INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

University of Alberta

The Generalized Mach-Zehnder Interferometer using Multimode Interference Couplers for Optical Communications Networks

by

Neil Satish Lagali ©

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

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Edmonton, Alberta

Spring, 2000

University of Alberta

Library Release Form

Name of Author: **Neil Satish Lagali**

Title of Thesis: **The Generalized Mach-Zehnder Interferometer using Multimode Interference Couplers for Optical Communications Networks**

Degree: **Doctor of Philosophy**

Year this Degree Granted: **2000**

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

_____

23 Lakeside Drive
Brantford, Ontario
Canada
N3R 5J5

Date: ___April 10, 2000___

"The voyage of discovery is not in seeking new landscapes,
but in having new eyes."

– Marcel Proust

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **The Generalized Mach-Zehnder Interferometer using Multimode Interference Couplers for Optical Communications Networks** submitted by **Neil Satish Lagali** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

R. I. MacDonald, Supervisor
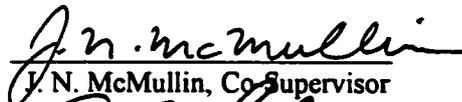
J. N. McMullin, Co-Supervisor
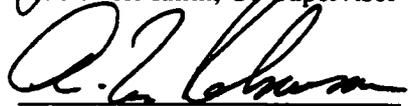
A. M. Robinson

R. P. W. Lawson

M. R. Freeman

D. V. Plant, External Examiner

H. G. Schmidt-Weinmar

Date: Feb. 1, 2000 .

For my Parents, Sudhakar and Kokila.

*Their love and support through the years
have made this work possible.*

Abstract

The explosive growth in telecommunications traffic in recent years has hastened the emergence of optical communications networks. As the volume and complexity of network traffic increases, efficient, flexible, and cost-effective methods of routing and distributing optical signals are being sought. One method for meeting these requirements is through the use of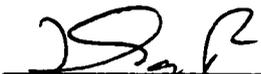 integrated optics technology, which enables the realization of miniaturized optical components and circuitry on a planar substrate surface.

In this work the integrated optic multiple-arm or generalized Mach-Zehnder interferometer (GMZI) using multimode interference (MMI) couplers is studied. Through an investigation into the principles governing the operation of this device, a mathematical theory describing both its idealized and observed properties is developed. This theory predicts the operation of the GMZI as a variable-ratio optical power splitter, which in specific instances functions as an optical switch. Both the power splitting and switching functions of the GMZI are useful for efficient optical signal distribution, routing, and protection switching applications in emerging optical communications networks.

The variable-ratio optical power splitting function allows an input optical signal to be distributed in a controlled, dynamic manner to a number of output channels. This provides an efficient and flexible means of broadcasting an optical signal to multiple users. The method of operating this novel device is described and its strengths and limitations are assessed.

A consideration of the optical switch function reveals that various levels of symmetry are present in the operation of the GMZI. This symmetry is used to develop new designs for $1 \times N$ optical switches with reduced operational requirements. Further consideration of the multiple-input switch properties leads to new designs for non-blocking $N \times N$ optical switches that possess the benefits of compactness and a simple design and fabrication procedure. The switches enable

optical signal routing and network protection and reconfiguration switching to be realized with compact components exhibiting low power consumption levels. The operation and performance issues of these switches are examined in detail.

To demonstrate the feasibility of the manufacture and operation of the GMZI and the validity of the theory upon which it is based, a number of integrated optical waveguide components including MMI couplers and GMZI devices have been realized. The devices were fabricated in a high refractive index contrast silicon oxynitride (SiON) based planar waveguide system, utilizing the thermo-optic effect for active operation.

The measured device performance is reported, and demonstrations of efficient thermo-optic operation, $1 \times N$ switching in 2×2 and 4×4 GMZI devices, and variable-ratio power splitting and multi-port switching in a 4×4 GMZI device have been achieved. The observed device characteristics validate the tenets of the mathematical theory developed in this work, and demonstrate the feasibility of using the GMZI to perform the complex optical functions required for state-of-the-art optical communications networks.

Acknowledgements

I would first like to express my most sincere thanks to my supervisor Dr. R. Ian MacDonald for introducing me into the field of optical communications and thereafter giving me the freedom to learn and grow, and for opening doors of opportunity that few students ever witness. I have had the extreme good fortune to share in his knowledge, insight, and vision both within and beyond the bounds of this work. I would also like to thank my co-supervisor Dr. James N. McMullin for his help and mentorship as I prepared this work.

I am indebted to my friend, colleague, and mentor Dr. Reza Paiam. He has provided the motivation for this research through pioneering efforts in the area of MMI couplers and has offered critical insights and direction in this work from its inception through to its completion. His hand is implicit in much of the work reported in these pages.

I wish to express my gratitude to Dr. Alfred Driessen and Dr. Kerstin Wörhoff, my collaborators in the experimental portion of this work. They have not only graciously provided me with the facilities of MESA and the Lightwave Devices Group at the University of Twente, The Netherlands, but their helpfulness, enthusiasm and friendship have brought much excitement into this work and provided me with a first-hand experience of the wonderful Dutch hospitality.

I would like to thank my colleagues in Edmonton, Alberta at the Telecommunications Research Laboratories (TRLabs), David Boertjes, Ray DeCorby, Alan Hnatiw, and Craig Unick for their help in the early stages of this work. I would also like to acknowledge the help of Dr. James Broughton at the Alberta Microelectronic Center.

I am very grateful to those at the University of Alberta, including my supervisory committee, for their diligence in examining this work. In particular, I would like to thank Michelle Lock for her tireless efforts in making things run smoothly while I've been thousands of kilometers away.

I would also like to acknowledge the help of many people at JDS Uniphase in Nepean, Ontario, in particular Liz Nowak, Catherine Murphy, Vincent Delisle, Rajiv Iyer, Malcolm Smith, Suresh Mistry, Cynthia Wilson, Jim Yang, and Barrie Keyworth.

At the MESA Research Institute, I would like to thank Nicole, Gerard, Peter, and Johnny for their training and help in the clean room, and Bert Otter for producing the SEM images.

In the Lightwave Devices Group at the University of Twente, I would like to thank the many people who have provided a fun and exciting environment for both research and pursuits beyond. I am particularly grateful to Anton Hollink for his constant support and technical wizardry in hardware and software, and diligence in all matters optical, electrical, mechanical, and procedural. I would like to thank Toon Andringa for offering his expertise in the clean room and for his artful work in wire bonding. Thanks are also due to Erwin Rikkers and Chris Roeloffzen for their help with the device characterization equipment. Thanks also to Sami Musa for graciously performing the transmission measurements for the annealed samples. Numerous others deserve mention: Marcel, Tjarko, Gregory, Dion, Simone, Rene, Harm, Ton, Gert, Wim, Robert, Libor, Lin, Hugo, Henk, and Remco.

I would also like to acknowledge the University of Alberta, TRLabs, JDS Uniphase Inc., and the Natural Sciences and Engineering Research Council of Canada for their generous financial support.

I would like to express a very special thanks to Theresa, for her wisdom, patience, understanding, and kindness, and for teaching me to focus on what is really important in life.

Finally, I am very grateful to my family, my Parents, and my sister Pamela. Their constant love, support, and encouragement is a continuous source of joy and strength.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

a	free design parameter for the lateral offset used in positioning the access ports of an MMI coupler
a_{ij}, a'_{ij}	real field amplitude transfer coefficient of an MMI coupler, MMI combiner
a_v	complex amplitude coefficient of the v-th normalized mode in an optical waveguide
A_1, A_2	real-valued constants
AB, ABC	binary logic control words
b	integer variable denoting the MMI coupler image number
c	speed of light in free space
C	complex normalization constant used in self-imaging analysis
d	etch depth during RIE process
d_s	GMZI arm perpendicular separation distance
e	exponential; $e = 2.7182818...$
E	electric field
E_i^{in}	complex field amplitude at MMI coupler input port i
E_j^{out}	complex field amplitude at MMI coupler output port j
f	time-dependent complex field or wave function
F, F_v, $\hat{F}_v$	real-valued amplitude distribution of a modal field, F of the v-th mode, normalized amplitude distribution
h, h_v	mode parameter, h of the v-th mode
h_h	height (maximum offset) of s-bend
H	magnetic field
I	identity matrix of order N
I_0	total input optical intensity
I^{out}	output optical intensity
i	index denoting the input ports of an MMI coupler or GMZI
i'	index denoting a specific MMI coupler input port
j	index denoting the output ports of an MMI coupler or the arms of a GMZI
j_{ref}	index denoting an MMI coupler reference output port
J	$\sqrt{-1}$
k	index denoting the output ports of a GMZI
k_1, k_2	complimentary sets of GMZI output ports

k'	index denoting a specific GMZI output port
$\bar{k}$	index denoting the GMZI output port from which switched light emerges
k_0	free space wavenumber
K	number of self-images in a passive GMZI when N is even, such that M and K have no common divisors
l_0	propagation length in a waveguide
l_b	length of s-bend
L	interaction length for a directional coupler
L_{GMZI}	imaging length of a passive GMZI
$L_{M,N}$	M-th, N-fold imaging length in an MMI coupler
$L_{M,N\,design}$	design length of MMI coupler based on MPA
L_π	beat length
$L_{\pi,\,TE}$, $L_{\pi,\,TM}$	beat lengths for TE and TM polarizations
L_t	thermo-optic heater length
M	period number of the periodic self-imaging length in an MMI coupler, such that M and N have no common divisors
m	number of guided modes in an optical waveguide (Chapter 2), fringe number (Chapter 6), integer variable
n, n_l	index of refraction, refractive index of the l-th layer in a slab waveguide
n_{ch}	slab effective index (transverse analysis) at the guiding region of a channel waveguide
n_{cl}	slab effective index (transverse analysis) at the cladding region of a channel waveguide
n_e	effective index (longitudinal analysis) of a channel waveguide
$n_{e,\,TE}$, $n_{e,\,TM}$	effective index of fundamental TE and TM waveguide modes
$n_{e,\,TE-TM}$	effective index birefringence of channel waveguide
$n_{eff,\,v}$	effective index of the v-th mode of a slab waveguide
n_m	complex refractive index of a metal
n_p	refractive index of a prism
n_r , n_{im}	real and imaginary parts of complex refractive index of a metal
N	number of ports in an MMI coupler, GMZI, or non-blocking switch
p, q, p_v, q_v	mode parameters (Chapter 2), internal GMZI arm index, p, q of the v-th mode
s_{ij} , s_{jk}	elements of MMI coupler transfer matrices
$S_{combiner}$	transfer matrix of an $N \times N$ MMI coupler used as a combiner in a GMZI
$S_{coupler}$	transfer matrix of an $N \times N$ MMI coupler
S_{shift}	transfer matrix of phase shifts applied to GMZI arms

$S_{splitter}$	transfer matrix of an $N \times N$ MMI coupler used as a splitter in a GMZI
t	time
t_{ik}	elements of the transfer matrix of a GMZI
T	complex field in an optical waveguide (Chapter 2), temperature (Chapter 5), matrix transpose operation (superscript), total transfer matrix of the GMZI
T_{ext}	periodically extended field at the input plane of an MMI coupler
w_1, w_2	half-widths of Gaussian field approximations to fundamental guided modes
W	thickness of a slab waveguide, width of a rib waveguide
W_e	effective width of a multimode waveguide
x	numerical phase shift iterate, solution vector
x_b	lateral position of self-image b at the output plane of an MMI coupler
x_{opt}	optimized phase shift solution vector
x, y, z	Cartesian spatial coordinates
X	matrix of idealized switching phase shifts
Y	vector containing functions to be numerically minimized

Greek Symbols

α	thermo-optic coefficient (Chapter 5), waveguide propagation loss
β, β_v	propagation constant, β of the v-th mode
$\delta\beta$	complex propagation constant perturbation for four-layer substrate slab analysis
δd	axial displacement between centers of fundamental fiber and waveguide modes
$\delta\varphi_{ij}$, $\delta\varphi'_{ij}$	phase deviation of a self-image in an MMI coupler, MMI combiner
$\delta\varphi_j$	phase error in a GMZI arm
δL	deviation from the optimum imaging length in an MMI coupler
δ_p	buffer layer thickness for maximum TM mode attenuation
Δ	phase difference of orthogonal polarization components of a reflected laser beam
$\Delta\beta$	change in propagation constant due to thermo-optic effect
$\Delta\varphi_j$	active phase shift applied to a GMZI arm
$\Delta L_{M,N}$	polarization-dependent MMI coupler imaging length variation
ΔL_π	difference between MMI coupler beat lengths for TE and TM polarization
ΔL_t	change in heater length due to thermal expansion
$\Delta\lambda$	wavelength channel spacing of the MMI phasar multiplexer

Δn	core-cladding refractive index contrast
ΔT	change in temperature
ε_0	electric permittivity of free space
ε_m	complex dielectric constant of a metal
ε_r, ε_{im}	real and imaginary parts of complex dielectric constant of a metal
ϕ_1	constant offset value for phases of self-images in an MMI coupler
ϕ_p	acute prism angle
φ_b	phases of self-image b at the output plane of an MMI coupler
φ_{ij}	phase of a self image in an MMI coupler
Φ	matrix of phases of self-images of an MMI coupler
γ	applied phase shift parameter
λ	free space wavelength
λ_b	wavelength channel of the MMI phasar multiplexer
λ_L	laser wavelength in RIE etch depth measurement
μ_0	magnetic permeability of free space
ν	mode number or order
π	ratio of the circumference of a circle to its diameter; $\pi = 3.14159\ldots$
θ	angular parameter for GMZI waveguide arm design
θ_i	incident beam angle at wafer surface in RIE etch depth measurement
$\theta_{i,\nu}$, $\theta_{e,\nu}$	internal, external resonant angles for prism coupling to the ν-th mode of a slab waveguide
σ_l	polarization parameter for the l-th layer in a slab waveguide
ω	angular frequency
$\psi_{i,k}$	function used in the derivation of GMZI cross-connect states
Ψ	amplitude ratio of orthogonal polarization components of a reflected laser beam

Abbreviations

AlGaAs	Aluminum Gallium Arsenide
APC	Angular Pressure Control
BPM	Beam Propagation Method
CCD	Charge-Coupled Device
CF_4	Carbon Tetrafluoride
Cr	Chromium

DC	Direct Current
EIM	Effective Index Method
GaAs	Gallium Arsenide
GMZI	Generalized Mach-Zehnder Interferometer (using multimode interference couplers)
H	Hydrogen
HeNe	Helium-Neon
HMDS	Hexamethyldisilizane
HNO_3	Nitric Acid
H_2O	water
IR	Infrared
$LiNbO_3$	Lithium Niobate
MESA	Microelectronics, Materials Engineering, Sensors & Actuators Research Institute
MFD	Mode Field Diameter
MMI	Multimode Interference
MPA	Modal Propagation Analysis
N, N_2	Nitrogen, Nitrogen Gas
NA	Numerical Aperture
NH_3	Ammonia
N_2O	Nitrogen Oxide
O_2	Oxygen
PECVD	Plasma Enhanced Chemical Vapor Deposition
phasar	Phased Array
RAE	Rotating Analyzer Ellipsometer
RIE	Reactive Ion Etching
SEM	Scanning Electron Microscope
Si	Silicon
SiH_4	Silane
Si_3N_4	Silicon Nitride
SiO_xN_y, SiON	Silicon Oxynitride
SiO_2, SiO_x, silica	Silicon Oxide
TE	Transverse Electric
TM	Transverse Magnetic
WDM	Wavelength Division Multiplexing

Chapter 1

Introduction

At the dawn of a new millennium, the trend towards an increasingly information-based society is firmly entrenched. Remarkable technological achievements have in recent years brought about unparalleled progress in the communication of information. As developments continue, an associated rise in the thirst for information is demanding an examination of new distribution mechanisms. This chapter defines the context of the present work within the broad field of communication, and discusses its role in facilitating the efficient dissemination of information.

1.1 Optical communications

In modern society there is an ever-increasing demand for the acquisition, processing, and sharing of information concerning the world and its future course. Major technological revolutions in the communication of knowledge have fueled this demand, and have continued to accelerate at an astounding pace in recent years.

The motivation for the rapid developments in the field of communications technology can be traced back to significant advances in the electronic processing of information signals achieved in the past few decades. Miniaturization and mass production of electronic circuits had for the first time provided the means for handling large quantities of data in a compact and cost-efficient manner. The ensuing advent of widespread voice, facsimile, video, computer, and internet services created an appetite for data communication that pushed beyond the capacity limits of the information transmission media in use, the copper wire pair and coaxial cable.

Fortunately, a parallel set of developments in the field of optics emerged to provide a solution to this capacity problem, by combining the tremendous information capacity inherent in optical signals with a new transmission medium, the optical fiber. The first major development began with experiments by Kao in 1966 [1], demonstrating that high-speed information could be sent down a narrow filament (fiber) of glass, confined by total internal reflection. At about the same time, in a seemingly unrelated series of events, the semiconductor diode laser was invented [2]. Remarkable progress made in both optical signal generation and transmission, such as the development of powerful, low-cost laser diodes and high-speed modulation techniques, and a reduction in optical fiber transmission losses by orders of magnitude, ushered in the era of modern optical communications.

The first generation of optical communication systems appeared in the 1970's and relied on light-emitting diode sources operating at visible wavelengths near 800 nm, transmitted through multimode optical fibers. These systems were capable of transmitting data at rates of about 100 Mbit/s over a single fiber span of about 10 km. For comparison, a standard telephone conversation requires a 64 kbit/s data rate, so a 100 Mbit/s signal can carry approximately 1500 simultaneous telephone calls. A switch to a transmission wavelength of 1300 nm in the early 1980's characterized the second generation of optical communication systems. At this wavelength, optical fibers exhibit lower signal attenuation and a zero dispersion. The elimination of modal dispersion effects by the use of single-mode fiber combined with the development of semiconductor laser sources and detectors operating at this wavelength enabled data rates of about 400 Mbit/s to be sent over a 40 km fiber link. The third generation of optical communication systems emerged in the mid-1980's to exploit the minimum attenuation regime of optical fiber occurring at a 1550 nm wavelength. Since this wavelength corresponds to a nonzero dispersion, special dispersion-compensated fibers and/or monochromatic laser sources were

utilized to realize single-mode optical fiber transmission rates up to the Gbit/s level over an optical fiber link length measured in the hundreds of kilometers.

Modern optical communications systems build upon the results achieved in the third telecommunication window around 1550 nm. The most notable advancements have been the introduction of rare-earth doped optical fiber amplifiers providing optical signal gain in the 1550 nm region, and the introduction of wavelength division multiplexing (WDM). The latter technology expands the capacity of a system by simply transmitting several information-carrying signals – each corresponding to a distinct optical wavelength – simultaneously through a single optical fiber. This has the dramatic effect of allowing every installed optical fiber cable to carry 10 or 100 times the traffic it was originally designed for, by using densely multiplexed wavelength channels and connecting each fiber to multiple-wavelength sources and receivers. These and other recent technologies have combined to expand the transmission rate and fiber link distance such that present state-of-the-art systems have successfully achieved transmission rates of hundreds of Gbit/s or a few Tbit/s over fiber link lengths in the hundreds and even thousands of km [3], [4].

1.2 Optical networks

The tremendous achievements in optical data transmission have resulted in a massive global deployment of optical fiber transmission lines, spanning a total distance that must now be measured in the hundreds of millions of kilometers.

Up until very recently, the story of the progress made in point-to-point optical fiber transmission has sufficed to capture all the interesting developments in modern optical communications. The dominant challenge was to build fiber links to support the growing amount of data traffic that needed to be transmitted.

In the last decade, however, an interesting evolution has been occurring in the world of optical communications. The major technology issues are now shifting from those of optical links to that of active control over optical signal paths through various information distribution topologies, as shown schematically in Fig. 1.1. This interest has been triggered by the explosive growth in high-data-rate multimedia traffic over the past few years. Increasing user demands for high-performance connections and an expanding number of users have initiated the development of 'fiber-to-the-home' or 'fiber-to-the-office' hub systems and 'local area' or 'metropolitan area' network systems.
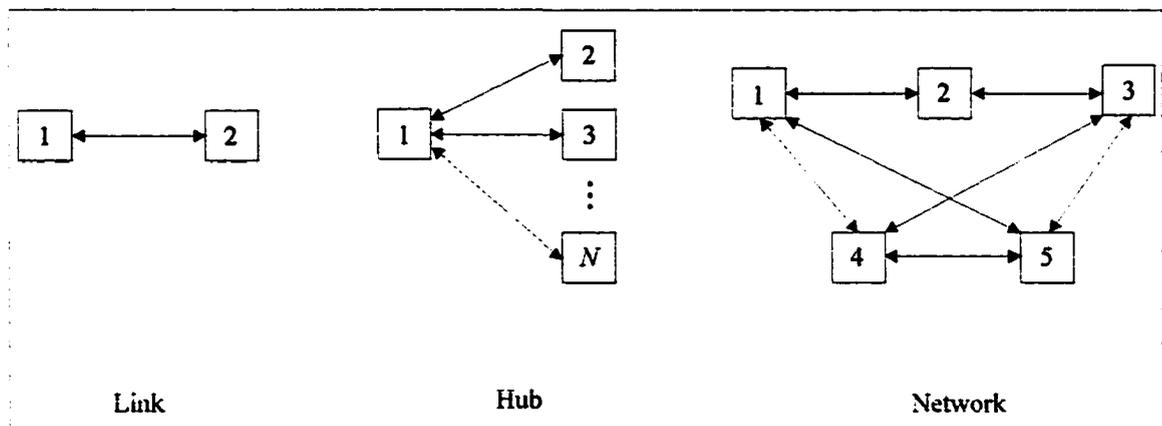
Figure 1.1 Three basic information distribution topologies, link, hub, and network.

The link topology in Fig. 1.1 represents the simplest use of optical technology in a communication system. The nodes transmit and receive optical signals and perform all other functions electronically. In a hub design, information is distributed to multiple nodes from a

single transmitter. Optical technologies are now being deployed to realize such hubs, and a good example of this is the hub-type passive optical network emerging for the distribution of cable television services. In these systems, fiber is used as the transmission medium, however, some optical functions such as passive signal distribution and wavelength multiplexing are beginning to permeate into the nodes. Similarly, in the full network structure shown in Fig. 1.1, optical technologies are not only providing the links between the nodes, but within the nodes they are beginning to be utilized for signal processing functions (e.g., channel add/drop) and also to provide some switching capability.

The hub and network topologies in Fig. 1.1 need not be mutually exclusive and can be combined; both topologies will hereafter be referenced by the term 'optical network'.

As optical networks develop, an increasing need for such vital functions as signal distribution, switching, protection, and reconfiguration is evident, due to the volume and complexity of network traffic and the high cost of lost high-throughput links. In the first generation of optical networks, these functions have been implemented electronically, with optical fiber simply replacing the old copper lines as a transmission medium. It is rapidly becoming apparent, however, that the information throughput of the network is limited by bottlenecks at the network nodes, where the electronic functions are performed.

The second generation of optical networks is currently under active development, and aims to achieve optically transparent paths through a network by removing the electrical components and exploiting the unique properties of optics. Active node functions in these networks can be performed by means of micro-optic components where light leaves the fiber and is manipulated through the use of elements such as lenses or mirrors while the beam is unguided, or by means of novel, *integrated* optical devices.

Unguided propagation technologies, however, suffer drawbacks such as optical coupling, reflection, and scattering losses as a beam passes through discrete components, sensitivity to environmental fluctuations, scalability to larger capacities, and an increasing cost and complexity of manufacture.

1.3 Integrated optics

To perform the required optical network node functions, an alternative technology has been sought. The technology should be capable of synthesizing components such as beam splitters/combiners for branching/combining optical signals, optical routers and switches for changing optical paths, wavelength multi/demultiplexers for combining/separating multiple wavelength channels, and channel add/drop elements for increased network flexibility. Furthermore, the components should meet practical requirements such as signal format and rate independence, low optical losses, polarization insensitivity, low-power active operation, small size, cost-effective manufacture, and reliability issues.

Recently, the technology of integrated optics is attracting much interest as a candidate to meet these demanding requirements. Integrated optical devices are superior to many of their micro-optic or fiber-optic counterparts in terms of size, freedom of design, ease of coupling between components, mass-production possibilities, device stability, and compatibility with integrated transceivers and electronic control circuitry [5].

The concept of integrated optics was first proposed by S. E. Miller in a landmark paper in 1969 [6]. The idea was that using well-established processes from the integrated electronics industry, planar lightwave circuits containing passive and active guided-wave optical components could be fabricated. Such circuits would contain many complex and diverse optical devices on a single wafer, linked together by transparent light conduits.

An analogy between integrated electronics and integrated optics is often drawn, and the significant gains and widescale deployment achieved in integrated circuit technology are expected to one day be paralleled in the realm of integrated optics. So far, integrated optics has

fallen somewhat short of these expectations, mainly due to difficulties in obtaining ultra-low losses, an immature fabrication technology, and the costs associated with fiber-to-chip connections and packaging [7]. Furthermore, the miniaturization of electronics cannot be similarly achieved in the optical realm, since optical circuitry requires component sizes of the order of the wavelength of light used, which are much larger than electronic dimensions.

Significant research addressing these areas continues and many gains have been made, leading to the widespread introduction of the first integrated optical components into the marketplace in the last two to three years. Research in the field of integrated optics continues to expand, exploiting the unique functions accessible through the use of planar optical guided-wave, or 'waveguide' components.

It is the goal of this work to continue to expand the sphere of knowledge and applicability of integrated optics by proposing new integrated optical components for emerging optical networks and demonstrating the feasibility of their manufacture and use. As the components dealt with in this work are based upon the principles of integrated optical couplers, a brief introduction to this area is appropriate.

1.4 Integrated optical couplers

A key component from which passive and active integrated optical waveguide devices can be constructed is the waveguide coupler, which performs the basic yet important function of controllably distributing optical power among waveguides. The directional coupler is perhaps the simplest example and is realized by bringing two single-mode waveguide channels in close proximity along a certain interaction length L, as shown in Fig. 1.2 [8]. As the input light from one waveguide propagates through the interaction region, the optical power is gradually transferred, or coupled into the other waveguide. By using a specific value of L determined by the strength of the interaction between the guides, a desired power transfer ratio can be achieved, such as 0 % power transfer to waveguide 2 (bar state), 100 % power transfer to waveguide 2 (cross state), or 50 % power transfer (a 3 dB coupler state).
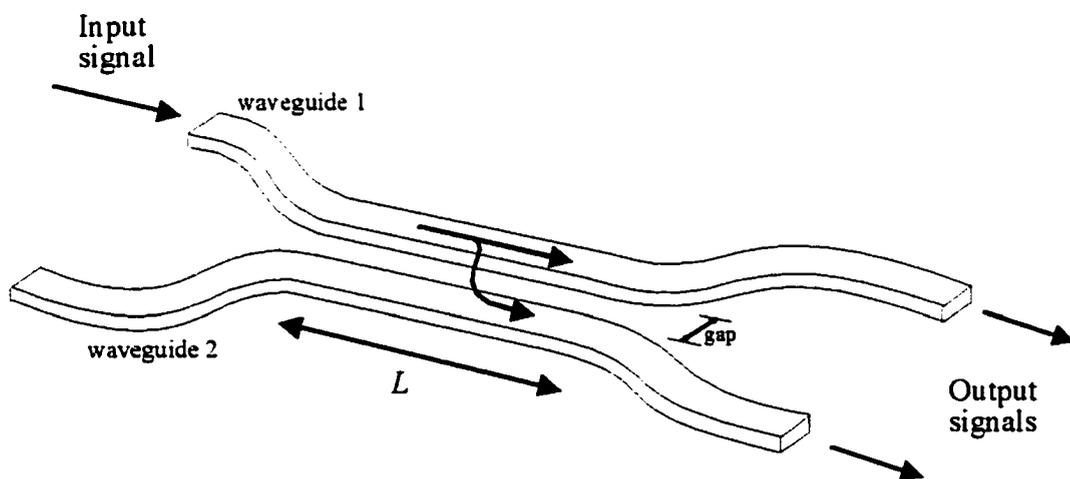
Figure 1.2 The integrated optical directional coupler.

The power transfer ratio is strongly dependent upon the physical and optical properties of the waveguides, the wavelength of the light, the gap between the waveguides, and the interaction length L. To fabricate a directional coupler, very accurate control of several of these design parameters is needed, which may be difficult to achieve in practice [9]. Also if a signal is to be

split into many signals or if many signals are to be combined, a number of directional couplers must be cascaded, resulting in large, lossy, and complex devices [10].

An alternate method of optical power transfer has gained widespread popularity in recent years and utilizes the self-imaging effect in multimode waveguides [11]. In this approach, the parallel single-mode waveguides and the gap separating them along the distance L are replaced by a wider multimode waveguide as shown in Fig. 1.3. The single-mode waveguides are spaced so that coupling between them is negligible, thereby confining the interaction between the waveguides to the multimode region. The input field from a single-mode waveguide propagates as a linear combination of the allowed modes in the multimode region. The multiple modes have different propagation constants and therefore move out of phase as they propagate along the multimode region. This causes the modes to interfere destructively, but at certain propagation distances a mode-beating phenomenon occurs, where the modes constructively interfere to partially or fully reconstruct the original input field. The result is that at certain propagation lengths in the multimode region, single or multiple self-images of the input field are formed. The device, by virtue of the interference between the multiple modes, is called a *multimode interference* coupler, or simply, an MMI coupler [11].
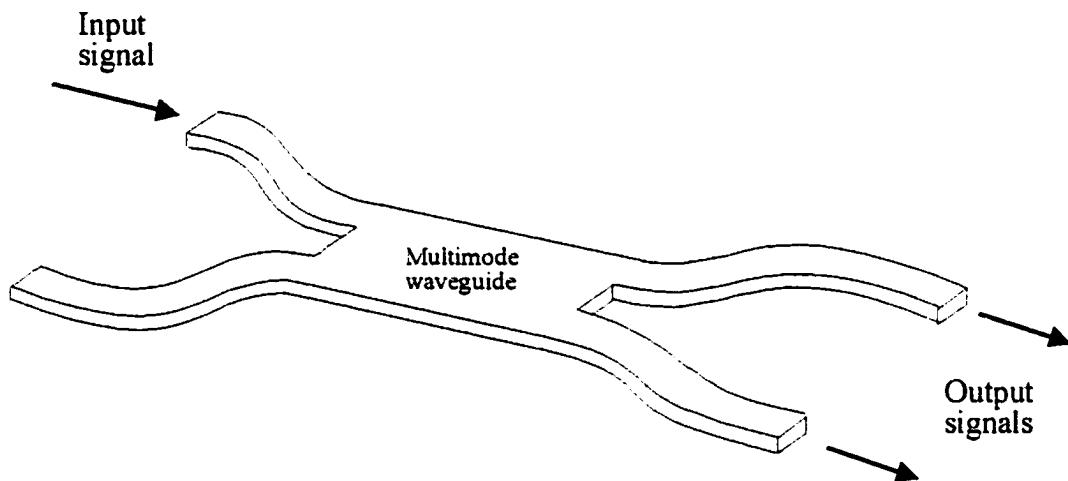
Figure 1.3 A schematic of the multimode interference coupler. Joining the waveguides in the interaction region forms a wider region within which a self-imaging phenomenon occurs.

Depending on the length of the multimode region, the two-port MMI coupler shown in Fig. 1.3 can function in the bar, cross, or 3 dB states, as in the case of the directional coupler. The advantages of the MMI coupler over the integrated directional coupler, however, are numerous: it is more compact, simpler to design, more fabrication tolerant, less polarization and wavelength sensitive, it can accommodate a large number of input/output waveguides in a small area, and it has favourable power and phase transfer properties [12].

Of particular importance in optical network applications are the scaleable nature of the MMI coupler and the phases of the self-images in the 3 dB (equal power splitting) state. In this state, the two input/two output device in Fig. 1.3 is termed a 2×2 MMI coupler, and can be generalized in a straightforward manner to an $N \times N$ MMI coupler.

1.5 Applications

The use of two $N \times N$ MMI couplers in a serial configuration forms the basis of a device called the multiple-arm or generalized Mach-Zehnder interferometer (GMZI). This device was first proposed by Ulrich [13], and is useful because it represents a simple and compact method of

realizing an integrated optical 1 x N switch that can be controlled by applying, for instance, electrical signals to optical waveguides.

In the present work the GMZI structure is investigated in detail, and mathematical models are developed to predict its behaviour both in theory and in practice. This has led to the development of efficient integrated optical switches and power splitters with functions expected to find various uses in the emerging generation of optical networks. Three possible applications of these devices are shown in Fig. 1.4.

(a)

(b)

(c)
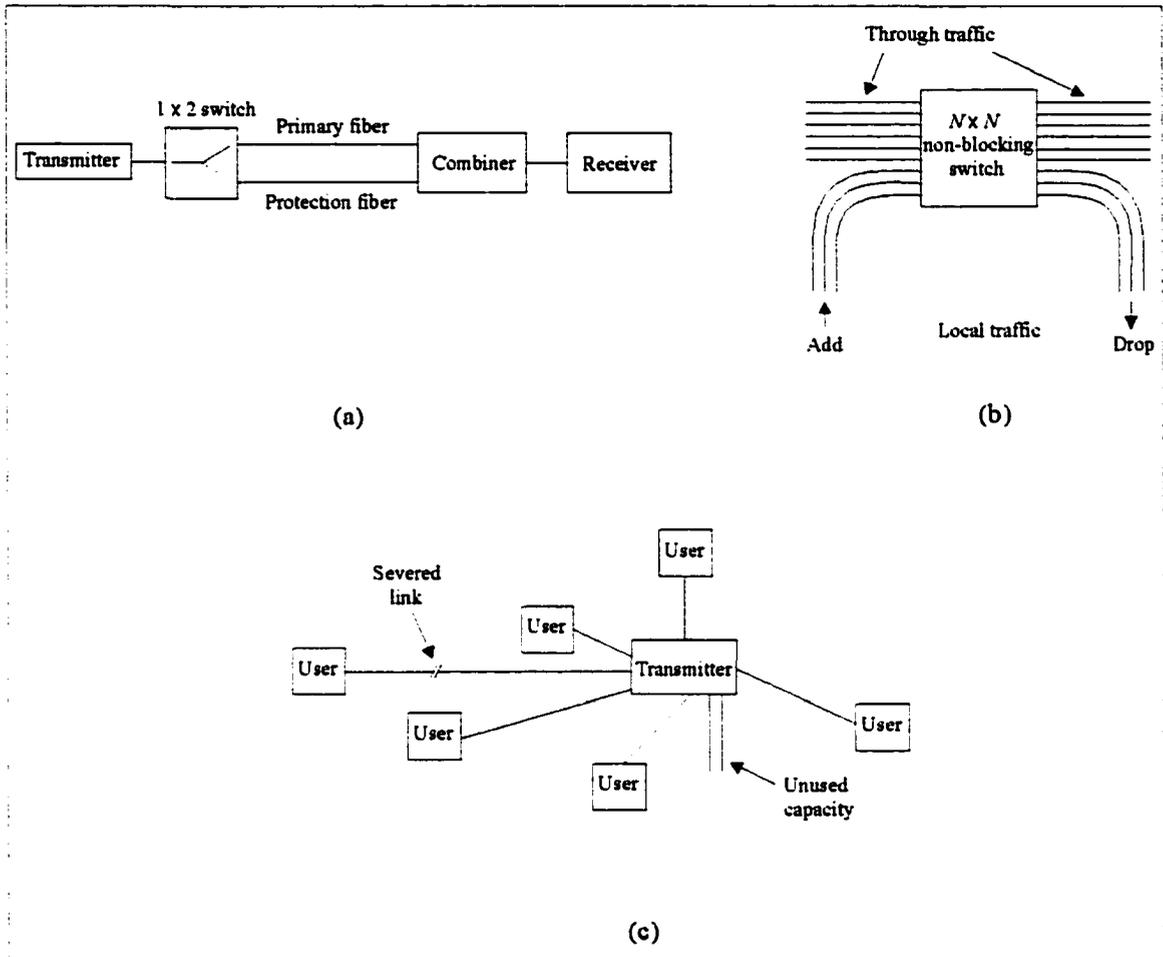
Figure 1.4 Optical network applications of the devices investigated in this work, indicating (a) optical protection switching, (b) optical cross-connect switching, and (c) dynamic variable optical power distribution.

In Fig. 1.4(a), 1 x N switches are useful in a protection scenario when a primary optical fiber link is severed. An optical cross-connect is an N x N non-blocking optical switch, and allows a traffic stream to be modified at a network node as pictured in Fig. 1.4(b). Such switches can also be composed of smaller switch units such as 1 x N switches [14]. The N x N switch is set in a state such that desired data stream(s) are added and dropped at a given node. Finally, a new device developed in this work allows the functionality pictured in Fig. 1.4 (c). In this scenario, an optical signal is distributed to various end users located at different distances from a transmitter. The optical power allocation to each user is based on a number of factors such as distance, user demand for services, and the presence of severed or unused links, and can be re-allocated as desired.

1.6 Overview of this work

Much of this work is devoted to the study of the integrated optical GMZI configuration using $N \times N$ MMI couplers. As a logical starting point for the discussion, the fundamental principles of the integrated optical devices used in this work are presented in Chapter 2.

This forms a solid theoretical foundation for a mathematical description of the general properties of the GMZI, which is presented in Chapter 3. This has led to the interpretation of the device as a generalized or variable-ratio power splitter, the theoretical characteristics of which are discussed in detail. In a specific case of variable-ratio power splitting, the GMZI functions as the previously known optical switch. The switching properties of the GMZI are investigated in Chapter 4, and lead to several new designs for $1 \times N$ and $N \times N$ switches.

Demonstration of the validity of the theory governing the devices and the practicality of their implementation is achieved through the design, realization, and performance measurement of the integrated optical GMZI. The design of a set of prototype integrated optical devices is described in Chapter 5. Much of this work was performed at the University of Alberta and the Telecommunications Research Laboratories, in Edmonton, Alberta.

Chapter 6 discusses the fabrication of the integrated devices which was performed during a period as a visiting scientist at the MESA Research Institute, University of Twente, in The Netherlands. Some initial device characterization has been performed as a student intern at JDS Uniphase Inc., in Nepean, Ontario.

The performance of the fabricated devices has been characterized during a second period as a visiting scientist at the University of Twente, using the facilities of the Lightwave Devices Group. The measurement results are presented in Chapter 7.

Finally, a general summary, future prospects for this work, and concluding remarks are given in Chapter 8.

Chapter 2

Theoretical foundations

This chapter provides a description of the fundamental physical principles upon which the integrated optical components in this work are based. The intent is not to reproduce a rigorous and detailed mathematical description of the underlying theory developed by others; for this the interested reader may refer to the references cited herein. The emphasis of this chapter will instead be to give the necessary mathematical and physical arguments to form an intuitive picture of light propagation in optical waveguides, the phenomenon of multimode interference, and the generalized Mach-Zehnder interferometer structure. This will serve to provide a proper foundation for an understanding of the concepts described in subsequent chapters.

2.1 Introduction

We begin the discussion by considering the propagation of light in a medium. Since light is a form of electromagnetic radiation, Maxwell's equations are the logical starting point in the analysis. In integrated optics, light is typically guided in media with structural dimensions of the order of the wavelength of light used. An application of Maxwell's equations leads to the result that light becomes quantized into a set of allowed spatially-invariant distributions, or modes, which can propagate through a light-guiding structure. The simplest such structure is the dielectric slab waveguide. An understanding of the slab waveguide permits the analysis of more complicated waveguide structures, as well as providing insights into the self-imaging phenomenon exhibited by multimode interference (MMI) couplers. The properties of the MMI coupler in turn enable a description of the operation of the generalized Mach-Zehnder interferometer (GMZI) configuration, which is at the heart of this work. We therefore turn our attention first to the slab waveguide.

2.2 The slab waveguide

A three-layer asymmetric ($n_1 \neq n_3$, in general) dielectric slab waveguide is shown schematically in Fig. 2.1. It consists of a core layer (n_2) with a higher refractive index than that of the surrounding cladding layers (n_1, n_3). Light propagates in the $+z$ direction and in the plane of n_2, confined by total internal reflection at the core-cladding interfaces. For the analysis of this structure, the $x = 0$ origin is chosen at the $n_2 - n_3$ interface and it is assumed without loss of generality that $n_3 \geq n_1$.
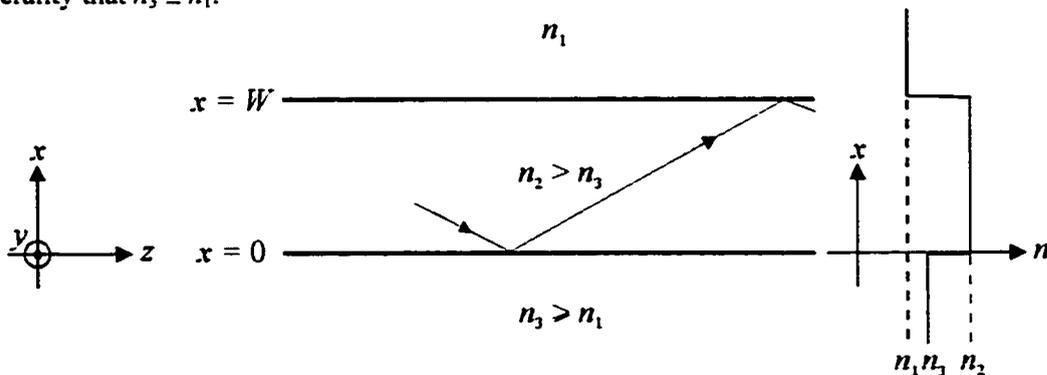
Figure 2.1 The asymmetric three-layer dielectric slab waveguide with the refractive index profile shown on the right.

Since light can be described as an electromagnetic wave, it is composed of complex-valued electric and magnetic fields of the form $f(x, y, z, t)$ which vary in the three spatial coordinates (f_x, f_y, f_z) and with time t. In linear, homogeneous, isotropic media, these fields satisfy the electromagnetic wave equation [15]

$$\nabla^2 f = \frac{n^2}{c^2}\frac{\partial^2 f}{\partial t^2} \tag{2.1}$$

where n is the refractive index of the medium and c is the speed of light in free space. The light in the slab can be described by fields of the form

$$f(x,y,z,t) = F(x) \cdot e^{J(\omega t - \beta z)} \tag{2.2}$$

where $J = \sqrt{-1}$, ω is the optical frequency, β is the propagation constant, and $F(x)$ represents the (real-valued) field amplitude distribution, or mode. Note that y-independence of the field is assumed due to the invariance of the slab in the y direction. Substitution of (2.2) into (2.1) yields the wave equation in each slab layer

$$\frac{d^2 F}{dx^2} + \left(k_0 n_l^2 - \beta^2\right) F = 0 \tag{2.3}$$

where $k_0 (= 2\pi/\lambda)$ is the free space wavenumber, λ is the free space optical wavelength, and n_l is the refractive index of the l-th layer. The set of solutions $F(x)$ to the wave equation (2.3) are the modes of the slab waveguide. For the slab shown in Fig. 2.1, the propagation constants of the guided modes will lie in the range

$$k_0 n_3 < \beta < k_0 n_2 \tag{2.4}$$

so from (2.3) it is evident that the modes will be sinusoidal in the core layer and exponential in the cladding layers. Radiation (unguided) modes also exist as solutions to (2.3); however, these are strongly attenuated in the z direction and are therefore omitted from the analysis.

The electric $(f \equiv E)$ and magnetic $(f \equiv H)$ field components which contribute to the guided modes must now be determined. Substituting (2.2) into the two Maxwell curl equations yields relations between the six spatial components of E and H:

$$H_x = -\frac{\beta}{\omega\mu_0} E_y \tag{2.5a}$$

$$H_y = -\frac{J}{\omega\mu_0}\left(\frac{dE_z}{dx} + J\beta E_x\right) \tag{2.5b}$$

$$H_z = \frac{J}{\omega\mu_0}\frac{dE_y}{dx} \tag{2.5c}$$

$$E_x = \frac{\beta}{\omega\varepsilon_0 n^2} H_y \tag{2.5d}$$

$$E_y = \frac{J}{\omega\varepsilon_0 n^2}\left(\frac{dH_z}{dx} + J\beta H_x\right) \tag{2.5e}$$

$$E_z = -\frac{J}{\omega\varepsilon_0 n^2}\frac{dH_y}{dx} \tag{2.5f}$$

where μ_0 and ε_0 are the free space permeability and permittivity, respectively. Note that (2.5a, c, e) relate the set of field components $\{E_y, H_z, H_x\}$ while (2.5b, d, f) relate an independent set of field components $\{H_y, E_z, E_x\}$. These sets are designated transverse electric or TE, and transverse magnetic or TM, polarizations, respectively, since the electric or magnetic fields are orthogonal to the plane of propagation (x-z plane in Fig. 2.1).

The solutions $F(x)$ of (2.3) can therefore be represented simply by $E_y(x)$ for the TE case and $H_y(x)$ for the TM case. The solutions are found by solving (2.3) subject to the boundary conditions from electromagnetic theory, which require tangential components of F to be continuous across a dielectric discontinuity. From Fig. 2.1 this implies for TE fields that E_y and H_z must be continuous, and for TM fields that H_y and E_z must be continuous. From (2.5c, f) this is equivalent to

$$F_y \quad \text{and} \quad \sigma_l \frac{dF_y}{dx} \quad \text{continuous} \tag{2.6}$$

across dielectric boundaries, where $\sigma_l = 1$ for TE and $\sigma_l = 1/n_l^2$ for TM. For the slab in Fig. 2.1 the wave equation (2.3) can then be written as

upper cladding $(x \geq W)$:
$$\frac{d^2 F_y}{dx^2} - q^2 F_y = 0 \tag{2.7a}$$

core$(0 \leq x \leq W)$:
$$\frac{d^2 F_y}{dx^2} + h^2 F_y = 0 \tag{2.7b}$$

lower cladding $(x \leq 0)$:
$$\frac{d^2 F_y}{dx^2} - p^2 F_y = 0 \tag{2.7c}$$

Comparing (2.3) and (2.7), the mode parameters q, h, and p are defined as

$$q^2 = \beta^2 - k_0^2 n_1^2 \tag{2.8a}$$

$$h^2 = k_0^2 n_2^2 - \beta^2 \tag{2.8b}$$

$$p^2 = \beta^2 - k_0^2 n_3^2 \tag{2.8c}$$

where (2.4) has been utilized to ensure that h, q, and p are real.

The solutions to (2.7) are the bound waveguide modes, written as

upper cladding $(x \geq W)$: $F(x) = A_1 e^{-qx}$ (2.9a)

core$(0 \leq x \leq W)$: $F(x) = \cos(hx) + A_2 \sin(hx)$ (2.9b)

lower cladding $(x \leq 0)$: $F(x) = e^{px}$ (2.9c)

where the constants A_1 and A_2 are found by applying the boundary conditions (2.6) to the fields in (2.9):

$$A_2 = \frac{\sigma_3 p}{\sigma_2 h} \tag{2.10a}$$

$$A_1 = e^{qW}\left[\cos(hW) + A_2 \sin(hW)\right] \tag{2.10b}$$

Applying the boundary conditions (2.6) to the $x = W$ boundary in (2.9) results in the following dispersion relation

$$\tan(hW) = \frac{\sigma_2 h(\sigma_3 p + \sigma_1 q)}{(\sigma_2 h)^2 - \sigma_1 \sigma_3 pq} \tag{2.11}$$

which is a transcendental equation for β, since q, h, and p are functions of β in (2.8). This can be solved using numerical or graphical methods. Since the tangent function is periodic, a number of solutions may exist, each value of β corresponding to a bound mode of the slab waveguide. The order of the mode is denoted by $v = 0, 1, 2, \ldots$, and the parameters β_v, q_v, h_v, and p_v determine the field F_v of the v-th mode in (2.9). As we shall see later, it is also useful to characterize each mode in terms of an effective refractive index $n_{eff.v}$ defined as

$$n_{eff.v} = \frac{\beta_v}{k_0} \tag{2.12}$$

The number of guided modes is limited by the relation (2.4), and increases as the core width W increases, as the wavelength decreases, and as the core-cladding refractive index difference

(contrast) increases [16]. From (2.4) and (2.12) it is apparent that the effective refractive indices of guided modes fall between the core and cladding values.

The mode fields for a symmetric ($n_1 = n_3$) multiple-mode (or multimode) slab waveguide are shown in Fig. 2.2. The v-th order mode field has v nulls and $v + 1$ extrema. Note that the fields are not confined strictly to the core layer, since the exponential tails penetrate into the surrounding cladding layers. The penetration depth ($1/q_v$) increases with increasing order v, however the modes can be assumed to be confined within an effective width W_e. The effective width is wider than the actual width W, and is approximated by the effective width of the fundamental ($v = 0$) mode [17]

$$W_e \cong W + \frac{\lambda}{\pi}\frac{\sigma_2}{\sigma_1}\left(n_2^2 - n_1^2\right)^{-1/2}$$

(2.13)

and approaches the actual width W as the refractive index contrast increases.
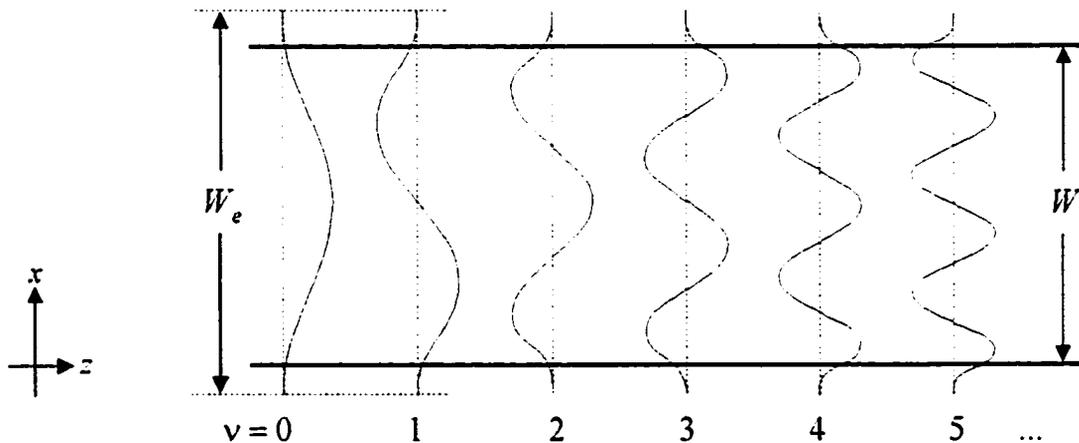
Figure 2.2 Mode field patterns $F_v(x)$ for the first six bound modes of the symmetric slab waveguide. The modes are assumed to be confined within an effective width W_e.

The three-layer asymmetric slab waveguide illustrates the concept of guided modes of propagation, however, strictly speaking, it is an idealization. The layers are y-z planes, infinite in extent, and the lower and upper boundaries of the cladding layers extend to $\pm \infty$. We now consider removing these restrictions in order to gain an understanding of the behavior of real waveguides. In practice, the core and cladding layers are thin films of material deposited on a substrate, usually a wafer. The lower cladding layer has a finite thickness, since it is bounded from below by a substrate material. The upper cladding also has a finite thickness, and is usually bounded from above by air. Finally, the slab waveguide is in general a poor conduit for light, since the light is confined only in the x direction and can propagate freely in the y-z plane. Practical waveguides consist of channels to guide light by means of two-dimensional confinement, analogous to the manner in which electrons are guided by the metal strip lines of integrated electronic circuitry. In this section we discuss the modifications to the basic slab theory necessary to describe an air cladding and semiconductor substrate, and in the following section we discuss the channel waveguide.

2.2.1 Air cladding

To include the effect of a top air cladding layer, the asymmetric slab waveguide theory must be extended to four layers. This is done in a straightforward manner, with sets of four equations replacing (2.7) – (2.9). The fields are sinusoidal in the core and exponential in the cladding and air regions. Applying the boundary conditions to the fields yields three systems of

two equations, which can be reduced to a single system of two equations through back substitution [16]. The resulting dispersion relation can be solved for the β_v and hence the effective indices of the modes. In general, as the thickness of the top cladding layer increases, the air cladding has a diminishing effect on the waveguide mode.

2.2.2 Substrate effects

Integrated optical devices are fabricated on various substrate materials, and in a number of cases the refractive index of the substrate is higher than the indices of the core/cladding films. Common examples are AlGaAs films on GaAs substrates and doped SiO_2 and SiO_xN_y films on Si substrates. For waveguiding to occur, the core layer must have a refractive index higher than that of the surrounding layers. The presence of a high index substrate may result in the leakage of light from the waveguide core into the substrate, attenuating the field as it propagates. A sufficiently thick lower cladding, however, can serve as an effective isolation layer and prevent undesired substrate attenuation.

To determine the required minimum thickness of the isolation layer, the treatment of the slab waveguide is first extended to a four-layer system, where the fourth layer is the substrate. The high index substrate changes β for the waveguide mode from its value in the three-layer slab, and using a perturbation approach [16] this change can be accurately estimated. In this approach, the propagation constant is separated into two components, the (real) propagation constant β_0 for the three-layer structure, plus a small change $\delta\beta$,

$$\beta = \beta_0 + \delta\beta \qquad (2.14)$$

where $\delta\beta$ is complex. This form for β is substituted into the dispersion relation of the four-layer structure, and keeping terms to first order in $\delta\beta$, an expression for $\delta\beta$ can be derived. The real part of $\delta\beta$ contributes to a change in the effective index, and the imaginary part contributes to an attenuation as the mode propagates in the z direction, as can be seen from (2.2). The minimum thickness for the isolation layer is the thickness for which the attenuation of the guided modes is negligible, i.e., less than the expected material attenuation in the core/cladding layers.

2.3 Channel waveguides - the effective index method

To efficiently guide light in integrated optics, it is necessary to use channel waveguides which confine light in the two directions (x and y) orthogonal to the propagation direction. Light can then be directed on a planar surface by simply defining the path of the waveguide channel. Two-dimensional field confinement is achieved through the use of any of a number of possible waveguide geometries, such as ridge, rib, buried channel, and strip loaded [16], which are realized through a combination of planar (slab) waveguide technology and lithographic techniques.

The solution of Maxwell's equations for the two-dimensional field profile in channel waveguides, however, does not have a simple, closed form solution. Numerical discretization schemes are generally used to solve for the modes; however, a simpler and more physically intuitive technique known as the effective index method (EIM) can also be used [18] – [20]. Values of $n_{eff,v}$ calculated using the EIM are applicable for weakly and strongly confining channel structures [21] and are reasonably accurate if they are not too close to the cladding indices, i.e., for well-guided modes.

To implement the EIM, it is necessary to recognize that the modes of the channel waveguide are no longer strictly TE or TM. The modes are hybrid, which means that the field components can no longer be grouped into two independent sets, so in general six field components are needed to describe each mode. A simplifying assumption, however, can be made. For core widths greater than their thicknesses, certain field components dominate such that the modes are similar to TE and TM. These are referred to as quasi-TE and quasi-TM modes. In

the following analysis, these modes will be referenced to the direction of the dominant component of the electric field (E_y for TE and E_x for TM), and we will continue to use the terms TE and TM, recognizing, however, that the modes are hybrid.

Consider a channel waveguide as shown in Fig. 2.3a, and the TM polarization. The EIM is implemented in two steps. In the first step, the ridge is divided into three slab regions. The asymmetric slab analysis is then used to find the effective index of the TM mode for the slab in the channel region, n_{ch}, and the slab in the cladding region, n_{cl}. These effective indices are then considered to form a slab in the horizontal direction as shown in Fig. 2.3b, with a core layer thickness equal to the channel width W. In the second step, a symmetric slab analysis is performed on the structure in Fig. 2.3b, using the effective indices n_{ch}, n_{cl}, and a TE polarization, since the dominant field component E_x is in a direction parallel to the slab interfaces. The result is a single effective refractive index n_e for the channel waveguide structure for each mode. The analysis for TE polarized light follows in an analogous manner.

(a)   (b)
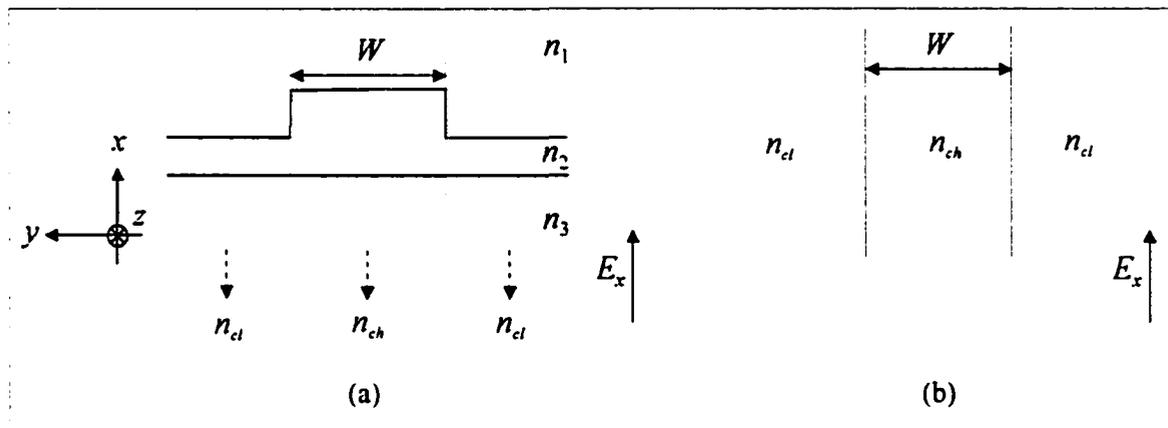
Figure 2.3 The effective index method, shown here for a TM polarization. In the first step (a) the waveguide is divided into three slab regions with TM effective indices n_{ch} and n_{cl}, which in the second step (b) form a slab in the horizontal direction, which must be analyzed for a TE polarization, to arrive at a single TM effective index n_e for each mode of the entire channel waveguide.

2.4 Modal propagation analysis

In a channel waveguide that supports only a single mode, the light propagates, by definition, without variation of the spatial mode field. This property is also true of each of the modes of the slab waveguide; however, at any point along the slab, the total field is a superposition of the individual mode fields, taking into account the phases of the modes, which in general differ. Interference therefore occurs between the modes, and results in a spatially-varying total field in the multimode waveguide. We now describe this interference quantitatively, as it forms the basis of self-imaging in MMI couplers.

For light to be guided by the slab waveguide, it is not necessary for the light to have the distribution of one of the modes. The mode fields $F_v(x)$ form an orthogonal set of solutions to the wave equation (2.3), and normalizing these fields,

$$\hat{F}_v(x) = \frac{F_v(x)}{\sqrt{\int_{-\infty}^{\infty} |F_v(x)|^2 \, dx}}$$ (2.15)

an orthonormal set of mode functions $\hat{F}_\nu(x)$ is obtained. An arbitrary input field $T(x, z_0)$ launched into the slab waveguide at the point $z = z_0$ can be written as a superposition of these mode functions,

$$T(x, z_0) = \sum_{v=0}^{m-1} a_v \hat{F}_v(x) \tag{2.16}$$

where m is the number of guided modes and a_v is the complex amplitude coefficient of the v-th mode, given by

$$a_v = \int_{-\infty}^{\infty} \hat{F}_v^*(x) \, T(x, z_0) \, dx \tag{2.17}$$

and $\hat{F}_v^*(x)$ is the complex conjugate of $\hat{F}_v(x)$. From (2.16), it is seen that the optical power is divided among the modes. In multimode waveguides, the number of guided modes is finite. Therefore, depending on the form of $T(x, z_0)$, some of the power at the input may not be coupled to a guided mode. Power not coupled to propagation modes is radiated at the transition into the multimode region. Usually, however, the radiated power is negligible and it is therefore ignored in the following analysis.

Since the different modes travel with different propagation constants β_v, the input field $T(x, z_0)$ changes its transverse distribution as it travels through the waveguide in the z direction. The field at any longitudinal point z can be expressed as

$$T(x, z) = \sum_{v=0}^{m-1} a_v \hat{F}_v(x) \, e^{-j\beta_v z} \tag{2.18}$$

where time dependence is implicit. This method of input field decomposition into a superposition of guided modes and the subsequent propagation of these modes is referred to as modal propagation analysis (MPA). Alternatively, various numerical beam propagation methods (BPM) can be used to accurately simulate light propagation through waveguide structures, but these methods do not provide physical insights into the propagation mechanisms. While MPA is a simpler and more physical approach, it is applicable only where the refractive indices or geometry in the x-y plane do not change significantly with z. Since the MMI coupler is essentially a rectangular multimode waveguide region with a spatially-invariant structure, MPA is well-suited to the analysis of the MMI coupler and serves as an excellent tool for the explanation of its behavior.

2.5 The $N \times N$ multimode interference coupler

An $N \times N$ multimode interference (MMI) coupler consists of a multimode waveguide with N single-mode waveguides at both the input and output planes, shown in Fig. 2.4. It operates upon the principle of self-imaging, defined as "a property of multimode waveguides by which an input field profile is reproduced in single or multiple images at periodic intervals along the propagation direction of the guide" [22].

Output guides

Multimode
Interference
Coupler

$z = L_{M,N}$

Input guides

$z = 0$

x y z
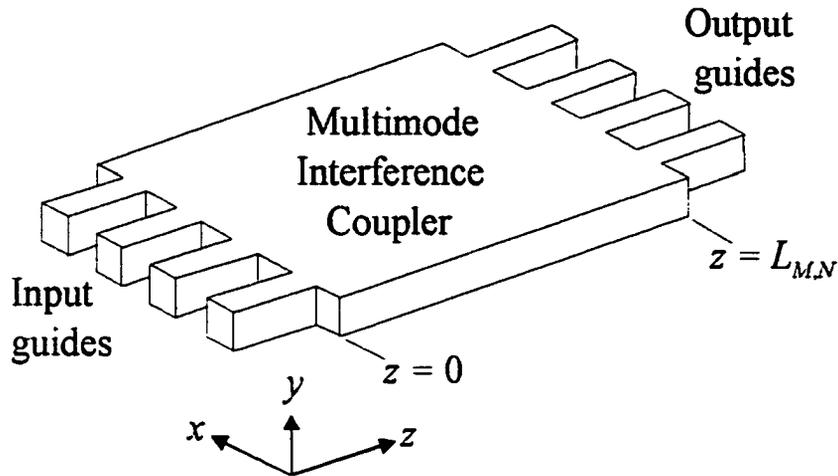
Figure 2.4 Layout of the $N \times N$ MMI coupler in a raised-channel (rib) waveguide geometry. The case $N = 4$ is shown.

The $N \times N$ MMI coupler is actually part of a much wider class of MMI couplers which include $1 \times N$ and $2 \times N$ MMI couplers [22], self-imaging based on different types of interference [11], couplers with a non-rectangular shape [23], [24], and couplers capable of non-uniform power splitting [23], [25]. In this work we restrict our attention to the $N \times N$ MMI coupler due to its ease of incorporation into a generalized Mach-Zehnder interferometer (GMZI) structure and favourable imaging properties, as will be described in the following chapter. It should be noted, however, that many of the concepts dealt with in this work can be adapted to allow the use of other types of MMI couplers.

To clarify the origins of self-imaging, the following physical interpretation can be offered. Consider an optical field at one of the inputs of the multimode waveguide shown in Fig. 2.4. It was shown in the previous section that this input field is expressed in the multimode waveguide as a sum of weighted mode field amplitudes, each with an associated phase factor $\beta_v z$. At the beginning of the multimode region, the modes are 'in phase' since their addition results in a representation of the input field. As the modes propagate, they move 'out of phase' by virtue of their differing propagation constants, and the sum of the modes at a point $z > z_0$ in general no longer results in a faithful representation of the input. The mode phases, however, are periodic and at certain distances will 'beat', meaning that the accumulated phase dispersion between the modes will equal an integer multiple of π. At these distances some or all of the modes will again come together 'in phase', with their addition resulting in a reconstruction of the input field in the form of single or multiple self-images.

To exploit the self-imaging principle, the distances at which mode beating occurs must be determined. This will depend upon the relationship between the propagation constants of the various modes, so we first turn our attention to determining expressions for the β_v. This, in turn, will enable a mathematical description of the self-imaging principle and some simple formulae for designing MMI couplers. Again, the emphasis of the discussion will be to highlight the physical phenomena rather than to provide rigorous derivations of the results.

Rearranging (2.8b), we find

$$\beta_v^2 = k_0^2 n_{ch}^2 - h_v^2 \tag{2.19}$$

where n_{ch} is the equivalent refractive index of the channel waveguide region calculated in the first step of the EIM, and replaces n_2 which is valid only for the infinite slab waveguide. In the slab waveguide analysis, it was shown that the h_v are obtained by solving (2.11) either graphically or numerically for the β_v, followed by substitution into (2.8b). To obtain analytic expressions that

will greatly simplify our analysis, we consider a 'strongly guiding' approximation, wherein the mode fields in the multimode region are confined to the channel layer with negligible penetration into the cladding layers. This is achieved in practice by using deeply etched rib waveguides to provide a high refractive index contrast, thereby confining the mode field within the bounds of the rib. An integer number of lobes are then present within the effective width W_e of the channel (refer to Fig. 2.2), and the mode fields are well described by functions of the form

$$F_v(x) \cong \sin\left[\frac{\pi(v+1)}{W_e}x\right]$$
(2.20)

where the approximation sign reflects the strongly guiding approximation used. Comparing (2.20) to (2.9b), the parameter h_v can be written as

$$h_v \cong \pi(v+1)/W_e$$
(2.21)

Substituting (2.21) into (2.19) and using the paraxial approximation $h_v^2 \ll k_0^2 n_{ch}^2$, the binomial expansion can be used to express β_v as

$$\beta_v \cong k_0 n_{ch} - \frac{\pi\lambda(v+1)^2}{4n_{ch}W_e^2}$$
(2.22)

indicating that the spacing of the propagation constants increases quadratically as the mode number increases. Self-imaging depends upon the length at which all the modes beat, which will be governed by the two modes which require the longest propagation length in order to beat. These are the modes with the most closely spaced propagation constants, which in view of (2.22) are the fundamental and first order modes. A characteristic beat length L_π for a multimode waveguide can therefore be defined as [11]

$$L_\pi \equiv \frac{\pi}{\beta_0 - \beta_1} \cong \frac{4n_{ch}W_e^2}{3\lambda}$$
(2.23)

where (2.22) has been used to arrive at the second expression.

The field at a distance z along the multimode region given can now be expressed in terms of this beat length by first factoring out a constant phase factor $e^{-j\beta_0 z}$ from (2.18) and substitution of (2.22) and (2.23) to arrive at

$$T(x,z) = \sum_{v=0}^{m-1} a_v \hat{F}_v(x) e^{\left[\frac{\pi v(v+2)}{3L_\pi}z\right]}$$
(2.24)

At distances $z = L_{M,N}$ where

$$L_{M,N} = \frac{M}{N}3L_\pi$$
(2.25)

with M and N positive integers with no common divisor, the argument of the phase term in (2.24) becomes an integer multiple of π. It can be shown that in these cases the field in the multimode region will be of the form [12]

$$T(x, L_{M,N}) = \frac{1}{C}\sum_{b=0}^{N-1} T_{ext}(x - x_b) e^{j\varphi_b}$$
(2.26)

where $T_{ext}(x)$ describes the field profile at the input of the multimode region, C is a complex normalization constant, and x_b and φ_b describe the positions and phases, respectively, of N self-images at the output of the MMI coupler. The self-images are positioned with a constant spacing between them. The parameter M describes a multiple of the imaging length that also results in N self-images, since the self-imaging phenomenon is periodic. Usually the shortest devices are obtained when $M = 1$.

The form of (2.26) is derived by noting that an arbitrary input field $T(x,0)$ can be represented by a Fourier sine series, provided that the input field and the functions (2.20) are extended antisymmetrically by a distance W_e in the $-x$ direction to form a virtual MMI section with a periodicity of $2W_e$. It can be verified from Fig.2.2 that this results in antisymmetric mode fields. The input field is unchanged in the real MMI section, however, it is denoted by $T_{ext}(x)$ to reflect the inclusion of the antisymmetric extension to the virtual MMI section.

The ability of the MMI coupler to form high-resolution self-images of the input field is related to the Fourier series representation of the input field in the multimode region. As the number of terms in the Fourier series increases, the self-image fields are reconstructed more accurately. The resolution of the self-images therefore improves as the number of modes in the multimode region increases. Another way of stating this is that since the self-images are formed from a linear combination of the guided mode fields, the narrowest image attainable in a multimode waveguide is roughly given by the lobe width of the highest supported mode (i.e., the spatial half-period, seen in Fig. 2.2) [22].

From (2.26), it is seen that the field at a distance $L_{M,N}$ in the multimode region is a sum of N fields $T_{ext}(x)$, each identical to the input field and spatially shifted by a different amount. The result is N self-images with locations corresponding to the positions of the output single mode waveguides in Fig. 2.4. The geometric layout of the $N \times N$ MMI coupler is shown in Fig. 2.5, and utilizes a free design parameter a that can take on any value in the range $0 < a < W_e / N$ [12]. For simplicity, the value of $a = W_e / 2N$ is usually chosen and results in equally spaced input and output waveguides (ports).

It is also noted in (2.26) that each self-image has an associated phase φ_b, which in general differs for each image. These phases have been derived by Bachmann et al., [26], and can be expressed more conveniently as [27]

$$\varphi_{ij} = \phi_1 - \frac{\pi}{2}(-1)^{i+j+N} + \frac{\pi}{4N}\left[i + j - i^2 - j^2 + (-1)^{i+j+N}\left(2ij - i - j + \frac{1}{2}\right)\right] \qquad (2.28)$$

where ϕ_1 is a constant phase given by

$$\phi_1 = -\beta_0 \frac{3L_\pi}{N} - \frac{9\pi}{8N} + \frac{3\pi}{4} \qquad (2.29)$$

and i and j represent the input and output port number, respectively, for the MMI coupler as shown in Fig. 2.5. φ_{ij} is interpreted as the phase of the image at output port j with respect to the phase of the input at port i. It will be seen in the next section that the differing phases of the self-images constitute an important property of the MMI coupler when it is used to form Mach-Zehnder type structures.

MMI couplers have experimentally been shown to function as efficient splitters and combiners of optical beams, while possessing a host of advantageous characteristics, such as low excess losses and low crosstalk [11], [28], accurate splitting ratios [16],[29], small device sizes [30], and good fabrication tolerances [31]. Furthermore, it has been shown that these features can be maintained independent of polarization, and for operation over a broad range of wavelengths and temperatures [28], [31].
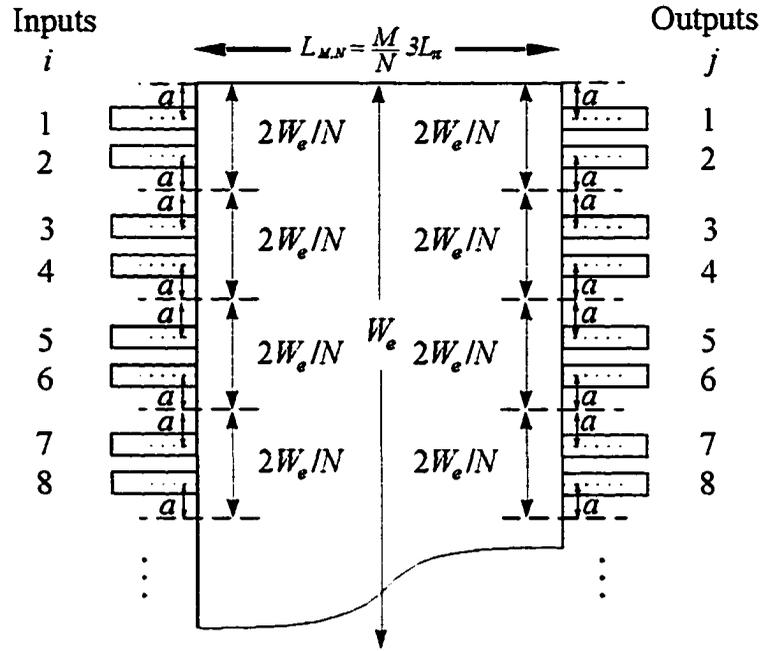
Figure 2.5 Geometry of the $N \times N$ MMI coupler, showing the positions and numbering of the input and output single mode access waveguides.

2.6 The $N \times N$ generalized Mach-Zehnder interferometer

To see how the phase properties of the $N \times N$ MMI coupler can be useful, consider as an example the operation of the 4×4 MMI coupler shown in Fig. 2.6a.
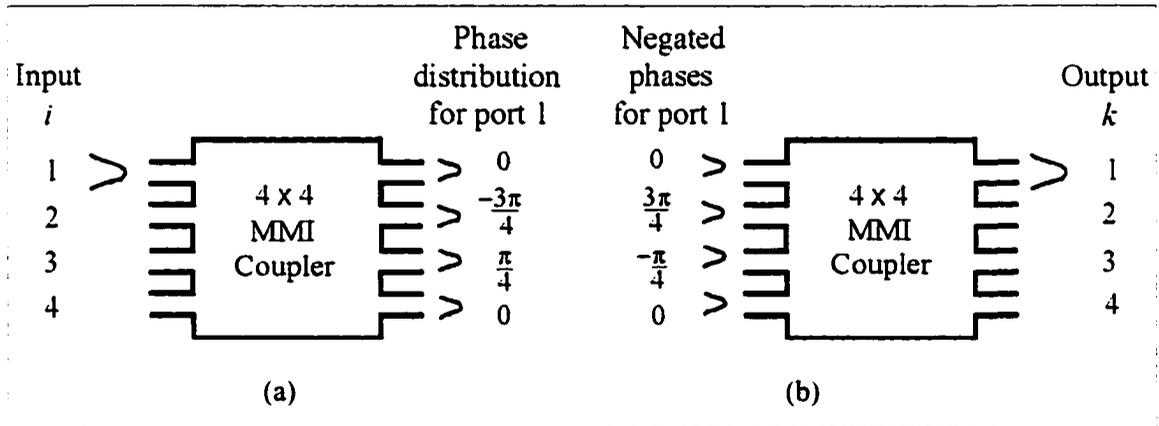
Figure 2.6 A 4×4 MMI coupler acting as (a) a 1×4 splitter, and (b) a combiner.

An optical field entering the MMI coupler from input port $i = 1$ (arbitrarily chosen) results in four self-image fields at the outputs with relative phases as shown, calculated using (2.28). Note that each input port imparts its own characteristic relative phase distribution to the self-images, and that the distribution differs in general for different input ports.

The MMI coupler acts as a passive uniform power splitter, but it is interesting to note that it also works in reverse, performing a combining operation. If the propagation of the self-image fields is retraced as the self-images propagate back through the coupler region, they will interfere

to re-establish the original input field. This principle of reciprocity applies generally to the propagation of light through an isotropic optical medium [8].

Now in order to route the input field back to the same port, instead of reversing the direction of the fields, which is a difficult function to achieve in integrated optics, consider the use of a second, identical MMI coupler, as shown in Fig. 2.6b. If an input of the first MMI coupler is to be routed to the same output port number of the second MMI coupler, the sign of the phases of the self-images must be reversed before they enter the second coupler. Reversing the sign results ensures the cancellation of the accumulated phase dispersion among the modes as they travel back through the multimode region, so that they reconstruct the original input field.

Similarly, to route a general input port i of the first MMI coupler to an output port k of the second MMI coupler, one need only adjust the phase distribution at the outputs of the first MMI coupler such that it is the negative of the phase distribution for port k. Through active adjustment of the phases, an input can be switched to any output port, thereby forming a $1 \times N$ switch. For example, if input 1 in Fig. 2.6a is to be routed to output 3 of the second MMI coupler in Fig. 2.6b, the phase distribution entering the second MMI coupler should be adjusted to be the negative of the phase distribution for port 3.

By connecting the output ports of the first coupler with the input ports of the second, waveguide arms j are formed and can be utilized as a region in which to actively shift the phases of the self-images. The result is the structure shown in Fig. 2.7, which is the multiple-arm, or generalized Mach-Zehnder interferometer configuration (GMZI). The structure is useful because it yields a simple and compact method of realizing an integrated optical $1 \times N$ switch, which can be controlled by applying, for instance, electrical signals in the arms to achieve the required phase shifts.
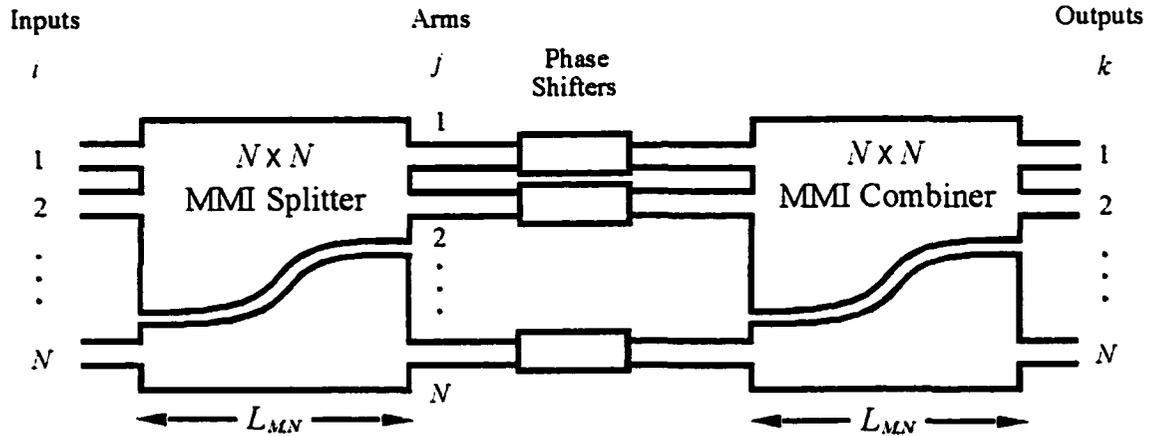
Figure 2.7 The generalized Mach-Zehnder interferometer (GMZI) structure.

Using this structure, 2×2 switches [32], a 1×4 switch [33], and a 1×10 switch [34] have all been realized. A full theoretical description of the GMZI, however, has not been put forth, so a principal aim of the present work is to analyze the GMZI structure in detail. The analysis begins in the following chapter with the development of a mathematical model with which to characterize GMZI structures. This model provides a fundamental insight into the nature of the GMZI, namely that it possesses much more general power splitting properties than previously thought. An understanding of these properties then enables the design of novel integrated optical routing and switching devices.

Chapter 3

Theory of variable-ratio power splitting

In this chapter transfer matrix techniques are used to gain a deeper understanding of the self-imaging properties in the generalized Mach-Zehnder interferometer (GMZI). In the analysis, expressions are derived and indicate that the GMZI is a generalized N-way power splitter. A method enabling this mode of operation is developed and results in a new integrated optic device based on the GMZI: the variable-ratio power splitter. The practical feasibility of realizing this device is assessed by modeling its operation in the presence of parameter variations and non-ideal effects stemming from technological issues. This is accomplished through the introduction of modifications into the basic transfer matrix theory to provide a model for the non-ideal GMZI. Portions of this chapter have been reported in Photonics Technology Letters [35]; *a paper to appear in the* Journal of Lightwave Technology [36]; *and in a* U.S. Patent Application [37].

3.1 Introduction

The operation of the GMZI is based primarily upon the phases of the self-images produced by the MMI coupler. Since the MMI coupler is capable of reproducing a field from any of the N inputs at all the N outputs, the phases of the self-images formed by an $N \times N$ coupler can be described by an $N \times N$ matrix of relative phases. This matrix is used to form a transfer matrix for the coupler that describes how it operates on an arbitrary input field.

The transfer matrix approach has in the past been applied to microwave directional couplers [38], [39], and has subsequently been used for integrated optical two-port directional couplers [40], [41]. Jenkins *et al.* [34] described the MMI coupler using a transfer matrix, but only gave the matrix for $N = 4$, and its use was limited to the determination of the phase shifts necessary for the switching function.

In this work, a simple and intuitive form for the transfer matrix of the $N \times N$ MMI coupler is developed. This matrix is then used to develop the transfer matrix for the GMZI. The GMZI transfer matrix enables the design of variable-ratio N-way active or passive power splitters. We therefore begin our discussion by describing the transfer matrix approach.

3.2 Transfer matrix of the $N \times N$ MMI coupler

Consider the $N \times N$ MMI coupler shown in Fig. 3.1. The coupler length given by (2.25) can be expressed as

$$L_{M,N} = \frac{M}{N} 3L_\pi \cong \frac{4 n_{ch} W_e^2 M}{\lambda N} \qquad (3.1)$$

where the approximation from (2.23) has been used for the beat length L_π.

Inputs   Outputs

i   *j*
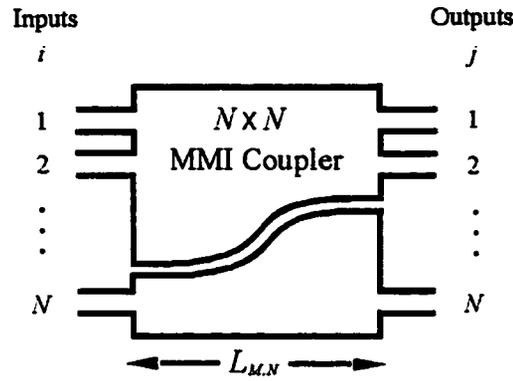
Figure 3.1 The $N \times N$ MMI coupler.

The phase associated with imaging an input i to an output j in an MMI coupler was given in (2.28). It contains two contributions – an absolute phase term and a relative phase term. In the present analysis only the relative phases at the coupler outputs are important, so the absolute term may be dropped without loss of generality to yield the relative phases which are again labeled φ_{ij}:

$$\varphi_{ij} = -\frac{\pi}{2}(-1)^{i+j+N} + \frac{\pi}{4N}\left[i+j-i^2-j^2+(-1)^{i+j+N}\left(2ij-i-j+\frac{1}{2}\right)\right] \qquad (3.2)$$

where the first subscript indicates the input port number and the second subscript indicates the output port number, corresponding to the numbering shown in Fig. 3.1. The symmetry relation $\varphi_{ij} = \varphi_{ji}$ can be found from (3.2), by simply interchanging i and j. Note that the φ_{ij} are not unique for a given N, since integer multiples of 2π can be added to any of them.

The φ_{ij} form a matrix Φ, with i representing the row number, and j representing the column number. By the symmetry relation, Φ is a symmetric matrix:

$$\Phi = \Phi^T \qquad (3.3)$$

where the superscript T indicates the transpose operation.

The transfer matrix for the MMI coupler is directly related to Φ, and is given by:

$$S_{coupler} = \left(\frac{1}{N}\right)^{\frac{M}{2}}\left(e^{J\Phi}\right)^M \qquad (3.4)$$

where $J=\sqrt{-1}$. Since Φ is symmetric, it follows that the transfer matrix is symmetric and unitary [42], i.e.,

$$S = S^T \quad \text{and} \quad S^\dagger S = S S^\dagger = I \qquad (3.5)$$

where † denotes the adjoint (complex conjugate matrix transpose) operation and I is the identity matrix of order N. The conditions in (3.5) are general properties of a transfer matrix S.

The output field distribution emerging from the MMI coupler is given by the following matrix equation

$$\begin{bmatrix} E_1^{out} \\ E_2^{out} \\ \vdots \\ E_N^{out} \end{bmatrix} = S_{coupler} \begin{bmatrix} E_1^{in} \\ E_2^{in} \\ \vdots \\ E_N^{in} \end{bmatrix} \qquad (3.6)$$

where E_i^{in} is the complex optical field amplitude at input port i and E_j^{out} is the optical field at output port j. The optical field intensity, which is the measurable quantity, is directly proportional to $|E|^2$.

3.3 The GMZI transfer matrix

3.3.1 Basic structure and operation of the GMZI

The basic layout of the $N \times N$ GMZI has been shown in Fig. 2.6. The structure consists of three components: an $N \times N$ MMI splitter, an active phase shifting region with N phase shifters, and an $N \times N$ MMI combiner. The MMI splitter forms N self-images of the input, which appear at the output ports, resulting in an equal distribution of the power to all arms of the phase shifting region.

The function of the active phase shifting elements is to modify the phases of the self-images such that a specific distribution of relative phases is obtained at the inputs to the MMI combiner. The active phase shifters are indicated by shaded regions on the internal arms, and can be operated by applying a driving signal (voltage, current, etc.) to shift the phases of the light in the arms by virtue of any of a number of effects, such as the electro-optic or the thermo-optic effect [43].

The MMI combiner performs the function of combining N inputs with equal intensities and different phases to form a specific power distribution at the outputs. The positions and intensities of the final output images depend upon the relative phases at the input to the MMI combiner. Therefore light from an input port i can be distributed in a controlled way to the output ports k simply by adjusting the phase shifters to achieve the required phase distribution at the combiner inputs.

3.3.2 The GMZI transfer matrix

The splitter is represented by a matrix $S_{splitter}$ given by (3.4) with elements s_{ij} and the combiner is represented an identical matrix $S_{combiner}$ with elements s_{jk}. The active phase shifts are applied independently and simply shift the phase φ_{ij} in a given arm j by an amount $\Delta\varphi_j$. The action of the phase shifters can therefore be described by a diagonal $N \times N$ transfer matrix [38]

$$S_{shift} = \begin{bmatrix} e^{j\Delta\varphi_1} & 0 & 0 & 0 \\ 0 & e^{j\Delta\varphi_2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & e^{j\Delta\varphi_N} \end{bmatrix} \tag{3.7}$$

and the optical fields at the output ports of the GMZI are given by

$$\begin{bmatrix} E_1^{out} \\ E_2^{out} \\ \vdots \\ E_N^{out} \end{bmatrix} = T \begin{bmatrix} E_1^{in} \\ E_2^{in} \\ \vdots \\ E_N^{in} \end{bmatrix} \tag{3.8}$$

where T is the total transfer matrix of the GMZI, given by

$$T = S_{combiner} \cdot S_{shift} \cdot S_{splitter} \tag{3.9}$$

3.4 Variable-ratio power splitting

The GMZI transfer matrix can be used to derive two important results. The first is that analytic expressions for the output field intensities in terms of the applied phase shifts can be derived by expanding (3.9). Secondly, the matrix equation (3.8) can be solved to yield the phase shifts $\Delta\varphi_j$ required to achieve a given output intensity distribution. This is the principle behind the operation of the variable-ratio power splitter.

Specification of the output intensity distribution is useful in such applications as the tap function, where only a small portion of the light is extracted from a channel for monitoring purposes [25], in ring lasers where the splitting ratio of the coupler dictates the operation of the laser [44], or in WDM devices where the spectral response can be improved through non-uniform power splitting [27].

Variable-ratio power splitting also has interesting potential applications in emerging optical networks as shown in Fig. 1.4(c), where the nodes must optimally allocate and distribute signal power to numerous paths based on span length and/or changing demand. Such an application is envisioned in metropolitan optical networks which will deliver a mix of broadband data services in a fiber-to-the-X infrastructure, where X may represent a region of a city, a neighborhood, a street (curb) or a home. In this scheme, a central transmission station may initially allocate signal power to various substations in an optimum manner, based on their distance and the number of customers demanding services at any given time. Alternatively, these allocations can be determined at the substation level or can be shared between both levels. Flexibility in setting the power ratio and the ability to quickly modify this ratio are paramount, and the variable-ratio splitter can satisfy both requirements.

Passive, N-way power splitters using MMI couplers have been realized [25]; however, only a few fixed ratios are possible with these devices. Variable-ratio power splitting has been achieved in passive two-port devices using angled MMI couplers [45] or the 'butterfly' MMI geometry [23], however these devices have not been developed for a general N and in many cases the ratios cannot be set arbitrarily. The GMZI structure overcomes these difficulties, in that it provides passive or active variable-ratio power splitting for general N.

The procedure for determining the optimum phase shifts needed to operate the variable-ratio power splitter is now given.

3.4.1 Output intensity expressions

The elements of the transfer matrix T, denoted by t_{ik}, are found by substituting (3.4) and (3.7) into (3.9):

$$t_{ik} = \frac{1}{N}\sum_{j=1}^{N} e^{J\left(\varphi_{ij}+\Delta\varphi_j+\varphi_{jk}\right)}$$

(3.10)

where $M = 1$ has been used in (3.4). Note the symmetry relation $t_{ik} = t_{ki}$. With a single input beam of optical intensity I_0 into port i, the intensity at an output port k is given by

$$I_k^{out} = |t_{ki}|^2 I_0$$

(3.11)

Substituting (3.10) into (3.11) and expanding the complex exponential term yields

$$I_k^{out} = \frac{I_0}{N^2}\left\{\left[\sum_{j=1}^{N}\cos\left(\varphi_{ij}+\Delta\varphi_j+\varphi_{jk}\right)\right]^2 + \left[\sum_{j=1}^{N}\sin\left(\varphi_{ij}+\Delta\varphi_j+\varphi_{jk}\right)\right]^2\right\}$$

(3.12)

Expanding the squared terms and simplifying the resulting expression yields

$$I_k^{out} = \frac{I_0}{N^2}\left\{N + 2\left[\sum_{p=1}^{N-1}\sum_{q=p+1}^{N}\cos\left[(\varphi_{ip} - \varphi_{iq}) + (\Delta\varphi_p - \Delta\varphi_q) + (\varphi_{pk} - \varphi_{qk})\right]\right]\right\} \quad (3.13)$$

where p and q refer to the internal arms j. Analytic expressions for the intensities at the various outputs in terms of the applied phase shifts $\Delta\varphi_j$ can be derived from (3.13). To illustrate this result, analytic expressions are explicitly developed for two specific cases.

First consider the 2×2 GMZI. Substituting $N = 2$ into (3.13) yields

$$I_1^{out} = \frac{I_0}{4}\left\{2 + 2\cos\left[(\varphi_{11} - \varphi_{12}) + (\Delta\varphi_1 - \Delta\varphi_2) + (\varphi_{11} - \varphi_{21})\right]\right\} \quad (3.14)$$

$$I_2^{out} = \frac{I_0}{4}\left\{2 + 2\cos\left[(\varphi_{11} - \varphi_{12}) + (\Delta\varphi_1 - \Delta\varphi_2) + (\varphi_{12} - \varphi_{22})\right]\right\} \quad (3.15)$$

where an input into port $i = 1$ has been (arbitrarily) chosen. The relative phase matrix is given by (3.2) with $N = 2$

$$\Phi = \begin{bmatrix} 0 & \dfrac{\pi}{2} \\ \dfrac{\pi}{2} & 0 \end{bmatrix} \quad (3.16)$$

where a global phase shift has been used for convenience. Substituting (3.16) into (3.14) and (3.15) yields

$$I_1^{out} = I_0\sin^2\left(\frac{\Delta\varphi_1 - \Delta\varphi_2}{2}\right) \quad (3.17)$$

and

$$I_2^{out} = I_0\cos^2\left(\frac{\Delta\varphi_1 - \Delta\varphi_2}{2}\right). \quad (3.18)$$

The intensity expressions in (3.17) and (3.18) are the standard result for the two-arm Mach-Zehnder interferometer (see, for example [8]). As expected, the classical Mach-Zehnder interferometer is equivalent to a GMZI with $N = 2$.

The same analysis can be used to determine the analytic intensity functions for arbitrary N. As N increases, however, the number of active phase shifting elements and the number of terms in (3.13) also increase, resulting in output intensities which are complicated functions of multiple variables. For instance, the intensities at the outputs of the 3×3 GMZI are given by

$$I_1^{out} = \frac{I_0}{9}\left\{3 + 2\left[\cos\gamma_1 - \frac{1}{2}\cos\gamma_2 + \frac{\sqrt{3}}{2}\sin\gamma_2 - \frac{1}{2}\cos\gamma_3 + \frac{\sqrt{3}}{2}\sin\gamma_3\right]\right\}$$

$$I_2^{out} = \frac{I_0}{9}\left\{3 + 2\left[-\frac{1}{2}\cos\gamma_1 - \frac{\sqrt{3}}{2}\sin\gamma_1 - \frac{1}{2}\cos\gamma_2 - \frac{\sqrt{3}}{2}\sin\gamma_3 + \cos\gamma_3\right]\right\} \quad (3.19)$$

$$I_3^{out} = \frac{I_0}{9}\left\{3 + 2\left[-\frac{1}{2}\cos\gamma_1 + \frac{\sqrt{3}}{2}\sin\gamma_1 + \cos\gamma_2 - \frac{1}{2}\cos\gamma_3 - \frac{\sqrt{3}}{2}\sin\gamma_3\right]\right\}$$

where $\gamma_1 \equiv \Delta\varphi_1 - \Delta\varphi_2$, $\gamma_2 \equiv \Delta\varphi_1 - \Delta\varphi_3$, $\gamma_3 \equiv \Delta\varphi_2 - \Delta\varphi_3$, and an input into port $i = 1$ has again been arbitrarily chosen.

While the output power distribution for any given input and set of phase shifts can be found using this analysis, the determination of the phase shifts needed to achieve a specified power splitting ratio becomes exceedingly difficult beyond the simplest ($N = 2$) case. For larger N, the phase shifts needed to achieve a desired output power distribution must be determined numerically. To address this issue a numerical approach is now introduced.

3.4.2 Numerical Analysis and Optimization Algorithm

In the variable-ratio power splitter with a single input into port i, the goal is to arbitrarily specify the intensities at the outputs. The problem is to solve the vector analogue of (3.11):

$$\begin{bmatrix} |t_{1i}|^2 \\ |t_{2i}|^2 \\ \vdots \\ |t_{Ni}|^2 \end{bmatrix} I_0 = \begin{bmatrix} I_1^{out} \\ I_2^{out} \\ \vdots \\ I_N^{out} \end{bmatrix} \tag{3.20}$$

where $|t_{ki}|^2$ is obtained from (3.13). Equation (3.20) is a system of N nonlinear equations with N unknowns: the phase shifts $\Delta\varphi_j$.

Equation (3.20) can be arranged to define a new vector Y:

$$Y \equiv \begin{bmatrix} |t_{1i}|^2 \\ |t_{2i}|^2 \\ \vdots \\ |t_{Ni}|^2 \end{bmatrix} I_0 - \begin{bmatrix} I_1^{out} \\ I_2^{out} \\ \vdots \\ I_N^{out} \end{bmatrix} \tag{3.21}$$

so that the problem can be recast as

$$Y(x) = 0 \tag{3.22}$$

where the required phase shifts are given by

$$x \equiv \begin{bmatrix} \Delta\varphi_1 \\ \Delta\varphi_2 \\ \vdots \\ \Delta\varphi_N \end{bmatrix} \tag{3.23}$$

and the vector notation is used for convenience.

The nonlinear system (3.22) has been solved numerically using an iterative procedure. The numerical algorithm is described here only in general terms; the details of the implementation are given in [46], [47]. An initial guess x_0 is made, after which the following occurs at each iteration: first, the partial derivatives of $Y(x_0)$ are calculated by a finite difference approximation, using a systematic perturbation of each of the elements of x_0. The partial derivatives are then used in the Levenberg-Marquardt set of linear equations, to establish a direction in which to search for decreasing values of the function Y. A step is then taken in this direction, and the iterate is updated.

A limitation of this numerical method is that an iterate of x may sometimes settle into a local rather than a global minimum. To overcome this, the initial guess x_0 is perturbed significantly and the analysis is repeated. Confidence in the numerical solution is gained when it is consistently obtained from a variety of initial guesses.

Although the number of iterations required to achieve a solution depends upon the initial guess, the desired output intensities, and the value of N, a solution is generally found in under $100N$ iterations.

It must be noted, however, that for some intensity distributions where $N > 2$, a real solution to the nonlinear system may not exist. This effect occurs for a class of degenerate distributions, where one output intensity is near zero, and the others are not intensity extrema (i.e., 0% or 100%). In these cases, the closest real solution is found, which in general results in a

small deviation of the output intensities from the specified distribution. This effect arises from the interdependence of the nonlinear equations. When an extremum in intensity emerges from one port, (3.13) dictates that intensity extrema must also emerge from all other ports.

Once the required phase shifts x are determined numerically, they are optimized by offsetting the phases by a constant amount and shifting individual phases by multiples of 2π. This allows the GMZI to be operated with a driving signal consuming the least possible power.

Once the optimum phase shifts x_{opt} are found, the GMZI can be implemented in a reconfigurable mode, where active phase shifters are used, or it can be implemented in a passive mode, where the phase shifts are achieved by varying the internal arm lengths. The intended application of the GMZI will dictate whether it should be implemented in an active or a passive mode.

As an example of the N-way power splitting function, consider a GMZI with $N = 7$. With a single input beam into port $i = 2$ (arbitrarily chosen), we wish to achieve the following (arbitrarily chosen) output intensity distribution:

$$I_k^{out} = \begin{bmatrix} 0.06 \\ 0.02 \\ 0.41 \\ 0.15 \\ 0.15 \\ 0.01 \\ 0.20 \end{bmatrix} I_0 \tag{3.24}$$

where the numbering of the outputs is from top to bottom, as in Fig. 2.7. Substituting (3.24) into (3.21) and solving for the optimum phase shifts yields:

$$x_{opt} = \begin{bmatrix} -0.24 \\ -0.54 \\ 0.66 \\ 1.26 \\ -0.85 \\ 1.52 \\ -1.52 \end{bmatrix} \text{radians} \qquad I_k^{out} = \begin{bmatrix} 0.0603 \\ 0.0172 \\ 0.4093 \\ 0.1506 \\ 0.1494 \\ 0.0138 \\ 0.1994 \end{bmatrix} I_0 \tag{3.25}$$

In this example, the maximum phase shift magnitude needed in any single arm is 1.52 radians = 87.1°. The optimization routine ensures that the maximum phase shift magnitude needed to achieve any intensity distribution in any N-port GMZI is always $\leq \pi$.

3.5 Tolerance analysis

In practice, a number of parameter variations and non-ideal effects will limit the performance of the variable-ratio power splitter. Specifically, the response of the GMZI to the effects of deviations in the applied phase shifts and non-ideal performance of the MMI couplers needs to be determined. The sensitivity of the response to these effects will dictate the design, fabrication, and operational steps needed to realize devices with acceptable performance levels. The effect of inaccuracies in the applied phase shifts is determined first, using the model described in the previous section. The effect of MMI coupler parameters is addressed next, by introducing a number of modifications to the transfer matrix theory to arrive at a model for the non-ideal GMZI.

3.5.1 Phase shift deviations

The phase shifts in (3.25) result in an intensity distribution with a maximum deviation of 0.38% from the distribution specified in (3.24). To achieve this result, the phase shifts must be applied to a precision of at least ± 0.01 radians, which yields a tolerable phase error of $\pm 0.01 \times 180/\pi = \pm 0.57$ °. As the precision of the applied phase shifts diminishes, the output intensities deviate further from the specified distribution; however, the deviation remains small for even a relatively large change in the applied phase shifts. In this example, it is found that if only a ± 0.1 radian (± 5.7°) precision in the phase shifts can be maintained, the intensity distribution will deviate by only 0.62% from the ideal values. In general, the accuracy of solutions varies with the chosen intensity distribution and the value of N.

This result is strictly only applicable for a single wavelength, because as multiple wavelengths travel through a fixed-length region of modified refractive index, the phases of the different wavelengths are shifted by different amounts. This additional phase shift deviation is proportional to the change in wavelength. For example, over an 80 nm wavelength range (bandwidth) from 1510 nm to 1590 nm, the maximum phase shift deviation is 2.6 % from the phase shift applied at 1550 nm. The effect is also relative, so as the phase shift in an arm approaches zero, it will be less wavelength sensitive. In the above example, it is found that over the entire 80 nm bandwidth the wavelength effect degrades the intensity distribution in (3.24) by a maximum of 0.74 % using an accuracy of ± 0.01 radians, and by a maximum of 1.63 % using an accuracy of ± 0.1 radians.

Although an arbitrary bandwidth has been used in the above example, it must be noted that the bandwidth of real devices cannot be arbitrarily specified. The actual bandwidth of the GMZI is limited by the finite bandwidth of the $N \times N$ MMI coupler and is a result of both material parameters and device design [31].

3.5.2 Theory of the non-ideal GMZI

Since the GMZI relies upon the interference of modes and interactions among self-images, its ultimate performance depends upon the accuracy of the amplitudes and phases of multiple beams within the device. This places strict requirements on the performance of the constituent MMI couplers. In practice, errors in the phases of the self-images and imbalance in the power splitting ratio are both present and can significantly degrade the performance and complicate the active control of devices based on the GMZI [33], [34], [48]. Furthermore, imperfect waveguide fabrication is known to cause additional phase errors [34], [49]. Since the arms of the GMZI are often relatively long to accommodate efficient phase shifting, the cumulative effects of these errors can be significant. While the MMI coupler performance is also a function of the degree to which the theoretical relations in Chapter 2 are satisfied, the analysis here is mainly concerned with sources of non-ideality in addition to these fundamental limits.

While an analysis of the non-ideal nature of the GMZI is important for determining the performance limits of the variable-ratio power splitter, the results derived actually have a much wider applicability. This is because the number and complexity of applications using the GMZI continues to expand. In addition to switches [26], [34], the GMZI is used in a number of theoretical constructions such as multi-frequency lasers [50], flat-response wavelength (de)multiplexers [51], and wavelength-selective switches [52].

In this section, a model for the non-ideal GMZI is developed by incorporating non-ideal parameters into the transfer matrix approach described in Sections 3.2 and 3.3. The model is then used to investigate the tolerances and scalability of the variable-ratio power splitter.

3.5.2.1 Modified GMZI transfer matrix

Deviations in the phases of the self-images and imbalance in the power splitting ratio are the principal sources of non-ideal behaviour in MMI couplers. These phenomena modify the transfer matrix (3.4) for the $N \times N$ MMI coupler $S_{coupler}$ such that the element s_{ij} can be described by

$$\tilde{s}_{ij} = a_{ij}e^{J(\varphi_{ij} + \delta\varphi_{ij})}$$

(3.26)

where a_{ij} is the (real) field amplitude transfer coefficient from input i to output j, $\delta\varphi_{ij}$ is the phase deviation of the image at output port j for an input into port i, and the tilde (~) is used to distinguish the parameters of the non-ideal device theory from the idealized theory. In a lossless device, the optical power at the output ports is conserved, i.e.,

$$\sum_{j=1}^{N} a_{ij}^2 = 1$$

(3.27)

The imbalance in an output port (measured in dB units) is given by

$$Imbalance = 20 \cdot \log_{10}\left(\frac{a_{ij}}{a_{ij_{ref}}}\right)$$

(3.28)

where j_{ref} is a reference output port.

The effect of phase errors due to imperfect waveguide fabrication can be incorporated into the theory by introducing a phase error $\delta\varphi_j$ in each arm. The matrix S_{shift} from (3.7) is then modified to give

$$\tilde{S}_{shift} = \begin{bmatrix} e^{J(\Delta\varphi_1 + \delta\varphi_1)} & 0 & 0 & 0 \\ 0 & e^{J(\Delta\varphi_2 + \delta\varphi_2)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & e^{J(\Delta\varphi_N + \delta\varphi_N)} \end{bmatrix}$$

(3.29)

The total transfer matrix $\tilde{T}$ of the non-ideal GMZI is then given by

$$\tilde{T} = \tilde{S}_{combiner} \cdot \tilde{S}_{shift} \cdot \tilde{S}_{splitter}$$

(3.30)

3.5.2.2 Output intensity expressions

The elements of the transfer matrix $\tilde{T}$ are denoted by $\tilde{t}_{ik}$, and are found by substituting (3.26) and (3.29) into (3.30) to yield

$$\tilde{t}_{ik} = \sum_{j=1}^{N} a_{ij}a'_{jk}e^{J(\varphi_{ij} + \delta\varphi_{ij} + \Delta\varphi_j + \delta\varphi_j + \varphi_{jk} + \delta\varphi'_{jk})}$$

(3.31)

where the primed notation is used to distinguish the parameters of the MMI combiner from that of the splitter. In general the amplitude transfer coefficients will differ and phase deviations will vary from coupler to coupler and arm to arm. With a single input beam of optical intensity I_0 into port i, the intensity at an output port k is given by

$$\tilde{I}_k^{out} = \left|\tilde{t}_{ki}\right|^2 I_0 = I_0\left(\sum_{j=1}^{N} a_{kj}^2 a_{ji}'^2\right) + 2I_0\left[\sum_{p=1}^{N-1}\sum_{q=p+1}^{N} a_{kp}a_{pi}'a_{kq}a_{qi}' \cos\left[(\varphi_{ip} - \varphi_{iq}) + \right.\right.$$

$$\left.\left.(\delta\varphi_{kp} - \delta\varphi_{kq}) + (\Delta\varphi_p - \Delta\varphi_q) + (\delta\varphi_p - \delta\varphi_q) + (\varphi_{pk} - \varphi_{qk}) + (\delta\varphi_{pi}' - \delta\varphi_{qi}')\right]\right]$$

(3.32)

where p and q refer to the internal arms j.

Analytic expressions for the intensities at the various outputs in terms of the applied phase shifts, amplitude transfer coefficients, and phase deviations can be derived from (3.32). If these parameters are known, (3.32) contains N equations (one for each desired output power) and N unknowns (the applied phase shifts), and can be solved numerically to determine the phase shifts necessary to operate the GMZI as a generalized power splitter.

In practice, however, these parameters are not known, resulting in $(2N^2 + 4N)$ unknowns in (3.32). Determination of these parameters is exceedingly difficult unless the GMZI is physically separated into its splitter, arm, and combiner regions, which is impractical. The result is that the response of the GMZI to various applied phase shifts has not been predictable [34], [48], so optimization schemes or iterative methods are needed for the GMZI to be useful in optical switching or power splitting applications. Furthermore, the presence of imbalance and phase deviations in the MMI couplers and GMZI arms indicate that the self-images do not appear exactly at the output waveguide positions and therefore result in higher excess losses.

Improvement of waveguide fabrication technology can in principle solve these problems. For example, it is known that small variations in the width of MMI couplers can result in imbalance and phase deviations among the output images [31]. Furthermore, imperfect waveguide fabrication can result in local variations in the waveguide width and core/cladding refractive indices, thereby introducing additional phase errors [49]. Since the arms of the GMZI consist of long, narrow, single-mode waveguides, the phase errors manifest themselves principally in the GMZI arms.

3.5.2.3 Non-ideal GMZI characteristic response

To gauge the performance of an active device, it is important to first determine its linear or characteristic response. In the case of a 2 × 2 GMZI, the response is measured by monitoring the output intensity from both ports while gradually increasing the applied power to a single phase shifter. This method can be extended to the $N \times N$ GMZI by choosing an arbitrary input port and setting all but one phase shifter equal to zero. The intensity at the N output ports can then be plotted as a function of the single phase shift. In this manner, a 'cross section' of the response of a non-ideal GMZI is obtained, and can be compared to that of an ideal one. The deviation between the two responses qualitatively indicates the degree to which phase deviations and imbalance are present in a device.

As an example, consider the $N = 4$ GMZI, input port $i = 2$ (arbitrarily chosen), and operation of the phase shifter on arm $j = 4$ (arbitrarily chosen). The response of the ideal GMZI given by (3.13) is plotted in Fig. 3.2 for a single period of $\Delta\varphi_4$. Note that this result is for perfect or ideal MMI couplers with no imbalance or phase deviation.

The model (3.32) is now used to determine the role imbalance and phase deviations in the MMI couplers play in determining the GMZI performance. To understand the origins of these non-idealities, consider (3.1), which indicates that the correct coupler length is proportional to the square of the width. For example, using typical values in (3.1), it is found that a deviation of ± 0.1 μm in the width of an MMI coupler typically results in an optimum imaging length deviation of $\delta L = \pm 3$-5 μm. The correct coupler length is therefore a sensitive function of the coupler width. Since the resolution of the lithographic techniques used to fabricate the MMI couplers is

finite, slight variations in the length of the MMI couplers typically occur [31]. The result is that the couplers are often longer or shorter than the optimum length for self-imaging. As a result, the self-images will be blurry and out of focus or will not form in the correct lateral positions.
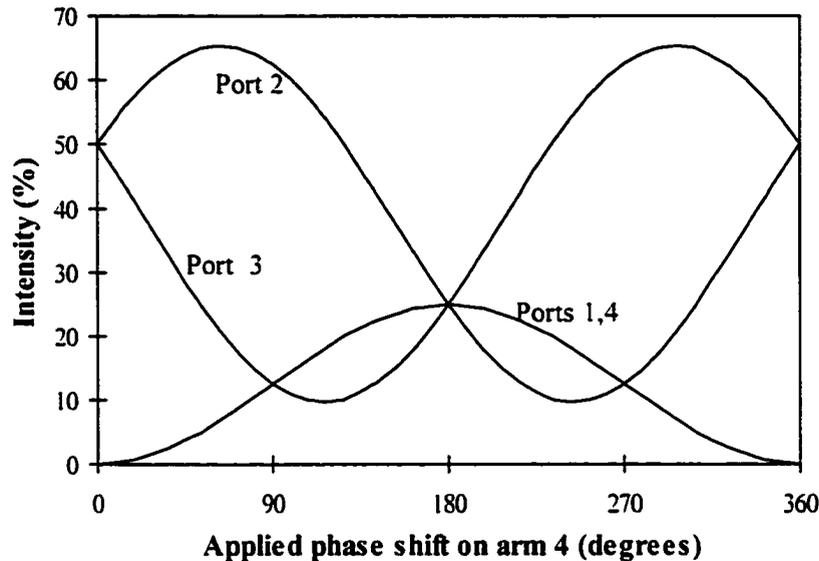
Figure 3.2 Response of an ideal 4 × 4 GMZI to a single phase shift applied to arm 4.

To determine the effect of MMI coupler length deviations on the imbalance and phase deviation among the self-images, a modal propagation analysis (MPA) simulation of a 4 × 4 MMI coupler was performed, using a waveguide design which will be discussed in detail in Chapter 5. The MPA result is given in Fig. 3.3 and shows the maximum imbalance and phase deviation among the self-images for various MMI coupler lengths, using output port 2 as a reference (arbitrarily chosen). A given level of imbalance and phase deviation can be compactly expressed as an equivalent coupler length deviation, δL.
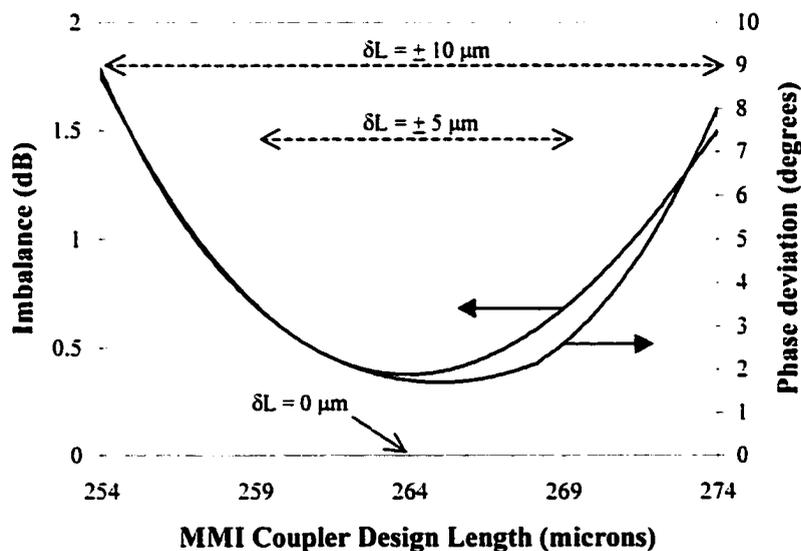
Figure 3.3 MPA simulation of a 4 × 4 MMI coupler showing the imbalance and phase deviation among the self-images for various coupler lengths.

It must also be kept in mind that the optimum MMI coupler imaging length (i.e., δL = 0 μm) still represents a departure from ideality, since the strongly guiding approximation results in only approximately ideal sine-like modes in the coupler, as seen from (2.20). Therefore any practical coupler will have some imbalance and phase deviation even at the optimum imaging length.

Worst-case values of imbalance and phase deviation for the three values of δL shown in Fig. 3.3 are found from performing multiple MPA simulations where each output port has been chosen as a reference. The values are then used as bounds within which values for a, a', $\delta\varphi$, and $\delta\varphi'$ are randomly chosen. These parameters are then introduced into (3.32), ensuring that the maximum imbalance and phase deviation are always present in at least one output port. The ideal curve for an arbitrarily chosen output port ($k = 2$) of the GMZI from Fig. 3.2 is then modified as shown in Fig. 3.4. Imbalance and phase deviations in the MMI couplers result in a shift in the phase of the ideal response curve, and changes in both the offset level and the amplitude.
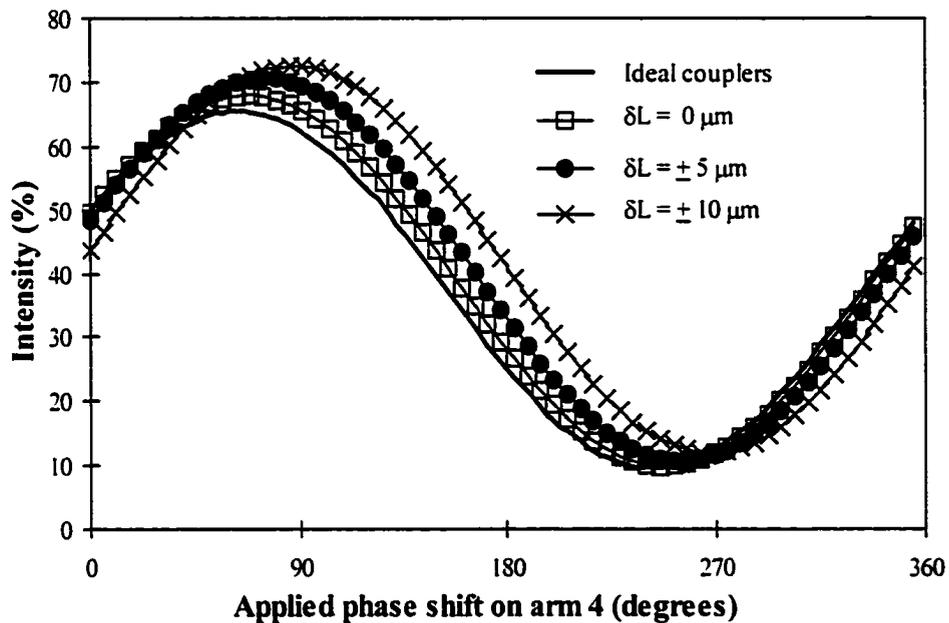
Figure 3.4 Calculated deviation of the ideal response of port 2 from Fig. 3.2 for various equivalent length deviations in the MMI couplers.

The model is now used to determine the effects of phase errors in the arms of the GMZI. Introducing random phase errors with various upper limits into (3.32) modifies the ideal curve for output port 2 from Fig. 3.2 in the manner shown in Fig. 3.5.
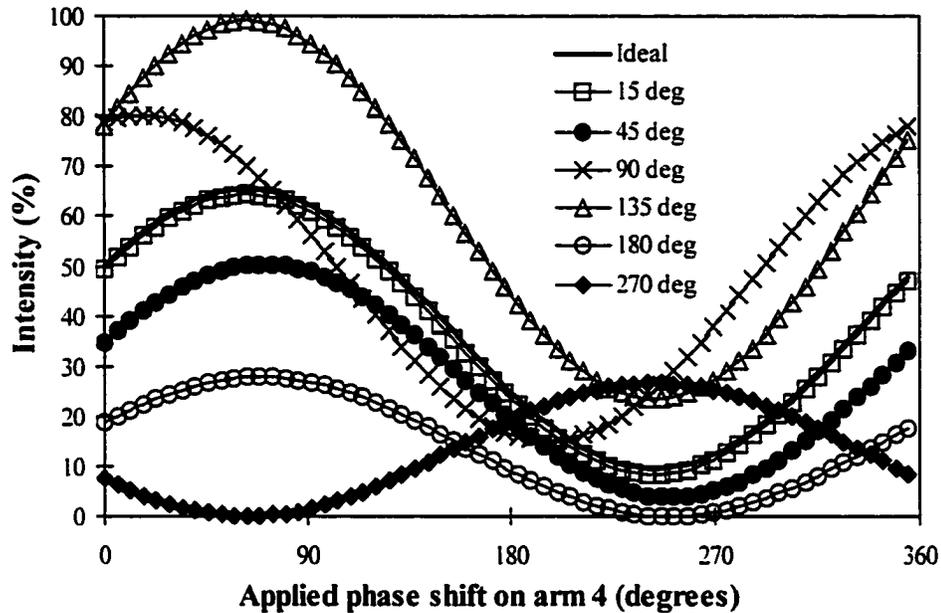
Figure 3.5 Calculated deviation of the ideal response of port 2 from Fig. 3.2 for various phase errors in the GMZI arms distributed randomly between 0° and various upper bounds as indicated.

It is seen from Fig. 3.5 that phase errors in the arms of the GMZI significantly alter the phase, amplitude and offset level of the ideal response. These errors are approximated by random values because material and waveguide parameters will vary locally from point to point in a given waveguide and arm to arm in a given device. If the magnitude of these random errors is limited to small fluctuations as shown in Fig. 3.5, the device response agrees reasonably well with the theory. In practice, the phase shifts applied to the arms can include a bias component to compensate for the phase errors in the arms.

3.5.3 Variable-ratio power splitter tolerances

In this section the variable-ratio power splitting response of the non-ideal GMZI is investigated by including the combined effects of non-idealities in the MMI couplers and the arms. The model given by (3.32) is used as in the characteristic response analysis of the previous section, but with one important difference: the N applied phase shifts are, in general, all nonzero. The sensitivity of the power splitting to various non-idealities is gauged by choosing an arbitrary set of phase shifts, and determining the resulting deviations in the output intensity ratios. This is best illustrated through the use of examples.

Consider variable-ratio power splitting with $N = 4$, using an arbitrarily chosen power splitting state of

$$I_k^{out} = \begin{bmatrix} 0.10 \\ 0.30 \\ 0.40 \\ 0.20 \end{bmatrix} I_0 \tag{3.33}$$

The required theoretical phase shifts are given by the solution (3.23) to the ideal system of nonlinear equations given by (3.13). The solution and various values for non-idealities are both used in (3.32) to calculate the resulting splitting ratio. Table 3.1 lists the deviations of this

ratio from the desired ratio given in (3.33). Various levels of imbalance and phase deviation in the MMI couplers have been introduced as equivalent deviations δL in the imaging length, given by the simulated results in Fig. 3.3. All values in the table were calculated using a random fluctuation of $\pm$ 3 % of the value of the phase shift. This allows for the combined effects of a finite precision with which the phase shifts can be applied, phase shift deviations due to wavelength variations, and fluctuations in the power applied to the phase shifters.

4 × 4 MMI coupler deviation (Fig. 3.2 - worst case from all reference ports)			Arm phase error (degrees)	Deviation in variable-ratio power splitting state above (+) and below (-) desired value (% of total output intensity)			
Equivalent δL	Imbalance (dB)	Phase Deviation (degrees)		Port $k = 1$	Port $k = 2$	Port $k = 3$	Port $k = 4$
Ideal coupler	0	0	0	- 0.04	- 0.37	+ 0.71	- 0.30
$\delta L = 0$ μm	0.4	1.5	0	- 0.11	+ 1.17	+ 0.54	- 1.60
			5	+ 0.07	- 1.10	+ 3.49	- 2.46
$\delta L = \pm 5$ μm	1.0	3.5	0	- 0.38	+ 1.34	- 3.17	- 4.13
			5	- 2.80	+ 4.57	+ 1.53	- 3.29
$\delta L = \pm 10$ μm	2.0	9.0	0	- 1.33	+ 4.10	+ 8.13	- 10.91
			5	- 3.07	+ 6.08	+ 8.79	- 11.80

Table 3.1 Simulated tolerances of $N = 4$ variable-ratio power splitter for power splitting ratio (3.33), including a $\pm$ 3 % fluctuation in the applied phase shifts. The 'arm phase error' refers to an upper limit for the uncompensated phase error in the GMZI arms.

It is seen in Table 3.1 that a degradation in the power splitting ratio occurs as the imaging in the MMI couplers worsens. In this case, an acceptable level of deviation in the power splitting ratio is achieved if δL for the 4 × 4 MMI couplers is kept below $\pm$ 5 μm. As the level of uncompensated phase errors in the GMZI arms increases from 0° to 5°, the effect on the deviation in the power splitting ratio is small when the MMI couplers are close to ideal; however, as the MMI coupler imaging degrades, the effects of errors in the arms become more significant. For the operation of the power splitter it is therefore of prime importance to ensure that the MMI couplers perform well.

To investigate the tolerance behaviour in larger devices, a similar analysis is performed for $N = 5$ up to $N = 10$ using an equal power splitting state in every case, i.e.,

$$I_k^{out} = \frac{I_0}{N} \qquad (3.34)$$

where $k = 1, 2, \ldots , N$. The results are shown in Table 3.2. To simulate scaling effects in real devices, a maximum MMI coupler imbalance of $\pm$ 1.0 dB and phase deviation of $\pm$ 3.5° are used for $N = 5,6,7$ and maximum imbalances and phase deviations of $\pm$ 2.0 dB and $\pm$ 9° , respectively, are used for $N = 8,9,10$. Imbalance values such as these have been measured in MMI couplers with large N [22], [29]. Table 3.2 also includes the phase shift fluctuations used in Table 3.1.

N	Deviation in equal power splitting state (% of total output intensity) Above (+) and below (-) desired value									
	Output Port k									
	1	2	3	4	5	6	7	8	9	10
5	- 0.53	- 2.36	- 2.16	+ 3.26	+ 1.80					
	- 0.17	- 2.22	- 1.40	+ 2.52	+ 1.27					
6	+ 0.12	+ 0.99	- 1.31	+ 3.19	+ 0.72	- 3.71				
	+ 0.58	+ 2.30	- 1.95	+ 0.45	+ 0.24	- 1.62				
7	- 2.07	+ 2.35	- 1.11	- 0.98	- 2.12	+ 2.34	+ 1.58			
	- 0.95	+ 2.92	- 0.35	- 1.24	- 1.77	- 1.00	+ 2.39			
8	- 2.14	+ 1.15	- 2.15	+ 2.15	- 0.60	- 0.97	- 1.95	+ 4.52		
	- 2.06	+ 1.60	- 2.80	+ 1.84	- 0.70	- 1.00	+ 0.63	+ 2.49		
9	- 2.10	- 3.15	- 0.62	- 1.15	+ 1.61	+ 1.83	- 0.19	+ 1.29	+ 2.47	
	- 0.13	- 2.81	+ 1.26	- 2.80	+ 2.03	+ 0.21	+ 0.14	+ 0.95	+ 1.15	
10	- 2.74	- 0.05	- 1.38	- 3.66	- 1.65	- 1.75	+ 0.24	+ 0.19	+ 5.74	+ 5.05
	- 2.45	+ 1.87	- 2.84	- 4.94	- 0.03	- 2.30	+ 0.71	+ 0.25	+ 6.47	+ 3.26

Table 3.2 Simulated tolerances of variable-ratio power splitter for $N = 5$ up to $N = 10$ for power splitting ratio (3.34), including a ± 3 % fluctuation in the applied phase shifts. The imbalance and phase deviation for $N = 5, 6, 7$ are ± 1.0 dB and ± 3.5° and for $N = 8, 9, 10$ they are ± 2.0 dB and ± 9°, respectively. The first and second lines for each value of N represent upper limits of 0° and ± 5° for the uncompensated phase error in the GMZI arms.

Deviations in the power splitting ratio seen in Table 3.2 increase slowly with increasing N. Maintaining a certain level of imbalance and phase deviations, however, becomes more difficult as N increases. This is because the tolerable length deviation δL required for a given level of imbalance and phase deviation in the MMI coupler shrinks as N increases. This effect is partially taken into account in the analysis by using larger imbalance and phase deviations for $N = 8,9,10$ than for $N = 5,6,7$. It is also seen in Table 3.2 that partially uncompensated phase errors in the GMZI arms have only a small effect upon the power splitting ratio N, sometimes even improving the ratio due to partial compensation of the inaccuracies in the applied phase shifts.

The results of the tolerance analysis performed in this section indicate that the variable-ratio power splitter can be operated even if the GMZI is non-ideal. Furthermore, the power splitting ratio remains stable and tolerant to parameter fluctuations even for moderate values of N.

One effect not considered in the preceding analysis, however, is that of a reduction in the range of available power splitting ratios as the MMI coupler imaging degrades. This effect is most easily understood by considering the special cases where the GMZI is operated as a switch. In this mode of operation, all the light emerges from a single output and there is ideally no light at the other outputs. In practice, however, the MMI coupler image plane does not perfectly coincide with the output ports (waveguides) and furthermore, the self-images have a limited resolution due to the finite number of modes available for self-imaging [53]. This results in some of the light emerging from unwanted ports, which is called crosstalk. The result is that a limited range of intensity values can emerge from a given port. Thus as the crosstalk increases, the range of available power splitting ratios decreases. The relation between crosstalk in the GMZI and various non-idealities is examined as part of the analysis in the next chapter, which deals with the special case of switching using the GMZI structure.

Chapter 4

Theory of Generalized Mach-Zehnder Interferometer Switching

Switching is a special case of the general power splitting capability of the GMZI. In this chapter the switching properties of the N × N GMZI are analyzed in detail. When multiple inputs and outputs are used, a number of novel switching and routing functions are possible. These functions enable the efficient routing and reconfiguration of optical paths within a network. Portions of this chapter have been reported in a paper to appear in the Journal of Lightwave Technology [36]; *a paper submitted to the* Journal of Lightwave Technology [54]; *and two* U. S. Patent Applications [55], [56].

4.1 Introduction

The transfer matrix approach described in the previous chapter is extended in this chapter to describe the switching capabilities of the GMZI. The GMZI operates as a 1 × N switch when certain sets of phase shifts are applied to the arms. These phase shifts can be derived from the transfer matrix theory and indicate that in principle simple digital switch control is possible for small values of N.

When multiple inputs are used, the GMZI provides a limited multi-port switching function that enables the design of larger 1 × N switches with relaxed control requirements. The multi-port switching can also be used to realize non-blocking switches comprised of stages of cascaded GMZI units. The simulated loss and crosstalk in these switches in the presence of non-ideal MMI coupler imaging highlights the practical challenges of realizing GMZI-based switches.

4.2 Idealized 1 × N GMZI switching

Consider the case where, for a given input port i the desired output distribution of the GMZI is set to a switch state, i.e.,

$$I_k^{out} = \begin{cases} I_0 & , \ k = \bar{k} \\ 0 & , \ k \neq \bar{k} \end{cases}$$ (4.1)

The output light emerges from port $\bar{k}$ and the switch state is denoted compactly as $(i, \bar{k})$. Substituting (4.1) into (3.21) and numerically solving the resulting system (3.22) using the idealized equations (3.13) results in a set of switching phase shifts. These can be optimized in the manner described in Section 3.4.2 to yield a set of phase shifts x_{opt}. Repeating this process for all switch states $\bar{k} \in \{1, 2, ..., N\}$, the resulting N sets of switching phase shifts x_{opt} form the columns of a matrix, called the ideal switching matrix X. For a given input port i, the desired output port is selected by the column k, and the element in row j gives the phase shift that must be applied to arm j in order to route (cross-connect) the light from input port i to output port k. Using a different input port amounts to interchanging the columns of the matrix X.

For example, the following ideal switching matrix is found for the 3 × 3 GMZI when $i = 1$:

$$X = \frac{2\pi}{3} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$ (4.2)

demonstrating that the 1 × 3 GMZI switch can be operated with phase shifts controlled by digital logic, as shown in Fig. 4.1. The phase shifters in Fig. 4.1 are designed so that the application of a voltage with a logic level '1' imparts a phase shift of magnitude $2\pi/3$ to the light travelling in the waveguide, while a logic level '0' imparts no phase shift. The matrix (4.2) is reproduced in the table in Fig. 4.1. The control word AB is applied to the logic circuitry and induces the appropriate phase shifts to route the light to the desired output port. In this case the circuitry uses a single exclusive-OR logic gate. If a center-fed 1 × 3 MMI coupler employing symmetric interference [22] is used as the splitter, X and the resulting control circuitry are further simplified, though this is not true for a general value of N.

Required 3 × 3 GMZI phase shifts $\left(x\frac{2\pi}{3}\right)$			
ΔP_1	1	0	1
ΔP_2	1	1	0
ΔP_3	0	1	1
Resulting 3 × 3 GMZI state	(1.1)	(1.2)	(1.3)
Control word AB	01	10	11
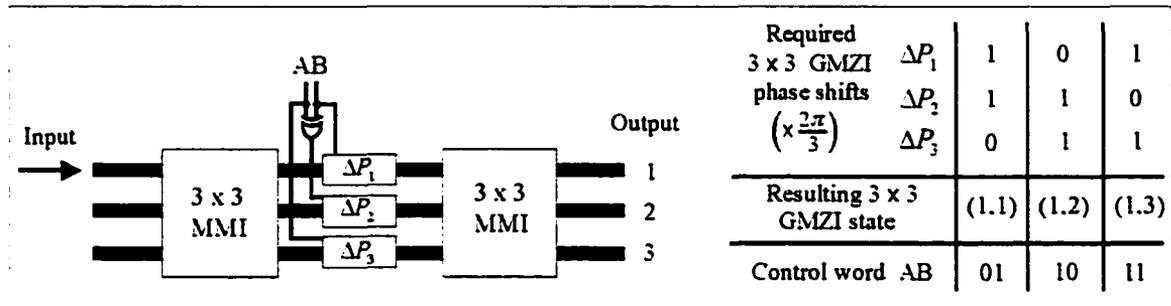
Figure 4.1 Digital control of a 1 × 3 GMZI switch using a control word AB to apply phase shifts from the ideal switching matrix (4.2).

Binary control is unique to the $N = 3$ GMZI switch. Generally, as N increases, X increases in complexity. For example, the ideal switching matrix for the $N = 5$ GMZI switch is given by

$$X = \frac{4\pi}{5}\begin{bmatrix} 0 & 1 & 2 & 0 & 2 \\ 0 & 2 & 1 & 2 & 0 \\ 2 & 2 & 0 & 0 & 1 \\ 2 & 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 2 & 2 \end{bmatrix} \tag{4.3}$$

As N increases, the number of distinct phase shift (logic) levels increases, while the maximum phase shift magnitude approaches 2π. A summary of the results obtained for switches with values of N up to 10 is shown in Fig. 4.2.

It should be emphasized that these results are idealizations, and assume that the non-ideal effects in a fabricated GMZI can be overcome. Even slight fabrication imperfections have been shown in the previous chapter to result in a non-ideal device response. While the GMZI can still be used as a variable-ratio power splitter in these cases, the switching requirements need to be investigated. This is done in the following section.
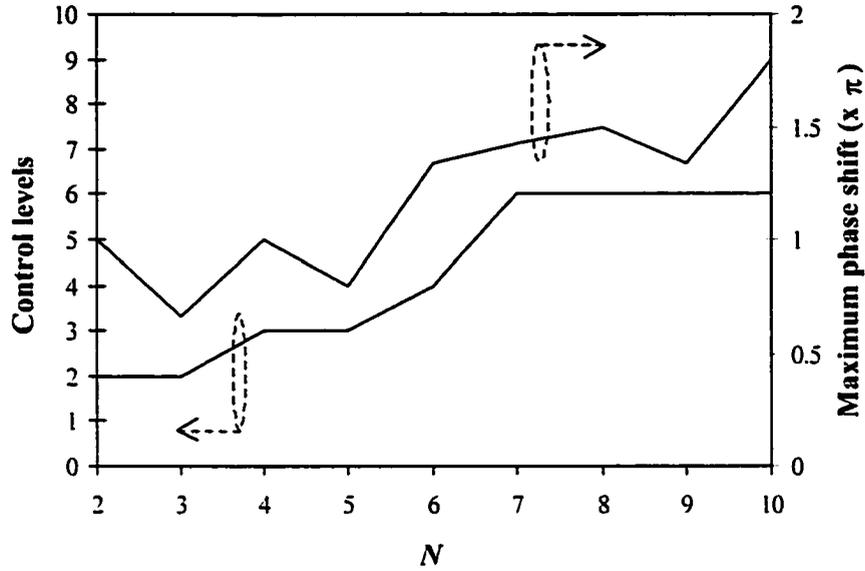
Figure 4.2 Variation in the number of distinct phase shift control levels and the maximum phase shift required for idealized 1 × N GMZI switch operation with the number of ports N.

4.3 Tolerances for GMZI switching

The performance of a 1 × N switch is gauged by two critical parameters: crosstalk and excess loss, defined by

$$\text{crosstalk} = 10 \cdot \log_{10}\left[\frac{\max\{I^{out}_{k \neq \bar{k}}\}}{I^{out}_{\bar{k}}}\right] \tag{4.4}$$

$$\text{excess loss} = 10 \cdot \log_{10}\left(\frac{\sum_{k=1}^{N} I^{out}_{k}}{I_0}\right) \tag{4.5}$$

where the max{ } function selects the element with the maximum value. A perfect switch has an infinitely low crosstalk (− ∞ dB) and a 0 dB excess loss. The crosstalk in the non-ideal GMZI is investigated in this section by using the model (3.32), while the excess losses will be investigated in Section 4.7 using the MPA technique.

To determine the crosstalk in the presence of various non-idealities, a technique similar to the analysis of the power splitter is used, but replacing (3.33) by an arbitrarily chosen switch state. In this case, $\bar{k} = 4$ in (4.1) and $i = 2$ are arbitrarily chosen. The crosstalk as a function of various MMI coupler deviations and fluctuations in the applied phase shifts are shown in Table 4.1, assuming the phase errors in the GMZI arms are fully compensated. Degradation in the imaging of the MMI couplers is seen to significantly increase crosstalk in the switch, in the same manner as it increased the error in the ratio from the variable-ratio power splitter in Chapter 3.

The trend in Table 4.1 suggests that good imaging in the MMI couplers is paramount to achieve low crosstalk levels. As the imaging approaches that of the ideal coupler, the sensitivity

of the crosstalk to the precision of the phase shifting increases, indicating that simultaneous betterment of both parameters is needed to obtain very low crosstalk levels.

MMI coupler deviation (Equivalent δL)	Phase Shift Fluctuation ($\pm$ %)	Typical Crosstalk (dB)
Ideal coupler	0	- ∞
	1	- 44.8
	2	- 38.6
$\delta L = 0$ μm	0	- 30.0
	1	- 28.7
	2	- 27.0
$\delta L = \pm 5$ μm	0	- 22.9
	1	- 22.4
	2	- 22.2
$\delta L = \pm 10$ μm	0	- 14.1
	1	- 13.9
	2	- 13.8

Table 4.1 Simulated tolerances of the crosstalk in a 1 x 4 GMZI switch assuming fully compensated phase errors in the arms.

A similar tolerance analysis has been performed for 1 x N GMZI switches up to $N = 10$, using a range of MMI coupler performances and phase shift fluctuations and again assuming compensation of the phase errors in the GMZI arms. The results are shown in Fig. 4.3.
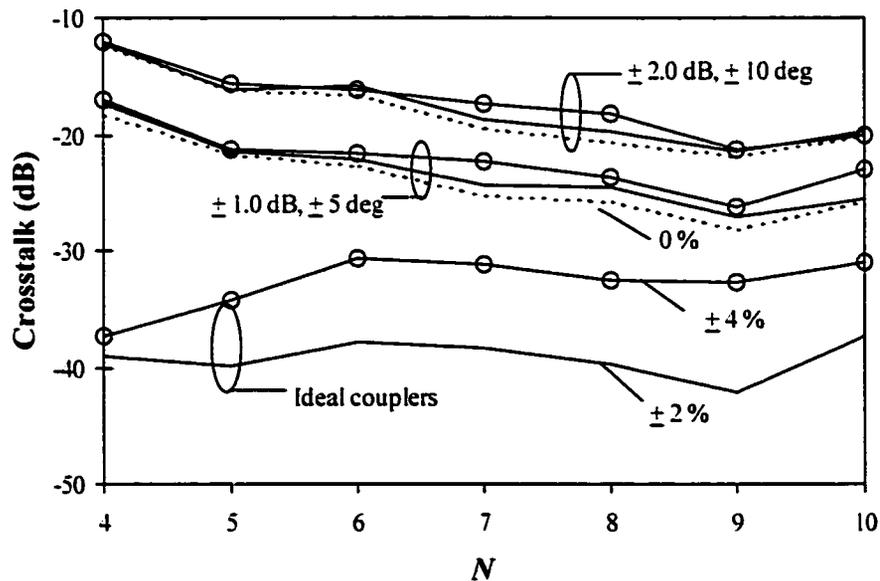
Figure 4.3 Crosstalk in 1 x N GMZI switches as a function of N for various MMI coupler imbalances, phase deviations, and fluctuations in the applied phase shifts. The dashed, solid, and circled lines are for applied phase shift fluctuations of 0 %, $\pm$ 2 %, and $\pm$ 4 %, respectively.

Since idealized switching demands that the phase shifts are exact values, additional fluctuations due to variations in the parameters of fabricated phase shifters (such as electrical resistance) are present. For this reason, an extra ± 2 % (a technologically feasible variation) is included to arrive at the ± 4 % fluctuation in Fig. 4.3.

In Fig. 4.3 three levels of MMI coupler performance are shown. If a certain level of performance can be maintained as N increases (usually requiring tightened fabrication tolerances), the level of crosstalk remains relatively constant. With perfect imaging in the MMI couplers (no imbalance or phase deviation) low crosstalk values are possible, but even slight inaccuracies in the applied phase shifts can degrade the crosstalk significantly. For example, a ± 2 % fluctuation in the applied phase shifts increases the crosstalk from $-\infty$ dB (with no fluctuation) to about – 40 dB.

For a maximum imbalance of ± 1.0 dB and phase deviation of ± 5°, the crosstalk is limited to about –28 dB, while increasing these values to ± 2.0 dB and ± 10° degrades the crosstalk to about –20 dB. Again as in Table 4.1, the sensitivity of the crosstalk to the accuracy of the applied phase shifts increases as the imaging in the MMI couplers improves.

It must be kept in mind that Fig. 4.3 shows typical results. Since the imbalance, phase deviation, and phase shift fluctuations are applied in a random manner, the exact position of a data point will vary slightly from simulation to simulation. The general trend from a large number of simulations indicates that the crosstalk improves slightly (< 5 dB) as N increases from 4 to 10. The reason for this is that the power in unwanted output ports is distributed among an increasing number of ports, thereby lowering the maximum crosstalk level. This effect is not seen for the ideal coupler since imbalance and phase deviations (the principle sources of the crosstalk) are not present in this case.

The requirements for idealized switching are more stringent than for variable-ratio power splitting. Practical switches provide crosstalk levels in the –20 to –30 dB range, and can only be improved with very accurate device fabrication and control methods. To exploit the simplicity of the idealized phase shifts for switching, it is thus imperative that the device response be as close as possible to the ideal response.

4.4 Multi-port blocking switches

When light is introduced into multiple input ports, the GMZI provides a limited $N \times N$ switching capability. This can be intuitively understood by considering the symmetric nature of the (ideal) GMZI. The GMZI structure shown in Fig. 2.7 is the same when viewed upside-down or from the front or behind the page. Physically, this means that input ports and output ports can be interchanged simply by launching light into the different ends of the device, and a port i can be considered to be port $(N + 1 - i)$, depending upon the direction from which the GMZI is viewed. Mathematically, this can be seen from (3.10); the transfer matrix for the GMZI is unchanged when i is replaced by k or when i and k are replaced by $(N + 1 - i)$ and $(N + 1 - k)$.

This leads to the conclusion that if the state (i, k) is achieved by applying a certain set of phase shifts, then the same set of phase shifts result in the state (k, i). If $i \neq k$, the switch is capable of supporting multiple simultaneous states (cross-connects). In this section the cross-connects available in the GMZI are derived from the GMZI transfer matrix. It is shown that the same phase shifts and cross-connects which govern this switch also apply to a wavelength division multiplexing (WDM) version of the switch, enabling both wavelength multiplexing and sorting functions to be performed with a single device.

4.4.1 Cross-connect pattern of the $N \times N$ GMZI

When the $N \times N$ GMZI is used as a $1 \times N$ switch, a single input can be cross-connected to any output using a set of optimum phase shifts given by the columns of the matrix X. In the case of multiple inputs, each column of X corresponds to a cross-connection of all input and output ports of the $N \times N$ GMZI in pairs. A different set of cross-connects is associated with each column of X, so an $N \times N$ GMZI has N distinct cross-connect sets. To determine these sets, consider the expression (3.13) for the intensities at the outputs of the GMZI, derived from the GMZI transfer matrix. A cross-connect (i, k) occurs when the argument of the cosine is an integer multiple of 2π, i.e.,

$$\left(\varphi_{ip} - \varphi_{iq}\right) + \left(\Delta\varphi_p - \Delta\varphi_q\right) + \left(\varphi_{pk} - \varphi_{qk}\right) = 2m\pi \qquad (4.6)$$

where m is an integer. It is easily verified that when (4.6) is substituted into (3.13), $I_k^{out} = I_0$. Considering the phase shift in arm $p = 1$ as a reference and setting $\Delta\varphi_1 = 0$ as a reference without affecting the generality of the analysis, (4.6) becomes

$$\Delta\varphi_q = \varphi_{i1} + \varphi_{1k} - \left(\varphi_{iq} + \varphi_{qk}\right) - 2m\pi \qquad (4.7)$$

where $\Delta\varphi_q$ are the phase shifts needed to cross-connect input i to output k. Since all input ports and output ports are cross-connected in pairs, the phase shift $\Delta\varphi_q$ must also connect an input i' to an output k', where $i', k' \in \{1, 2, \ldots, N\}$. Without loss of generality, we may choose $i' = 1$, which yields

$$\Delta\varphi_q = \varphi_{11} + \varphi_{1k'} - \left(\varphi_{1q} + \varphi_{qk'}\right) - 2m\pi \qquad (4.8)$$

Equating (4.7) and (4.8) and substituting for φ using (3.2), after some manipulation we obtain:

$$\left[(-1)^i(2i-1) + (-1)^k(2k-1) - (-1)^{k'}(2k'-1) + 1\right] \bmod(4N) = 0 \qquad (4.9)$$

To find the output port k' to which input port 1 is cross-connected, we define a useful function ψ as follows:

$$\psi_{i,k} \equiv (-1)^i(2i-1) + (-1)^k(2k-1) + 1 \qquad (4.10)$$

We note from (4.10) the symmetry relation, i.e., the cross-connect set for (i, k) and (k, i) are the same. Note also that $\psi_{i,k}$ always lies in the range $-4N < \psi_{i,k} < 4N$. From (4.9) and (4.10) an analytic expression for k' is obtained:

$$k' = \frac{\left|\psi_{i,k}\right| + 1}{2} \qquad \text{for } \left|\psi_{i,k}\right| < 2N$$

$$k' = \frac{4N - \left|\psi_{i,k}\right| + 1}{2} \qquad \text{for } 2N \le \left|\psi_{i,k}\right| < 4N \qquad (4.11)$$

The complete set of cross-connects for the $N \times N$ GMZI switch can be determined by substituting any (i, k) pair into (4.10), and using (4.11) to determine the output port k' to which input port 1 (i') is cross-connected. As an example, Table 4.2 shows the set of cross-connects possible in the 5×5 GMZI switch.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)
(2,3)	(2,1)	(2,5)	(2,2)	(2,4)
(3,2)	(3,4)	(3,1)	(3,5)	(3,3)
(4,5)	(4,3)	(4,4)	(4,1)	(4,2)
(5,4)	(5,5)	(5,2)	(5,3)	(5,1)

Table 4.2 The set of cross-connects in the 5 × 5 GMZI switch. A cross-connect of an input i to an output k is given by the pair (i, k), and each column represents the simultaneous cross-connects in the switch when N inputs are used. The states in each column are achieved by applying the phase shifts in the corresponding column of (4.3) to the GMZI arms.

4.4.2 Tunable $N \times N$ MMI Phasar Multiplexer

The $N \times N$ GMZI shown in Fig. 2.7 has arms with equal lengths. Allowing the arm lengths to differ results in relative phase delays among the self-images that are proportional to the arm length differences. In this case the optical power at each input to the combiner is distributed to all outputs k, where the relatively delayed optical fields interfere. If the arm lengths are chosen so that constructive interference occurs for different wavelengths at different outputs of the combiner, then the device is capable of passively multiplexing or de-multiplexing N wavelengths. The set of arms with differing lengths are referred to as a phased array, or phasar [57], and the entire device is called an MMI-phasar multiplexer [27]. The concept of this device is shown in Fig. 4.4.
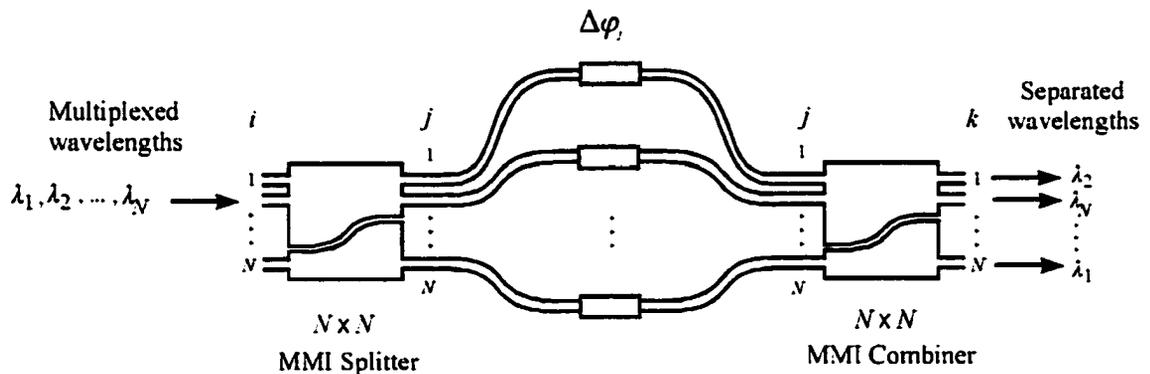
Figure 4.4 General layout of a tunable $N \times N$ MMI phasar multiplexer. The arms have different lengths, chosen to achieve wavelength separation at the outputs. The phase shifting regions are used to change the output wavelength assignment.

The phase shifting elements which allow the operation of the $N \times N$ GMZI switch now enable the MMI phasar multiplexer to operate in an active, or tunable mode [58]. Once the arm lengths are chosen such that a multiplexed signal at an input is passively demultiplexed, different wavelengths can be made to emerge at an output k by adjusting the phase shifters to achieve the switching states of the GMZI switch. Each column of X therefore corresponds to a different wavelength pattern at the MMI phasar multiplexer outputs. The individual wavelength channels are denoted by λ_b, where b is an integer in the range $1 \leq b \leq N$, and the wavelength increases with the subscript; that is, $\lambda_{b-1} - \lambda_b = \Delta\lambda$.

As an example, consider the tunable 5 × 5 MMI phasar multiplexer. The phase shifts required to achieve the switch states of the 1 × 5 GMZI switch are given by the columns of the ideal switching matrix X in (4.3). Now instead of switching the signal, in the MMI phasar device these switch states result in a different ordering of the demultiplexed wavelength channels at the output ports, for a fixed input. The resulting wavelength channel positions can be found using (4.11) and have been given in Table 4.2, however the interpretation of the columns must be made clear. For a fixed input, each column in Table 4.2 corresponds to a different wavelength. For example using the design described in [27] when $i = 1$, column 1 of Table 4.2 corresponds to the wavelength λ_2. Then (1,1) indicates λ_2 is sent to port 1 when the switch state corresponds to (1,1). The next state (2,3) means that λ_2 is sent to port 3 when the switch state corresponds to (1,2), the next state (3,2) will send λ_2 to port 2 when the switch state corresponds to (1,3), etc. The resulting wavelength routing table is derived in this manner, and given as Table 4.3.

| Output | 1 × 5 GMZI switch state | | | | |
k	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)
1	λ_2	λ_5	λ_4	λ_3	λ_1
2	λ_5	λ_3	λ_2	λ_1	λ_4
3	λ_4	λ_2	λ_1	λ_5	λ_3
4	λ_3	λ_1	λ_5	λ_4	λ_2
5	λ_1	λ_4	λ_3	λ_2	λ_5

Table 4.3 Wavelength routing table for the tunable 5 × 5 MMI phasar multiplexer for $i = 1$. Each switch state of the 1 × 5 GMZI is denoted by a column, and corresponds to a re-ordering of the wavelength positions at the output ports.

4.5 Phase shifting requirements in the GMZI

Returning for a moment to Fig. 4.2, note that as N increases, the number of distinct phase shift levels required for switching also increases. Beyond the simplest cases (i.e., $N > 4$), the control circuitry for the switches becomes increasingly complex. When N is large, the added control requirements result in switches that are physically larger, require more power to operate, and are less economical to produce.

A method to improve the control requirements of the 1 × N GMZI switch is therefore sought. To find this, the origins of the switching phase shifts given by X need to be investigated more closely. The patterns in (4.2) and (4.3) are not random, and since the GMZI has been shown to have multiple symmetries, it may be postulated that the switching phase shifts also contain symmetries. These could be used to relax the control requirements of the switch.

Symmetry in the GMZI is investigated by first considering the operation of the passive device. This yields insights that explain the active operation of the switch, thereby allowing improved designs for active GMZI-based switches to be developed.

4.5.1 Passive GMZI

Consider the $N × N$ GMZI shown in Fig. 2.7, in the passive case where the phase shifters are not operated. The M-th N-fold imaging length of each MMI coupler given by (2.25) is repeated here for convenience:

$$L_{M,N} = \frac{M}{N} 3L_\pi \qquad (4.12)$$

where L_π is the beat length of the coupler given by (2.23), and M and N are any two positive integers without a common divisor. The shortest coupler length is obtained when $M = 1$.

Since all arms in the GMZI have equal lengths, no relative phase shift is imparted to the light in the arms of the passive device. The presence of the arms may therefore be disregarded entirely, resulting in the merging of the two MMI couplers into a single MMI coupler with a length equal to twice the original coupler length:

$$L_{GMZI} = \frac{M}{N} 3L_\pi + \frac{M}{N} 3L_\pi = \frac{M}{N} 6L_\pi \qquad (4.13)$$

Rearranging (4.13), two alternate expressions for the GMZI length can be found,

$$L_{GMZI} = \frac{2}{N} 3L_\pi \qquad \text{for } N \text{ odd,} \qquad (4.14)$$

where $M = 1$ has been used in (4.13), and

$$L_{GMZI} = \frac{M}{(N/2)} 3L_\pi = \frac{M}{K} 3L_\pi \quad \text{for } N \text{ even,} \qquad (4.15)$$

where

$$N = 2K, \qquad K = 1,2,3,\ldots, \qquad (4.16)$$

and M and K are any two positive integers without a common divisor. Comparing (4.12) and (4.14) it is clear that when N is odd, the passive GMZI acts like a single MMI splitter with $M = 2$ and N output images. By contrast, upon comparison of (4.12) and (4.15) for N even, the passive GMZI functions as a single MMI splitter with a reduced number (K) of output images.

In both cases, there are N output ports. For N odd, the action of the passive GMZI is trivial: the N output images emerge from the N output ports, since (4.14) is identical to (4.12) if $M = 2$ is used in (4.12). Since this does not yield any new information, our analysis will be restricted to the case of N even. When N is even, there are N output ports and only K images. The output ports from which the light emerges will depend upon both the input port and the value of K. Two cases must therefore be considered: K even and K odd.

A. K even

Since the passive GMZI behaves as an MMI coupler with K output ports, the geometry of an MMI coupler with an even number of output ports is analyzed. This geometry has been described by Bachmann et al. [12], and is shown in Fig. 4.5.

The spacing of the output images by intervals of length $2W_e/N$ in the $N \times N$ coupler are in the $N \times K$ coupler described by intervals of length:

$$\frac{2W_e}{K} = \frac{4W_e}{N} \qquad (4.17)$$

where (4.16) has been used, and W_e is the effective width of the coupler, defined by (2.13). The corresponding geometry for the $N \times K$ coupler is shown in Fig. 4.6.
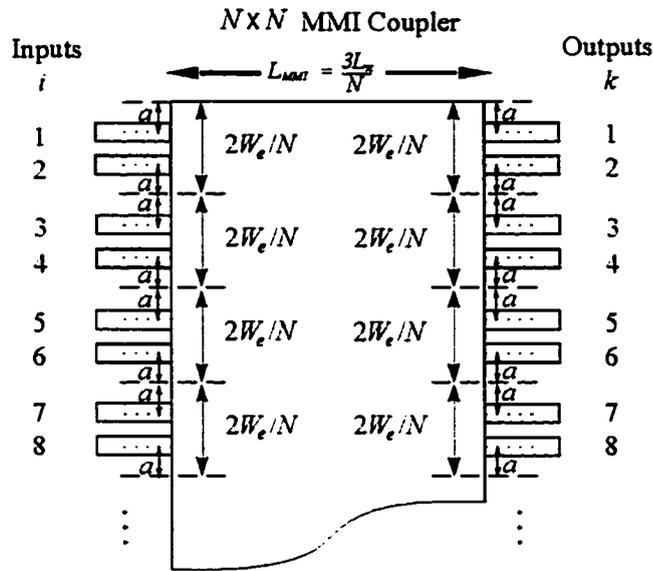
Figure 4.5 Geometry of the $N \times N$ MMI coupler illustrating the input and output port positions for the case where N is even.

In the $N \times N$ coupler of Fig. 4.5, the first input port ($i = 1$) is located a distance a (a free design parameter) from the top edge of the coupler. The corresponding output images are located the same distance a inside both edges of every interval of length $2W_e/N$. The result is N self-images at the outputs.
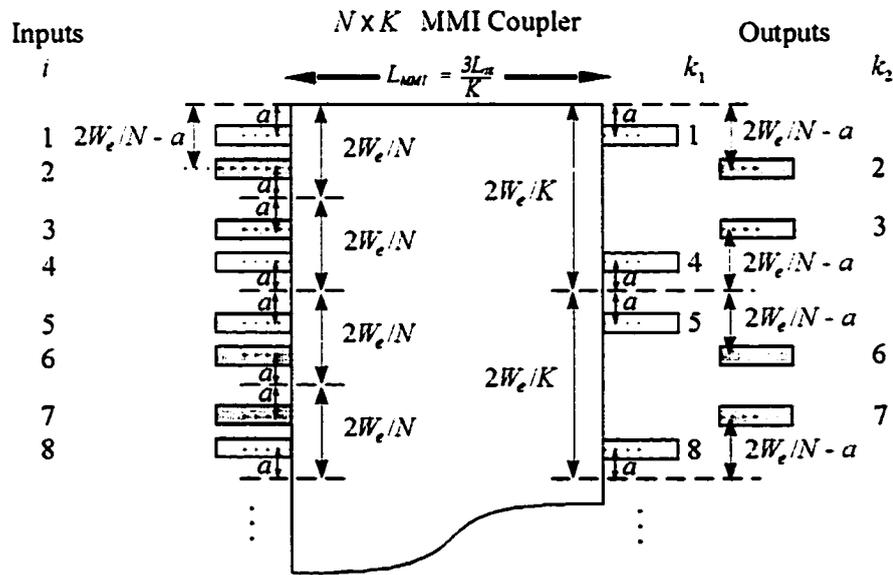
Figure 4.6 Geometry of the $N \times K$ MMI coupler illustrating the input and output port positions for the case where K is even.

In the $N \times K$ coupler, the same relation holds; however, the interval lengths are given by (4.17). Observing Fig. 4.6, the result is that for an input into port $i = 1$, K self-images emerge from the set of output ports:

$$k_1 = \{1,4,5,8,9,12,13, ...\} \tag{4.18}$$

which are located a distance a inside both edges of every interval of length given by (4.17).

Similarly, for an input into port $i = 2$ located a distance $2W_e/N - a$ from the top edge of the coupler in Fig. 4.5, the output images are located the same distance inside both edges of every interval of length $2W_e/N$, again corresponding to N self-images at the outputs.

In the $N \times K$ coupler of Fig. 4.6, however, input port $i = 2$ corresponds to the set of output ports:

$$k_2 = \{2,3,6,7,10,11,14, ...\} \tag{4.19}$$

which represents the complement of the set of ports given by k_1.

Continuing in the same manner, it is confirmed that for the $N \times N$ coupler, light from each input port forms self-images at all N outputs. In the $N \times K$ coupler, however, light from an input port in the set k_1 forms K output images in the set k_1, and light from an input port in the set k_2 forms K output images in the set k_2. This can be written more compactly as:

$$i \in \{k_1\} \rightarrow k = \{k_1\}$$
$$i \in \{k_2\} \rightarrow k = \{k_2\} \tag{4.20}$$

This indicates that when K is even, the passive GMZI functions as two superimposed $K \times K$ MMI couplers, each with input and output ports in the same set. This is shown schematically in Fig. 4.7. The validity of (4.20) is confirmed using both MPA and transfer matrix techniques.
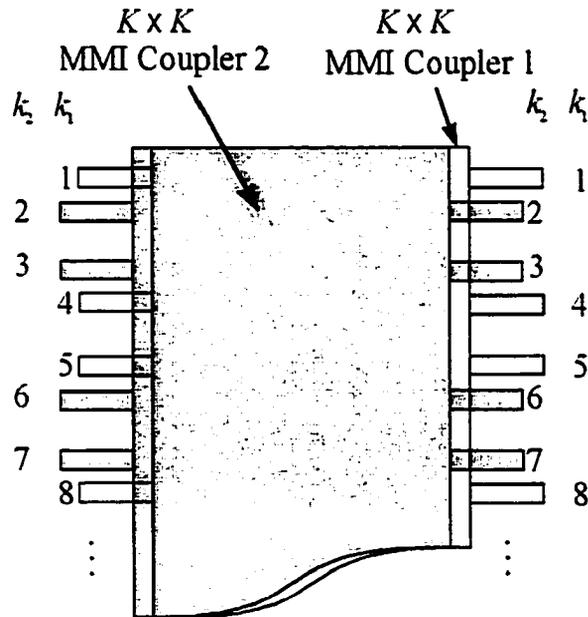
Figure 4.7 When K is even, the passive $N \times N$ GMZI functions as two superimposed $K \times K$ MMI couplers with input and output ports in the same port set.

B. K odd

Now let us consider the case where K is odd. The geometry of an $N \times N$ MMI coupler with an odd number of output ports is shown in Fig. 4.8. The spacing of the N output images are described by intervals of length $2W_e/N$, except for the first interval which has a length W_e/N.
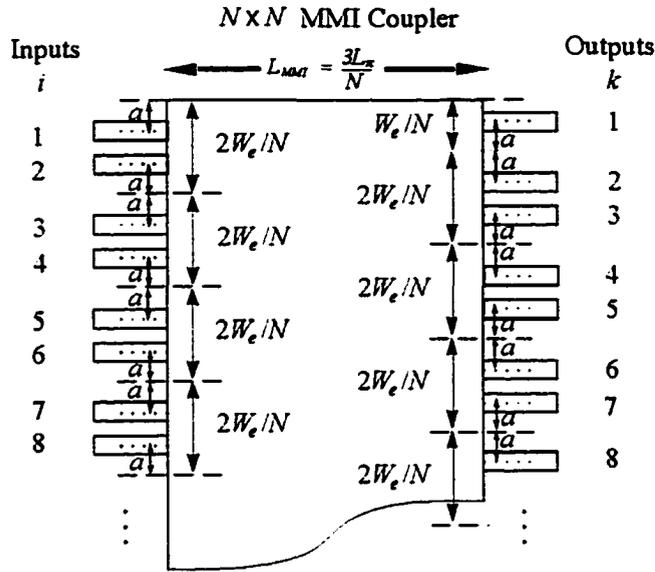
Figure 4.8 Geometry of the $N \times N$ MMI coupler illustrating the input and output port positions for the case where N is odd.

For the corresponding $N \times K$ coupler shown in Fig. 4.9, the spacing of the K images are described by intervals of length given by (4.17), except for the first interval, which has a length:

$$\frac{W_e}{K} = \frac{2W_e}{N} \qquad (4.21)$$

Similar to the case of K even, the positions of the input and output images are again described by the free parameter a. In this case, however, the presence of an interval with a length
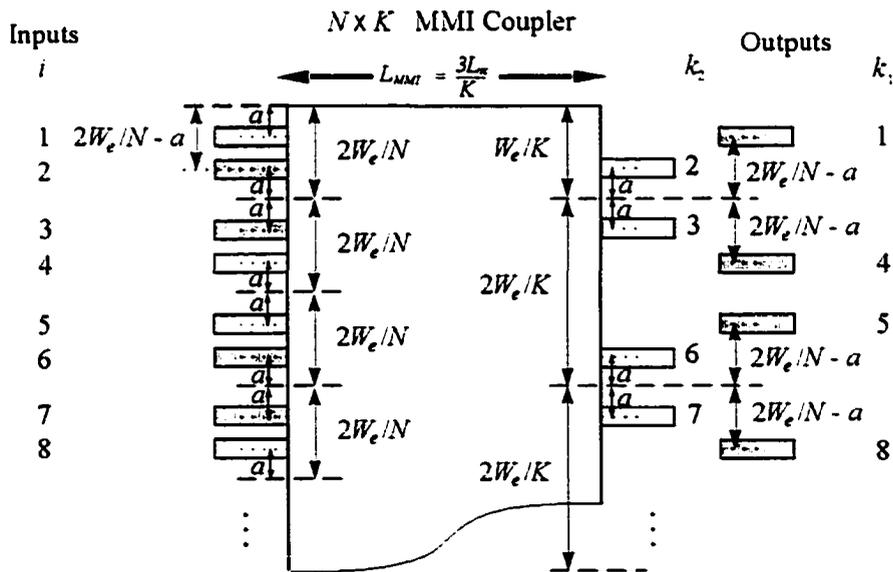
Figure 4.9 Geometry of the $N \times K$ MMI coupler illustrating the input and output port positions for the case where K is odd.

given by (4.21) results in light from inputs in the set k_1 forming K output images in the set k_2, and light from inputs in the set k_2 forming K output images in the set k_1, or

$$i \in \{k_1\} \rightarrow k = \{k_2\}$$
$$i \in \{k_2\} \rightarrow k = \{k_1\}$$

(4.22)

indicating that when K is odd, the passive GMZI functions as two superimposed $K \times K$ MMI couplers, each with input and output ports in different sets. This is shown in Fig. 4.10, and is verified using both MPA and transfer matrix methods.
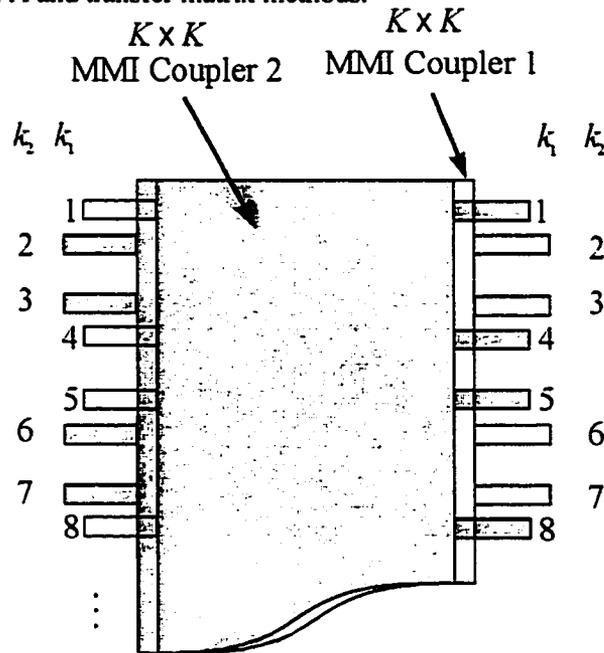
Figure 4.10 When K is odd, the passive $N \times N$ GMZI functions as two superimposed $K \times K$ MMI couplers with input and output ports in different port sets.

Realizing that the passive GMZI functions as two smaller MMI couplers occupying the same physical region now permits an intuitive understanding of the active device behavior.

4.5.2 Active GMZI

The active GMZI has been shown to function as a $1 \times N$ switch by applying N sets (columns of the matrix X) of N phase shifts (rows of X) to the arms. As N increases, the number of distinct phase shift levels needed increases; however, a simplification of the phase shifts is possible. When N is even, the phase shift sets can be divided into two groups of K sets each. Let us call one group "simple" and the other "complex". The sets in the simple group possess one or more of the following properties:

1) The phase shifts in each set in the group are symmetric in pairs about a central bisection of the N phase shifts, resulting in only K independent phase shifts (i.e., the two outermost phase shifts (on arms 1, N) are equal, and so on, until the two innermost).
2) The resulting K sets of K independent phase shifts are themselves anti-symmetric in pairs, (i.e., phase shift 1 from one set in a pair is equal to phase shift K from the other, etc).
3) The phase shifts in the group have the smallest maximum phase shift of any of the N sets.
4) The group has the smallest number of distinct phase shift values of any of the N sets.

Specifically, for K even, the simple group possesses at least the first two properties, and for K odd, the simple group possesses all four properties. By contrast, the complex group possesses properties opposite to those stated above: they are in general neither symmetric nor anti-symmetric, nor do they necessarily have the smallest maximum phase shift or the smallest number of distinct values of any of the sets. These properties are best illustrated through the use of examples, which will be given in the following section.

Since each set of phase shifts cross-connects an input i to an output k, the simple and complex groups correspond to groups of simple and complex output ports. These can be interpreted as the output ports which can be accessed using simple phase shifts and complex phase shifts, respectively. The simple and complex port numbers can be compared to the analysis of the passive GMZI to arrive at the following design rule:

Cross-connecting light from any of the inputs of a given $K \times K$ coupler in a GMZI to any of the outputs of the same $K \times K$ coupler can be accomplished using simple phase shifts, while cross-connecting light from any of the inputs of a given $K \times K$ coupler to any of the outputs of a different $K \times K$ coupler requires complex phase shifts.

This rule is a simple, yet powerful tool which can be used to dramatically improve the control requirements of $1 \times N$ GMZI switches. This is described in the following section.

4.6 Improved $1 \times N$ Switch

4.6.1 Theory and Design

The above design rule states that the simple and complex outputs correspond to either of the sets k_1 or k_2. Then for a single input and when N is even, the $N \times N$ active GMZI can be operated as a $1 \times K$ switch, using only simple phase shifts. If, however, the restriction of a single input can be removed - allowing the capability of switching between inputs from different sets - then the alternate K outputs can also be accessed. The result is a $1 \times N$ switch that can be operated using only the K sets of simple phase shifts. This is the principle of the improved $1 \times N$ switch.

In practice, the restriction of a single input is removed by utilizing a 1×2 switch to connect a single input to any two GMZI inputs from different sets, corresponding to the set numbering given by (4.18) and (4.19). Any desired output is then accessed with simple phase shifts, by operating the 1×2 switch such that the input port of the GMZI corresponds to the same $K \times K$ coupler as the output port. This structure is shown schematically in Fig. 4.11.
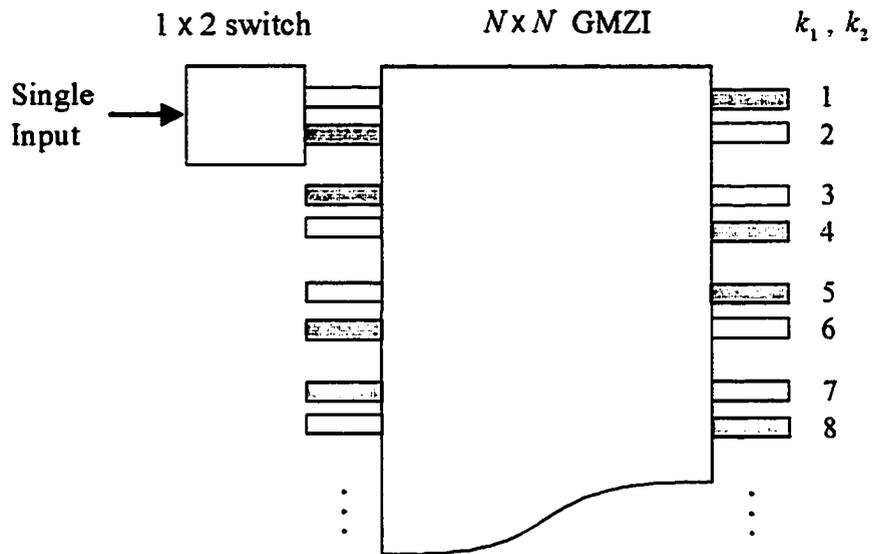
Figure 4.11 Improved 1 × N switch design. The case where K is odd is shown. Ports with a similar shading belong to the same K × K MMI coupler. By switching a single input between GMZI inputs from different port sets, the input can be cross-connected to any output port using only simple phase shifts.

The improved 1 × N switch has a number of advantages over the standard 1 × N GMZI switch. Regardless of the value of K, the simple sets of phase shifts are symmetric, as stated in property 1, above. This means that the phase shifters or their controllers can be physically connected in pairs, so that only K rather than N independent phase shifters are needed (an additional phase shifter is needed to operate the 1 × 2 switch if it is an integrated switch).

From property 2, the anti-symmetric nature of the K sets of K phase shifts results in a further relation between the phase shifts. It will be seen that this relation can be used to further simplify the switch control logic. Furthermore, when K is odd or in certain cases when K is even, from properties 3 and 4, the phase shift magnitudes are minimized, and the number of distinct phase shift values is smaller than in the standard switch.

These concepts are illustrated by means of two examples.

4.6.2 Examples of Improved 1 × N switch operation

4.6.2.1 Improved 1 × 4 Switch

The smallest improved 1 × N switch occurs when N = 4. Following the procedure described in Section 4.2, the ideal switching matrix for an input into port i = 1 is found:

$$X = \frac{\pi}{2}\begin{bmatrix} 0 & -1 & -1 & 1 \\ 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

(4.23)

In this case K is even. Using (4.20) and the design rule above, i belongs to the set k_1, so the simple group corresponds to the outputs k_1, or columns 1 and 4 in (4.23). It is easily verified that these sets satisfy at least the first two properties stated above. The phase shifts in columns 1 and 4 are mirrored in a horizontal line drawn between rows 2 and 3. The resulting two independent phase shifts in columns 1 and 4 (given in bold) are anti-symmetric with respect to

each other. Since K is even, properties 3 and 4 are in general not satisfied, however, in this case both properties are satisfied.

The simple phase shifts can cross-connect the input $i = 1$ to either output 1 or 4. From (4.20), however, the remaining output ports 2 and 3 can be accessed simply by changing the input to a member of the set k_2. For simplicity, we choose the input $i = 2$. Using the analytic result for the cross-connects given in (4.11), it can be shown that the same set of phase shifts corresponding to column 1 in (4.23), or the cross-connect pair $(i, k) = (1,1)$ also yield the cross-connect pair (2,3). Similarly, column 4 in (4.23) corresponds to the cross-connect pairs (1,4) and (2,2).

All four states of the 1×4 switch are therefore accessible if a 1×2 switch is used to switch between the input ports 1 and 2. The operation of this improved 1×4 switch is shown in Fig. 4.12. As a consequence of the symmetry properties, the 4×4 GMZI can be controlled using only a single set of two phase shifts. If this set is represented by a logic level '0' or '1', then control of the 4×4 GMZI portion of the switch is possible using a single binary bit. An additional binary bit is used to control the 1×2 switch, so the improved 1×4 switch shown in Fig. 4.12 can be controlled with a single 2-bit binary word.

Required 1×2 GMZI phase shift $(\times \pi)$ ΔP	0	1
Resulting 1×2 GMZI states	(1,2) CROSS	(1,1) BAR
Binary control bit A	0	1

Required 4×4 GMZI phase shifts $\left(x\frac{\pi}{2}\right)$	ΔP_1	0	1
	ΔP_2	1	0
	ΔP_2	1	0
	ΔP_1	0	1
Resulting 4×4 GMZI states		(1,1) (2,3)	(1,4) (2,2)
Binary control bit B		0	1
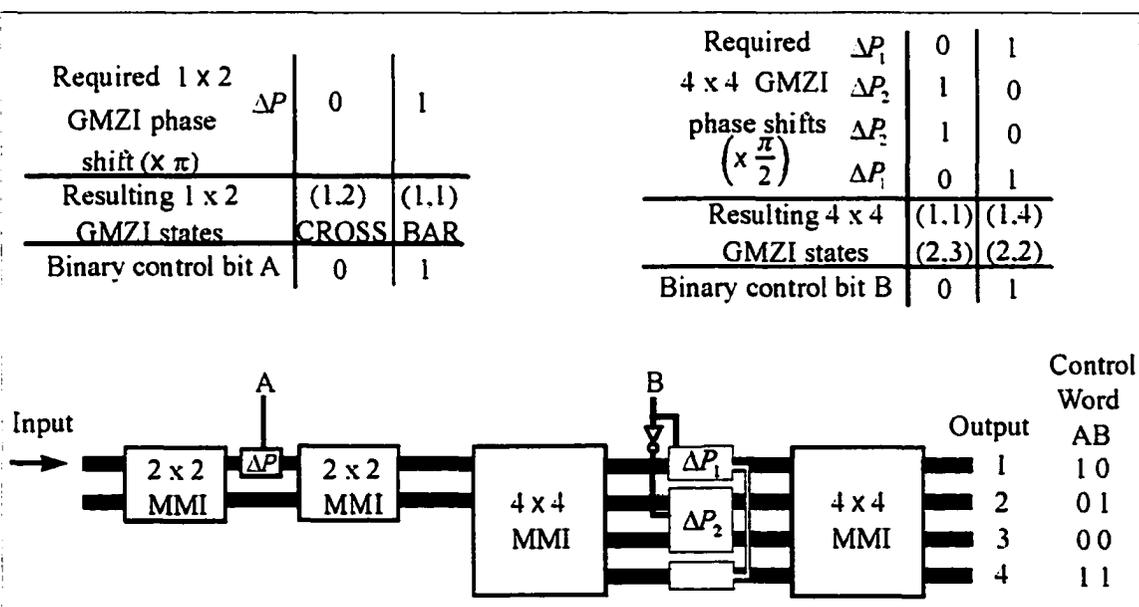
Figure 4.12 Operation of the improved 1×4 switch. The 4×4 GMZI is controlled using a single binary bit B. An additional bit A is required to control the 1×2 switch. A cross-connect to any output is therefore possible by applying a two-bit control word AB to the device. The derivation of the control logic is given in tabular form.

4.6.2.2 Improved 1×6 Switch

In this case K is odd, so from (4.22) or Fig. 4.10 it can be concluded that with an input into port $i = 1$, cross-connects to outputs $k = 2,3,6$ are possible using simple phase shifts. Using (4.11) it is found that these same phase shifts will cross-connect input port $i = 2$ to the outputs $k = 1,4,5$. Since K is odd, all four properties of the simple group above apply. From property 1, only half of the phase shifts in the three simple sets from the matrix X are required,

$$X = \frac{\pi}{3} \begin{bmatrix} 1 & -1 & -1 & 3 & 3 & 1 \\ 2 & -1 & 1 & 2 & 0 & -1 \\ 2 & 1 & -1 & 0 & 2 & -1 \\ -1 & 1 & -1 & -3 & -1 & -1 \\ -1 & -1 & 1 & -1 & -3 & -1 \\ -2 & -1 & -1 & 0 & 0 & 1 \end{bmatrix} \qquad (4.24)$$

which are given in bold. The first and third columns in bold satisfy property 2, namely, they are anti-symmetric with respect to each other. The simple phase shifts in (4.24) also satisfy properties 3 and 4, namely, the maximum phase shift magnitude ($\pi/3$) is the smallest maximum of the six sets, and the phase shifts have the minimum number of distinct values (two) of any of the six sets.

Simplified digital control is possible, since two binary bits are needed to control the 6 × 6 GMZI when the phase shifters are connected with their mirror opposites. Again only a single bit is needed to control the 1 × 2 switch, so the improved 1 × 6 switch can be controlled with a single 3-bit binary word. A diagram of this switch is given in Fig. 4.13. Note that a value of $\pi/3$ has been added to the phase shifts in bold in (4.24) to convert them to logic levels '0' and '1'.

Larger improved 1 × N switches are designed in a similar manner, and generally require control circuitry with a complexity characteristic of a 1 × K switch.

Required 6 × 6 GMZI phase shifts $(x\frac{2\pi}{3})$		0	0	1
ΔP_1		0	0	1
ΔP_2		0	1	0
ΔP_3		1	0	0
Resulting 6 × 6 GMZI states		(1,2) (2,1)	(1,3) (2,5)	(1,6) (2,4)
Binary control BC		10	00	01

Output	Control Word ABC
1	0 1 0
2	1 1 0
3	1 0 0
4	0 0 1
5	0 0 0
6	1 0 1
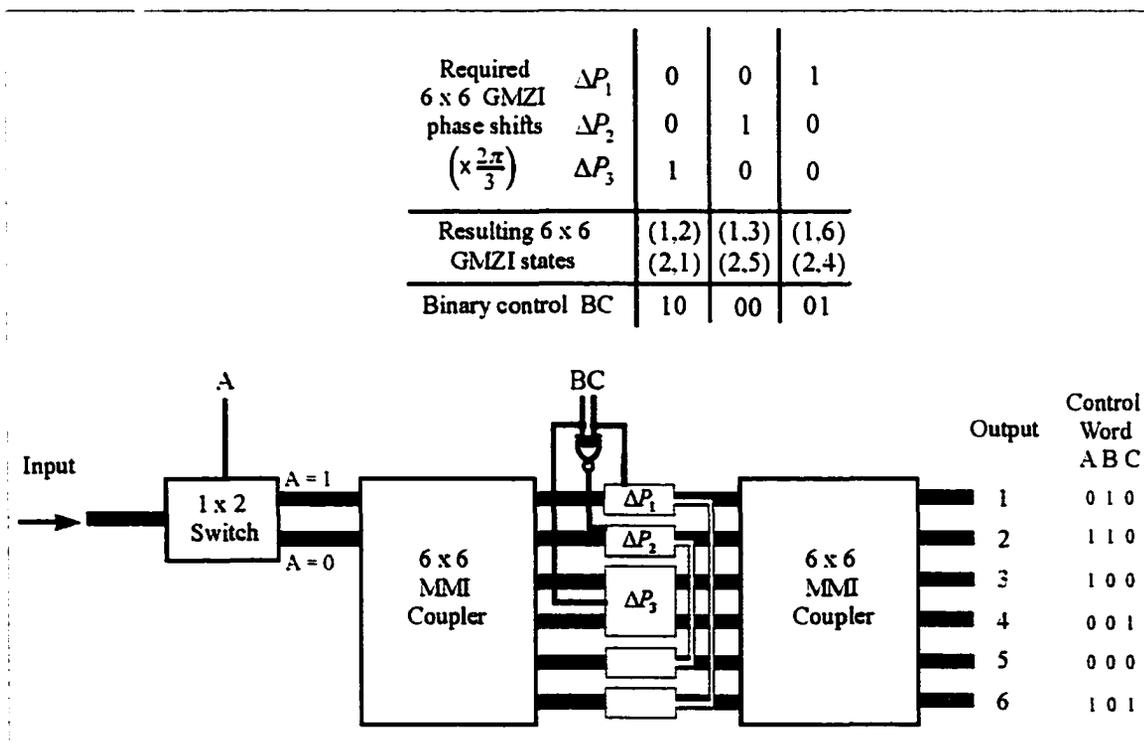
Figure 4.13 Operation of an improved 1 × 6 switch using a three-bit control word ABC.

4.7 N × N non-blocking switching

The N cross-connect states of the N × N GMZI given by (4.11) provide a limited multi-port switching function which was shown in the previous section to be useful in the development of improved 1 × N switches. The distinct cross-connect configurations are limited, however, to

only N possible permutations of the inputs to be achieved at the outputs. An example of these permutations for the case of $N = 5$ was given in Table 4.2. Such switches, with limited sets of cross-connect permutations, are commonly referred to as *blocking* switches. To understand this term, consider the situation in the switch where some input and output ports are cross-connected (occupied) and others are free. Free input ports cannot be cross-connected at will to free output ports, because their paths are set by the phase shifts currently being applied to the arms. Some cross-connects are therefore blocked, owing to the limited number of permutations available.

In a *non-blocking* switch architecture, however, the entire set of $N!$ permutations of cross-connects are possible and thus any two free ports can be cross-connected while preserving the cross-connect pattern of the occupied ports. Non-blocking switches are therefore of prime importance in network applications, where the maximum flexibility in cross-connect possibilities is desired [14].

Dilated versions of non-blocking switches have been proposed, in which a number of independently controlled $1 \times N$ GMZI switches are used [33]. This design, however, requires $4N$ MMI couplers, $2N^2$ phase shifters, and numerous waveguide bends and crossings, resulting in a large switch with complicated design and control requirements.

It has been shown in the previous section that using a smaller (1×2) switch in conjunction with the $N \times N$ GMZI greatly increases the flexibility of the $1 \times N$ switch. In a similar manner, it is shown in this section that smaller GMZIs can be used in conjunction with larger GMZIs to achieve $N \times N$ non-blocking operation in a simple and compact structure. Using this idea, two designs for the non-blocking switch are proposed: a cascaded design and a balanced design.

The tradeoff for the simplified non-blocking designs mentioned in this section is that the switches are only *rearrangably non-blocking*, meaning that any free pair of ports can be cross-connected only by momentarily disrupting the existing cross-connects. This arises because the multiple cross-connects in the $N \times N$ GMZI depend upon the state of a single phase shifter. Although numerous phase shifters are used in the dilated switch mentioned above, each phase shifter is associated with only one cross-connect. The result is that the dilated switch is *strictly non-blocking* (free ports can be cross-connected without disrupting existing cross-connects). Strictly non-blocking switching may not be required, however, in certain network restoration and re-configuration applications where momentary disruptions may be tolerated. Furthermore, it has recently been shown [59] that *blocking* GMZI switches can be used as the building blocks of larger strictly non-blocking switch architectures.

4.7.1 Cascaded Design

The simplest GMZI is the 2×2 version, which is strictly non-blocking since a free input and output port can be connected without disrupting an existing cross-connect, and the full set of two permutations (bar and cross states) are attainable.

To create a 3×3 non-blocking switch, the six permutations listed on the left-hand side of Table 4.4 are required. The three columns list the output ports to which inputs 1, 2, and 3 are cross-connected, respectively. The 3×3 GMZI is capable of only three distinct output permutations, listed on the right-hand side of Table 4.4. We only require the further capability of switching the two latter outputs of the GMZI (second and third columns on the right-hand side) from ascending to descending order, and vice-versa, to obtain the entire set of six permutations. Fortunately, this is exactly the function provided by a 2×2 GMZI.

Output Permutations	3 × 3 GMZI States
1 2 3	
1 3 2	1 3 2
2 1 3	2 1 3
2 3 1	
3 1 2	
3 2 1	3 2 1

Table 4.4 Design of a 3 × 3 non-blocking switch. The left-hand side lists the cross-connect permutations needed, and the right-hand side lists the permutations available from a 3 × 3 GMZI. The three columns list the output ports to which inputs 1, 2, and 3 are cross-connected, respectively.

The 3 × 3 non-blocking switch can therefore be realized using a 3 × 3 GMZI with two adjacent output ports fed into a 2 x 2 GMZI. Control can be achieved using simple digital logic, as in the 1 × N switches. For instance, the 3 × 3 non-blocking switch can be controlled using a three-bit binary word.

To create larger non-blocking switches, the same approach as above can be extended. For instance, a 4 × 4 non-blocking switch is realized by cascading a 4 × 4 GMZI, a 3 × 3 GMZI, and a 2 × 2 GMZI. In general, $N - 1$ cascaded GMZIs are required to realize an N × N non-blocking switch, as shown in Fig. 4.14. Operation of the cascaded switch is straightforward. The first N × N GMZI is set such that the port to appear at the N-th position of the desired permutation appears at the N-th output of the first GMZI. Next, the port to appear at the $(N-1)$-th position is cross-connected to the $(N-1)$-th output port of the second GMZI, and so forth.
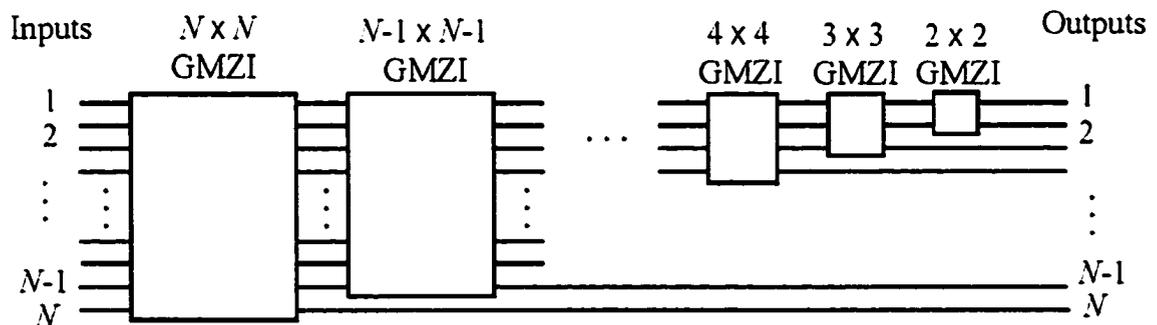
Figure 4.14 General layout of the N × N cascaded non-blocking switch. The switch is composed of smaller GMZIs arranged in descending order.

The cascaded design is efficient since the number of possible switch states is equal to the number of possible output permutations, which ensures that there is no redundancy in the design. This enables the smallest number of MMI couplers, phase shifters, and waveguide crossings to be used to achieve the $N!$ permutations. Specifically, $2N - 2$ MMI couplers, $(N^2 + N - 4)/2$ phase shifters, and no waveguide bends or crossings are needed.

A disadvantage of the cascaded design is that the paths traversed by the optical signals are not identical. The signal at the N-th output has traversed two MMI couplers, while the signal at the $(N - 1)$-th output has traversed four MMI couplers, and so forth, while the signal at the first output has traversed all $(2N - 2)$ MMI couplers. Since each MMI coupler has a finite excess loss,

the signal at the N-th output will be the strongest, while the signal at the first output will be the weakest.

Because this power imbalance can severely limit the usefulness of the switch, an alternative design is proposed.

4.7.2 Balanced Design

To balance $N \times N$ non-blocking switches requires removal of the restriction that smaller GMZIs are placed at only the output ports of the largest ($N \times N$) GMZI. By utilizing smaller switches at both the inputs and the outputs of the $N \times N$ GMZI, the permutations necessary for non-blocking operation can be achieved while maintaining a symmetric, balanced design.

To illustrate this concept, consider the 4 × 4 switch shown in Fig. 4.15. The switch is balanced, since each path from an input to an output traverses two 4 × 4 couplers and four 2 × 2 couplers.
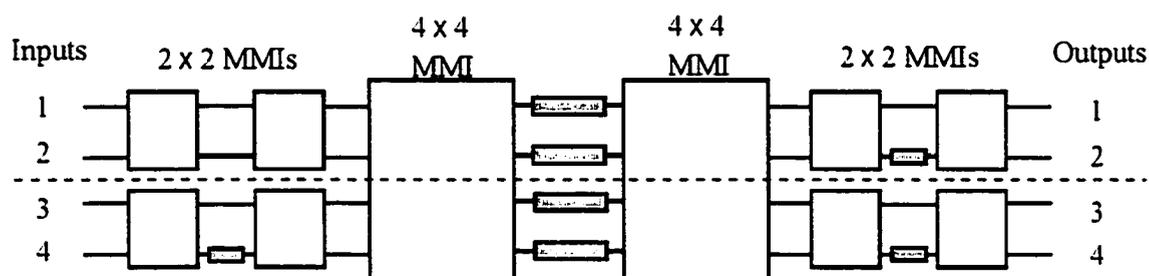
Figure 4.15 The 4 × 4 balanced non-blocking switch. The shaded regions indicate phase shifters. Three active 2 × 2 GMZIs and a central 4 × 4 GMZI allow $4 \cdot 2^3 = 32$ output permutations to be achieved, with redundancy. The location of the passive 2 × 2 GMZI is chosen arbitrarily.

Non-blocking operation is possible since the two 2 × 2 switches at the outputs extend the four permutations of the 4 × 4 GMZI to 16 permutations and an additional 2 × 2 switch at the input provides another factor of two, for a total of 32 permutations, which is greater than the 24 permutations required. In this case, three 2 × 2 switches are needed for non-blocking operation, however, four 2 × 2 switches are needed to maintain balance in the switch. The fourth switch is therefore passive. From symmetry considerations, the location of the passive 2 × 2 switch can be chosen arbitrarily among any of the four peripheral positions.

In this switch both balanced and non-blocking operation are achieved, but at the expense of increased complexity of control. In the 4 × 4 balanced switch, the 32 possible permutations are not unique, so there are numerous ways to achieve the 24 permutations of the 4 × 4 switch. Determination of the GMZI states required to achieve a given permutation is no longer trivial.

To determine how the balanced switch can be operated, it is necessary to first understand why it is capable of non-blocking operation. Consider in Fig. 4.15 the imaginary horizontal line (dashed line) that bisects the switch. This line bisects the 4 × 4 GMZI but does not cross any of the other components. The cross-connect states for the 4 × 4 GMZI found from (4.11) are given in Table 4.5, where the bisection of the 4 × 4 GMZI is again represented by a dashed line.

4 × 4 GMZI Switch State (columns)	(1,1) (2,3) (3,2) (4,4)	(1,2) (2,1) (3,4) (4,3)	(1,3) (2,4) (3,1) (4,2)	(1,4) (2,2) (3,3) (4,1)
Number of Crossovers	1	0	2	1

Table 4.5 The cross-connect states (i, k) for the 4 × 4 GMZI, where each state corresponds to a single column. The set of all possible crossovers of the inputs i from one side of the bisection line (dashed line) to the outputs k on the other side is obtained within the four GMZI states.

For non-blocking operation, the switch must be capable of cross-connecting every possible number of inputs from the top half of the switch to the bottom half, and vice-versa. This is because the 4 × 4 GMZI is the only region in the switch where inputs into the top half of the switch can be cross-connected to the bottom half, and vice-versa. The 2 × 2 switches are confined to the top or bottom half of the device, and therefore cannot perform this "crossover" function. In this example, the GMZI must be capable of 0,1, and 2 crossovers of inputs from the top half to outputs in the bottom half, and vice-versa. From the last row in Table 4.5 it is seen that the four states of the GMZI provide this full set of crossovers. This property of the GMZI results in the non-blocking operation of the balanced switch.

To operate the switch in Fig. 4.15, the required number of crossovers is determined and the 4 × 4 GMZI is set to the corresponding state in Table 4.5. The 2 × 2 GMZIs are then set in either the bar or cross state to achieve the desired permutation. In this manner an operation table for the 4 × 4 switch is derived. This is shown in Table 4.6.

Now returning our attention for a moment to the bottom row of Table 4.5, it is seen that another design is possible. Note that one of the four GMZI states is redundant, in that it results in the same number of crossovers as another state. We may therefore choose either of these states, and operate the 4 × 4 GMZI with only three states. The efficiency of this approach, however, is obtained at the expense of operating the peripheral switches. Because we no longer have the flexibility of all four possible 4 × 4 GMZI states, all four 2 × 2 GMZIs must now be capable of active operation.

Design of larger balanced non-blocking switches follows in a similar manner. A bisection line is first drawn through the $N \times N$ GMZI. The cross-connect capability of the GMZI ensures that every possible number of crossovers is achievable within the N states. The $N \times N$ balanced non-blocking switch is then formed by simply connecting the ports on both sides of the bisection line to $N/2 \times N/2$ non-blocking switches.

Output Port Permutation				4 × 4 GMZI State	2 × 2 GMZI States		
Port 1	Port 2	Port 3	Port 4		Lower Left	Top Right	Lower Right
1	2	3	4	(1,2)	CROSS	BAR	BAR
1	2	4	3	(1,2)	BAR	BAR	BAR
1	3	2	4	(1,4)	CROSS	CROSS	CROSS
1	3	4	2	(1,4)	BAR	CROSS	CROSS
1	4	2	3	(1,4)	CROSS	CROSS	BAR
1	4	3	2	(1,4)	BAR	CROSS	BAR
2	1	3	4	(1,2)	BAR	CROSS	CROSS
2	1	4	3	(1,2)	CROSS	CROSS	CROSS
2	3	1	4	(1,4)	CROSS	BAR	CROSS
2	3	4	1	(1,4)	BAR	BAR	CROSS
2	4	1	3	(1,4)	CROSS	BAR	BAR
2	4	3	1	(1,4)	BAR	BAR	BAR
3	1	2	4	(1,1)	BAR	BAR	BAR
3	1	4	2	(1,1)	CROSS	BAR	BAR
3	2	1	4	(1,1)	BAR	CROSS	BAR
3	2	4	1	(1,1)	CROSS	CROSS	BAR
3	4	1	2	(1,3)	CROSS	CROSS	CROSS
3	4	2	1	(1,3)	BAR	CROSS	CROSS
4	1	2	3	(1,1)	BAR	BAR	CROSS
4	1	3	2	(1,1)	CROSS	BAR	CROSS
4	2	1	3	(1,1)	BAR	CROSS	CROSS
4	2	3	1	(1,1)	CROSS	CROSS	CROSS
4	3	1	2	(1,3)	CROSS	CROSS	BAR
4	3	2	1	(1,3)	CROSS	BAR	BAR

Table 4.6 Operation table for the 4 × 4 balanced non-blocking switch in Fig. 4.15. The 4 × 4 GMZI state corresponds to the relevant column in Table 4.5, and is abbreviated by the first pair (i, k).

As an example of the balanced switch configuration for larger N, consider the 8 × 8 GMZI with the set of cross-connects given in Table 4.7. The ports are bisected as indicated by the dashed line, and the number of crossovers for each cross-connect state are counted. Since the ports are divided into groups of four, balanced 4 × 4 non-blocking switches can be connected to both sides of the 8 × 8 GMZI to obtain the necessary permutations for the 8 × 8 balanced non-blocking switch. The device is shown schematically in Fig. 4.16.

8 × 8 GMZI Switch State (columns)	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)	(1,8)
	(2,3)	(2,1)	(2,5)	(2,2)	(2,7)	(2,4)	(2,8)	(2,6)
	(3,2)	(3,4)	(3,1)	(3,6)	(3,3)	(3,8)	(3,5)	(3,7)
	(4,5)	(4,3)	(4,7)	(4,1)	(4,8)	(4,2)	(4,6)	(4,4)
	(5,4)	(5,6)	(5,2)	(5,8)	(5,1)	(5,7)	(5,3)	(5,5)
	(6,7)	(6,5)	(6,8)	(6,3)	(6,6)	(6,1)	(6,4)	(6,2)
	(7,6)	(7,8)	(7,4)	(7,7)	(7,2)	(7,5)	(7,1)	(7,3)
	(8,8)	(8,7)	(8,6)	(8,5)	(8,4)	(8,3)	(8,2)	(8,1)
Number of Crossovers	1	0	2	1	3	2	4	3

Table 4.7 The cross-connect states (i, k) for the 8 × 8 GMZI, where each state corresponds to a single column. The set of all possible crossovers of the inputs i from one side of the bisection line (dashed line) to the outputs k on the other side is obtained within the eight GMZI states.

Larger switches are more complicated to operate than in the 4 × 4 case; however a simple algorithm can be written to determine the individual GMZI states needed to achieve a desired output permutation. The algorithm must first determine which of the $N \times N$ GMZI states yields the number of crossovers indicated by the desired output permutation. Once this state is determined, operation tables for the smaller non-blocking switches can be used to arrive at the desired permutation. As mentioned previously, the permutations are not unique, and it may be possible to operate a number of the smaller component GMZIs in a passive mode.
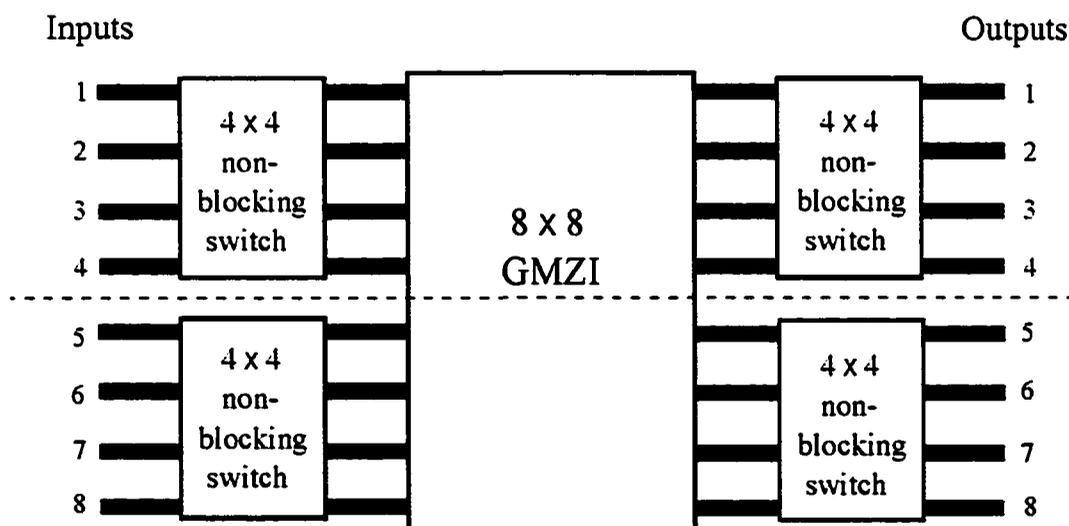
Figure 4.16 A balanced 8 × 8 non-blocking switch. Since the 8 × 8 GMZI is capable of every possible number of crossovers of the imaginary bisection line (dashed line), non-blocking operation is ensured by connecting the ports on either side of this line by smaller balanced non-blocking switches.

By considering the multi-port cross-connect capabilities of the $N \times N$ GMZI, it has been shown in this section that using combinations of larger and smaller GMZIs can result in non-blocking operation. Two different designs for non-blocking switches have been presented. In general, the cascaded approach yields switches which are simple to design and operate, but at the expense of having unbalanced losses. The balanced approach equalizes the switch losses, but at the expense of complex operation.

The concepts discussed in this section are generally applicable, independent of a particular waveguide or material technology and device design. In real devices, however, these factors will play a critical role in determining the ultimate performance of the non-blocking switches. It has been shown in Section 4.3 that non-idealities can severely limit the crosstalk attainable in a GMZI switch. It is therefore expected that when multiple GMZIs are used, the tolerances to parameter fluctuations are even more stringent. This is investigated further in the following section.

4.7.3 Tolerance analysis of a balanced non-blocking switch

To realize GMZI switches with good crosstalk performance and low optical losses, it is essential that the constituent MMI couplers impart accurate power splitting ratios and phases to the self-images. In this section the effects of MMI coupler imbalance and phase deviations are quantified, enabling a tolerance analysis of a 4 × 4 balanced non-blocking switch to be made.

The balanced 4 × 4 non-blocking switch shown in Fig. 4.15 is comprised of 2 × 2 and 4 × 4 GMZIs, each with a characteristic loss and crosstalk which is a function of the imbalance and phase deviation present in the MMI couplers. These effects arise from imperfect imaging in the coupler, which can occur when the fabricated length of the coupler differs from the designed value. It was shown in Chapter 3 that the length of an MMI coupler is proportional to the square of its width, so it is therefore critical to control the coupler width during the fabrication process [31].

The performance of 2 × 2 and 4 × 4 GMZIs for various deviations of the MMI coupler length has been simulated using MPA. While the analytic expression (3.32) can be used to find the crosstalk, the MPA approach has been chosen since it can also be used to determine excess losses. The details of the waveguide design used in the analysis will be given in the next chapter. The resulting worst-case channel loss and crosstalk values are given in Table 4.8, along with values for the 4 × 4 non-blocking switch, calculated by noting that each path through the switch traverses two 2 × 2 GMZIs and a 4 × 4 GMZI.

MMI Coupler Length Deviation	2 × 2 GMZI		4 × 4 GMZI		4 × 4 balanced non-blocking switch	
	Channel Loss (dB)	Crosstalk (dB)	Channel Loss (dB)	Crosstalk (dB)	Channel Loss (dB)	Crosstalk (dB)
$\delta L = 0 \ \mu m$	0.90	-26.8	-2.19	-22.0	3.99	-19.8
$\delta L = \pm 5 \ \mu m$	0.96	-23.9	-2.22	-19.3	4.14	-17.0
$\delta L = \pm 10 \ \mu m$	1.04	-17.4	-2.43	-13.3	4.51	-10.8

Table 4.8 Tolerance analysis for the component GMZIs and the aggregate 4 × 4 balanced non-blocking switch in Fig. 4.15, calculated using the MPA technique.

With the waveguide design chosen, the non-blocking switch is limited to a loss of 4 dB and a crosstalk level of -20 dB, even with optimum coupler lengths. This is the result of a small imbalance and phase deviation present even at the optimum coupler length, due to the finite number of modes available for imaging in the multimode region. When the imaging in the couplers degrades, the loss and crosstalk both increase. The situation becomes worse for larger switches, because the region of acceptable MMI coupler length deviation for a given excess loss and crosstalk shrinks as N increases, placing stricter requirements on the fabrication technology [31]. To realize non-blocking switches with good performance, it is therefore imperative that MMI couplers with precise imaging and low excess losses are used.

Chapter 5

Device Design

The theory of self-imaging in multimode waveguides and variable-ratio power splitting and switching in the GMZI have been presented in previous chapters. To verify this theory and demonstrate the practical feasibility of achieving the predicted device functions, a number of prototype devices have been designed. This chapter describes the device design and modeling performed at the Telecommunications Research Laboratories in Edmonton, Alberta, which culminated in the development of a photomask from which the devices can be fabricated. Portions of this chapter have been reported in Photonics Technology Letters *[35];* and in a paper presented at the *International Conference on the Applications of Photonic Technology (ICAPT '98), in Ottawa, Canada [60].*

5.1 Introduction

To verify the theoretical results developed in the earlier chapters, it is necessary for the GMZI to be designed and fabricated. Perhaps the most important determinant of the performance of an integrated optical device, and a necessary prerequisite for its design, is the choice of material system from which it is to be fabricated. From the growing number of technology possibilities in which to realize integrated optical devices, a silica-based waveguide system has been chosen, for reasons which will be discussed in this chapter.

A strongly guiding rib waveguide design has been used, and a number of structures such as single mode waveguides, 2 × 2 and 4 × 4 MMI couplers, and 2 × 2 and 4 × 4 GMZIs have been designed. The design requirements of these devices, the motivation for various choices and tradeoffs, the final design parameters chosen, and simulations of the device performance are described in this chapter. This work serves the basis for the development of a photomask, which is used to specify the device designs during the fabrication process. The fabrication of the integrated devices is the subject of the next chapter.

5.2 Choice of technology

Since the idea of integrated optics was first proposed in 1969 [6], extensive research has been performed on a variety of optical materials such as semiconductors, lithium niobate, silica-based glasses, and polymers, and to date much discussion exists over which system is the most suitable for integrated optical switching components. Each material system has its own advantages and drawbacks.

Semiconductors typically have a high refractive index, allowing ultra-compact devices to be made. Large-scale single-wafer integration of semiconductor switches with sources, detectors, and optical amplifiers is also possible. Furthermore, through the carrier-induced electro-optic effect, very fast active operation is achieved using a relatively low electrical power [61]. The drawbacks of semiconductor devices include high coupling losses to optical fiber, high material absorption losses, an expensive and complex material processing technology, and problems with reproducibility, long-term stability, and reliability [10].

Lithium niobate ($LiNbO_3$) exhibits lower losses than semiconductors, has very good electro-optic properties resulting in low power switching, has a very fast dynamic response, and exhibits useful nonlinear properties [8]; however the materials and processes involved make device fabrication difficult, the device size tends to be large, device stability and reliability are low, and coupling losses to optical fiber are still relatively high [10].

Silica-based waveguide technology is a promising alternative for integrated optical devices. Silica (SiO_2) is widely available, relatively inexpensive, has a very low transmission

loss, low fiber coupling loss, low polarization sensitivity, a simple and flexible fabrication technology, and a high stability and reproducibility [9], [10], [22]. Drawbacks of silica-based materials are larger device sizes compared with semiconductors, and active devices which are operated using the thermo-optic effect, which is weak in silica-based materials and results in relatively low-speed and high-power switching [5].

Polymer-based waveguide technology has been gaining considerable interest in recent years, due to the ease of fabrication (polymer layers can be spin-coated), lower power active switching compared with silica, and good coupling to standard single mode fibers [62]. The drawbacks are the high material losses in polymers compared with silica-based glasses, a large thermal expansion coefficient, and the unproven long-term stability and reliability of these materials.

In the present work, a very high switching speed is not required, because the reconfiguration function for optical networks will generally occur on a time scale of milliseconds [62]. Instead, achieving low-loss, low-power, polarization-insensitive operation for the given technology are more important, so the integrated optical devices will be fabricated in a silica-based system. Furthermore the device processing technology is well-developed and understood, and devices can be fabricated in a relatively straightforward manner [63].

5.3 Waveguide Design

5.3.1 Design requirements

Integrated optical switches and power splitters to be used within an optical fiber network must satisfy rigorous loss, polarization, power, and functional requirements. To meet each of these demands a number of choices and tradeoffs must be made, resulting in an optimum waveguide design for a given application. Each of these requirements will be discussed.

Since planar optical waveguide propagation losses typically exceed the losses in an optical fiber by a factor of 10^4 or greater, it is imperative that integrated devices are kept small. Recall from Chapter 2 that waveguides with a large core-cladding refractive index contrast support a larger number of modes for a given width. Small single mode and multimode waveguides can therefore be produced if the refractive index contrast is large. Furthermore, it will be shown that integrated optical devices typically require waveguide bends; a high index contrast allows compact, low-loss bends to be designed.

Integrated optical devices such as power splitters and switches are typically connected to optical fibers at both the input and output. The input light from a fiber typically has an unknown state of polarization, so an integrated device must operate independent of polarization to minimize polarization-dependent loss and response characteristics. Effects such as waveguide geometry, core and cladding refractive indices, and intrinsic material properties determine the degree of polarization dependence in a waveguide. These parameters must therefore be optimized in the waveguide design.

Active integrated components should be operable using low power levels, to minimize the cost of operation and the total power dissipation on a wafer, and to reduce the overall size of a packaged device. The choice of a silica-based technology rules out the possibility of active control through the electro-optic or magneto-optic effects, because the strength of these effects is very low in silica-based materials [43]. The thermo-optic effect in silica, however, is moderate, so this is the method chosen for active device operation in this work. The thermo-optic effect uses heat to change the optical properties of a waveguide, so a good waveguide design must direct heat flow and confine it to desired locations in order to maximize speed while minimizing power consumption levels.

Since the devices to be realized are based upon multimode interference effects, a waveguide design which results in good imaging behavior in MMI couplers must be used in order

to minimize excess losses and phase deviations. Recall from Chapter 2 that the resolution of self-images in the MMI coupler increases as the number of modes in the coupler increases. This indicates that a high refractive index contrast should be used to maximize the number of modes available for imaging. Furthermore, it has been shown in Chapters 3 and 4 that in order to fabricate devices with acceptable performance levels, it is essential that the theoretical relations for the MMI coupler imaging length and the phases of the images are satisfied. Since these relations have been derived based upon the approximation of sinusoidal field distributions given by (2.20), it is essential that this approximation is accurate. A large refractive index contrast is therefore required to ensure that the strongly-guiding approximation is valid in the multimode region.

The requirements described above are numerous and sometimes conflicting. For example, a metal heater is typically used for thermo-optic devices and must be located sufficiently far from the optical field to prevent interaction with the optical field which will lead to attenuation of the light by the metal. If the heater is far from the optical field, however, the waveguide heating is inefficient and the heat can potentially spread to unwanted regions. This conflict will be addressed in Section 5.3.4.

Another example is the design of polarization-independent MMI-based devices. Devices such as the GMZI rely upon self-imaging in MMI couplers and propagation in single-mode waveguide arms. If a polarization-independent single-mode waveguide is designed, then the MMI couplers will be polarization dependent since their geometry differs from that of the single-mode waveguide. Conversely, if polarization-independent MMI couplers are designed, the single-mode waveguides will exhibit polarization-dependent behavior. To resolve this issue while maintaining a simple waveguide design, a tradeoff must be made. In this work, a polarization-independent single-mode waveguide design is chosen because the polarization dependence in MMI couplers is typically low [9], and the light in the devices propagates for only a short distance in the couplers (hundreds of μm), while the propagation distance in single-mode waveguides is much longer (mm to cm). Moreover, in a waveguide with a high refractive index contrast, the light is strongly confined within the core region, allowing adjacent single-mode waveguides to be located in close proximity with negligible coupling between them. This results in MMI couplers which can be designed with closely-spaced input and output waveguides, thus minimizing the width and length of the couplers and in so doing improving their polarization behavior.

As a final example illustrating conflicting design requirements, consider all the benefits of small devices with a high refractive index contrast described above. In contrast to this, an optical fiber has a low refractive index contrast and dimensions that are typically larger than those of planar single-mode waveguides. Since low-loss fiber-to-chip coupling requires matching the size and shape of the mode in a single-mode optical fiber to the waveguide mode, large mode-mismatch losses will occur at fiber-waveguide interfaces. To overcome this problem a number of methods are available, such as increasing the waveguide mode size through adiabatic waveguide tapers in one or two dimensions [64], [65], using complex configurations of partially segmented waveguides [66], or using special fibers with reduced mode sizes [67]. In this work, the latter approach is chosen for its simplicity; however, in principle, other methods can be applied to the designs developed here.

5.3.2 Choice of material system

To achieve the confinement and waveguiding of light, regions of differing refractive indices are needed. Control over the refractive index in SiO_2 is typically achieved by doping with a small amount of phosphorus, germanium, or titanium; however, the refractive index contrast attainable by these methods is typically very low [5], [68], [69]. As a result, silica-based devices tend to have a relatively large size and exhibit weak optical confinement.

In recent years, silicon oxynitride (SiON) has received considerable interest as a promising material for use in silica-based waveguide applications (see, for example, [63], [70]–[72]). The method of plasma-enhanced chemical vapor deposition (PECVD) offers an easy way of depositing core layers of SiO_xN_y by nitrogen/silicon doping of silica at low temperatures. Furthermore, the material exhibits favourable optical properties such as tuneability of the refractive index in the range between 1.45 (SiO_2) and 2.0 (Si_3N_4), and low absorption losses for visible and near infrared wavelengths.

Unfortunately as a result of the PECVD growth process, SiO_2 and SiON layers contain hydrogen, which causes absorption of optical wavelengths around 1500nm, near the telecommunication wavelength region of interest [63], [70], [72], [73]. A common method to reduce the loss is by performing a high-temperature (>1100 °C) annealing step to drive off excess hydrogen in the layers, thereby resulting in very low propagation losses around the 1550nm wavelength region [70], [73], [74]. The drawback of this method, however, is that the layer thicknesses and refractive indices change upon annealing, and the changes themselves differ with the refractive indices used [75]. The compensation of these effects for a particular waveguide structure is largely a material processing task.

Since it is more important in this work to demonstrate the functional operation of devices than to achieve the lowest possible losses, the layers to be used in the fabrication were not annealed. This decision allows the device functionality to be emphasized. The applicability of the low losses attained by others [75], [76] to the present devices was instead demonstrated by measuring an annealed SiON layer with the same characteristics as the core layer used in the devices.

5.3.3 Waveguide structure – optical properties

In this work a novel waveguide structure has been chosen to exploit the advantages of PECVD SiO_2/ SiON technology. A cross-section of this structure is shown in Fig. 5.1.
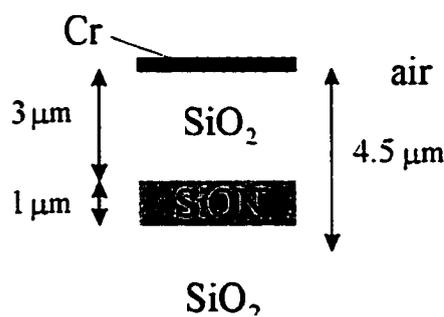
Figure 5.1 Cross-section of the rib waveguide design used in this work. The chromium layer is used for thermo-optic phase shifting in the arms of GMZI devices.

The waveguide consists of a core layer of SiON surrounded by upper and lower cladding regions of SiO_2. The structure is a strongly guiding rib waveguide, with a very high lateral index contrast due to the lateral air interfaces and a high transverse index contrast ($\Delta n = 0.101$) between the core and cladding regions. The core ($n_2 = 1.561$) and cladding ($n_1 = 1.46$) refractive indices are designed for a 1550nm wavelength and are chosen so as to maximize the index contrast while maintaining single-mode waveguide dimensions large enough to facilitate fiber coupling. The rib height is 4.5 μm, and a 6 μm SiO_2 buffer layer separates the core from the underlying silicon substrate. This buffer layer thickness results in a substrate attenuation on the order of 10^{-6} dB/cm, calculated using the method described in Section 2.2.2.

The rib is formed by a deep etch through the core and into the lower cladding region. This serves a number of useful purposes. First, the lateral air-film interfaces result in a very high refractive index contrast. In multimode waveguides, this yields a good approximation to (2.20) and therefore a better MMI coupler performance. Secondly, etching through the waveguide core reduces the sensitivity of the waveguide performance to fluctuations in the etch depth arising during fabrication. An etch depth of $0.5\mu m$ into the lower cladding has been chosen so that a complete etch through the core is assured. Finally, the deep etch results in favourable thermal properties, which will be discussed in detail in the following section.

The effective indices for the TE and TM modes of the single mode waveguide can be calculated using the EIM described in Section 2.3. However, this method requires a separation of the rib waveguide into three slab regions as shown in Fig. 2.3. Note that in Fig. 5.1 the regions adjacent to the rib do not constitute slabs, so the total effective index n_e for this structure cannot be calculated using the EIM. Since the MPA and the resulting MMI coupler design equations require a value for n_{ch}, a modified-EIM approach is utilized [77]. In this approach, the total effective index n_e for the modes of the waveguide are first calculated by an accurate method. The vector finite element mode solver FWave [78] has been chosen for this. Next, the second step of the EIM is performed using $n_{cl} = 1$ and a value of n_{ch} that results in a total n_e (for a single-mode waveguide) or L_π (for a multimode waveguide) that matches the value calculated by the mode solver. The use of the unity index is justified since the regions adjacent to the core in Fig. 5.1 are substantially air.

The TE and TM single mode field profiles and effective indices calculated using FWave are shown in Fig. 5.2, for a single-mode waveguide width of 2.7 μm.

$n_{e,TE} = 1.4877489$ $n_{e,TM} = 1.4877239$
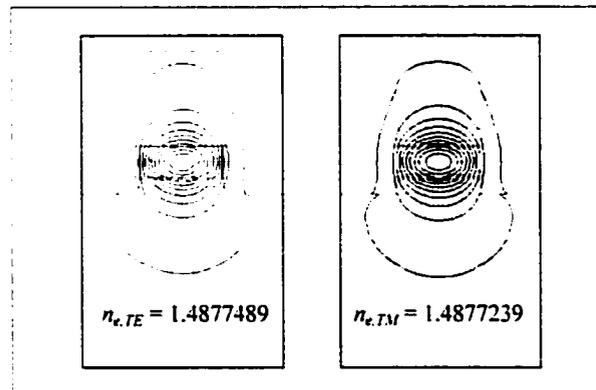
Figure 5.2 Single-mode field profiles for TE and TM modes for the waveguide structure in Fig. 5.1. The contours represent field amplitude values in steps of 10%, while the outermost contour represents a 1% field amplitude. The figure is to scale.

The waveguide width has been chosen to ensure single-mode, polarization-independent propagation in the waveguide. The polarization dependence of a waveguide is typically characterized by its birefringence, which is the difference between the TE and TM effective refractive indices, i.e.,

$$n_{e,TE-TM} = n_{e,TE} - n_{e,TM} \qquad (5.1)$$

Note that in Fig. 5.2 the effective indices for the TE and TM modes are nearly identical, resulting in a birefringence of 2.5×10^{-5}. This low value has been obtained by calculating the effective TE and TM indices for various waveguide widths using FWave, and noting that at a width of 2.7 μm the values nearly overlap. Typically, symmetric waveguide structures consisting of a square waveguide core surrounded by a uniform cladding region are used to ensure polarization-independent propagation [5], [67], [69]. While these "buried-core" waveguides achieve

polarization independence through symmetry in their geometry, the design in Fig. 5.1 uses a combination of geometry, etching, and refractive index contrast to achieve the same effect. Furthermore, as will be discussed in the following section, the present geometry has more favourable thermal properties than the buried-core structure.

The upper SiO_2 cladding layer thickness in Fig. 5.1 is 3 μm. This value has been chosen for two reasons. First, the relatively thick top cladding serves to balance the thick lower cladding, maintaining symmetry in the waveguide. The result is that the mode fields shown in Fig. 5.2 are nearly circular in the rib, which will result in lower coupling losses between the waveguide mode and the circular mode field of an optical fiber. Secondly, the thick top cladding serves to effectively isolate the waveguide mode from the chromium metal region used for active thermo-optic device control. This will be discussed further in the following section.

5.3.4 Waveguide structure – thermal properties

In the GMZI devices described in Chapters 3 and 4, the intensity in an output port is a function of the phase shifts applied to the light in the waveguide arms. In silica-based waveguides, this phase shifting is most efficiently achieved by the thermo-optic effect, wherein the phase of a mode in a waveguide arm is controlled by means of a temperature change in that arm.

SiO_2 and SiON are thermo-optic materials, which means that the refractive index is a function of the temperature of the material. The strength of this dependence is gauged by the magnitude of the thermo-optic coefficient α, which measures the change in refractive index in a material for a unit temperature change,

$$\alpha \equiv \frac{dn}{dT} \tag{5.2}$$

A change in refractive index, following the discussion in Chapter 2, modifies the effective index of the mode in the waveguide. From (2.12), it is seen that this results in a modified propagation constant $\Delta\beta$. Since the phase $\varphi = \beta z$ in (2.2), the resulting phase shift is given by

$$\Delta\varphi = L_t \cdot \Delta\beta + \beta \cdot \Delta L_t \tag{5.3}$$

where L_t is the propagation length in the region of modified temperature. The phase shift in (5.3) contains two terms: the first is a change in propagation constant due to the thermo-optic effect, and the second is a change in the physical length of the waveguide due to thermal expansion. Since the thermal expansion coefficient for silica is very low for moderate temperature increases [5], the second term in (5.3) may be ignored, while the first term may be expanded to yield

$$\Delta\varphi = \frac{2\pi\alpha L_t \Delta T}{\lambda} \tag{5.4}$$

where λ is the wavelength and ΔT is the temperature increase in the waveguide. Using the generally accepted value of $\alpha = 1 \times 10^{-5}\ °C^{-1}$ for SiO_2 and SiON [5], [71], [79], and a $\lambda = 1550$ nm in (5.4), a phase shift of π is expected for temperature increases of $\Delta T = 38.8$ °C and 25.8 °C in heater lengths of $L_t = 2$ mm and $L_t = 3$ mm, respectively.

In Fig. 5.1 a 200 nm thick chromium layer is situated on top of the rib to serve as a thermo-optic heating element in the waveguide arms of the GMZI. The ends of the heater are connected to electrodes and a voltage is supplied to induce an electrical current to flow through the chromium. The heater is thus a resistor that dissipates heat by an amount proportional to the square of the current flowing through it, or by an amount linearly proportional to the electrical power applied.

Since the effective index of the waveguide mode must be altered to induce a phase shift, the temperature in the region where the mode is guided must be increased. From Fig. 5.2, this corresponds to the waveguide core region and the immediately adjacent cladding regions. This

suggests that the most efficient heat transfer (and hence the lowest required operating power) can be achieved by locating the heater as close as possible to the waveguide mode. Metals, however, have complex refractive indices with an imaginary part contributing to the attenuation of light. The metal heater is thus an optical absorber or attenuator, and must be located a safe distance from the waveguide mode to ensure low-loss phase shifting [43].

For this reason, a relatively thick upper cladding (or buffer layer) has been used in the design shown in Fig. 5.1. A thickness of 3 μm has been determined based upon an analysis of an asymmetric three-layer slab waveguide with a fourth metal cladding layer covering the structure. The analysis is exactly the same as the four-layer slab described in Section 2.2.1, if the refractive index of the air layer is replaced with the complex refractive index n_m of a metal given by

$$n_m^2 = (n_r - Jn_{im})^2 = \varepsilon_m = \varepsilon_r + J\varepsilon_{im} \qquad (5.5)$$

where $J = \sqrt{-1}$, ε is the dielectric constant, ε_r, $\varepsilon_{im} < 0$, and the subscripts r and im denote the real and imaginary parts, respectively. A rigorous analysis of the slab structure is presented elsewhere [43]; only the results will be discussed.

The presence of the metal layer affects TM modes to a significantly greater degree than TE modes due to the continuity of tangential field components at an interface. The analysis is therefore restricted to TM modes. The result of the analysis is that the propagation loss of the fundamental TM mode is a maximum at a certain buffer layer thickness δ_p. This thickness is approximated by

$$\delta_p \cong \frac{\lambda}{2\pi\sqrt{n_2^2 - n_1^2}} \tanh^{-1}\left[\left(\frac{n_1^2}{|\varepsilon_m|}\right)\sqrt{\frac{n_2^2 - \varepsilon_r}{n_2^2 - n_1^2}}\right] \qquad (5.6)$$

where n_2 and n_1 are the core and upper cladding indices, respectively, as defined in Fig. 2.1. At thickness values below δ_p, the fundamental TM mode is transformed into a surface plasma mode.

In a single-mode waveguide, it is found that the propagation loss in the slab can be reduced to a value well below 0.1 dB/cm if an upper cladding layer thicker than $10 \times \delta_p$ is used [43]. Using the complex dielectric constant for chromium $\varepsilon_m = -41 -J$ 42, the layer indices $n_1 = 1.46$ and $n_2 = 1.561$, and $\lambda = 1550$ nm in (5.6), a value of $\delta_p = 0.206$ μm is found. The buffer layer should therefore be thicker than 2.06 μm. For safety, a value of 3 μm has therefore been chosen.

Silicon has been chosen as a substrate material primarily because high-purity, inexpensive wafers are widely available. An added benefit is that it can be considered as a perfect heat sink, since the thermal conductivity of silicon is much higher than that of SiO_2/SiON [80]. This means that any heat applied to the waveguides will be quickly drawn to the substrate, resulting in a potentially fast device response. Unfortunately, this also means that a greater applied power is needed to achieve a given temperature increase in a waveguide, resulting in high-power operation.

This tradeoff between low-power and fast operation is typical of silica-based thermo-optic devices [80], [81]. Complicated techniques such as etching grooves alongside buried-core waveguides or under-etching the silicon region below a waveguide have been used to improve the time-power product of the thermo-optic effect, with limited success [80]–[82].

The waveguide in Fig. 5.1 is designed to minimize the time-power product, which can be considered as a figure of merit for thermo-optic devices. The thickness of the lower cladding plays an important role in this respect. In general, a thick lower cladding increases the heat capacitance in the SiO_2/SiON layers resulting in low-power but slow operation, while a thin lower cladding results in fast but high-power operation. It is thus believed that the choice of an optimum lower cladding thickness in combination with the waveguide geometry presented can enable a low time-power product to be achieved.

An added benefit of the present design is that the deeply-etched rib virtually eliminates the polarization dependence of the thermo-optic effect. Though silica is an isotropic material (it possesses directionally-independent properties), anisotropic stress-optical effects occur in the material when it is non-uniformly heated [43], [82]. It has been found that etching stress-releasing grooves alongside buried-core waveguides improves their polarization behavior when heated [80]. In contrast, the structure shown in Fig. 5.1 has intrinsically low stresses since the core and cladding regions are directly exposed to air and are therefore free to expand. Furthermore, the thermal isolation of adjacent rib waveguides is high, since the thermal conductivity of air is very low [83]. This ensures that heat is directed down towards the substrate heat sink and does not move laterally toward an adjacent waveguide.

The thermal behavior of the waveguide design has been simulated using the software PDEase [84] to solve the two-dimensional heat equation in the rib waveguide using the finite element method. Source code has been written to specify the waveguide geometry, boundary conditions, and material parameters such as thermal conductivity and heat transfer coefficients for Si, SiO_2, SiON, and air. Simulations have been performed to determine the required heater power to induce a given temperature increase in the waveguide core, for various buffer layer thicknesses and waveguide separation distances.

Based on the simulations, an optimum lower cladding thickness of 6 μm has been calculated. The results of the simulation for the optimum structure are shown in Fig. 5.3. Fig. 5.3(a) shows the heat flow through the rib region and Fig. 5.3(b) shows the resulting ΔT distribution. The temperature increase in the waveguide is found to be proportional to the applied heater power. From (5.4), the phase shift is thus proportional to the applied power. A temperature increase of $\Delta T = 45$ °C in the core region is seen for an applied power of 100 mW in a 2 mm long heater, or for an applied power of 50 mW in a 3 mm long heater. Based on the calculation from (5.4), this temperature increase is sufficient to induce a π phase shift in the light travelling through the waveguide.

	ΔT
n	85.000
m	80.000
l	70.000
k	65.000
j	55.000
i	45.000
h	40.000
g	30.000
f	25.000
e	20.000
d	15.000
c	10.000
b	5.000
a	0.000 (ambient)

(a)   (b)
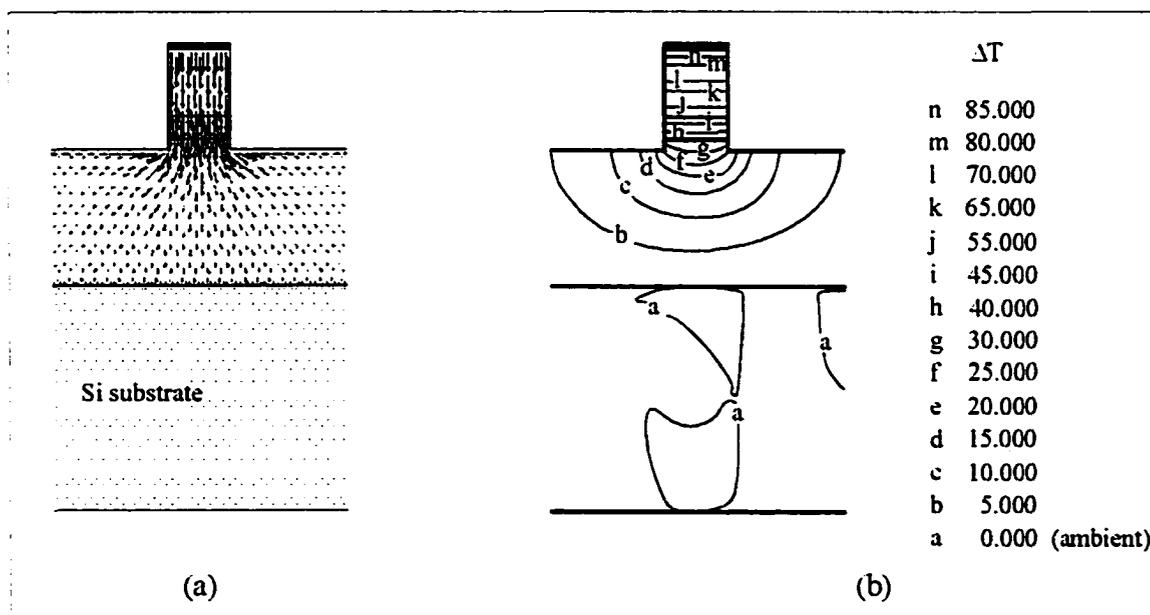
Figure 5.3 Thermal simulation results for an optimized waveguide structure with an applied heater power of 100 mW, showing (a) the heat flow, and (b) the ΔT distribution in the single-mode rib waveguide. The figures are to scale.

The heat flow diagram in Fig. 5.3(a) indicates that most of the heat is confined to the rib region and is quickly drawn down to the silicon substrate which has a high thermal conductivity and acts as a heat sink. Since adjacent ribs are separated by air, the lateral heat dissipation is small and therefore large temperature gradients can be maintained over small lateral distances.

The temperature distribution in the rib waveguide in Fig. 5.3(b) indicates that the ambient temperature is maintained at a lateral distance of < 10 μm from the rib. This absence of lateral heat conduction allows adjacent rib waveguides to be placed in close proximity, with a negligible temperature increase in an unheated waveguide. This allows for the design of compact active devices.

5.4 MMI coupler design

5.4.1 MMI coupler dimensions

The optimum imaging length of an MMI coupler must satisfy the relation given in (2.25). The coupler length is a function of the beat length of the two lowest order modes in the multimode region, which from (2.23) depends upon the propagation constants and hence the effective indices for these modes. The effective indices in turn, depend upon the waveguide geometry, and in particular the physical width W. Since the width of a multimode waveguide is larger than the width for polarization-independent propagation shown in Fig. 5.2, polarization dependence among the TE and TM modes arises in the multimode region. This results in slightly different beat lengths and thus different imaging lengths for the two polarizations.

To minimize the polarization dependence, the coupler width must be kept as small as possible. The coupler should, however, be wide enough to accommodate the N input and output single-mode (access) waveguides as shown in Fig. 2.5, with a negligible power coupling between them. The choice of a suitable coupler width is therefore critical. The following design procedure has been used to determine the optimum coupler width.

First, FWave has been used to find the effective indices of the two lowest order TE and TM modes of the MMI coupler, for various physical widths W. These are then used to find the beat length of the coupler for both polarizations, using (2.12) and (2.23). To minimize the polarization-dependent coupler imaging length requires the difference between these beat lengths, ΔL_π,

$$\Delta L_\pi = \left| L_{\pi.TE} - L_{\pi.TM} \right| \qquad (5.7)$$

to be minimized. Allowing a width to support at least N single-mode waveguides, optimum coupler widths of W = 8.0 μm and W = 16.0 μm have been found for N = 2 and N = 4, respectively. This results in minimized beat length differences of ΔL_π = 5.44 μm (N = 2) and ΔL_π = 11.04 μm (N = 4). For short coupler lengths a value of M = 1 is chosen, so from (2.25), the polarization-dependent coupler imaging length variations are $\Delta L_{M.N}$ = 8.16 μm (N = 2) and $\Delta L_{M.N}$ = 8.28 μm (N = 4). These values represent the result of the tradeoff between polarization independence in the single-mode waveguides and in the MMI couplers discussed earlier.

Next, the effective width W_e is calculated from (2.13) for both TE and TM polarizations, and an average value is found. The waveguide positions are then calculated according to Fig. 2.5, using a free parameter value of a = W_e /2N so the waveguides are spaced as far apart as possible. This results in an access waveguide spacing of 1.45 μm (N = 2) and 1.36 μm (N = 4).

Finally, as mentioned earlier, the separation between adjacent access waveguides must be large enough to ensure negligible power coupling between the waveguides. To verify this, FWave is used to find the two lowest order TE and TM effective indices of a system of two adjacent single-mode waveguides with the structure shown in Fig. 5.1, separated by a gap (in air) of 1.36 μm, as a worst case. The coupling length (or beat length), where the optical power in one waveguide is completely transferred to the other waveguide, is found for both polarizations using

the definition of beat length given in (2.23). The result is a coupling length of more than 15 mm for both TE and TM polarizations. The choice of a rib waveguide structure with a strong refractive index contrast results in relatively large coupling lengths for only a small air gap between waveguides.

It must also be noted that the gap value should be large enough to allow for accurate definition of the waveguide structure on a photomask and in the lithographic patterning of the devices. This and other device fabrication issues will be discussed in detail in the next chapter.

5.4.2 Access waveguide bends

From the above discussion of coupling between access waveguides, it may be thought that a coupling length of 15 mm is insufficient to ensure that the power remains in the access guides without coupling to an adjacent guide. This is true if the access guide separation remains at a constant value of 1.36 μm for a relatively long propagation length, however, in the present design, the guides remain parallel for only a short distance (< 100 μm) at the entrance/exit of the MMI coupler region. The coupling between the single-mode access waveguides is virtually zero if a fan-in/fan-out of the waveguides is employed, as shown in Fig. 5.4. Besides separating the access waveguides to a safe lateral distance to minimize coupling, fanning provides a suitable pitch between the waveguides at the point where optical fibers are coupled to the chip. This aids in the coupling of light to a single waveguide, and, if a proper pitch is chosen, ribbons or arrays of optical fibers can be conveniently coupled to the waveguides.
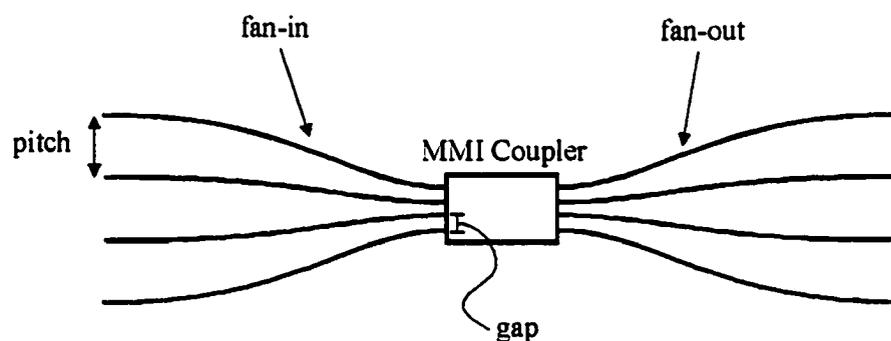
Figure 5.4 Use of fan-in and fan-out sections to increase the separation gap between access waveguides to a pitch suitable for optical fiber coupling.

The fanning must be gradual, so as to avoid excess losses incurred in the bent waveguide regions [85]. Only small bend radii can be tolerated if losses are to be kept negligible. The origin of these losses can be intuitively understood by the following argument. As light propagates around a bend, the portion of the mode towards the inside of the bend has a shorter distance to travel than the portion towards the outside of the bend. For the mode to propagate, the outer portion has to therefore "catch up" to the inner portion, which implies that the outer portion of the mode must travel at a speed exceeding the speed of light in the waveguide medium. When this occurs, part of the power carried by the mode is radiated away from the bend into the surroundings, and is thus lost.

Furthermore, as light travels through a bent single-mode waveguide, the main lobe of the mode field is shifted towards the outside of the bend [86], [87]. When a bent section is joined to a straight waveguide section a mismatch in the shape of the mode fields results, causing additional losses. These transition losses can be virtually eliminated by ensuring that there are no discontinuities in the radius of curvature along a waveguide path [88].

To satisfy this condition, the s-bend geometry originally proposed by Marcuse [89] and later used by Minford *et al.* [90] is adopted. This geometry is effective because it eliminates discontinuities in the radius of curvature and has a simple form, given by [91]

$$x(z) = \frac{W}{2} + \frac{h_b z}{l_b} - \frac{h_b}{2\pi} \sin\left(\frac{2\pi z}{l_b}\right)$$

(5.8)

where h_b is the height or maximum offset of the s-bend, l_b is the length of the s-bend, and W is the waveguide width. The geometry is shown schematically in Fig. 5.5, along with the radius of curvature for the bend, both plotted as a function of z. Note that the radius of curvature is infinite at both endpoints and in the centre of the bend, so straight waveguides can be joined to the s-bend at any of these points without transition losses.
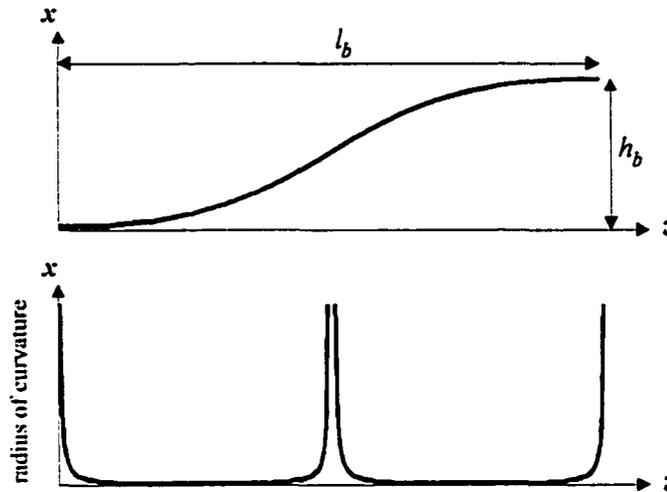
Figure 5.5 The S-bend geometry given by (5.8) is shown in the upper plot, while the lower plot shows the radius of curvature as a function of z.

The abruptness, or steepness of the bend is given by the ratio h_b / l_b. This ratio must be chosen to satisfy the required fan out distance from the coupler to the desired pitch, which is based upon multiple considerations such as device placement on the wafer, device size, and locations of cleaving points on the wafer. The ratio also determines the magnitude of the losses in the s-bend, and should be small enough to minimize these losses. From these various considerations, a ratio of ¼ has been chosen for the $N = 2$ (2 × 2) MMI coupler s-bends and the inner pair of s-bends for the 4 × 4 MMI coupler, while outer pairs of s-bends for the 4 × 4 MMI coupler have a ratio of ⅑. This results in a final pitch of 250 μm between the access waveguides, which coincides with the spacing between adjacent cores of standard single-mode optical fiber.

The losses for these s-bends using the waveguide geometry shown in Fig. 5.1 have been simulated using a scalar BPM [92]. For both bends, the excess losses were below 0.1 dB. Since this value is lower than the expected material propagation loss, the effect of the bends is assumed to be negligible. The simulated result is in harmony with the fact that in general, bend losses decrease as the optical confinement in a waveguide increases. This is another benefit of the deep etch and high refractive index contrast of the waveguide structure chosen.

5.4.3 MMI coupler simulation

Optimum MMI coupler widths and access waveguide locations have been determined in the Section 5.4.1. The corresponding MMI coupler length can be found from (2.23); however, recall that the theory upon which the relation (2.23) is based was formulated under a number of

approximations. These include the approximated effective width of the fundamental mode in (2.13), the strongly-guiding approximation yielding (2.20), and the quadratic spacing of the propagation constants in (2.22). In reality, the modes in the multimode region do not strictly follow these relations, resulting in a discrepancy between the calculated and the actual optimum coupler imaging length.

The MPA method described in Section 2.4 utilizes the mode fields determined by the modified EIM to determine the total field in the MMI coupler, and is therefore considered to be more accurate. For this reason, the performance of 2 × 2 and 4 × 4 MMI couplers for various imaging lengths have been simulated using MPA. The imbalance and excess losses are calculated by first integrating the square of the total field in the coupler given by (2.18) over the range of x-values corresponding to the output access waveguide positions determined in Section 5.4.1, and then using the relations (3.28) and (4.5). Results of the MPA simulations for the MMI couplers are given in Fig. 5.6.

(a)

(b)
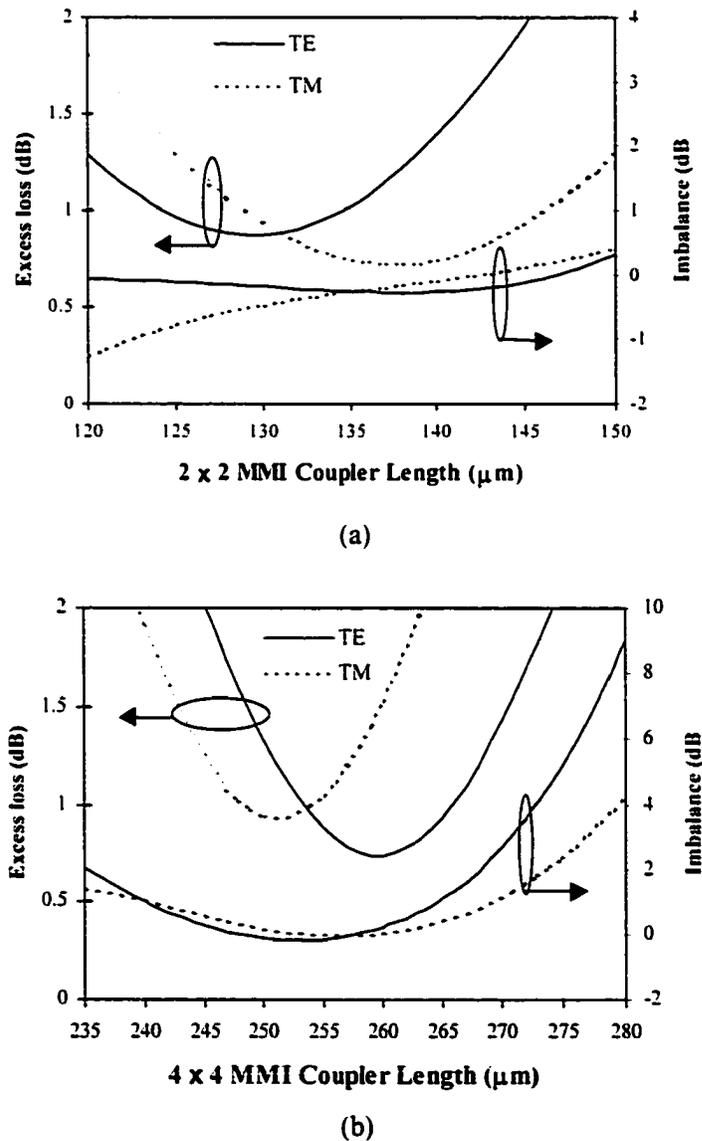
Figure 5.6 MPA simulation results for (a) 2 × 2 and (b) 4 × 4 MMI couplers, used in the determination of the optimum coupler imaging length $L_{MN\,design}$.

The polarization dependence in Fig. 5.6 manifests itself typically as a relative shift in the curves. Note that the differences in optimum imaging length for the different polarizations are very close to the values of about 8 μm predicted in Section 5.4.1.

From Fig. 5.6, it is clear that there is no single optimum coupler length where excess loss, imbalance, and polarization dependence are minimized. Tradeoffs between the parameters must be made to arrive at a suitable imaging length. The final MMI coupler design lengths chosen are $L_{M,N\,design}$ = 130 μm for the 2 × 2 MMI coupler and $L_{M,N\,design}$ = 253 μm for the 4 × 4 MMI coupler.

A summary of the design parameters for the MMI couplers, along with the excess losses and imbalances simulated using MPA, is given in Table. 5.1. The theoretical MMI coupler lengths $L_{M,N}$ calculated using (2.23) are also given for comparison. It is interesting to note that the MMI coupler dimensions in Table 5.1 are actually smaller than MMI couplers fabricated in high-contrast semiconductor waveguides [93], [94].

Parameter	2 × 2 MMI coupler	4 × 4 MMI coupler
Gap distance (μm)	1.45	1.36
Coupler width, W	8.0	16.0
Effective width, W_e	8.3	16.3
Beat Length, L_π	86.66 (TE) 92.11 (TM)	338.35 (TE) 349.40 (TM)
ΔL_π (μm)	5.44	11.04
$\Delta L_{M,N}$ (μm)	8.16	8.28
$L_{M,N\,design}$ (μm)	130.0	253.0
$L_{M,N}$ (μm)	130.0 (TE) 138.2 (TM)	253.8 (TE) 262.0 (TM)
MPA Simulation Results		
Imbalance (dB)	0.18 (TE) 0.32 (TM)	0.24 (TE) 0.40 (TM)
Excess loss (dB)	0.94 (TE) 0.90 (TM)	0.90 (TE) 0.83 (TM)

Table 5.1 Design parameters and simulation results for the 2 × 2 and 4 × 4 MMI couplers in this work.

The total light intensity pattern in the 4 × 4 MMI coupler calculated using MPA is shown in Fig. 5.7 for the TM polarization.
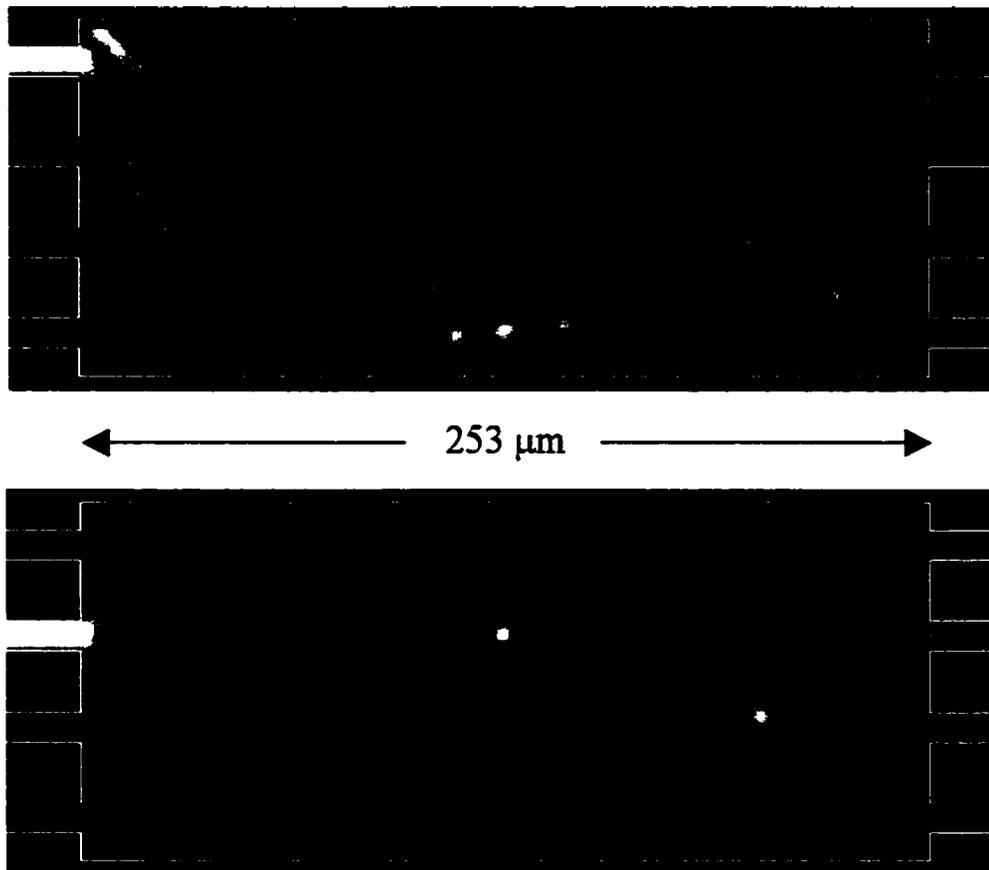
Figure 5.7 MPA simulation of the light intensity pattern in the 4 × 4 MMI coupler used in this work. The case of TM polarization for two different inputs is shown.

5.5 GMZI design

5.5.1 Phase shifting region

While the 2 × 2 and 4 × 4 GMZI designs will utilize MMI couplers and access waveguide bends with parameters identical to those described in the previous section, the design of the GMZI arms requires careful consideration. In Fig. 2.7, the arms of the GMZI are schematically pictured as straight, parallel waveguides, however, this representation is unrealistic for a number of reasons.

First, the results of the thermal analysis in Fig. 5.3(b) indicate that the temperature increase immediately adjacent to the waveguide rib is significant. If the GMZI arms are straight, parallel waveguides separated by a gap identical to the gap between the access waveguides, then significant heating will occur in immediately adjacent waveguide arms. This 'thermal crosstalk' will result in a reduced phase shifting efficiency, and more importantly, a violation of the fundamental property of independent phase shifts, upon which the operation of the GMZI is based (see Section 3.3.2).

Second, the calculation of the required heater length from (5.4) indicates that efficient phase shifting is possible for moderate temperature increases if heater lengths on the order of a few millimeters are used. If the GMZI arms are parallel for this length and are separated by only a small gap, then from the discussion in Section 5.4.1, significant power coupling among the

waveguides will occur. This violates the fundamental property of distinct self-images in the interferometer arms.

Finally, if the waveguide arms are closely spaced for a distance of a few millimeters, the probability of fabrication errors occurring in this region is high. This is due to limitations in the resolution of the photolithographic and etching processes used to form the waveguides. Examples of typical fabrication errors for closely spaced parallel waveguides include:

1) a physical connection being formed (zero gap) at certain points along the waveguide length;
2) a shallow etch depth between the waveguides due to reduced visibility of the area between the waveguides over a wide angular range; and
3) imperfect removal of chromium near the heater regions, resulting in "shorted-out" connections between the waveguides.

These errors violate both fundamental properties of the GMZI stated above.

For these reasons, the GMZI arms must be separated by a sufficient lateral distance. The most demanding criteria to satisfy is the thermal isolation between waveguide arms; however, as indicated in Fig. 5.3(b), a separation of > 10 μm effectively eliminates any thermal crosstalk. In the GMZI design to be used in this work a severely overestimated arm separation distance of $d_s = 250$ μm is chosen to ensure that the waveguides are completely optically and thermally isolated.

As in the case of the access waveguides, this separation requires a fan-out and fan-in, and can only be achieved through the use of waveguide bends. In contrast to the access waveguides, however, the requirements for the waveguide arms are very strict. The waveguide arms must remain equal in length and in loss, so the bends used must be identical. Furthermore, it is desirable to eliminate transition losses between the bent and straight waveguide sections. To satisfy these requirements, the GMZI arm geometry shown in Fig. 5.8 is used.
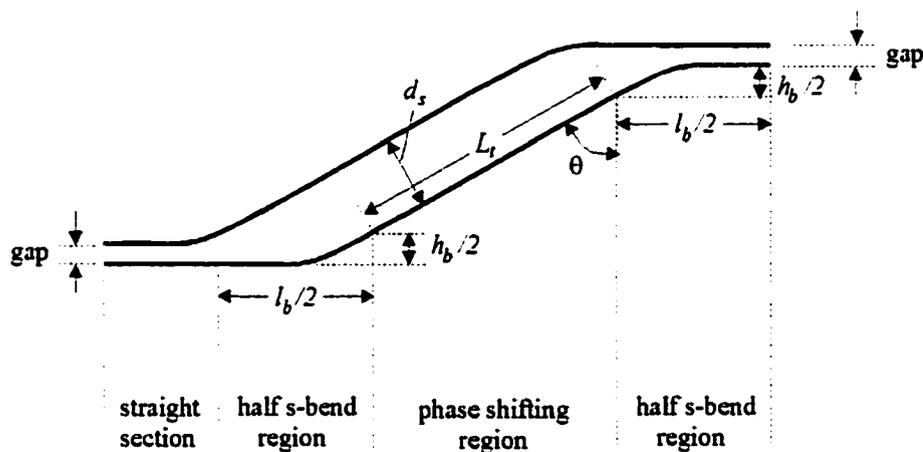
Figure 5.8 Schematic diagram of two adjacent GMZI arms. The design is used to achieve a waveguide separation of d_s in the phase-shifting region while eliminating transition losses, equalizing total losses among the arms, and maintaining equal arm lengths.

The arm design uses an s-bend described by (5.8) with a height-to-length ratio of ¼, but split in half to form two identical half s-bend sections. Since the radius of curvature of the s-bend given by (5.8) is infinite at the centre, a natural transition into and out of the straight thermo-optic heater region occurs, thereby eliminating transition losses. The design also uses buffer sections of straight waveguides 50 μm in length to separate the half s-bend from the MMI coupler and the phase shifters, and a straight waveguide section to equalize the total arm lengths.

The parameter θ determines the angle of the waveguides in the thermo-optic phase shifting region, and must be chosen such that the slope of the straight waveguide is equal to the slope of the tangent to the s-bend at its centre. This condition results in an equation which can be solved for θ:

$$\theta = \tan^{-1}\left(\frac{I_b}{2h_b}\right) \tag{5.9}$$

resulting in a value of $\theta = 63.4°$.

5.5.2 GMZI simulation – variable ratio power splitting

The switch behavior (loss and crosstalk) in the 2×2 and 4×4 GMZI with the parameters listed in Table 5.1 have been simulated using MPA, and the results have been given in Table 4.8. The variable-ratio power splitting function can also be simulated using MPA, and an example of this function is now given.

As an example of the N-way power splitting function, consider a GMZI with $N = 4$. With a single input beam into port $i = 2$ (arbitrarily chosen), we wish to achieve the following (arbitrarily chosen) output intensity distribution:

$$I_k^{out} = \begin{bmatrix} 0.27 \\ 0.01 \\ 0.56 \\ 0.16 \end{bmatrix} I_0 \tag{5.10}$$

Solving for the optimum phase shifts x_{opt} using the numerical procedure outlined in Section 3.4.2 yields, after 187 iterations:

$$x_{opt} = \begin{bmatrix} -1.37 \\ 0.01 \\ 1.37 \\ 0.13 \end{bmatrix} \text{radians} \tag{5.11}$$

In this example, the maximum phase shift magnitude needed in any single arm is 1.37 radians = 78.5°. The phase shifts given in (5.11) are applied to the arms of the GMZI, and the resulting optical field profile at the output for both TE and TM polarized light at 1550 nm is shown in Fig. 5.9. The calculated excess loss of the device is 1.73 dB for TM and 2.01 dB for TE polarization, while a maximum deviation of 1.5 % in the power splitting ratio is observed. The corresponding intensity pattern in the GMZI is shown in Fig. 5.10.
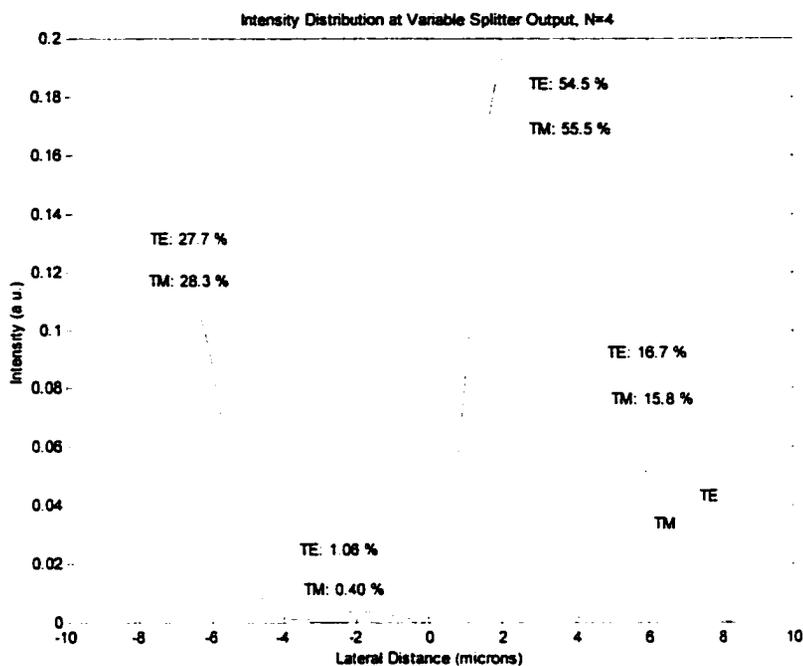
Figure 5.9 MPA simulation of the intensity profile at the output of a 4 × 4 GMZI, for the variable-ratio power splitting state in (5.10). The intensity distribution in the output waveguides for both polarizations is shown.
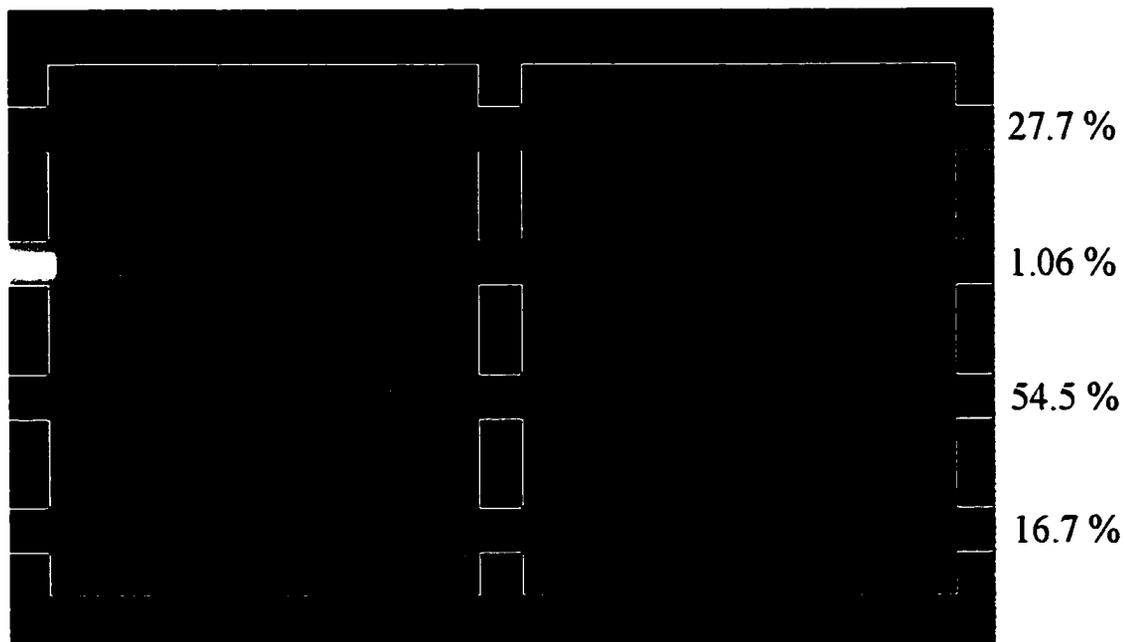
Figure 5.10 MPA simulation of the intensity pattern in the 4 × 4 GMZI with applied phase shifts given by (5.11). The result shown is for TE polarized light.

5.6 Photomask Layout

Recall from the discussion in Section 3.5.2.3 that the optimum coupler imaging length is a sensitive function of the coupler width. In the present MMI coupler designs, a ± 0.1 μm deviation in the MMI coupler width results in an imaging length deviation of $\delta L = \pm 3$ μm. To compensate for this effect, five different MMI coupler lengths are placed on the mask. These bracket and include the design length, so that width variations of ± 0.2 μm during the fabrication process can be tolerated.

The device designs have been coded into a MATLAB [95] software environment. The various design parameter values and the device positions on the mask are specified in the MATLAB source code. This code is used to generate a data file that can be read by the mask layout editor software L-Edit [96]. L-Edit enables the devices to be viewed as they will appear on the final photomask. The layout of the MMI couplers and GMZIs as they appear in the mask layout editor are shown in Fig. 5.11. Straight waveguides alongside the devices serve as reference waveguides that enable excess losses in the devices to be determined.

2 x 2 GMZIs   4 X 4 GMZIs

2 x 2, 4 x 4 MMI
couplers
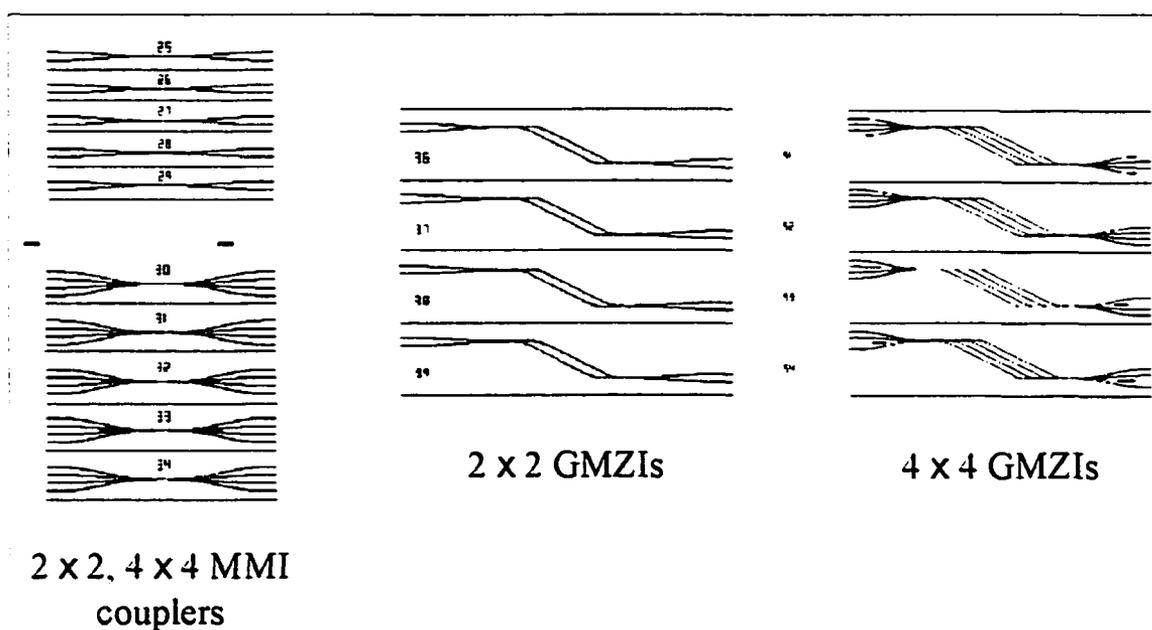
Figure 5.11 Final mask layout of the 2 × 2 and 4 × 4 MMI couplers, 2 × 2 and 4 × 4 GMZIs, and reference waveguides. The devices are fabricated with a number of different MMI coupler imaging lengths, and the numbers are used for device identification.

Excluding access waveguide fan-in and fan-out distances, the dimensions of the 2 × 2 and 4 × 4 GMZIs are approximately 3.5 × 1 mm and 6 × 2 mm, respectively.

The mask layout editor has been used to generate a file that enables a foundry to fabricate the photomask. The photomask was produced by the Photomask Department of Nortel (now Nortel Networks) in Ottawa, Ontario. The mask is a transparent glass slide with chromium metal regions defining the device geometry. The chromium has been evaporated onto the glass slide using an electron beam with a resolution of approximately 0.1 μm. The fabrication of the devices using this photomask will be discussed in detail in Chapter 6.

Chapter 6

Device Fabrication

This chapter details the process used to fabricate the integrated optical devices in this work. The various materials, equipment, and techniques employed, and results of qualitative and quantitative characterization during different phases of the fabrication process are presented. The device fabrication has been performed during a period as a visiting scientist at the MESA Research Institute, University of Twente, The Netherlands, and a portion of the characterization has been performed during a student internship at JDS Uniphase Inc., Nepean, Ontario. Portions of this chapter have been reported in a paper presented at the ICAPT '98 Conference [60].

6.1 Introduction

Fabrication of the integrated optical devices in this work can be separated into three broad stages: film processing, photolithography and etching, and post-processing. The first stage involves the deposition of high-quality thin films of optical and electrical materials on a wafer substrate. In the second stage, the device patterns are transferred onto these thin films using the photomask designed in Chapter 5. The devices are then defined by etching away unwanted film regions. Finally, in the third stage, the individual devices are separated, mounted, and (in the case of the active devices) connected to a power source.

The major steps in the fabrication process are shown schematically in Fig. 6.1 on the following page. Each of the steps will be discussed in detail in the following sections.

6.2 Film processing

6.2.1 Substrate preparation

High-purity 3-inch diameter silicon wafers with a <100> crystal orientation are chosen as the substrate material. These wafers have crystal axes oriented such that they can be easily cleaved in a direction either parallel or perpendicular to the wafer flat. Before processing, the wafers are cleaned in a standard cleaning procedure involving immersion in solutions such as fuming 100 % nitric acid (HNO_3), de-ionized water, and 70 % HNO_3 heated to 80 °C.

The first layer to be deposited on the wafer is the lower cladding; however, a tradeoff must be made for the practical realization of this layer. Recall that an SiO_2 lower cladding layer thickness of 6 μm has been calculated in Section 5.3.4. This is a relatively thick film for the PECVD process. When a film of such thickness is deposited onto the silicon substrate using PECVD, the differing crystalline structures of the materials gives rise to stresses which tend to bend the wafer. This bending can result in errors during the lithography stage, especially when fine structures are present, as is the case for typical integrated optical devices. To overcome this problem, the silicon wafer is first thermally oxidized to form a thin top layer of SiO_2 as shown schematically in Fig. 6.1(a). The thermal oxidation process has the advantage that an oxide layer forms on both sides of the silicon wafer, thereby rendering it stress-free and flat.

The oxidation is performed by placing the silicon wafers in an oven with an environment of N_2 and H_2O vapour at 1150 °C. Since thermal oxidation is a diffusion process, the thickness of the oxide layer increases with the square root of time. For example, an oxidation time of 28 hours results in a 3 μm thick thermal oxide layer. To oxidize the full 6 μm lower cladding layer, however, would require an unreasonably long time. Furthermore, since the refractive index of thermal SiO_2 is slightly lower than that of PECVD SiO_2, a thermal oxide lower cladding would
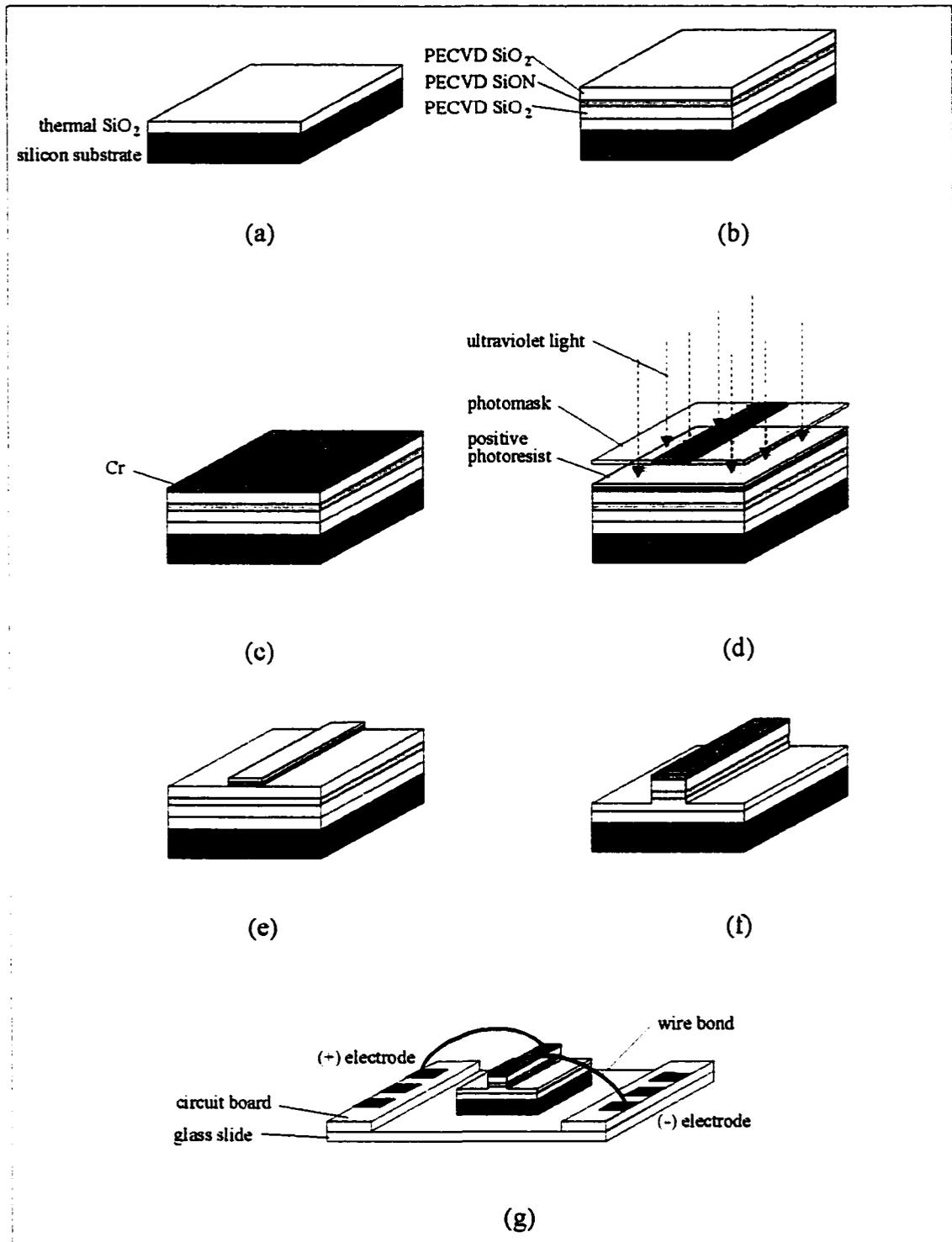
Figure 6.1 Steps in the fabrication of the integrated optical devices in this work. The silicon substrate is (a) thermally oxidized, (b) dielectric thin films are then deposited using PECVD, (c) a chromium metal layer is deposited, (d) a photoresist is spun onto the films and exposed to ultraviolet light through a photomask, (e) the photoresist is developed and excess chromium is etched away, (f) the dielectric layers are etched and the photoresist is removed, and finally (g) the devices are cleaved, mounted, and wire bonded.

result in an asymmetry between the upper and lower cladding regions, altering the modes of propagation in the devices and therefore the device designs. For these reasons, a compromise must be made, resulting in a lower cladding layer consisting of both thermal oxide and PECVD oxide.

For the lower cladding layer, a thermal oxide thickness of 3 μm covered by 3 μm of PECVD oxide has been chosen. This design sufficiently alleviates wafer bending while maintaining practical oxidation times. Furthermore, the slight asymmetry in the layer structure has been calculated to have a negligible effect upon the propagating mode, which is confined almost entirely within the top 3 μm of the lower cladding (see Fig. 5.2).

6.2.2 PECVD film growth

Plasma-enhanced chemical vapor deposition (PECVD) is a well-known and popular technique for realizing highly transparent thin-films of SiO_2 and SiO_xN_y [97]. With this method, films are deposited directly on an oxidized silicon wafer at a relatively low temperature (< 400 °C) from a plasma of N_2, SiH_4, NH_3, and N_2O. In this work a parallel-plate Electrotech Plasmafab 310 PECVD reactor has been used, and is shown schematically in Fig. 6.2.
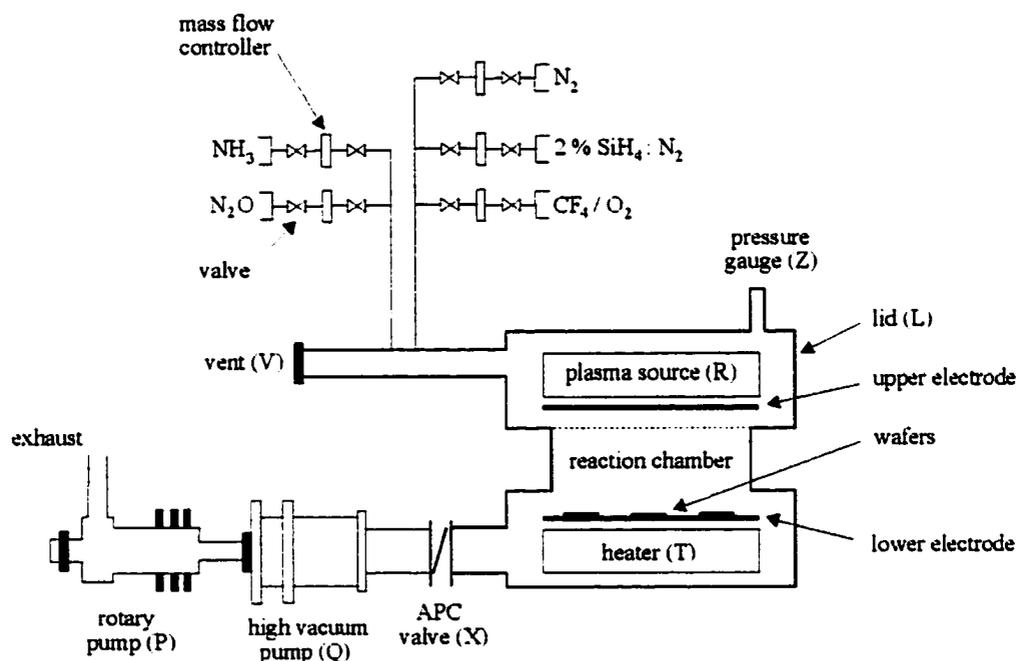
Figure 6.2 Schematic diagram of the PECVD system used to deposit the dielectric films in this work.

Up to seven 3-inch wafers are placed in the reaction chamber on an anodized circular lower electrode plate, while the lid of the reaction chamber contains the upper electrode. The electrodes are circular with a 300 mm diameter, and when the lid (L) is closed the electrode spacing is 35 mm ± 1 % over the electrode surface. The lower electrode is heated to a temperature of 300 °C by a heater (T) which is normally kept on. The two parallel-plate electrodes are required to maintain and direct the gaseous plasma.

The vent (V) is closed and the chamber is evacuated from atmospheric pressure to below 100 mTorr first using a rotary roughing pump (P), and then a high-vacuum diffusion pump (Q). The gauge (Z) monitors the pressure in the chamber. By opening the relevant valves, the process gasses (N_2O, NH_3, and 2 % SiH_4 diluted in N_2) are introduced into the reaction chamber through a

showerhead located in the chamber lid. The influx of the gasses temporarily increases the chamber pressure, but the vacuum pumps quickly lower this value.

For the films deposited in this work, a chamber pressure of 650 mTorr is required. To maintain this pressure in the chamber, the angular pressure control (APC) valve (X) is activated. The valve is a metal plate which partially blocks the channel between the pumps and the chamber, reducing the pumping efficiency and thereby increasing the pressure in the chamber. The angle of the plate determines the effective cross-sectional area exposed to the pumps and a desired chamber pressure is thus obtained by setting the valve at the appropriate angle. An automatic feedback mechanism is employed for this, in which angular adjustments are made to maintain this pressure throughout the deposition process.

Finally, the plasma source (R) is activated, and the deposition process begins. The frequency of the plasma power generator is 187.5 kHz, and a plasma power of 60 W is used. The N_2 gas is used for purging the chamber after the deposition process and the CF_4/O_2 mixture is used for periodic chemical cleaning of the reaction chamber.

6.2.3 Film optimization

6.2.3.1 Parameter optimization

The refractive indices of PECVD layers are tuned by changing the ratio of flow rates between the processing gases. Standard flow rates for SiO_2 deposition and experimental curves for SiON deposition were available for determining the required flow rates to achieve layers with desired refractive indices (see, for example, [75]). Based on this previous experimental data, rough flow rate values were determined. These flow rates were used to deposit "test layers" of SiO_2 and SiON, which consisted of a short (10-15 min.) deposition, followed by ellipsometer measurements to determine the refractive indices and layer thicknesses. The flow rates and deposition time were then modified to target the desired refractive indices and control the layer thicknesses. New layers were then deposited with the modified parameters. This iterative procedure was repeated until the desired refractive indices could be approximately obtained in a reproducible manner. The resulting optimized PECVD process parameters are given in Table 6.1.

Process	Parameter	SiO_2	SiON
PECVD	Pressure	650 mTorr	650 mTorr
	Plasma power	60 W	60 W
	Temperature	300 °C	300 °C
	Flow rates		
	NH_3	0 sccm	6 sccm
	N_2O	710 sccm	500 sccm
	2 % $SiH_4 : N_2$	200 sccm	1000 sccm
Ellipsometry	Deposition rate	21.7 nm/min	36.5 nm/min
	Refractive index	1.471	1.572

Table 6.1 Process parameters for optimized deposition of thin films of PECVD SiO_2 and SiON.
Note: 1 sccm = 1 standard cm^3/min.

Note that the refractive indices in Table 6.1 differ from the values specified in Chapter 5. It is typically difficult to deposit layers with refractive index values to a high degree of precision; however, it is the refractive index contrast Δn and not the absolute indices, that is the critical parameter in determining the device performance. So long as the desired Δn is maintained, minor deviations in the absolute film indices play an insignificant role in determining the device

performance. For example, using the waveguide design in this work (see Fig. 5.1), FWave calculations yield – for a relatively significant core/cladding absolute index shift of 0.01 while maintaining a contrast of Δn = 0.101 – a birefringence change of < 1 × 10^{-5} in the single mode waveguide and a < 2 μm change in the beat lengths of both 2 × 2 and 4 × 4 MMI couplers for both polarizations.

An ellipsometer determines the refractive index and thickness of a thin-film dielectric layer by probing the layer with laser light and analyzing the polarization properties of the reflected beam [98]. The ratio of the amplitudes of orthogonal polarization components yield a parameter Ψ while their phase difference is represented by the parameter Δ. Of the very large number of possible refractive index and film thickness combinations, a computer program is used to determine the combination that results in the measured Ψ and Δ values for the given laser wavelength and incident beam angle. In this work, a PLASMOS SD 2000 rotating analyzer ellipsometer (RAE) has been used, with a laser diode operating at λ = 1546.5 nm and an incident beam angle of 70°. The system is equipped with a personal computer for data acquisition and calculation of the film parameters. To gain confidence in the calculated index and thickness values, measurements were typically performed at a number of different wafer positions.

The accuracy of these measurements is a function of refractive index and thickness and it has been found that in some cases the results obtained may be grossly inaccurate [99]. Nevertheless, ellipsometry provides a quick and simple way to test the film parameters during the iterative film optimization process. Once the films are optimized in this manner, the exact indices and growth rates of the films are checked using the more accurate (and time-intensive) prism coupling method, which will be described in the following section.

6.2.3.2 Film characterization

The refractive index and thickness of the layers optimized using ellipsometry must be accurately determined to ensure that the index contrast does not differ significantly from the design value, and that the layer thicknesses can be accurately controlled. The prism coupling method is utilized for this purpose.

The prism coupling method [100]–[102] is based upon measurement of the effective refractive indices of the allowed modes of propagation in a slab waveguide, $n_{eff,v}$ from (2.12). Incident light is coupled into the slab by means of frustrated total internal reflection at the base of a glass prism (with index n_p) placed in contact with the film surface. A schematic diagram of the prism coupling principle is shown in Fig. 6.3.
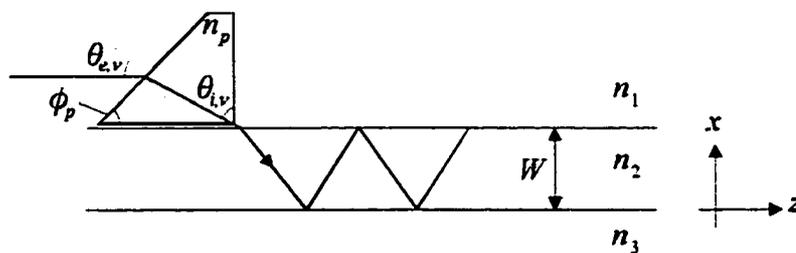
Figure 6.3 Illustration of the prism coupling principle for a slab waveguide.

Coupling into a guided slab mode depends upon a resonance condition, wherein the angle of incidence at the film surface $\theta_{i,v}$ is such that the component of phase velocity in the propagation (z) direction matches the phase velocity of a guided mode,

$$\frac{c}{n_p \sin\theta_{i,v}} = \frac{c}{n_{eff,v}} \tag{6.1}$$

where c is the speed of light in a vacuum (3×10^8 m/s). In practice it is more convenient to utilize the external prism angle $\theta_{e,v}$, so from (6.1) and Fig. 6.3,

$$n_{eff,v} = n_p \sin\left[\phi_p + \sin^{-1}\left(\frac{\sin\theta_{e,v}}{n_p}\right)\right]$$

(6.2)

The effective indices $n_{eff,v}$ are thus found by determining the resonance angles $\theta_{e,v}$. When the cladding indices and the wavelength of the light used are known, the effective index of a given mode has a well-known relation to the refractive index (n_2) and thickness (W) of the core layer, given by (2.11) and (2.12). Both parameters can be determined from the effective index values (resonance angles) of at least two different modes.

In practice, the prism parameters are supplied by the prism manufacturer (prism D9, $\phi_p = 59.59°$, $n_p = 1.8524$), the upper cladding is air ($n_1 = 1$), and the lower cladding is usually thermal oxide ($n_3 = 1.444$ at $\lambda = 1550$ nm). Both the refractive index and the thickness of a film can be obtained with good accuracy by this method, because it requires only the measurement of the incoupling (resonance) angles, which can be done conveniently and with high precision using the setup shown in Fig. 6.4.
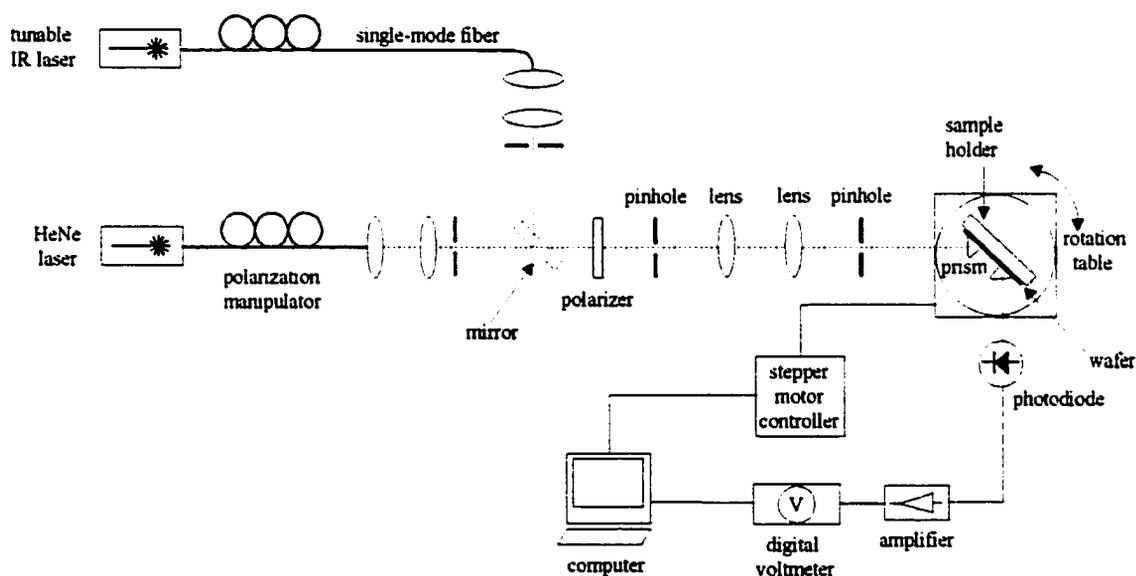
Figure 6.4 Diagram of the prism coupling setup used for the experimental determination of refractive index and thickness of PECVD films.

In this system, two lasers are used, a Helium-Neon (He-Ne) laser at a wavelength of $\lambda = 632.8$ nm for visual alignment and a tunable laser (HP 8168C) for measurements in the 1550 nm wavelength region. Using mirrors, lenses, and pinholes for adjustment, the input beam is collimated and sent to the incoupling prism. A second prism is used to couple the light out of the slab, where it is detected by a germanium photodiode placed close to the prism. The sample and prisms are fixed to a sample holder which is mounted on a rotation table that can be controlled with a high angular precision using a stepper motor. A polarization manipulator and polarizer are used together to select either the TE or TM polarization. To set a desired polarization state, the polarizer is rotated to pass the desired component and the manipulator is adjusted to maximize the intensity of the beam incident at the prism (visually, using a fluorescent card).

The computer is used to control the stepper motor and plot the signal level detected by the photodiode versus the angular displacement of the rotation table. Guided modes correspond

to peaks in the detected signal. The corresponding resonance angles are used, along with the slab and prism data in a software program [103] to calculate the refractive index and layer thickness. The accuracy of the prism coupling method has been studied [99], [104] and errors of 2–6 × 10^{-4} in refractive index, 0.1 % in layer thickness, and 2 × 10^{-4} in birefringence are found to be present.

The index and thickness of layers with the optimized PECVD parameters in Table 6.1 have been verified in the following manner. First, to satisfy the slab waveguide requirement for prism coupling, oxidized silicon wafers have been used. Secondly, the optimized layers of SiO_2 and SiON have been deposited on separate oxidized wafers, with thicknesses chosen to ensure that at least two guided modes are present for both polarizations in the 1550 nm wavelength region. An SiO_2 layer thickness of about 6 μm (4.5 hour deposition time) and an SiON layer thickness of about 3 μm (80 min. deposition time) have been targeted.

The prism coupling setup has been subsequently used to determine the refractive index and thickness values for the layers using both TE and TM polarized light at wavelengths in the 1550 nm region. These results, along with the index contrast values and their associated uncertainties, are given in Table 6.2

Parameter	Wavelength (nm)	Refractive index n_2		Layer thickness W (μm)	
		TE	TM	TE	TM
SiO_2 layer	1548	1.4587	1.4594	6.054	6.180
	1550	1.4588	1.4595	6.069	6.149
	1552	1.4587	1.4595	6.028	6.116
	1554	1.4587	1.4595	6.077	6.164
SiON layer	1548	1.5586	1.5593	3.036	3.106
	1550	1.5587	1.5593	3.031	3.106
	1552	1.5588	1.5594	3.026	3.106
	1554	1.5588	1.5593	3.017	3.105
		TE		TM	
Index contrast Δn	1548	0.100 ± 0.001		0.100 ± 0.001	
	1550	0.100 ± 0.001		0.100 ± 0.001	
	1552	0.100 ± 0.001		0.100 ± 0.001	
	1554	0.100 ± 0.001		0.100 ± 0.001	

Table 6.2 Parameters of PECVD SiO_2 and SiON layers used in the fabrication of integrated devices, measured using the prism coupling technique.

Table 6.2 indicates that the desired index contrast ($\Delta n = 0.101$) is achieved with a maximum error of 2 × 10^{-3}, the layer thicknesses can be controlled to within 3.5 %, and the parameters are insensitive to wavelength in the 1550 nm region. Furthermore, the birefringence in the index contrast value is below the uncertainty limits of the measurement. Since the results are close to the design values, the PECVD parameters in Table 6.1 are adopted for layers to be used in the fabrication of devices.

Using these parameters, the three-layer slab shown schematically in Fig. 6.1(b) is deposited on a number of oxidized silicon wafers in a single, continuous PECVD run. The slab consists of a 3 μm thick SiO_2 lower cladding, a 1 μm thick SiON core, and a 3 μm thick SiO_2 upper cladding.

6.2.3.3 Heat treatment of SiON layer

Though the deposited films are highly transparent for visible wavelengths, there is significant absorption of light in the films for wavelengths in the infrared region, as discussed in

Section 5.3.2. This is due to the incorporation of hydrogen in the layers during the deposition process [73]. The hydrogen forms Si–H and N–H bonds in the films, which have a resonant energy coinciding with the energy of photons at a wavelength of 1508 nm. Thus the light at this and nearby wavelengths is absorbed by the films. This results in significant optical propagation loss in the films, which is defined as

$$\alpha = \frac{1}{l_0} \cdot 10 \log_{10}\left(\frac{I_0}{I_{z=l_0}}\right) \qquad (6.3)$$

The loss is measured in units of dB/cm, I is the measured optical intensity along the propagation (z) direction, and l_0 is the propagation length in the waveguide.

For PECVD SiO_2 and SiON layers, typical propagation losses at $\lambda = 1550$ nm range from 3 – 10 dB/cm [105]. In Section 5.3.2, however, it was noted that a heat treatment process, or annealing, can significantly reduce these losses. In this process, the wafers are placed in an oven at a temperature typically > 1100 °C to drive off excess hydrogen from the deposited layers. It has been confirmed that annealing reduces the number of Si–H and N–H bonds in the films [75], resulting in a corresponding decrease in the optical losses. Optical loss values of annealed SiO_2 and SiON films are typically below 0.2 dB/cm [70], [75] and losses below 0.1 dB/cm have been reported [76].

As mentioned in Section 5.3.2, the films to be used for devices in this work have not been annealed. To qualitatively determine the potential improvement of the losses by annealing, the SiON film used in the prism coupler measurements above has been annealed. First, the wafer was cleaved into three sections. One section was not annealed, one section was annealed at 1150 °C for one hour, and the final section was annealed at 1150 °C for two hours. The sections to be annealed were placed in an oven with an N_2 flow rate of 1200 sccm at a temperature of 400 °C. The temperature was then gradually increased to 1150 °C over a 30 min. period, where it was maintained for the duration of the anneal, followed by a similar 30 min. cool-down period.

The spectral transmission in the films has been measured in a prism coupling setup using a broadband light source and a spectrum analyzer. The coupling angle is fixed at a given mode, and the outcoupled light intensity is measured as a function of wavelength. Next, the outcoupling prism is moved further from the incoupling prism, and the measurement is repeated. In this way, the spectral response of the film can be isolated. The results for the SiON films are given in Fig. 6.5.

The sample which has not been annealed shows a large absorption for wavelengths near 1500 nm. For these wavelengths, the transmitted intensity is near the noise limit of the spectrum analyzer, resulting in the spurious fluctuations seen in Fig. 6.5. By comparison, the sample annealed at 1150 °C for one hour shows a significant reduction in the losses near 1500 nm, indicating that hydrogen has been effectively removed from the layers. Annealing at 1150 °C for two hours induced significant bending in the wafer, so that prism coupling measurements for this sample could not be performed reliably. The result shown in Fig. 6.5 demonstrates that a simple heat treatment can dramatically reduce the optical losses in the PECVD films used in this work.
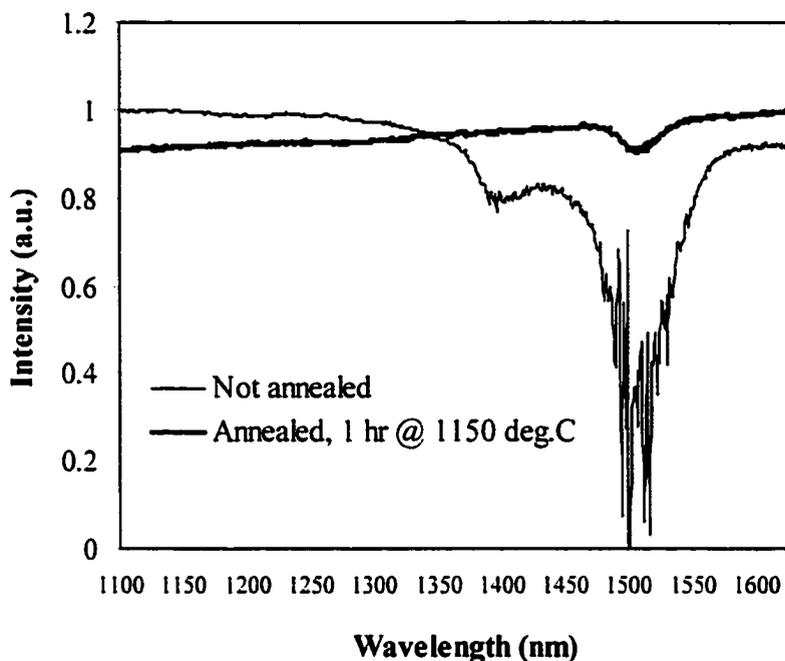
Figure 6.5 Spectral transmission characteristic of un-annealed and annealed SiON films.

6.2.4 Metal sputtering

The final step in the film processing stage is the deposition of a layer of chromium metal on the dielectric films as shown schematically in Fig. 6.1(c). This layer serves as the thermo-optic heater material and is also used as a masking layer during the etch process in Fig. 6.1(f). Thin films of chromium are simple to deposit, exhibit good adhesion to silica, and are relatively hard, making this a good choice for thermo-optic devices. The major disadvantage of chromium is the formation of oxides when the metal is exposed to air, resulting in a relatively high electrical resistance.

For a given applied electrical power, high resistances imply the use of large voltages, which can be impractical for an integrated device. The resistance of the thermo-optic heater, however, is inversely proportional to the cross-sectional area, so the layer thickness can be increased in order to decrease the resistance. For the chosen deposition method, layers can be safely deposited to an upper thickness limit of about 200 nm, so this is the layer thickness used.

The chromium is deposited by the well-known sputtering method [106], using a sputter-gun system called 'Sputterke' made at the University of Twente. A schematic of the system is shown in Fig. 6.6.
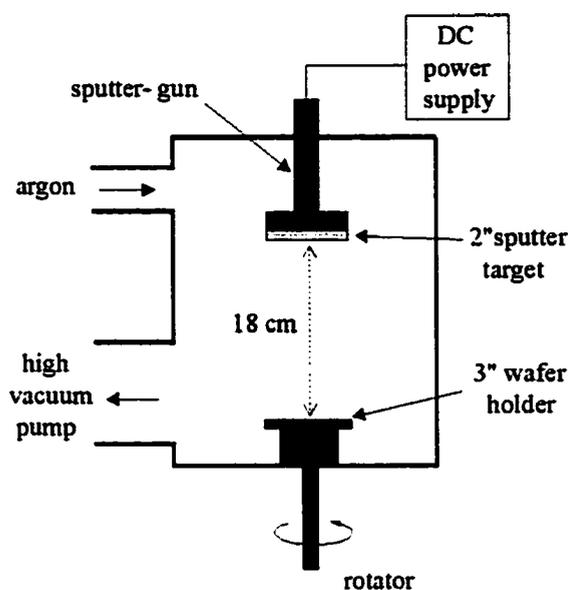
Figure 6.6 Schematic diagram of the 'Sputterke' system used for the deposition of chromium layers.

The wafer is placed on the holder, a 2-inch chromium target is fixed to the sputter-gun, and the chamber is evacuated to a background pressure of about 2×10^{-7} mTorr, using a high-vacuum turbomolecular pump. Argon gas is introduced into the chamber and a plasma is initiated using a DC power supply. The argon ions bombard the chromium sputter target, releasing chromium which coats the wafer surface. The wafer is located on a rotating holder situated 18 cm below the target, and is water-cooled during the sputter process. The pressure in the chamber during sputtering is 5 mTorr, a 200 W DC power maintains the plasma, and the deposition rate of chromium is 10 nm/min.

Sputtering of the chromium layer completes the first stage of film processing. Fig. 6.7 shows a cross-sectional view of a cleaved wafer at this stage, taken using an optical microscope.
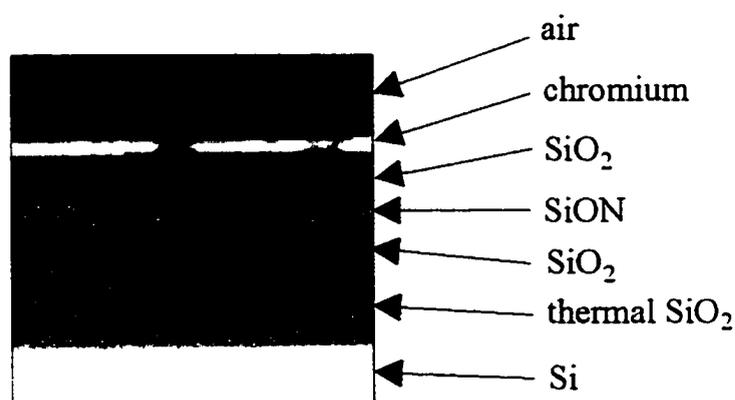
Figure 6.7 Optical microscope photograph of the wafer cross-section after the film processing stage. A smooth transition between the thermal and PECVD oxide layers is evident. The chromium and SiON layers appear thicker than they actually are, due to the finite microscope resolution and the large amount of light reflected by these regions.

6.3 Photolithography and Etching

The second stage in the fabrication process involves transferring a two-dimensional pattern of integrated devices from a photomask to the wafer (photolithography), and the subsequent structural definition in the vertical dimension by removal of unwanted material (etching). These processes are now described.

6.3.1 Photolithography

The photolithography process has been performed in the MESA clean room facility at the University of Twente. A laminar flow bench is used, illuminated by yellow light to prevent unwanted exposure of photosensitive materials to the high-energy (short-wavelength) components of white light. At MESA a standard lithography process is available; however, in practice a procedure is often specific to the devices and films used in a particular design. In addition, the condition of the chemicals used can vary on a daily basis, contamination of chemicals by other users is possible, and chemical suppliers are periodically changed resulting in slightly differing product characteristics. The process to be described has therefore been developed by a certain amount of trial and error, and is specific to the devices in this work and the time period in which they have been fabricated. The relevant process steps are now described.

First, to improve adhesion of photoresist to the wafer, a layer of hexamethyldisilizane (HMDS) is spun on the wafer at 4000 rpm for 30 sec. using a Headway resist spinner. Shipley S1813 positive photoresist (resist) is then spun onto the wafer at 4000 rpm for 30 sec. The result is a resist layer with a 1.3 μm thickness, which has been determined to be optimum for the present devices. If a resist layer is too thin it may be easily etched away during the etch processes, however if the layer is too thick, the fine structures in the photomask become ill-defined when transferred to the resist layer.

At this point, the wafer is visually inspected for inhomogeneities in the resist layer. Minute particles which may be present in the photoresist cause dots to appear in the resist layer which could cause errors in the device definition. If at any point in the photolithography process the wafer is deemed to be unacceptable, the photoresist can be removed by rinsing it with acetone, isopropanol, and de-ionized water, spinning it dry, and repeating the previous photolithography steps.

Next, the wafer is pre-baked on a 90 °C hot plate for 5 min to harden the resist layer. This is followed by alignment of the photomask over the wafer using a Karl Süss MJB3 mask aligner. This equipment moves the photomask pattern over the wafer, allowing for coarse or fine adjustments to be made by the user. Once the mask is aligned, hard contact of the mask with the wafer is made (the pre-bake was necessary so the resist layer could withstand this physical contact). The wafer is then exposed to an ultraviolet light source through the photomask, as pictured schematically in Fig. 6.1(d). The source has a wavelength of 350 nm and an intensity of 15 mW/cm^2, and an exposure time of 6.0 sec. has been used. After exposure, the photomask is cleaned using acetone, isopropanol, and de-ionized water, to remove any remnants of photoresist.

The wafer is subsequently developed in metal-free developer MF351 for positive resist, diluted to a developer:de-ionized water ratio of 1:5. A development time of 30 sec. has produced the best results. Immediately following development, the wafer is rinsed thoroughly in de-ionized water (a quick-dump rinser is used for this purpose) to halt the development process. The wafer is then spun dry. The developer should remove all resist corresponding to the glass areas on the photomask, leaving a photoresist replica of the dark pattern from the photomask. The structures in this resist replica are viewed with an optical microscope to ensure that fine features such as the gap between closely spaced access waveguides and the definition of the MMI coupler input and output regions are satisfactory. The wafer is also inspected for any significant errors, dirt particles, or other undesirable features in the resist.

The wafer is then post-baked for 1.5 hours on a 120 °C hot plate, to further harden the resist layer prior to etching.

6.3.2 Wet chemical etching

Before the structures on the wafer can be defined in the vertical direction by reactive ion etching, excess chromium must first be removed from wafer regions not covered by the resist. Chromium is then present only under the resist, thereby serving as a secondary mask for the reactive ion etch while at the same time ensuring that the resulting waveguides are completely covered by the required thermo-optic layer. The secondary chromium mask has been shown [107] to improve the smoothness and reproducibility of the vertical walls (sidewalls) created by the reactive ion etch.

The excess chromium is removed by a wet chemical etching procedure. The wafer is repeatedly immersed in the chromium etch solution $Ce(NH_4)_2(NO_3)_6$ / $HClO_4$ / H_2O (mixed in the ratio 165g / 42 mL / 1 litre) for a duration of 2 min. The etch rate of this solution is 10 nm/min. The end of the etch process is readily visible since the underlying SiO_2 layer (which has a much lower reflectivity) is exposed. The wafer is then rinsed in de-ionized water and baked for 5 min. on a 120 °C hot plate. The wafer at this stage is shown schematically in Fig. 6.1 (e).

6.3.3 Reactive ion etching

Reactive ion etching (RIE) is a dry etch procedure applied to define dielectric films in the vertical direction in order to realize channel waveguide structures. To perform the RIE process, a PlasmaFab Electrotech PF340 machine has been used. The etch process is computer-controlled, and the general configuration of the equipment is similar to the PECVD system. Although up to three wafers can be placed on the lower electrode, in this work wafers have been placed centrally on the electrode and etched one at a time to improve the etch uniformity. The substrate is cooled, the chamber is evacuated, and the etching gas is introduced. The background pressure is set to a desired level by an APC valve, and a plasma is initiated. The etching gas dissociates into radicals that then perform the etching.

The selectivity of the process, defined as the ratio of the etch rates of the resist to the dielectric films, is 1:5 to 1:10 for silica-based films [107]. This indicates that in the worst case, an etch depth of 6.5 μm is reached before the 1.3 μm resist layer is completely etched. From Fig. 5.1, however, an etch depth of only 4.5 μm is required. The chromium layer is used as a mask material under the resist for the purpose of increasing the channel definition and creating smooth sidewalls. However, it does not serve as an effective etch mask once the resist layer has been completely etched. This has been confirmed by using a set of process parameters with a faster resist etch rate. Once the resist is completely etched, the chromium layer is attacked by the plasma and re-deposited around the wafer in a sputter-like process. The re-deposited chromium islands serve as miniature masks, resulting in a severely scarred and pitted wafer surface. An adequate resist thickness is therefore crucial in the etch process.

The waveguide structures described in earlier chapters consist of rectangular channels with vertical sidewalls. However, the RIE technology is capable of both isotropic (directionally-independent) and anisotropic (directionally-dependent) etching. Since the etch procedure must be highly directional in order to realize perfect verticality, anisotropic etching is required. A detailed investigation of the variation of RIE parameters in the etching of PECVD SiO_2 and SiO_xN_y has been performed [108], and a set of process parameters has been developed to achieve an anisotropic etch process. These parameters are described below.

The substrate is cooled to a temperature of 10 °C, 10 sccm of CHF_3 gas is introduced into the chamber, a 10 mTorr background pressure is reached, and the plasma source is activated with a power of 60W. The etch rate using these parameters is about 2.5 μm/hour; however, this rate is only approximate and can vary significantly depending upon the wafer used, the previous

processes performed in the chamber, and the general cleanliness of the chamber. An accurate etch rate, however, is needed to achieve the desired 4.5 μm etch. Etch depth values which differ from this desired value alter the mode shape and introduce birefringence in the single mode waveguides, and modify the beat lengths in the MMI couplers. Control of the etch depth during the process is therefore essential.

Accurate etch depth control has been achieved by monitoring the progress of the RIE in real-time using the interference technique pictured in Fig. 6.8. When the etch process begins, a diode laser mounted with a beam splitter is placed on top of the RIE chamber. The beam passes through the beam splitter and through a glass window on top of the chamber (1), is reflected from the wafer being etched and passes through the window again (2), and is deflected at the beam splitter (3). The signal (3) is detected by the photodiode and displayed with a multimeter and simultaneously plotted versus time using a personal computer.
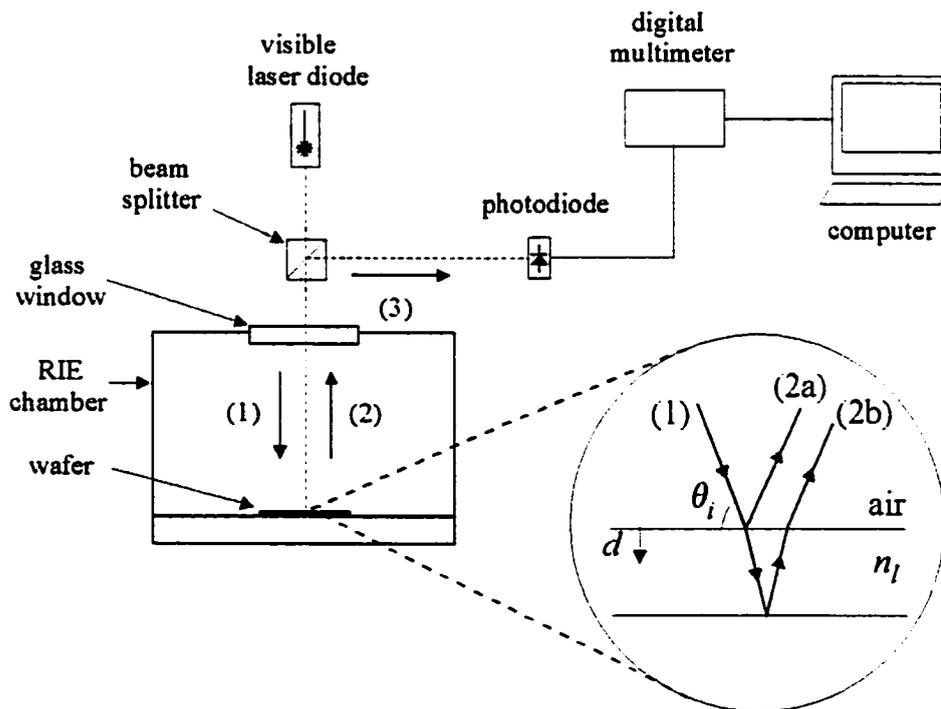
Figure 6.8 Schematic of the interference technique used for monitoring and controlling the etch rate of a wafer during the RIE process.

The orientation of the beam splitter is adjusted to make the paths (1) and (2) coincide, corresponding to normal incidence at the wafer surface (i.e., θ_i = 90° in the magnified inset of Fig. 6.8). The two components of the reflected beam (2), labeled (2a) and (2b) in Fig. 6.8, interfere, forming a characteristic fringe pattern. When the etch depth d is equal to a half-wavelength of the laser light in the film (i.e., $\lambda_L/2n_l$), the fringe pattern reproduces itself. This is summarized in the following relation [109],

$$m\lambda_L = 2n_l d \sin\theta_i \qquad (6.4)$$

where m is the fringe number, λ_L = 674.53 nm is the laser wavelength, n_l is the refractive index of the l-th layer at the laser wavelength, and θ_i is the angle of incidence at the wafer surface. Since the angle of incidence is 90°, the sine factor in (6.4) is unity. Using refractive index data obtained from prism coupling measurements at the HeNe wavelength, (6.4) has been used to calculate the

expected etch rates d/m. The results are 229 and 215 nm/fringe for the cladding and core layers, respectively. A typical interference plot obtained using the setup in Fig. 6.8 is shown in Fig. 6.9.
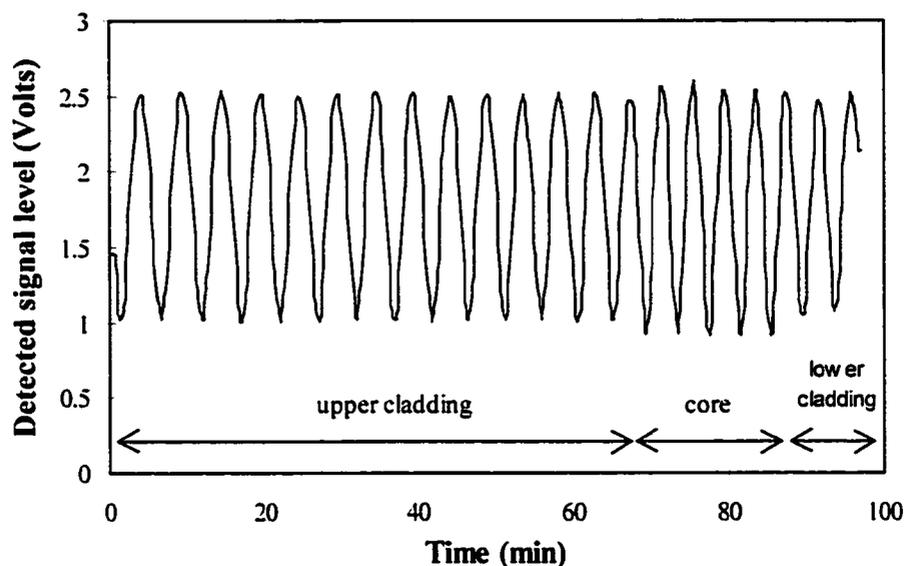
Figure 6.9 Interference plot obtained from the etch rate control setup for the RIE process, showing the approximate fringe locations of the SiO_2 cladding and SiON core layers.

Since the core (higher index) layer has a greater reflectance than the cladding layers, the interference between beams (2a) and (2b) in Fig. 6.8 is more complete, so the amplitude of the fringes are greater in this region, as seen in Fig. 6.9. The visible core-cladding transition allows the determination of both the core layer thickness and the point at which the etch process should be terminated. The core layer thickness is calculated by counting the number of high-amplitude fringes and multiplying this value by the etch rate determined from (6.4). From Fig. 6.9, about 5 SiON fringes are present, resulting in a core layer thickness of approximately 5 fringes $\times$ 215 nm/fringe = 1.075 μm, which is close to the desired value of 1 μm. To determine the point at which the RIE process must be terminated, an upper cladding thickness of 3 μm (assumed) is first divided by the number of fringes for this layer, resulting in a better approximation of the actual etch rate. Using Fig. 6.9, an etch rate of 218 nm/fringe is determined. The desired lower cladding etch depth of 0.5 μm therefore corresponds to 2.3 fringes, so the RIE process is terminated when 2.3 fringes are counted in the lower cladding layer.

After the RIE process is completed, the remnants of the resist layer are removed using acetone, to arrive at the final step in the second fabrication stage, pictured in Fig. 6.1(f). To determine the accuracy of the RIE process, a measure of the total etch depth is needed. This has been done by performing a surface profile scan of test structures on the wafer, using a Sloan DEKTAK 3030 surface profiler. The profiler consists of a sensitive needle that makes contact with the etched lower cladding layer. The needle is then drawn horizontally across the surface. When the needle encounters a rib waveguide channel, it deflects vertically as well, until it reaches the top of the channel. The total vertical deflection is measured with a resolution of 0.1 nm, providing an accurate indication of the height of the rib, which is the total etch depth. A typical profile scan of a rib structure on the wafer is shown in Fig. 6.10.
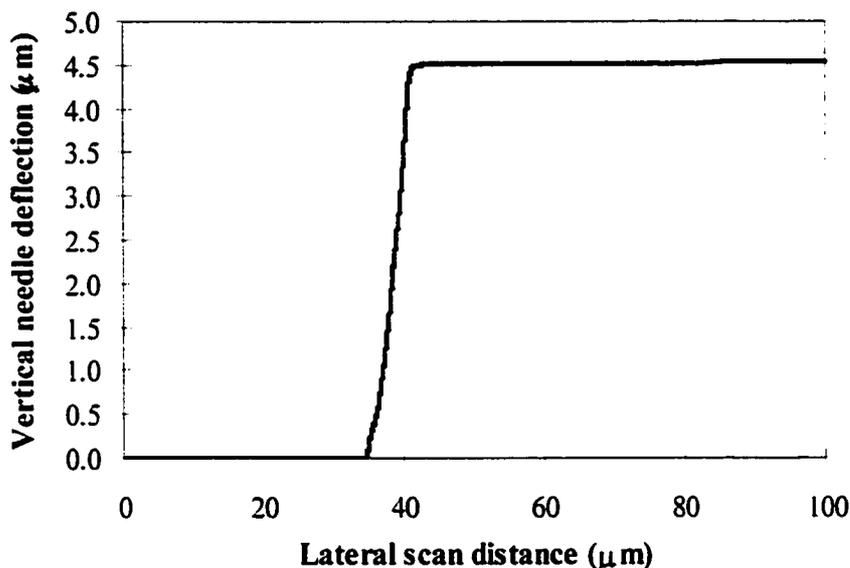
Figure 6.10 Surface profile scan of a rib waveguide channel after the RIE process, indicating that the desired etch depth of 4.5 µm has been achieved.

Note that the vertical deflection is not perfectly 90°. This is an effect arising from the triangular point of the needle, which prevents it from accessing the corner region of the rib. The height of the rib measured by this method is 4.52 ± 0.02 µm, which is very close to the desired value of 4.5 µm, indicating that the RIE has been performed successfully. The height of the chromium layer has not been considered because the needle pushes the relatively softer chromium metal aside as it is scanned horizontally, producing a scratch in the chromium layer. Using the surface profiling method, an etch depth variation of less than 50 nm has been measured for channel structures in a 3-inch wafer area.

6.3.4 Visual characterization

Upon completion of the second stage of fabrication, the resulting devices have been characterized using optical microscopes and a scanning electron microscope (SEM). This section presents some of the images of the device geometry obtained by these methods.

A cross-sectional view of a single-mode waveguide as seen under an optical microscope (1000× magnification) is shown in Fig. 6.11. The image has been taken using a CCD camera, and the scale shown is approximate. The result can be qualitatively compared to the design shown in Fig. 5.1. The etch is nearly vertical, indicating that an anisotropic RIE process has been achieved. It must be noted, however, that a limitation of the optical microscope is that edges and transitions between regions are ill-defined. Because of this, the waveguide appears wider than the designed value of 2.7 µm and as in Fig. 6.7, the SiON and chromium layers appear thicker than they actually are due to the high reflectivity of these layers. Still, much qualitative information is provided by the images, such as the shape of the structures and relative core/cladding positions.
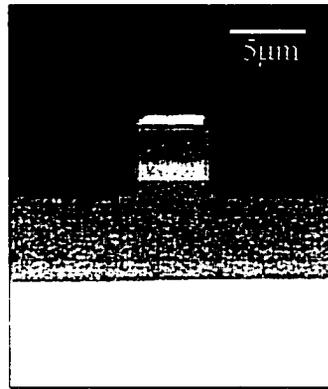
Figure 6.11 Cross-sectional view of a fabricated single-mode waveguide.

In Fig. 6.12, two views of a fabricated 2 × 2 MMI coupler are shown. Fig. 6.12(a) shows a top view of the region where the access waveguides enter the MMI coupler, and indicates that the desired sharp edges in this region are slightly rounded by the photolithography/RIE processes. In Fig. 6.12(b) a cross-section of the single-mode access waveguides is shown, indicating that the gap (1.40 μm design value) has been achieved with a good rectangular etch profile down to the bottom of the etch.

(a)   (b)
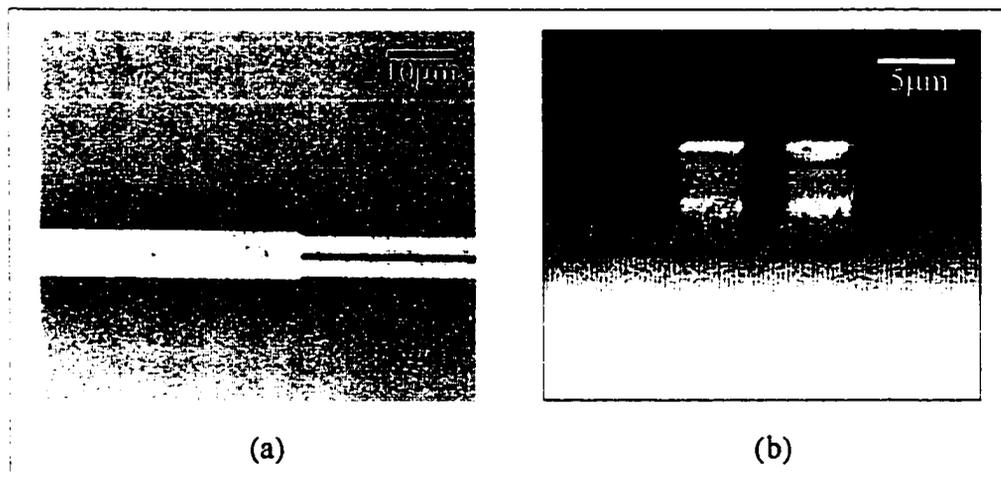
Figure 6.12 Two views of a fabricated 2 × 2 MMI coupler. A top view (a) shows the access waveguides entering the coupler region, and a cross-section (b) of the access waveguides indicates the rectangular nature of the etch in this region.

Fig. 6.13 shows two views of a fabricated 4 × 4 MMI coupler that has been cleaved. The top view in Fig. 6.13(a) shows detail of the region where the access waveguides enter the MMI coupler, and again a slight rounding effect is seen. Fig. 6.13(b) is an image of the MMI coupler from Fig. 6.13(a) viewed in cross-section. The layers are clearly distinguishable; however, the image sharpness is reduced due to an imperfect cleave.
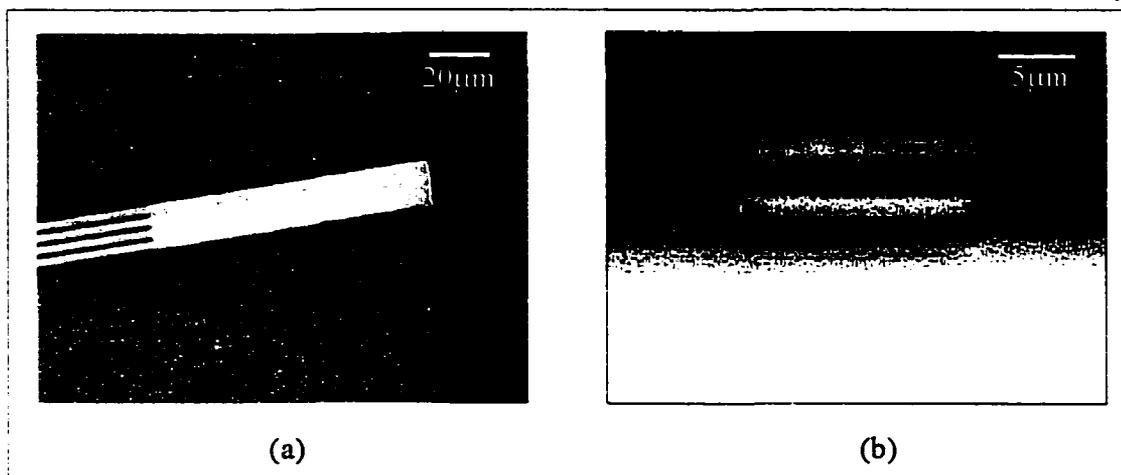
Figure 6.13 Two views of a fabricated 4 × 4 MMI coupler. The MMI coupler in (a) is cleaved near the region where the access waveguides enter the coupler, and in (b) the MMI coupler is viewed in cross-section.

Although the above images provide a significant amount of qualitative device information, the dimensions of structures cannot be determined reliably. Furthermore, the magnification in an optical microscope cannot reveal fine structural details, and only top and cross-sectional views are practical. To overcome these limitations and gain quantitative insights into the fabricated devices, a scanning electron microscope (SEM) has been utilized.

In this technique, the sample is first coated with a thin layer (~10 nm) of gold. The sample is then mounted on a holder and placed in a chamber, which is pumped down to a vacuum. The SEM operates by bombarding the sample with a scanned beam of electrons accelerated by a potential of up to 40 kV, and detecting either the backscattered electrons or the secondary electrons emitted from the gold layer. The relative intensity and locations of these electrons are used to deduce the image characteristics. Magnification of a few times to greater than 10^5× is possible using a SEM.

Fig. 6.14 shows SEM images of two different single-mode waveguides. Using the scale provided, a rib width of 2.6 – 2.9 μm and a rib height of 4.4 μm are found. Note the verticality of the etch produced by the RIE process.
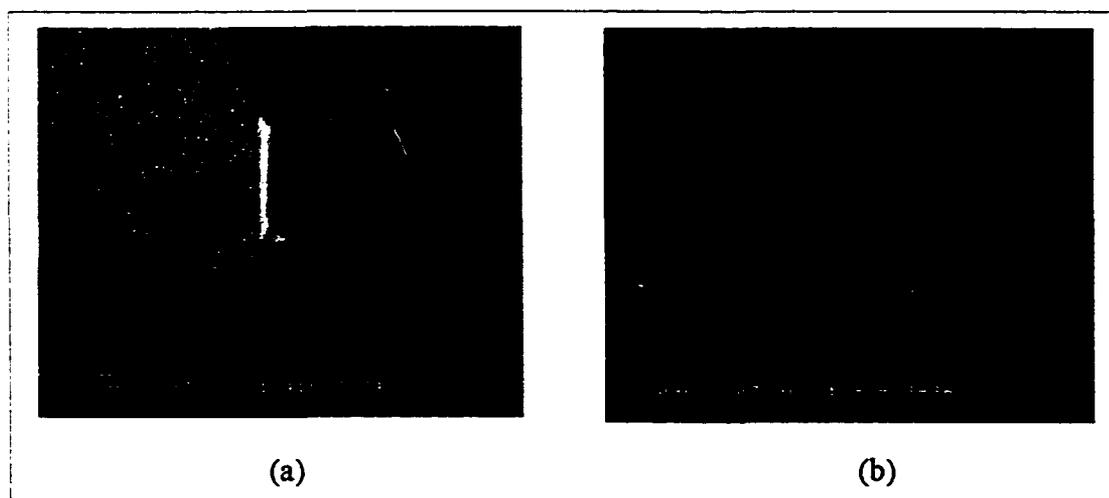
Figure 6.14 SEM images of two different single-mode waveguides.

Fig. 6.15 shows SEM images of two different 4 × 4 MMI couplers. The width of the coupler in Fig. 6.15(a) is 15.8 μm. A magnified view of the access waveguides entering the coupler region is given in Fig. 6.15(b). The details of the etch can also be seen. The etch process has resulted in some roughness in the waveguide sidewalls near the bottom of the rib. The magnitude of this roughness is estimated to be on a scale of about 10–100 nm. Note that the gap between the waveguides in Fig. 6.15 is larger than the desired value of 1.36 μm. This is the result of poor lithography for the particular wafers shown. This problem was corrected in subsequent wafers, as can be seen in Figs. 6.12(a) and 6.13(a).

(a)   (b)
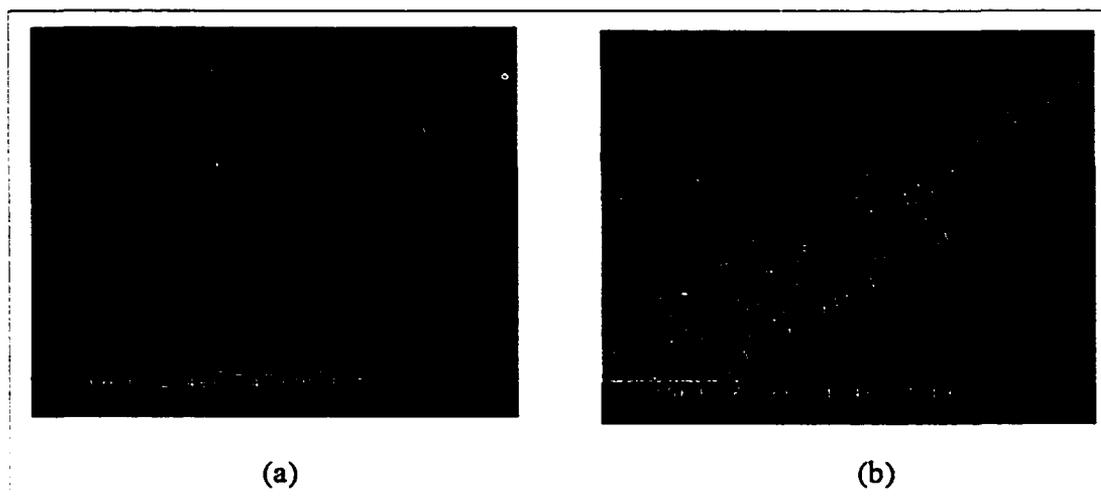
Figure 6.15 SEM images of two different 4 × 4 MMI couplers, showing (a) a top view of the MMI coupler and access waveguides, and (b) detail of the sidewalls resulting from the etch process.

6.4 Post-processing

The final stage in the fabrication process involves the preparation of devices for the experimental measurements to be described in the following chapter. The first part of this process involves dividing the wafer into small pieces that contain devices of a similar nature. The division points should cross the single-mode access waveguides and the resulting cross-section of the waveguides should be smooth and flat as shown in Fig. 6.14, to facilitate the coupling of light to and from the waveguide. Two methods of dividing the wafer are typically employed in integrated optics. Both methods have been studied to determine their suitability for the devices in the present work.

The first method involves sawing the wafer at desired locations, followed by a mechanical polish of the sawed facets. This method offers the advantage of precision, since a special wafer saw can make cuts to within 50 μm of the desired location. A MicroAutomation 1006 wafer saw has been used to divide a wafer. The wafer is first coated with a protective resist layer and pre-baked for a few minutes. An adhesive foil is then placed over the wafer to keep the devices in place during the sawing process. The saw is guided with the aid of a microscope, and de-ionized water is continuously flushed over the wafer to keep the devices clean. A SEM image of a single-mode waveguide after a saw cut is shown in Fig. 6.16(a). The waveguide is completely destroyed, clearly indicating the harshness of the sawing process.

The second method of dividing the wafer is by simply cleaving it along the crystal planes. This method has been used to produce the images shown in the previous section. In principle, this method can produce straight edges and very smooth facets, however, the 4.5 μm rib height makes the single-mode waveguides very fragile. Considerable care is therefore necessary to

avoid chipping of the waveguides during the cleaving process. A SEM image of a cleaved waveguide that has been chipped at the end is shown in Fig. 6.16(b).

To alleviate this problem, a number of different cleaving methods have been investigated. These included using a cleaving block, coating the waveguides with waxes or epoxies and cleaving it from underneath, and various manual cleaving techniques using different types of diamond-tipped cleaving pens. The technique which produced the best and most reproducible results for the devices in this work was to use a sharp, diamond-tipped pen held perpendicular to the wafer surface, and make repeated short scratches at the end of a long wafer section until the wafer cleaved. Generally, it was found that as the piece to be cleaved became shorter, a good cleave was more difficult to obtain. An optical microscope photograph of waveguides cleaved using this method is given in Fig. 6.16(c).

(a) (b) (c)
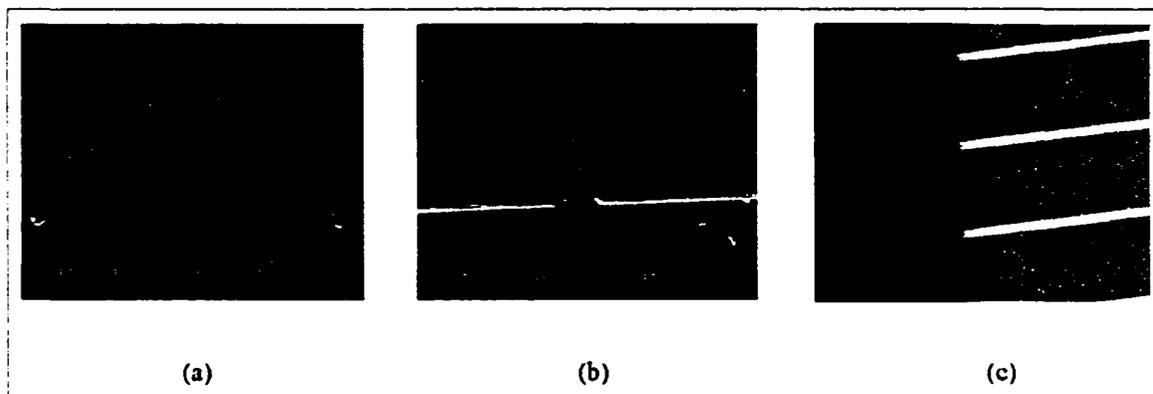
Figure 6.16 Microscope images of single-mode waveguides after (a) wafer sawing, (b) an imperfect cleave, and (c) a good cleave (top view).

After cleaving the devices, a number of optical microscope images of the GMZIs have been taken. Fig. 6.17 shows a 2 × 2 GMZI a 4 × 4 GMZI after the devices have been separated by cleaving. The final device sizes (not including access waveguides) are 3.5 × 1 mm (2 × 2 GMZI) and 7 × 2 mm (4 × 4 GMZI).

(a) (b)
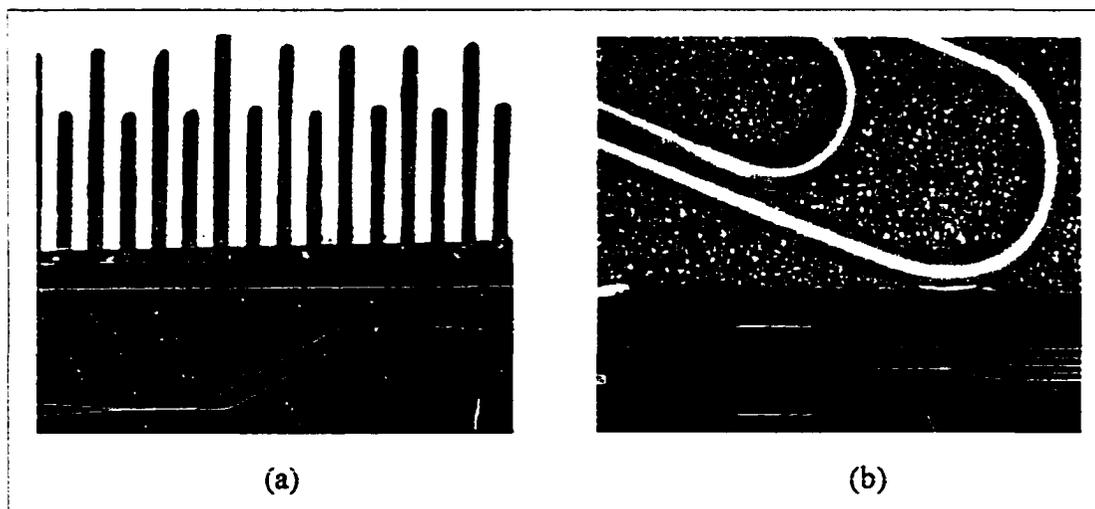
Figure 6.17 Optical microscope images taken after cleaving, showing (a) a 2 × 2 GMZI next to a ruler with adjacent markings separated by 500 μm, and (b) a 4 × 4 GMZI next to a standard paperclip.

In order to characterize the active devices, a special mount is needed, to ensure that light can be coupled to and from the input and output access waveguides (ports), and that electrical signals can be simultaneously applied to the thermo-optic phase shifters. This is depicted in Fig. 6.1(g). The mount is made by first cleaving a standard microscope slide to a size dictated by the sample stage in the experimental setup. The wafer is attached to this base using double-sided tape. Small pieces of circuit board containing copper electrodes are glued to either side of the base, and wire bonds are made from the electrode to a location adjacent to either end of a thermo-optic heater. The final electrical connections from the wire bond to the thermo-optic heater are made by a manual application of conductive (silver) paint. This is necessary because the relatively fragile rib waveguides developed in this work could not withstand the forces generated during the wire bonding process. Optical microscope photographs showing a mounted and bonded 4 × 4 GMZI are given in Fig. 6.18.

(a)   (b)
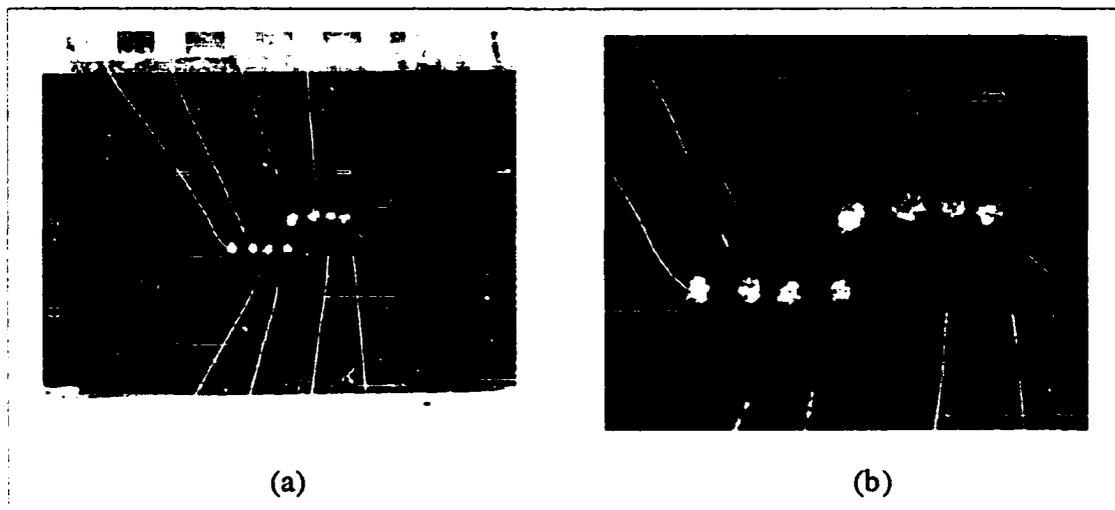
Figure 6.18 Optical microscope photographs showing (a) a mounted and wire-bonded 4 × 4 GMZI, and (b) a magnified view of the thermo-optic heater contacts, made using wire bonds and conductive paint.

The disadvantage of the bonding technique used in this work is that the light in the arms of the GMZI will suffer additional attenuation due to the conductive paint coating the waveguide arms. Furthermore, additional phase and amplitude errors may arise in the self-images, thereby degrading the performance of the GMZI. These and other experimental performance issues will be discussed in the following chapter.

Chapter 7

Device Performance

In this chapter the measured performance of the integrated devices fabricated in this work is presented. Through demonstrations of the device operation, the principles of the theory developed in earlier chapters are verified. The work presented in this chapter was performed during a period as a visiting scientist in the Lightwave Devices Group at the University of Twente, The Netherlands. Portions of this chapter have been reported in a post-deadline paper presented at the European Conference on Integrated Optics (ECIO '99) in Turin, Italy [110]; *in a paper to appear in the* Journal of Lightwave Technology [36]; *and in a paper submitted to the* Journal of Lightwave Technology [54].

7.1 Introduction

The realization of single-mode waveguides, MMI couplers, and GMZIs in a high refractive index contrast design using PECVD SiO$_2$/SiON technology has been described in the previous chapter. To assess the validity of the theory of the GMZI and the integrated device designs developed in earlier chapters, the experimental operation of the fabricated devices must be demonstrated.

A principal aim of this chapter is to provide insights into the feasibility of attaining multimode interference-based devices providing new and useful functions for optical network applications. A secondary goal is to show that the various choices made in the design and realization of the devices together provide the possibility of achieving compact, low-loss, polarization-insensitive silica-based devices that have a fast active response with low power consumption.

The results presented here are technology dependent and in many cases the device performance is sub-optimal with respect to some parameter. For example, relatively high propagation losses are seen because the films used have not been annealed. Imperfections in the fabricated waveguide sidewalls and thermo-optic contacts degrade the performance of the MMI couplers and GMZI devices, but such issues are amenable to improvement through development of the fabrication technology. This work attempts to tread a 'middle ground' between demonstrating the validity of theoretical concepts such as variable-ratio power splitting and multi-port switching in the GMZI on the one hand, while on the other hand utilizing a new waveguide design to investigate the possibility of polarization-insensitive, fast, and low-power device operation.

Often a number of iterative stages of device design, fabrication, and performance measurement are employed to meet the increasingly demanding requirements for the commercial viability of a new integrated optical device. It is the goal of this work to show that in a single iteration, the feasibility of realizing new types of devices can be assessed while simultaneously providing information into the technology possibilities available. It is the hope that this will considerably reduce the time frame for which these new devices can be employed in practical optical networks.

7.2 Experimental procedure

The performance of the devices in this work has been measured using the waveguide characterization setup in the Lightwave Devices Group at the University of Twente. The features of this setup will be described in this section. Also discussed here is the theory that has been used to estimate fiber-to-chip coupling losses in the setup. The coupling of light from a fiber into a planar waveguide chip typically results in some level of optical loss. Estimation of the amount of

this coupling loss enables the true waveguide propagation loss to be isolated. This is the subject of Section 7.3.

7.2.1 Experimental setup

A diagram of the experimental setup used in this work is given in Fig. 7.1.
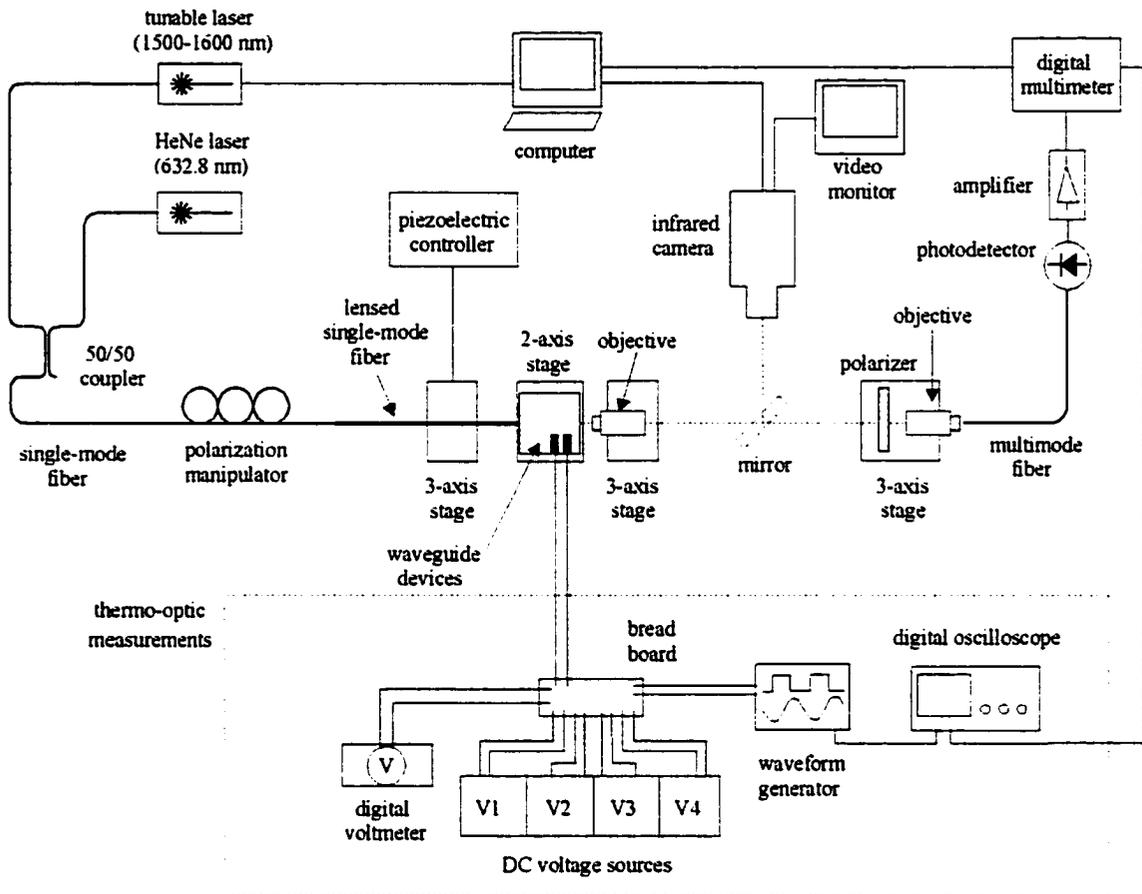
Figure 7.1 Diagram of the experimental setup used to test the integrated devices in this work. For passive device measurement, the equipment in the dashed box is not used.

The setup is located on an optical breadboard with vibration compensation, in a temperature-controlled room illuminated under yellow light. The light from a HeNe laser (used for alignment) and an infrared tunable laser (HP8168C, power used = 1-2 mW) are coupled together using a 50/50 optical fiber coupler, the output of which is sent into a polarization manipulator. The wavelength of the tunable laser can be computer-controlled, allowing a particular wavelength to be set or a range of wavelengths in the 1500-1600 nm region to be scanned. One end of a single mode fiber is connected to the output of the manipulator, and the other end (which is tapered into a lens) is mounted on a three-axis stage controlled by piezoelectric actuators. The lensed fiber is positioned such that its focal spot coincides with the input facet of a planar waveguide. The planar waveguide chip is fixed on a two-axis stage with a vacuum holder. A microscope objective lens (Spindler-Hoyer 20×, NA = 0.65), fixed on a manually-controlled three-axis stage, is used to outcouple and collimate the light from the planar waveguide.

At this point, two beam paths are possible. If the mirror, which is fixed on a rotatable holder, is rotated such that it intercepts the beam path, then the beam is deflected to an infrared camera. This camera is connected to a video monitor which enables the outcoupled light from a device to be viewed in real time. The camera is also connected to the computer, where image processing software is used to capture the video image for quantitative image analysis.

If the mirror does not intercept the beam path, then the beam passes through a polarizer and a microscope objective lens (Melles-Griot 10x, NA = 0.25) fixed on a three-axis stage. The lens focuses the beam into a multimode optical fiber (50 μm dia. core) with an output connected to a germanium photodetector. The resulting signal is fed into an electrical amplifier capable of small-signal gain values up to a factor of 10^9. The amplifier output is fed into a digital multimeter, which is also in communication with the computer.

Prior to device measurements, a reference measurement is made with the waveguide chip removed and the outcoupling objective placed in the focal plane of the lensed fiber. The detected signal is first maximized without the polarizer. The polarizer is then inserted and rotated to pass the desired polarization, and the manipulator is adjusted to maximize the detected signal. A wavelength scan of the lensed fiber output is then performed.

For thermo-optic measurements, electrode leads from the waveguide devices are connected to an electrical bread board, which in turn is connected with up to four independent DC voltage sources. A digital voltmeter is used to verify the applied voltages. To measure the dynamic response of the thermo-optic devices, a waveform generator (Philips PM 3217) capable of a $\pm$ 10 volt amplitude is substituted in place of one of the voltage sources. The pulsed voltage signal (input) is connected to one channel of a four-channel digital oscilloscope (Tektronix), and the device response (output) from the multimeter is connected to another oscilloscope channel.

7.2.2 Fiber-to-chip coupling

A major problem in integrated optics is the light coupling loss between a single-mode optical fiber and a planar waveguide device. This is due to the disparity between the mode field sizes in fibers and waveguides, as discussed in Section 5.3.1. An important parameter used in estimating the expected coupling losses in a given system is the mode field diameter (MFD). This is defined as the width of the fundamental mode field amplitude when it is approximated by a Gaussian curve (typically a good approximation), and is measured between points where the curve has dropped to a value of $1/e$ of the maximum (central) value [111].

When the light in an optical fiber is coupled to a waveguide, the dissimilar mode sizes result in a mode-mismatch and therefore a coupling loss. This loss can be calculated by determining the two-dimensional power overlap integral between two Gaussian fields with different half-widths w_1 and w_2, and is given by

$$\text{Coupling loss} = 20 \cdot \log\left[\frac{2w_1 w_2}{w_1^2 + w_2^2} \cdot e^{-\frac{|\delta d|^2}{w_1^2 + w_2^2}}\right] \tag{7.1}$$

where the effect of an axial displacement δd between the centers of the fiber and waveguide modes due to imperfect alignment has also been included in the calculation, and MFD = $2w$. The definition of the parameters in (7.1) is illustrated in Fig. 7.2.
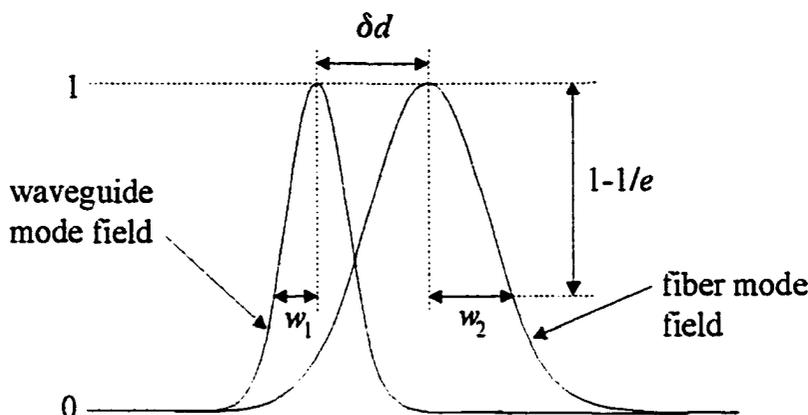
Figure 7.2 Definition of the parameters used in the coupling loss calculation (7.1).

Since the fields in Fig. 5.2 are approximately Gaussian in shape (verified by the EIM), the waveguide MFD is about 2.5 μm. By contrast, the MFD of a standard single-mode optical fiber is about 10.5 μm [112]. Using (7.1), the coupling losses between the waveguide mode field (w_1 = 1.25 μm) and various optical fiber mode fields (w_2) have been calculated for a number of axial displacement values. The results are plotted in Fig. 7.3.
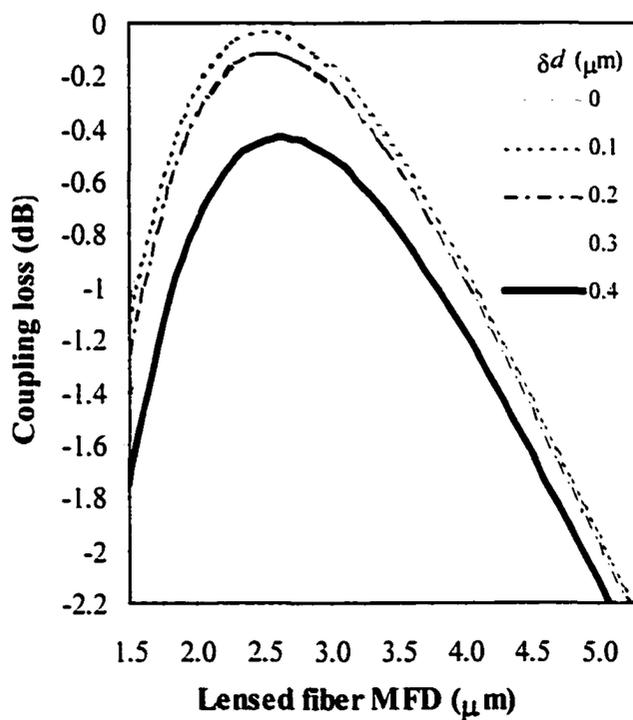
Figure 7.3 Coupling losses between the mode field of the single-mode waveguide used in this work (w_1 = 1.25 μm) and single-mode lensed fibers of varying field sizes, calculated using (7.1). Included are the coupling losses for various axial displacements δd between the mode fields.

The calculations indicate that an optical fiber MFD of less than 5 μm is required to maintain a coupling loss below 2 dB. The MFD of a standard single-mode optical fiber is therefore too large so special fibers must be used. Fortunately, a number of companies [113],

[114] produce optical fibers which are tapered and polished at one end to produce a lens effect, whereby the mode field of the fiber is focused to a small spot. Use of such 'lensed' fibers is common for coupling light into waveguides with small dimensions [67].

The lensed fibers used (SEFA Fujikura SM 10/125, made by TECOS) had a MFD of 4.2 μm. From Fig. 7.3, this results in a coupling loss in the range of 1.1 dB – 1.4 dB for an axial misalignment of less than 0.4 μm, which is achievable with the sensitive piezoelectric controllers used to position the lensed fiber.

7.3 Single-mode waveguide performance

The estimation of coupling losses enables the true propagation losses in the single mode waveguides to be experimentally determined. Since the dielectric films used in the fabrication have not been annealed, relatively high propagation losses are expected. The waveguide response, however, can be used as a reference level for determining the additional loss (excess loss), wavelength dependence, and polarization dependence introduced in more complex waveguide devices.

In this section, the polarization- and wavelength-dependent propagation losses in the fabricated waveguides are determined by the cutback technique using the setup in Fig. 7.1. First, however, the mirror is rotated into the beam path so that the mode field of the single-mode waveguide can be examined using an infrared camera image.

7.3.1 Mode field

An infrared camera image of the light ($\lambda = 1550$ nm) emerging from a single-mode waveguide is shown in Fig. 7.4. The optical intensity variation in the vertical and horizontal directions is obtained using image processing software. The mode field is roughly circular, which qualitatively confirms the predicted behaviour (i.e., the 20 % field contour line in Fig. 5.2 is almost circular).
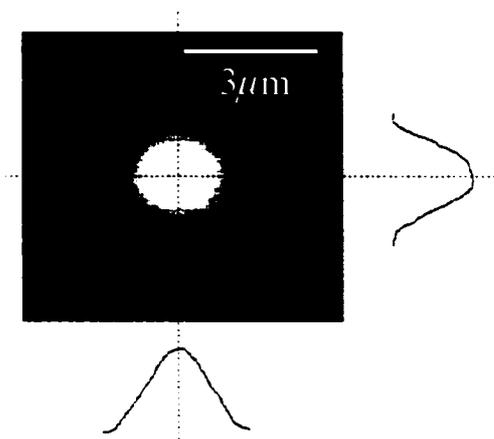
Figure 7.4 Infrared camera image showing the light emerging from a fabricated single-mode waveguide. Cross-sections of the intensity variation through the centre of the mode in the vertical and horizontal directions are also shown.

The image processing software has been used to count the number of pixels in the mode field in the vertical and horizontal directions, and this is compared to the dimensions of the video monitor in pixels. The mode field is then scanned across the video monitor by adjusting the micrometer positioners on the outcoupling stage. In this manner, the pixel dimensions are obtained, resulting in the scale shown in Fig. 7.4.

A slight horizontal asymmetry in the mode can be seen. The source of this is believed to be an imperfect RIE in some waveguides, resulting in a vertical waveguide sidewall on one side, and a slight angle in the sidewall on the other (see Fig. 6.14(a), for example). The measured MFD of the waveguide mode in Fig. 7.4 is 2.3 μm (vertical direction) and 2.4 μm (horizontal direction). Note that since the MFD is being determined from a field intensity scan, the values must be taken at the $1/e^2$ intensity points.

7.3.2 Propagation losses

Propagation losses in the single-mode waveguides have been measured using the cutback technique. In this technique, the ratio of the transmitted intensity through a waveguide to the intensity from the input fiber (reference) is measured, to arrive at the insertion loss:

$$\text{Insertion loss} = 10 \cdot \log_{10}\left(\frac{I^{out}_{waveguide\ output}}{I^{out}_{fiber}}\right) \tag{7.2}$$

The waveguide is subsequently cleaved to a shorter length and the new insertion loss is measured. This procedure is repeated several times. The insertion losses are then plotted versus the waveguide length, and the slope of the line is the propagation loss α in the waveguide, defined by (6.3) and typically measured in units of dB/cm.

Though the cutback technique is straightforward in principle, it has the drawback that the accuracy of the method is dependent upon the reproducibility of both the waveguide cleaves and the alignment used to couple light to and from the waveguide. This uncertainty is, however, partially offset by repeating the cutback measurement a number of times, and using the slope of a line fitted to the experimental points. To reduce the uncertainty further, the intercept of the fitted line with the loss axis has been fixed at a value based upon the fiber-to-chip coupling losses estimated from Fig. 7.3.

The cutback measurement has been performed on a number of straight single-mode waveguides by cleaving the waveguides and performing the insertion loss measurement three times. The measured insertion losses for a typical single-mode waveguide at λ = 1550 nm are plotted in Fig. 7.5 for both polarizations. The waveguide has been cleaved from its original length of 21 mm to 17.5 mm, 14 mm, and 10.5 mm. Also included are the least squares regression lines fitted to the data points, and their equations.

A value of 1.5 dB has been chosen for the coupling losses (y–intercept), based on the values in Fig. 7.3 and additional reflection losses at the waveguide input facet and at the 20× microscope objective lens. The large numerical aperture of this lens ensures that most of the light emerging from the output facet of the waveguide is collected, so further losses are assumed to be negligible. Using this value, the slopes of the fitted lines yield waveguide propagation losses of 3.417 dB/cm (TE) and 3.408 dB/cm (TM). These values indicate that the waveguide is relatively polarization insensitive, with a polarization dependent loss of less than 0.01 dB/cm. As expected, the losses in these waveguides are high since the films have not been annealed.
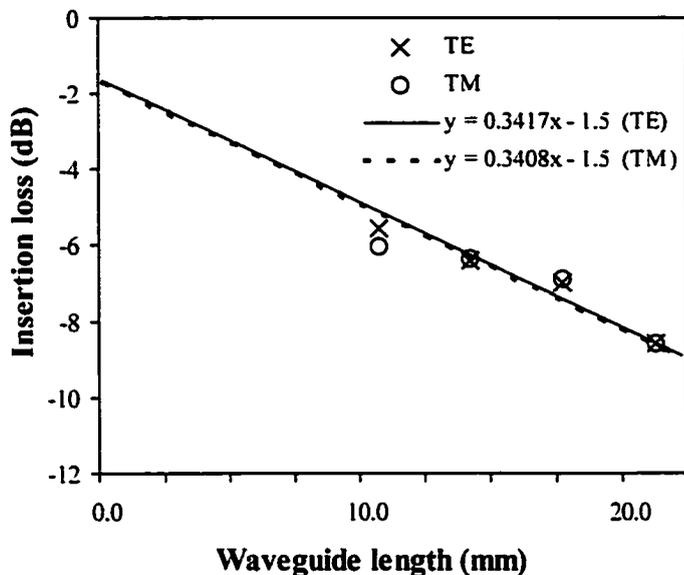
Figure 7.5 Measured insertion losses (data points) of a single-mode waveguide at $\lambda = 1550$ nm for four different cleaved lengths, for the determination of propagation losses using the cutback technique. The lines are fitted to the data points by a least-squares regression.

The data in Fig. 7.5 are plotted for a fixed wavelength of 1550 nm, however, a spectral scan ($\lambda = 1500 - 1580$ nm, in 2 nm steps) has been performed for each data point. The regression analysis is repeated at each wavelength to determine the spectral propagation losses, shown in Fig. 7.6.
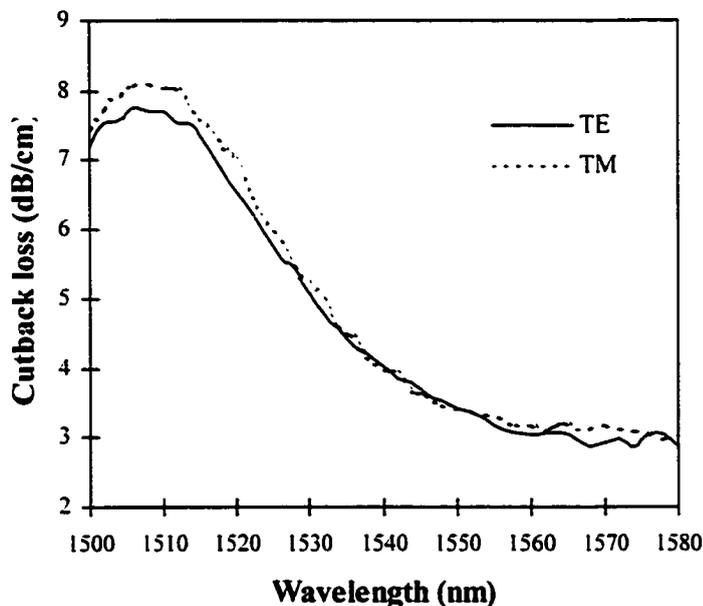
Figure 7.6 Spectral dependence of the propagation losses in a single-mode waveguide, measured using the cutback technique.

Propagation in the waveguides is nearly polarization independent over the entire wavelength range, but the polarization dependence becomes more pronounced at the shorter

wavelengths. The propagation losses in the waveguides also increase at these shorter wavelengths. A loss maximum is evident, and corresponds to the well-known absorption peak at $\lambda = 1508$ nm due to N-H bond resonances in the un-annealed SiO_2 and SiON films [70], [72]. In fact, the general shape of the curve in Fig. 7.6 is very similar to the transmission characteristic for the un-annealed SiON film in Fig. 6.5 in the same wavelength region. The curve is also similar to results reported by Wörhoff *et. al.* [64], where it has been shown that annealing can reduce the propagation losses by an order of magnitude at $\lambda = 1550$ nm.

A knowledge of the propagation losses in the fabricated single-mode waveguides enables the performance characteristics of more complex devices such as MMI couplers, to be determined.

7.4 MMI coupler performance

7.4.1 Definitions

In previous chapters, it has been established that MMI couplers with low excess losses and an equal power splitting ratio (low imbalance) are required for good GMZI performance. To gain an understanding of the GMZI response, it is thus necessary first to investigate the characteristics of the constituent MMI couplers. In this section, the measured excess losses and imbalance of 2×2 and 4×4 MMI couplers are compared to predicted values, thereby providing insights into the limitations of the device design and the fabrication process.

Imbalance is a measure of the deviation in the power splitting ratio of the $N \times N$ MMI coupler from the ideal case of equal power splitting. This has been defined in (3.28) in terms of field amplitude coefficients and is redefined here in terms of the measured intensity,

$$\text{Imbalance} = 10 \cdot \log_{10}\left(\frac{I_j^{out}}{I_{j_{ref}}^{out}}\right) \tag{7.3}$$

where j refers to the output port number in Fig. 3.1 and j_{ref} is a reference output port. Some ambiguity exists in (7.3), because an $N \times N$ MMI coupler would have $N - 1$ imbalance values and these values differ depending upon the reference port chosen. To circumvent this problem and arrive at a single imbalance value for an $N \times N$ MMI coupler, in the results presented here the largest magnitude from the set of $N(N-1)$ imbalances (where all ports are used as references) is used.

The excess loss for a switch was defined in (4.5), however, the excess loss in the MMI coupler is defined as the ratio of the total intensity at the MMI coupler outputs to the output intensity from a reference single-mode waveguide running parallel to the MMI coupler, i.e.,

$$\text{Excess loss} = 10 \cdot \log_{10}\left(\frac{\sum_{j=1}^{N} I_j^{out}}{I_{ref}^{out}}\right) \tag{7.4}$$

This definition is used because it isolates the loss due to imperfect MMI coupler imaging and does not include losses arising from the access waveguides. This enables a meaningful comparison of measured and the simulated coupler losses to be made, since the loss values from the MPA simulation performed in Chapter 5 assume no access waveguide losses.

7.4.2 Performance vs. imaging length

As mentioned in Section 5.6 and shown in Fig. 5.11, MMI couplers with five different imaging lengths near $L_{M,N \ design}$ have been fabricated. The excess loss and imbalance measured in the 2×2 and 4×4 MMI couplers is shown in Fig. 7.7, along with the simulated results from Fig. 5.6.
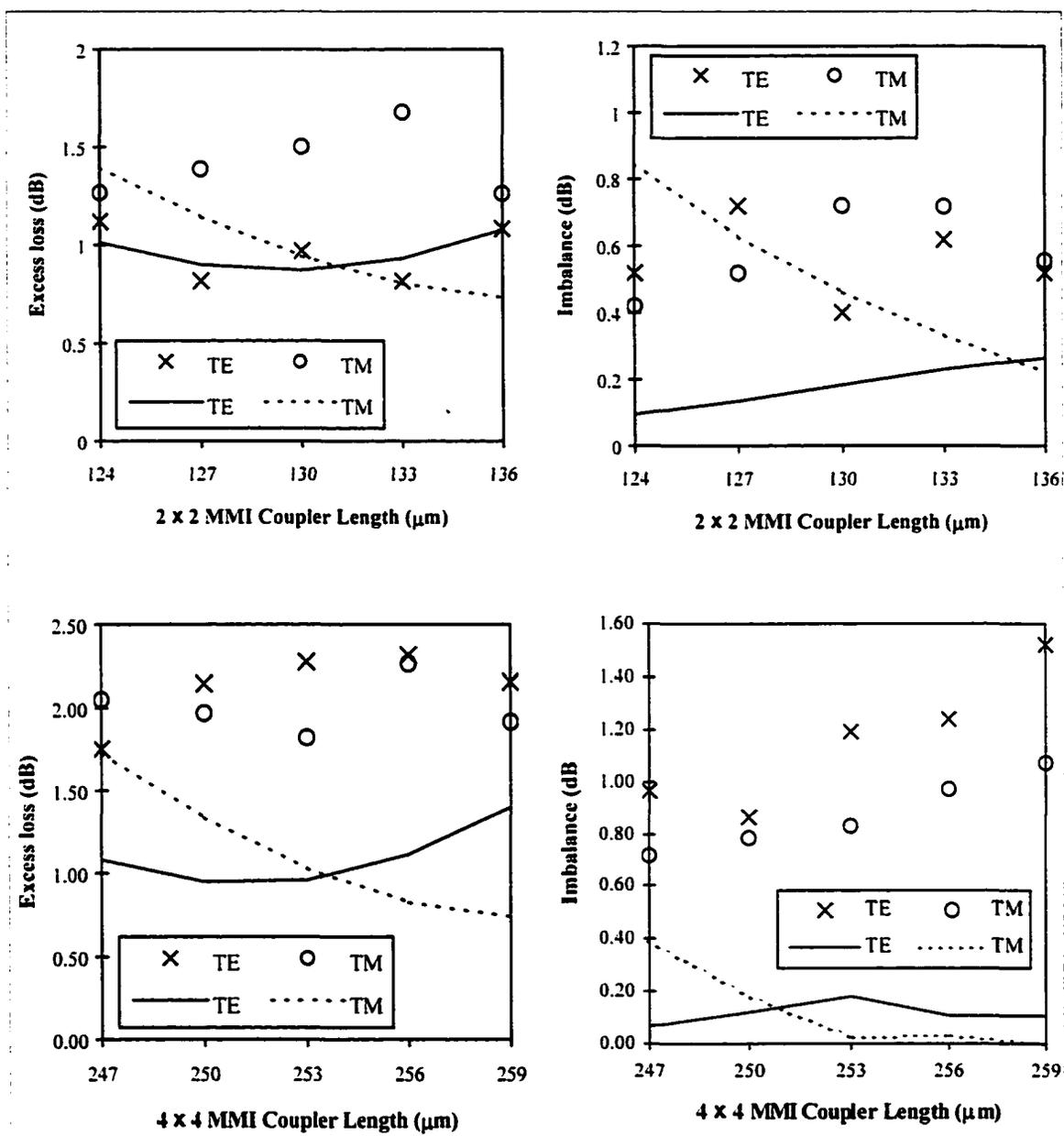
Figure 7.7 Measured ($\times$ and o) excess loss and imbalance in 2×2 and 4×4 MMI couplers for five different coupler lengths, along with MPA simulation results (solid and dashed lines) reproduced from Fig. 5.6.

In all cases, the magnitude of the polarization dependence in the couplers is similar to the simulated results. The measured values also show the general trend that the coupler performance is relatively insensitive to the imaging length near the optimum design length.

The measured excess loss and imbalance in the 2 × 2 MMI couplers show a general agreement with the calculations. The measured performance of the 4 × 4 MMI couplers, however, is somewhat poorer than predicted. Measured excess loss and imbalance exceed the simulated values by about 1 dB and 0.7 dB, respectively in the 4 × 4 device. One possible source of this is the different s-bends used for the access waveguides to the 4 × 4 MMI coupler, which have been described in Section 5.4.2. Although in BPM simulations the losses incurred in both bends are negligible, in practice it is commonly seen that a finite amount of bend loss is always present [88], [104]. Since the bend loss is not included in the straight reference waveguide, the measured excess loss value is higher than expected. Since the different s-bends result in differing losses, the imbalance at the coupler outputs is also higher than expected.

Another possible source of imbalance is the imperfect lithographic definition of the MMI couplers. The rounding effect seen in the region near the access waveguides in Figs. 6.12 and 6.13 can lead to a partial leakage of light to an adjacent output port, thereby affecting the balance.

7.4.3 Performance vs. wavelength

Emerging optical communications networks aim to exploit the transmission capacity of optical fibers through the use of WDM techniques. In such networks, the signal to be routed may consist of a discrete wavelength channel or a number of multiplexed channels. It is therefore desirable that power splitting and switching devices for these networks have a relatively wide bandwidth of operation.

The bandwidth of the GMZI devices in this work will be discussed in the following section, however the results are ultimately dependent upon the bandwidth of the individual MMI couplers. For this reason, the wavelength-dependent properties of the MMI couplers have been investigated. Rather than presenting a rigorous analysis of the wavelength-dependent properties in the MMI coupler which can be found elsewhere [31], [115] it is the goal of this section to present the general behaviour of the fabricated MMI couplers over a range of wavelengths.

The excess loss and imbalance in the MMI couplers at the design length $L_{M,N\ design}$ has been measured for the wavelength range 1530 – 1570 nm, and the results are shown in Fig. 7.8.
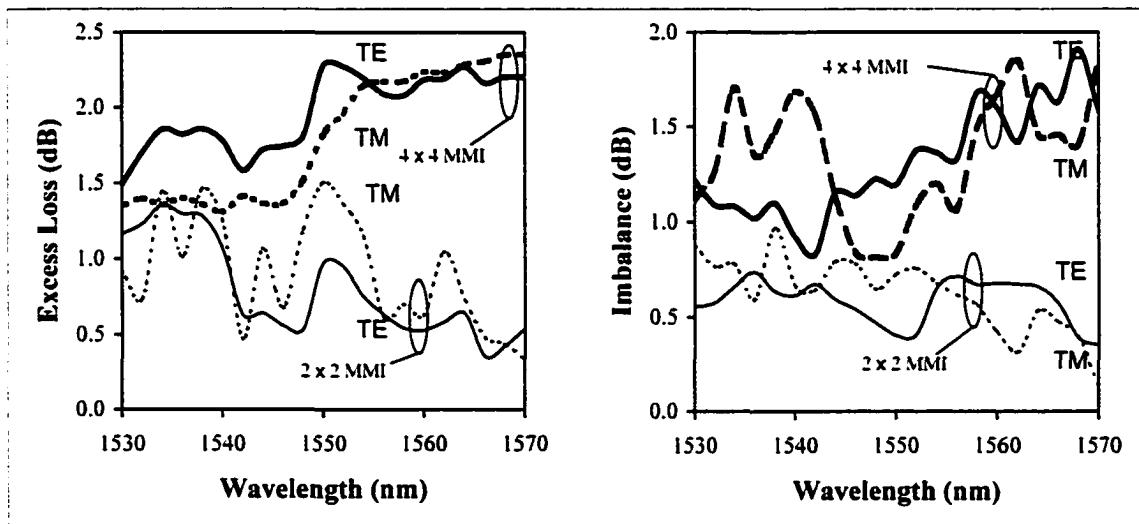
Figure 7.8 Measured wavelength variation of excess loss and imbalance in the fabricated MMI couplers. Results for the 2 × 2 and 4 × 4 MMI couplers are given by regular and bold lines, respectively. Results for the TE and TM polarizations are given by solid and dashed lines, respectively.

A typical variation of about ± 0.5 dB in excess loss and imbalance over the result at 1550 nm is evident for the 40 nm measurement range. Although both polarizations exhibit the same general behavior, some wavelength-dependent effects are evident. This is expected, since the propagation constants of the modes in the coupler given by (2.22) are both wavelength and polarization dependent. Hence, the coupler beat lengths also vary with these parameters.

7.4.4 Visual characterization

The optimization of fiber-to-chip coupling was often aided by utilizing an infrared camera image on a video monitor. The positioning of the equipment in Fig. 7.1 is configured such that when the image of a planar waveguide output is in focus and centered on the monitor, the corresponding signal level at the multimeter is maximized when the mirror is rotated out of the beam path.

The infrared camera also allows the light emerging from all output ports of the MMI couplers (i.e., all I_j^{out}) to be observed simultaneously. An example of this is shown in Fig. 7.9, which shows an 8-bit grayscale image of the outputs from a 4 × 4 MMI coupler and the resulting intensity scan.
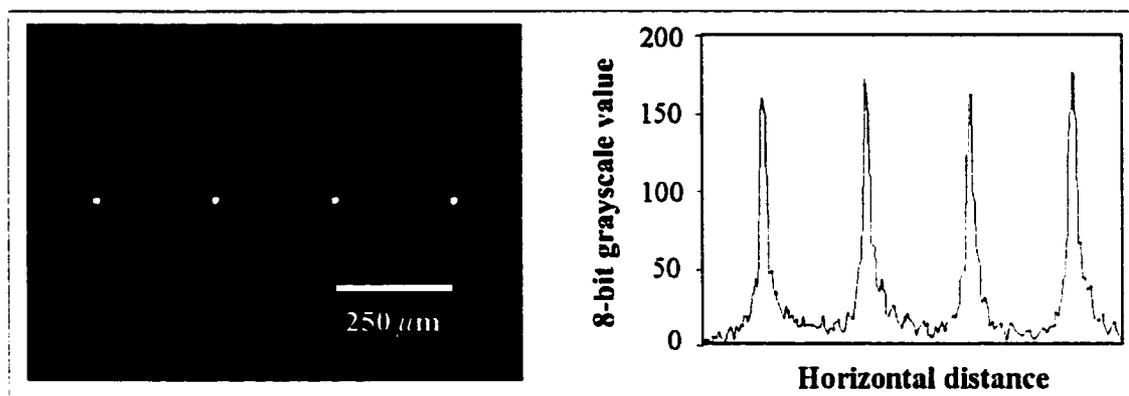
Figure 7.9 Infrared camera image of the output plane of a 4 × 4 MMI coupler, and the resulting intensity scan through a line crossing the centers of the waveguide mode fields.

7.5 GMZI performance

7.5.1 The 2 × 2 GMZI

A schematic diagram of the 2 × 2 GMZI is shown in Fig. 7.10.
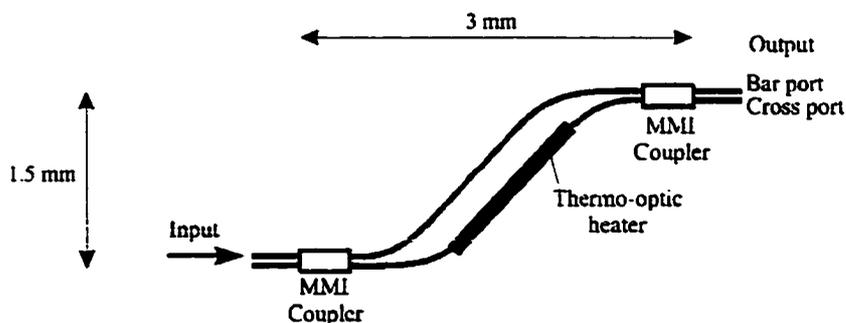
Figure 7.10 Diagram of the layout of the fabricated 2 × 2 GMZI.

The measured electrical resistance of the heater region varied from device to device due to different heater lengths, which were defined by the manual application of conductive paint to connect the heater to an adjacent wire bond. Typical measured resistances of the 2 mm long heater were in the 2.0 kΩ – 2.5 kΩ range.

The measured excess loss of the best switch over a straight, single-mode reference waveguide was –3.0 dB for both polarizations, not including the loss due to the conductive paint contacts. After the contacts were made, additional excess losses of 10 dB (TM) and 15 dB (TE) were measured. Recalling the discussion in Section 5.3.4, this result is expected since direct contact of the metal with the waveguide sidewall increases the absorption of the TE mode preferentially.

For active operation, light at λ = 1550 nm is coupled into an input port and the intensity at the 'bar' and 'cross' output ports are measured as the power applied to the heater is increased. The waveguide chip is mounted on a water-cooled stage maintained at a temperature of 15 °C. The resulting switch characteristic for the 2 × 2 GMZI is shown in Fig. 7.11, where the excess loss contribution from the conductive paint has been removed.
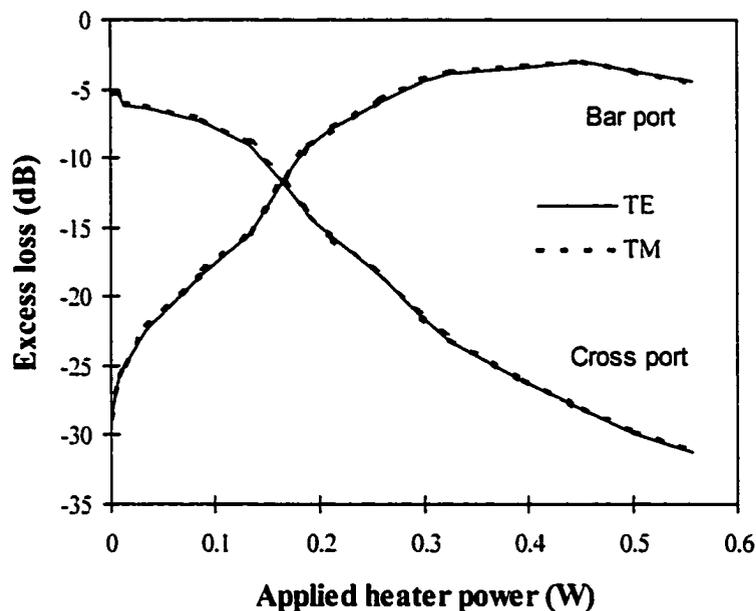
Figure 7.11 Measured switch characteristic of a fabricated 2 × 2 GMZI.

The extinction ratio is a measure of the dynamic range of power in an output port k and is defined as

$$\text{Extinction ratio} = 10 \cdot \log_{10} \left(\frac{\max\{I_k^{out}\}}{\min\{I_k^{out}\}} \right) \qquad (7.5)$$

where the functions max{ } and min{ } represent the maximum and minimum values, respectively, from the continuum of values I_k^{out} can assume. Ideally these states correspond to 100 % and 0 % of the light emerging from an output port k, respectively. For the 2 × 2 GMZI in Fig. 7.11, the extinction ratios are 25.8 dB for both the bar and cross ports.

The crosstalk at the output ports is –23.5 dB in the passive state, and improves to –26.7 dB with a switching power of 557 mW, achieved by applying a DC potential of 34.1 V across the 2.10 kΩ heater resistance. This applied power corresponds to a phase shift of π and is

significantly larger than the value of 100 mW predicted in Section 5.3.4. The major cause of this is believed to be the poor quality of the applied conductive paint connections, resulting in excess power dissipation at the heater contacts. As a secondary effect, the dissimilar metals used for the thermo-optic heater, the conductive paint, the wire bonds, and the electrodes are expected to further reduce the switch efficiency.

Since the phase shift is proportional to the applied heater power (see Section 5.3.4), the expected switching characteristic of a 2 × 2 GMZI is given by (3.17) and (3.18). The source of the asymmetry in Fig. 7.11 is believed to be differing attenuation and phase errors in the GMZI arms and imbalance and phase deviations in the MMI couplers, which alters the switching characteristic in a manner similar to that shown in Figs. 3.4 and 3.5.

Dynamic or time-dependent switching in the 2 × 2 GMZI has been measured by applying a square-wave voltage to the phase shifter and monitoring the intensity response at the output ports using a digital oscilloscope. The resulting oscilloscope trace is shown in Fig. 7.12.

Tek Run: 10.0kS/s

Ch1   10.0 V   Ch2   250mV   M5.00ms   Ch1 ∫   5.0 V
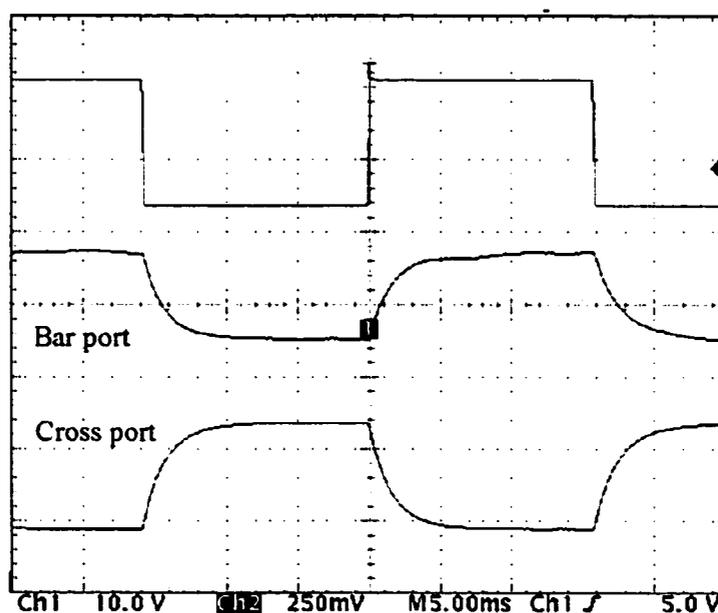
Figure 7.12 Oscilloscope trace showing the optical response of the 2 × 2 GMZI outputs to an applied square-wave heater voltage (top waveform). The horizontal time scale is 5 ms/div.

The bar and cross ports are complimentary and follow the response predicted by (3.17) and (3.18). Measured rise and fall times for the 2 × 2 GMZI were typically in the 2 – 4 ms range. The switch is sufficiently fast to be used for optical network restoration functions, however, the total switching time will increase if a number of switches are used in a cascaded configuration.

In general, the fabricated 2 × 2 GMZI shows good crosstalk and polarization behavior. Based on the measurements of fabricated 2 × 2 MMI couplers, the principal source of excess loss in the 2 × 2 GMZI is the MMI couplers. Additional losses arise from the longer single-mode waveguide length and the bends in the GMZI, which were not present in the reference waveguide. The thermo-optic switching is fast, but requires a high power level and thus cooling of the devices.

7.5.2 The 4 × 4 GMZI

A schematic diagram of the 4 × 4 GMZI is shown in Fig. 7.13.
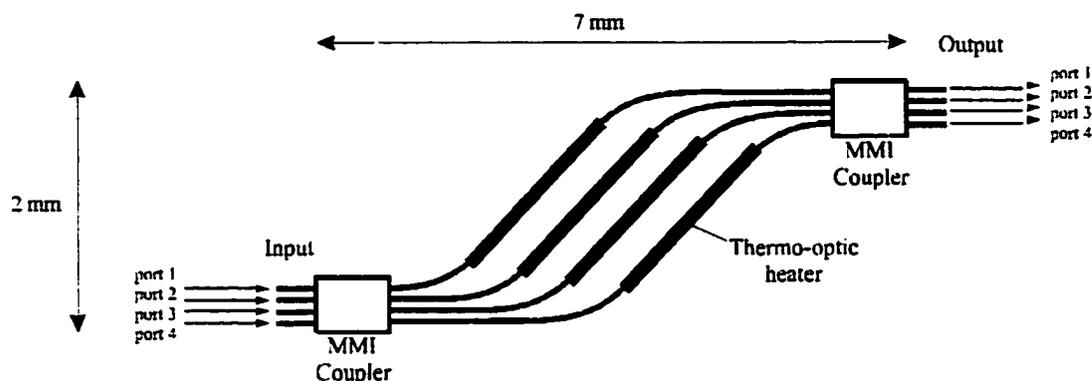
Figure 7.13 Diagram of the layout of the fabricated 4 × 4 GMZI.

For the 4 × 4 GMZI devices, an improved procedure for applying the conductive paint to the heaters resulted in smaller and more uniform heater contacts and resistances. For the device pictured in Fig. 6.18, the 3 mm long heater resistances are 2.91, 3.22, 3.29, and 3.31 kΩ for $j = 1,2,3,4$, respectively.

The measured excess loss in the device over an adjacent reference waveguide (also seen in Fig. 6.18) is 4.8 dB (TE) and 5.7 dB (TM). The source of the polarization dependence is believed to be the imaging in the MMI couplers and a differing birefringence in the two different access waveguide bends used. The excess loss was measured in the passive state before the final heater connections were made. Additional excess losses from the conductive paint were about 3 dB (TM) and 8 dB (TE). This improvement over the 2 × 2 GMZI result is a direct result of the smaller conductive paint contacts obtained with the improved application procedure.

The switch characteristic for the 2 × 2 GMZI has been given in Fig. 7.11, but the equivalent characteristic for the 4 × 4 GMZI is not well defined since four phase shifters may be operated independently. Recall, however, that in Section 3.5.2.3 a cross-section of the characteristic response of a non-ideal 4 × 4 GMZI was simulated and compared to the response of an ideal device. This procedure is now repeated to provide insights into the performance of the fabricated 4 × 4 GMZI.

To obtain a plot similar to Fig. 3.2, an input port $i = 2$ is used and the output ports of the 4 × 4 GMZI are monitored while increasing the electrical power applied only to the arm $j = 4$. Since the intensity at all four outputs must be examined simultaneously, line scans of infrared camera images similar to the one in Fig. 7.9 are used to determine the relative intensity at the output ports while the heater power is varied. The result is shown in Fig. 7.14 for the TE polarization.

The periodicity in Fig. 7.14 is characteristic of an interferometer and enables the efficiency of the phase shifter to be determined. From Fig. 7.14, an applied heater power of 204 mW results in a complete period, or a 2π phase shift. Measurements similar to this have been performed for all four phase shifters. The resulting power required for a π phase shift in the four arms is 115.5, 121.5, 103, and 102 mW, for $j = 1,2,3,4$, respectively.

These results can be compared to the thermal simulation prediction of an applied power of 50 mW for a π phase shift. Though the measured power is twice this value, the results nevertheless represent a significant improvement over the 2 × 2 GMZI result. Note also that the curves in Fig. 7.14 are symmetric and sinusoidal, similar to the ideal GMZI response in Fig. 3.2.

Finally, the result in Fig. 7.14 does not change when the devices are water-cooled, indicating that excess heat dissipation in the heater contacts is low. By contrast, the switching power for the 2 × 2 GMZI increases by about 50 mW when the device is not cooled.
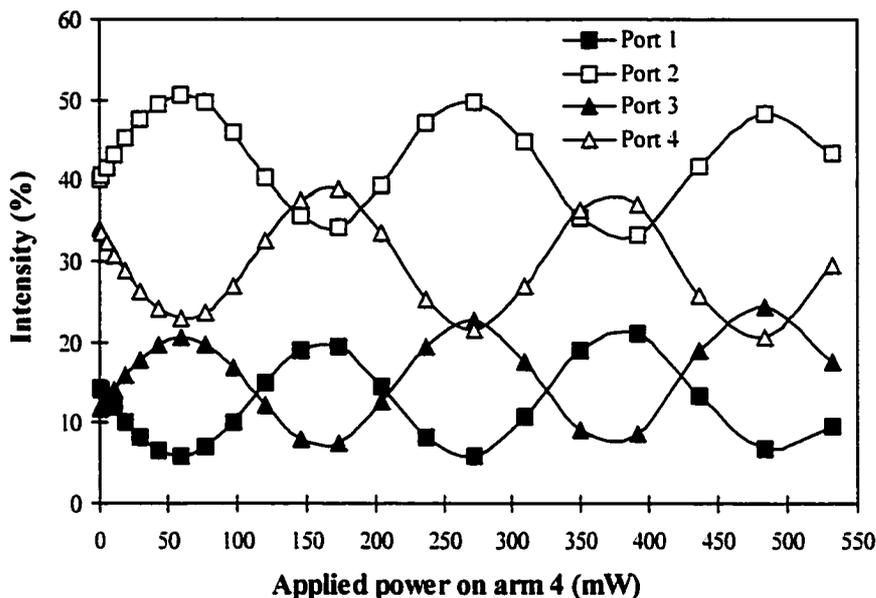
Figure 7.14 Measured response from a 4 × 4 GMZI to a single phase shift applied to arm 4. The lines are fitted to the data points.

The experimental curves of Fig. 7.14 are reproduced for a single period in Fig. 7.15 (dashed lines) and plotted with the response of the ideal 4 × 4 GMZI (solid lines) reproduced from Fig. 3.2.
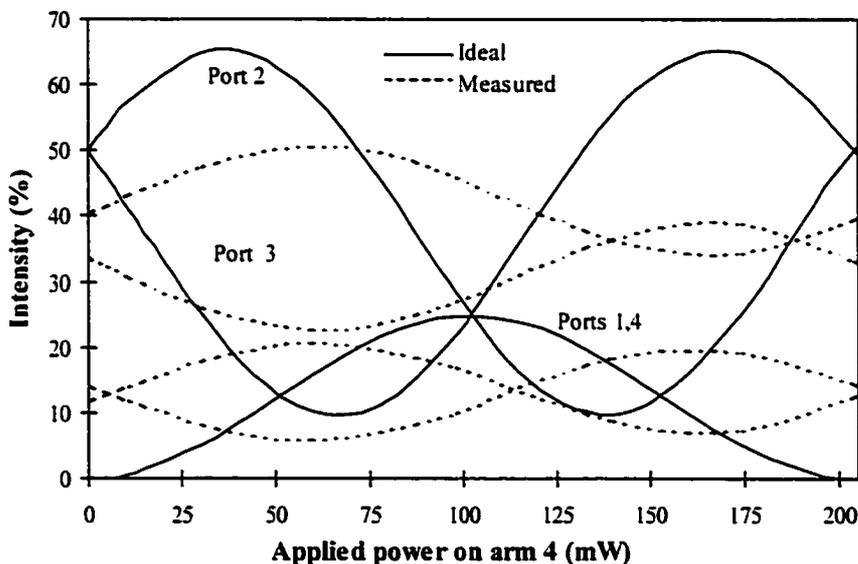
Figure 7.15 The measured response of the 4 × 4 GMZI from Fig. 7.14 (dashed lines) superimposed over the response of the ideal device from Fig. 3.2 (solid lines), for an applied phase shift of up to 2π.

The measured device response is in general phase shifted, the amplitude of intensity variations is diminished, and the level about which the intensity varies sinusoidally is offset with respect to the ideal device. Similar results were noted for the 2 × 2 GMZI. Consider for a moment a mathematical description of this result. Arbitrarily choosing the port $k = 2$ in Fig. 7.15, in the ideal GMZI the relative intensity in this port should vary according to the result predicted by (3.13), i.e.,

$$\frac{I_2^{out}}{I_0} = \frac{1}{8} \cdot \left[3 + 2\sin(\Delta\varphi_4) + \cos(\Delta\varphi_4) \right] \tag{7.6}$$

The equation of the measured response for this port in Fig. 7.15, however, is given by

$$\frac{I_2^{out}}{I_0} = \frac{1}{8} \cdot \left[3.32 + 0.63\sin(\Delta\varphi_4 - 0.45) + 0.3\cos(\Delta\varphi_4 - 0.45) \right] \tag{7.7}$$

To understand the sources of this discrepancy, the result (7.7) should be compared to the theoretical model of the non-ideal GMZI in (3.32). Phase deviations in the MMI couplers and phase errors in the GMZI arms both contribute to the −0.45 radian shift, the imbalance in the MMI couplers contributes to the amplitude attenuation in the sine and cosine terms, and a combination of phase and imbalance effects contribute to the offset of 0.32 in the constant term in (7.7).

Note that phase errors and imbalance arising from the conductive paint in the arms and the differing output access waveguide lengths have not been included in this analysis. Clearly, these factors also contribute to the observed response.

A qualitative comparison of Fig. 7.15 with Fig. 3.4 confirms that MMI coupler deviations are present in the fabricated device. Phase errors in the arms are also present and a comparison of Fig. 7.15 with Fig. 3.5 suggests they are distributed approximately randomly among a range of values between 0 and 2π.

7.5.2.1 1 × N GMZI switching

The unknown (and likely large) random phase errors in the arms of the fabricated GMZI result in an unpredictable response for a given set of applied phase shifts. In Section 3.5 it was shown that the phase shifts applied to the GMZI arms can include a bias component to compensate for the phase errors in the arms. The response of the fabricated GMZI is thereby brought closer to the ideal response, so that the device behaves more predictably with an applied set of phase shifts. This enables the practical operation of the GMZI.

Since the phase errors in the GMZI arms are not known, the following experimental procedure utilized by Jenkins et. al. [34] is used to compensate for these phase errors. The device is first brought into a switching state, where ideally all of the output light emerges from a single port. This has been done in the 4 × 4 device by adjusting the voltage applied to the four thermo-optic phase shifters while monitoring the optical intensity emerging from the output ports using an infrared camera. The camera signal is sent to a computer, where image processing software is used to provide a line scan of the image. An iterative manual phase shift adjustment technique is then used to bring the device to a switch state, corresponding to maximizing the peak on the line scan for a given output. The bias phase shifts are then the difference between the applied phase shifts and the theoretical phase shifts given by the relevant column of the ideal switching phase shift matrix X given in (4.23).

Using these bias values, the phase shifts necessary for the other switching states have been calculated and applied to the device, to obtain the remaining switch states. In practice, however, only approximate switch states are achieved by this method. It has been found that fine-tuning of the phase shifts is often needed to improve the switch state. This can be understood intuitively by noting that the phase shifts from the ideal switching matrix are calculated by solving (3.13). The non-ideal device, however, includes MMI coupler deviations,

so the system (3.32) must be solved, setting the $\delta\varphi_j = 0$. This cannot be done, however, since the exact imbalances and phase deviations in the MMI couplers are unknown. Unless these parameters can be rendered negligible, the procedure of operating a GMZI will always have some associated uncertainty.

Line scans of the four switch states obtained using the described method are shown superimposed in Fig. 7.16, for the TE polarization. The scans for the TM polarization are similar.
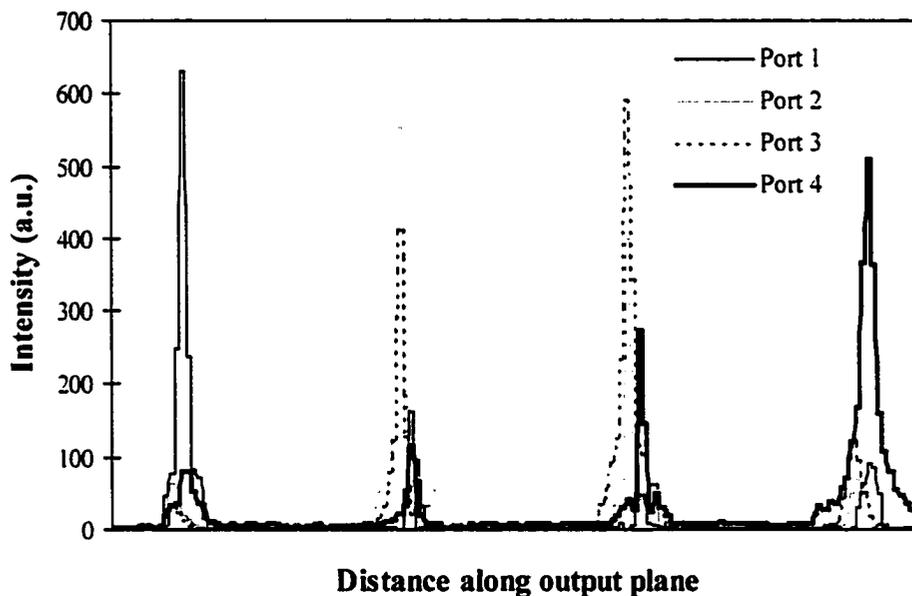
Figure 7.16 Superimposed line scans of the four switch states of the 4 × 4 GMZI taken from an infrared camera image. The phase shifter voltages were adjusted manually using an iterative technique to maximize the output light emerging from each port.

As mentioned above, although compensation of the phase errors in the GMZI arms is possible, imbalance and phase deviations in the MMI couplers cannot in general be compensated by the applied phase shifts, resulting in the observed crosstalk and non-uniformity in the switch states. Typical measured crosstalk values were –7 dB and a best result of –11 dB has been obtained. Measured extinction ratios for all ports were 15 – 18 dB.

The poor performance of the switching function in the GMZI shown in Fig. 7.16 can be compared to the theoretical results obtained in Table 4.1. Crosstalk values in the 1 × 4 switch for various MMI coupler deviations and phase shift fluctuations were shown in Table 4.1, assuming fully compensated phase errors in the arms. Degradation in the imaging of the MMI couplers is seen to significantly increase crosstalk in the switch. If a small amount of uncompensated phase errors in the arms are added to this, the result is the poor measured performance.

The trend in Table 4.1 shows that the imaging in the MMI couplers must first be improved to lower the crosstalk level. As the imaging approaches that of the ideal coupler, however, the sensitivity of the crosstalk to the accuracy of the phase shifting increases, indicating that simultaneous improvement of both the fabrication and control are needed to obtain good switch performance.

The dynamic response of the switch has been measured by applying a pulsed voltage to different heaters and using the digital oscilloscope to monitor the response in an output port. In Fig. 7.17 the oscilloscope trace for two arbitrarily chosen output ports is given.

Tek Run: 25.0kS/s
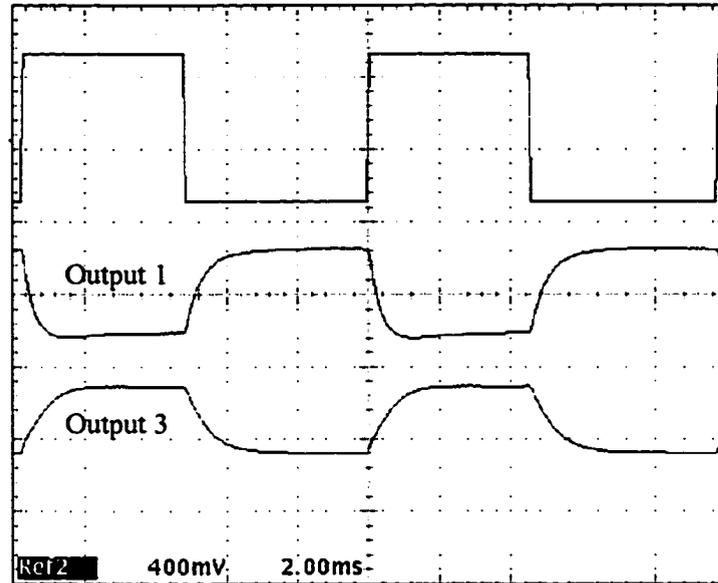
Figure 7.17 Oscilloscope trace showing the optical response of two 4 × 4 GMZI outputs to an applied square-wave heater voltage (top waveform). The horizontal time scale is 2 ms/div.

Note that in the 4 × 4 GMZI the response is faster than in the 2 × 2 GMZI. The rise and fall times are below 1 ms in this device, and switching times as low as 460 μs (rise) and 650 μs (fall) have been measured. This fast response is the result of the improved thermo-optic heater contacts.

As mentioned in Section 5.3.4, a goal in this work has been to minimize the switching time × switching power product in the thermo-optic switches. For the 2 × 2 GMZI described in the previous section this product is 1114 mW·ms. The best value obtained for the 4 × 4 GMZI is 46.9 mW·ms. These values are put into perspective by a comparison with the performance of other silica-based thermo-optic switches published in the literature, given in Table 7.1. It must be noted that the switches are designed to meet a number of requirements such as low loss, low crosstalk, and polarization insensitivity in addition to the efficient thermo-optic operation, so some of the switches in Table 7.1 have not been optimized for this latter requirement.

Switch type	Material technology	Power × time product (mW·ms)	Year	Ref.
2 × 2	Ti: SiO₂	440	1990	[80]
2 × 2	Ge: SiO₂	96	1993	[81]
2 × 2	SiON/SiO₂	1500	1997	[68]
2 × 2	SiOₓ/SiO₂	16.5	1998	[67]
2 × 2	Ge: SiO₂	900	1999	[116]
2 × 2	SiON/SiO₂	1114	this work	
1 × 4	SiON/SiO₂	46.9	this work	

Table 7.1 Comparison of the performance of the devices in this work with other published results for silica-based thermo-optic switches.

The switch in [67] has a rib structure similar to the one used in this work, and a continuous thin-film low-resistance aluminum heater electrode to which a voltage is directly applied. The result is fast (150 μs rise time) and low power (110 mW) thermo-optic operation. A drawback of the approach, however, is an increase in the complexity of fabrication as it requires an additional photomask to define the heater electrode regions.

The performance of the present switches is expected to improve with a similar electrode design, which would remove the necessity for wire bonding and the application of conductive paint. Based on thermal simulations, the expected time × power product with optimized electrodes is about 5 mW·ms.

7.5.2.2 Multi-port switching

The $N \times N$ blocking function in the GMZI has been described in Section 4.4. When a GMZI is set to various switch states, N inputs appear in certain permutations at the output ports, given by (4.11). This multi-port switching function has been demonstrated in the fabricated 4 × 4 GMZI. The switch is first set to route input port 1 to output port 1, or (1,1), using heater voltages determined by the method described in the previous section. An infrared camera image of the outputs is then taken. Next, the lensed fiber is moved to input ports 2, 3, and 4 and camera images are taken, while the same applied heater voltages are maintained. The four images along with their corresponding line scans are shown in Fig. 7.18.
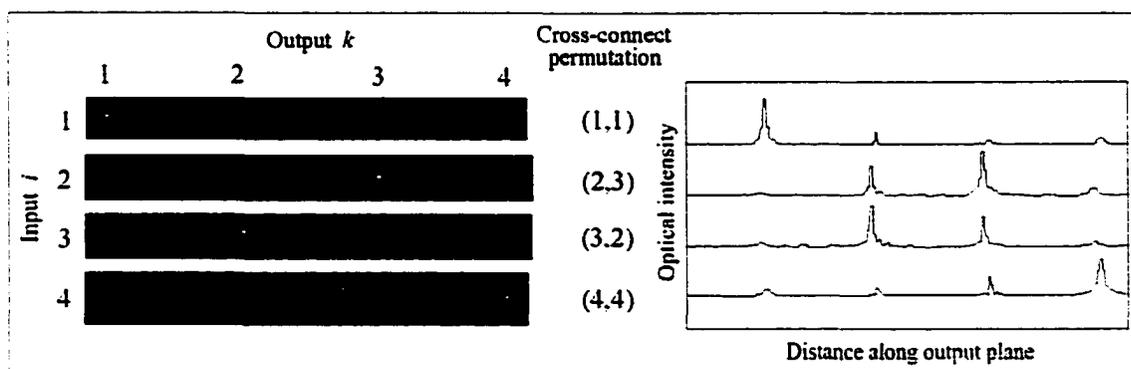
Figure 7.18 Infrared camera images and line scans of the outputs from a 4 × 4 GMZI set in a switch state, for light launched into each of the input ports. The result is one of the N cross-connect permutations of the blocking switch, given in the first column of Table 4.5.

Though crosstalk is evident, the multi-port switching capability of the GMZI is clearly discernable from Fig.7.18. The relatively high crosstalk level of the 1 × N switch for ports 2 and 3 shown in Fig. 7.16 is also seen in the line scans of the multi-port switch in Fig. 7.18. Again, the main cause is the imbalance and phase deviations present in the MMI couplers.

7.5.2.3 Variable ratio power splitting

The procedure used to demonstrate the switching function in the GMZI has also been used to demonstrate the variable-ratio power splitting function. An arbitrary power-splitting state is first chosen. The required theoretical phase shifts are then calculated using the method described in Section 3.4.2. The bias phase shifts calculated in Section 7.5.2.1 are then added to the theoretical phase shifts, and the equivalent voltages are applied to the heaters of the 4 × 4 GMZI.

As in the case of the switch, some fine-tuning to the heater voltages is generally required to obtain an output intensity distribution close to the desired ratio. The tuning is performed while viewing a line scan of the GMZI outputs superimposed on the simulated intensity scan, calculated using MPA. The results for four power splitting states are shown in Fig. 7.19.
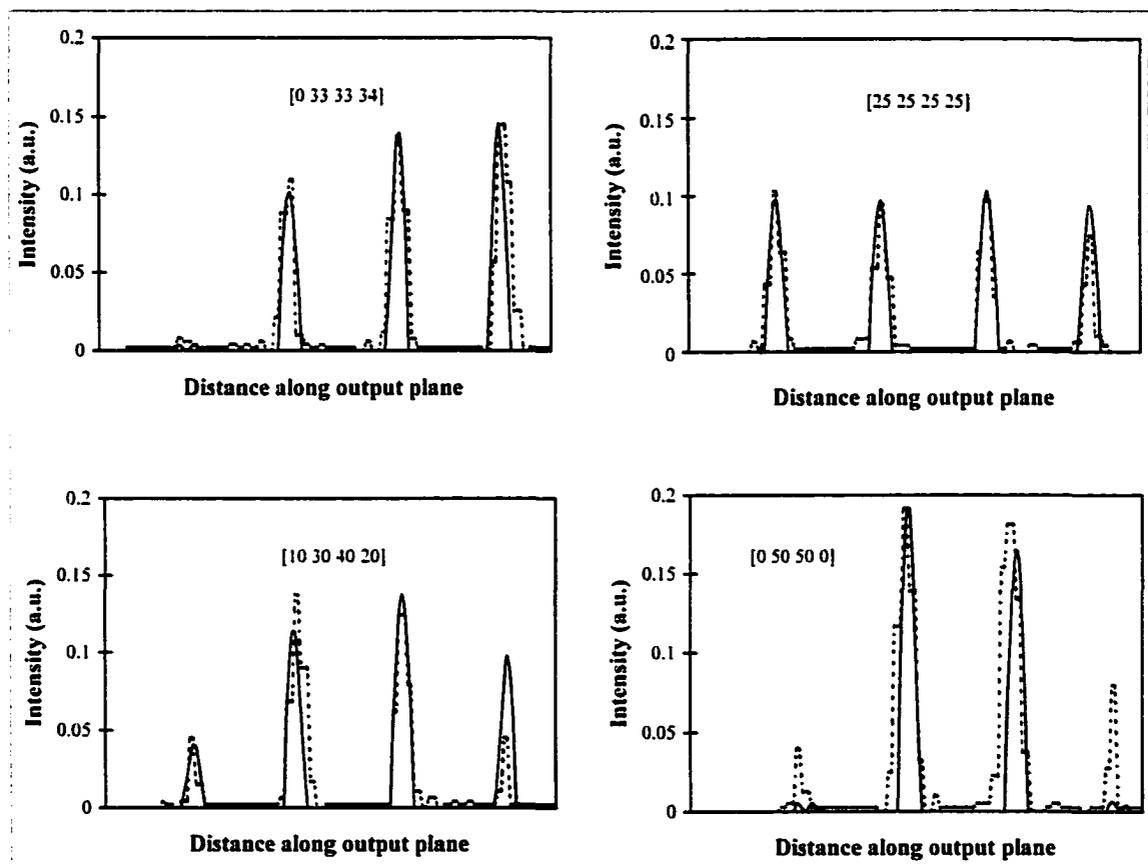
Figure 7.19 Measured (dashed line) and MPA simulated (solid line) intensity variation at the outputs of a 4 × 4 GMZI for four variable-ratio power splitting states. The desired power ratios in percentages are given in square brackets, for output ports k = 1,2,3,4, respectively.

The result in Fig. 7.19 is an experimental verification of the theory of variable-ratio power splitting developed in Chapter 3. It is evident From Fig. 7.19 that in the fabricated device some power splitting ratios can be obtained accurately while others cannot. As discussed in the

theory of the non-ideal GMZI in Section 3.5.3, a reduction in the available power splitting ratios is an effect arising from crosstalk in the GMZI. Imperfect imaging in the MMI couplers and uncompensated phase errors in the GMZI arms are thus the cause of the discrepancies seen in Fig. 7.19. As these non-ideal effects are reduced, a wider range of power splitting ratios are attainable.

For the GMZI to be useful as a variable-ratio power splitter in optical network applications, stability of the power splitting ratio over a range of wavelengths is necessary. Since the GMZI arms have equal lengths, the dominant dispersive element in the device is the MMI coupler. The measured spectral response of the MMI coupler imbalance in Fig. 7.8 is not flat. This indicates that imbalance and the associated phase deviation (arising from the coupler length deviation δL) both vary with wavelength. Then from (3.32), some wavelength sensitivity in the power splitting ratio of the GMZI is expected.

This wavelength sensitivity has been measured by first setting the GMZI to an arbitrary power splitting state. The spectral response of each output of the 4 × 4 GMZI is then measured and expressed as a ratio of the total GMZI output. The results are shown in Fig. 7.20, for a wavelength region of 1530 – 1570 nm.
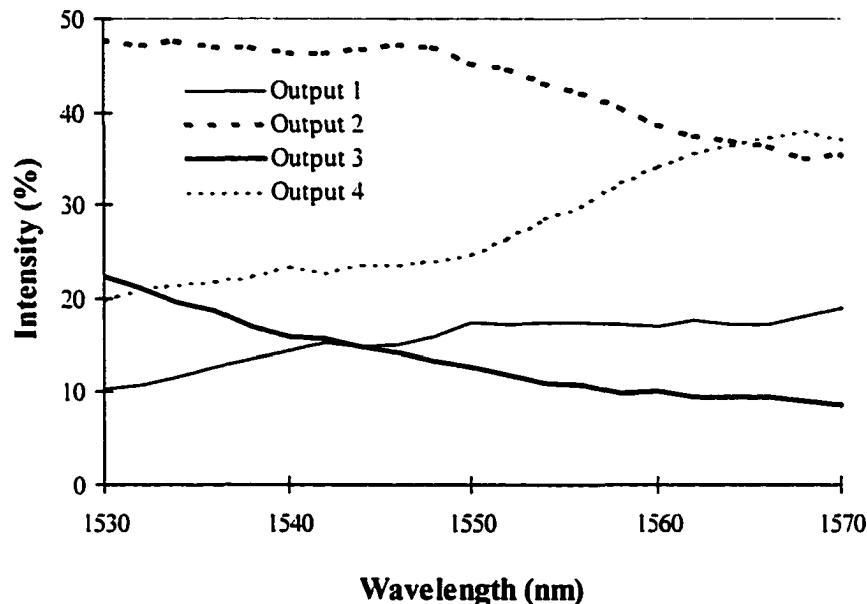
Figure 7.20 Measured wavelength sensitivity of an arbitrary power splitting state in the 4 × 4 GMZI.

The power splitting ratio varies gradually over the wavelength region shown, with a maximum drift of about 15 % in output port 4 over the 40 nm measurement region. This result will improve as the wavelength sensitivity of the imaging in the MMI couplers decreases.

Chapter 8

Summary and Future Directions

A concise summary of the important results obtained in this work is given in this chapter. Throughout the work a number of areas for modification and improvement have emerged, and can be applied to the devices developed in order to facilitate their timely incorporation into optical networks. Several of these possibilities are discussed, and provide the motivation for continuation of this work along a number of research and development directions.

8.1 Summary

The previous chapters have dealt with the basic principles and theory, design, fabrication, and characterization of new integrated optical devices based on the phenomenon of self-imaging exhibited in dielectric optical waveguides. This phenomenon has been exploited in a generalized Mach-Zehnder interferometer (GMZI) configuration employing multimode interference (MMI) couplers to develop devices that perform novel optical power splitting and switching functions useful for emerging optical networks.

Specifically, the integrated optical $N \times N$ GMZI that is at the heart of this work can be operated as a variable-ratio power splitter, a $1 \times N$ switch, and an $N \times N$ multi-port switch. In Chapter 1 it has been indicated that these functions are useful for re-routing signals through unaffected areas in the event of an interrupted link in an optical fiber network. Dynamic redistribution of optical power within a network can also be achieved using the GMZI, thereby allowing the available optical power to be used in an efficient manner.

In Chapter 2 the basic principles of dielectric optical waveguides were presented. The effective index method (EIM), which enables the two-dimensional analysis of propagation in optical channel waveguides to be reduced to a single dimension, has been discussed. The EIM facilitates implementation of the modal propagation analysis (MPA) technique, which is a simple yet accurate tool that can be used to study the properties of multimode waveguides.

Beating between the modes in a multimode waveguide leads to a self-imaging phenomenon, whereby an input optical field is manifested in the form of multiple copies. This phenomenon can be exploited, and simple analytic equations detailing the design of a multimode interference (MMI) coupler have been presented. One embodiment, the $N \times N$ MMI coupler, is capable of producing self-images of an optical field from any of N input waveguides at all N output waveguides. The phase transfer properties of this coupler were presented, and enable the design of an $N \times N$ generalized Mach-Zehnder interferometer (GMZI) structure which can spatially route and switch individual or multiple signals by means of active phase shifting.

The description of the specific contributions in this work begins with a formal framework for the analysis of the GMZI developed in Chapter 3. The operation of the $N \times N$ GMZI has been described mathematically using transfer matrices, leading to the analytic expression (3.13) describing its optical response to a set of applied phase shifts.

The converse mathematical problem, namely, finding the phase shifts necessary to achieve a specified optical response, has been solved by numerical techniques aided by the compact matrix expressions. The physical interpretation of the solution is that the optical intensity distribution at the N output ports of the $N \times N$ GMZI can be controlled through an application of an appropriate set of phase shifts. This new integrated optical device is called a variable-ratio power splitter. The methods derived here for controlling the GMZI to realize this device are a novel contribution of this work that is believed to be of considerable practical significance.

A tolerance analysis of a typical device has revealed that with phase shift and wavelength variations of $\pm 5.7\,^\circ$ and ± 40 nm, respectively, the power-splitting ratio is altered by only $1 - 2\,\%$. A more comprehensive tolerance analysis, however, would include inaccuracies stemming from the non-ideal behaviour of a fabricated GMZI. The non-ideal behaviour is the result of both technology limitations in the fabrication of integrated optical devices and fundamental theoretical limitations in the performance of MMI couplers.

Though the fundamental theoretical limitations of the MMI coupler expressions were not addressed in this work, the technologically related non-ideal effects have been incorporated into the transfer matrix framework to develop the expression (3.32), which describes the behavior of a non-ideal GMZI. Using this expression, imbalance and phase deviation among the self-images in the MMI couplers and phase errors arising in the GMZI arms have been introduced, to determine their effect on the performance of the GMZI. The result is that the accuracy of the power-splitting ratio degrades as the imaging performance in the MMI couplers worsens. For N up to 10, the MMI coupler imbalance and phase deviation must be kept to below ± 1.0 dB and $\pm 3.5°$, respectively, for maximum splitting ratio deviations of a few percent. Furthermore, phase errors in the arms of the GMZI often occur randomly during device fabrication and must also be kept low to maintain an acceptable performance level. As the non-idealities in the GMZI increase, the power splitting ratio not only degrades, but the predictability of the device response and the set of available ratios also diminishes.

In a special case of variable-ratio power splitting, the GMZI functions as a $1 \times N$ switch. In Chapter 4 the switch properties were investigated using the transfer matrix framework, and a matrix describing the optimum phase shifts required for switching was derived. The crosstalk in non-ideal GMZI switches has been calculated using the theory developed in Chapter 3, and it was found that to achieve low crosstalk levels, a low imbalance and phase deviation in the MMI couplers is of critical importance. As the MMI coupler operation improves, the crosstalk then becomes a sensitive function of the precision of the applied phase shifts. A crosstalk analysis for $1 \times N$ switches for N up to 10 was performed, with the result that the crosstalk is limited to the -20 dB to -30 dB range for MMI couplers with imbalance and phase deviation levels of ± 1.0 dB and $\pm 5°$, respectively. Even with perfect MMI coupler imaging, a $\pm 2\,\%$ deviation in the applied phase shifts increases the crosstalk in the GMZI from $-\infty$ dB to the -40 dB level.

When multiple inputs to the GMZI are used, a switch state of the $1 \times N$ switch corresponds to a permutation of the N inputs at the output ports. Since N distinct permutations are possible, the multi-port GMZI switch is of the $N \times N$ blocking type. The analytic result (3.13) has been used to derive an expression (4.11) for the set of permutations possible in the blocking switch. This expression not only describes the operation of the blocking switch, but has been shown to predict the wavelength assignments in a tunable MMI phasar wavelength multiplexer.

A detailed investigation of the imaging properties in the passive $N \times N$ GMZI has led to an interpretation of this device as two smaller, superimposed MMI couplers. The phase shifting requirements for the active $N \times N$ GMZI were then analyzed and shown to be linked to the superimposed coupler interpretation. This resulted in the development of a design rule that has been used to create $1 \times N$ switches with dramatically improved phase shifting requirements. By connecting adjacent $N \times N$ GMZI inputs to a 1×2 switch, the number of independent phase shifts required for control are reduced by a factor of two and have a minimal number of distinct levels. Examples of the design of these improved switches have shown that the phase shifts required for 1×4 and 1×6 switches can be applied using two- and three-bit binary control words, respectively.

A deeper understanding of the multi-port switching capabilities of the GMZI has led to the development of new designs for $N \times N$ non-blocking integrated optical switches. These are based upon configurations utilizing multiple, smaller GMZI units in a building block fashion. While the switches are rearrangably and not strictly non-blocking, they nevertheless have the

advantage that they can be realized without the use of waveguide bends or crossings, leading to very compact and potentially low-loss devices. Two embodiments of these switches have been conceived: a cascaded design and a balanced design. The cascaded design yields switches which are simple to design and operate, but have unbalanced losses. The balanced approach equalizes the losses, but at the expense of an increased complexity in the switch operation. The performance of a 4 × 4 balanced non-blocking switch has been simulated using MPA, and was shown to have an excess loss of 4 dB and a crosstalk level of −20 dB, even with accurate MMI coupler imaging.

The primary goal of the latter half of this work has been to experimentally demonstrate the validity of the theoretical concepts presented. A secondary goal has been to show that compact, high-performance integrated optical devices can be realized through a careful consideration of various device design and technology alternatives. To meet these objectives, integrated optical devices such as MMI couplers and GMZIs have been designed, fabricated, and characterized.

The design and simulation of these devices is the subject of Chapter 5. The devices in this work have been designed for fabrication in a silicon oxynitride (SiON) based waveguide technology. This represents a simple, flexible, and cost-effective process capable of realizing compact, low-loss, and polarization-insensitive devices with efficient thermo-optic active control.

To exploit the many potential advantages of this technology, the devices have been designed using a rib waveguide structure with a high refractive index contrast in both the lateral and transverse directions. Simulations have shown that this provides a nearly circular single-mode field and a TE-TM effective index birefringence of 2.5×10^{-5}. The structure is also designed to direct heat to the rib region for low power active operation, and to efficiently remove heat from this region for a fast active response. Simulation of the thermal behavior of the waveguide has predicted a π phase shift in 2 mm and 3 mm-long waveguides under an applied electrical power of 100 mW and 50 mW, respectively.

The high refractive index contrast has also resulted in the design of compact 2 × 2 and 4 × 4 MMI couplers with low polarization sensitivity. The simulated excess loss and imbalance levels for both couplers were below 1.0 dB and 0.5 dB, respectively. The dimensions of the MMI couplers are $8 \times 130~\mu m$ (2 × 2) and $16 \times 253~\mu m$ (4 × 4). Using these MMI couplers, 2 × 2 and 4 × 4 GMZIs have also been designed. Variable-ratio power splitting in the 4 × 4 GMZI has been simulated for the waveguide design chosen, with the results indicating excess losses below 2 dB, a maximum splitting ratio deviation of 1.5 %, and polarization-insensitive behaviour. Careful consideration of the design of the GMZI arms has resulted in the development of equal-length, equal-loss, thermally isolated waveguide arms with gradual bends and no transition losses.

The final device designs have been transferred to a photomask to be used in the device fabrication. The device fabrication process has been carried out at the MESA clean room facility at the University of Twente, in The Netherlands, and is described in detail in Chapter 6. The first stage of fabrication involved thin-film processing using plasma-enhanced chemical vapor deposition (PECVD) for the dielectric films and sputtering for the metal film. The process parameters for the PECVD have been optimized by an iterative technique, to ensure the desired refractive index contrast of $\Delta n = 0.101$ and film thickness control could be achieved reproducibly. The film properties have been determined using prism coupling, and resulted in a measured refractive index contrast of 0.100 ± 0.001 for both polarizations in a region of wavelengths around 1550 nm.

PECVD SiON films exhibit a high optical absorption loss in the 1500–1600 nm region; however, it has been shown qualitatively in this work that an annealing treatment for 1 hour at 1150 °C can significantly reduce these losses.

The second stage of fabrication involved photolithography and etching. The photolithography process has been optimized for the present devices, and the detailed procedural

steps performed in this process have been presented. The devices have been defined in the vertical direction by reactive ion etching (RIE), and the process parameters and etch depth characterization have been discussed. Optical microscope and scanning electron microscope (SEM) images of the fabricated devices have been given, enabling physical device characteristics to be determined both qualitatively and quantitatively. Single-mode waveguides with rib dimensions of 2.7×4.5 μm were designed, while values of $2.6 - 2.9$ μm $\times 4.4$ μm were measured. The lithographic process has, however, introduced rounding at the MMI coupler input and output facets, and the RIE has produced waveguide sidewalls with a roughness on the order of 100 nm.

The final stage of the fabrication process involved cleaving and mounting the devices, followed by wire-bonding and the application of the final heater contacts. Microscope images of devices in these stages have been given, and indicate two major drawbacks of the fabricated devices. These are the fragile nature of the rib structure, resulting in difficulty in device cleaving, and the absence of integrated thin-film heater contacts, resulting in additional optical attenuation and excess power dissipation.

The fabricated devices have been tested using the experimental characterization setup of the Lightwave Devices Group at the University of Twente. The measured performance results have been given in Chapter 7. An infrared camera image of the mode field emerging from a single-mode waveguide confirmed the nearly circular shape, and measured MFD values of $2.3 - 2.4$ μm can be compared with the simulated value of 2.5 μm. The propagation loss in the fabricated single-mode waveguides has been measured using the cutback technique. Loss values of 3.417 dB/cm (TE) and 3.408 dB/cm (TM) have been determined, and polarization insensitivity is confirmed over the wavelength region 1520 – 1580 nm.

The measured excess loss and imbalance in the fabricated 2×2 MMI couplers were 1.0 – 1.5 dB and 0.4 – 0.8 dB, respectively. The polarization sensitivity in these couplers is minimal and the results are similar to those predicted by MPA. The measured excess loss and imbalance in the 4×4 MMI couplers were 2.0 – 2.5 dB and 0.8 – 1.4 dB, respectively. These levels are higher than those predicted by MPA and are believed to be caused by imperfect lithographic definition of the couplers and bend losses in the access waveguides. The excess loss and imbalance of the 2×2 and 4×4 MMI couplers exhibited some wavelength dependence, with a variation of about 1 dB in both parameters over the wavelength region 1530 – 1570 nm for both couplers.

The performance of the 2×2 GMZI has been investigated, and a minimum crosstalk level of –26.7 dB and an extinction ratio of 25.8 dB were measured. Excess losses were 3 dB for both polarizations. An applied electrical power of 557 mW induced a π phase shift in a 2 mm long heater and switched the light from the cross to the bar port. The cause of this high switching power is believed to be excess power dissipation in the thermo-optic heater contacts which were applied manually using a crude procedure. The dynamic response of the 2×2 GMZI yielded rise and fall times of 2 – 4 ms.

The 4×4 GMZI exhibited excess losses of 4.8 dB (TE) and 5.7 dB (TM). An improved contact application procedure resulted in a π phase shift in an arm for an applied electrical power as low as 102 mW. Response times below 1 ms and as low as 460 μs (rise) and 650 μs (fall) have been measured. The measured optical response of the 4×4 GMZI mimics the theoretical result (3.32) describing a non-ideal GMZI. This indicates the presence of MMI coupler deviations and phase errors in the GMZI arms. The result is unpredictable device behavior, which was only partially compensated through the application of bias phase shifts. The GMZI performance is ultimately limited by the quality of the imaging in the MMI couplers.

Operation of a 1×4 GMZI switch has been demonstrated, with typical measured crosstalk values of –7 dB and a lowest value of –11 dB. Extinction ratios of 15 – 18 dB were

measured in the four ports. The high crosstalk is a direct consequence of imbalance and excess loss in the MMI couplers.

The multi-port switching capability of the 4 × 4 GMZI has been investigated, and the validity of (4.11) has been confirmed by demonstrating a cross-connect state. This result represents an experimental confirmation of the principle upon which the new switch designs described in Chapter 4 are based.

The variable-ratio power splitting function has also been demonstrated in the 4 × 4 GMZI. The theory developed in Chapter 3 has been used along with bias voltage values to achieve desired power splitting states. The flexibility of the power splitting function has been demonstrated by the attainment of a number of different power splitting ratios. Drawbacks of the fabricated device were a small uncertainty in the applied phase shifts, and the inaccessibility of some power splitting ratios. The cause of this is believed to be non-ideal imaging in the MMI couplers and additional phase errors and imbalance introduced in the GMZI arms as a result of the heater contacts. Finally, the wavelength dependence for an arbitrarily chosen power splitting ratio was measured, and a maximum variation of 15 % was observed over a 40 nm wavelength region.

8.2 Future Directions

Throughout this work a number of improvements, modifications, and avenues for continued research have arisen. Perhaps the next logical step in continuing the work presented here is the optimization of the device performance. An annealing study of the PECVD films would provide the required process parameters to achieve the desired refractive index contrast and layer thickness control, enabling in the realization of waveguides with a propagation loss of below 0.2 dB/cm. Single-mode reference waveguides closely matching the bends and propagation length in the devices can also be implemented to provide more accurate excess loss values. Optimization of the RIE process to produce smooth waveguide sidewalls would ensure that scattering losses, bend losses, and random phase errors remain low.

It has also been shown recently [64] that adiabatic tapering of the waveguides in a single direction can significantly reduce the coupling losses of SiON waveguides to standard single-mode optical fiber, thereby allowing the integrated devices to be easily packaged.

It has been mentioned that the use of integrated thin-film thermo-optic heaters could significantly reduce the power consumption and response time of the active devices. Through a suitable design of the heater contacts, the attenuation and phase errors introduced by the metal in the GMZI arms can be completely eliminated. The chromium metal can also be replaced by aluminum or a titanium-gold alloy to reduce the heater resistance and achieve low-voltage, high-current heating [67], [107].

The MMI couplers fabricated in this work have relatively high excess loss and imbalance, even in simulations. This is partially due to the narrow coupler widths used, which can only support a small number of modes, thereby sacrificing the imaging resolution. The small couplers were used to maintain polarization insensitivity, however, accurate MMI coupler imaging is more critical to the performance of the GMZI. Wider couplers with more widely spaced access waveguides can be used to improve the imaging. If the access waveguide spacing is greater than about 10 μm the thermal isolation is sufficient and the bent arms of the GMZI are not required. Straight waveguide arms significantly simplify the GMZI design and can reduce polarization sensitivity and the overall device size, which is dominated by the phase shifting region.

To improve the fabrication tolerances of the MMI coupler, wider single-mode access waveguides can be utilized [9]. This can be achieved by tapering the waveguides as they enter and exit the MMI coupler region.

MMI couplers and single-mode waveguides can possibly both be designed to have polarization insensitive behaviour, by utilizing a cross-sectional waveguide geometry for the multimode waveguide region that is different from that of the single-mode access waveguides.

Employing many of these suggested improvements may improve the GMZI characteristics to a level where the variable-ratio power splitter, the improved $1 \times N$ switches, and the $N \times N$ non-blocking switches presented in Chapters 3 and 4 become practical. This work can then be extended to the development of such devices, where values of $N > 4$ are utilized. Of course, the theory presented in this work is independent from the material system or active phase shifting method employed. The devices can therefore be implemented in a different waveguide technology, which may possess certain advantages over the silica-based system used for this work.

A concern for large-port GMZI devices, however, is the active phase shifting requirements. Though the phase shifts applied to the GMZI have been controlled manually in this work, the task becomes impractical for devices with a larger number of phase shifters. For such devices the phase shifts should be computer controlled, which offers the advantage that iterative algorithms for phase shift optimization can be easily implemented.

The ultimate goal for integrated GMZI-based devices is to realize a self-contained, packaged device. This would include user input of the desired state of the device in electrical form, which would be used with on-board processor electronics to determine, apply, and maintain the necessary phase shifts to the GMZI arms. The algorithms developed in this work can be coded into the processor electronics to determine appropriate phase shifts. For the devices based on idealized switching, the phase shifting can be reduced to the application of one of a set of quantized voltage levels. Voltages of only a few volts can induce a π phase shift in a waveguide if low-resistance thermo-optic heaters are used [67], suggesting that standard integrated electronic circuitry is suitable for the application of the phase shifts.

Maintenance of the desired optical output state of a device in the presence of environmental fluctuations is a key requirement for the practical feasibility of these devices, and a more general requirement for all interferometric devices. This necessitates monitoring the optical output of a device and providing an appropriate feedback mechanism to determine and apply corrections to the phase shifts. This must be achieved in a minimally invasive manner, so as to protect the integrity of the optical signals. One method of monitoring utilizes optical tap ports to divert a small portion of the light from each output waveguide to a detector, which then drives the feedback electronics. This principle has recently been demonstrated for an integrated thermo-optic attenuator [117].

The above description of areas for continuation of this work is not exhaustive, but is designed to convey a flavour of the many doors of investigation which have been opened. Finally, in addition to the principal application of optical networking, it is the hope that the devices described in this body of work will find a wider and more far-reaching applicability, thereby furthering the dream of integrated optics and fueling the evolution of communication.

8.3 Conclusion

In this work, a theory of the GMZI is developed and has been used to predict the ideal and non-ideal response of this device. The theory has led to the development of new integrated optical devices: the variable-ratio power splitter, improved $1 \times N$ switches, and $N \times N$ non-blocking switches. To determine the feasibility of the practical operation of these devices, GMZI devices have been fabricated in a high refractive index contrast silicon oxynitride waveguide system. The devices have been designed to provide the functionality required to verify the theoretical predictions while taking full advantage of the waveguide technology to demonstrate the possibility of achieving compact, low-loss GMZI-based devices that exhibit high-performance passive and active operation. The fabricated devices have been characterized, and the theory of the GMZI has been verified through demonstrations of variable-ratio power splitting and $1 \times N$ and multi-port switching. Measurement of the device response has also confirmed aspects of the theory of the non-ideal GMZI, which has provided insights into the sources of the non-ideal behaviour. The optical and thermal waveguide characteristics confirm the validity of the simulations used for their design and demonstrate the possibility of achieving compact, low-loss, polarization insensitive integrated devices with low power consumption levels and a fast response. Deviations in the performance of the fabricated devices from the theoretical predictions indicate that imperfect imaging (imbalance and phase deviation) in the MMI couplers is the principal mechanism responsible for poor GMZI performance. For the realization of commercially viable devices based on the GMZI, it is therefore essential that mechanisms for the improvement of the performance of the MMI couplers are investigated further.

It is the belief then that with an improved robustness of the self-imaging phenomenon in multimode interference couplers, the GMZI-based devices described herein will be invaluable components in the realization of efficient optical networks.

References

[1]   C. K. Hao and G. A. Hockham, "Dielectric-fiber surface waveguides for optical frequencies," *Proc. IEEE*, vol. 113, no. 7, pp. 1151-1158, 1966.

[2]   R. Dupuis, N. Holonyak Jr., M. I. Nathan and R. H. Rediker, "Papers on early history of the semiconductor laser diode," *IEEE J. Quantum Electron.*, vol. 23, no. 6, pp. 651-695, 1987.

[3]   S. Kawanishi, H. Takara, K. Uchiyama, I. Shake, and K. Mori, "3 Tbit/s (160Gbit/s × 19 channel) optical TDM and WDM transmission experiment," *Electron. Lett.*, vol. 35, no. 10, pp. 826-827, May 1999.

[4]   T. Matsuda, M. Murakami, and T. Imai, "340 Gbit/s (34 × 10 Gbit/s) WDM transmission over 8,514 km using broadband gain equalisation technique for transoceanic systems," *Electron. Lett.*, vol. 35, no. 13, pp. 1090-1091, June 1999.

[5]   M. Kawachi, "Silica waveguides on silicon and their application to integrated-optic components," *Opt. Quantum Electron.*, vol. 22, pp. 391-416, 1990.

[6]   S. E. Miller, "Integrated optics: An introduction," *Bell Syst. Tech. J.*, vol. 48, no. 7, pp. 2059-2069, 1969.

[7]   E. Pennings, G. D. Khoe, M. K. Smit, and T. Staring, "Integrated-optic versus microoptic devices for fiber-optic telecommunication systems: a comparison," *IEEE J. Select. Topics Quantum Electron.*, vol. 2, pp. 151-164, June 1996, invited paper.

[8]   B. E. A. Saleh and M. C. Teich, *Fundamentals of photonics*. New York: Wiley, 1991.

[9]   L. H. Spiekman, *Compact integrated optical components for telecommunication networks*. PhD thesis, Delft University of Technology, Delft, the Netherlands, 1996.

[10]   M. Okuno, N. Takato, T. Kitoh, and A. Sugita, "Silica-based thermo-optic switches," *NTT Review*, vol. 7, no. 5, pp. 57-63, Sept. 1995.

[11]   L. B. Soldano, F. B. Veerman, M. K. Smit, B. H. Verbeek, A. H. Dubost, and E. C. M. Pennings, "Planar Monomode Optical Couplers Based on Multimode Interference Effects," *J. Lightwave Technol.*, vol. 10, no. 12, pp. 1843-1849, Dec. 1992.

[12]   M.Bachmann, P.A. Besse, and H. Melchior, "General self-imagaing properties in $N \times N$ multimode interference couplers including phase relations," *Appl. Opt.*, vol. 33, no. 18, pp. 3905-3911, June 1994.

[13]   R. Ulrich, U. K. Patent Application No. 1 525 492, 1978.

[14]   J. E. Midwinter, Ed., *Photonics in switching*. San Diego: Academic Press, 1993.

[15]   D. J. Griffiths, *Introduction to Electrodynamics*, 2nd Ed., Prentice-Hall, NJ, 1989.

[16]   J. T. Boyd, "Photonic Integrated Circuits," in *Photonic Devices and Systems*, R. G. Hunsperger, Ed., Marcel Dekker, NY, 1994.

[17] N. S. Kapany and J. J. Burke, *Optical Waveguides*, Academic Press, NY, 1972.

[18] K. S. Chiang, "Dual effective-index method for the analysis of rectangular dielectric waveguides," *Appl. Opt.*, vol. 25, no. 13, pp. 2169-2174, 1986.

[19] A. Kumar, D. F. Clark, and B. Culshaw, "Explanation of errors inherent in the effective-index method for analyzing rectangular-core waveguides," *Opt. Lett.*, vol. 13, no. 12, pp. 1129-1131, 1988.

[20] K. S. Chiang, "Performance of the effective-index method for the analysis of dielectric waveguides," *Opt. Lett.*, vol. 16, no. 10, pp. 714-716, 1991.

[21] P. C. Kendall, M. J. Adams, S. Ritchie, and M. J. Robertson, "Theory for calculating approximate values for the propagation constant of an optical rib waveguide by weighting the refractive indices," *IEE Proceedings*, vol. 134, Pt. A, no. 8, pp. 699-702, 1987.

[22] L. B. Soldano and E. C. M. Pennings, "Optical multi-mode interference devices based on self- imaging: principles and applications," *J. Lightwave Technol.*, vol. 13, no. 4, pp. 615-627, Apr. 1995, invited paper.

[23] P. A. Besse, E. Gini, M. Bachmann, and H. Melchior, "New 2×2 and 1×3 multimode interference couplers with free selection of power splitting ratios," *J. Lightwave Technol.*, vol. 14, no. 10, pp. 2286-2293, Oct. 1996.

[24] D. S. Levy, R. Scarmozzino, and R. M. Osgood, Jr., "Length reduction of tapered $N \times N$ MMI devices," *IEEE Photon. Tech. Lett.*, vol. 10, no. 6, pp. 830-832, June 1998.

[25] M. Bachmann, P. A. Besse, and H. Melchior, "Overlapping-image multimode interference couplers with a reduced number of self-images for uniform and nonuniform power splitting," *Appl. Opt.*, vol. 34, no. 30, pp. 6898-6910, Oct. 1995.

[26] P. A. Besse, M. Bachmann, and H. Melchior, "Phase relations in multi-mode interference couplers and their application to generalized integrated Mach-Zehnder optical switches," in *Proc. ECIO '93*, pp. 2.22-2.23, Apr. 1993.

[27] M. R. Paiam and R. I. MacDonald, "Design of phased-array wavelength division multiplexers using multimode interference couplers," *Appl. Opt.*, vol. 36, no. 21, pp. 5097-5108, July 1997.

[28] E. C. M. Pennings, R. J. Deri, A. Scherer, R. Bhat, T. R. Hayes, N. C. Andreadakis, M. K. Smit, L. B. Soldano, and R. J. Hawkins, "Ultracompact, low-loss directional couplers on InP based on self-imaging by multimode interference," *Appl. Phys. Lett.*, vol. 59, pp. 1926-1928, 1991.

[29] M. A. Fardad and M. Fallahi, "Sol-gel multimode interference power splitters," *IEEE Photon. Tech. Lett.*, vol. 11, no. 6, pp. 697-699, June 1999.

[30] L. H. Spiekman, Y. S. Oei, E. G. Metaal, F. H. Groen, I. Moerman, and M. K. Smit, "Extremely Small Multimode Interference Couplers and Ultrashort Bends on InP by Deep Etching," *IEEE Photon. Tech. Lett.*, vol. 6, no. 8, pp. 1008-1010, Aug. 1994.

[31] P. A. Besse, M. Bachmann, H. Melchior, L. B. Soldano, and M. K. Smit, "Optical Bandwidth and Fabrication Tolerances of Multimode Interference Couplers," *J. Lightwave Technol.*, vol. 12, no. 6, pp. 1004-1009, June 1994.

[32] M. Bachmann, M. K. Smit, P. A. Besse, E. Gini, H. Melchior, and L. B. Soldano, "Polarization-insensitive low-voltage optical waveguide switch using InGaAsP/InP four-port Mach-Zehnder interferometer," in *OFC/IOOC '93 Technical Digest*, pp. 32-33, paper ThH3.

[33] M. Bachmann, Ch. Nadler, P. A. Besse, and H. Melchior, "Compact Polarization-Insensitive Multi-Leg 1 x 4 Mach-Zehnder Switch in InGaAsP/InP," in *Proc. ECOC*, Firenze, Italy, pp. 519-522, 1994.

[34] R. M. Jenkins, J. M. Heaton, D. R. Wight, J. T. Parker, J. C. H. Birbeck, G. W. Smith, and K. P. Hilton, "Novel $1 \times N$ and $N \times N$ integrated optical switches using self-imaging multimode GaAs/AlGaAs waveguides," *Appl. Phys. Lett.*, vol. 64, no. 6, pp. 684-686, Feb. 1994.

[35] N. S. Lagali, M. R. Paiam, and R. I. MacDonald, "Theory of variable-ratio power splitters using multimode interference couplers," *IEEE Photon. Technol. Lett.*, vol. 11., no. 6, pp. 665-667, June 1999.

[36] N. S. Lagali, M. R. Paiam, R. I. MacDonald, K. Wörhoff, and A. Driessen, "Analysis of generalized Mach-Zehnder interferometers for variable-ratio power splitting and optimized switching," *J. Lightwave. Technol.*, to appear in Dec. 1999 issue.

[37] N. S. Lagali, R. I. MacDonald, and M. R. Paiam, "Optical variable-ratio power splitter," U. S. Patent Application No. 09/227939, Priority Date: Jan. 30, 1998, filed Jan. 11, 1999.

[38] C. G. Montgomery, R. H. Dicke, and E. M. Purcell, Eds., *Principles of Microwave Circuits*, 1st Ed., McGraw-Hill, NY, 1948.

[39] R. Levy, in *Advances in Microwaves*, L. Young, Ed., Academic Press, NY, 1966.

[40] B. H. Verbeek, C. H. Henry, N. A. Olsson, K. J. Orlowsky, R. F. Kazarinov, and B. H. Johnson, "Integrated Four-Channel Mach-Zehnder Multi/Demultiplexer Fabricated with Phosphorous Doped SiO_2 Waveguides on Si," *J. Lightwave Technol.*, vol. 6, no. 6, pp. 1011-1015, June 1988.

[41] K. Jinguji and M. Kawachi, "Synthesis of Coherent Two-Port Lattice-Form Optical Delay-Line Circuit," *J. Lightwave Technol.*, vol. 13, no. 1, pp. 73-82, Jan. 1995.

[42] G. Arfken, *Mathematical Methods for Physicists*, 3rd Ed., Academic Press, San Diego, 1985.

[43] H. Nishihara, M. Haruna, and T. Suhara, *Optical Integrated Circuits*, McGraw-Hill, NY, 1989.

[44] E. C. M. Pennings, R. van Roijen, M. J. N. van Stralen, P. J. de Waard, R. G. M. P. Koumans and B. H. Verbeek, "Reflection Properties of Multimode Interference Devices," *IEEE Photon. Tech. Lett.*, vol. 6, no. 6, pp. 715-718, June 1994.

[45] Q. Lai, M. Bachmann, W. Hunziker, P. A. Besse and H. Melchior, "Arbitrary ratio power splitters using angled silica on silicon multimode interference couplers," *Electron. Lett.*, vol. 32, no. 17, pp. 1576-1577, Aug. 1996.

[46] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," *Numerical Analysis*, ed. G. A. Watson, Lecture Notes in Mathematics 630, Springer-Verlag, pp. 105-116, 1977.

[47] J. E. Dennis, Jr., and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[48] P. A. Besse, M. Bachmann, Ch. Nadler, H. Melchior, "The integrated prism interpretation of multileg Mach-Zehnder interferometers based on multimode interference couplers," *Opt. Quant. Electron.*, vol. 27, pp. 909-920, 1995.

[49] T. Goh, S. Suzuki, and A. Sugita, "Estimation of waveguide phase error in silica-based waveguides," *J. Lightwave Technol.*, vol. 15, no. 11, pp. 2107-2113, Nov. 1997.

[50] M. R. Paiam and R. I. MacDonald, "A 12-channel phased-array wavelength multiplexer with multimode interference couplers," *IEEE Photon. Tech. Lett.*, vol. 10, no. 2, pp. 241-243, Feb. 1998.

[51] C. K. Madsen, "A multiport frequency band selector with inherently low loss, flat passbands, and low crosstalk," *IEEE Photon. Tech. Lett.*, vol. 10, no. 12, pp. 1766-1768, Dec. 1998.

[52] T. Augustsson, "Theoretical investigation of a wavelength selective switch architecture based on a Bragg-grating-assisted MMIMI configuration," *IEEE Photon. Tech. Lett.*, vol. 11, no. 7, pp.839-841, July 1999.

[53] R. Ulrich and T. Kamiya, "Resolution of self-images in planar optical waveguides," *J. Opt. Soc. Am.*, vol. 68, no. 5, pp. 583-592, May 1978.

[54] N. S. Lagali, M. R. Paiam, R. I. MacDonald, K. Wörhoff, and A. Driessen, "1 x N and non-blocking N x N optical space switches based on multimode interference," submitted to *Journal of Lightwave Technology*, received Oct. 1998, revised Sept. 1999.

[55] N. S.Lagali, R. I.MacDonald, and M. R. Paiam, "An integrated 1 x N optical switch," U. S. Patent Application No. 09/227856, Priority Date: Jan. 30, 1998, filed Jan. 11, 1999.

[56] N. S. Lagali, R. I. MacDonald, and M. R. Paiam, "An N x N non-blocking optical switch," U. S. Patent Application No. 09/227857, Priority Date: Jan. 30, 1998, filed Jan. 11, 1999.

[57] B. H. Verbeek, A. A. M. Staring, E. J. Jansen, R. van Roijen, J. J. M. Binsma, T. van Dongen, M. R. Amersfoort, C. van Dam, and M. K. Smit, "Large bandwidth polarization independent and compact 8 channel PHASAR demultiplexer/filter," in *Optical Fiber Communication Conference*, Vol. 4 of 1994 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1994), pp. PD13-1-PD13-4.

[58] Jean-Pierre Weber, B. Stoltz, and O. Oberg, "A new type of tunable demultiplexer using a multi-leg Mach-Zehnder interferometer," *ECIO '97 Proceedings*, Stockholm, pp. 272-275, 1997.

[59] R. I. MacDonald, "Large modular, expandable optical switching matrices," *IEEE Photon. Technol. Lett.*, vol. 11, no. 9, June 1999.

[60] N. S. Lagali, R. I. MacDonald, M. R. Paiam, K. Wörhoff, and A. Driessen, "Silica-based rib waveguides for integrated optics using multimode interference," in *Applications of photonic technology 3: closing the gap between theory, Development, and application*, G. A. Lampropoulos, and R. A. Lessard, Eds., *Proc. SPIE*, vol. 3491, pp. 380-385, 1998.

[61] T. Uitterdijk, D. H. P. Maat, F. H. Groen, H. van Brug, C. G. M. Vreeburg, J. W. Pedersen, and I. Moerman, "Dilated, polarization insensitive InP-based space switch," in *Proc. ECIO '97*, Stockholm, pp. 551-554, Apr. 1997.

[62] T. A. Tumolillo, Jr., M. Donckers, and W. H. G. Horsthuis, "Solid state optical space switches for network cross-connect and protection applications," *IEEE Comm. Magazine*, pp. 124-130, Feb. 1997.

[63] R. M. de Ridder, K. Wörhoff, A. Driessen, P. V. Lambeck, and H. Albers, "Silicon oxynitride planar waveguiding structures for application in optical communication," *IEEE J. Select. Topics Quantum Electron.*, vol. 4, pp. 930-937, 1998.

[64] K. Wörhoff, P. V. Lambeck, and A. Driessen, "Design, tolerance analysis, and fabrication of silicon oxynitride based planar optical waveguides for communication devices," *J. Lightwave Technol.*, vol. 17, no. 8, pp. 1401-1407, Aug. 1999.

[65] C. T. Lee, M. L. Wu, L. G. Sheu, P. L. Fan, and J. M. Hsu, "Design and analysis of completely adiabatic tapered waveguides by conformal mapping," *J. Lightwave Technol.*, vol. 15, no. 2, pp. 403-410, Feb. 1997.

[66] Z. Weissman and I. Hendel, "Analysis of periodically segmented waveguide mode expanders," *J. Lightwave Technol.*, vol. 13, no. 10, pp. 2053-2058, Oct. 1995.

[67] Q. Lai, W. Hunziker, and H. Melchior, "Low-power compact 2 x 2 thermooptic silica-on-silicon waveguide switch with fast response," *IEEE Photon. Technol. Lett.*, vol. 10, no. 5, pp. 681-683, May 1998.

[68] M. Hoffmann, P. Kopka, and E. Voges, "Low-loss fiber-matched low-temperature PECVD waveguides with small-core dimensions for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 9, no. 9, pp. 1238-1240, Sept. 1997.

[69] K. Jinguji, "Optical devices for processing optical signals as light without transforming them into electrical signals," *NTT Review*, vol. 7, no. 1, pp. 80-86, Jan. 1995.

[70] B. J. Offrein, G. L. Bona, R. Germann, F. Krommendijk, I. Massarek, and H. W. M. Salemink, "High contrast and low loss SiON optical waveguides by PECVD," in *Proc. IEEE/LEOS Symp. Benelux Chapter*, 1996, pp. 290-293.

[71] D. Peters, K. Fischer, and J. Müller, "Integrated optics based on silicon oxynitride thin films deposited on silicon substrates for sensor applications," *Sens. Act. A*, vols. 25-27, pp. 425-431, 1991.

[72] M. Hoffmann and E. Voges, "Low temperature, nitrogen doped waveguides on silicon with small core dimensions fabricated by PECVD/RIE," in *Proc. 7th Eur. Conf. Integrated Optics (ECIO '95)*, pp.299-302.

[73] H. Bezzaoui and E. Voges, "Integrated optics combined with micromechanics on silicon," *Sens. Act. A.*, vol. 29, pp. 219-223, 1991.

[74] H. Albers, L. T. H. Hilderink, E. Szilágyi, F. Paszti, P. V. Lambeck, and Th. J. A. Popma, "Birefringence compensation in double-core optical waveguides," in *Proc IEEE/LEOS '95*, San Francisco, CA, pp. 88-89.

[75] K. Wörhoff, A. Driessen, P. V. Lambeck, L. T. H. Hilderink, P. W. C. Linders, and Th. J. A. Popma, "PECVD silicon oxynitride optimized for application in integrated optics," *Sens. Act. A*, vol. 74, pp. 9-12, 1999.

[76] F. Horst, H. W. M. Salemink, R. Germann, B. J. Offrein, and G. L. Bona, "High quality ring resonators in high refractive index contrast SiON waveguides," in *Proc. IEEE/LEOS Symp. Benelux Chapter*, 1998, pp.33-36.

[77] M. R. Paiam, *Applications of MMI couplers in WDM.* PhD Thesis, University of Alberta, Edmonton, 1997.

[78] M. Taylor, *FWave: two-dimensional mode solver*, public domain, 1995.

[79] B. J. Offrein, G. L. Bona, F. Horst, H. W. M. Salemink, R. Beyeler, and R. Germann, "Wavelength tunable optical add-after-drop filter with flat passband for WDM networks," *IEEE Photon. Technol. Lett.*, vol. 11, no. 2, pp. 239-241, Feb. 1999.

[80] A. Sugita, K. Jinguji, N. Takato, K. Katoh, and M. Kawachi, "Bridge-suspended silica-waveguide thermooptic phase shifter and its application to Mach-Zehnder type optical switch," *Trans. IEICE*, vol. E73, no. 1, pp. 105-109, Jan. 1990.

[81] B. A. Møller, L. Jensen, C. Laurent-Lund, and C. Thirstrup, "Silica-waveguide thermooptic phase shifter with low power consumption and low lateral heat diffusion," *IEEE Photon. Technol. Lett.*, vol. 5, no. 2, pp. 1415-1418, Dec. 1993.

[82] Y. Inoue, K. Katoh, and M. Kawachi, "Polarization sensitivity of a silica waveguide thermooptic phase shifter for planar lightwave circuits," *IEEE Photon. Technol. Lett.*, vol. 4, no. 1., pp. 36-38, Jan. 1992.

[83] D. R. Lide, Ed., *CRC Handbook of Chemistry and Physics*, 79th Ed., New York: CRC Press, 1998.

[84] SPDE Inc., *PDEase2: Finite element analysis for partial differential equations.* Bass Lake, CA, Copyright 1995.

[85]   D. Marcuse, *Light Transmission Optics*. 2nd Ed., New York: Van Nostrand Reinhold Co., 1982.

[86]   E. A. J. Marcatili, "Bends in optical dielectric guides," *Bell Syst. Tech. J.*, vol. 48, no. 7, pp. 2103-2132, Sept. 1969.

[87]   F. Ladouceur and P. Labeye, "A new general approach to optical waveguide path design," *J. Lightwave Technol.*, vol. 13, no. 3, pp. 481-492, March 1995.

[88]   V. Subramaniam, G. N. De Brabander, D. H. Naghski, and J. T. Boyd, "Measurement of mode field profiles and bending and transition losses in curved optical channel waveguides," *J. Lightwave Technol.*, vol. 15, no. 6, pp. 990-997, June 1997.

[89]   D. Marcuse, "Length optimization of an S-shaped transition between offset optical waveguides," *Appl. Opt.*, vol. 17, pp. 763-768, 1978.

[90]   W. J. Minford, S. K. Korotky, and R. C. Alferness, "Low-loss $Ti:LiNbO_3$ waveguide bends at $\lambda = 1.3$ μm," *IEEE J. Quantum Electron.*, vol. QE-18, pp. 1802-1806, 1982.

[91]   M. T. Kauffman, C. J. Radens, and J. T. Boyd, "Length minimization in integrated optical circuits incorporating directional couplers and curved sections," *J. Lightwave Technol.*, vol. 8, no. 10, pp. 1504-1508, Oct. 1990.

[92]   D. W. Boertjes and J. N. McMullin, "Novel Wide-Angle 3D-BPM Implemented on a Path-based Grid", Microwave and Optical Technology Letters, Vol. 20, No. 5, pp. 287-290, March 5, 1999.

[93]   L. H. Spiekman, Y. S. Oei, E. G. Metaal, F. H. Groen, I. Moerman, M. K. Smit, and B. H. Verbeek, "Extremely small, fabrication-tolerant InP-based power-splitting and combining structures by deep etching," in *Proc. European Conf. Opt. Commun. (ECOC)*, Firenze, Italy, Sept. 1994, paper We.C.2.3.

[94]   E. C. M. Pennings, R. J. Deri, R. Bhat, T. R. Hayes, and N. C. Andreadakis, "Ultra-compact all-passive optical 90° hybrid on InP using self-imaging," *IEEE Photon. Technol. Lett.*, vol. 6, no. 5, pp. 701-703, 1993.

[95]   The Math Works, Inc., *MATLAB 4.0*, numerical computation software.

[96]   Tanner Research Inc., *L-Edit*, Version 4.0.

[97]   C. E. Morosanu, *Thin films science and technology*. Vol. 7, Amsterdam: Elsevier, 1990.

[98]   R. M. A. Azzam and N. M. Bashara, *Ellipsometry and polarized light*. Amsterdam: North-Holland, 1977.

[99]   K. Wörhoff, *Optimized LPCVD SiO_xN_y waveguides covered with calixarene for non-critically phase-matched second harmonic generation*. Ph.D. Thesis, University of Twente, Enschede, The Netherlands, 1996.

[100] P. K. Tien and R. Ulrich, "Theory of prism-film coupler and thin-film light guides," *J. Opt. Soc. Am.*, vol. 60, no. 10, pp. 1325-1337, Oct. 1970.

[101] R. Ulrich, "Theory of the prism-film coupler by plane-wave analysis," *J. Opt. Soc. Am.*, vol. 60, no. 10, pp. 1337-1350, Oct. 1970.

[102] R. Ulrich and R. Torge, "Measurement of thin film parameters with a prism coupler," *Appl. Opt.*, vol. 12, no. 12, pp. 2901-2908, Dec. 1973

[103] O. F. J. Noordman, *INDEXFIT*, calculation software.

[104] B. Sikken, *Tuneable integrated optic wavelength filter using Mach-Zehnder interferometers*. M.Sc. Thesis, University of Twente, Enschede, The Netherlands, 1998.

[105] M. V. Bazylenko, M. Gross, P. M. Allen, and P. L. Chu, "Fabrication of low-temperature PECVD channel waveguides with significantly improved loss in the 1.50–1.55 μm wavelength range," *IEEE Photon. Technol. Lett.*, vol. 7, no. 7, pp.774-776, July 1995.

[106] J. L. Vossen and W. L. Kern, Eds., *Thin film processes*. Orlando, FL: Academic Press, 1978.

[107] F. Krommendijk, *An integrated thermo-optical switch based on lateral index contrast modulation*. M.Sc. Thesis, University of Twente, Enschede, The Netherlands, 1997.

[108] R. Legtenberg, *Electrostatic actuators fabricated by surface micromachining techniques*. Ph.D. Thesis, University of Twenete, Enschede, The Netherlands, 1996.

[109] F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*. 2nd Ed., New Jersey: Prentice-Hall, 1993.

[110] N. S. Lagali, M. R. Paiam, R. I. MacDonald, K. Wörhoff, and A. Driessen, "4 x 4 variable-ratio integrated optical power splitter," in *Proc. Eur. Conf. on Int. Optics (ECIO '99)*, Torino, Italy, Postdeadline papers, pp. PD9-PD12, Apr. 1999.

[111] J. Hecht, *Understanding Fiber Optics*. 2nd Ed., New Jersey: Prentice-Hall, 1993.

[112] Corning Optical Fiber Product Information Sheet PI1036, "Corning SMF-28 CPC6 single-mode optical fiber," http://pro.corningfiber.com/library/pi1036.htm, Feb. 1998.

[113] TECOS product information sheet, http://www.tecos.de/products/lens.html

[114] Opto Speed, Monomode taper lensed fiber TAP 1000 product information sheet, www.optospeed.ch

[115] L. B. Soldano, *Multimode interference couplers: design and applications*. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1994.

[116] T. Goh, A. Himeno, M. Okuno, H. Takahashi, and K. Hattori, "High extinction ratio and low-loss silica-based 8 x 8 strictly nonblocking thermooptic matrix switch," *J. Lightwave Technol.*, vol. 17, no. 7, pp. 1192-1199, July 1999.

[117] S. S. Lee, Y. S. Jin, Y. S. Son, and T. K. Yoo, "Polymeric tunable optical attenuator with an optical monitoring tap for WDM transmission network," *IEEE Photon. Technol. Lett.*, vol. 11, no. 5, pp. 590-592, May 1999.