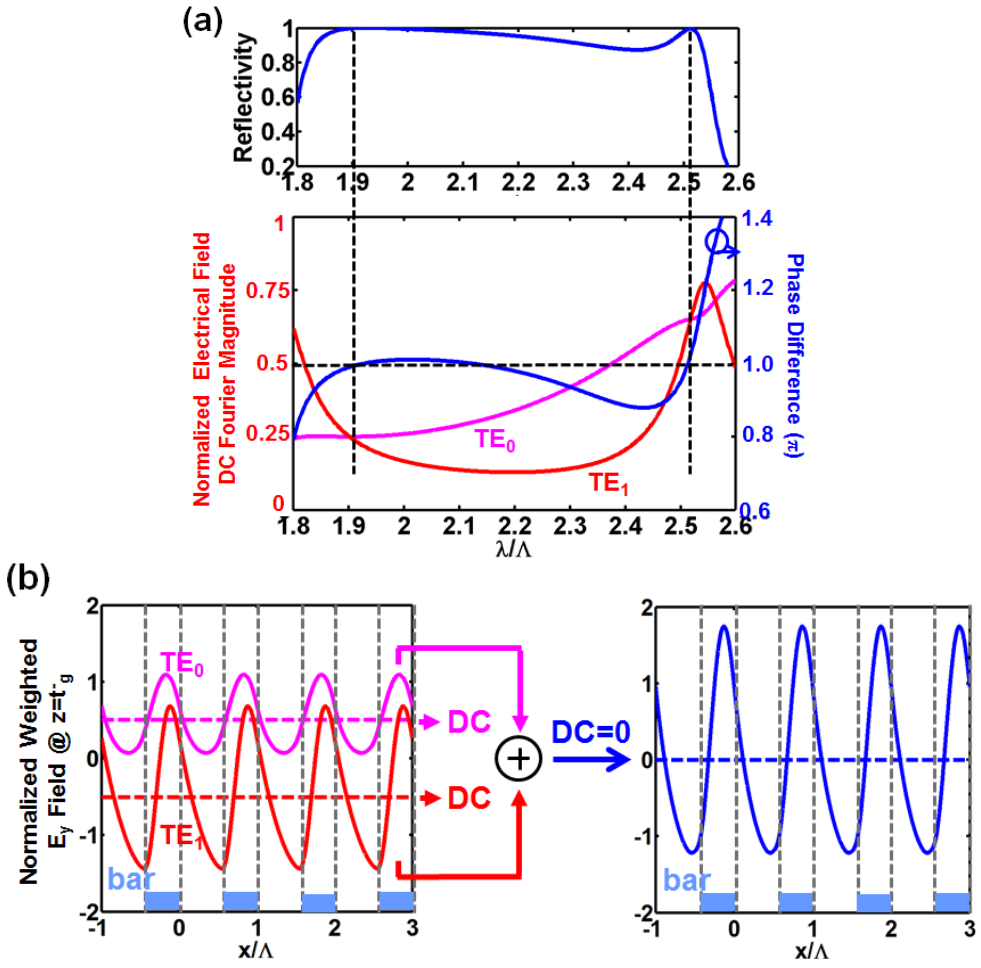
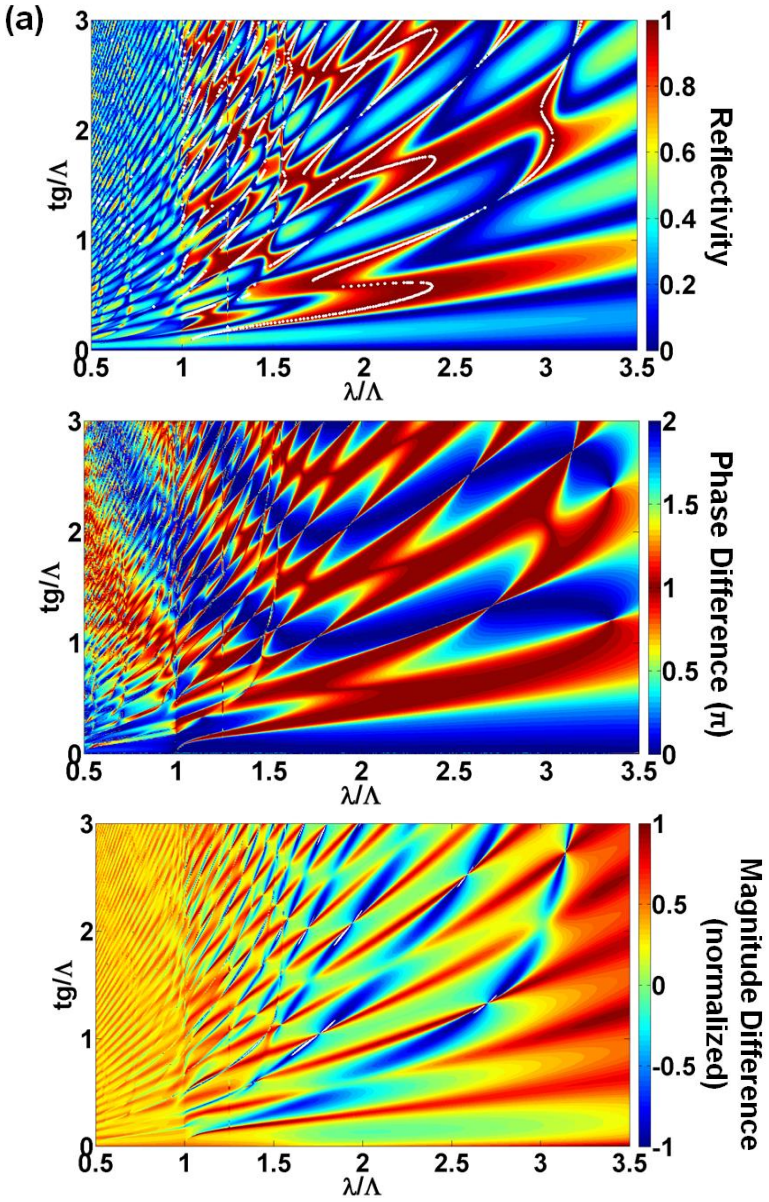


Figure 2.12 (a) Two-mode solution exhibiting perfect cancellation at the HCG exit plane leading to 100% reflectivity for TE polarized light. At the wavelengths of 100%

reflectivity (marked by the two vertical dashed lines), both modes have the same magnitude of the "DC" lateral Fourier component, but opposite phases. (b) Two-mode solution for the field profile (weighted by $e^{j(kosin\theta)x}$) at the HCG exit plane in the case of perfect cancellation. The cancellation is shown to be only in terms of the "DC" Fourier component. The higher Fourier components do not need to be zero, since subwavelength gratings have no diffraction orders other than 0^{th} order. The left plot shows the decomposition of the overall weighted field profile into the two modes, whereby the DC-components of these two modes cancel each other. HCG parameters are the same as (a), and λ/Λ=2.516, corresponding to the dashed line on the right-hand side in (a).



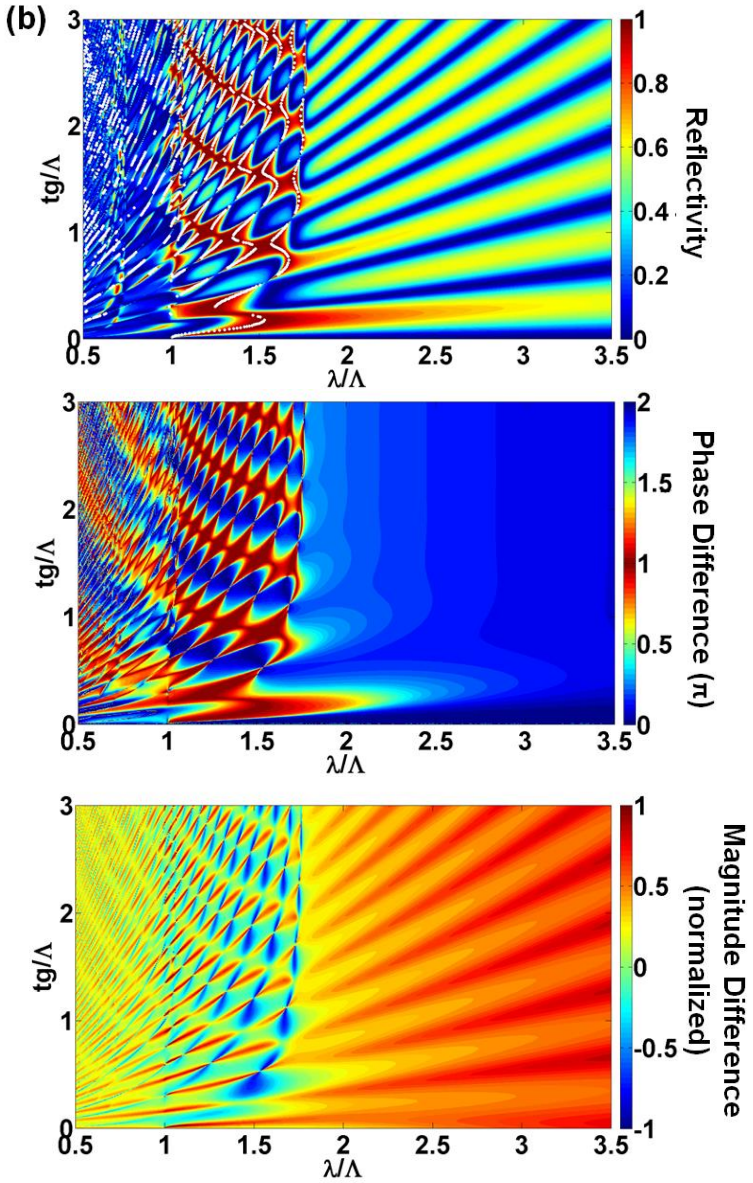


Figure 2.13 Reflectivity contour plot compared with the two modes' DC-component "phase difference" at the HCG output plane and "magnitude difference" at the HCG output plane. The white curves overlaid onto the reflectivity contour plot indicate the HCG conditions where Eq. 2.25 is satisfied. At the two modes region, except at the resonance curves, the input plane wave couples relatively equal to both modes. Thus their phase difference dominantly determines the reflectivity. (a) TM HCG, η=0.75, n_{bar}=3.48. (b) TE HCG, η=0.45, n_{bar}=3.48. Surface normal incidence.

2.3.2 The mechanism of 100% transmission

Similar to the analysis for the 100% reflection, the 100% transmission in the two-mode region can also be explained by the interference concept. With Eqs. 2.1a, 2.1b for the input, and Eq. 2.1d for the output, we get the following condition for 100% transmission

$$|\mathbf{T}_{00}| = 100\% \quad \leftrightarrow \quad |\tau_0| = |[\mathbf{E}(\boldsymbol{a}+\boldsymbol{b})]_0| = \left|\sum_m \mathbf{E}_{0m}(a_m + b_m)\right| = 1 \qquad (2.26)$$

Similar as Eq. 2.24, τ_0 can be rewritten as follows (for TM light):

$$\left|\sum\nolimits_m (a_m + b_m)\Lambda^{-1}\int_0^\Lambda \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,0}^{III,*}(x)\,dx\right| = \left(\gamma_0/k_0\sqrt{\frac{\mu_0}{\varepsilon_0}}\right)^2$$

$$\dashrightarrow \left|\sum\nolimits_m (a_m + b_m)\Lambda^{-1}\int_0^\Lambda \mathcal{E}_{x,m}^{II}(x)e^{j(k_0 sin\theta)x}\,dx\right| = \gamma_0/k_0\sqrt{\frac{\mu_0}{\varepsilon_0}} \qquad (2.27)$$

The underlying concept in Eq. 2.27 is that at the HCG exit plane, the lateral average of all the grating modes weighted by the tilted wavefront $e^{j(k_0 sin\theta)x}$, equals to the amplitude of the input plane wave. In other words, the DC Fourier component of the grating modes in the HCG output plane has the same magnitude as the input plane wave in Region I. When these grating modes out-couple to the 0th order plane wave in Region III, it has the same magnitude as the input plane wave. This translates into a 100% transmission. The phase of the transmission light will be determined by the phase of $\sum_m (a_m + b_m)\Lambda^{-1}\int_0^\Lambda \mathcal{E}_{x,m}^{II}(x)e^{j(k_0 sin\theta)x}\,dx$.

So far, we focus our discussion at the HCG exit plane. When 100% transmission happens, there is also interesting phenomenon at the HCG input plane. 100% transmission means 0% reflection; this implies that the input plane wave would fully couple into the grating. This means the DC Fourier component of the grating modes should equal to that of the input plane wave:

$$\sum\nolimits_m \left[a_m e^{+j\beta_m t_g} + b_m e^{-j\beta_m t_g}\right]\Lambda^{-1}\int_0^\Lambda \mathcal{E}_{x,m}^{II}(x)e^{j(k_0 sin\theta)x}\,dx = \gamma_0/k_0\sqrt{\mu_0/\varepsilon_0} \qquad (2.28)$$

Note that the left side of Eq. 2.28 does not contain a modulus sign, meaning not only the magnitude but also the phase should match. This results from the fact that in order to vanish any reflection at the input plane, the impedance has to be completely matched, not only magnitude, but also phase.

To testify the above concept, at the two-mode region, we plot the magnitude and phase spectrum of the two grating modes' lateral average (weighted by the tilted wavefront $\boldsymbol{e^{j(k_0 sin\theta)x}}$) at the HCG exit plane and input plane, shown in Figure 2.14. The HCG parameters are $n_{bar} = 3.48$, $s/\Lambda = 0.6$, $t_g/\Lambda = 1.38$, $\theta = 50^o$, TM polarization. It is clearly seen that at 0% reflectivity points (i.e. 100% transmission points), the two grating modes' lateral average matches with the input plane wave at the HCG input plane, both the magnitude and phase; whereas at the HCG exit plane, its magnitude is the same as that of the input plane wave, indicating 100% transmission.

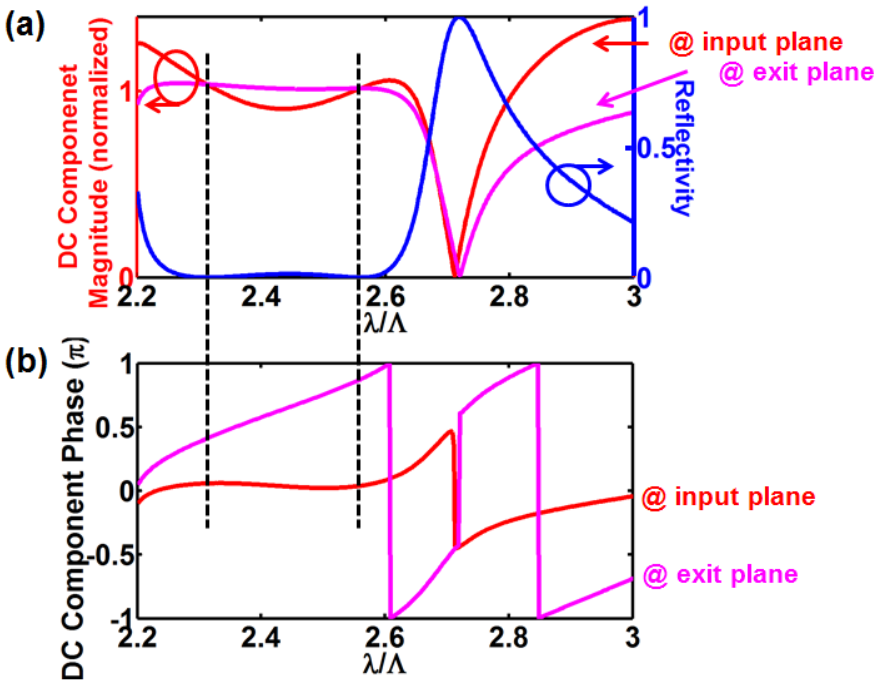


Figure 2.14 Magnitude (a) and phase (b) spectrum of the two grating modes' lateral average at the HCG exit plane, and at the input plane, compared with the reflection spectrum. At the 0% reflection points (i.e. 100% transmission points), marked by the two vertical dashed lines, Eq. 2.27 and 2.28 are satisfied.

2.3.3 Crossings and Anti-crossings

The phenomena of band anti-crossings are commonly observed in physics and, more specifically, optics. Examples include coupled quantum-dot cavities [61, 62], semiconductor alloys [63], photonic crystals [64], optical modulators [65], and many others. In all cases, anti-crossings are known to be indicators of strong coupling, while crossings typically indicate weak or no coupling.

Figure 2.11 shows the crossing and anti-crossings phenomena in the HCG reflectivity contour plot ($t_g - \lambda$ diagram). Within optics and electromagnetism, crossings and anti-crossings are typically presented in $\omega - \beta$ diagrams, rather than in a $t_g - \lambda$ diagram, which we chose because the effect can be more clearly seen. However, the two representations are indeed very similar, since as we mentioned above, $\psi_m = n\pi$ are merely generalizations of the simpler condition $t_g = n\pi\beta_m^{-1}$, and therefore the resonance curves in Figure 2.11 are equivalent to a $\beta^{-1} - \omega^{-1}$ diagram. Hence, all intuitions behind crossings and anti-crossings in standard texts apply.

Figure 2.15(a) shows the resonance curves in the t_g-λ diagram, for an HCG with $\eta = 0.70$, $n_{bar} = 3.48$, TE polarization light, surface normal incidence. It is clearly seen that when two resonance curves intersect each other, anti-crossing happens if $|\Delta\psi| = |\psi_0 - \psi_2| = 2l\pi$; whereas crossing happens if $|\Delta\psi| = |\psi_0 - \psi_2| = (2l+1)\pi$. We expect high-Q resonance phenomenon around the anti-crossing points: when the two

intersecting resonance lines are in-phase, the FP resonances have the same *parity* (both odd, or both even) and their intensity profiles overlap significantly, which leads to strong coupling between them and results in an anti-crossing. In contrast, an opposite parity between the two resonances leads to a weak coupling, resulting in a crossing.

Figure 2.15(b)(c) present the intensity profiles in the grating for the cases of an anti-crossing and a crossing, respectively. Anti-crossing corresponds to a very strong resonance, where a very significant (10^7-fold) intensity buildup occurs within the grating. On the other hand, crossing is shown to correspond to a fairly weak (25x) energy buildup.

Figure 2.16 shows the field profile of the anti-crossings conditions, with the same HCG parameters in Figure 2.15. In general, the number of the fields' zero crossing in the z direction is related to the phase value of the higher order resonance curve; whereas the number in the x direction is related to the larger number of m_1 and m_2. By a careful design, large field enhancement can occur either in the air gap or the grating bar. This opens up various applications such as sensing [35] or surface emitting laser [36].

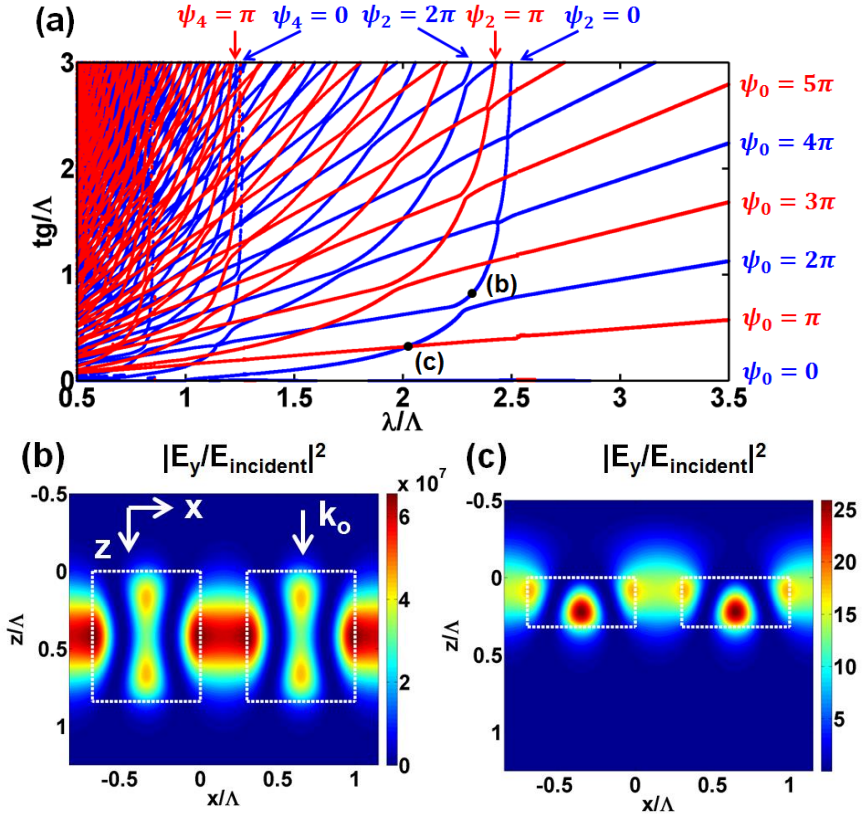


Figure 2.15 (a) Resonance lines on t_g-λ diagram. Depending on $|\Delta\psi|$, the intersecting resonance curves either crosses or anti-crosses. (b) Intensity profile inside the grating for an anti-crossing, showing 10^7-fold resonant energy buildup. (c) Intensity profile for a crossing, showing only weak energy buildup. The HCG conditions for (b) and (c) are labeled on t_g-λ diagram in (a). HCG parameters: η=0.70, n_{bar}=3.48, TE polarization light, surface normal incidence.

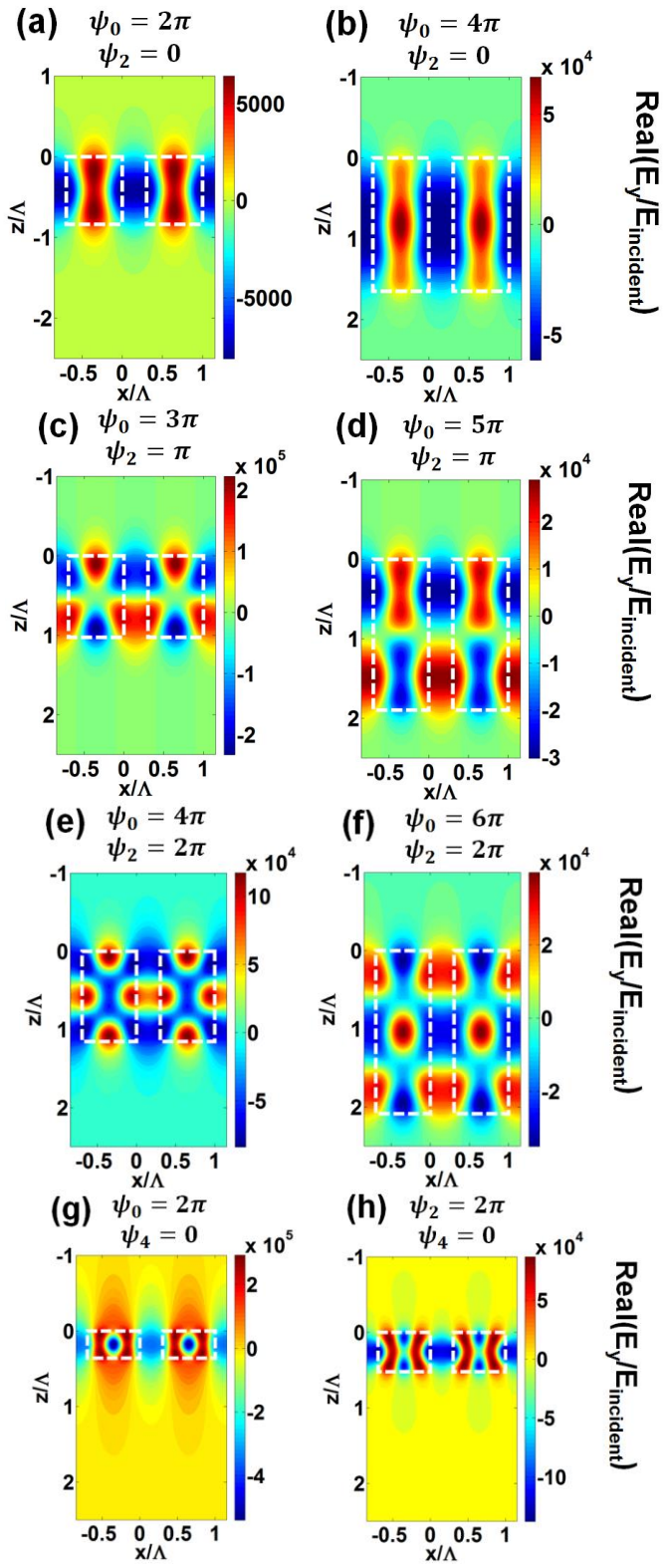


Figure 2.16 Field profile [real($E_y/E_{incident}$)] of various HCG resonance at anti-crossings. The resonance conditions are labeled in the figures. HCG parameters: η=0.70, n_{bar}=3.48, TE polarization light, surface normal incidence.

2.3.4 Resonator without mirrors

As seen from the above discussion, the resonance curves shown in Figure 2.11 and Figure 2.15 represent the HCG resonance. The HCG resonance mechanism is greatly different from a plain, simple FP cavity. In a simple FP cavity, the FP modes represent *orthogonal* longitudinal modes of a cavity, which do not couple into each other at the exit planes (reflectors). The uniquely interesting case of the HCG resonator is that the propagating modes inside the grating actually *strongly couple into each other* at the input/exit planes due to the *abrupt change in the spatial distribution of refractive index*. This makes the FP analysis a matrix one rather than scalar. We will have a more insightful and quantitative discussion on the mechanism of HCG resonance in this section.

The HCG resonance occurs when the supermode is self-sustainable. For a self-sustainable supermode $\mathbf{C}$, it satisfies the round trip condition $\mathbf{C} = \boldsymbol{\rho\varphi\rho\varphi C}$, or equivalently: $(\mathbf{I} - \boldsymbol{\rho\varphi\rho\varphi})\mathbf{C} = \mathbf{0}$, $\mathbf{I}$ being a unit matrix . Clearly, for a non-zero solution to exist on such equation, the following condition must be satisfied:

$$det[I - (\rho\varphi)^2] = 0 \tag{2.29}$$

Eq. 2.29 is the HCG resonance condition. The ability to reduce the elaborate nature of HCGs into a single "half-roundtrip" matrix $\boldsymbol{\rho\varphi}$ is a significant advantage of Eq. 2.29. When this requirement is satisfied, the boundary conditions at both HCG input and exit planes are satisfied solely by the internal supermode, leaving no need for outside excitation. This description is simply a matrix-form paraphrase of the already well-known round-trip description of a regular Fabry-Perot resonance.

While Eq. 2.29 is useful to identify wavelength and designs for an idealized ("infinite" Q) resonance, a more general expression is used to find the finite HCG quality factors [9, 58]. The supermode with the highest Q_j will dominate and represent the Q value of the entire structure. Q_j value of each HCG supermode can be expressed as:

$$Q_j = 2\pi t_g\left(n_{gr}/\lambda\right)\left|r_j/\left(1 - r_j^2\right)\right| \tag{2.30}$$

The supermodes are Eigen-solutions of HCG, and therefore r_j in Eq. 2.30 is the *j*th eigenvalue of the "half-roundtrip" matrix $\boldsymbol{\rho\varphi}$. n_{gr} is the effective group index of the FP cavity. The overall HCG Q factor is the *maximal* among the Q_j in Eq. 2.31:

$$Q_{HCG} = max\, Q_j \tag{2.31}$$

There are two major advantages of using an HCG resonator over traditional FP resonator. First of all, the entire cavity can be ultra-thin, much less than half wavelength. Secondly, the HCG modes couple into each other upon reflection from the HCG input and exit planes. This coupling allows a completely excellent optical resonator without traditional high reflectivity mirrors!

2.4 HCG Band Diagram

Although HCG is a 1D grating and therefore resembles a conventional DBR or 1D photonic crystal, the physics is distinctly different. As the incident angle θ increases from 0^o to 90^o, the incidence beam propagation direction for a transverse HCG has a significant projection along the periodicity direction, and the grating periodicity places an important role on the boundary conditions. However, this is completely different for a longitudinal HCG, where the incidence direction has no projection at all along the periodicity direction. And yet, both transverse and longitudinal HCGs exhibit very similar, extraordinary properties, which were not discovered in photonic crystals or conventional gratings. The unique characteristics of HCG come from the rich interactions of the waveguide-array modes and large contrast in refractive indices.

The strong dependence on HCG thickness and duty cycle further distinguishes an HCG from a photonic crystal or conventional gratings. In the HCG analytic treatment, we consider incident waves coming from Region I and having a real propagation constant (k_z) in z-direction ($k_{x0} < k_0$ and thus $k_z^2 > 0$ in Region I with $k_{x0}^2 + k_z^2 = k_o^2$). In photonic crystals, on the other hand, the consideration is for waves incident and propagating in the x-direction, while exponentially decaying in z-direction ($k_{x0} > k_0$, and thus $k_z^2 < 0$ in Region I). Hence, a transverse HCG and 1D photonic crystal is distinguished from each other by the different operating regimes. This can be best illustrated by the HCG band diagram, as shown in this section.

We start with the reflectivity contour plot of a transverse TE HCG with an incident angle of 50^o and duty cycle of 0.45, as shown in Figure 2.17. This is the same as Figure 2.11 (b), but with a different color code for a better display of the resonance curves and labeling. We choose an HCG thickness $t_g = 0.89\Lambda$ and draw a horizontal line, which cuts across the various resonance curves of different modes, at different wavelengths. The crossing point regions are labeled as A~D. In region A, $t_g = 0.89\Lambda$ crosses the 1^{st} order resonance ($\psi_1 = 0$); in region B, 0^{th} order resonance ($\psi_0 = 2\pi$); in region C, 1^{st} order resonance ($\psi_1 = \pi$); in region D, 0^{th} and 1^{st} order resonance ($\psi_0 = 3\pi, \psi_1 = 2\pi$). These crossing points would later become part of the band diagram. If we repeat the same process for all the incident angles from 0^o to 90^o, and trace the wavelength of these crossing points for different incident angle, then we can plot out the whole photonic bands.

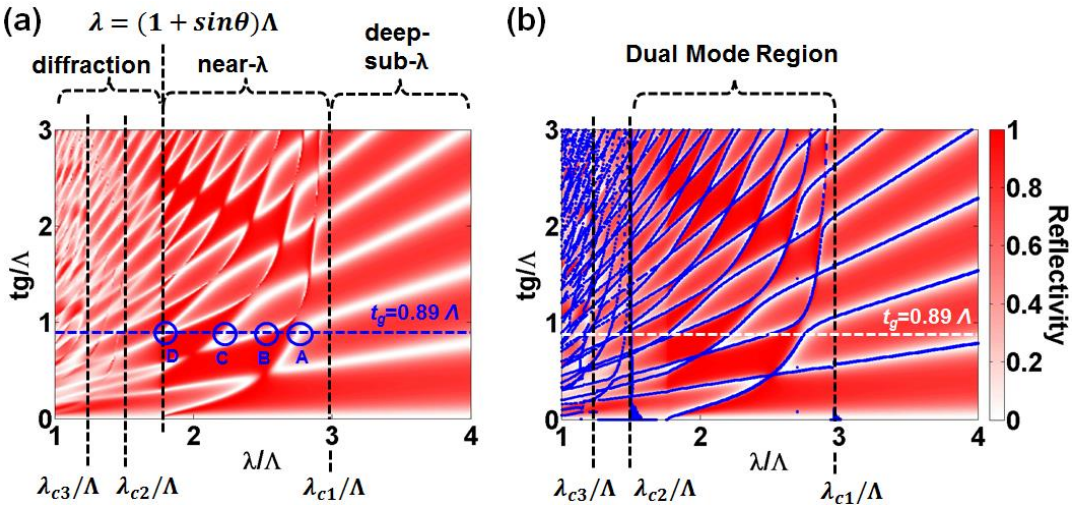


Figure 2.17 (a) Reflectivity contour of a TE HCG as a function of wavelength and grating thickness. The incident wave is TE polarized, with an incident angle θ=50^o (b) Analytical solutions of Fabry-Perot resonance conditions of the individual supermodes, shown by the blue curves, superimposed on the reflectivity contour in (a). A horizontal line t_g=0.89Λ is plotted for the chosen thickness, which cuts across various resonance curves of different modes. The crossing point regions are labeled A–D. The incident angle is then varied from 0^o to 90^o, and these crossing points can be traced along the wavelength, forming a photonic band, as shown in Figure 2.18 (a).

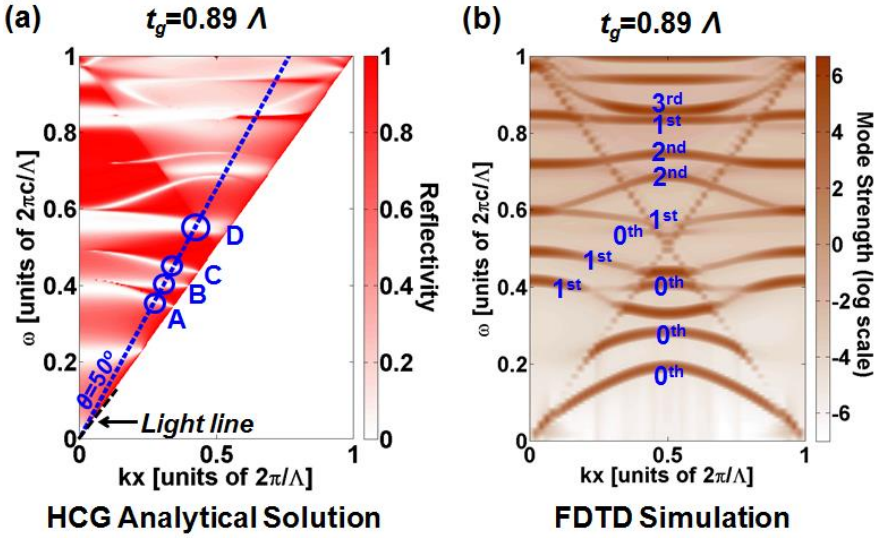
The above process can be simplified into the calculation of the HCG reflectivity versus different optical frequencies and incident angles θ, for a fixed HCG thickness and duty cycle ($t_g = 0.89\Lambda$ and $\eta = 0.45$ in this case). With θ subsequently converted into k_{x0}, the HCG reflectivity can be plotted versus frequency and k_{x0} (for simplicity, we label k_{x0} as k_x in this section), as shown in Figure 2.18 (a). This is essentially a band diagram with different photonic bands illustrated with lines across which sharp reflectivity change happens or lines with full transmission. We can associate these different photonic bands with the mth order supermode resonance in HCG. A line $k_x = (\omega/c)sin50^o$ is drawn in Figure 2.18 (a), corresponding to an incident angle of 50^o. This is essentially the horizontal lines shown in Figure 2.17, but with a different scale in wavelength / optical frequency. This line cuts across the various photonic bands; the crossing point regions A~D corresponds to those in Figure 2.17. These discrete bands above the light line shown here indicate the resonating conditions, where relatively high Q can be achieved. One interesting result from Figure 2.18 (a) is that an HCG can be designed to have a very wide-angle high reflection or transmission in a certain wavelength range. Note that only the photonic bands above the light line can be calculated, simply because we use a plane wave to excite these modes from the top of HCG.

Alternatively, a full band diagram can be obtained by FDTD simulation with Bloch boundary condition on a single HCG period. This is simulated by placing random dipole sources in the unit cell to excite all the possible modes supported by the HCG. The time domain signal from FDTD is recorded, and subsequently Fourier transformed into frequency domain. The Eigen modes are indicated by the large peaks in the frequency

signal. The Bloch boundary conditions are swept from $k_x = 0 \sim 2\pi/\Lambda$, and for each condition, the frequency domain signals are recorded. The results are shown in Figure 2.18 (b), which has the same HCG parameters with Figure 2.18 (a). The band diagram is symmetric between $k_x = \pi/\Lambda$ because of Brillouin zone folding. Above the light line, this band diagram shows exactly the same band structure as in Figure 2.18 (a). This strongly indicates that the photonic bands originate from the resonance of the HCG supermodes. Below the light line lie the bands typically seen in (1D) photonic crystal ($k_x > k_o$). Given that the band is continuous across the light line if it appears both above and below light line, the HCG supermode resonance can be used to signify the whole photonic band, not only above the light line, but also below the light line!

Figure 2.18 (a) and (b) provides a new perspective on the photonic bands: the mth order band originates from the mth order HCG supermode resonating in the *z* direction, as described in Eq. 2.21. For the same mth order HCG supermode, there are various resonating conditions corresponding to different longitudinal Fabry-Perot modes (*n*=0, 1, 2…in Eq. 2.21). This translates to different subbands belonging to the same mth order band. The thicker the HCG is, the denser the subbands would be, as a result of a smaller free spectral range. This is illustrated in Figure 2.18 (c), where the HCG thickness is 1.5Λ.

A special case is when the HCG thickness increases to infinity. This, in fact, is the case for a *pure 1D* photonic crystal. Lacking the boundaries at *z*=0 and $z = t_g$, there are no longer mode interaction and reflection. Thus the HCG supermodes resonating in the *z* direction reduce themselves to the pure waveguide array modes. It is interesting to note that these waveguide array modes do not propagates in the *z* direction, and thus they are at the mode cutoff conditions (β=0). This corresponds to the typical consideration in the pure 1D photonic crystal, where the wave vector in the *z* direction is zero. Instead of there being a large amount of subbands, there would be only one single band for each mth order mode [Figure 2.18 (d)], which corresponds to the cutoff frequency trace for the mth order waveguide array modes.



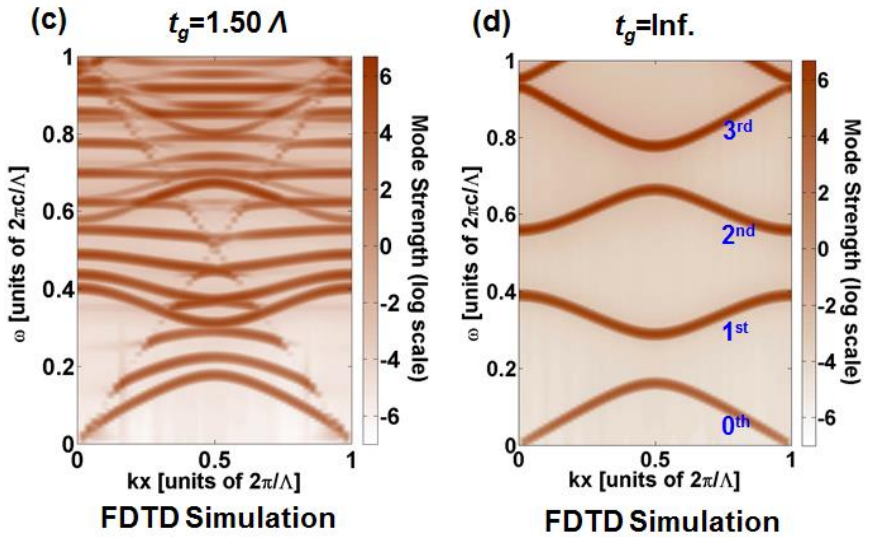


Figure 2.18 Band diagram analysis of HCG and 1D photonic crystal. (a) HCG band diagram calculated with HCG analytical solution, for t_g=0.89Λ. The photonic bands are signified with lines across which sharp reflectivity change happens or lines with full transmission. A dashed curve indicating θ=50^o is plotted and crosses the band. This curve corresponds to the horizontal curve t_g=0.89Λ in Figure 2.17(a). The crossing point regions are labeled A–D, corresponding to those in Figure 2.17(a). (b) Full band diagram simulated with FDTD for HCG thickness t_g=0.89Λ. The bands are symmetric along k_x=0.5(2π/Λ) due to Brillouin zone folding and the light line is indicated by the dotted line. In comparison with (a), one can associate different bands with different orders of supermodes in HCG. Because of the low quality factor of the zeroth-order supermodes above the light line, they do not show up clearly above the light line in (b). (c) The FDTD simulated band diagram for an HCG thickness of t_g=1.5Λ. (d) The FDTD simulated band diagram for a pure 1D photonic crystal, where t_g=∞. HCG parameters: n_{bar}=3.48, η=0.45, TE HCG.

This band analysis unifies the HCG and the 1D photonic crystal analyses within the same theoretical architecture. Figure 2.19 summaries the connections between HCG and photonic crystal. As previously discussed, the diffraction region is $\lambda_D \leq (1 + sin\theta)\Lambda$, which can be translated to frequency $\omega_D/(2\pi c/\Lambda) \geq 1 - k_x/(2\pi/\Lambda)$. Thus the diffraction line is basically the folded light line, and the diffraction region is the region above the light line and the folded light line. The HCG operates above the light line but below the diffraction line; whereas photonic crystal typically operates below light line. The dual mode region and the cutoff frequency traces for HCG 1st, 2nd, 3rd and 4th supermodes are also plotted. These cutoff frequency traces are the same as the 1D pure photonic crystal bands in Figure 2.18 (d).

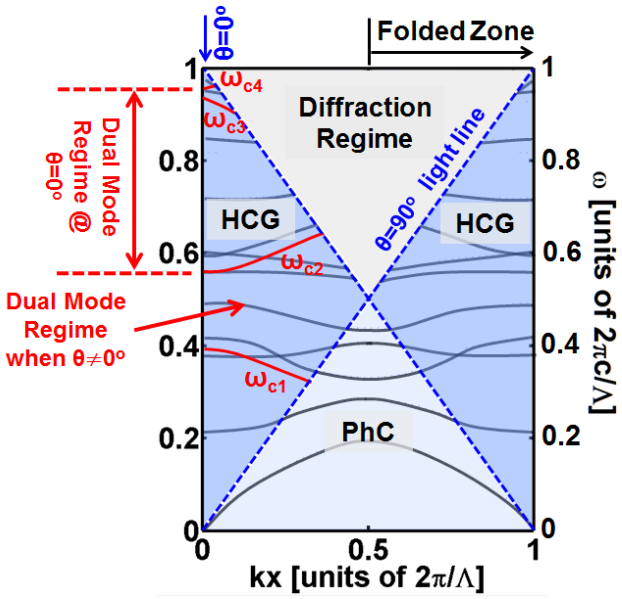


Figure 2.19 Schematic showing the different operation regime for HCG and photonic crystal (PhC), separated by the light line. The bands are extracted from Figure 2.18 (a)(b). The diffraction line is basically the folded light line, and the diffraction region is the region above the light line and the folded light line. The HCG operates above the light line but below the diffraction line; whereas photonic crystal typically operates below light line. The dual mode regime for HCG with cutoff frequencies for HCG 1st, 2nd, 3rd and 4th order supermodes are also shown. HCG parameters: n_{bar}=3.48, η=0.45, t_g=0.89Λ, TE HCG.

The supermode formulation and band diagram developed for HCG is more powerful than the FDTD analysis typically used for photonic crystal. The resonance modes manifest themselves as full transmission, as in a FP resonance or, most peculiarly, as sharp transient from high to low reflections and vice versa as in a Fano resonance. All the resonance modes can be revealed, whereas in the FDTD band analysis, typically only those bands associate with high Q resonances can be recorded. The HCG band diagram provides not only intuitive physical insights, but also quantitative information on the reflection and transmission coefficients. This is especially useful to design the various optical components with different functionalities based on HCG. Figure 2.20 shows the band diagram for various HCG with different thickness and duty cycle, from where one can find HCG designs for omni-directional high reflection, omni-directional high transmission, and boardband reflection etc.

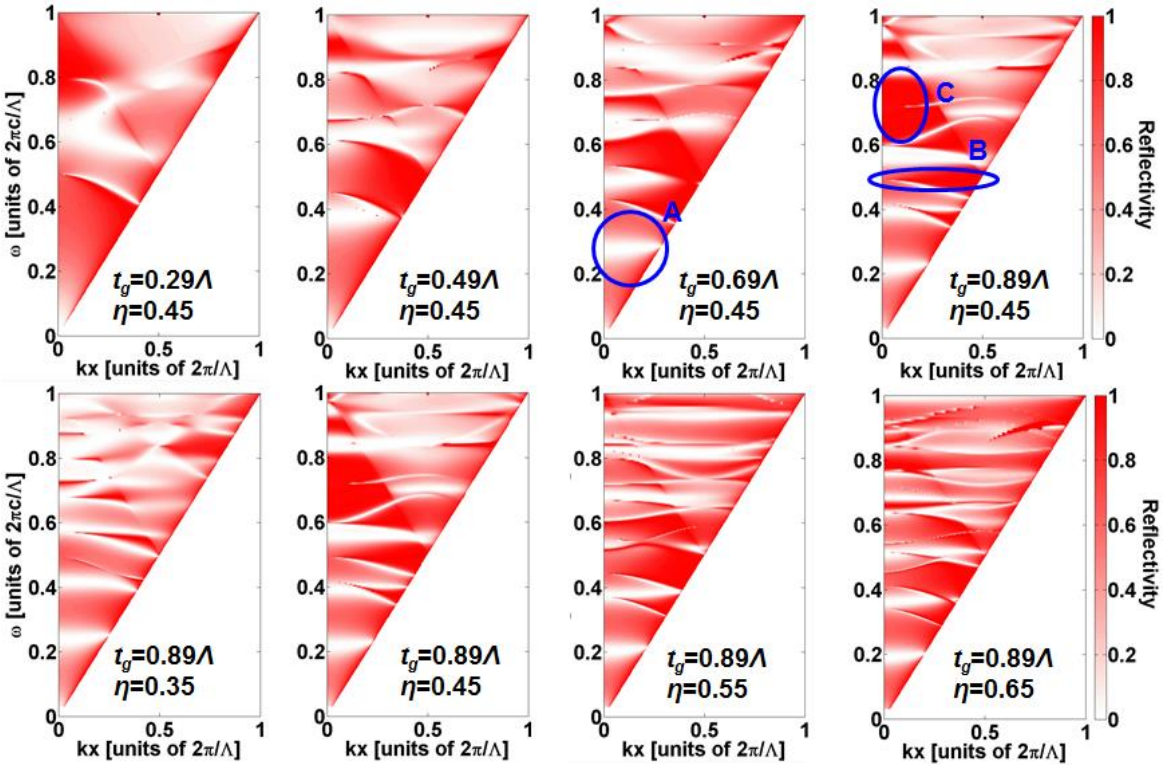


Figure 2.20 Band diagrams for various HCG with different thickness and duty cycle. HCG designs for different optical functionalities can be found, such as the omni-directional high transmission, omni-directional high reflection, and boardband reflection etc, indicated by the circles A, B, and C respectively. HCG parameters: n_{bar}=3.48, TE polarized.

2.5 Longitudinal HCG

In the previous discussion, we focus on the transverse HCG, i.e. the light propagation direction is perpendicular to the grating bars, $\varphi = 0$. The analytical treatment of the longitudinal HCG, i.e. $\varphi = 90^0$, is similar. To solve the electromagnetic property of the longitudinal HCG, we can first find the Eigen modes of each region, and then match the boundary conditions at the interface between different regions. In contract to the transverse HCG where there is only three non-zero field components (e.g. H_y, E_x, E_z for TM polarization; and E_y, H_x, H_z for TE polarization) inside the HCG, all the six field components are non-zero in the longitudinal HCG. This increases the complexity in finding the Eigen modes. The general solution would be to first write down the expression of H_y and E_y of the Eigen mode in each region, and then use them to derive the expression of other field components with Maxwell equation. The rich characteristics of HCG, such as broadband reflection, high Q resonance, still remain. The distinct difference from the transverse HCG is that the diffraction region will be $\lambda \leq \Lambda$ regardless of the incident angle θ. This expands the near-wavelength region, where most of the peculiar phenomena happen.

2.6 Summary

In this chapter, the analytical treatment of the HCG is presented and discussed. In general, the key to solve or analyze the electromagnetic properties of a specific structure is to find a set of modes which can serve as basis to express a more general mode in the structure. The selection of this mode basis is important, and one should always try to choose the Eigen modes as this mode basis. In this way, the physical picture would be clear and the analysis would be simple. As a general solution, RCWA uses plane wave as the mode basis, which is general enough so that it can tackle all the periodic structure. The drawback is that the plane waves are usually quite different from the Eigen modes of the structure, and thus many orders are required to construct the mode. The presented analytical solution uses waveguide-array modes as the mode basis, which captures the physics of the HCG. This makes the subsequent analysis straight forward and leads to the intuitive explanations of the various phenomena. Another beauty of this analytical solution is that its formulation can be formatted in matrix form. Matrix algorithm can thus be used, and this greatly simplifies the calculation. Though the analytical solution is derived to solve one single layer of HCG, it can be readily expanded to more complicated structures, such as multilayer HCG structure. More elegantly, a transmission matrix method is proposed to solve these general HCG structures. This is very similar to the transmission matrix method used to solve the multilayer structure where each layer is homogenous material. However, instead of there being just a plane wave in each layer, there is a set of Eigen modes (such as waveguide array modes in the HCG layer) for each layer, and the size of the transmission matrix is 2M x 2M instead of 2 x 2, where M is the number of Eigen modes.

The HCG properties are governed by the interaction of the waveguide-array mode, or more concisely, the interference between the different supermodes. The extraordinary features of the HCG locate at its two mode region. It is this two modes interaction lead to its 100% reflection and 100% transmission properties. More interestingly, the HCG can be a high Q resonator – a high Q resonator without mirrors! In other words, it is the mode interference at the HCG input and exit plane makes these two planes a perfect reflective mirror. It should be note that the rich properties of the HCG originates from the near-wavelength operation region as well as a large index contrast between the grating bars and its surrounding media.

A band diagram is developed for transverse HCG. The bands refer to the supermodes of the HCG. The band diagram shines light on a better understanding on the difference between 1D photonic crystal and HCG. HCG operates above the light line, whereas 1D photonic crystal typically operates under the light line. The operation of the HCG largely depends on the mode interface. The light propagation direction can be either perpendicular or in parallel to the grating periodicity; whereas in photonic crystal, the light propagation direction is always in parallel to the periodicity.

Chapter 3
High Contrast Grating based Hollow-core Waveguide and Its Application

The rich characteristics of HCG enable a wide range of applications. The most intuitive one is the replacement of DBRs, where broadband high reflectivity is desirable. One of the applications is hollow-core waveguides (HCWs).

Guiding light through a low-index core surrounded by high refractive index cladding layers has emerged as a new tool for applications, in contrast to conventional high-index waveguides. In particular, hollow-core optical waveguides / fibers are desirable for gas sensors and gas-based nonlinear optics because of the increased lengths for light-matter interaction [66, 67], and for laser surgery to guide light in mid- to far- infrared wavelength regimes that lack low-absorption materials [68, 69]. Chip-scale hollow-core waveguides are desirable because they enable cost-effective manufacturing of on-chip systems with the potential to monolithically integrate light sources, detectors and electronics. In particular, due to the elimination of the core material, light propagation properties in HCW are temperature insensitivity. This opens up a lot of applications in optoelectronics, such as precise optical delay lines, and RF filters. Chip-scale HCWs have been reported using metal [70], distributed Bragg reflectors [71, 72] and anti-resonant reflection layers [73, 74] as the guiding reflectors. However, their use is limited due to large optical losses because of insufficient reflection.

HCG provides a superiorly high reflectivity, making it an ideal guiding reflector in HCWs. As mentioned before, light propagation direction can be either parallel or perpendicular to the grating bars; this leads to two basic types of HCW configurations, as shown in Figure 3.1. The optical beam is guided by zig-zag reflections from the guiding walls – an intuitive ray optics model for HCW. The HCG provides a strong optical confinement in the transverse direction; novel scheme is needed for lateral light confinement. In this chapter, we demonstrate various HCG HCWs with different novel lateral confinement schemes. The first one is constructed by just two planar, parallel, silicon-on-insulator wafers with HCGs. By spatially chirping the HCG dimensions, two-dimensional light confinement (both in transverse and lateral direction) can be achieved though there is only one-dimensional confinement in the geometry (transverse direction). In other words, lateral confinement is realized without any physical sidewalls! We experimentally demonstrated that its propagation loss is as low as 0.37 dB/cm [46], the lowest for HCW mode-matched to a single mode fiber. The second type of HCG HCW involves placing DBR on the side for lateral confinement. The core size of the HCW is 100 μm by 100 μm. A <0.1 dB/m propagation loss is estimated based on the loss measurement. The third type of HCG HCW is a novel cage-like HCW [47], where four HCG reflectors serve as four waveguide walls. Strong light confinement is experimentally demonstrated [47]. The HCG HCWs establish a new platform for integrated optoelectronics circuits, and facilitate various new functionalities and system-on-a-chip applications, such as gas sensing, slow light [47, 48], coupler [49], optical switch [50], etc.

In section 3.1~3.3, we discuss the three types of HCG HCWs. The optical switch based on the HCW would be discussed in section 3.4, and the advantage of using HCG HCW for gas sensing application would be discussed in section 3.5. Section 3.6 summarizes this chapter.

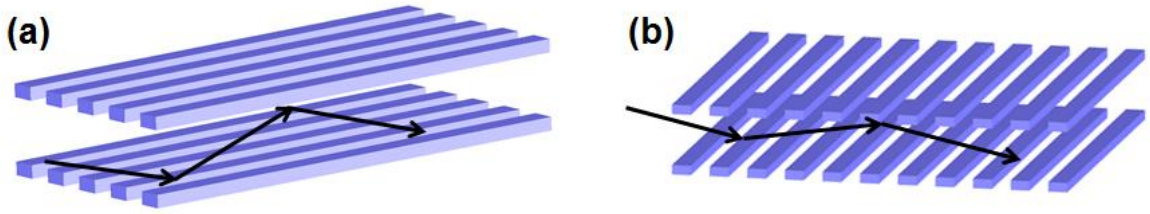


Figure 3.1 Two basic types of HCG HCWs. Light propagation direction can be either parallel (a) or perpendicular to the grating bars (b). The arrows illustrate that light is guided by the zig-zac reflection of the grating walls.

3.1 High Contrast Grating Hollow-core Waveguide with Novel Lateral Confinement

In this section, we discuss the design and experimental demonstration of a new class of two-dimensional (2D) HCW where the guided wave propagates along the HCG grating bars. The transverse guiding is provided by HCG reflections, whereas the lateral confinement is achieved by varying the HCG dimensions laterally. Thus, 2D waveguiding is achieved without any physical boundaries or sidewalls, but merely with variations of HCG dimensions to create a lateral effective refractive index. Both straight and curved waveguides are demonstrated using two parallel planar wafers, each containing a single layer of HCGs. Such a 2D waveguide has never been reported before. This is also the first experimental verification of a 2D HCG HCW design using the one-dimensional $\omega - k$ diagram [75] in combination with a very new use of the effective index method [76]. The design method is simple and intuitive. The propagation loss in a straight waveguide with a 9-μm waveguide height is measured to be 0.37 dB/cm, the lowest reported loss for an HCW with such a small core. This sidewall-less waveguide is superior in gas sensing applications compared to other HCWs. Gases or fluids can be flown into the waveguide from the side openings, instead of from the two ends of the waveguide. The cross section is thus increased by a factor of L/d, where L is the waveguide length, and d is the core height. With Fick's laws of diffusion and depending on the exact configuration, the dynamic detection speed can be increased at least by a factor of L^2/d^2, a number that is easily 8-12 orders of magnitude. The HCG HCW demonstrated here opens up a new scheme of light guiding for low-loss HCWs, allowing for an entirely new range of applications.

3.1.1 Design

To design a 2D waveguide, we should begin with a 1D slab waveguide, as shown in Figure 3.1. We focus on the case where grating bars run along the direction of guided wave, i.e. Figure 3.1(a); however the principle is the same for the other type, where the grating bars are perpendicular to the direction of the guided wave, i.e. Figure 3.1(b). The HCGs are fabricated on silicon-on-insulator (SOI) wafers with the gratings formed on a silicon (Si) layer above silicon dioxide (SiO_2), as shown in Figure 3.2 (a). The HCW is formed by placing two HCGs in parallel, separated by an air-gap d. Wave guiding is illustrated with ray optics [45], where the optical beam is guided by zig-zag reflections from the HCG. The HCG is designed to have very high reflectivity, so that the light is well confined in the x direction. In this grating design, the polarization of the electrical field is along y direction, i.e. the gratings are TM-HCGs that are designed to be highly reflective. This notation of TM is relative to the gratings and is actually orthogonal to the typical definition for waveguide modes.

With simple ray optics [45], the propagation loss and the effective refractive index n_{eff} of the fundamental mode are calculated using the following formulae:

$$Loss\ \left[\frac{dB}{m}\right] = -\frac{10tan\theta}{d_{eff}} \cdot \log_{10}|r|^2 \tag{3.1}$$

$$n_{eff} = cos\theta = \frac{k_z}{k_0} \tag{3.2}$$

Here θ is the angle between the ray and the waveguide, k_z is the propagation constant, k_0 is the wave vector of the light in free space, and d_{eff} is the effective waveguide height. d_{eff} takes into account both the physical waveguide height d and the reflection phase φ_r, which is approximately π in general. The parameter d_{eff} can be calculated by the round-trip phase condition of the fundamental mode:

$$2k_x d + 2\varphi_r = 4\pi \tag{3.3}$$

$$2k_x d + 2\varphi_r = 2k_x d_{eff} + 2\pi \tag{3.4}$$

Next, we explain the scheme to obtain lateral guiding in a 2D HCG hollow-core waveguide. For solid-core waveguides, a typical lateral guiding design employed is the effective index method [76], which uses different k_z values in the core and cladding region. Here, we also propose the same -- obtaining lateral confinement (y direction) by using different HCG designs for the core and cladding region so that the effective refractive index of the core is higher than that of the cladding [77], shown in Figure 3.2 (b). This can be achieved by varying the HCG reflection phase, φ_r, which determines the effective index n_{eff} of the 1D-slab waveguide in Eq. 3.2-3.4. To ensure overall low propagation loss, both HCGs should have high reflectivities. From a practical point of view, the HCG thickness should be fixed for both the core and cladding to facilitate the fabrication process. We choose an SOI wafer with a device layer thickness of 340 nm, and a buried oxide thickness of 2 μm. HCG reflectivity is calculated for different HCG periods Λ and silicon grating bar widths s. The incidence angle of the light on the HCG is 85.06^o, corresponding to the angle between the light ray and the normal of the HCG

reflector in a 9 μm height waveguide, and an HCG reflection phase φ_r of π. The wavelength of the light is 1550 nm, and the light polarization is TM from the perspective of the grating and TE from the perspective of the waveguide. Based on the ray optics for an HCW (Eq. 3.1~3.4), the HCG reflectivity is converted into the propagation loss, as well as the effective refractive index n_{eff} of the 1D slab waveguide's fundamental mode. Equivalently, finite element method (FEM) can be used to simulate the propagation mode of the 1D slab HCG HCW, and propagation loss and effective refractive index can be extracted. Figure 3.3 shows the contour plot of loss and effective refractive index calculated by FEM. It is seen that for the fixed HCG thickness t_g, HCG designs with different periods Λ and grating widths s can provide remarkably large differences in φ_r while maintaining a high reflectivity; this results in a variation in effective refractive index between HCG designs on a flat surface.

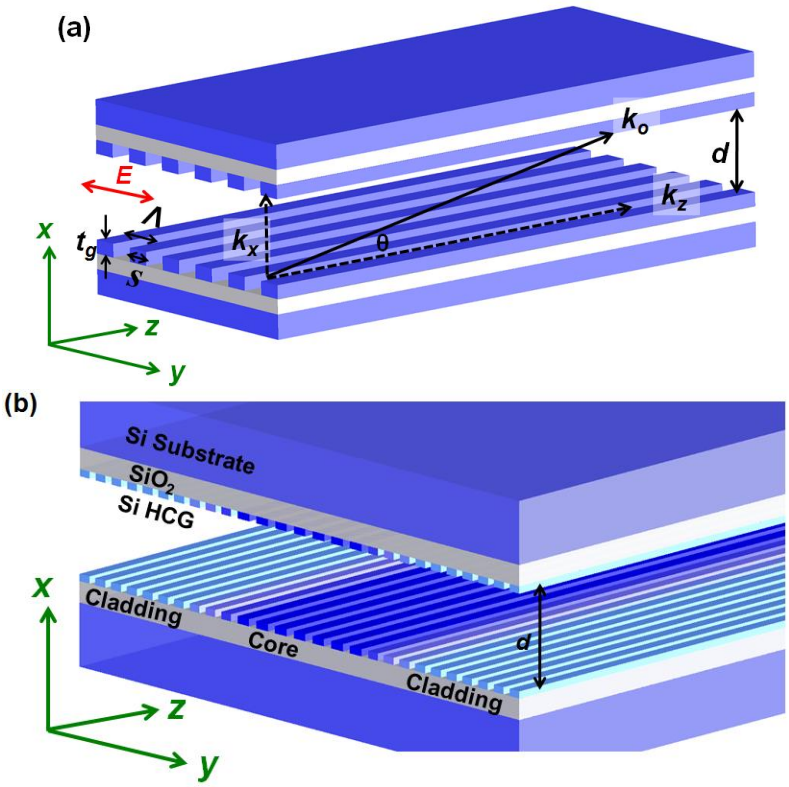


Figure 3.2 (a) Schematic of a 1D slab HCG HCW. The silicon HCG sits on top of a SiO_2 layer and silicon substrate. The two HCG chips are placed in parallel with a separation gap *d*, forming an HCW. Ray optics illustrates how light is guided: the optical beam is guided by zig-zag reflections from the HCG. The HCG is designed to have very high reflectivity, so that the light can be well confined in the *x* direction. (b) Schematic of a 2D HCG HCW. In the lateral direction, the core and cladding are defined by different HCG parameters to provide lateral confinement.

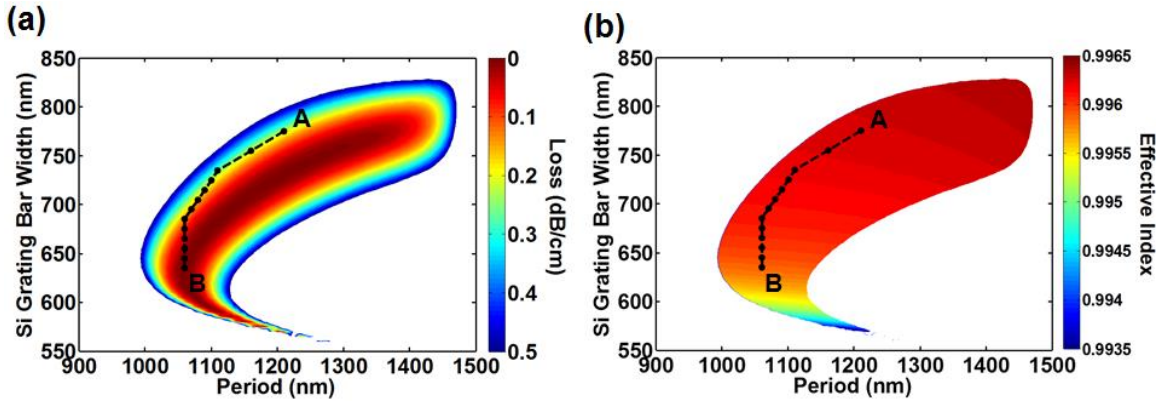


Figure 3.3 Loss contour plot (a) and effective index contour plot (b) of a 1D slab HCG HCW with a 9-μm waveguide height. The contour plots provide the design template for the waveguide. Different HCG periods Λ and silicon grating bar widths s are chosen for the core (point A) and cladding (point B), as well as the transition region (dots linked with dashed lines). The HCG thickness t_g is fixed at 340 nm, and the buried oxide thickness is set to 2 μm. The wavelength of the light is 1550 nm.

The loss tolerance on HCG thickness is very large, over a wide range of operation wavelength. Figure 3.4 shows the loss contour plot against HCG thickness and operation wavelength, calculated for a 9 μm 1D slab HCG HCW using FEM. The HCG period Λ and silicon grating bar width s is 1210 nm and 730 nm respectively. This large HCG thickness tolerance lowers the specification requirement of the SOI wafers, and makes the fabrication more feasible.

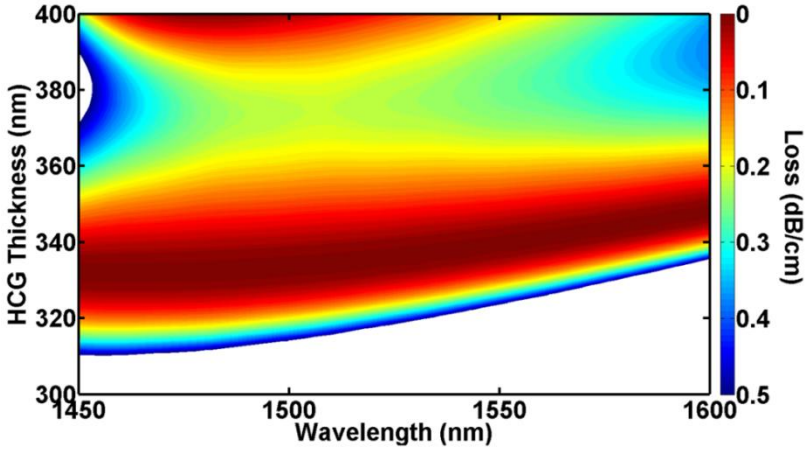


Figure 3.4 Loss contour plot against HCG thickness and operation wavelength, for a 1D slab HCG HCW with a 9 μm waveguide height. The loss tolerance is very large over a wide range of HCG thickness and wavelength. The HCG period Λ and silicon grating bar width s is 1210 nm and 730 nm respectively.

We design the HCG period and silicon bar width to be 1210 nm and 775 nm for the core region, and 1060 nm and 635 nm for the cladding region. The waveguide n_{eff} is $n_{core} = 0.99626$ and $n_{cladding} = 0.99589$ respectively for the core and cladding. The resulting effective refractive index difference between the core and cladding is $4\text{x}10^{-4}$.

Also widely known for solid core waveguides, graded-index waveguides typically exhibit lower loss than step-index waveguides [78]. A graded effective-index profile is introduced in HCWs with chirped HCG dimensions on the order of tens of nanometers, as shown in Figure 3.3. The core width W_c and the transition region width W_t are 10.9 μm and 11.9 μm, respectively. The cladding width of the waveguide is 42.7 μm on each side. This waveguide configuration ensures a single lateral mode operation, which can be evaluated by the following condition, typically used for a solid-core slab waveguide.

$$V \approx \frac{k_o \cdot W_{eff\_core}}{2} \sqrt{n_{core}^2 - n_{cladding}^2} < \frac{\pi}{2} \tag{3.5}$$

W_{eff_core} is the effective core width of the HCG HCW, which is estimated to be the sum of the core width W_c and half the transition region width W_t on both sides. In the HCG HCW design, $W_c = 10.9\ \mu m$, $W_t = 11.9\ \mu m$, and thus $W_{eff\_core} = 22.8\ \mu m$. V is calculated to be 0.4π, confirming a single mode operation.

The propagation mode profile of the 2D HCG HCW is simulated in FEM, shown in Figure 3.5 (a). The effective refractive index and propagation loss of the HCW is simulated to be 0.9961 and 0.35 dB/cm respectively at 1550 nm. The minimum loss is 0.31 dB/cm at 1535 nm. It is truly remarkable to note that, although the guided mode has very little energy in the HCGs, the effective index method can be realized with a simple and small parameter change of the HCG.

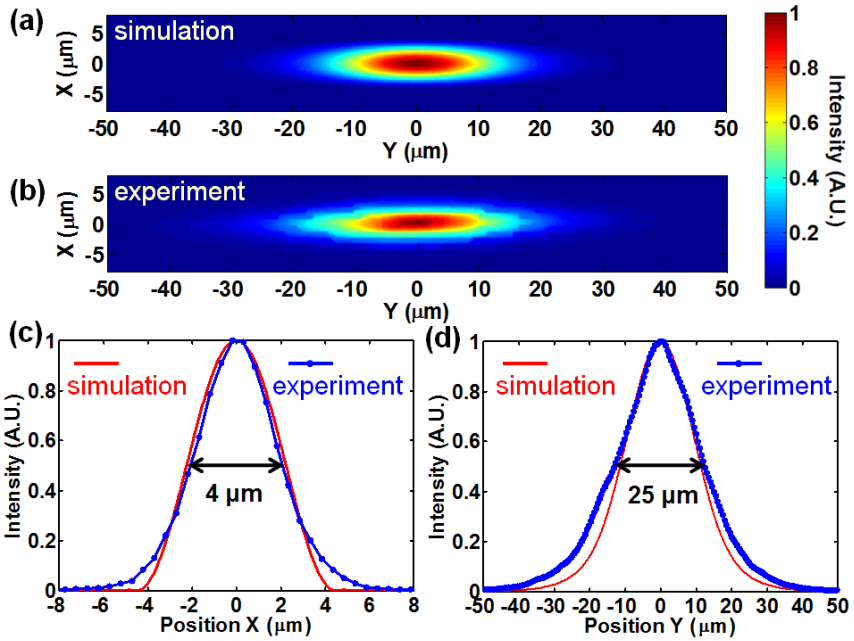


Figure 3.5 Optical mode of an HCG HCW. (a) Propagation mode profile simulated by FEM. (b) The measured mode profile from the fabricated device. (c) Transverse and (d) lateral mode profile. The simulation (red curve) agrees well with experiment (blue line). The full width at half maximum (FWHM) is 4 μm in the transverse direction, and

25 μm in the lateral direction. The wavelength in both simulation and measurement is 1550 nm.

3.1.2 Device Fabrication

The HCG HCW is fabricated using deep ultra-violet lithography on 6-inch SOI wafers, followed by a standard silicon inductively coupled plasma reactive-ion etching process. The great advantage of this lateral confinement scheme is that only a single etching step is required. Figure 3.6 shows the top-view optical microscope image of the fabricated chips (a) as well as the scanning electron microscope images of the HCG in the core (b) and cladding (c) regions. The core, transition, and cladding regions of the waveguide can be clearly distinguished under the optical microscope. The HCG grating bars have a smooth surface and a sidewall roughness of approximately 10 nm. The period and silicon grating bar width of the HCG are in agreement with the design values, and the silicon grating bar width varies by $<\pm 1.5\%$ across the 6-inch wafer. The silicon device layer thickness is measured across the 6-inch SOI wafer using an ellipsometer, showing a very uniform thickness 340 $\pm$ 1 nm. The two pieces of patterned HCG chips are then mounted onto two translation stages, and brought close to each other to form the waveguide.

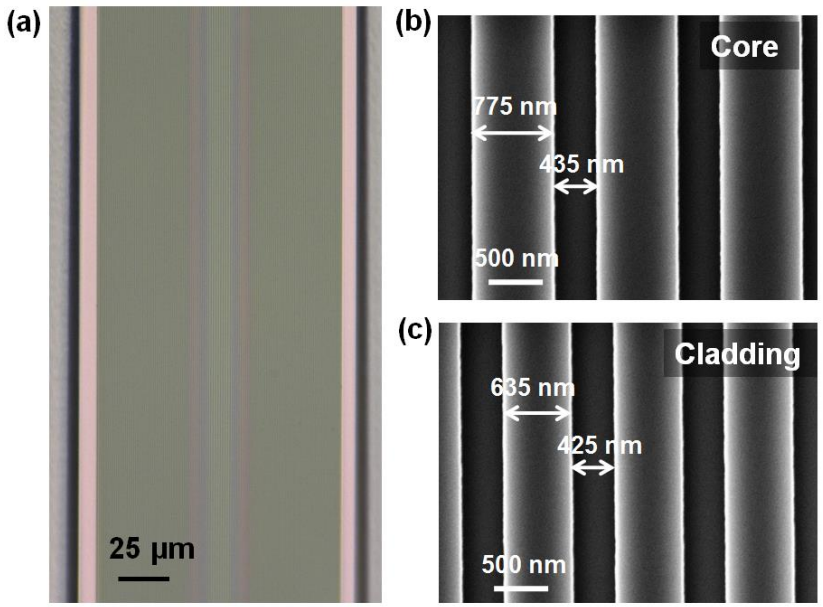


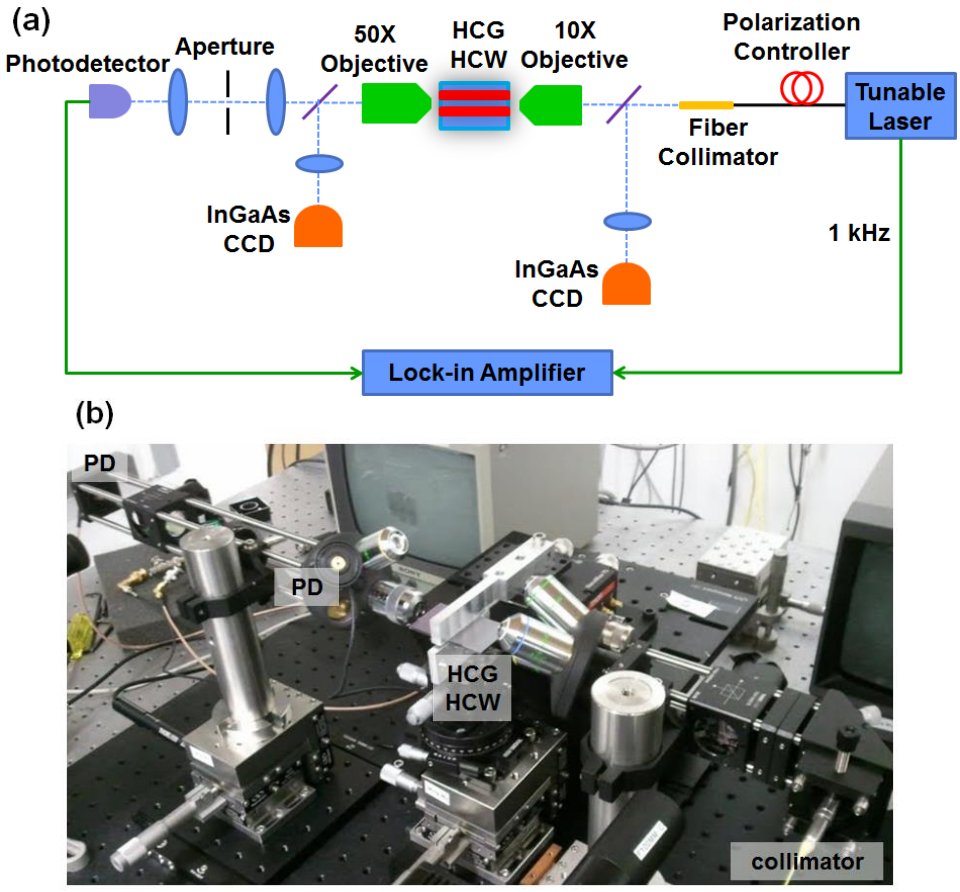
Figure 3.6 Fabricated HCG HCW chips. (a) Optical microscope image of the HCG chip. The core, transition and cladding regions are clearly distinguished by their diffracted colors. The SEM image of the HCG grating bars in the core region and cladding regions are shown in (b) and (c).

3.1.3 Device Characterization

Light guiding in the HCG HCW is experimentally confirmed by launching a laser beam into the waveguide and measuring the intensity profile at the output facet. The

experimental setup is shown in Figure 3.7. A laser beam from a tunable laser source (Agilent 81680A) is first polarization adjusted and then collimated by a fiber collimator, and launched into the HCG HCW sample by a 10X (N.A.=0.25) objective. The Gaussian beam waist size at the waveguide's input facet is optimized such that only the fundamental mode is excited. A 50X (N.A.=0.5) objective is used to collect the light for output facet imaging. With precise alignment of the two chips, an optical mode can be seen at the output facet.

Figure 3.5 (b) shows the output image with the waveguide height d set to 9 μm. The waveguide operates at fundamental mode. The measured profiles in the transverse and lateral direction are shown in Figure 3.5(c) and (d) with 4 μm and 25 μm full width at half maximum (FWHM), respectively, at a wavelength of 1550 nm. Excellent agreement is obtained between simulation and experiment.



(c)

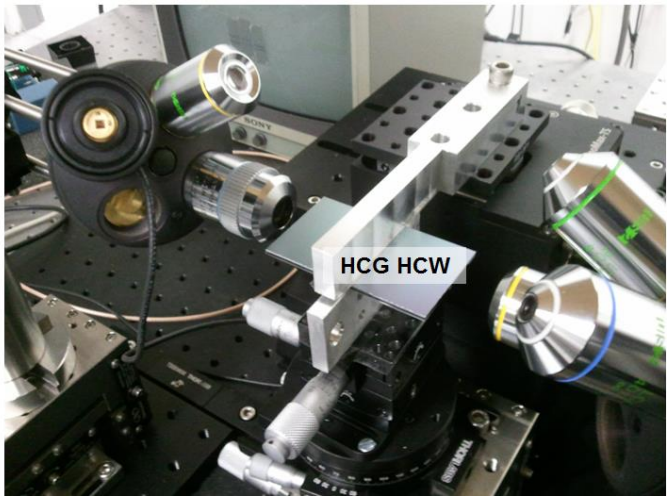


Figure 3.7 (a) Schematic of the measurement setup to characterize the mode profile and loss of the HCG HCW. (b) Image of the real experiment setup. (c) Zoom-in view of the coupling system and the HCG HCW.

Waveguides are cut into different lengths for loss measurement. To increase the signal-to-noise ratio, the laser is internally modulated at 1 kHz, and a lock-in amplifier is used in the detection system. The 50X objective is replaced with a photodetector that butt-couples the light from the waveguide in order to allow the optical power to be measured with the lock-in amplifier (Stanford Research Systems SR850). Alternatively, the 50X objective can be used to collect the light. The light is then spatially filtered by a pinhole and focused onto a photodetector. These two light detection methods are equivalent when the waveguide height d is small and lateral light leakage is negligible (detailed to be discussed in the following). However, as d increases such that lateral leakage becomes substantial, the spatial filter ensures that only the light in the waveguide is detected, which is essential for measuring the loss as a function of d.

A cut-back method is applied to extract the net propagation loss and the coupling loss of the HCG HCW. The loss spectrum of the whole optical path is first measured without the HCG HCW. The total loss spectrum is then measured for different lengths of waveguides. Due to the fabrication variation across the 6-inch SOI wafer, the HCG dimensions are not identical on all different pieces of waveguides. As mentioned in the previous section, the silicon grating bar width varies by $<\pm 1.5\%$ across the 6-inch wafer. This leads to a shift of the loss spectra (<20 nm with respect to the wavelength) between different pieces of waveguides. To improve the accuracy of the cut-back method, the loss spectra for the four different waveguide lengths are aligned in a range according to their minimum loss values. The net propagation loss is extracted for each wavelength based on a linear fit for the total loss of the four different lengths, and the propagation loss spectrum is obtained.

The waveguide height d is set to be 9 μm to facilitate the mode-matching with the optical fiber. Figure 3.8 (a) shows the measured total loss spectrum for straight waveguides with lengths of 18 mm, 38 mm, 58 mm and 78 mm. The extracted propagation loss spectrum agrees well with the results of the FEM simulation [Figure 3.8

(b)]. The minimum loss value from experiment is 0.37 dB/cm at 1535 nm, slightly higher than the simulated value. This difference is attributed to a slight warping of the two HCG chips across their length, which leads to a variation in core height on the order of ±1 μm. Furthermore, HCG surface roughness scattering may cause additional loss. The coupling loss is estimated to be 4 dB. This can be further reduced by improving the coupling design.

To better understand the loss of the 2D-confined propagation mode, we use FEM to simulate the loss spectra for a 1D slab HCG HCW with uniform HCG, both for the core and cladding design, shown in Figure 3.8(b). It can be seen that the loss spectrum of the 2D-confined propagation mode follows and agrees well with the 1D propagation mode. With further optimization of the HCG dimensions, an even lower loss can be expected.

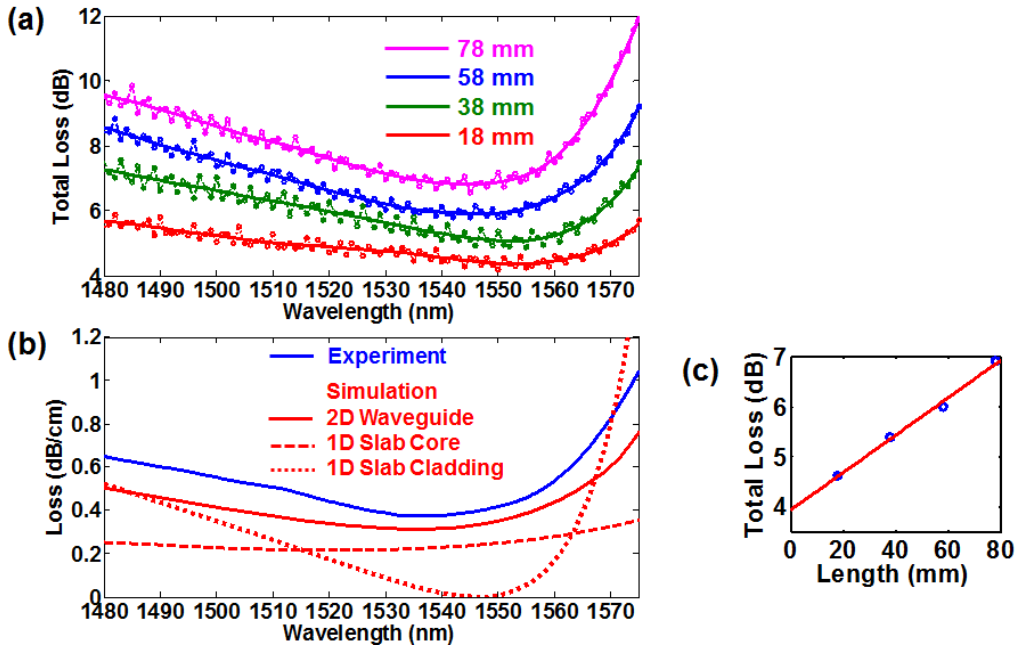


Figure 3.8 Loss spectrum of the HCG HCW for a 9-μm-high waveguide. (a) Total loss spectrum for an HCG HCW with four different lengths. The dashed dot line is the measured data. The oscillation is due to the laser and a residual Fabry-Perot cavity in the optical path of the measurement system. To remove this noise, a smoothing spline method is applied. The solid curves show the smoothed spectra. (b) The experimental extracted propagation loss as a function of wavelength (blue) and the simulated loss spectrum obtained by FEM (red, solid). The simulated loss spectrum for the 1D slab HCG HCWs with pure core (red, dashed) or cladding (red, dotted) design are also shown. (c) The linear curve fitting (red curve) to the measured data (blue dots) used to extract the propagation loss and coupling loss at 1535 nm. The slope shows the propagation loss, and the intersect at the y-axis shows the coupling loss.

3.1.4 Control of Lateral Confinement

The effective index method is the main concept for the proposed lateral confinement scheme. It is further tested and illustrated by varying the waveguide height d. As seen in

Figure 3.2 and Eq. 3.3-3.4, for a round trip in the transverse direction, the beam acquires phase through two components: interaction with the HCG (associated with a phase of $2\varphi_r$) and travel through the air trajectory (associated with a phase of $2k_x d$). Since d is the same for both the core and cladding regions, it is the HCG phase that creates the effective index difference (labeled as $\Delta n\ /n_{core}$). As d is reduced, the contribution from the HCG increases relative to the air trajectory, and thus $\Delta n/n_{core}$ becomes more pronounced. This distinction results in a stronger lateral confinement and a narrowing of the mode profile with reduced d. Figure 3.9 (a) and (b) show the experimental measured mode profiles and the lateral FWHM versus d, respectively. Excellent agreement is obtained with the simulated $\Delta n/n_{core}$ value. The propagation loss, on the other hand, is not monotonically dependent on d, as shown in Figure 3.9 (c). This is because the waveguide loss consists of both transverse and lateral components. A larger d leads to a weaker lateral confinement, and thus a larger lateral loss. However, the transverse loss decreases since it is proportional to $1/d^2$. Thus there is an optimal d that corresponds to the lowest loss at a specific wavelength.

To characterize the waveguide loss as a function of different waveguide height d, the total loss is first measured at a fixed wavelength as d varies, and then a smoothing spline curve is fitted to the loss as a function of d. This is done for each waveguide length (18 mm, 38 mm and 58 mm in this experiment); for each d, the cut-back method is then applied to extract the propagation loss based on the three different lengths, and propagation loss against d is thus plotted, shown in Figure 3.9 (c), for the operation wavelength of 1535 nm. When d is 9 μm, a minimum loss of 0.37 dB/cm is achieved, consistent with Figure 3.8. As d increases to 10 μm, the lateral loss dominates and the overall loss increases. In general, the experimental loss agrees with the simulated loss. For a larger core height, the discrepancy is more pronounced, a consequence attributed to coupling into higher order modes.

To further illustrate lateral index guiding, we fabricated various waveguides on the same chip with uniform HCG design (with the core and cladding sharing the same HCG design) and anti-guiding design (with the core and cladding designs swapped from the original configuration). The output mode profiles are presented in Figure 3.9 (d), and they show distinct differences with light dispersed in the waveguide without the appropriate HCG design. These lateral confinement measurements demonstrate the effectiveness of the effective index method for an HCW for the first time. It is truly remarkable that with little optical energy in the HCG, lateral guiding can be obtained with a planar structure. This enables light to be guided in an HCW without the aid of physical side reflectors, and opens up a new regime of optical waveguiding.

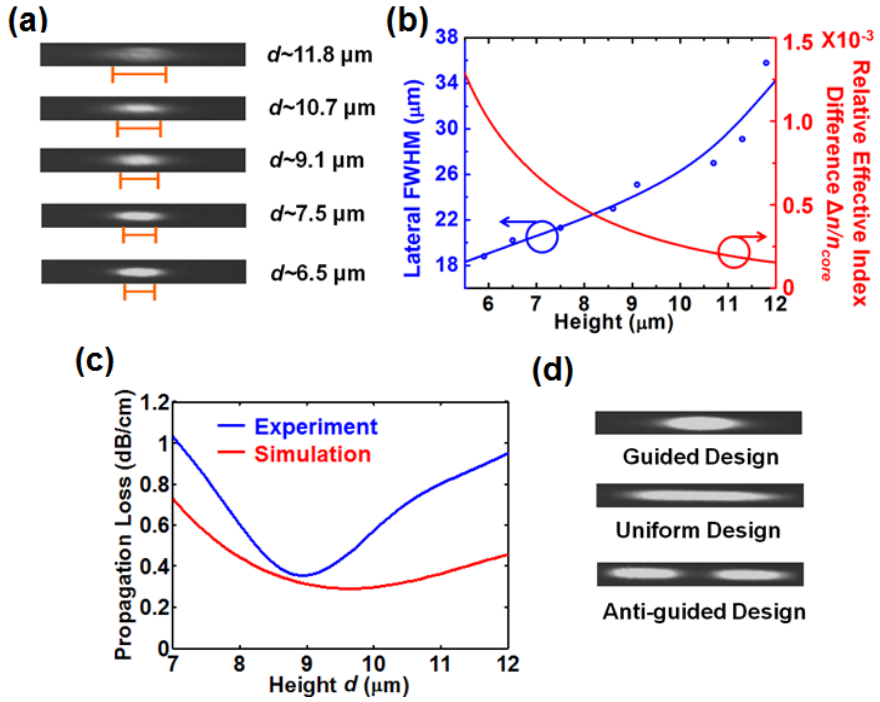


Figure 3.9 Lateral confinement in the HCG HCW. (a) Mode profile at different waveguide heights d. As d decreases, $\Delta n/n_{core}$ increases, and the mode is more confined with reduced lateral leakage. The guidelines indicate the FWHM of the mode in the lateral direction. (b) Experimental lateral FWHM of the fundamental mode versus waveguide height d (blue dots as experiment sampling points, curve-fitted with blue curve). The $\Delta n/n_{core}$ value of the fundamental mode as simulated by FEM is shown in red. The wavelength for the measurement is 1550 nm. (c) Propagation loss versus waveguide height d at a wavelength of 1535 nm. At the optimized waveguide height of 9 μm, an optimal tradeoff between lateral leakage and transverse leakage is achieved. The FEM simulated loss for the fundamental mode is also plotted, in reasonable agreement with experiment. (d) Mode profiles for three side-by-side HCWs, with lateral guiding (top), uniform design (middle) where the core and cladding share the same HCG design and anti-guided design (bottom) where the core and cladding designs are swapped from the guiding design. For the mode profiles, the output power of the mode is kept constant and d is constant ~9 μm. The image window is 140 μm by 16 μm. The wavelength is set to 1550 nm.

3.1.5 Light guiding in curved HCG Hollow-core Waveguides

The waveguide discussed so far is straight waveguide. Light can also stay guided around curves by this sidewall-less waveguides. We designed an "S-shape" and "double-S-shape" curved waveguide. Figure 3.10 (a) shows a top view of the layout. For the "S-shape" layout, curved waveguides A-A' and B-B' are parallel, and the input port of A-A' is aligned with the output port of B-B'. Light is launched into port A. Light guiding by the bend is demonstrated with the output observed in A' rather than B'. The "double-S-shape" waveguide is used to extract the bending loss, with various waveguide length combinations. The mode profile in the bending section is similar to that of a straight waveguide but with its center shifted towards the waveguide edge that is farther from the

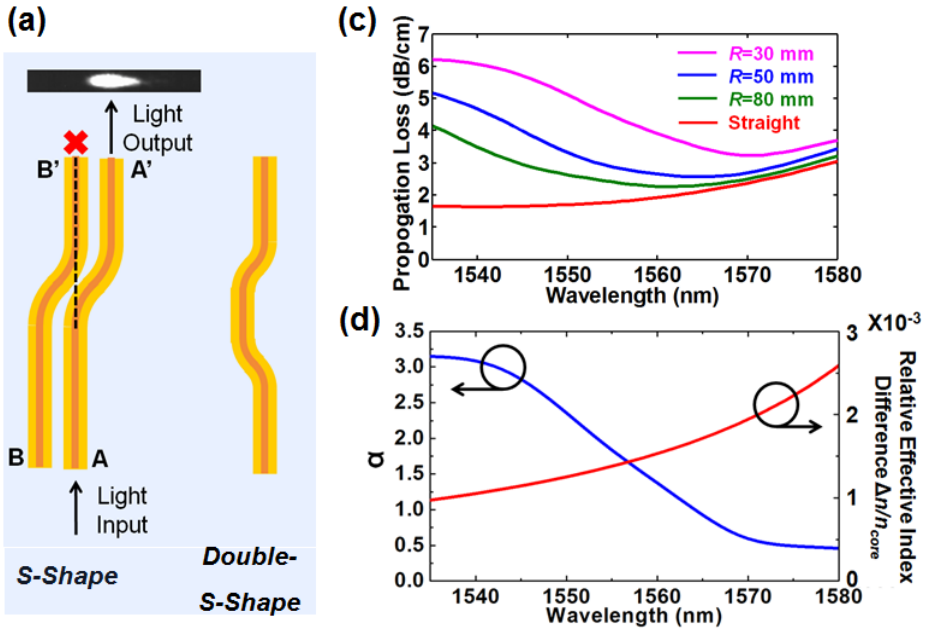
center of the curving radius, as shown in the FEM simulation results in Figure 3.11. This is the same case as a solid-core curved waveguide [79] with the conventional index guiding mechanism. After the bend, the mode would adjust its central position back to the center of the waveguide, as shown in Figure 3.10 (b), which are the near field intensity profiles at the waveguide output for the "S-shape" and "double-S-shape" waveguide with a waveguide height d~6 μm. Figure 3.10(c) shows the loss spectra for various bending radii R extracted from the 18 mm and 38 mm long waveguides, with a waveguide height d of 6 μm. The loss includes both the propagation loss and the mode coupling loss at the four bending junctures for an 18-mm long waveguide. It is seen that as the bending radius of curvature decreases, the loss spectrum slightly red-shifts. This is because lateral confinement tends to be stronger at longer wavelengths, as indicated by the FEM simulated spectrum of $\Delta n/n_{core}$ shown in Figure 3.10(d). This also explains why the loss difference between the straight and the curved waveguides becomes smaller as the wavelength increases.

In general, the bending loss can be expressed by the following formulae:

$$Loss\ [/cm] = Ae^{\frac{\alpha}{R}} \tag{3.6}$$

$$Loss\ \left[\frac{dB}{cm}\right] = 10\log_{10} A + \frac{10}{R}\alpha log_{10} e \tag{3.7}$$

Here R is the radius of curvature and A is a coefficient related to the straight waveguide loss. α indicates the degree to which light is confined in the curved waveguide, and is strongly correlated with $\Delta n/n_{core}$. A linear fit is applied to the loss as a function of R^{-1} for each wavelength, and the slope α is extracted, as plotted in Figure 3.10(d). α decreases with increasing wavelength, consistent with the simulated spectrum of $\Delta n/n_{core}$.



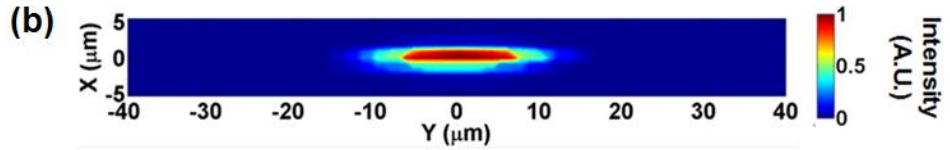


Figure 3.10 A curved HCG HCW. (a) Layout of the "S-shape" and "double-S-shape" curved waveguides. Light guiding by the bend is demonstrated with the output observed in A' rather than B' in the "S-shape" waveguide. (b) The near field intensity profiles recorded at the output of the "S-shape" and "double-S-shape" waveguide, with waveguide height d~6 μm. (c) Experimental loss spectrum for waveguides with different radii of curvature R, extracted from various waveguide length combinations of the "double-S-shape" curved waveguide layout. (d) The slope of the linear fit of loss versus R^{-1}, α, as a function of wavelength; this is consistent with the FEM simulated $\Delta n/n_{core}$ spectrum.

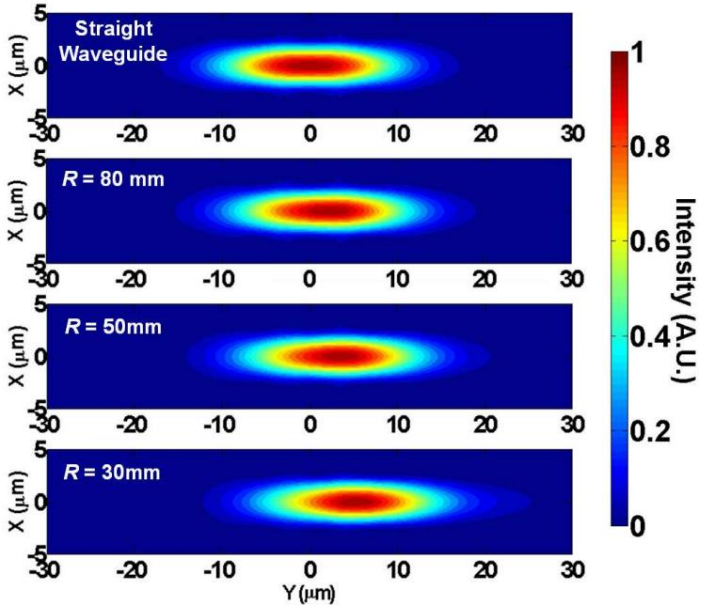


Figure 3.11 Mode profile for straight HCG HCW and curved HCG HCW with different radius of curvature R, simulated in FEM. The waveguide bends towards –y direction, with the center of the bending curvature at y=-R. The HCW height d=6 μm.

3.1.6 Discussion

The demonstration of lateral confinement in planar HCG HCWs opens up a new scheme of waveguide engineering through surface phase manipulation. Keeping the HCG thickness fixed, different periods and grating bar widths are able to provide different reflection phases, while maintaining a high reflectance.

The planar structure of the waveguide makes fabrication simple. The HCG thickness is kept constant, and thus only one step of etching is required. While the waveguides

presented here are formed from two pieces of HCG chips and offer a proof-of-concept, monolithic integration is possible by flip-chip bonding or by processing on a multi-stack silicon oxide wafer.

The presented HCG HCW is the first experimental device showing lateral confinement in a low-loss planar HCW structure. The effective index method is demonstrated for the first time in a planar waveguide structure, as well as in an HCW. The measured propagation loss is the lowest among all HCWs that are mode-matched to a single-mode optical fiber. Further optimization of the HCG dimensions based on the loss contour and effective index contour map can lower the loss to less than 0.1 dB/cm for a straight waveguide in FEM simulations. This is a result of the high reflectivity of the HCG, as well as the unique lateral confinement scheme. Radius of curvature and loss of the curved waveguide can also be reduced by further optimization of the effective index. As an example, we design $\Delta n/n_{core} \sim 2.1 \times 10^{-3}$ for a 5 μm height waveguide. The straight waveguide loss is 0.442 dB/cm, whereas the bending loss is 1.026 dB/cm and 8.839 dB/cm for a radius of curvature R=15 mm and 10 mm respectively in FEM simulation. Further optimization of the waveguide configuration is possible to reduce R to below 10 mm while maintaining a low loss.

The HCG-HCW discussed above offers transverse confinement by HCG, and lateral confinement by effective index method. In general, transverse confinement is stronger, due to a high reflectivity from the HCG. The bending, however, happens in-plane, and thus relies on the lateral confinement. If one can switch the two confinement schemes, so as to make the stronger confinement happens in-plane, then it would make the bending much easier – a smaller radius of curvature while maintaining a low loss. We can thus rotate the HCG 90^o into posts, as shown in Figure 3.12. The light would thus be confined laterally by the HCG posts. As long as the HCGs are designed and optimized such that they support high reflection over a large incident angles, the low propagation loss can be maintained in a small bending radius. In the transverse direction (vertical direction), light can be confined in the center region through effective index method by choosing different materials for the center and the two ends of the posts. This defines the core and cladding vertically; though their HCG parameters are the same, the different refractive index of the material can induce different effective indices. As a preliminary simulation result, Figure 3.12 shows the mode profile of a vertical HCG-HCW after 300 μm propagation in finite-difference time-domain method (FDTD) simulation. The core region is 3 μm thick ($Al_{0.1}Ga$)InAs lattice matched to the 4 μm InP cladding region on each side. The waveguide width is 4 μm. The period, bar width, and thickness of the HCG is optimized to be 675 nm, 285 nm and 400 nm respectively. This simulation verifies the effective index method. It is expected that a small bending radius and small bending loss can be achieved in this structure.

We believe that the effective index method developed here and the two unique waveguide structures presented in this section would establish a new regime of waveguiding in HCWs.

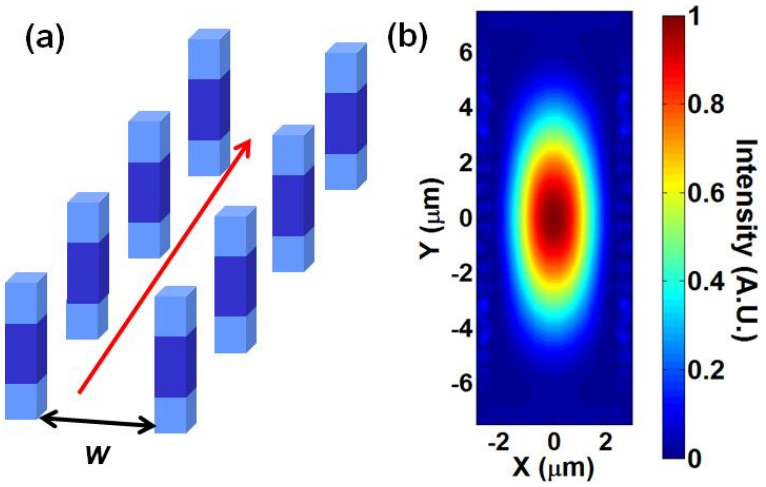


Figure 3.12 (a) Schematic of the proposed vertical HCG-HCW. The arrays of HCG posts form the HCW with width *w*. Each HCG post is composed of three sections vertically: high refractive index material for the center region functions as core, and low refractive index material for the two end region functions as cladding. Vertical confinement is achieved by effective index method. (b) Mode profile of the vertical HCG-HCW simulated by FDTD. The propagation length is 300 μm. The core region is 3 μm thick ($Al_{0.1}Ga$)InAs lattice matched to the 4 μm InP cladding region on each side. The waveguide width is 4 μm. The period, bar width, and thickness of the HCG is optimized to be 675 nm, 285 nm and 400 nm respectively.

3.2 HCG-DBR hybrid Hollow-core Waveguide

As is mentioned above, a larger core yields a lower transverse mode loss; however the lateral loss would be large due to a weaker confinement. To overcome this, distributed Bragg reflector (DBR) is utilized to confine the light laterally. The waveguide is 100 μm by 100 μm in core size, and is schematically shown in Figure 3.13. A <0.1 dB/m loss is estimated based on the loss measurement.

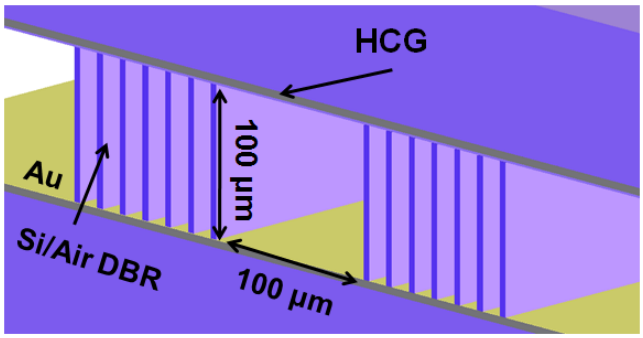


Figure 3.13 HCG-DBR hybrid hollow-core waveguide. Light is confined transversely by top HCG layer and bottom Au layer on the SiO_2/Si structure, which can also be

replaced by HCG through further processing; laterally, light is confined by Si/Air DBR. The core size is 100 μm by 100 μm.

3.2.1 Design and fabrication

The DBR is created by a deep etch on an SOI wafer with 100 μm device layer thickness. To overcome the fabrication error during the deep etch, the period as well as the Si thickness of each DBR pair slightly differs from each other. This chirped profile provides a much broader high reflection bandwidth. After the deep etch, Au is subsequently deposited onto the exposed SiO_2 layer on the SOI to increase the bottom reflectivity. Another SOI wafer with thin device layer is processed with DUV lithography for the HCG layer. The two SOIs are then placed together to form the waveguide. Figure 3.14 shows the SEM images of the fabricated DBR through Bosh process. The etch profile angle is $<0.5^o$.

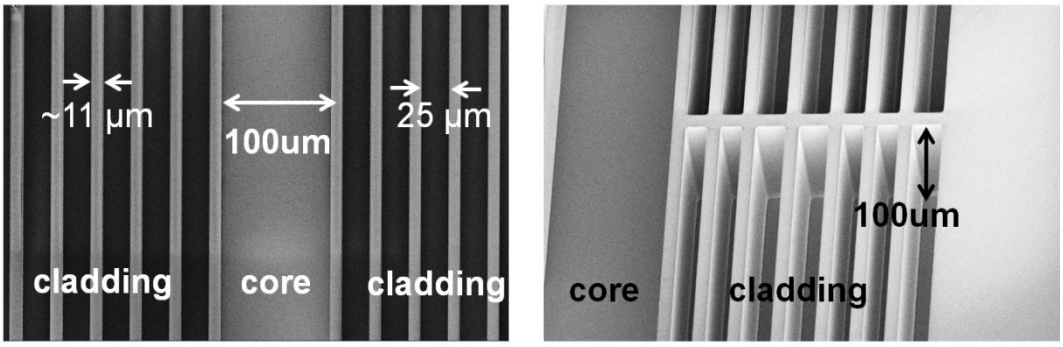


Figure 3.14 SEM image of the HCG-DBR hybrid HCW. The DBR structure is fabricated on an SOI wafer by deep etch Bosh process. Au will be subsequently deposited onto the exposed SiO_2 layer. Another SOI wafer is patterned with HCG and would be placed on top of this wafer to form the 100 μm by 100 μm HCW.

3.2.2 Characterization

The loss measurement setup is similar to the previous one discussed above, shown in Figure 3.7. The only difference is that the waist of the launched beam is shaped to be ~35 μm, which is optimized to have a lowest coupling loss ~0.09 dB into the 100 μm by 100 μm HCW. Figure 3.15 shows the loss spectrum of two identical 10 cm long straight waveguides. A minimum loss of ~0.08 dB is achieved at 1563 nm over a 10 cm waveguide including coupling loss. This total loss value is compatible with the 0.09 dB coupling loss, and with the consideration that the measurement uncertainty of the system is ~0.08 dB, it is estimated that the propagation loss of this HCW is <0.1 dB/m. The mode profile of the HCW measured from the experiment matches well with that from the simulation.

The waveguide length can be extended by spiral patterns through continuous curving or folded patterns through 45^o turning mirror pairs on the 6'' SOI wafer, as shown in Figure 3.16. The HCG wafer, on the other hand, can be uniform, as its reflectivity is maintained high with different incident directions. The propagation loss is characterized

by an optical backscatter reflectometer (OBR). Figure 3.17 shows the measurement results on an HCW with 45° turning mirror pairs. Although the propagation loss for the straight waveguide is very low, it increases dramatically after the first turning. This is due to the fundamental mode scattering into high order modes after the turning, and high order modes have much higher loss. To reduce this loss, a DBR structure with full photonic bandgap for all incident angles is needed. This insures low loss on high order modes. Si/SiO_2 DBR structure can fulfill this requirement [80].

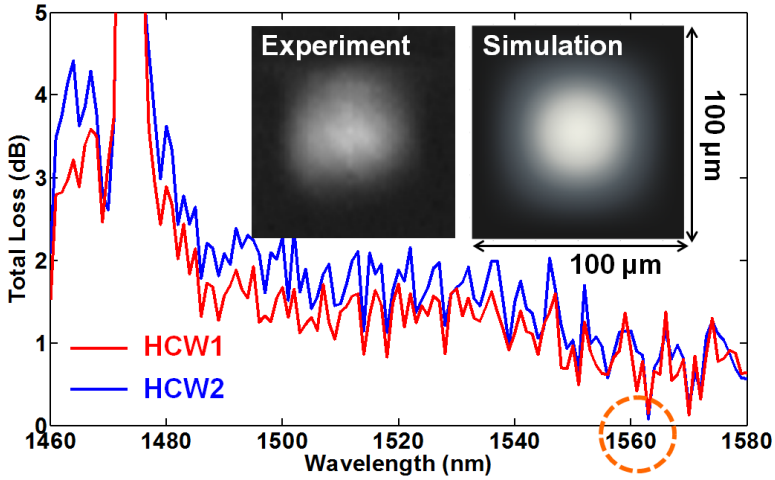


Figure 3.15 Total loss spectrum is measured for two identical 100 μm by 100 μm core 10 cm long HCG-DBR hybrid HCWs. A minimal total loss of ~0.08 dB is obtained at 1563 nm. The experimental mode profile matches well with the simulation.

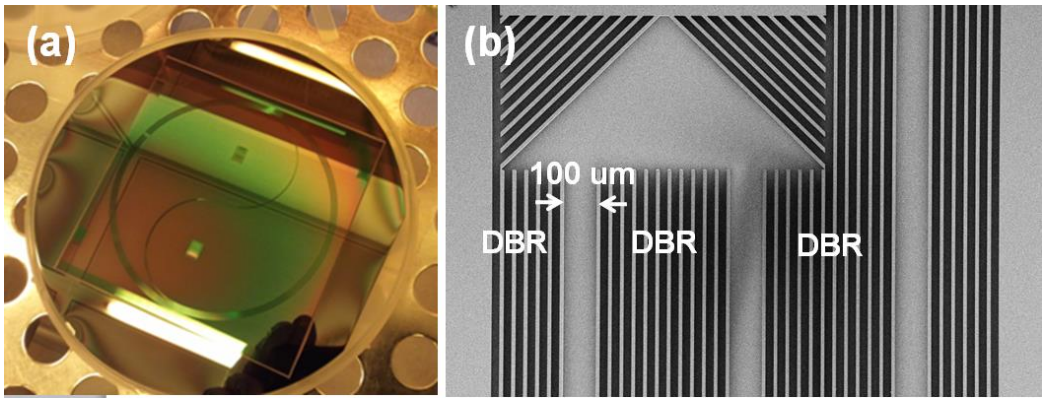


Figure 3.16 (a) Image of an HCG-DBR hybrid HCW with spiral pattern through continuous curving. The total length is 1.34 m, with >40 mm radius of curvature in the outer rings and 21.6 mm in the handoff region. Au is deposited on the bottom of the exposed SiO_2 to enhance the reflectivity. (b) SEM image of a 45° turning mirror pair. This folds the straight HCW and thus aggressively extends the length.

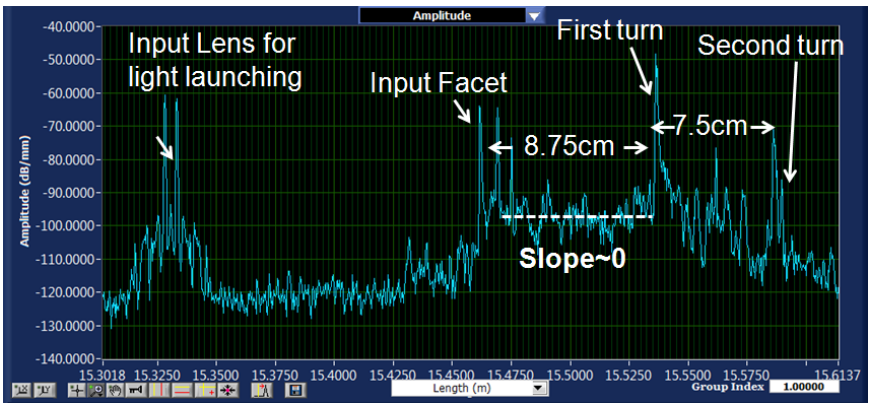


Figure 3.17 Loss measurement through optical backscatter reflectometer on an HCG-DBR hybrid HCW. The spikes indicate large reflections, typically because of the interfaces between different optics in the optical path. The slope of the signal indicates propagation loss. The average slope is around zero for the first 8.75 cm straight waveguide, indicating an extremely low loss. After the first turning however, the slope becomes signification, indicating a high loss. This is due to the fundamental mode scattering into high order modes, which have much higher propagation losses. The noise is due to environmental vibration and defects along the waveguide.

3.3 Cage-like High Contrast Grating Hollow-core Waveguide

The HCG HCWs discussed in the previous section have innovative structures and a very low propagation loss. Besides those, one can also build up an HCG HCW with four HCG reflectors serving as four waveguide walls, as shown in Figure 3.18. Since the waveguide is constructed by four highly reflective HCG mirrors, strong optical confinement is expected. The mode profile of the fundamental mode is also shown in Figure 3.18. The propagation loss is simulated to be <0.05 dB/cm in finite-difference time-domain method (FDTD) for an HCW with core size of 12 μm x 8 μm. Low loss and strong optical guiding in this cage-like HCW has recently demonstrated [47]. Slow light application in this cage-like HCG HCW is currently a research topic.

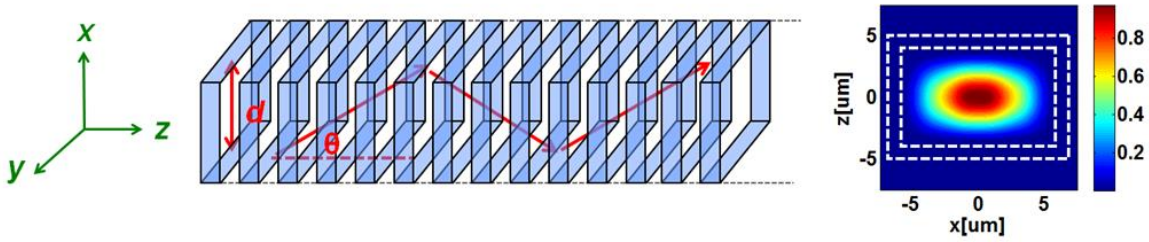


Figure 3.18 Schematics of a 3D cage-like HCG HCW and the mode profile of the fundamental mode. The color code indicates the normalized electrical field intensity.

3.4 Optical Switch based on High Contrast Grating Hollow-core Waveguide

The HCG HCWs discussed in the previous sections establish a new platform for integrated optics. Various applications are promising on this platform, such as gas sensing, slow light, optical delay line, RF filter, etc. One interesting application is the optical switch. We propose an optical switch based on HCG HCW in this section. Simulation results show that the switching length is ~25x smaller than the conventional 2x2 multimode interference coupler.

Optical switch is a key component in integrated optics. Low switching energy and small footprint are critical for future large scale integrated optics. Various approaches such as ring resonator [81] and photonic crystal nanocavity [82] have been demonstrated. All these optical switches are based on conventional solid-core waveguide, and there are no reports on optical switch based on hollow-core waveguide. HCG HCW offers a unique opportunity. HCG can be highly reflective; meanwhile, if it operates at a resonance regime, its reflectivity can change from 1 to 0 with a small refractive index change of the HCG bars. With the high reflectivity of HCG, two HCWs can be placed adjacent to each other and completely isolated by ONE single layer of HCG. By a small change of the refractive index of the shared HCG layer (either by optical induced carrier or by electrically injected carrier), the HCG becomes effectively transparent, and light can thus switch from one HCW to the adjacent one, as shown in Figure 3.19. This is analogous to removing a fence between two car lanes and thus cars can switch lanes.

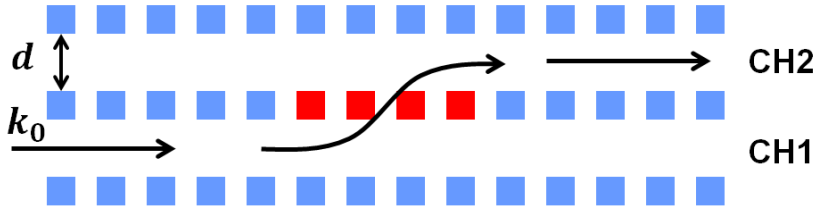


Figure 3.19 Schematics of an HCG-HCW optical switch. Light propagating in the two slab HCG-HCWs can be completely isolated by ONE single HCG layer. By a small change of the HCG bars' refractive index (region labeled in red), the HCG can become transparent, and light can switch from CH1 to CH2.

Figure 3.20 (a) and (b) shows the reflectivity of a TE HCG as a function of incident angles and wavelengths, for the grating bar index of 3.483 and 3.490 respectively. The HCG has a period of 0.77 μm, thickness of 0.47 μm, and grating bar width of 0.5775 μm. In both case of the grating bar index, the HCG's reflectivity is almost independent of the incident angle θ, at the wavelength range from 1.54 μm to 1.56 μm. Due to the resonance effect, the reflectivity can experience a large change (from fully reflective to fully transparent) with a small change of wavelength. Alternatively, at the same wavelength, a small refractive index change can transfer HCG from almost fully reflective to almost fully transparent, as seen in Figure 3.20 (a) and (b), at the wavelength of 1.553 μm, shown by the white dashed line. This HCG property provides a foundation for optical switch.

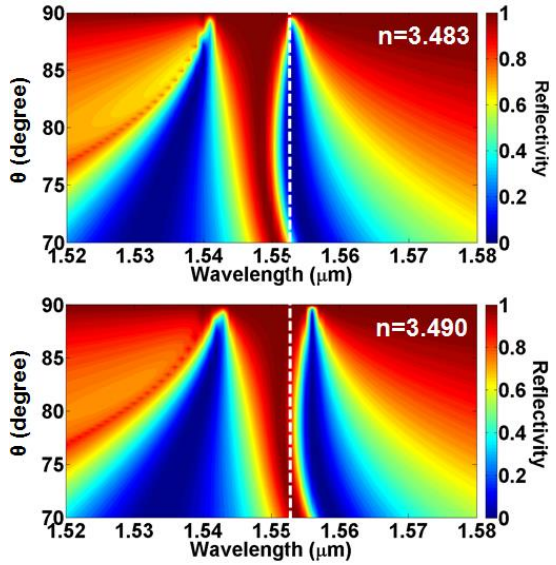


Figure 3.20 Reflectivity contour of an HCG resonance design as a function of incident angles θ and wavelengths, for the grating bar refractive index of 3.483 (a) and 3.490 (b). At the same wavelength (1.553 μm, indicated by the white dashed line), the HCG can change from highly transparent to highly reflective, with the grating bar index change from 3.483 to 3.490.

The reflectivity contour plot provides an intuition of the switching mechanism. To have a more comprehensive understanding, the dispersion relation ($\omega - k$ diagram) of the HCW modes is calculated. The waveguide dispersion relation can be extracted by calculating the reflectivity of a plane wave that is externally incident on the waveguide, versus its incident angle at a fixed optical frequency ω [75]. To find the waveguide propagation modes, one can look for the sharp resonance dips in this reflectivity vs. incidence angle plot. A resonance dip at a specific incident angles θ indicates the phase matching condition in the waveguide is satisfied, i.e. the optical modes propagating in z-direction having standing wave pattern in the transverse x-direction. The wave vector k_x of the mode in the x direction satisfies $2k_x d + 2\varphi_r = 2m\pi$, where φ_r is the reflection phase from the HCG, and m is an integer. The propagation constant k_z of the mode is the same as the z-wavenumber of incident plane wave, i.e. $k_o sin\theta$, where k_o is the wave vector in vacuum. Using the above method, one can find the Eigen modes of any hollow-core waveguide system (either single HCW or two HCWs running in parallel).

There are various methods to calculate this reflectivity. Since the waveguide system is essentially composed of various HCGs, separated by air layers (e.g. double stack of HCG separated by an air layer for a single HCW, or triple stack of HCG interleaved by two air layers for the 2x2 optical switch), one can use the T-matrix method described in Section 2.2.4, or use RCWA to calculate the reflectivity. Alternatively, one can first calculate the reflection coefficients (0^{th} order) of the single HCG, and then calculate the reflectivity of the whole system with the transmission matrix method. This is accurate in the subwavelength regime and when the waveguide core is large, since there is no higher order reflection and transmission, and the evanescent wave of the HCG can be ignored. This is illustrated in Figure 3.21. The whole system is divided into several layers,

separated by the HCG. The HCG is treated as a pure boundary (with equivalent zero thickness) between two layers, with its property described by its transmission and reflection coefficients.

The electrical field at a specific layer *i* can be described by the following expression,

$$E_i(x,z) = \left[E_+ e^{-ik_o cos\theta x} + E_- e^{ik_o cos\theta x}\right] e^{i(\omega t - k_o sin\theta z)}$$

$$= [A_i(x) + B_i(x)] e^{i(\omega t - k_o sin\theta z)} \quad (3.8)$$

where $A_i(x)$ is the forward propagation wave and $B_i(x)$ is the backward propagation wave in the x direction.

At the HCG boundary *j* between two layers, the transmission matrix can be described by the following,

$$T_{boundary j} = \frac{1}{t_j}\begin{pmatrix} 1 & -r_j \\ r_j & t_j^2 - r_j^2 \end{pmatrix} \quad (3.9)$$

where t_j and r_j is the transmission and reflection coefficient of the HCG as the jth boundary. This matrix relates the *A*, *B* coefficients at the two sides of the HCG boundary. For example, we have the following equation to relate the A_1, B_1, A_2' and B_2' in Figure 3.21 (a):

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \frac{1}{t_1}\begin{pmatrix} 1 & -r_1 \\ r_1 & t_1^2 - r_1^2 \end{pmatrix}\begin{pmatrix} A_2' \\ B_2' \end{pmatrix} \quad (3.10)$$

The transmission matrix for the *A*, *B* coefficients at the two boundary of a specific layer *i* (with thickness d_i) can be expressed in the following,

$$T_{layer i} = \begin{pmatrix} e^{ik_o cos\theta d_i} & 0 \\ 0 & e^{-ik_o cos\theta d_i} \end{pmatrix} \quad (3.11)$$

A_2, B_2, A_2' and B_2' in Figure 3.21 (a) can be related with the above transmission matrix,

$$\begin{pmatrix} A_2' \\ B_2' \end{pmatrix} = \begin{pmatrix} e^{ik_o cos\theta d_i} & 0 \\ 0 & e^{-ik_o cos\theta d_i} \end{pmatrix}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \quad (3.12)$$

With the above equations, we can relate A_1, B_1, A_3' and B_3' in Figure 3.21 (a) in Eq. (3.13a), and A_1, B_1, A_4' and B_4' in Figure 3.21 (b) in Eq. (3.13b).

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = T_{boundary1} T_{layer2} T_{boundary2}\begin{pmatrix} A_3' \\ B_3' \end{pmatrix}$$

$$= \begin{pmatrix} M1_{11} & M1_{12} \\ M1_{21} & M1_{22} \end{pmatrix}\begin{pmatrix} A_3' \\ B_3' \end{pmatrix} \quad (3.13a)$$

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = T_{boundary1} T_{layer2} T_{boundary2} T_{layer3} T_{boundary3} \begin{pmatrix} A_4' \\ B_4' \end{pmatrix}$$
$$= \begin{pmatrix} M2_{11} & M2_{12} \\ M2_{21} & M2_{22} \end{pmatrix} \begin{pmatrix} A_4' \\ B_4' \end{pmatrix} \qquad (3.13b)$$

The transmission and reflection coefficient of the whole system for Figure 3.21 (a) would thus be

$$t = \frac{1}{M1_{11}} \qquad (3.14)$$

$$r = \frac{M1_{21}}{M1_{11}} \qquad (3.15)$$

The same applies to that in Figure 3.21 (b).

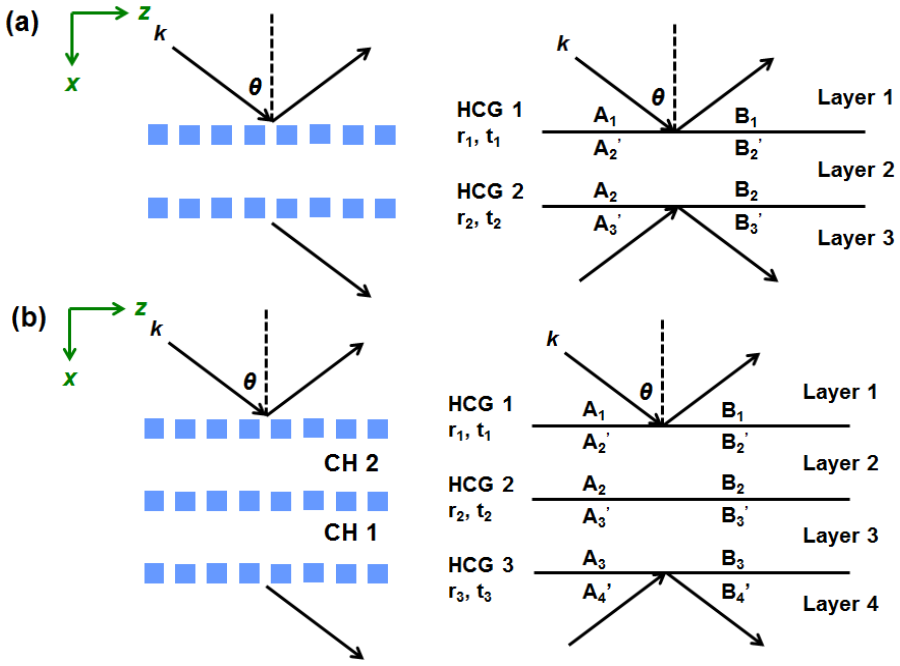
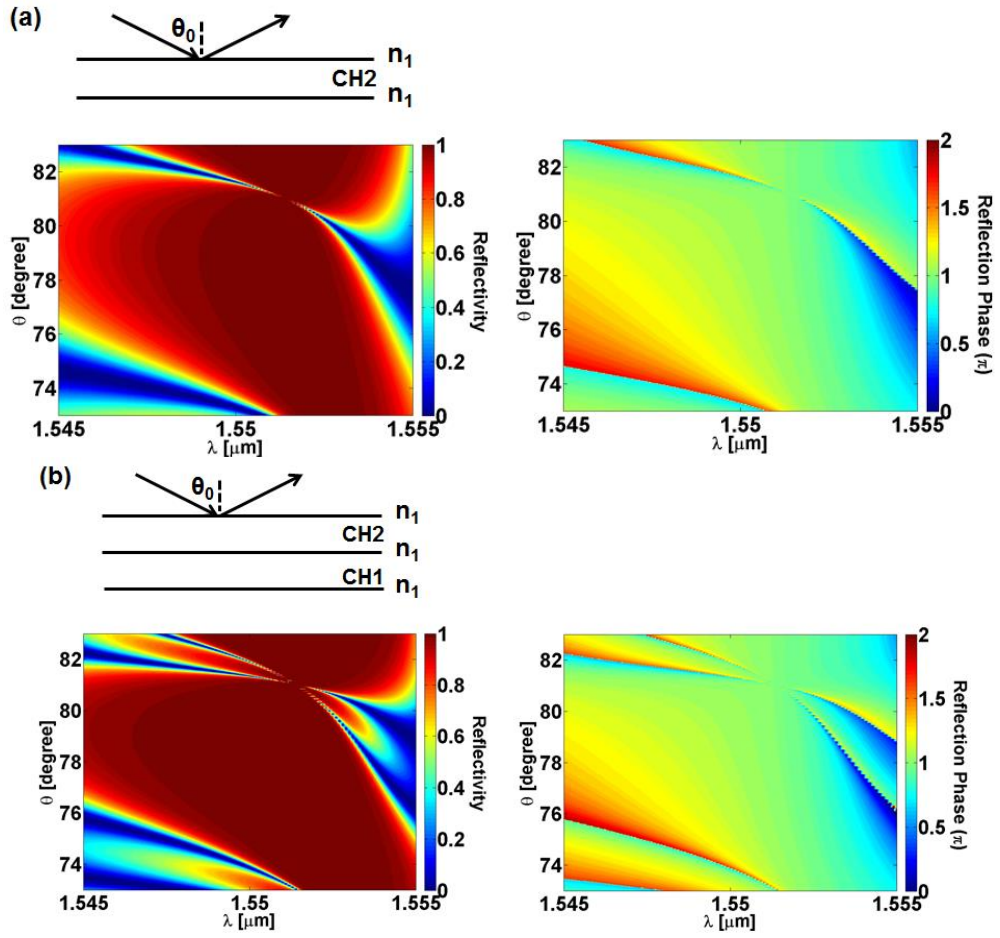


Figure 3.21 schematics for the HCG HCW and the definition of the layers and boundaries. (a) shows a single HCG HCW; (b) shows two parallel HCG HCWs, which can be a 2x2 optical switch.

With the above transmission matrix method, one can plot out the reflection contour map against incident angle and wavelength, as shown in Figure 3.22. The core size of all the waveguides is 5 μm. Figure 3.22 (a) shows the case of the HCW where the HCG's grating bars have a refractive index of 3.490; (b) shows the two parallel HCWs where the three HCG's grating bars have the same refractive index of 3.490; (c) shows the same as (b) but with the middle HCG's grating bar index changed to 3.483. Figure 3.22 (b) is essentially the 2x2 optical switch in the OFF state, where the two channels are isolated; (c) is the 2x2 optical switch in the ON state, where the light from one channel can switch to the other channel. The reflection dips indicate the waveguide modes. The horizontal and

vertical axis in Figure 3.22 is wavelength λ and incident angle θ respectively, and they can be converted to $k_o = \frac{2\pi}{\lambda}$ and $k_z = k_o sin\theta$. The $\omega - k$ diagram can thus be extracted from Figure 3.22, shown in Figure 3.23. The blue curve shows the ω-k diagram of the fundamental mode of a single HCW [Figure 3.22 (a)]. When two such identical HCWs are running in parallel [Figure 3.22 (b)], this fundamental mode splits into two modes, an odd mode and even mode with its propagation constant k_{z_odd} and k_{z_even} respectively. The two lines cross each other showing no coupling. When the switch is turned on [Figure 3.22 (c)], however, these two modes have a much larger split and no crossing. This is because the light gains a large extra phase at the transparent HCG, which is typical at the resonance condition. The two split modes interference (beat) with each other, and thus energy can transfer from one HCW to the other. The switching length can be expressed as $\pi/|k_{z_{odd}} - k_{z\_even}|$. Thus the larger the modes spilt, the shorter the switching length. As an example, the odd mode and even mode for a specific wavelength are plotted in Figure 3.24 (a), and their beating along the waveguide is calculated in (b). The mode profile, depicted by Eq. 3.8, is calculated by the transmission matrix. Note that the evanescent wave is ignored. For a more rigorous treatment, one can use the full transmission matrix method described in Section 2.2.4. Nevertheless, the transmission matrix described above provides a greatly simplified and accurate approach to analyze the switching behavior and predicts well the field profile in the hollow-core region.



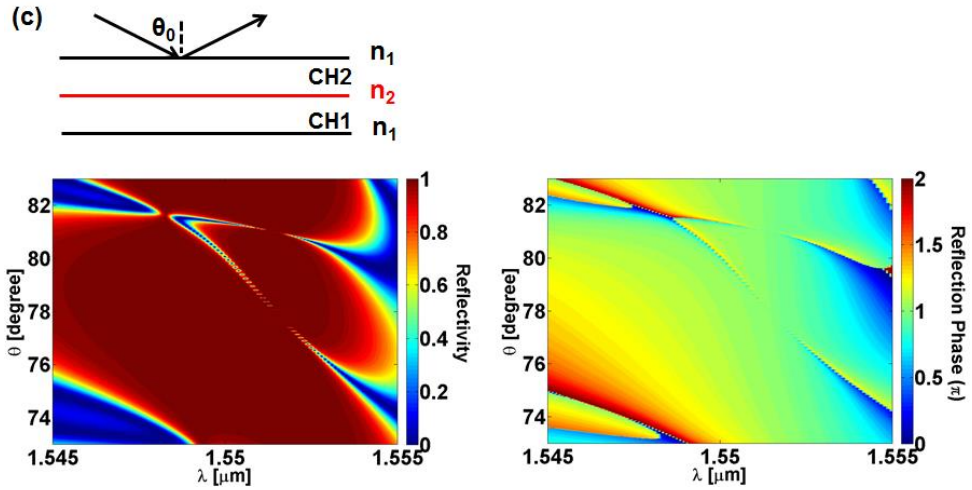


Figure 3.22 Reflection and reflection phase contour map against incident angle and wavelength, for (a) a single HCG HCW, (b) two identical HCG HCWs running in parallel sharing a common HCG, and (c) two HCG HCWs running in parallel, but with the refractive index of the middle HCG layer changed. n_1=3.490, n_2=3.483.

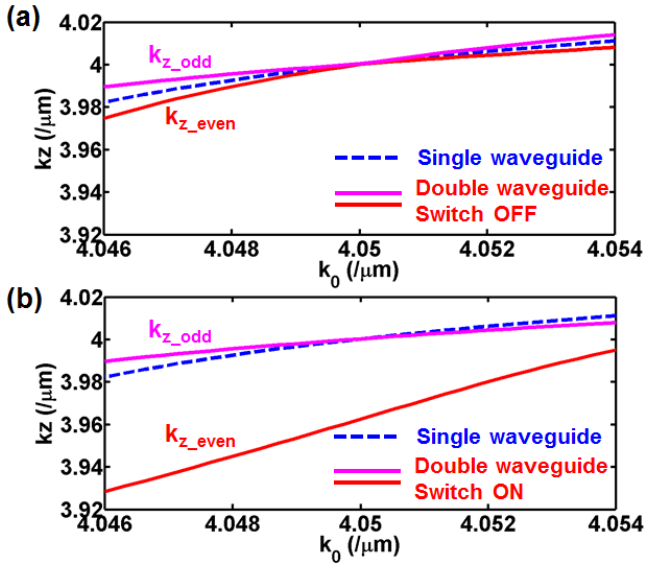


Figure 3.23 ω-k diagram of the HCG-HCW extracted from Figure 3.22. The blue dashed curve is the ω-k diagram of the fundamental mode of a single HCW. When two such identical HCWs run in parallel, this mode splits into an odd mode and even mode, indicated by the magenta and red curves. (a) Switch OFF state; (b) Switch ON state.

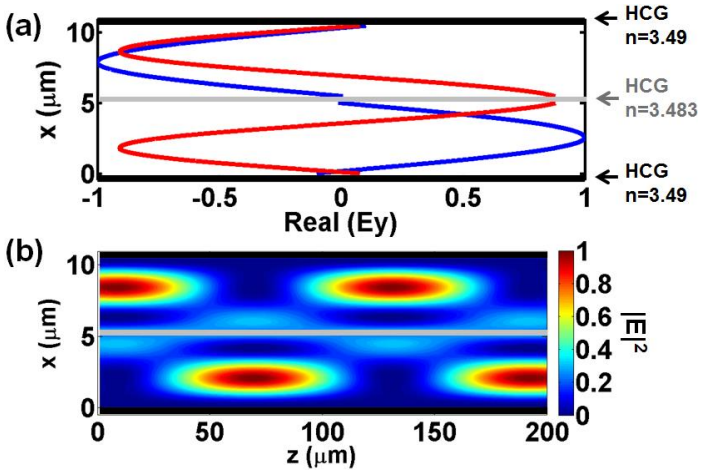


Figure 3.24 (a) Mode profile of the even mode (red) and odd mode (blue) for a specific wavelength at the switch ON state. (b) The two modes beating field pattern along the HCG-HCW. The energy is switched in between the two HCWs.

To further verify the HCG-HCW switch, finite-difference time-domain (FDTD) method is used to simulate the switching behavior. A fundamental waveguide mode is launched at one HCW, depending on the refractive index of the switching region (~60 μm in length), light either stays at the original HCW or switches to the adjacent one, as shown in Figure 3.25. The required refractive index change Δn is very small, -7x10^{-3}, making it very promising for achieving very low switching energy. The insertion loss is 0.03 dB and 0.3 dB for the switch OFF and ON respectively, and the crosstalk ratio is -21dB for both states. It is interesting to note that the switching length is <4d^2/λ, where 4d^2/λ corresponds to the length of one full optical bounce in the HCW predicted by ray optics. This surprisingly small switching length is due to the rapid phase change of the HCG at the resonance. In the conventional 2x2 coupler using multimode interference (MMI) (e.g. using silicon photonic technology), the coupling length is ~100d^2/λ [49]. The proposed HCG-HCW switch thus represents a ~25x length reduction. This length reduction is thanks to the ultra-thin HCG that isolates the two waveguides. One can estimating the switching energy by calculating how much energy is needed to change the refractive index of the grating bars in the switching region. If we assume the refractive index change is due to the optical induce carriers, we can calculate the switching energy in the following formulae,

$$E_{switch} = \frac{carrier\ density * switching\ volume}{absorptivity} * photon\ energy \qquad (3.16a)$$

$$carrier\ density = \Delta n/\sigma_c \qquad (3.16b)$$

$$switching\ volume = L * t_g * w * \eta = \frac{4d^2}{\lambda} * t_g * w * \eta \qquad (3.16c)$$

where Δn is the required refractive index change, σ_c is the index change with carrier density, d is the core height, t_g is the grating thickness, w is the waveguide width, η is the grating duty cycle. σ_c is material dependent, we set it be to 8.2x10^{-20} cm^3,

corresponding to that of InGaAsP, which is one of the materials have high σ_c value. w is set to be 5 μm, and absorptivity is set to be 1. The switching energy is calculated to be 1.25 pJ. This value can further decrease with smaller waveguide height d.

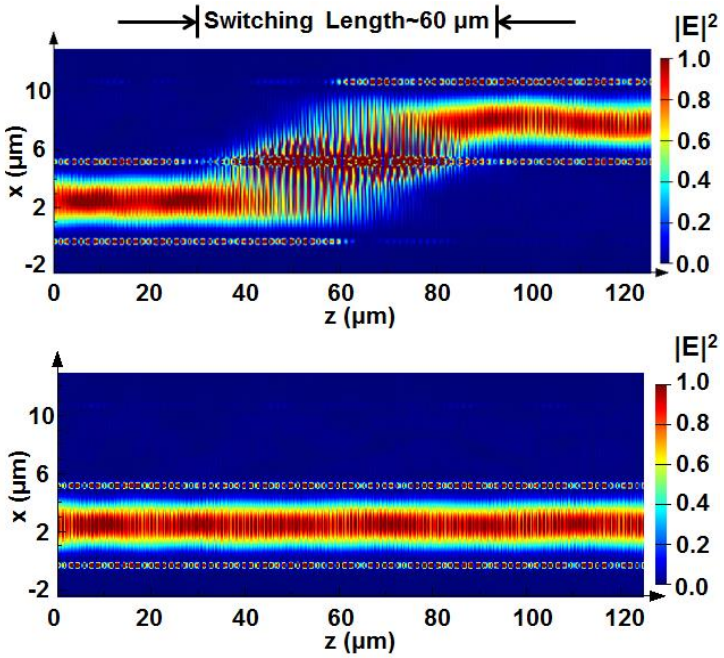


Figure 3.25 FDTD simulation of the HCG-HCW switch. The switching length is ~ 60 μm, and the refractive index change is $-7\text{x}10^{-3}$ at this switching region. The top panel shows the switch ON state, where as the bottom shows the switch OFF state.

Using HCG resonance to switch, this structure shows a strong dependence on wavelengths, as shown in Figure 3.26, which can serve as a wavelength multiplexer (MUX) or demultiplexer (DEMUX). Different wavelengths can be designed to join and/or leave the HCW at different positions. This makes the HCG-HCW array particularly useful as a compact DWDM tunable add-drop MUX and DEMUX.

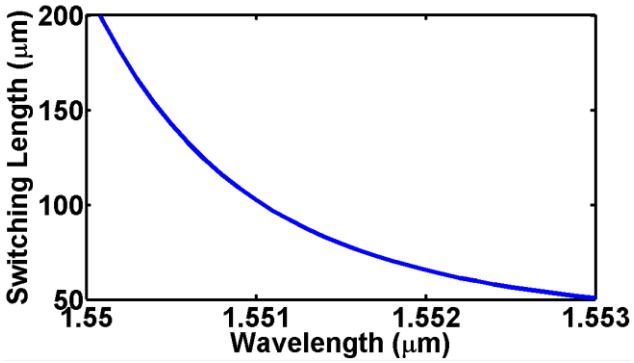


Figure 3.26 Switching length versus different wavelength for the HCG HCW switch.

3.5 Application of High Contrast Grating Hollow-core Waveguide in Gas sensing

HCG HCW has superior advantages in gas sensing application over other conventional HCWs. With no sidewalls, gaseous or fluidic molecules can flow in and out from the side rapidly without latency, instead of being pushed from one end of the waveguide to the other. This dramatically reduces the response time in gas detection. With Fick's first law, the diffusion length (a measure of the distance l that the concentration has propagated over by diffusion in time t) can be derived to be proportional to $\sqrt{Dt}$, where D is diffusion coefficient. Thus t is proportional to l^2/D. This is derived in an assumption of continuum flow without any boundary. In a gas sensing experiment with an HCW, the diffusion time t can be approximated through the above equation, as long as the mean free path of the gas molecule is much smaller than the waveguide core size, which is the typical case. Thus it takes L^2/d^2 times faster for gaseous molecules to fully diffuse into the HCG HCW than the conventional HCWs, where L is the length of HCW and d is the core diameter. This increase can be easily 8-12 orders of magnitude with typical values of L on the order of centimeter to meter, and d on the order of micrometer.

3.6 Summary

In this chapter, various types of HCG based hollow-core waveguides are discussed. The key advantage of using HCG is its high reflection, while its simple geometry and reflection phase engineering enables novel device structures. The sidewall-less hollow-core waveguide and the caged hollow-core waveguide represent such innovation. This builds up a new platform for integrated optics with hollow-core waveguide. This platform is desirable for applications where hollow-core is required, such as for low-loss, high power handling, athermal operation etc. Thus this platform would be ideal for low-loss delay line, RF filtering, gas sensing, and lab-on-a-chip systems. In this chapter, an optical switch is proposed and discussed in detailed. The main difference of the proposed optical switch from the conventional MMI based optical switch is that the two channels are separated by a very thin layer of HCG, which enables the fast switching between the two channels.

Other on-going research topics on this hollow-core waveguide platform includes slow-light and gas sensing. The slow-light takes advantages of the phase engineering of the HCG, while the motivation of the gas sensing is that the gas can diffuse into the hollow-core waveguide orders of magnitude faster than others, due to the unique geometry of the HCG. We believe that by further optimizing the HCG design and waveguide structure, lower loss, lower bending radius of curvature and longer waveguide distance can be achieved. It is expected that this hollow-core waveguide would play an important role in integrated optics.

Chapter 4
High Contrast Grating Optical Phased Array

The previous chapter devotes to the hollow-core waveguide, one of the applications of HCG at oblique incident angle. In this chapter, we explore the applications of HCG at surface normal incidence. One of the novel applications is to use HCGs to construct all-pass optical filters, and use them for optical phased array, which is the focus of this chapter.

An optical phased array is an array of optical antennas that can emit light or direct light into desired directions and suppress in undesired directions. The output phase of the light from each antenna element can be actively controlled, and the whole array forms a near field phase pattern. By Fourier transform, a desirable far field pattern can be obtained.

Optical phased arrays have enabled free-space beam steering for a wide range of applications, such as imaging, display, chemical-bio sensing, precision targeting, surveillance, etc. A chip-scale and high-speed optical phased array is of particular desire. It matches the high-integration-density, small-footprint, low-power-consumption and high-speed requirement of advanced applications such as optical circuit switching, light detection and ranging (LIDAR) etc. The central element of the phased array is the individual phase tuner. The phase tuner can either actively emit light with different phase, or passively transmit or reflect incoming light while modifying its transmission phase or reflection phase. Several phase tuning mechanisms have been demonstrated for optical phased arrays, such as using electro-mechanical [83, 84], electro-optic [85-88], thermo-optic [89, 90] effect. The liquid crystal represents the most mature technology; however its response time is very slow, typically in the order of hundreds of Hz to tens of kHz [85-87]. Thermo-optic is slow in response as well. Electro-mechanical (i.e. MEMS) based phase tuner can be fast, but the MEMS structure is usually sophisticated in design and fabrication [83, 84]. Electro-optics effect based on dielectric is fast as well [88]; however for two dimensional beam steering, a tunable laser with high tuning speed is required, which largely increases the complexity of the whole system.

The need for a high speed optical phased array offers an opportunity for HCG. HCG is light-weight and can be made into a fast MEMS structure. Mechanical resonance frequency as high as 27 MHz has been previously demonstrated in an HCG VCSEL [22]. With this motivation, we explored applying HCG in optical phased array. In this chapter, we discussion a demonstration of a novel 8x8 optical phased array based on HCG all-pass filters (APF) with high speed micro-electro-mechanical actuation. Highly efficient phase tuning is achieved by actuation of the HCG to tune the cavity length of the APFs. Using APF phase-shifters allows a large phase shift with an actuation distance range of only tens of nanometers. The ultrathin HCG further ensures a high tuning speed (0.53 MHz). Fast beam steering is experimentally demonstrated. Besides one-dimensional HCG, two-dimensional HCG APF are also designed and fabricated, and demonstrated for fast beam steering. The material platform can be either III-V or silicon. In this chapter, devices

developed on both material platforms will be discussed, but we will mainly focus on the III-V GaAs based devices.

The concept and design of HCG APF optical phased array is introduced in section 4.1, followed by the experimental characterization and demonstration in section 4.2. The device is fabricated on GaAs substrate. In section 4.3, we briefly discuss the development of HCG optical phased array on silicon platform. Section 4.4 provides a summary of this chapter and an outlook of the HCG APF optical phased array.

4.1 HCG MEMS Mirror as a Phase Tuner

In this section, we describe two ways of using the HCG MEMS mirror as a phase tuner. The first approach is the most intuitive way, using HCG as a piston mirror. While this approach is simple and straight-forward, the required actuation voltage is large in order to achieve a high actuation speed. The second approach is to construct an all-pass filter, and use HCG as its top mirror. This approach is shown to have a high efficiency in the phase tuning, i.e. small actuation voltage, small HCG displacement to achieve a large and high speed phase shift.

4.1.1 Piston Mirror Approach

HCG is a very good reflector and can be electrostatically actuated. Thus an intuitive way to design a phase tuner is to use the HCG as a piston mirror. A laser beam comes at surface normal incident angle onto the HCG and gets reflected. By actuating the HCG (out of plane translation movement), the reflection light would experience a physical delay, and thus its phase can be tuned, shown schematically in Figure 4.1. The HCG is connected to the substrate with a spring, and can be electrostatically actuated. This is a thus a parallel plate MEMS configuration. The phase shift $\Delta\varphi$ with respect to the HCG displacement d is $\Delta\varphi = 2k_o\Delta d$.

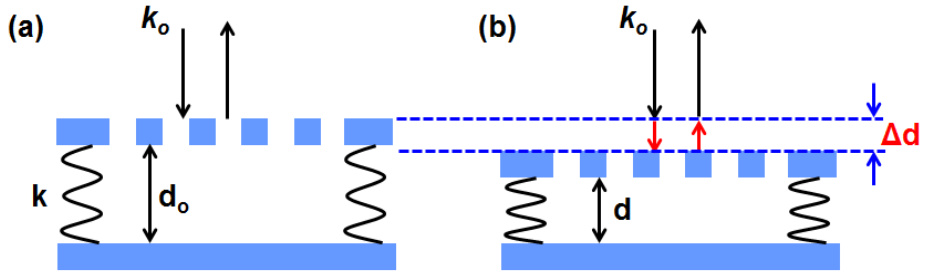


Figure 4.1 Schematic of using HCG as a piston mirror for phase tuner. The HCG is connected to the substrate with a spring. The HCG can be electrostatically actuated. By actuating the HCG for a displacement Δd, the reflection light experiences a physical delay and thus a phase shift of $\Delta\varphi=2k_o\Delta d$.

The piston mirror provides a straight forward phase tuning mechanism. The drawback is that in order to have a 2π phase shift, a very large actuation distance is

needed ($\Delta d = \lambda/2$), and thus a large voltage. On the other hand, in order to have a fast actuation speed, the spring constant k needs to be large. This further increases the actuation voltage. It will be shown in the following that using an HCG as the piston mirror instead of a DBR can greatly reduce the actuation voltage; on this other hand, the required voltage to reach high speed operation is still high.

In general, for MEMS parallel plate, we have the following formulation,

$$F_{spring} = k(d - d_0) \tag{4.1}$$

where d_0 is the initial gap distance between the two plates (i.e. HCG and the substrate). When there is an actuation voltage V between the two plates, the gap distance becomes d. The two plates form a capacitor. With an actuation voltage V, the energy stored in the capacitor is

$$W_e = \frac{1}{2}CV^2 = \frac{\varepsilon_0 A}{2d}V^2 \tag{4.2}$$

Thus the electrostatic force is

$$F_e = \frac{\partial W_e}{\partial d} = -\frac{\varepsilon_0 A}{2d^2}V^2 \tag{4.3}$$

At the steady state

$$F_{spring} = F_e \tag{4.4}$$

$$(d - d_0)d^2 = -\frac{\varepsilon_0 A}{2k}V^2 \tag{4.5}$$

To reach 2π phase shift, $\Delta d = \lambda/2$, and thus the required voltage is

$$V_{2\pi} = \sqrt{\frac{\lambda k}{\varepsilon_0 A}}\left(d_o - \frac{\lambda}{2}\right) \tag{4.6}$$

It is seen that the smaller the initial gap d_0 is, the lower the required voltage. However, to make the MEMS stable, one has to satisfy $(\partial F_{spring} - \partial F_e)/\partial d > 0$, or else the net force increases as the gap is narrowed, and the MEMS would snap. This imposes a minimum gap size, $d_m = 2/3d_0$. To achieve $\Delta d = \lambda/2$, the minimum d_0 is $3\lambda/2$, this corresponds to the minimum voltage in the following expression,

$$V_{2\pi} = \sqrt{\frac{\lambda k}{\varepsilon_0 A}}\lambda \tag{4.7}$$

On the other hand, the mechanical resonance frequency f_r is related to the mirror mass and spring constant, in the following equation,

$$f_r = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{4.8}$$

Eq. 4.7 and 4.8 indicates a linear relationship between $V_{2\pi}$ and f_r. To achieve high speed phase tuning, a larger actuation voltage is required. $V_{2\pi} \propto \sqrt{t_g}$, where t_g is the mirror thickness; this shows the advantage of using HCG over DBR as the piston mirror.

$$V_{2\pi} = 2\pi\sqrt{\frac{\lambda m}{\varepsilon_0 A}}\lambda f_r = 2\pi\sqrt{\frac{\lambda \rho t_g}{\varepsilon_0}}\lambda f_r \tag{4.9}$$

Figure 4.2 and Figure 4.3 show the relationship between $V_{2\pi}$, f_r, and k. The material is silicon. The HCG is 20 μm by 20 μm, with a duty cycle of 0.5, and a thickness of 200 nm. It is seen that in order to reach 0.5 MHz resonance frequency (i.e. the phase tuning speed), the required voltage is more than 40 V (at 1MHz, this voltage becomes more than 80 V). Such a large voltage is not feasible for many integrated circuits. There is thus a need to increase the phase tuning efficiency and reduce the actuation voltage.

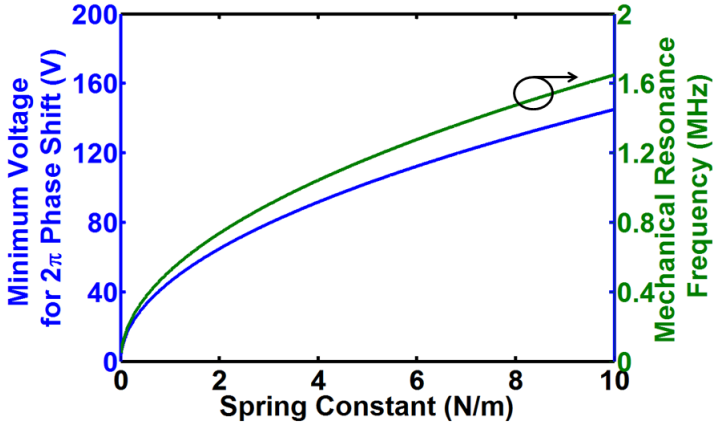


Figure 4.2 Relationship between $V_{2\pi}$, f_r and the spring constant k for a silicon HCG piston mirror designed for 1550 nm operation wavelength.

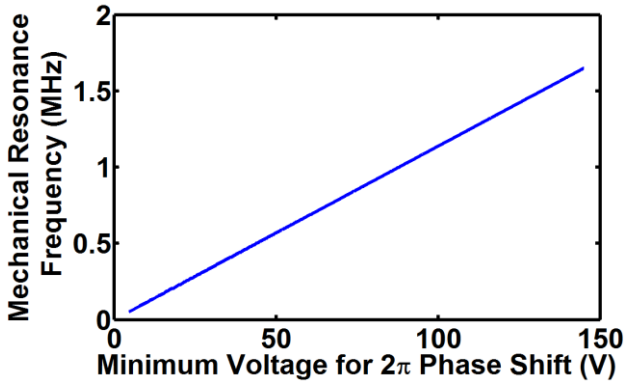


Figure 4.3 Linear relationship between $V_{2\pi}$ and f_r, for a silicon HCG.

4.1.2 All-pass Filter Approach

To reduce the actuation voltage, one needs to design a new scheme such that the required actuation distance is small. In order words, an increase in the phase tuning efficiency is desired. Resonator structure can meet this requirement. This is similar with the fact that the footprint of a ring optical modulator is much smaller than a Mach-Zehnder modulator. By replacing the substrate shown in Figure 4.1 with a highly reflective mirror, an FP etalon can be constructed. Figure 4.4 (a) shows the typical reflection spectrum, and reflection phase spectrum of an FP etalon. The reflection dip indicates the resonance. Across the resonance, a large phase change can be obtained. Alternatively, we can fix the wavelength of incident light, and plot the reflectivity and reflection phase versus the cavity length, shown in Figure 4.4 (b). As the cavity length goes across the cavity resonance, a large reflection phase shift is obtained. In the calculation, the reflectivity of the top HCG and bottom DBR is 0.99 and 0.995 respectively. The cavity length is set to be 775 nm in Figure 4.4 (a), and the central wavelength is set to be 1550 nm in Figure 4.4 (b). Note that in the calculation for Figure 4.4 (a), the phase dispersion of the mirror is not considered. The reflection phase of both mirrors is set to be π.

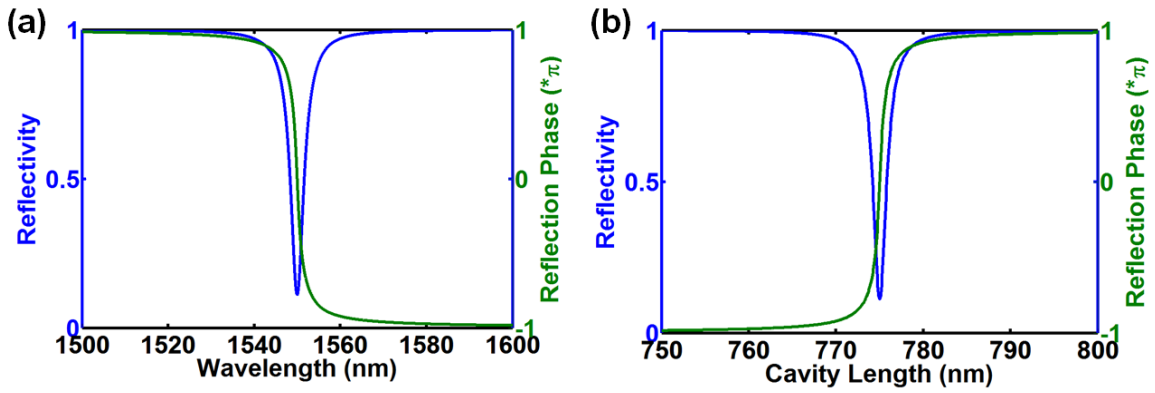


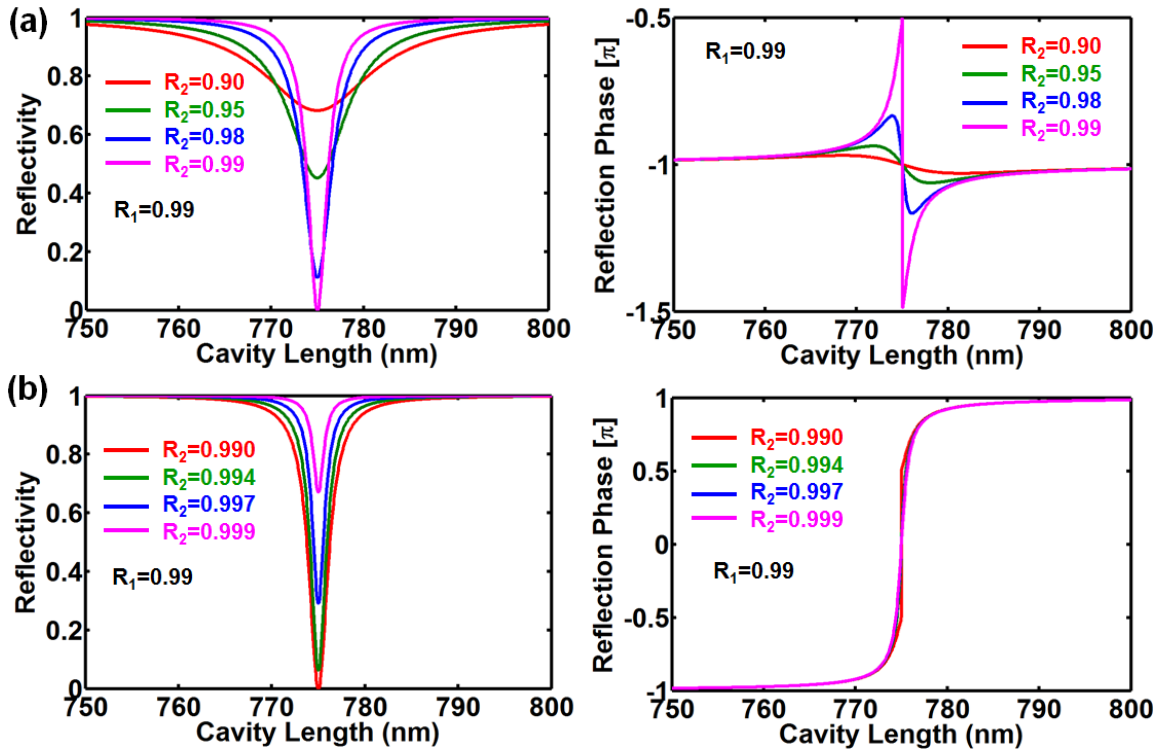
Figure 4.4 (a) Reflection spectrum and reflection phase spectrum of a FP etalon. (b) Reflectivity and reflection phase versus the cavity length, at a fixed incident wavelength of 1550 nm.

As shown in Figure 4.4(b), a small cavity length change (i.e. top mirror displacement) leads to a very large reflection phase change. This sets up a foundation for efficient phase tuning. As the cavity length is tuned, different reflection phase is obtained, however, the reflectivity is also different. In particular, the reflectivity at the cavity resonance point is low. This translates to the non-uniform reflectivity among the different pixels in the phased array, since each pixel might be assigned to provide different phase. This non-uniform reflectivity effectively scatters the light into the background, and lowers the contrast in the far field beam steered pattern. Thus a uniform reflectivity is desired for different reflection phase value.

To solve this, we can adjust the reflectivity of the top mirror R_1 and bottom mirror R_2 (R_1 and R_2 refers to the reflectivity in power). We first investigate the regime where $R_2 \leq R_1$. Figure 4.5 (a) plots the reflectivity and reflection phase versus the cavity length,

for a fixed $R_1 = 0.99$, and various R_2 from 0.9 to 0.99. In this regime, the total range of the accessible reflection phase value is small, less than π. This is thus not the desirable configuration.

Next, we explore the regime where $R_2 > R_1$. We fix $R_1 = 0.99$, and varies R_2 from 0.99 to 0.999. The results are shown in Figure 4.5 (b). As R_2 increases, the reflection dip at the resonance becomes shallower, and thus the reflectivity is more uniform across different cavity lengths; meanwhile the total range of the accessible reflection phase is 2π. The drawback in the regime is that the phase changes too rapidly with respect to the cavity length. This is not desirable for the practical system as the phase is too sensitive to the mirror displacement. To solve this, we keep R_2 at a high value (0.997), and decrease R_1. The results are shown in Figure 4.5 (c). As R_1 decreases, the reflection dip at the resonance gets shallower, and meanwhile the phase change becomes less sensitive to the cavity length change. This is thus a good operation regime. To get a full picture on how R_1 and R_2 influences the reflection property, the reflectivity at the cavity resonance (reflectivity dip) is plotted as a function of R_1 and R_2, shown in Figure 4.6(a); this can quantify the uniformity of the reflectivity across different cavity length. To quantify the phase change sensitive towards the cavity length change, we calculate the minimum mirror displacement in order to reach a phase shift of 1.6π. This is plotted as a function of R_1 and R_2, shown in Figure 4.6(b). From these results, it looks like that keeping R_2 above 0.997, and R_1 in between 0.9 to 0.95 can achieve a good balance between the reflectivity uniformity and phase change sensitivity. In this case, the reflectivity at the resonance is larger than 0.79, and the required actuation distance to reach 1.6π phase change is less than 40 nm. This looks to be a good operation regime for the optical phased array.



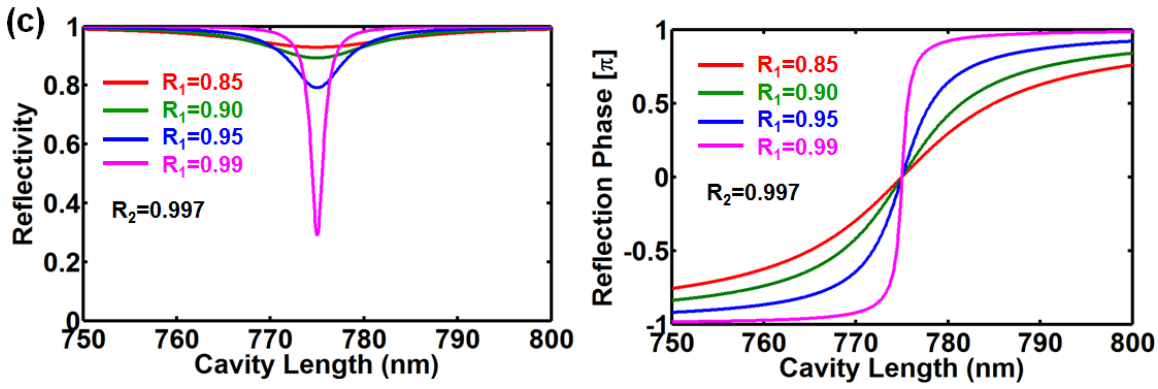


Figure 4.5 FP cavity reflectivity and reflection phase versus cavity length. The incident light wavelength is 1550 nm. R_1 and R_2 is the reflectivity of the top mirror and bottom mirror in power respectively.

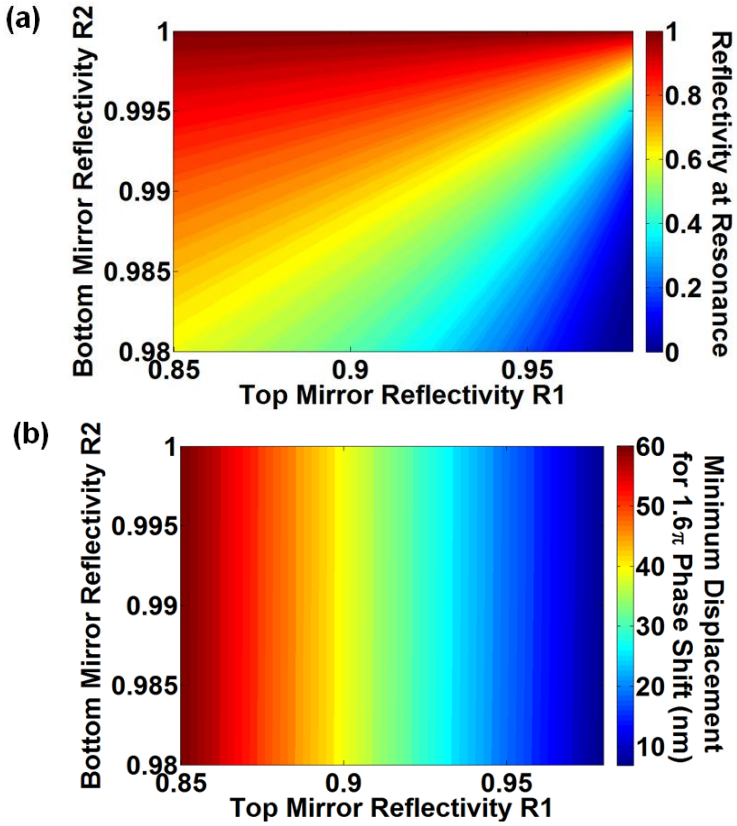


Figure 4.6 (a) Reflectivity contour of the FP cavity at resonance as a function of the top mirror reflectivity R_1 and bottom mirror reflectivity R_2. (b) Minimum mirror displacement to reach a reflection phase shift of 1.6π, as a function of R_1 and R_2.

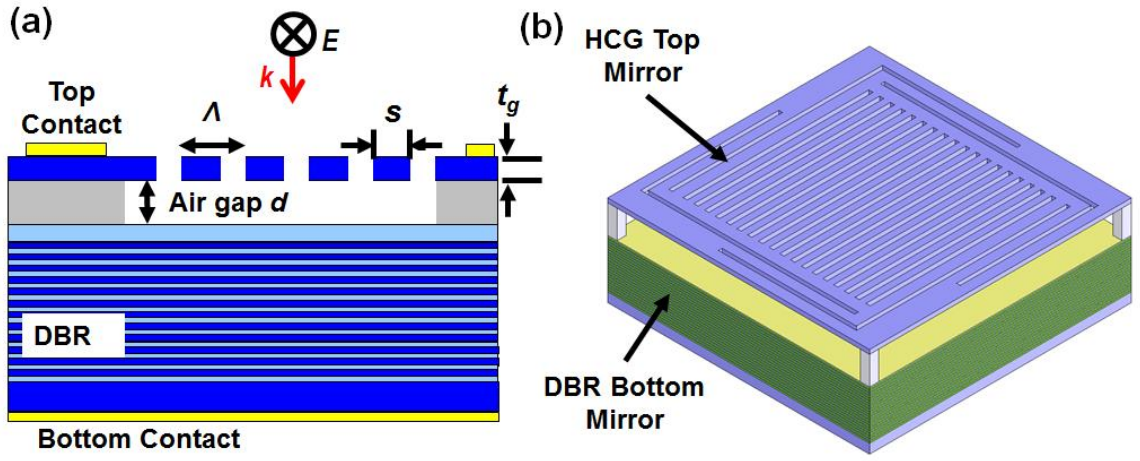
The above selection of R_1 and R_2 for the FP cavity leads to the concept of all-pass filter (APF). The all-pass filter is essentially an asymmetric Fabry-Perot (FP) etalon with carefully designed reflectivity of the top and bottom mirror. By actuating the top mirror to tune the length of the etalon across its FP resonance, the reflection phase of the surface normal incident light experiences a continuous phase change approaching 2π, while the

reflection beam power can maintain nearly the same as the incident light. To have an APF, the reflectivity of the bottom mirror should be as high as possible; while the top mirror should not be too high.

To implement the APF, HCG is used as the top reflector and DBR is used as the bottom reflector. Figure 4.7 (a)(b) show the schematic of an individual pixel of the optical phased array. Figure 4.7 (c) shows the schematic of the 8x8 phased array. The device is fabricated on a GaAs epitaxial wafer. The HCG is defined by electron beam lithography, and followed by a reactive ion etch on a p-doped $Al_{0.6}Ga_{0.4}As$ epitaxial layer, which is on top of an intrinsic GaAs sacrificial layer and 22 pairs of $GaAs/Al_{0.9}Ga_{0.1}As$ n-doped DBR. The top metal contacts (p-type contact), fanned-out lines and the wire-bonding metal pad are defined for each pixel on the HCG layer. A deep dry etch is then applied to electrically isolate the pixels from each other. Metal is uniformly evaporated on the backside to form the n-type contact. Wet chemistry is used to selectively etch the sacrificial layer (GaAs) against the AlGaAs HCG and the etch-stop layer, followed by a critical point drying. This thus forms FP cavities with the suspended HCG as top mirror and DBR as bottom mirror.

To form the all-pass filter, the reflectivity of the DBR is designed to be >0.9975 and the HCG ~ 0.9. We first design and fabricate the one-dimensional HCG. The HCG period, bar width and thickness is designed to be 1150 nm, 700 nm and 450 nm respectively. The incident light polarization is TE, i.e. electrical field along the HCG bars. The static cavity length is 700 nm; with the reflection phase response of the designed HCG, this corresponds to the cavity resonance wavelength of ~ 1550 nm. Each HCG mirror is 20 μm by 20 μm in size, and 8x8 individual pixels form the whole optical phased array, with the pitch ~33.5 μm. Each pixel is electrically addressable through the metal fanned-out lines. Figure 4.8 shows the SEM image of the fabricated device.

Indium is used to bond the whole chip onto a chip carrier, and each individual bonding pad (for each individual pixel) on the chip is wired-bonded to the contacts on the chip carrier. A printed circuit board is used to electrically interface with the chip carrier and thus the phased array. Figure 4.9 shows the image of the assembly of the optical phased array system.



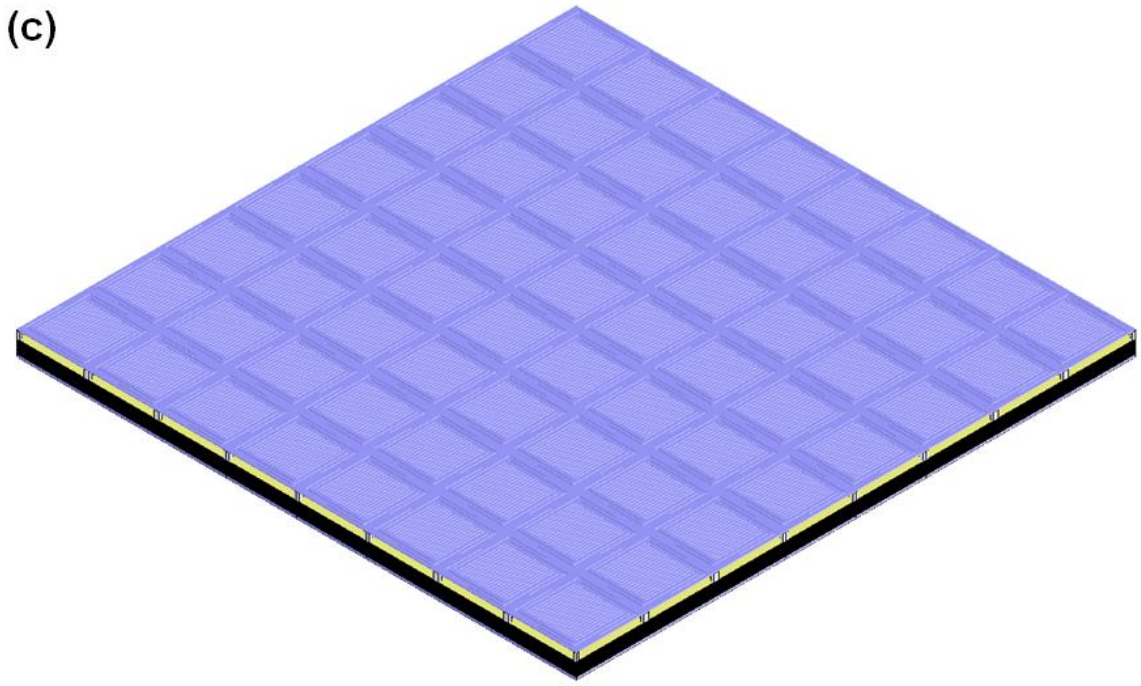


Figure 4.7 (a)(b) Schematic of an individual pixel of the optical phased array (with one-dimensional HCG). The $Al_{0.6}Ga_{0.4}As$ HCG and 22 pairs of $GaAs/Al_{0.9}Ga_{0.1}As$ DBR serve as the top and bottom reflector of the Fabry-Perot etalon. The incident light is surface normal to the etalon, and polarized parallel to the grating bar. (c) Schematic of an 8x8 optical phased array.

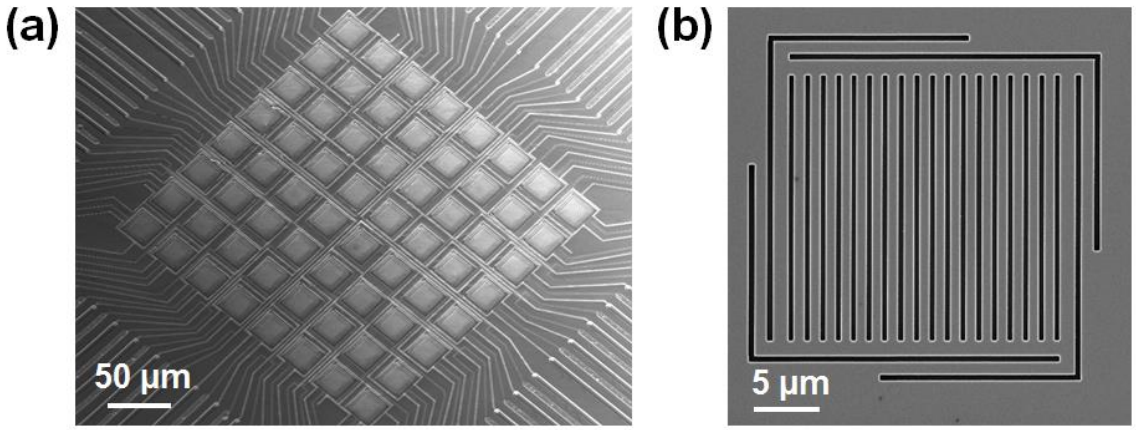


Figure 4.8 SEM image of an 8x8 optical phased array. Each pixel is an HCG-APF, which can be individually electrically addressed by the fanned-out metal contacts. The pitch of the HCG mirror is ~33.5 µm. (b) Zoom-in view of the HCG mirror in a single pixel. The HCG mirror size (without the MEMS) is 20 µm by 20 µm.

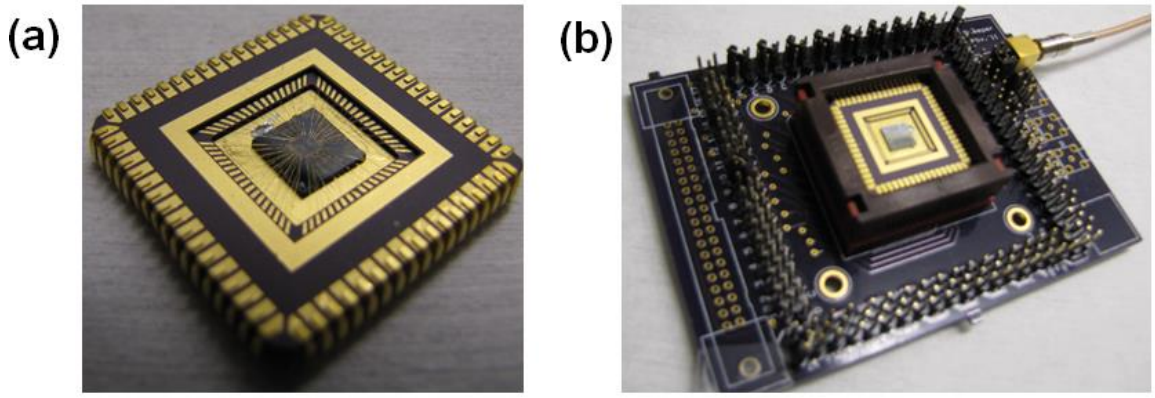


Figure 4.9 Image of an assembly of the optical phased array system. The chip is bonded on the chip carrier (a), which is hosted by a printed circuit board (b).

4.2 Device Characterization and Beam Steering Experiment

4.2.1 Small Actuation Distance for Large Phase Shift

To characterize the performance of the HCG-APF, the reflection and reflection phase versus incident light wavelength or cavity length is measured. A 1550 nm tunable laser is used as the light source; the cavity length of the APF can be changed by applying a reversed electrical bias between the p-doped HCG and n-doped DBR such that the HCG is actuated. Figure 4.10 shows the reflection spectrum of a single HCG APF of the array, for different reversed bias. As the reversed bias increases, the cavity length decreases, resulting in a blue-shift of the resonance wavelength. The measured reflection spectrum is fitted with the standard FP etalon reflection formulation, and the top mirror and bottom mirror's reflectivity is extracted. The reflectivity of the bottom DBR is extracted to be 0.9952~0.9977, and the reflectivity of the top HCG increases from 0.955 to 0.976 as the wavelength decreases. This HCG reflectivity is higher than the design value due to an inadvertent inaccuracy in electron beam lithography and etching process.

Figure 4.10 provides a way to evaluate the HCG displacement versus the actuation voltage. To do this, we need to convert the cavity resonance wavelength to the displacement of the HCG. The relationship between the cavity length and the resonance wavelength can be calculated with the consideration of the phase dispersion of the mirror reflection. Figure 4.11 (a)(b) shows the reflection and reflection phase spectrum of the designed DBR and the designed HCG; Figure 4.11 (c) shows the relationship between the cavity length and the resonance wavelength. It is seen that their relationship follows the phase dispersion of the DBR and HCG. Around the central reflection band of the DBR, the cavity length and the resonance wavelength has a linear relationship. Note that the linear relationship is valid as we operate the HCG-APF at the central reflection band of the DBR. However, this slope depends on the exact HCG dimension. This exact slope can be fitted based on the other characterizations. Later, we will show that the undamped mechanical resonance frequency of the HCG MEMS is ~0.53 MHz. Based on Eq. 4.5 and 4.8, one can plot out the relationship between the mirror displacement and actuation

voltage. This can be used to compare the data obtained here, and fit the slope. The slope is fitted to be 1.25. Figure 4.12 shows the HCG displacement versus actuation voltage, extracted by the two methods. They match with each other well.

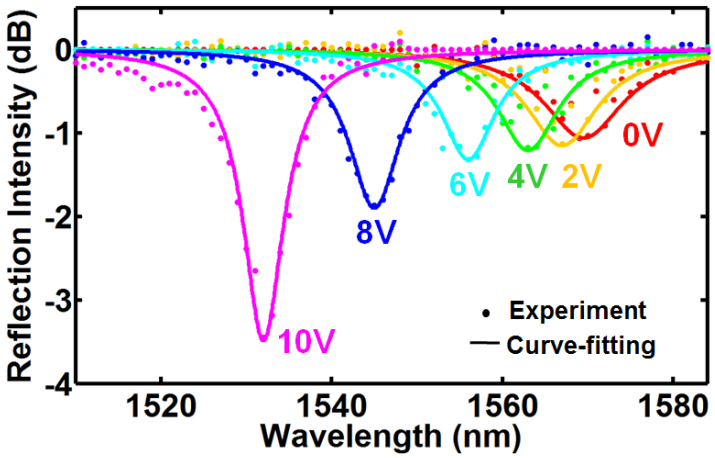


Figure 4.10 Reflection spectrum of an HCG-APF with different actuation voltages. As the reversed bias increases, the cavity length decreases, resulting in a blue-shift of the resonance wavelength.

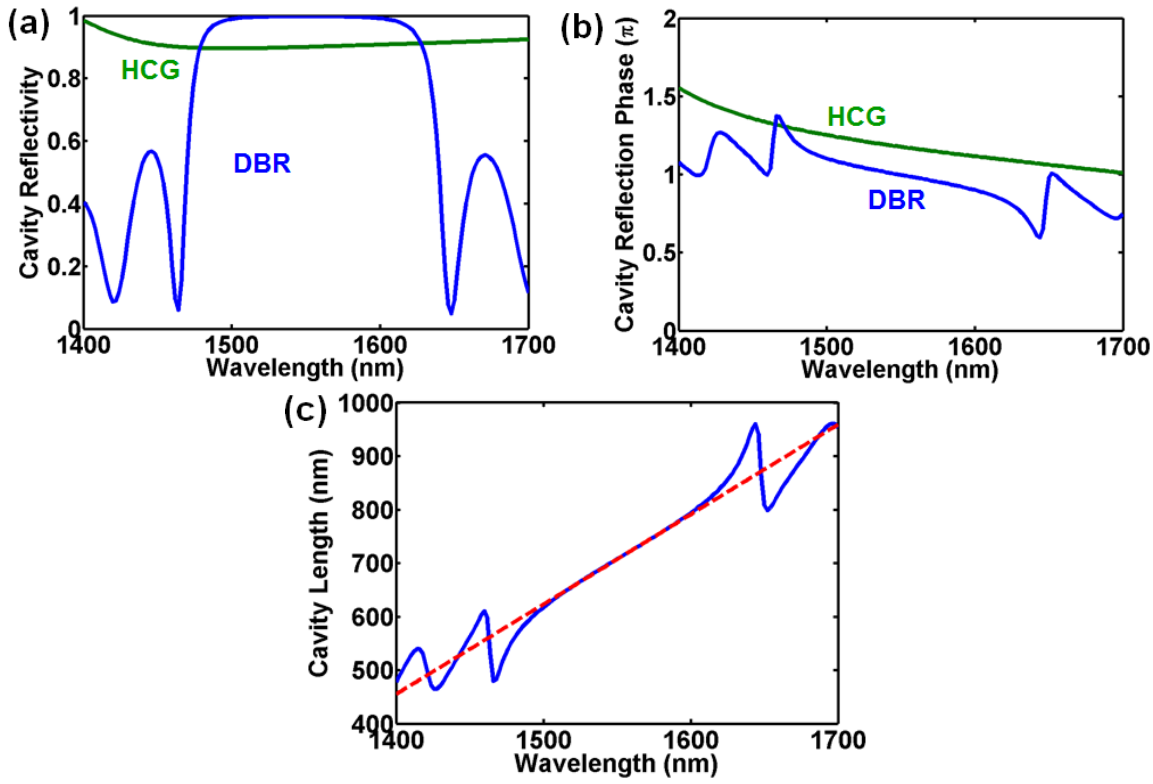


Figure 4.11 (a) Reflection spectrum and (b) reflection phase spectrum of the designed DBR and HCG. (c) Relationship between the FP cavity length and the wavelength. Their relationship follows the phase dispersion of the DBR and HCG. Around the

central reflection band of the DBR, the cavity length and the resonance wavelength has a linear relationship.

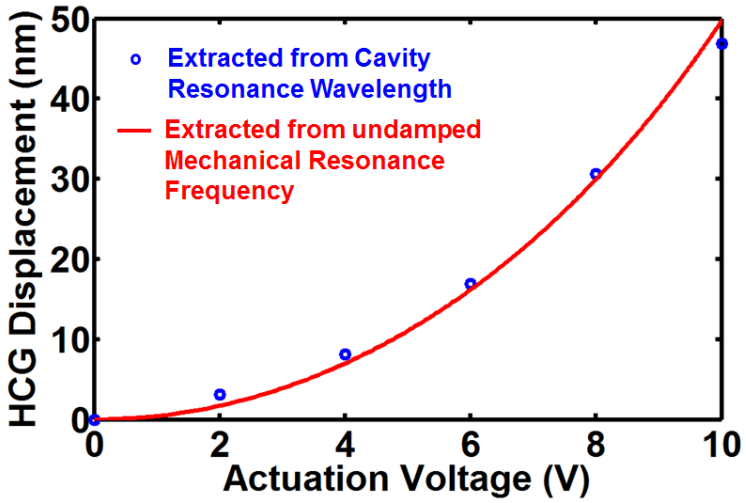


Figure 4.12 HCG displacement versus actuation voltage. The blue dots are extracted from the cavity resonance wavelength measured by the optical reflection spectrum. The red trace is extracted from the undamped mechanical resonance frequency and Eq. 4.5 and Eq. 4.8.

The reflection phase of the HCG APF is characterized with a Michelson interferometer, schematically shown in Figure 4.13. This setup also characterizes the beam steering performance. In Figure 4.13, the blue line and red line indicates that the light polarization is in x and y direction respectively. The polarization beam splitter 1 is used to split light into two parts. One part is delivered to the HCG phased array, which will carry the information of its reflection phase; it thus serves as the signal beam. The other part serves as the reference beam. These two beams combine at polarization beam splitter 4, and goes into a CDD. The phase information of the signal beam is effectively converted into intensity information at the CDD. The reflection phase of the whole HCG APF array can thus be measured at the same time. Besides going into the interference path, part of the signal beam is used for beam steering measurement. In this setup, $\lambda/2$ half waveplate 1 and 3 is used to adjust the power ratio between x and y polarization. The $\lambda/2$ half waveplate 2 and 4 is used to switch the light polarization from x to y or y to x. The HCG phased array is aligned 45^{o} with respect to x and y. The Faraday rotator is used to rotate the y polarization by 45^{o}, so the polarization of the incident light for the TE HCG is aligned to grating bar direction. This method ensures that there is no power lost due to the beam splitting cube.

Figure 4.14 shows the reflection phase of a single pixel versus applied voltage. The actual displacement of the HCG is also labeled in the horizontal axis. A total phase change of ~1.7 π is achieved within 10V actuation voltage range at a wavelength of 1556 nm; this corresponds to a displacement of ~50 nm of the HCG. This demonstrates the high phase tuning efficiency of the APF. The reflectivity of the DBR and HCG can be extracted from this phase measurement by a curve fitting. They are extracted to be 0.9977

and 0.935 respectively, in a reasonable match with the value extracted from the reflection spectrum measurement.

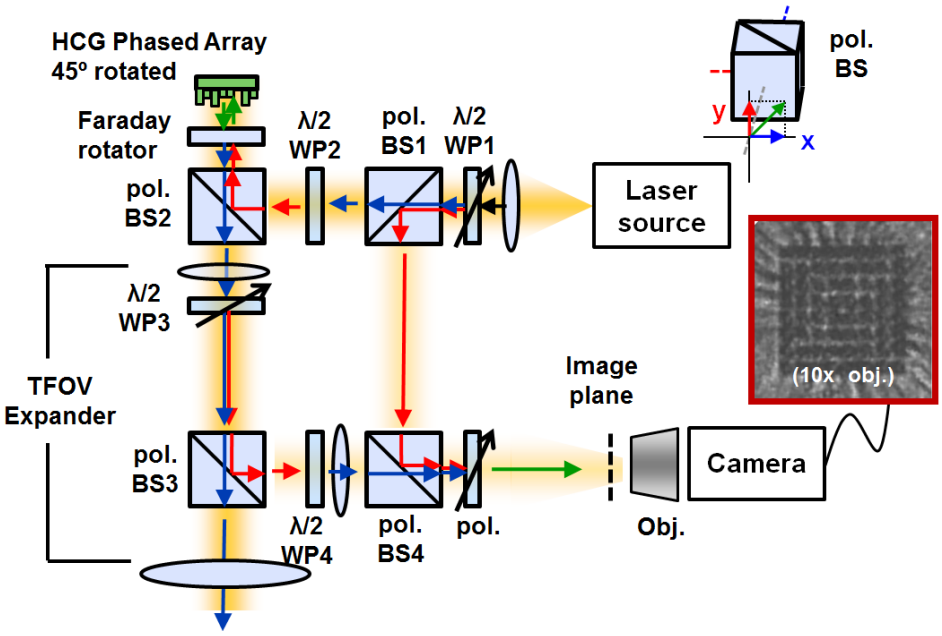


Figure 4.13 Experimental setup to characterize the reflection phase of the HCG APF phased array, as well as the beam steering performance. WP, wave plate; pol. BS, polarization beam splitter; Obj., objective; pol., polarizer; TFOV, total field of view.

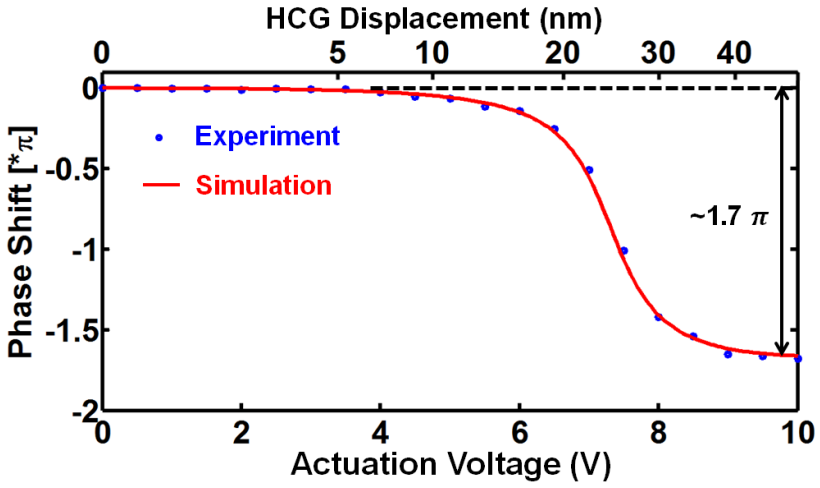


Figure 4.14 Reflection phase shift versus applied voltage on a single HCG-APF of the phased array. ~1.7 π phase shift is achieved within 10V actuation voltage range at a wavelength of 1556 nm; this corresponds to a displacement of ~50 nm of the HCG. This APF design enables a small actuation distance for a large phase change. The experimental data (blue dots) are fitted with the simulation results with the DBR and HCG reflectivity of 0.9977 and 0.935 respectively (red curve).

4.2.2 High Speed Phase Tuning

The phase tuning speed can be evaluated by the mechanical resonance frequency f_r of the HCG MEMS mirror. Two different methods are used to characterize the mechanical resonance frequency. The first one is the laser Doppler velocimetry (LDV). A white-noise electrical signal is used to actuate the HCG. A laser is incident onto the HCG, and the Doppler shift of the reflection beam is recorded with respect to time, followed by a Fourier transform to reveal the information in the frequency domain, shown in Figure 4.15 (a). The mechanical resonance frequency f_r is 0.53 MHz. Alternatively, a step voltage is applied to actuate the HCG mirror, and time resolved phase measurement can be used to extract the mechanical resonance frequency. This time resolved phase trace is shown in Figure 4.15 (b). The applied voltage changes from 6 V to 7 V at 13.7 μs and back to 6 V at 33.8 μs. A damped second harmonic oscillator model is used to analyze the trace and extract the mechanical resonance frequency. Eq. 4.10 shows the equation governing the harmonic oscillator.

$$\frac{\partial^2}{\partial t^2}x(t) + 2\gamma\frac{\partial}{\partial t}x(t) + \omega_o^2 = 0 \tag{4.10}$$

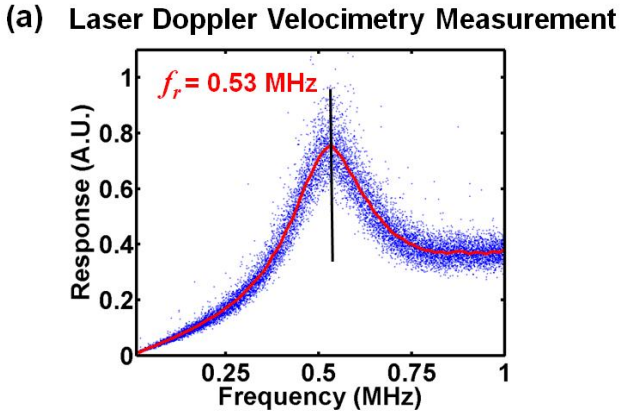
where x(t) describes the displacement of the HCG mirror, γ is the damping factor, and ω_o is the undamped mechanical resonance frequency. The solution of Eq. 4.10 can be written into the following form,

$$x(t) = Ae^{i\omega t-\gamma t} + Be^{-i\omega t-\gamma t} \tag{4.11}$$

where ω is the underdamped mechanical resonance frequency, and $\omega^2 = \omega_o^2 - \gamma^2$. The boundary condition for Eq. 4.10 can be $\frac{\partial x(t)}{\partial t}|_{t=0} = 0$, and thus Eq. 4.11 becomes

$$x(t) = (A+B)[\cos(\omega t) + \gamma/\omega \sin(\omega t)]e^{-\gamma t} \tag{4.12}$$

Together with the relationship between HCG reflection phase and HCG displacement in Figure 4.14, Eq. 4.12 can then be used to fit the time-resolved phase trace in Figure 4.15 (b). ω_o and γ is fitted to be $2\pi * 0.53$ MHz and 0.6993 MHz respectively, leading to $\omega = 2\pi * 0.52$ MHz. This is in a good match with the LDV measurement.



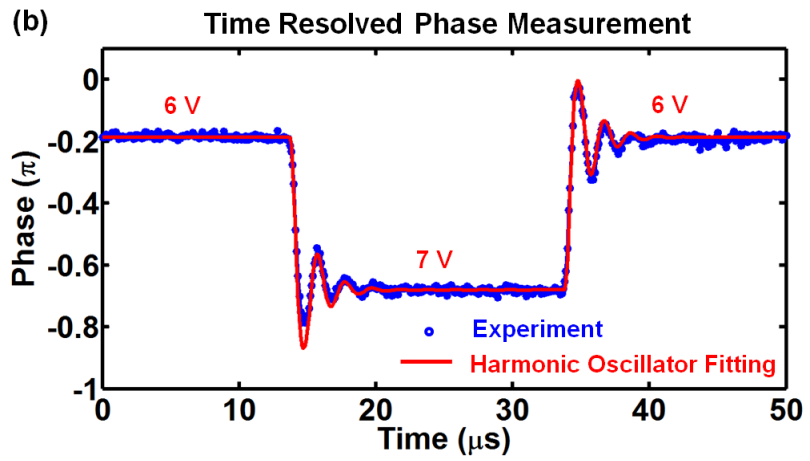


Figure 4.15 (a) Laser Doppler velocimetry measurement to characterize the mechanical resonance frequency of the HCG MEMS mirror. (b) Time resolved phase measurement of the HCG APF with a step voltage actuation signal. The blue dots are recorded in the experiment, and red traces are the simulated fitting curve from the second harmonic oscillator model.

In Figure 4.15 (b), ringing in the phase response is seen. This is not ideal in the practical system. To reduce the ringing, one can break the single voltage step into two steps. Instead changing directly from 6 V to 7 V in the above example, we first change the voltage from 6 V to 6.5 V, hold it for 1 μs, and then change from 6.5 V to 7 V. The same applies for case when the voltage changes form 7 V to 6 V. 1 μs corresponding to half of the ringing period, and thus the individual ringing from these two separate steps would have destructive interference, leading to an overall reduced ringing. This comparison is shown in Figure 4.16. The phase settles down within 3.4 μs with the two-step voltage control.

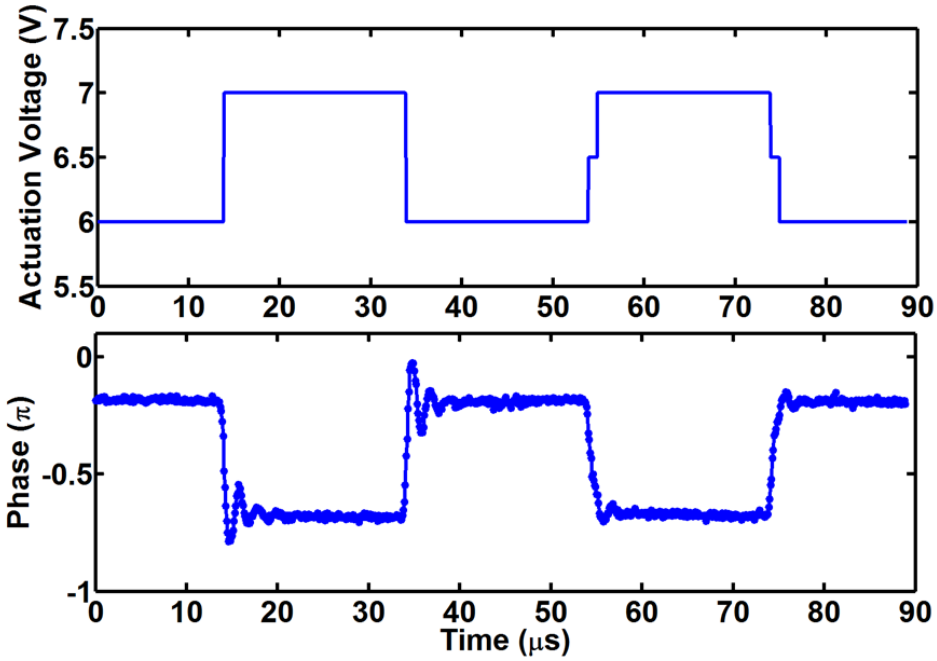


Figure 4.16 Comparison of the ringing between a one step and two step voltage control. In the two step voltage control case, the time interval between the two different step is 1 μs, corresponding to half of the ringing period. The individual ringing from these two separate steps would have destructive interference, leading to an overall reduced ringing.

4.2.3 Beam Steering Experiment

Beam steering in far-field is achieved by creating the desired near-field phase front of the reflection beam on the whole 8x8 phased array. By controlling the applied voltage on each individual pixel of the HCG-APF array, the near-field phase pattern can be generated. The experiment setup is shown in Figure 4.13. The maximum steering angle is achieved when the phase is alternative between the pixels. This is also defined as half the total field of view (TFOV). This angle is determined by the smallest individually addressable phase element.

$$TFOV = 2sin^{-1}[\lambda/(2pitch)] \tag{4.13}$$

Since the pitch of the pixel of the 8x8 array is 33.5 μm, the TFOV is 2.65^o, for the operation wavelength of 1.55 μm. To increase this, a two-lens system is used to magnify the beam steering angle. This angular magnification ratio is set to be 3.45 in the current experiment, and thus the TFOV is 9.14^o.

The full width beam divergence at the half power point (FWHM) is determined by the total size of the array. This can be approximately expressed as

$$FWHM \sim \frac{\lambda}{D} \tag{4.14}$$

Since the array size is 8x8, the total size D^2 of the array is $(8 \times 33.5\ \mu m)^2$. Thus the FWHM is 0.33^o, which is magnified to be 1.14^o. The measured value is ~1^o, matching well with this calculation.

Figure 4.17 shows various near-field phase patterns on the phased array, and the corresponding measured far field pattern. The strong zeroth order beam is due to the relatively low filling factor of the phased array (~36%). Quite a large portion of light gets directly reflected from the background without phase shift, contributing to the zeroth order beam. With this consideration and the near-field phase patterns, Fourier optics is applied to calculate the far-field patterns, also shown in Figure 4.17. The experiment is in good agreement with the simulation.

Figure 4.18 shows the time-resolved beam steering measurement so as to characterize how fast the beam can be steered. At t=0, actuation voltage is applied to the phased array, and the intensity of the steered beam is measured versus time. Similar to the time-resolved phase measurement, ringing is seen at a period of ~2 μs. This is in good agreement with the resonance frequency measured by LDV and time-resolved phase measurement. This truly demonstrates the fast beam steering of this phased array.

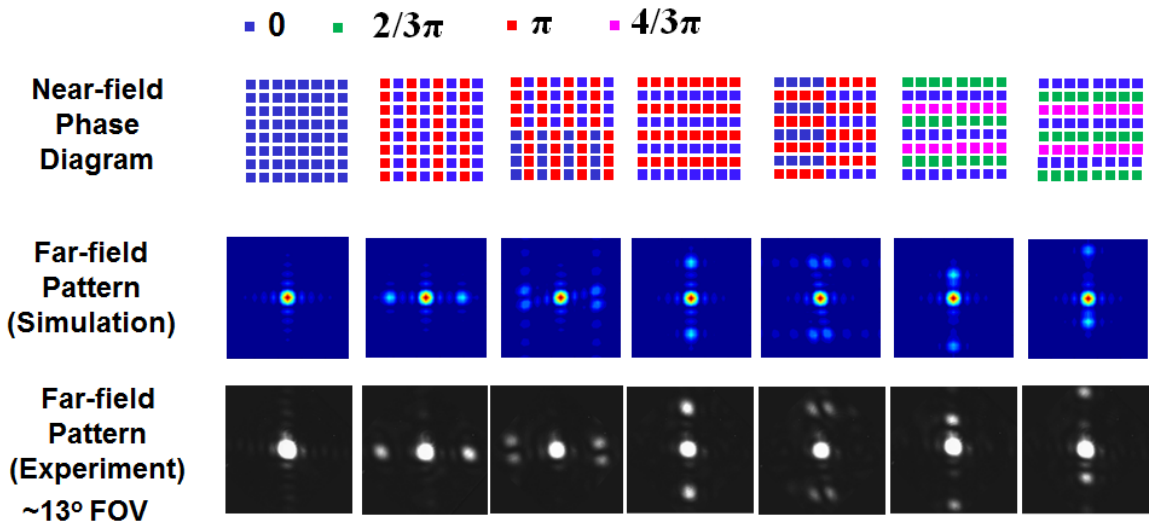


Figure 4.17 Beam steering experiment. (a) Near-field phase pattern created by the HCG-APF optical phased array. (b) The corresponding far-field pattern calculated by Fourier optics. (c) Experimentally measured far-field pattern, in good agreement with the calculation. The strong zeroth order beam is due to the relatively low filling factor of the phased array (~36%). The light that does not hit on the HCG-APF gets directly reflected without phase shift, contributing to the zeroth order beam. The field of view (FOV) of the image window is 13^{o} x 13^{o}. The wavelength is 1570 nm.

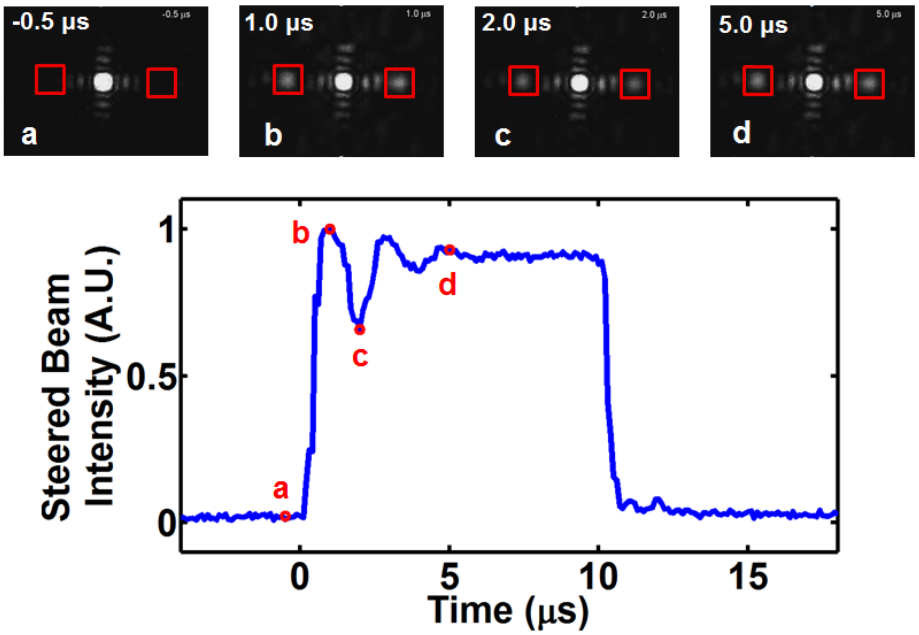


Figure 4.18 Time resolved beam steering experiment to characterize the beam steering speed. At t=0, actuation voltages are applied to the phased array, and the intensity of the steered beam (indicated by the red box) is plot versus time. Ringing is seen at a period of ~2 μs, in good agreement with the resonance frequency measured by LDV and time-resolved phase measurement. The far field images a~d are taken at t=-0.5 μs, 1 μs, 2 μs, and 5 μs, respectively.

4.2.4 Two-dimensional HCG

To suppress the zeroth order beam, a high-filling factor phased array is designed by grouping four pixels together without isolation trenches in between. Furthermore, two dimensional HCG is designed and fabricated. The symmetric HCG can overcome the polarization sensitivity of the one-dimensional HCG. The two-dimensional HCG grid can also be actuated more uniformly. Figure 4.19 shows the SEM image of the two-dimensional HCG mirror for the HCG-AFP array. The HCG mirror size is 20 μm by 20 μm, and the filling factor is ~47%.

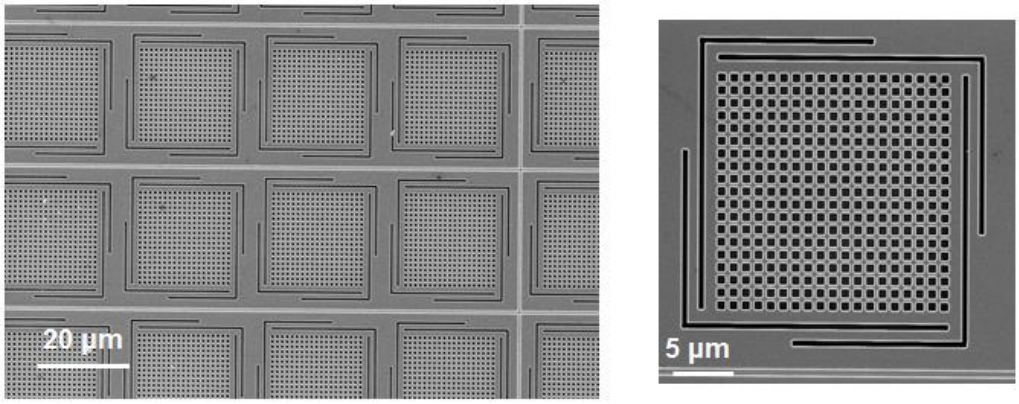


Figure 4.19 SEM image of the two-dimensional HCG mirror for HCG-AFP array. The individual pixel is shown on the right. The HCG mirror size (without MEMS) is 20 μm by 20 μm.

Figure 4.20 shows the reflection spectrum of the two-dimensional HCG-AFP for two orthogonal polarizations (x, y direction corresponds to the two directions of the grid in the two-dimensional HCG). The reflection spectrum of the two polarizations matches well with each other, indicating a symmetric geometry of the HCG. Their slightly difference is due to an inadvertent asymmetric in electron beam lithography. Beam steering is experimentally demonstrated with the two-dimensional HCG-AFP, shown in Figure 4.21. Compared to Figure 4.17 using the 8x8 array, more power is being beam steered in this case.

To further optimize the beam steering performance, a mirco-lens array can be placed in front of the phased array, which can focus the input beam onto the HCG mirror of each pixel. This effectively increases the filling factor to 100%, and will ultimately suppress the zeroth order beam.

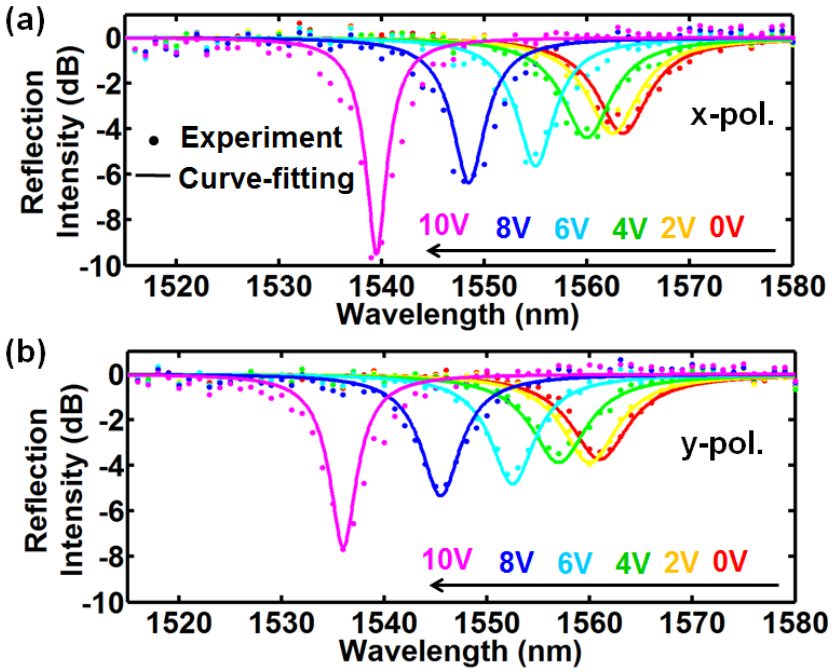


Figure 4.20 Reflection spectrum of a two-dimensional HCG-APF with different actuation voltages. The polarization of the incident light is aligned to x polarization and y polarization in (a) and (b) respectively. The x, y direction corresponds to the two directions of the grid in the two-dimensional HCG. The reflection spectrum of the two polarizations has a good match with each other. The slightly difference is due to an inadvertent asymmetric in electron beam lithography.

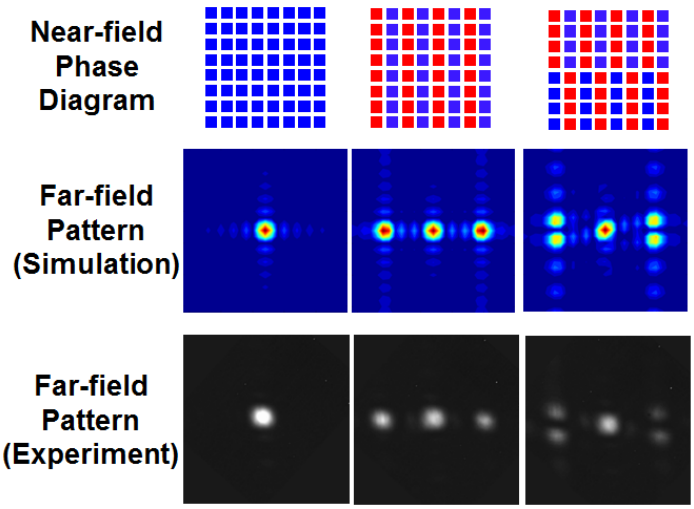


Figure 4.21 Beam steering experiment of the optical phased array using two-dimensional HCG as the top mirrors of the APF. (a) Near-field phase pattern created by the HCG-APF optical phased array. (b) The corresponding far-field pattern calculated by Fourier optics. (c) Experimentally measured far-field pattern, in good agreement with the calculation.

4.3 HCG Optical Phased Array on Silicon Platform

The previous discussion focuses on III-V material platform. The advantage of using III-V material platform is that the wafer can be epitaxial grown, and thus the thickness of each layer can be well controlled. While this is good for demonstration of concept, silicon material platform is desirable to scale up the array size. Phased array using HCG as piston mirror has been demonstrated with excellent beam steering performance [91]. For the APF, Si/SiO_2 DBR can be deposited on the Si substrate, followed by the deposition of SiO_2 as sacrificial layer and Si as the HCG layer. DUV lithography is used for to define the HCG. This greatly reduces the fabrication cost compared with electron beam lithography. The fabrication and characterization of a 32 by 32 array with a filling factor of 75% is currently a research topic.

4.4 Summary

In this chapter, optical phased array using HCG all-pass filters are discussed. The key advantage of using HCG APF is its high efficient phase tuning, i.e. small HCG MEMS mirror displacement (~50 nm) for large phase change (~1.7 π), and small voltage actuation (<10 V) for fast beam steering (0.53 MHz). 0.53 MHz beam steering is experimentally demonstrated with the 8x8 phased array. This high speed optical phased array will be extremely useful for high speed imaging, LIDAR and optical switch. Both one-dimensional and two-dimensional HCG is fabricated and demonstrated for beam steering.

The on-going research in this topic includes the dramatically up-scaling of the array size, improving its filling factor, and integrated mirco-lens array on top the phased array. To truly make the effective filling factor to 100% for the phased array, the filling factor of the micro-lens array needs to be 100%. This is not easy to achieve in the conventional micro-lens array. One way to solve this problem is to use HCG lens [38-40] array. Thanks to the planar geometry of the HCG lens, its filling factor can be 100%. This is conceptually shown in Figure 4.22. This is promising for high-speed, large-scale, high-performance beam former and beam steering.

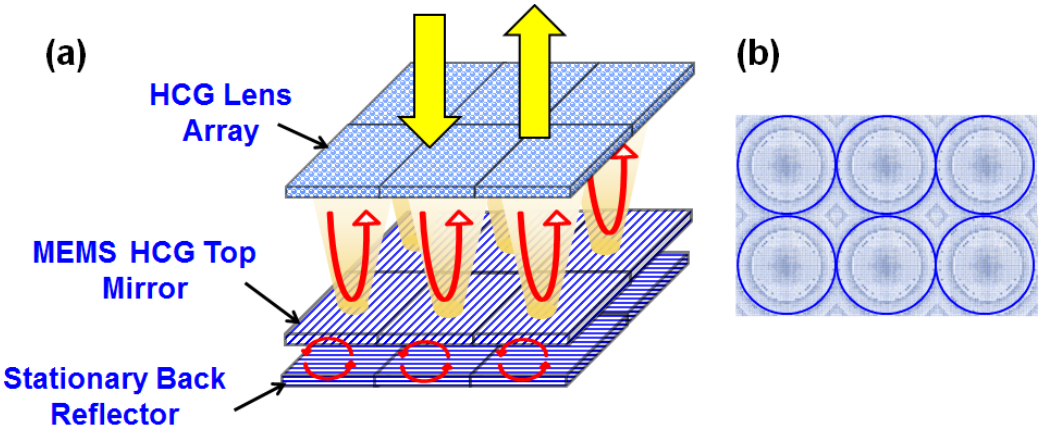


Figure 4.22 (a) Integrated optical phased array with HCG lens array. The effective filling factor of the phased array can be made to be 100%. (b) Layout of the HCG lens array.

Chapter 5
Tunable High Contrast Grating Detector

High contrast grating vertical-cavity surface-emission lasers (HCG VCSELs) is the best device that truly demonstrates the broadband high reflection of HCG, since extremely high mirror reflectivity is required for the laser to lase. The big success of HCG tunable VCSEL proves the HCG to be a robust integrated component, and in return, motives the optimization of HCG for a broader tuning range, lower lasing threshold, higher tuning speed, and larger fabrication tolerance. 26.3 nm wavelength tuning has been demonstrated in 1550 nm HCG tunable VCSEL, with >1.4 mW single-mode output under room temperature continuous-wave (CW) operation over most of the tuning range [29]. Furthermore, transmission of directly modulated 10 Gbps on-off-keying signal over 100 km dispersion-compensated single-mode fiber is demonstrated with the HCG VCSEL [29]. The great advantage of 1550 nm HCG tunable VCSEL is its simple coupling scheme, continuous tuning characteristics, low power consumption and low manufacturing cost. Besides the conventional application in optical communications, the widely tunable HCG VCSEL can also be applied as a wavelength-swept source in optical coherent tomography, light ranging and detection, and gas sensing etc.

While a lot of focus is concentrated on the lasing characteristics of the HCG VCSEL, it is also worthwhile to look at the other side of the laser, i.e. its photodetector characteristics. By applying a reversed bias across the laser junction, the VCSEL becomes a photodetector. Due to the high quality factor of the cavity, this is a resonance enhanced photodetector, i.e. the detector only has a large responsivity at the resonance wavelength of the cavity. By actuating the HCG, the resonance wavelength changes, and thus this acts as a tunable detector. This tunable detector can be controlled such that it only detects the desirable wavelength over a large wavelength range and rejects the other wavelengths.

Tunable detector provides a great flexibility and reconfigurability to the optical network. It is a key component for the time and wavelength division multiplexed (TWDM) passive optical network (PON), which is a second generation of PON, where time division multiplexed (TDM) and wavelength division multiplexed (WDM) scheme coexists. The first generation of PON is based on TDM scheme, where the central office sends out TDM data stream carried by a single wavelength. This data stream splits to multiple end users and each end user is assigned a time slot to receive the data and transmit its own upstream data. To increase the data bandwidth, WDM is added on top of the original TDM system. This is the second generation of PON, termed as TWDM-PON. In this system, the existing optical distribution network (ODN), which is the power splitter built in the first generation, will be reused. Thus one end-user might receive multiple wavelength channels, and the tunable detector is necessary to select the desired channel and block out others. The competing technologies against tunable detector would be the tunable filters in silicon photonic platform or standalone tunable filters. Either of this is less desirable due to the insertion loss. The HCG tunable detector essentially monolithically integrates the tunable filter and detector, which offers a smaller footprint and higher sensitivity due to a lower loss.

In this chapter, we experimentally demonstrate that a tunable HCG VCSEL also functions as a tunable photodetector by simply changing the bias polarity. The tunable detector shows a wide tuning range of 33.5 nm at 1550 nm, and a minimum spectrum width of 1.2 nm. We demonstrate that it can select a specific data channel at 1 Gb/s and reject adjacent channels that is 4 nm away. Higher detection speed is promising, as up to 10 Gb/s data rate is demonstrated with a fixed-wavelength HCG detector, which has a smaller device footprint.

The interesting part of this device is its bifunctionality. By changing the bias polarity, it can switch between laser mode and detector mode. Such dual functionality device with transmitter and receiver all-in-one is very unusual, as lasers and detectors are typically designed and fabricated separately for optimized performance. For low-cost applications, it is interesting to explore the use of one single device as both laser (transmitter) and detector (receiver) to increase routing flexibility in datacenters and reduce costs for fiber-to-the-home applications. We show an error-free 1 Gb/s link over 25 km single-mode fiber between two such tunable VCSELs, one is forward biased as a transmitter and being directly modulated, while the other is reversed biased as a receiver, or vice versa.

Another interesting part of the tunable detector is that it can be configured to self-track the wavelength of the incident light with an internal electrical feedback. As the detection wavelength can follow that of the incident light, the detector is resistant to the drifting of the incident light wavelength.

Yet another novel application of the tunable detector is high speed on-chip optical spectrometer. By sweeping the resonance wavelength, it can be used to characterize the optical spectrum of the incident light. The sweeping speed can be as high as the mechanical resonance frequency of the HCG MEMS mirror, from 100 kHz to MHz.

Section 5.1 is devoted to the basic design and characterization of the tunable HCG detector. Section 5.2 discusses the high speed characteristics of the tunable HCG detector, and its bifunctionality. Section 5.3 discusses the tracking detector and section 5.4 reports the high speed on-chip optical spectrometer. Section 5.5 summarizes this chapter.

5.1 Tunable High Contrast Grating Detector

The 1550 nm tunable HCG detector shares the same structure as the tunable HCG VCSEL, schematically shown in Figure 5.1. The structure is grown on an InP expitaxy wafer, which consists of 40-55 pairs of bottom n-doped AlGaInAs/InP DBR, an active region with multiple quantum wells, p-type cladding, tunnel junction, sacrificial layer, p-doped InP HCG layer and top contact layer.

The HCG is defined by electron beam lithography, followed by anisotropic vertical etching. Wet chemical etching and critical point drying is used to release the HCG structure so it is fully suspended in air as the top mirror for the cavity. The HCG is designed to have a broadband high reflectivity. For TE HCG, the duty cycle is ~0.35, and the desirable period and thickness is ~1050 nm and ~200 nm; for TM HCG, the duty

cycle is ~0.65, and the desirable period and thickness ~ 700 nm and 420 nm respectively. Similar with the HCG all-pass filter discussed in Chapter 4, the HCG structure and the semiconductor layer beneath it acts as a parallel capacitor transducer, and the HCG can be actuated towards the semiconductor layer by applying voltage between the tuning contact and the laser / photodetector contact.

The AlGaInAs multiple quantum wells are placed at the anti-node position of the optical field in the cavity. This facilitates a large overlap of the electrical field and the quantum wells, leading to a large absorption of the light at the resonance wavelength. The HCG mirror size is 16x16 μm^2, and a lensed fiber or a cleaved fiber is used to couple light into the cavity. Due to the one-dimensional HCG design, the photodetector is polarization sensitive. A polarization controller is used to adjust the polarization of the input light to maximize the photocurrent. The polarization dependence can be eliminated by designing the HCG with two-dimensional symmetricity.

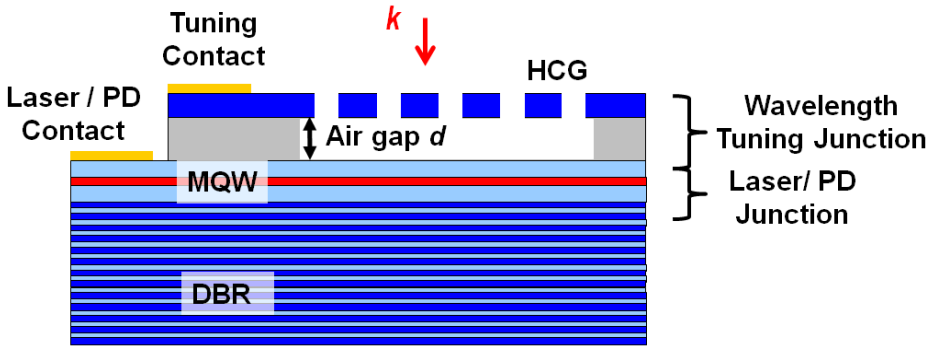


Figure 5.1 Schematics of a tunable HCG detector. It comprises of an electrostatically actuable HCG as the top mirror and a DBR as the bottom mirror, with a multiple quantum well structure as the absorption layer. By switching the bias polarity on the laser / PD junction, this device can operate as a laser or detector. PD, photodetector. MQW, multiple quantum well.

Figure 5.2 shows the responsivity of an HCG detector versus input light powers, at the cavity resonance wavelength of 1558.6 nm. Thanks to the resonance enhanced absorption, a responsivity as large as 1 A/W is obtained, corresponding to a quantum efficiency of 80%.

As is mentioned before, the HCG can be electrostatically-actuated so as to change the cavity length and thus the resonance wavelength. Due to the high quality factor of the Fabry-Perot cavity, the photodetector has a strong wavelength selectivity. Figure 5.3 (a) shows the relative responsivity spectrum versus different tuning voltages and Figure 5.3 (b) summarizes the resonance wavelength and spectrum width versus tuning voltage. A tuning range as large as 33.5 nm is achieved with a tuning voltage range of 6.1 V. A minimum spectrum width (full-width-at-half-maximum, FWHM) is 1.2 nm at 0 V tuning voltage. As the tuning voltage increases, the spectrum linewidth broadens and then shrinks again. This is due to the mechanical deformation of the HCG mirror, leading to a change of the quality factor of the cavity. With an optimization of the MEMS of the HCG, a uniform spectrum width across the tuning voltage is expected.

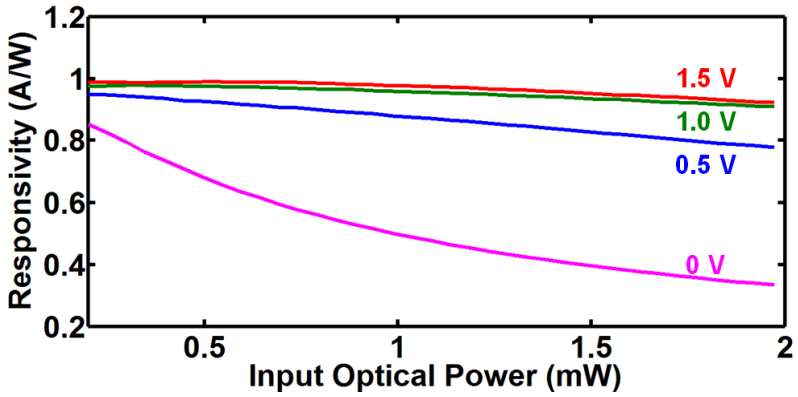


Figure 5.2 Responsivity of an HCG detector versus input light power at its resonance wavelength of 1558.6 nm, for various reversed bias across the photodiode junction. Light is coupled through a cleaved fiber into the detector. A maximum responsivity of 1 A/W is achieved.

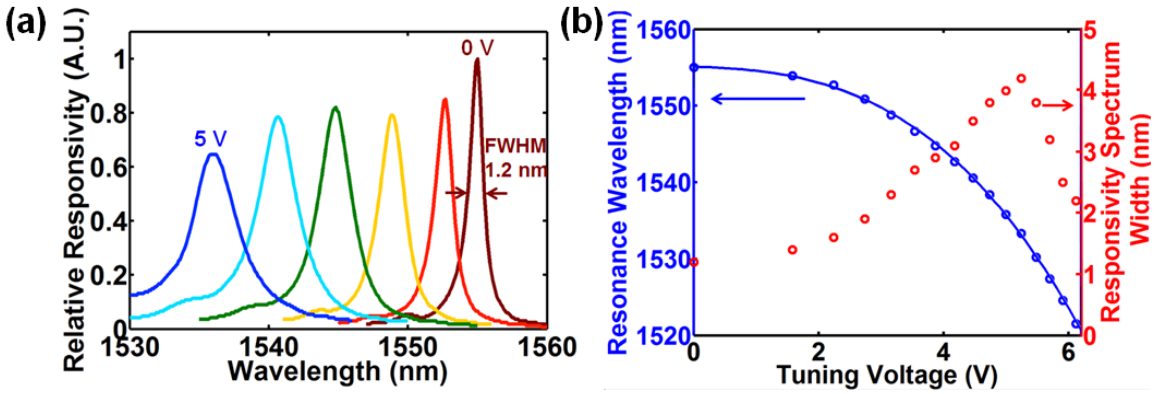


Figure 5.3 (a) Relative responsivity spectrum of an HCG detector versus tuning voltage. A minimum spectrum width (FWHM) of 1.2 nm is achieved at 0 V tuning voltage. (b) Resonance wavelength (blue) and responsivity spectrum width (red) versus tuning voltage. A 33.5 nm tuning range is achieved with a tuning voltage range of 6.1 V. The dots are experimental measured data, and the blue curve is a parabolic fitted curve for the resonance wavelength and tuning voltage. The non-uniformity in the responsivity spectrum width is due to the deformation of the HCG mirror when being actuated, and can be much improved with an optimization of the MEMS design for the HCG.

5.2 High Speed Tunable High Contrast Grating Detector and VCSEL

For high speed detection, it is important to reduce the parasitic RC. By shrinking the size of the mesa, parasitic capacitance can be reduced. After the mesa etch, BCB is used to planarize the surface to the height of the semiconductor mesa, and the backside contact is expanded to the surface of the mesa, so that it is in a ground-signal (GS) contact configuration for high speed operation.

To characterize the high speed detection capability of the HCG detector, the CW light from an instrumentation tunable laser is modulated by a Mach-Zehnder modulator (MZM), with 2^7-1 pseudo-random binary sequence (PRBS) data pattern from a bit error rate test bed (BERT). The electrical signal is amplified such that it is large enough to drive the MZM. The modulated light is polarization adjusted and incident on the HCG detector. A reversed voltage bias is applied to the detector through the DC port of the bias-tee. The photocurrent is extracted through the AC port of the bias-tee, and amplified by a RF amplifier. The signal is then detected by the BERT, and bit error rate (BER) is measured. Figure 5.4 (a) shows the experimental setup.

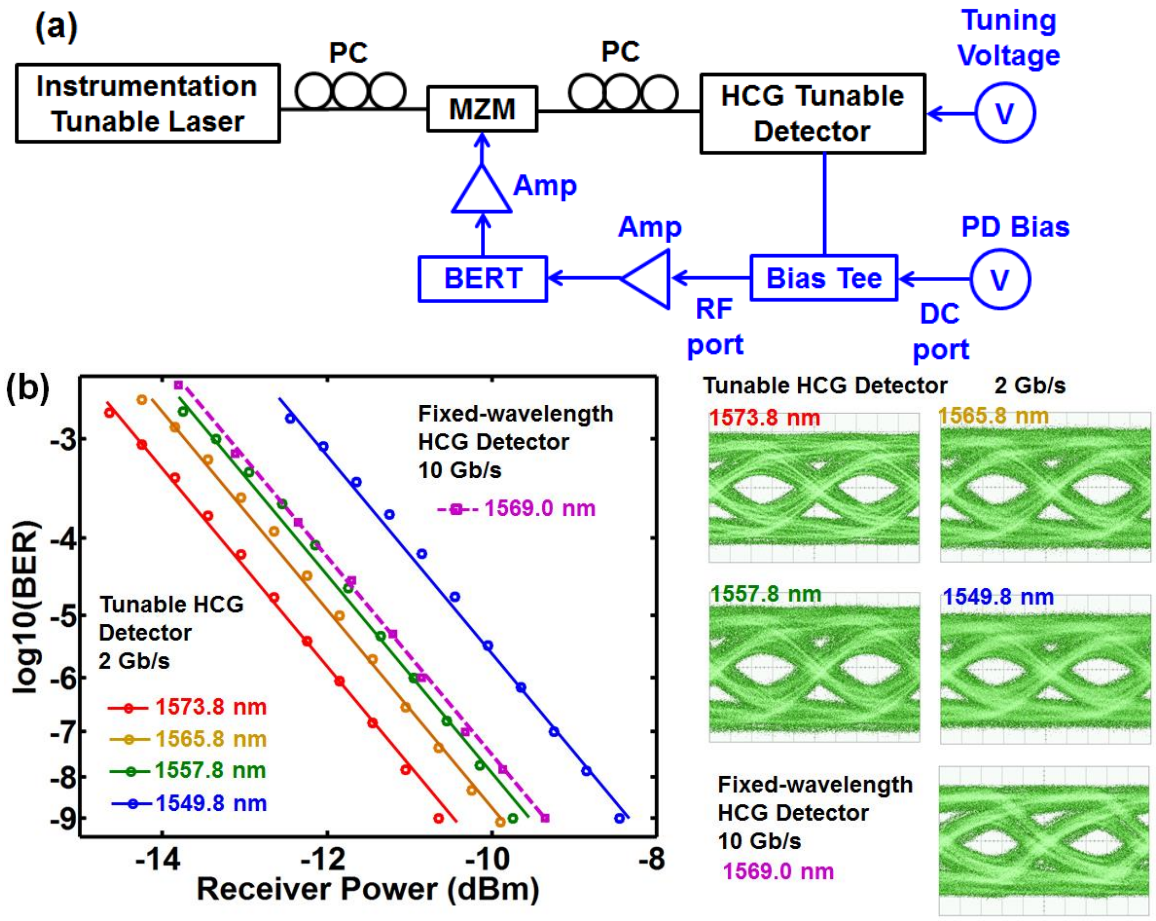


Figure 5.4 (a) Experimental setup to characterize the BER of the HCG detector. PC, polarization controller; MZM, Mach-Zehnder modulator; Amp, RF amplifier; BERT, bit error rate test bed. (b) BER vs. receiver power for a tunable detector at 2 Gb/s data rate, for four different channels, separated by 8 nm. The performance of a fixed-wavelength detector with a smaller device footprint at 10 Gb/s data rate is also shown for comparison. Eye diagrams of the receiver are shown for different conditions.

Figure 5.4 (b) shows the BER versus the receiver power for four wavelength channels over 24 nm wavelength range, for a data rate of 2 Gb/s. Eye diagram for these different wavelength channels are also shown. The detector is tuned according to the input wavelength. As the input wavelength decreases, higher reversed bias is applied to the tuning junction such that the resonance wavelength blue shifts. It is seen that as the

wavelength is blue tuned, the sensitivity decreases. This is associated with a drop in the responsivity of the detector. The best receiver sensitivity for 2 Gb/s data is -10.65 dBm. The overall sensitivity can be improved by reducing the RF cabling loss, and integrated a trans-impedance amplifier with the detector. The detection speed is still RC limited, and can be improved by further shrinking the mesa size of the device. Initial result shows that 10 Gb/s detection speed is achieved in a fixed-wavelength detector, which has a smaller mesa size. Its BER versus receiver power is also shown in Figure 5.4 (b). This promises a higher detection speed in the tunable device.

The wavelength selectivity of the tunable detector indicates that the detector can select a specific data channel while rejecting the others. To demonstrate this, two data channels at 1 Gb/s with different wavelengths are sent into the detector, and the BER for the selected channel is measured. Figure 5.5 (a) shows the experiment setup. Channel one comes from a tunable HCG VCSEL, which is directly modulated at 1 Gb/s by BERT 1. Channel two comes from an instrumentation tunable laser with external modulation (MZM) at the same data rate but from a different signal generator from BERT 2. Data from channel one transmits over 2.5 km single mode fiber (SMF), and combines with the light from channel two. The combined light, which contains two data channels, is coupled to the tunable detector. The detection wavelength of the tunable detector is tuned to be that of channel one, i.e. the data from the VCSEL. Channel two thus serves as a disturbance. The goal of this experiment is to test how well the tunable detector can receive the signal from channel one with the disturbance of an adjacent channel (i.e. channel two).

First, only channel one is turned on. The HCG tunable detector is tuned to the wavelength of this channel, and BER versus the receiver power is measured. The receiver power at BER of 1e-9 is -12.95 dBm. This serves as a reference measurement, since there is no disturbance from other channels.

Channel two is then turned on but at a different wavelength. BER is measured for channel one. At a BER of 1e-9, the power of the two channels is set to be the same. The power of channel one is then attenuated for the full BER curve measurement, while the power of channel two is fixed. Figure 5.5 (b) shows the result with various wavelength separations of the two channels. The detector completely rejects the disturbed channel (channel two) when it is at +4 nm away. When it is at -4 nm and -8 nm away, the power penalty at a BER of 1e-9 is 0 dB and 1.5 dB respectively. The slopes of the traces are different because of the fixed power of channel two. The different performance of +-4 nm channel separation is due to the asymmetric responsivity spectrum of the detector.

The HCG detector shares the same structure as the HCG VCSEL, and thus the same device can be either a transmitter or receiver. This is already shown in the experiment in Figure 5.5, where a tunable VCSEL is directly modulated as one of the data channel. To further demonstrate that the same individual device can operate both as a transmitter and receiver, we employ two identical HCG tunable VCSELs, and establish a 25 km single-mode fiber link between them. One device is forward biased, operating as a laser and being directly modulated, while the other is reversed biased, operating as a detector, and vice versa, schematically shown in Figure 5.6 (a). The tuning voltage of the laser and detector is adjusted for a best match of the lasing wavelength and detection wavelength.

Figure 5.6 (b) shows the BER versus receiver power for back-to-back (B2B) transmission and over 25 km SMF, at 1 Gb/s data rate. Figure 5.6 (c) shows the eye diagrams. Almost no power penalty is seen between B2B and 25 km SMF. It is worth to point out that at 25 km SMF transmission, error free operation can be achieved without using an isolator in the link, though at a cost of a few dB power penalty. In this case, only one fiber is required to connect the two VCSELs. This dual function device, together with its wavelength tunability, provides new flexibility and reconfigurability for optical network designers in both access networks and data center.

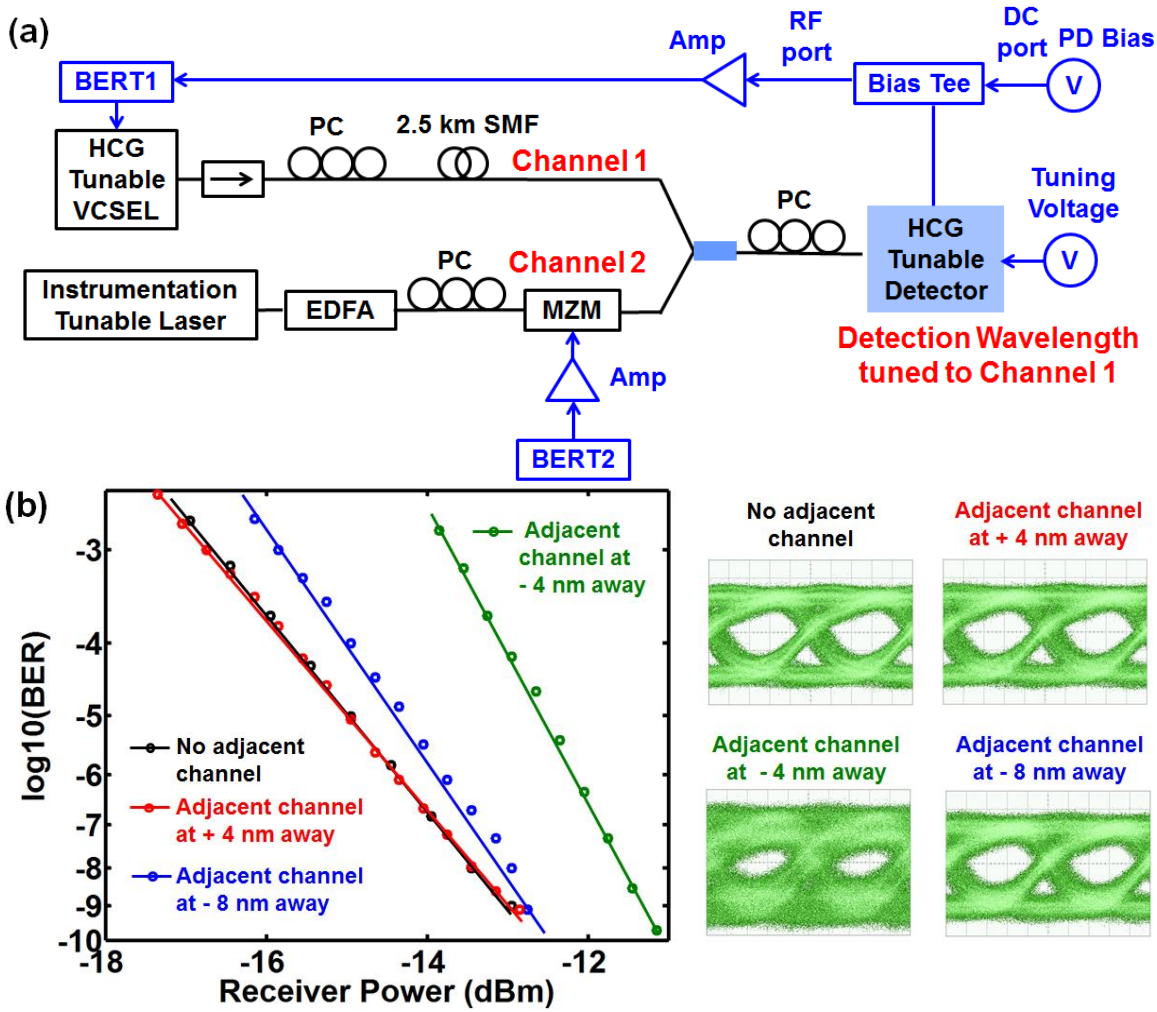


Figure 5.5 (a) System diagram where there are two data channel input to the tunable HCG detector, which is tuned to select channel 1 and reject channel 2. MZM, Mach-Zehnder modulator; PC, polarization controller; SMF, single mode fiber; Amp, RF amplifier; BERT, bit error rate test bed. (b) BER vs. receiver power for a selected channel when there is adjacent channel with a certain separation in wavelength. Eye diagrams of the receiver are shown for different conditions.

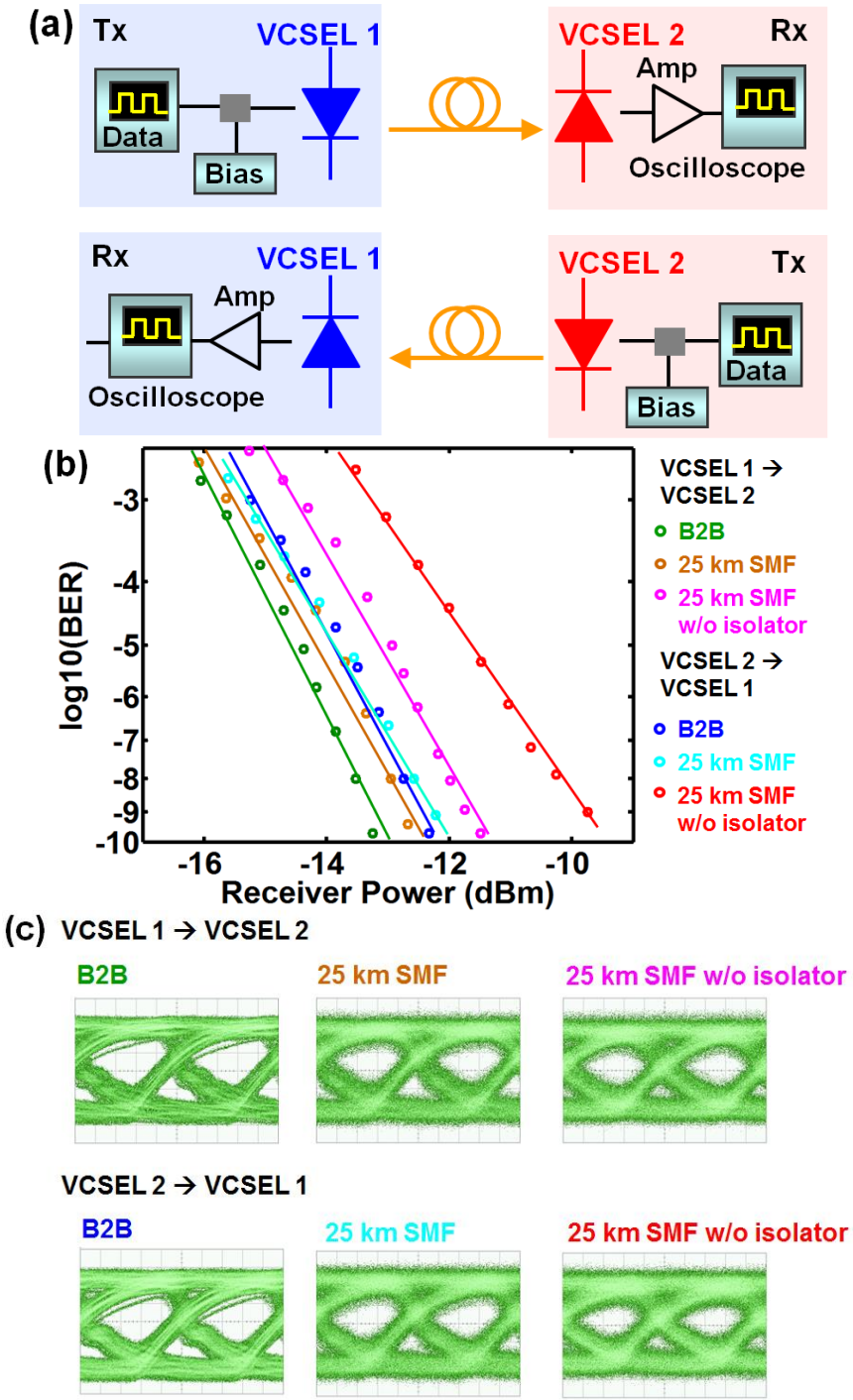


Figure 5.6 (a) Schematic of VCSEL-to-VCSEL communication. A link can be established between two identical HCG tunable VCSELs, one being forward biased and operating as a transmitter, while the other being reversed biased and operating as a receiver, and vice versa. (b) BER vs. receiver power at 1 Gb/s data rate for B2B and over 25 km SMF transmission. (c) Eye diagram for the various condition. Unless labeled, there is an isolator used in the system.

The HCG tunable VCSEL and detector will play an important role in the second generation of PON, i.e. TWDM-PON. As is mentioned earlier in this chapter, the first generation of PON uses TDM, where light of a single wavelength carries the TDM signal and broadcasts to all the end users using a power splitter. Each end user takes turn to receive the data, according to their time slot assignment. To enable a larger bandwidth, WDM comes into play in the second generation of PON. This, however, has to be compatible with the original TDM system, and does not influence the old users. In this scheme, all the old users still take turn to receive the data in their own time slot. All the new users will use a large time slot in common, but with different wavelengths. This makes sure that the original installed power splitters can stay unchanged. Channel selection and rejection is thus the key for the new users, as they will receiver different wavelength channels at the same time. Thus tunable lasers and detectors are the key elements for this new PON system. The HCG tunable detector presented here has demonstrated a good performance in channel selection and rejection. Its dual functionality further adds in flexibility in the network design. It is expected that the HCG tunable VCSEL and detector will be a main player for the transceiver markets in this new generation of PON.

5.3 Tracking Detector

In the previous section, the tuning voltage of the HCG tunable detector is set up manually to match the wavelength of the incident light. It will be interesting to make this matching automatically with some simple circuits. Since the photocurrent junction and the wavelength tuning junction shares a common contact (shown in Figure 5.1), it is possible to use the photocurrent to construct a feedback loop to adjust the tuning voltage automatically. Figure 5.7 (a) shows the circuit configuration. The tuning junction and the photodetector junction are in series connection. The voltage at the anode of the tuning junction, the common of the two junctions, and the cathode of the photodiode junction is labeled as V_T, V_F, V_D respectively. A feedback resistor R is connected between the anode of the photodetector junction and the Ground. Since the leakage current through the tuning junction is small and thus can be ignored, we have $V_F = I_{PD}R$, where I_{PD} is the photo current. The reverse voltage across the tuning junction is $V_F - V_T$. Figure 5.7 (b) shows the operation principle of the tracking detector. The three color traces indicates the photocurrent versus the reverse tuning voltage for three different input wavelengths. The black curve, with a slope of $1/R$, shows the load line from the resistor. For each input light wavelength, the photocurrent response trace and the load line crosses twice. Only one of these crossing points – the one where the two traces have opposite slopes – is stable. This is labeled by the dots in Figure 5.7 (b). As the input wavelength moves, this stable crossing point follows the load line. This is thus the tracking mechanism.

Figure 5.8 shows the measured photocurrent versus the tuning voltage for various input wavelengths. The input light power is - 11 dBm. The voltage across the photodiode junction is 2 V. For each input light wavelength, the trace has a longer tail on the small voltage side. This comes from the relationship between the tuning voltage and HCG displacement (and thus resonance wavelength), denoted by Eq. 4.5. If $\Delta\lambda_0$ is the change

of resonance wavelength due to the change of tuning voltage ΔV, then for the same ΔV, $\Delta\lambda_0$ will have a larger value at a larger tuning voltage. This explains why the trace has a longer tail on the small voltage side. A resistor can then be designed for the feedback.

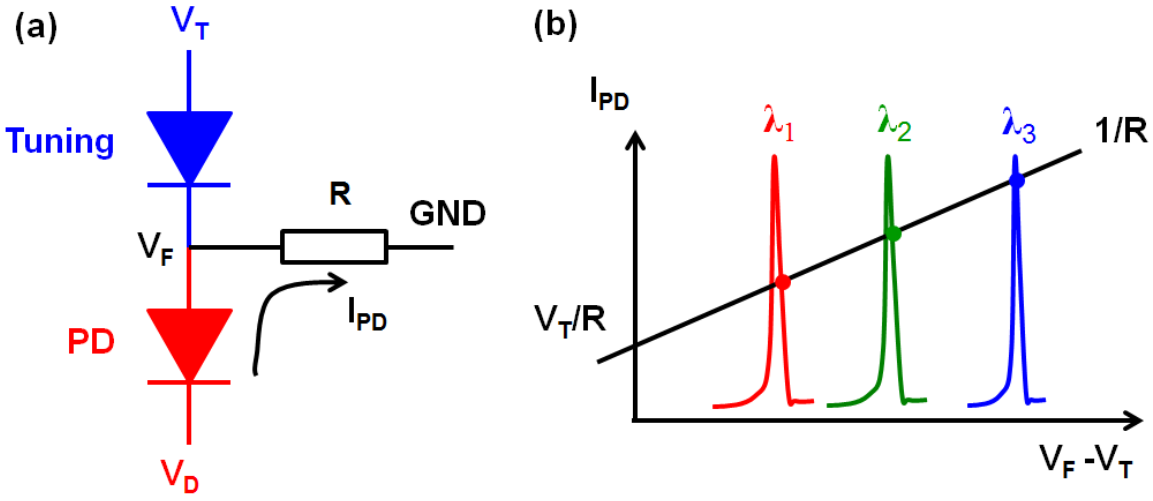


Figure 5.7 (a) Schematic of the circuit configuration of the tracking detector. A feedback resistor is connected to the common of the two junctions and ground (GND). (b) Operation principle of the tracking detector. The crossing point indicated by the dot is the stable operation point. As the input wavelength moves, this crossing point follows the load line, thus tracks the wavelength.

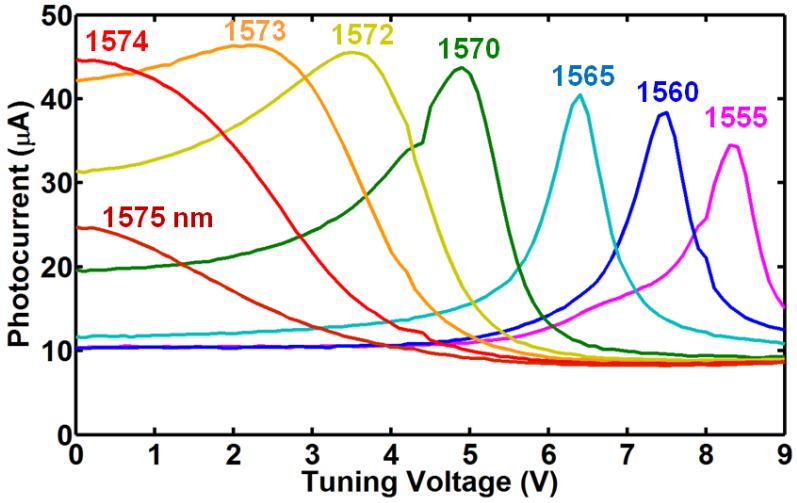


Figure 5.8 Photocurrent versus tuning voltage for different incident light wavelengths.

To test the performance of the tracking detector, a feedback resistor is chosen, and BER is measured for different incident light's wavelength, with a fixed incident light power of -11 dBm. This is compared with the non-tracking case. Figure 5.9 shows the results. It is shown that there is a larger wavelength range for a low BER with the

tracking scheme compared to non-tracking. With an optimization of the V_T, V_D and R, an even better performance is expected.

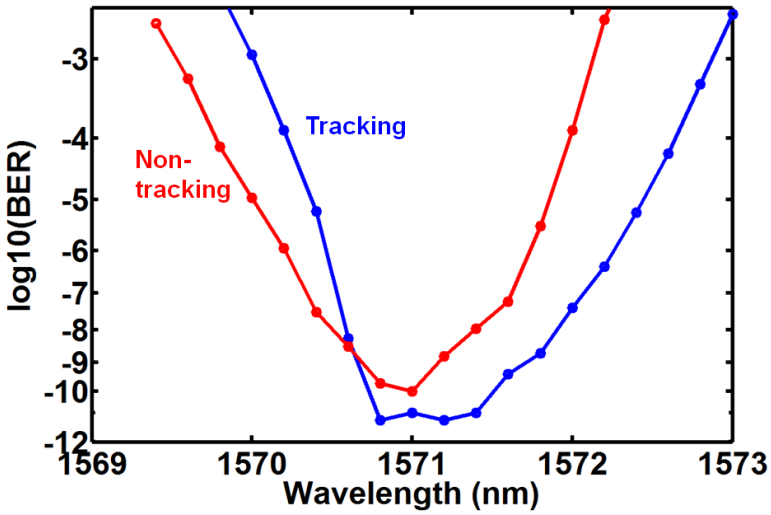


Figure 5.9 BER versus incident light wavelength for the tracking case and non-tracking case. The power of the incident light is fixed at -11 dBm.

5.4 Chip-scale Optical spectrometer

While the previous section focus on the application of the HCG tunable detector in optical communication network, there are other applications. One of them is using the tunable detector as an on-chip spectrometer, which is the topic of this section.

Chip-scale optical spectrometers have been drawing more and more research interests for its increasing applications in various lab-on-a-chip systems, e.g. medical diagnostics, biochemical sensing, and environmental monitoring, etc. It is also desirable in optical networks as a compact wavemeter. Significant progress has been made and chip-scale spectrometers have been demonstrated with various configurations, such as superprism or arrayed waveguide grating [92, 93], Fabry-Perot filter array [94], micro-resonator array [95], etc. However, there is a fundamental tradeoff between the resolution and footprint on these dispersive components or resonance array. Tunable Fabry-Perot filter, on the other hand, turns the spectrum splitting from the spatial domain into temporal domain. It requires only one photodetector, and can achieve high resolution and small footprint at the same time [96]. To further reduce the footprint of the whole system, monolithic integration of the tunable Fabry-Perot filter with the photodetector into one single device is necessary. This is exactly what HCG tunable detector is. We term it as the swept-wavelength detector.

To operate the swept-wavelength detector as a spectrometer, a periodic tuning voltage $V(t)$ is applied to the detector, and the photocurrent $I(t)$ is recorded versus time. As the tuning voltage is being swept, the detector experiences different resonance wavelengths $\lambda(t)$, thus recording the corresponding spectrum component of the input

light. The schematic of the setup is shown in Figure 5.10 (a). To convert the time domain photocurrent $I(t)$ into optical spectrum $P(\lambda)$ as shown in Figure 5.10 (b), one needs to correlate the resonance wavelength with time (so as to convert t to λ), and calibrate the responsivity for each resonance wavelength (so as to convert I to P). Note that we cannot use the DC characteristic shown in Figure 5.3 for this purpose, as the AC tuning characteristics of the MEMS HCG is different from that measured in DC. To do this conversion, we can first apply a periodic tuning voltage $V(t)$ to the HCG tunable detector, e.g. a sinusoidal waveform. Then we can employ a widely tunable laser as the input light source, and set the input wavelength of the tunable laser as λ_i with an optical power of P_i, and record the photocurrent $I(t)$, in one tuning period of time. There should be a specific time t_i, such that the photocurrent $I(t)$ is maximum. This maximum photocurrent is labeled as I_i. The tuning voltage at this specific time t_i is recorded as V_i. Thus we can correlate $\lambda_i \sim t_i \sim V_i \sim I_i \sim P_i$. Then we can repeat the above procedure for all the other wavelengths within the tuning range of the HCG tunable detector under the tuning voltage $V(t)$. With these steps, we can obtain the resonance wavelength versus time $\lambda(t)$, as well as the responsivity versus resonance wavelength $R(\lambda)$.

When this swept-wavelength detector is used to measure the spectrum of an input light with an arbitrary spectrum $P(\lambda)$, the recorded photocurrent $I(t)$ is first converted to $I(\lambda)$ with $\lambda(t)$, and then converted to $P(\lambda)$ with $R(\lambda)$. The functionality of the optical spectrometer is thus realized.

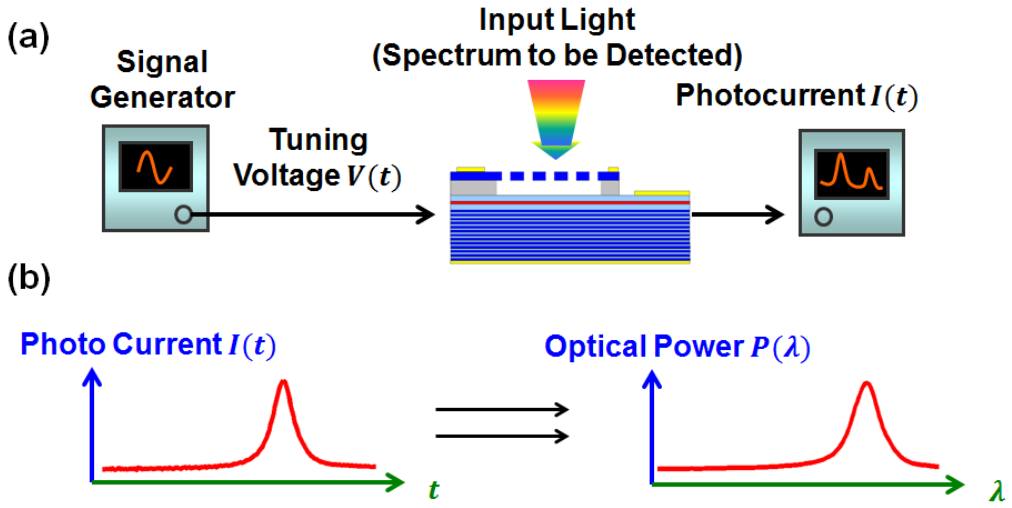


Figure 5.10 (a) Configuration of the optical spectrometer using the HCG swept-wavelength detector. A tuning voltage $V(t)$ is applied to the HCG swept-wavelength detector, and the photocurrent $I(t)$ is recorded versus time. $I(t)$ is subsequently converted to $P(\lambda)$, shown in (b).

As an example, a 1 kHz sinusoidal waveform $V(t)$ is used as a tuning voltage. Figure 5.11 (a) shows the raw data $I(t)$ for five different input light wavelength from a tunable laser. Within one sinusoidal sweeping cycle, the detector sweeps through the same resonance wavelength twice. Thus the light can be detected twice if its wavelength falls into the sweeping range. Figure 5.11 (b) zooms in the data in half of the sweeping period. By sweeping through the wavelength across the whole tuning range, and

recording the data, one can calibrate the relationship between the tuning voltage, responsivity and the resonance wavelength. The horizontal time axis and the vertical photocurrent axis in Figure 5.11 (b) can be converted to wavelength and optical power respectively, shown in Figure 5.11 (c).

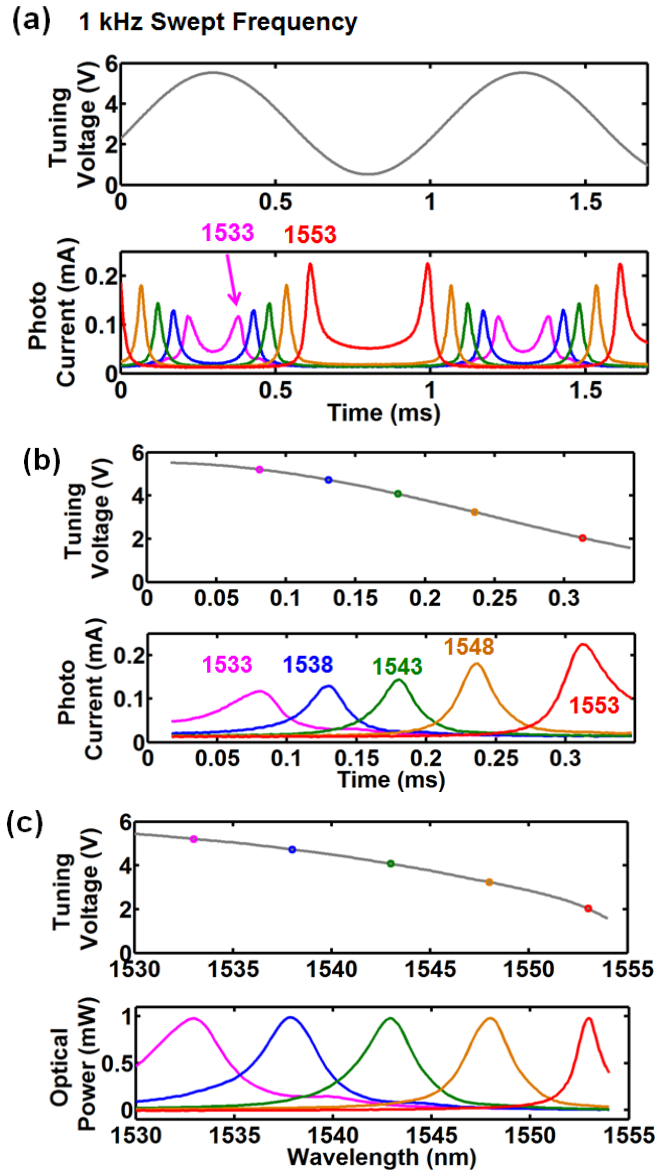


Figure 5.11 (a) Tuning voltage and recorded current versus time for five different single wavelength inputs. The tuning voltage is a 1 kHz sinusoidal waveform. (b) Zoomed-in view of (a), at half of the sweeping cycle. (c) The converted optical spectrum for these five different single wavelength inputs.

Figure 5.12 shows the result using a 100 kHz sinusoidal waveform as the tuning voltage. A delay is seen between the tuning voltage and the HCG displacement (indicated by the resonance frequency). This is because of this frequency (100 kHz) is closed to the mechanical resonance frequency of the MEMS HCG.

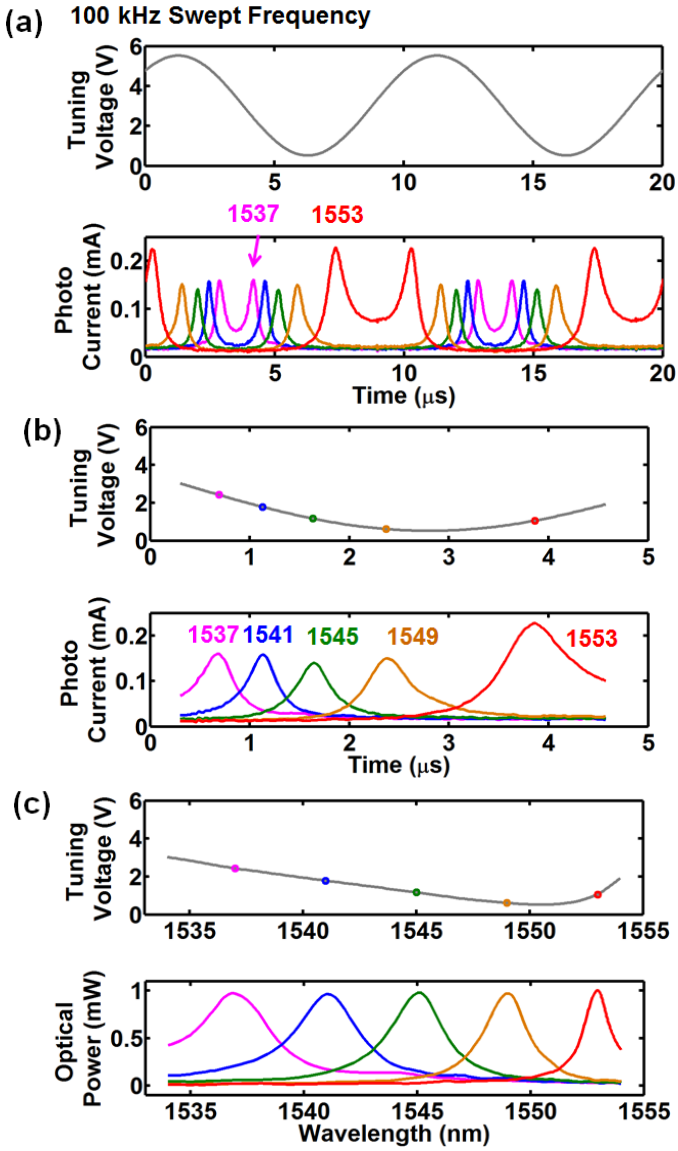


Figure 5.12 (a) Tuning voltage and recorded current versus time for five different single wavelength inputs. The tuning voltage is a 1 kHz sinusoidal waveform. (b) Zoomed-in view of (a), at half of the sweeping cycle. (c) The converted optical spectrum for these five different single wavelength inputs.

Figure 5.13 shows the spectrum recorded by the HCG swept-wavelength detector where the input light contains two wavelength components, 1544nm and 1552nm. The waveform of the tuning voltage is 100 kHz sinusoidal wave. Both the two wavelengths and their optical powers are resolved correctly.

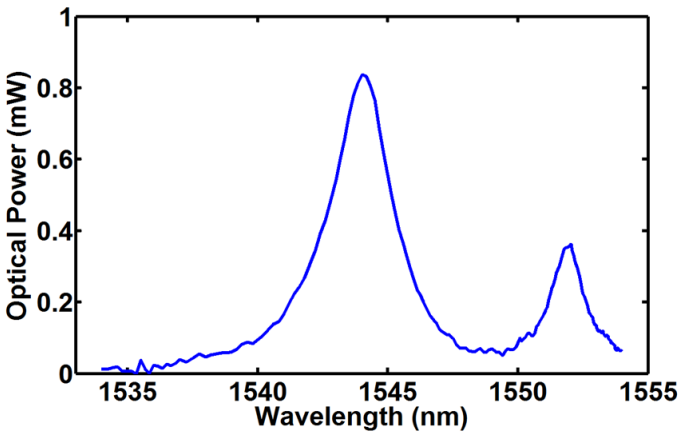


Figure 5.13 Example of the output spectrum of the spectrometer using the HCG swept-wavelength detector. The input light contains two wavelength components. Both the two wavelengths and their optical powers are resolved correctly.

The data in Figure 5.12 and Figure 5.13 can be acquired within half of the period of the sinusoidal waveform, and thus the operation speed of this spectrometer is essentially 200 kHz. This high operation speed of the swept-wavelength detector benefits from the light weight of the HCG, whose mechanical resonance frequency can be very high. Bandwidth in excess of 1 MHz can be achieved by optimization of the HCG size and the MEMS structure. A 27 MHz 3 dB mechanical tuning bandwidth has been previously demonstrated in an HCG tunable VCSEL [22].

To increase the resolution, the FWHM of the responsivity spectrum needs to be small across the whole tuning range. The presented HCG detector has a minimum responsivity spectrum width of 1.2 nm; however there is large fluctuation of this value across the 33.5 nm tuning range. It is expected this can be solved with an optimization of the HCG MEMS. The spectrum width can also be reduced with an optimization of the optical coupling scheme. Besides resolution, the total tuning range is another figure of merit. By operating the HCG MEMS at its mechanical resonance frequency, the tuning range can be essentially doubled, as the HCG can not only be pulled towards the substrate, but also pushed away from the substrate. This thus substantially increases the detection bandwidth.

5.5 Summary

In this chapter, a tunable 1550 nm HCG VCSEL is experimentally demonstrated as a photodetector. A large responsivity of 1A/W, a wide tuning range of 33.5 nm, a narrow spectrum width of 1.2 nm, and a high detection speed of 10 Gb/s are experimentally demonstrated. This device has an excellent wavelength channel selection and rejection, and will play an important role in the TWDM-PON. This single device can operate both in laser mode and detector mode, by simply switching the polarity of the bias. An error-free 1 Gb/s link is demonstrated between two such devices. This novel HCG VCSEL-detector fills in the void of needs for greatly simplified device and fiber packaging by

using a single device both as a laser and photodetector. It provides new elements and extra flexibility for the configuration of PON and data center. The tunability of the device would be useful for any wavelength-division-multiplexing network. Furthermore, the HCG VCSEL-detectors can be fabricated into array, facilitating the usage of fiber ribbon or multi-core fibers.

Tracking detection and swept-wavelength detection for chip-scale optical spectrometer are two novel functions of this HCG tunable detector. The functionality of an optical spectrometer will be important for various lab-on-a-chip systems.

The on-going research in this topic includes pushing the device to a higher detection speed, stronger wavelength selectivity and higher rejection for off-resonance wavelength channels. It is expected that the tunable HCG VCSEL and detector will be an important player in the transceiver markets for the new generation of PON system and the data centers.

Chapter 6
Summary and Outlook

In this dissertation, the rich physics of HCG is studied, and various applications in integrated optoelectronics are explored. The theoretical analysis of HCG provides a design guidance for applications; in return, the applications motivate the further development of the HCG physics.

The HCG theoretical analysis discussed in this dissertation provides an intuitive physical picture of HCG and a simple guideline for HCG designs. It explains various extraordinary properties of HCG, such as broadband high reflection, broadband high transmission, and the high Q resonance. The analysis is straight-forward; it decomposes the HCG into the input plane, exit plane and a set of waveguide array in between. Waveguide array modes are used as the mode basis to expand the electromagnetic field inside the HCG, which represents the key difference of this analysis from the other approaches. Reflection matrix for the waveguide array modes is then developed for the two boundaries. The waveguide array modes interfere at the input and exit plane, giving rise to the various properties. The thickness of HCG is thus one of the most important design parameters since it determines the interference effect. This is a unique feature of HCG, in sharp contrast to diffraction gratings, guided mode resonances, Wood's anomaly and photonic crystals.

This analytical treatment of the HCG is very similar to that of the FP cavity, however in a matrix format, and using the waveguide array modes as the mode basis. This also prompts the general transmission matrix method for the more complicated structure of HCGs. Band diagram is developed for HCG, which clearly outlines the different operation region between HCG and photonic crystal.

Novel applications of HCG in both oblique incident angle and surface normal angle are discussed. HCG hollow-core waveguide explores the high reflectivity and reflection phase properties of the HCG at oblique incident angle. A hollow-core waveguide is designed and demonstrated to have two-dimensional optical confinement, but with only one-dimensional geometric confinement. This seemingly impossible physics is enabled by the reflection phase variation of the chirped HCG design, which brings in the optical confinement in the second dimension. Record low propagation loss is demonstrated for this hollow-core waveguide. Other types of HCG hollow-core waveguides are explored. The HCG hollow-core waveguide itself has built up a new platform for integrated optics with novel functionalities, such as slow-light, gas sensing, and optical switch etc.

High speed optical phased array and tunable detector both uses the HCG as an electrostatically actuable top mirror of a FP cavity. Due to its light weight and ultra-small thickness, the MEMS can have a very high operation speed. All-pass filter is used to construct the phase tuner, which shows a very high phase tuning efficiency. Fast optical beam steering (0.53 MHz) is demonstrated with an 8x8 optical phased array. This fast beam steering or beam forming is particularly promising for LIDAR, optical switch, and fast biological imaging.

The strong wavelength selectivity, and high speed detection is the hidden capability of the tunable HCG VCSEL. These hidden treasures are explored here. The tunable HCG detector is shown to be able to select a desired wavelength channel and reject the others. This is extremely suitable for TWDM-PON system, and has a very large market once commercialized. More interestingly, the single device can exhibit high performance in both laser mode and detector mode; this opens up a great flexibility and opportunity for the network designers. Other application innovation with the HCG detector includes tracking detection and chip-scale optical spectrometer.

While this dissertation has discussed the various novel applications of HCG, there are a lot others await for more explorations and developments. The phase engineering of HCG is one of them. By spatially chirping the HCG dimension, the HCG with the same thickness can exhibit different transmission phase and reflection phase. HCG focusing reflectors and lenses have already been designed and fabricated [38-40]. With the same principle, other waveform patterns such as vortex beams, airy beams, beams with optical angular momentum can be generated; this has applications in imaging, and free space communication etc. They can also be designed for beam routing; this can provide a platform for compact photonic network, either in free-space, or stacking layers [41]. Besides this, the chromatic dispersion of the spatially chirped HCG is also worth to study. Dichromatic mirror, wavelength splitting may be able to realize with the chirped HCG.

Another interesting topic is the further experimental realization of the ultra-high Q HCG resonator. This new type of HCG resonator facilitates the surface normal coupling, and will find its application in laser [36], surface-enhanced Raman spectroscopy [35] and gas sensing etc.

Light source in integrated photonic circuits is an important element. Integration of laser onto silicon photonics or flexible substrates is a hot research topic. The HCG tunable VCSEL facilitates such an integration. The III-V VCSEL can be flip-chip bonded to the photonics layer, and a vertical-to-in-plane coupler [52] can be used to couple light from the VCSEL to the in-plane waveguides on the photonic network. Alternatively, the HCG itself can be fabricated on the silicon photonics, and the other half of the VCSEL can be bonded onto the HCG. This seamless integration scheme between III-V and silicon is expected to thrive in future photonic device and systems.

In the theoretical development of HCG, the reflection and transmission phase of HCG is not yet well modeled, and the properties are not yet intuitively explained. More work is required on this topic.

So far, only one dimensional grating is analyzed. Using the same principle, it is possible to develop the analysis for two dimensional gratings. This will for sure bring new properties and opportunity for applications, such as polarization independent beam forming, etc.

Since its invention in 2004, HCG has seen rapid progress both in theory and experiments. More and more research groups around the world pay their attention and conduct research on this field. A conference named “high contrast metastructures” has been devoted to this field since 2012 in SPIE Photonic West. HCG has not only been a

new element in integrated optoelectronics, but also provide a new path for device innovation and formed a new platform for the further development of integrated optoelectronics. It is expected HCG will continue to play an important role in integrated optoelectronics, and enable more and more devices and systems with novel functionality.

Bibliography

[1] E. G. Loewen and E. Popov, Diffraction Gratings and Applications, CRC Press, 1997.

[2] M. Born and E. Wolf, Principles of Optics, 7th (expanded) ed., Cambridge University Press, 1999.

[3] B. C. Kress and P. Meyrueis, Applied Digital Optics: from Micro-optics to Nanophotonics, Wiley, 2009.

[4] "High Contrast Metastructures," *edited by C. J. Chang-Hasnain, F. Koyama, A. E. Willner and W. Zhou, Proceedings of SPIE, 0277-786X, v. 8270,* 2012.

[5] "High Contrast Metastructures," *edited by C. J. Chang-Hasnain, F. Koyama, A. E. Willner and W. Zhou, Proceedings of SPIE, 0277-786X, v. 8633,* 2013.

[6] C. F. R. Mateus, M. C. Y. Huang, Y. Deng, A. R. Neureuther, and C. J. Chang-Hasnain, "Ultra-broadband mirror using low index cladded subwavelength grating," *IEEE Photonics Technology Letters,* vol. 16, pp. 518-520, 2004.

[7] C. F. R. Mateus, M. C. Y. Huang, L. Chen, and C. J. Chang-Hasnain and Y. Suzuki, "Broadband mirror (1.12-1.62 um) using single-layer sub-wavelength grating," *IEEE Photonics Technology Letters,* vol. 16, pp. 1676-1678, 2004.

[8] C. F. M. a. M. C. H. C. J. Chang-Hasnain, "Ultra broadband mirror using subwavelength grating". US Patent 7,304,781, 2007.

[9] C. J. Chang-Hasnain and W. Yang, "High-contrast gratings for integrated optoelectronics," *Advances in Optics and Photonics,* vol. 4, pp. 379-440, 2012.

[10] S. Astilean, P. Lalanne, P. Chavel, E. Cambril, and H. Launois, "High-efficiency subwavelength diffractive element patterned in a high-refractive-index material for 633 nm," *Optics Letters,* vol. 23, pp. 552-554, 1998.

[11] S. Goeman, S. Boons, B. Dhoedt, K. Vandeputte, K. Caekebeke, P. Van Daele, and R. Baets, "First demonstration of highly reflective and highly polarization selective diffraction gratings (GIRO-gratings) for long-wavelength VCSELs," *IEEE Photonics Technology Letters,* vol. 10, pp. 1205-1207, 1998.

[12] T. Glaser, S. Schroter, H. Bartelt, H.-J. Fuchs, and E.-B. Kley, "Diffractive optical isolator made of high-efficiency dielectric gratings only," *Applied Optics,* vol. 41, pp. 3558-3566, 2002.

[13] D. Rosenblatt, A. Sharon, and A. A. Friesem, "Resonant grating waveguide

structures," *Quantum Electronics, IEEE Journal of,* vol. 33, p. 2038–2059, 1997.

[14] R. Magnusson and S. S. Wang, "New principle for optical filters," *Applied Physics Letters,* vol. 61, pp. 1022-1024, 1992.

[15] A. Haglund, S. J. Gustavsson, J. Vukusic, P. Jedrasik, and A. Larsson, "High-power fundamental-mode and polarisation stabilised VCSELs using sub-wavelength surface grating," *Electronics Letters,* vol. 41, pp. 805-807, 2005.

[16] L. Zhuang, S. Schablitsky, R. C. Shi, and S. Y. Chou, "Fabrication and performance of thin amorphous Si subwavelength transmission grating for controlling vertical cavity surface emitting laser polarization," *Journal of Vacuum Science & Technology B,* vol. 14, pp. 4055-4057, 1996.

[17] M. C. Y. Huang, Y. Zhou, and C. J. Chang-Hasnain, "A surface-emitting laser incorporating a high index-contrast subwavelength grating," *Nature Photonics,* vol. 1, pp. 119-122, 2007.

[18] S. Boutami, B. Ben Bakir, J.-L. Leclercq, and P. Viktorovitch, "Compact and polarization controlled 1.55 μm vertical-cavity surface emitting laser using single-layer photonic crystal mirror," *Applied Physics Letters,* vol. 91, p. 071105, 2007.

[19] Y. Zhou, M. C. Y. Huang, and C. J. Chang-Hasnain, "Large fabrication tolerance for VCSELs using high contrast grating," *IEEE Photonics Technology Letters,* vol. 20, pp. 434-436, 2008.

[20] M. C. Y. Huang, Y. Zhou, and C. J. Chang-Hasnain, "A nanoelectromechanical tunable laser," *Nature Photonics,* vol. 2, pp. 180-184, 2008.

[21] I.-S. Chung, J. Mørk, P. Gilet, and A. Chelnokov, "Subwavelength grating-mirror VCSEL with a thin oxide gap," *IEEE Photonics Technology Letters,* vol. 20, p. 105–107, 2008.

[22] C. Chase, Y. Zhou, C.J. Chang-Hasnain, "Size effect of high contrast gratings in VCSELs," *Optics Express,* vol. 17, p. 24002 – 24007, 2009.

[23] C. J. Chang-Hasnain, Y. Zhou, M. C. Y. Huang, and C. Chase, "High-contrast grating VCSELs," *Selected Topics in Quantum Electronics, IEEE Journal of,* vol. 15, pp. 869-878, 2009.

[24] C. Chase, Y. Rao, W. Hofmann, and C. J. Chang-Hasnain, "1550 nm high contrast grating VCSEL," *Optics Express,* vol. 18, pp. 15461-15466, 2010.

[25] W. Hofmann, C. Chase, M. Müller, Y. Rao, C. Grasse, G. Böhm, M.-C. Amann, and Connie Chang-Hasnain, "Long-Wavelength High-Contrast Grating Vertical-Cavity Surface-Emitting Laser," *IEEE Photonics Journal,* vol. 2, pp. 415-422, 2010.

[26] P. Gilet, N. Olivier, P. Grosse, K. Gilbert, A. Chelnokov, I.-S. Chung, and J. Mørk, "High-index-contrast subwavelength grating VCSEL," in *Proceedings of SPIE Vol. 7615, 76150J, Vertical-Cavity Surface-Emitting Lasers XIV*, 2010.

[27] Y. Rao, C. Chase, C. J. Chang-Hasnain, "Multiwavelength HCG-VCSEL Array," in *International Semiconductor Laser Conference*, Kyoto, Japan, 2010.

[28] T. Ansbaek, I.-S. Chung, E. S. Semenova, and K. Yvind, "1060-nm tunable monolithic high index contrast subwavelength grating VCSEL," *IEEE Photonics Technology Letters*, vol. 24, pp. 455-457, 2013.

[29] Y. Rao, W. Yang, C. Chase, M. C. Y. Huang, D. P. Worland, S. Khaleghi, M. R. Chitgarha, M. Ziyadi, A. E. Willner, and C. J. Chang-Hasnain, "Long-Wavelength VCSEL Using High Contrast Grating," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 19, p. 1701311, 2013.

[30] C. Sciancalepore, B.B. Bakir, X. Letartre, J. Harduin, N. Olivier, C. Seassal, J. Fedeli, P. Viktorovitch, "CMOS-compatible ultra-compact 1.55- μ m emitting VCSELs using double photonic crystal mirrors," *IEEE Photonics Technology Letters*, vol. 24, pp. 455-457, 2012.

[31] W. Yang, L. Zhu, Y. Rao, C. Chase, M. Huang, and C. J. Chang-Hasnain, "Tunable 1550-nm high contrast grating VCSEL detector," in *Optical Communication (ECOC 2013), 39th European Conference and Exhibition on* , London, UK, 2013.

[32] W. Yang, T. Sun, Y. Rao, M. Megens, T. Chan, B.-W. Yoo, D. A. Horsley, M. C. Wu, and C. J. Chang-Hasnain, "High-speed optical phased array using two-dimensional high-contrast grating all-pass filters," in *Proceedings of SPIE 8633, High Contrast Metastructures II,*, San Francisco, 2013.

[33] T. Stöferle, N. Moll, T. Wahlbrink, J. Bolten, T. Mollenhauer, U. Scherf and R. F. Mahrt, "Ultracompact Silicon/polymer laser with an absorption-insensitive nanophotonic resonator," *Nano Letters*, vol. 10 , pp. 3675-3678, 2010.

[34] Y. Zhou, M. Moewe, J. Kern, M. C. Y. Huang, and C. J. Chang-Hasnain, "Surface-normal emission of a high-Q resonator using a subwavelength high-contrast grating," *Optics Express*, vol. 16, pp. 17282-17381, 2008.

[35] V. Karagodsky, T. Tran, M. Wu, and C. J. Chang-Hasnain, "Double-Resonant Enhancement of Surface Enhanced Raman Scattering Using High Contrast Grating Resonators," in *CLEO:2011 - Laser Applications to Photonic Applications, OSA Technical Digest*, 2011.

[36] T.-T. Wu, S.-H. Wu, T.-C. Lu, and S.-C. Wang, "GaN-based high contrast grating surface-emitting lasers," *Applied Physics Letters*, vol. 102, p. 081111, 2013.

[37] V. Karagodsky, B. Pesala, C. Chase, W. Hofmann, F. Koyama and C.J. Chang-Hasnain, "Monolithically integrated multi-wavelength VCSEL arrays using high-contrast gratings," *Optics Express,* vol. 18, pp. 694-699, 2010.

[38] F. Lu, F.G. Sedgwick, V. Karagodsky, C. Chase and C.J. Chang-Hasnain, "Planar high-numerical-aperture low-loss focusing reflectors and lenses using subwavelength high contrast gratings," *Optics Express,* vol. 18, pp. 12606-12614, 2010.

[39] D. Fattal, J. Li, Z. Peng, M. Fiorentino, and R. G. Beausolei, "Flat dielectric grating reflectors with focusing abilities," *Nature Photonics,* vol. 4, pp. 466-469, 2010.

[40] A. B. Klemm, D. Stellinga, E. R. Martins, L. Lewis, G. Huyet, L. O'Faolain, and T. F. Krauss, "Experimental high numerical aperture focusing with high contrast grating," *Optics Letters,* vol. 38, pp. 3140-3143, 2003.

[41] D. Fattal, J. Li, Z. Peng, M. Fiorentino, and R. G. Beausoleil, "A silicon lens for integrated free-space optics," in *Integrated Photonics Research, Silicon and Nanophotonics, OSA Technical Digest*, 2011.

[42] F. Koyama, "High contrast grating for spatial mode filtering and mode control of VCSELs," in *Proceedings of SPIE Vol. 8270, 827005, High Contrast Metastructure*, San Francisco, 2012.

[43] F. Brückner, D. Friedrich, T. Clausnitzer, M. Britzger, O. Burmeister, K. Danzmann, E. B. Kley, A. Tünnermann and R. Schnabel, "Realization of a monolithic high-reflectivity cavity mirror from a single silicon crystal," *Physical Review Letters,* vol. 104, p. 163903, 2010.

[44] S. Kroker, S. Steiner, T. Käsebier, E-B. Kley, and A. Tünnermann, "High contrast gratings for high-precision metrology," in *Proceedings of SPIE Vol. 8633, 86330M, High Contrast Metastructure II*, San Francisco, 2013.

[45] Y. Zhou, V. Karagodsky, B. Pesala, F.G. Sedgwick and C.J. Chang-Hasnain, "A novel ultra-low loss hollow-core waveguide using subwavelength high-contrast gratings," *Optics Express*, vol. 17, pp. 1508-1517, 2009.

[46] W. Yang, J. Ferrara, K. Grutter, A. Yeh, C. Chase, Y. Yue, A. E. Willner, M. Wu, and C. J. Chang-Hasnain, "Low loss hollow-core waveguide on a silicon substrate," *Nanophotonics,* vol. 1, pp. 23-29, 2012.

[47] W. Zhou, G. Dang, M. Taysing-Lara, V. Karagodsky, T. Sun and C. Chang-Hasnain, "Slow-light high contrast metastructure hollow-core waveguides," in *Proceedings of SPIE 8270, High Contrast Metastructures, 827009*, San Francisco, 2012.

[48] T. Sun, W. Yang, V. Karagodsky, W. Zhou, and C. Chang-Hasnain, "Low-loss slow light inside high contrast grating waveguide," in *Proceedings of SPIE 8270, High*

Contrast Metastructures, 82700A, San Francisco, 2012.

[49] B. Pesala, V. Karagodsky and C.J. Chang-Hasnain, "Ultra-compact optical coupler and splitter using high-contrast grating hollow-core waveguide," in *Integrated Photonics Research, Silicon and Nanophotonics, OSA Technical Digest*, 2010.

[50] W. Yang and C. J. Chang-Hasnain, "Ultra-compact optical switch using high contrast grating hollow-core waveguide," in *Conference on Lasers and Electro-Optics, OSA Technical Digest*, San Jose, USA, 2013.

[51] C. Sieutat, R. Peretti, J.-L. Leclercq, P. Viktorovitch, and X. Letartre, "Strong confinement of light in low index materials: the photon cage," *Optics Express,* vol. 21, pp. 20015-20022, 2013.

[52] L. Zhu, V. Karagodsky, and C. Chang-Hasnain, "Novel high efficiency vertical to in-plane optical coupler," in *Proceedings of SPIE 8270, High Contrast Metastructures, 82700L,* San Francisco, USA, 2012.

[53] V. Karagodsky, F.G. Sedgwick and C.J. Chang-Hasnain, "Theoretical analysis of subwavelength high contrast grating reflectors," *Optics Express,* vol. 18, pp. 16973-16988, 2010.

[54] V. Karagodsky, and C. J. Chang-Hasnain, "Physics of near-wavelength high contrast gratings," *Optics Express,* vol. 20, pp. 10888-10895, 2012.

[55] M. G. Moharam and T. K. Gaylord, "Rigorous coupled wave analysis of planar grating diffraction," *Journal of the Optical Society of America,* Vols. 71, 811, pp. 811-818, 1981.

[56] S. T. Peng, "Rigorous formulation of scattering and guidance by dielectric grating waveguides: general case of oblique incidence," *Journal of the Optical Society of America A,* vol. 6, pp. 1869-1883, 1989.

[57] L. Li, "A modal analysis of lamellar diffraction gratings in conical mountings," *Journal of Modern Optics,* vol. 40, pp. 553-573, 1993.

[58] V. Karagodsky, C. Chase, and C. J. Chang-Hasnain, "Matrix Fabry–Perot resonance mechanism in high-contrast gratings," *Optics Letters,* vol. 36, pp. 1704-1706, 2011.

[59] P. Lalanne, J.P. Hugonin and P. Chavel, "Optical properties of deep lamellar gratings: a coupled Bloch-mode insight," *Journal of Lightwave Technology,* vol. 24, pp. 2442-2449, 2006.

[60] P. C. Magnusson, G. C. Alexander, V. K. Tripathi, and A. Weisshaar, Transmission Lines and Wave Propagation, 4th ed., CRC Press, 2001.

[61] C. Weisbuch, M. Nishioka, A. Ishikawa and Y. Arakawa, "Observation of the coupled exciton-photon mode splitting in a semiconductor quantum microcavity," *Physcal Review Letters ,* vol. 69, pp. 3314-3317, 1992.

[62] J. P. Reithmaier, G. Sęk, A. Löffler, C. Hofmann, S. Kuhn, S. Reitzenstein, L. V. Keldysh, V. D. Kulakovskii, T. L. Reinecke and A. Forchel, "Strong coupling in a single quantum dot–semiconductor microcavity system," *Nature,* vol. 432, pp. 197-200, 2004.

[63] W. Shan, W. Walukiewicz, J. W. Ager III, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz, "Band anticrossing in GaInNAs alloys," *Physical Review Letters,* vol. 82, pp. 1221-1224, 1999.

[64] John D. Joannopoulos, Steven G. Johnson, Joshua N. Winn, and Robert D. Meade, Photonic Crystals: Molding the Flow of Light, 2nd edition, Princeton University Press, 2008.

[65] W. M. J. Green, M. J. Rooks, L. Sekaric and Y. A. Vlasov, "Optical modulation using anti-crossing between paired amplitude and phase resonators," *Optics Express,* vol. 15, pp. 17264-17272, 2007.

[66] Y. Saito, T. Kanaya, A. Nomura, and T. Kano, "Experimental trial of a hollow-core waveguide used as an absorption cell for concentration measurement of NH3 gas with a CO2 laser," *Optics Letters,* vol. 18, pp. 2150-2152, 1993.

[67] F. Benabid, J. C. Knight, G. Antonopoulos, and P. St. J. Russell, "Stimulated Raman scattering in hydrogen-filled hollow-core photonic crystal fiber," *Science,* vol. 298, pp. 399-402, 2002.

[68] B. Temelkuran, S. D. Hart, G. Benoit, J. D. Joannopoulos, and Y. Fink, "Wavelength-scalable hollow optical fibres with large photonic bandgaps for CO2 laser transmission," *Nature,* vol. 420, pp. 650-653, 2002.

[69] M. Shurgalin, and C. Anastassiou, "A new modality for minimally invasive CO2 laser surgery: flexible hollow-core photonic bandgap fibers," *Biomedical Instrumentation & Technology,* vol. 42, pp. 318-325, 2008.

[70] R. Bicknell, L. King, C. E. Otis, J.-S. Yeo, N. Meyer, P. Kornilovitch, S. Lerner, and L. Seals, "Fabrication and characterization of hollow metal waveguides for optical interconnect applications," *Applied Physics A,* vol. 95, pp. 1059-1066, 2009.

[71] T. Miura, F. Koyama, and A. Matsutani, "Modeling and fabrication of hollow optical waveguide for photonic integrated circuits," *Japanese Journal of Applied Physics,* vol. 41, pp. 4785-4789, 2002.

[72] T. Miura, and F. Koyama, "Low-loss and polarization-insensitive semiconductor hollow waveguide with GaAs/AlAs multi-layer mirrors," *Japanese Journal of*

Applied Physics, vol. 43, pp. L21-L23, 2004.

[73] D. Yin, H. Schmidt, J. P. Barber, and A. R. Hawkins, "Integrated ARROW waveguides with hollow cores," *Optics Express,* vol. 12, pp. 2710-2715, 2004.

[74] W. Yang, D. B. Conkey, B. Wu, D. Yin, A. R. Hawkins and H. Schmidt, "Atomic spectroscopy on a chip," *Nature Photon.,* vol. 1, pp. 331-335, 2007.

[75] V. Karagodsky, B. Pesala, F. G. Sedgwick, and C. J. Chang-Hasnain, "Dispersion properties of high-contrast grating hollow-core waveguides," *Optics Letters,* vol. 35, pp. 4099-4101, 2010.

[76] R. M. Knox, and P. P. Toulios, "Integrated circuit for the millimeter through optical frequency range.," in *Proceedings of the Symposium on Submillimeter Waves, J. Fox and M. H. Schlam, eds. Polytechnic Institute of Brooklyn, p. 497-516*, 1970.

[77] B. Pesala, V. Karagodsky, F. Koyama, and C. J. Chang-Hasnain, "Novel 2D high-contrast grating hollow-core waveguide," in *Conference on Lasers and Electro-Optics/International Quantum Electronics Conference, OSA Technical Digest (CD) (Optical Society of America, 2009), paper CMQ7.*, 2009.

[78] J. Hu, N.-N. Feng, N. Carlie, L. Petit, J. Wang, A. Agarwal, K. Richardson, and L. Kimerling, "Low-loss high-index-contrast planar waveguides with graded-index cladding layers," *Optics Express,* vol. 15, pp. 14566-14572, 2007.

[79] E. A. J. Marcatili, "Bends in optical dielectric guides," *Bell System Technical Journal,* vol. 48, pp. 2103-2132, 1969.

[80] Y. Fink, J. N. Winn, S. Fan, C. Chen, J. Michel, J. D. Joannopoulos, and E. L. Thomas, "A dielectric omnidirectional reflector," *Science,* vol. 282, pp. 1679-1682, 1998.

[81] Y. Vlasov, W. M. J. Green and F. Xia, "High-throughput silicon nanophotonic wavelength-insensitive switch for on-chip optical networks," *Nature Photonics,* vol. 2, pp. 242-246, 2008.

[82] K. Nozaki, T. Tanabe, A. Shinya, S. Matsuo, T. Sato, H. Taniyama, and M. Notomi, "Sub-femtojoule all-optical switching using a photonic-crystal nanocavity," *Nature Photonics,* vol. 4, pp. 477-483, 2010.

[83] U. Krishnamoorthy, K. Li, K. Yu, D. Lee, J. P. Heritage, and O. Solgaard, "Dual-mode micromirrors for optical phased array applications," *Sensors and Actuators A,* Vols. 97-98, pp. 21-26, 2002.

[84] P. F. Van Kessel, L. J. Hornbeck, R. E. Meier, and M. R. Douglass, "A MEMS-based projection display," *Proceedings of the IEEE,* vol. 86, pp. 1687-1704, 1998.

[85] P. F. McManamon, T. A. Dorschner, D. L. Corkum, L. J. Friedman, D. S. Hobbs, M. Holz, S. Liberman, H. Q. Nguyen, D. P. Resler, R. C. Sharp, and E. A. Watson, "Optical phased array technology," *Proceedings of the IEEE,* vol. 84, pp. 268-298, 1996.

[86] B. Wang, G. Zhang, A. Glushchenko, J. L. West, P. J. Bos, and P. F. McManamon, "Stressed liquid-crystal optical phased array for fast tip-tilt waveform correction," *Applied Optics,* vol. 44, pp. 7754-7759, 2005.

[87] D. Engstrom, M. J. O'Callaghan, C. Walker, and M. A. Handschy, "Fast beam steering with a ferroelectric-liquid-crystal optical phased array," *Applied Optics,* vol. 48, pp. 1721-1726, 2009.

[88] M. Jarrahi, R. F. W. Pease, D. A. B. Miller, and T. H. Lee , "Optical switching based on high-speed phased array optical beam steering," *Applied Physics Letters,* vol. 92, p. 014106, 2008.

[89] J. K. Doylend, M. J. R. Heck, J. T. Bovington, J. D. Peters, L. A. Coldren, and J. E. Bowers, "Two-dimensional free-space beam steering with an optical phased array on silicon-on-insulator," *Optics Express,* vol. 19, pp. 21595-21604, 2011.

[90] J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-scale nanophotonic phased array," *Nature,* vol. 493, pp. 195-199, 2013.

[91] B-W. Yoo, M. Megans, T. Chan, T. Sun, W. Yang, C. J. Chang-Hasnain, D. A. Horsley, and M. C. Wu, "Optical phased array using high contrast gratings for two dimensional beamforming and beamsteering," *Optics Express,* vol. 21, pp. 12238-12248, 2013.

[92] B. Momeni, E. S. Hosseini, M. Askari, M. Soltani, and A. Adibi, "Integrated photonic crystal spectrometers for sensing applications," *Optics Communications,* vol. 282, p. 3168–3171, 2009.

[93] K. Kodate and Y. Komai, "Compact spectroscopic sensor using an arrayed waveguide grating," *Journal of Optics A: Pure and Applied Optics,* vol. 10, p. 044011, 2008.

[94] S.-W. Wang, C. Xia, X. Chen, W. Lu, M. Li, H. Wang, W. Zheng, and T. Zhang, "Concept of a high-resolution miniature spectrometer using an integrated filter array," *Optics Letters,* vol. 32, p. 632–634, 2007.

[95] Z. Xia, A. A. Eftekhar, M. Soltani, B. Momeni, Q. Li, M. Chamanzar, S. Yegnanarayanan, and A. Adibi, "High resolution on-chip spectroscopy based on miniaturized microdonut resonators," *Optics Express,* vol. 19, pp. 12356-12364, 2011.

[96] R. A. Crocombe, D. C. Flanders, and W. Atia, "Micro-optical instrumentation for process spectroscopy," in *Proceedings of SPIE 5591, Lab-on-a-Chip: Platforms, Devices, and Applications*, 2004.

Appendix A Closed form of $\mathbf{H}_{n,m}$ and $\mathbf{E}_{n,m}$

The closed form of $\mathbf{H}_{n,m}$ and $\mathbf{E}_{n,m}$ shown in Eq. 2.5a can be derived with Eq. 2.1b, 2.1d, 2.2b and 2.2d. The following presents the results.

TM polarization

$$\mathbf{H}_{n,m} = \Lambda^{-1}\int_0^{\Lambda} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)\,dx$$

$$= \Lambda^{-1}\int_0^{a} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)\,dx + \Lambda^{-1}\int_a^{\Lambda} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)\,dx$$

$$= \mathbf{H}_{n,m,a} + \mathbf{H}_{n,m,s}$$

$$\mathbf{H}_{n,m,a} = \Lambda^{-1}\int_0^{a} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)\,dx$$

$$\mathbf{H}_{n,m,s} = \Lambda^{-1}\int_a^{\Lambda} \mathcal{H}_{y,m}^{II}(x)\mathcal{H}_{y,n}^{III,*}(x)\,dx$$

$$\mathbf{E}_{n,m} = \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}|\gamma_n|/k_0\right)^{-2}\int_0^{\Lambda} \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,n}^{III,*}(x)\,dx$$

$$= \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}|\gamma_n|/k_0\right)^{-2}\int_0^{a} \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,n}^{III,*}(x)\,dx + \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}|\gamma_n|/k_0\right)^{-2}\int_a^{\Lambda} \mathcal{E}_{x,m}^{II}(x)\mathcal{E}_{x,n}^{III,*}(x)\,dx$$

$$= \left(\mathbf{H}_{n,m,a} + n_{bar}^{-2}\mathbf{H}_{n,m,s}\right)(\beta_m/\gamma_n)$$

$$\mathbf{H}_{n,m,a} = \frac{1}{\Lambda}\left(\frac{1}{k_{xn}^2 - k_{a,m}^2}\right)\Big[-\left(1 + e^{ik_{xn}a}\right)\sin\left(k_{a,m}\frac{a}{2}\right)\left(iB_m k_{xn} + A_m k_{a,m}\right) + \left(-1 + e^{ik_{xn}a}\right)\cos\left(k_{a,m}\frac{a}{2}\right)\left(-iA_m k_{xn} + B_m k_{a,m}\right)\Big]$$

$$\mathbf{H}_{n,m,s} = \frac{1}{\Lambda}\left(\frac{1}{k_{xn}^2 - k_{s,m}^2}\right)\Big\{D_m\Big[-i\left(e^{ik_{xn}a} + e^{ik_{xn}\Lambda}\right)\sin\left(k_{s,m}\frac{s}{2}\right)k_{xn} - \left(e^{ik_{xn}a} - e^{ik_{xn}\Lambda}\right)\cos\left(k_{s,m}\frac{s}{2}\right)k_{s,m}\Big] + C_m\Big[i\left(e^{ik_{xn}a} - e^{ik_{xn}\Lambda}\right)\cos\left(k_{s,m}\frac{s}{2}\right)k_{xn} - \left(e^{ik_{xn}a} + e^{ik_{xn}\Lambda}\right)\sin\left(k_{s,m}\frac{s}{2}\right)k_{s,m}\Big]\Big\}$$

TE polarization

The formulations for TE polarization are similar as those for TM polarization, except for the following change.

Change $\mathcal{H}_y$ into $\mathcal{H}_x$, and $\mathcal{E}_x$ into $\mathcal{E}_y$. Change $\mathbf{E}_{n,m}$ into the following:

$$\mathbf{E}_{n,m} = \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}k_0/|\gamma_n|\right)^{-2}\int_0^{\Lambda}\mathcal{E}_{y,m}^{II}(x)\mathcal{E}_{y,n}^{III,*}(x)\,dx$$

$$= \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}k_0/|\gamma_n|\right)^{-2}\int_0^{a}\mathcal{E}_{y,m}^{II}(x)\mathcal{E}_{y,n}^{III,*}(x)\,dx$$

$$+ \Lambda^{-1}\left(\sqrt{\mu_0/\varepsilon_0}k_0/|\gamma_n|\right)^{-2}\int_a^{\Lambda}\mathcal{E}_{y,m}^{II}(x)\mathcal{E}_{y,n}^{III,*}(x)\,dx$$

$$= \left(\mathbf{H}_{n,m,a} + \mathbf{H}_{n,m,s}\right)(\gamma_n/\beta_m)$$

Appendix B HCG T-Matrix and Fresnel's Law

In Section 2.2.4, the expressions of $\boldsymbol{H}$, $\boldsymbol{E}$, $\boldsymbol{\rho}$, $\boldsymbol{\rho}'$, $\boldsymbol{t}$ and $\boldsymbol{t}'$ are derived for HCG. However, they are general and can be applied to other systems as well. For example, if the two regions are homogeneous media, and there is only a plane wave mode in each region, then $\boldsymbol{\rho}$, $\boldsymbol{\rho}'$, $\boldsymbol{t}$ and $\boldsymbol{t}'$ are reduced to the plane wave reflection and transmission coefficient between two media, denoted by the Fresnel's Law. The following provides the derivation. We consider light incidence from Region I, and transmit into Region II. The refractive index, relative permittivity and relative permeability is n_1, ε_1 and 1 respectively in Region I, and n_2, ε_2 and 1 in Region II. At the boundary between Region I and II, the light incident angle is θ_1, and refraction angle is θ_2.

TM polarization

The magnetic and electrical field in Region I can be expressed as

$$H_{y1} = e^{-ik_{z1}z}e^{-ik_{x1}x}$$

$$E_{x1} = \frac{1}{\varepsilon_1}\frac{k_{z1}}{k_0}\sqrt{\frac{\mu_0}{\varepsilon_0}}e^{-ik_{z1}z}e^{-ik_{x1}x} = \frac{1}{n_1}cos\theta_1\sqrt{\frac{\mu_0}{\varepsilon_0}}e^{-ik_{z1}z}e^{-ik_{x1}x}$$

Likewise, the magnetic and electrical field in Region II can be expressed as

$$H_{y2} = \left(\frac{n_2 cos\theta_1}{n_1 cos\theta_2}\right)e^{-ik_{z2}z}e^{-ik_{x2}x}$$

$$E_{x2} = \left(\frac{n_2 cos\theta_1}{n_1 cos\theta_2}\right)\frac{1}{n_2}cos\theta_2\sqrt{\frac{\mu_0}{\varepsilon_0}}e^{-ik_{z2}z}e^{-ik_{x2}x}$$

where

$$k_{zi} = k_i cos\theta_i = n_i k_0 cos\theta_i$$

$$k_{xi} = k_i sin\theta_i = n_i k_0 sin\theta_i$$

Following Eq. 2.16, we have

$$H = \frac{\int_0^\Lambda \mathcal{H}_{y2}(x)\left[\mathcal{H}_{y1}(x)\right]^* dx}{\int_0^\Lambda \left|\mathcal{H}_{y1}(x)\right|^2 dx} = \left(\frac{n_2 cos\theta_1}{n_1 cos\theta_2}\right)$$

$$E = \frac{\int_0^\Lambda \mathcal{E}_{x2}(x)[\mathcal{E}_{x1}(x)]^* dx}{\int_0^\Lambda |\mathcal{E}_{x1}(x)|^2 dx} = 1$$

$$\boldsymbol{\rho}' = -\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}(\boldsymbol{H}-\boldsymbol{E})\boldsymbol{H}^{-1} = \frac{-cos\theta_1 n_2 + cos\theta_2 n_1}{cos\theta_1 n_2 + cos\theta_2 n_1} = r_{21}$$

$$\boldsymbol{\rho} = (\boldsymbol{H}+\boldsymbol{E})^{-1}(\boldsymbol{H}-\boldsymbol{E}) = \frac{cos\theta_1 n_2 - cos\theta_2 n_1}{cos\theta_1 n_2 + cos\theta_2 n_1} = r_{12}$$

$$\boldsymbol{t}' = 2(\boldsymbol{H}+\boldsymbol{E})^{-1} = \frac{2n_1 cos\theta_2}{n_1 cos\theta_2 + n_2 cos\theta_1}$$

Note that

$$\boldsymbol{t}' = \frac{E_{x2}}{E_{x1}} = \frac{E_2 cos\theta_2}{E_1 cos\theta_1}$$

where

$$\boldsymbol{E_i} = \boldsymbol{E_{xi}} + \boldsymbol{E_{zi}}$$

Thus

$$\frac{E_2}{E_1} = \boldsymbol{t}'\frac{cos\theta_1}{cos\theta_2} = \frac{2n_1 cos\theta_1}{n_1 cos\theta_2 + n_2 cos\theta_1} = t_{21}$$

$$\boldsymbol{t} = 2\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}\boldsymbol{E} = \frac{2n_2 cos\theta_1}{n_1 cos\theta_2 + n_2 cos\theta_1}$$

Thus

$$\frac{E_1}{E_2} = \boldsymbol{t}\frac{cos\theta_2}{cos\theta_1} = \frac{2n_2 cos\theta_2}{n_1 cos\theta_2 + n_2 cos\theta_1} = t_{12}$$

The above formulation agree well with r_{21}, r_{12}, t_{21}, t_{12} given by the Fresnel's Law for p polarization light.

TE polarization

$$H_{x1} = e^{-ik_{z1}z}e^{-ik_{x1}x}$$

$$E_{y1} = -\frac{k_0}{k_{z1}}\mu_1\sqrt{\frac{\mu_0}{\varepsilon_0}}e^{-ik_{z1}z}e^{-ik_{x1}x} = -\frac{1}{n_1 cos\theta_1}\sqrt{\frac{\mu_0}{\varepsilon_0}}e^{-ik_{z1}z}e^{-ik_{x1}x}$$

$$H_{x2} = \left(\frac{n_2 cos\theta_2}{n_1 cos\theta_1}\right)e^{-ik_{z2}z}e^{-ik_{x2}x}$$

$$E_{y2} = -\left(\frac{n_2 cos\theta_2}{n_1 cos\theta_1}\right)\frac{1}{n_2 cos\theta_2}\sqrt{\frac{\mu_0}{\varepsilon_0}}e^{-ik_{z2}z}e^{-ik_{x2}x}$$

$$H = \frac{\int_0^{\Lambda}\mathcal{H}_{x2}(x)[\mathcal{H}_{x1}(x)]^* dx}{\int_0^{\Lambda}|\mathcal{H}_{x1}(x)|^2 dx} = \left(\frac{n_2 cos\theta_2}{n_1 cos\theta_1}\right)$$

$$E = \frac{\int_0^{\Lambda}\mathcal{E}_{y2}(x)\left[\mathcal{E}_{y1}(x)\right]^* dx}{\int_0^{\Lambda}\left|\mathcal{E}_{y1}(x)\right|^2 dx} = 1$$

$$\boldsymbol{\rho}' = -\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}(\boldsymbol{H}-\boldsymbol{E})\boldsymbol{H}^{-1} = \frac{-cos\theta_2 n_2 + cos\theta_1 n_1}{cos\theta_2 n_2 + cos\theta_1 n_1} = r_{21}$$

$$\boldsymbol{\rho} = (\boldsymbol{H}+\boldsymbol{E})^{-1}(\boldsymbol{H}-\boldsymbol{E}) = \frac{cos\theta_2 n_2 - cos\theta_1 n_1}{cos\theta_2 n_2 + cos\theta_1 n_1} = r_{12}$$

$$\boldsymbol{t}' = 2(\boldsymbol{H}+\boldsymbol{E})^{-1} = \frac{2n_1 cos\theta_1}{n_2 cos\theta_2 + n_1 cos\theta_1} = t_{21}$$

$$\boldsymbol{t} = 2\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}\boldsymbol{E} = \frac{2n_2 cos\theta_2}{n_1 cos\theta_1 + n_2 cos\theta_2} = t_{12}$$

The above formulation agree well with r_{21}, r_{12}, t_{21}, t_{12} given by the Fresnel's Law for s polarization light.