Gebruik van hoog indexcontrast in het III-V-golfgeleiderplatform
op silicium-op-isolator voor toepassingen van optischesignaalverwerking

Leveraging High Index Contrast in the III-V Waveguide Platform
on Silicon-on-Insulator for Optical Signal Processing Applications

Martijn Tassaert

Promotoren: prof. dr. ir. D. Van Thourhout, prof. dr. ir. G. Roelkens
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Fotonica

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. D. De Zutter
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2013 - 2014

UNIVERSITEIT
GENT

ISBN 978-90-8578-686-3
NUR 959
Wettelijk depot: D/2014/10.500/32

Promotor:

Prof. dr. ir. G. Roelkens	Universiteit Gent, INTEC
Prof. dr. ir. D. Van Thourhout	Universiteit Gent, INTEC

Examencommissie:

Prof. dr. ir. R. Van de Walle (voorzitter)	Universiteit Gent - iMinds, Multimedia Lab
Prof. dr. ir. G. Roelkens (promotor)	Universiteit Gent, INTEC
Prof. dr. ir. D. Van Thourhout (promotor)	Universiteit Gent, INTEC
Dr. ir. O. Raz	Technische Universiteit Eindhoven, ECO
Dr. ir. S. Lardenois	Caliopa
Prof. dr. ir. K. Neyts	Universiteit Gent, ELIS
Prof. dr. ir. G. Morthier	Universiteit Gent, INTEC
Prof. dr. ir. R. Baets	Universiteit Gent, INTEC

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur

Vakgroep Informatietechnologie (INTEC)
Sint-Pietersnieuwstraat 41
9000 Gent
België

Tel.: +32-9-264.33.19
Fax: +32-9-264.35.93
http://www.intec.ugent.be

Dit werk kwam tot stand in het kader van een aspirantenbeurs van het Bijzonder Onderzoeksfonds (BOF)

UNIVERSITEIT GENT

Preface - Voorwoord

Martijn Tassaert
Gent, 1 januari 2014

Contents

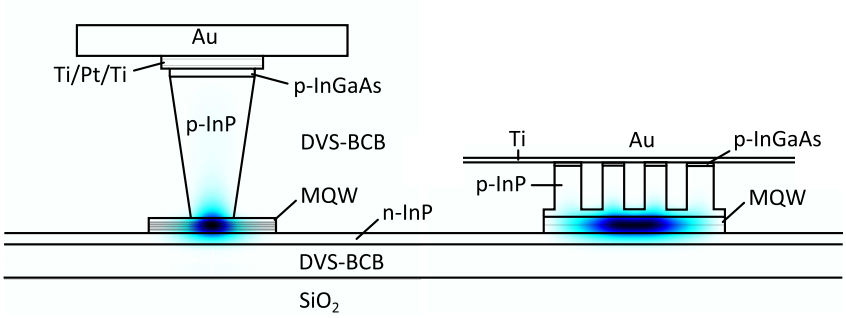
Nederlandstalige tekst

Nederlandstalige samenvatting

In dit werk worden verschillende toepassingen bestudeerd waar een laag vermogenverbruik belangrijk is.

Optische verbindingen

De rekenkracht van geïntegreerde circuits neemt steeds verder toe. Hierdoor moet ook steeds meer data rond gestuurd worden. Tot nu toe hadden elektrische verbindingen voldoende bandbreedte om dit te doen. Er is echter een probleem: alhoewel transistoren steeds sneller worden wanneer ze verder geminiaturiseerd worden, is dit niet het geval voor elektrische connecties. Het gevolg hiervan is dat de elektrische verbindingen een beperking dreigen te worden op de snelheid van geïntegreerde circuits. Een mogelijke oplossing voor dit probleem is het gebruik van geïntegreerde optische verbindingen [1], aangezien deze geen last hebben van dergelijke beperkingen. Om integratie toe te laten met elektronische circuits, moet de gebruikte optische technologie voor de realizatie van dergelijke optische verbindingen CMOS compatibel zijn. Een veelbelovend platform in deze context is het silicium-op-isolator golfgeleider platform, aangezien de gebruikte materialen en fabricageprocessen dezelfde zijn als voor de fabricage van de elektronische circuits. Voor de realizatie van een optische verbinding zijn echter ook actieve componenten nodig, wat niet triviaal is gezien het feit dat silicium een halfgeleider met indirecte bandgap is. De beste oplossing voor deze uitdaging is op dit moment de heterogene integratie van een III-V epitaxiaal lagenstructuur bovenop de silicium-op-isolator golfgeleider circuits [2, 3, 4].



Figuur 1: Voorgestelde elektrisch geïntegreerde golfgeleiders: de V-golfgeleider (links) and the gallerij golfgeleider (rechts).

Om echter competitieve optische verbindingen te kunnen maken, zijn de vereisten voor de geïntegreerde lasers en modulatoren erg uitdagend: ze moeten in verhoogde temperaturen werken, een laag vermogenverbruik hebben en bovendien een zo klein mogelijk oppervlak innemen. Om aan deze vereisten te voldoen, stellen we in dit werk voor om het hoge brekingsindex contrast tussen een gebonde III-V golfgeleider en de omgevende oxide- en polymeercladding uit te buiten. Twee golfgeleiderstructuren worden besproken. Om de optische confinement in de actieve regio van de golfgeleider te optimalizeren, wordt de gallerij golfgeleiderstructuur voorgesteld. De gallerij golfgeleider is een ridge golfgeleider waarbij groeven geëetst zijn in de bovenste p-InP cladding laag. Door deze groeven wordt de effectieve index van de bovenste cladding laag verkleind, waardoor de confinement in de actieve laag verhoogd wordt (zie figuur 1). Door van deze techniek gebruik te maken, kan de netto winst voor de fundamentele mode in de golfgeleider verdubbeld worden in vergelijking met de klassieke ridge golfgeleiderstructuur. Zo'n golfgeleider zou bijgevolg ideaal zijn voor de realizatie van een on-chip microlaser, aangezien de hoge netto winst compactere lasers met een lagere drempelstroom toelaat.

Een tweede golfgeleiderstructuur die voorgesteld wordt in dit werk, is de V-golfgeleider. In deze elektrisch geïntegreerde golfgeleider is de bovenste p-InP cladding laag zodanig geëetst dat deze de vorm van een trapezium krijgt zoals te zien is in figuur 1. Dit is erg handig, aangezien het ons toelaat om een smalle p-InP pilaar te definiëren van slechts $500\,\mathrm{nm}$ breed, terwijl de bovenste contactlaag toch nog $1\,\mu\mathrm{m}$ breed is. Dit betekent dat deze structuur toch nog lithografisch gedefinieerd kan worden met contact lithografie. Een ander voordeel is

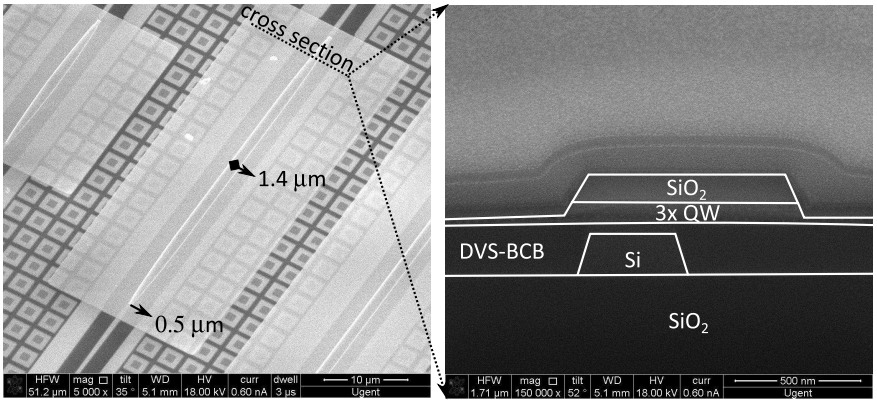
dat deze golfgeleiderstructuur een erg lage specifieke capaciteit heeft. Vandaar dat deze golfgeleiderstructuur in dit werk gebruikt is om een elektroabsorptiemodulator te ontwerpen. Via optimalisatie van de actieve lagenstructuur, is een hoog-performante modulator ontworpen op basis van de V-golfgeleider met een laag vermogenverbruik. De ontworpen modulator heeft slechts een piek-tot-piek spanning nodig van $0.45\,\mathrm{V}$ om een signaalcontrast van meer dan $10\,\mathrm{dB}$ te bereiken, zonder dat een DC spanning noodzakelijk is. Hierdoor is het vermogenverbruik erg laag: aan een bit rate van $40\,\mathrm{Gbit/s}$ bedraagt het vermogengebruik slechts $48\,\mathrm{fJ/bit}$. Daarnaast heeft de ontworpen modulator ook een hoge bandbreedte van $26.5\,\mathrm{GHz}$ gebruik makend van een lumped elektrode ontwerp.

Optische netwerken

Ook in optische netwerken neemt de benodigde bandbreedte snel toe, door de introductie van nieuwe diensten zoals video-on-demand en cloud computing. Hierdoor wordt de zogenaamde 'laaste mijl' tussen de providers en gebruikers vervangen door optische vezel. Om historische redenen werd hiervoor het reeds aanwezige koperen kabel- of telefoonnetwerk gebruikt. Tot nu toe zijn vooral passieve optische netwerken uitgerold. Deze netwerkarchitectuur heeft als voordeel dat ze goedkoop is om te implementeren, aangezien verschillende gebruikers een optische vezel delen voor de verbinding met de provider, in combinatie met een simpele vermogensplitter. Er zijn echter ook verschillende nadelen. De architectuur is niet erg energie-efficiënt, aangezien de verstuurde optische signalen vanuit de provider gewoon gesplitst worden tussen de verschillende gebruikers. Bovendien betekent dit ook dat de gebruikers ook data ontvangen die niet voor hen bedoeld is. Hierdoor wordt veel energie verspild aan de hogesnelheidssignaalverwerking van data die toch niet gebruikt zal worden door de gebruiker.

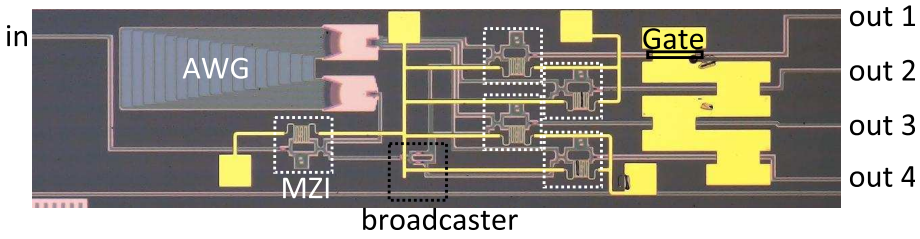
Om dit probleem op te lossen, stellen we een optische schakelarchitectuur voor. In dergelijke architectuur, ontvangt elke optische netwerkeenheid enkel nuttige data, waardoor heel wat energie uitgespaard kan worden, bijvoorbeeld door een slaapmodus te gebruiken wanneer geen data ontvangen wordt. Daarnaast kan het insertieverlies van een dergelijke schakelaar kleiner gemaakt worden dan het intrinsieke verlies in een splitter.

Om een dergelijke schakelaar te realiseren, wordt eerst een optisch controleerbare gate ontwikkeld. De optische controle is belangrijk,



Figuur 2: SEM opname van membraam golfgeleider en doorsnede door de taper sectie.

aangezien de schakelaars geplaatst zullen worden op locaties waar typisch geen elektrische aansluiting aanwezig is. Omwille van het feit dat er geen elektrische injectie is, is er een volledige vrijheid in het ontwerp van de golfgeleiderstructuur. Bijgevolg kan de optische confinement in de QWs van deze golfgeleiderstructuur gemaximaliseerd worden. Dit zou moeten leiden tot een sterke licht-materiaal interactie, en bijgevolg een hoge absorptie en winst in een korte golfgeleider. Dit is het geval wanneer de golfgeleider ontworpen wordt met een dikte van slechts $100\,\mathrm{nm}$ (zie figuur 2). In een eerste experiment wordt aangetoond dat dergelijke membraamgolfgeleiders inderdaad een groot potentieel hebben voor optische winst, aangezien een gepulste winst van meer dan $8\,\mathrm{dB}$ werd bereikt in een golfgeleider van slechts $100\,\mu\mathrm{m}$ lang. Echter, in continuë werking kon slechts een winst van ongeveer $2\,\mathrm{dB}$ aangetoond worden, door verhitting van de component, zelfs na optimalisatie van de golfgeleiderstructuur om de thermische karakteristieken te verbeteren. Blijkbaar wordt de grootste beperking voor optische winst gevormd door de kwaliteit van de geëetste zijwanden van de QWs, aangezien de etskwaliteit een directe impact heeft op de oppervlakterecombinatiesnelheid. Het gevolg hiervan is dat de niet-radiatieve levensduur van vrije ladingsdragers slechts $500\,\mathrm{ps}$ bedraagt. Doordat de levensduur zo kort is, wordt veel hitte gedissipeerd door niet-radiatieve recombinatie. Door de passivatie van de geëtste zijwanden te verbeteren, kan de levensduur sterk verhoogd worden. Dit wordt aangetoond in een simulatie, en de eerste fabricageresultaten tonen een verbetering van $20\,\%$ aan in



Figuur 3: Optische microscopie opname van een gefabriceerde 1x4 schakelaar.

de levensduur. Op basis van de resultaten gepubliceerd in [5] kan de levensduur nog verder verhoogd worden tot meer dan $2\,\mathrm{ns}$ met een geoptimaliseerd passivatierecept. Met zo'n lange levensduur zou een veel lager vermogenverbruik mogelijk zijn, aangezien populatie inversie al gerealiseerd zou zijn bij een lager input pompvermogen. Bijgevolg kan het oververhittingsprobleem vermeden worden, en zou een hoge optische winst ook in continuë werking mogelijk moeten zijn.

Gebruik makend van de ontworpen membraam gate, werd een 1x4 optische schakelaar ontwikkeld (zie figuur 3). Om de werking van de gate als schakelingselement te controleren, zijn twee experimenten uitgevoerd. In het eerste experiment wordt aangetoond dat een alleenstaande gate toelaat om een erg hoog aan-uit contrast can meer dan $30\,\mathrm{dB}$ te bereiken. Bovendien is er geen vermogen penalty gerelateerd aan de transmissie door de gate. In een tweede experiment wordt aangetoond dat de transparante golflengteband voor golflengtes langer dan de bandgap golflengte gebruikt kan worden om data, zoals bijvoorbeeld televisiekanalen, naar alle geconnecteerde gebruikers te sturen. Daarnaast kan ook een kloksignaal over het netwerk gestuurd worden in zo'n langer golflengtekanaal, waardoor de optische netwerkeenheden bij de ontvangers een stuk minder complex kunnen worden, aangezien er in dat geval geen circuit voor klokrecuperatie meer nodig is. In een laatste experiment wordt ook een volledige 1x4 optische schakelaar gedemonstreerd. De broadcast-and-select architectuur wordt hiervoor gebruikt, aangezien dit de meest robuuste architectuur blijkt voor fabricage gerelateerde variaties in de passieve silicium componenten. Met de gerealiseerde schakelaar wordt transmissie naar alle geconnecteerde gebruikers aangetoond, zonder een meetbare vermogen penalty ten gevolge van de schakelaar. Bovendien wordt een hoge isolatie tussen de uitgangspoorten van meer dan $25\,\mathrm{dB}$ aangetoond over een brede golflengteband. Het insertieverlies van de schakelaar is echter nog te

hoog: het verlies in vergelijking met een simpele passieve splitter is $20\,\mathrm{dB}$. Dit verlies is een gevolg van het gebruik van niet geoptimaliseerde passieve componenten, en de slechte werking van de gebruikte gates. In simulatie kan echter aangetoond worden dat dit grote verlies kan omgebogen worden in een netto winst indien we aannemen dat de huidige state-of-the-art passieve componenten gebruikt kunnen worden, in combinatie met een goed gepassiveerde gate zoals hierboven besproken.

Naast toepassing in access netwerken kan de membraam gate ook een rol spelen in optische pakket schakelaars voor toepassing in metro netwerken. Op dit moment worden namelijk elektische schakelaars gebruikt, waardoor er twee maal een electro-optische convertie nodig is. Hierdoor wordt heel wat energie en tijd (hogere latency) verspild. Dit is de reden waarom er ook onderzoek gevoerd wordt naar de implementatie van optische schakelaars die deze elektische schakelaars kunnen vervangen. Om dit doel te bereiken zijn er echter complexe optische signaalverwerkingscircuits nodig. Een van de componenten die een rol kan spelen in een dergelijk cicruit, is een on-chip regenerator. Blijkbaar is de membraam golfgeleider ideaal om deze taak te vervullen: door de sterke licht-materie interactie heeft de component een sterk niet-lineaire transmissie. Wanneer er nu een signaal door de membraamgolfgeleider gestuurd wordt met een gemiddeld vermogen die binnen die niet-lineaire regio van transmissie valt, dan zal het membraam het nul-niveau van het signaal veel sterker absorberen dan het een-niveau. Bijgevolg wordt het signaal contrast sterk verbeterd. Dergelijke regeneratie wordt experimenteel aangetoond aan een bit rate van $2.5\,\mathrm{Gbit/s}$, met een signaal contrast verbetering van $2\,\mathrm{dB}$ aan de input tot $6.2\,\mathrm{dB}$ aan de output in een $100\,\mu\mathrm{m}$ lange waveguide. Dit leidt tot een sensitiviteitsverbetering aan de ontvanger van $3.6\,\mathrm{dB}$ over de gehele C-band, tot $4.5\,\mathrm{dB}$ in de golflengteband $1530 - 1540\,\mathrm{nm}$. Alhoewel het membraam voor deze toepassing niet gepompt hoeft te worden door een extern pompsignaal, is er toch nog steeds een kost voor regeneratie: wanneer men een hoger signaal contrast wil bereiken, zal daar een hoger insertieverlies tegenover staan.

We kunnen besluiten dat er in dit werk aangetoond is dat er nog veel mogelijkheden zijn om III-V op silicium componenten te verbeteren in verschillende toepassingen, door relatief eenvoudige ingrepen door te voeren in het design van de gebruikte golfgeleiderstructuren.

English text

Summary

In this work, several applications are highlighted where power efficiency is extremely important.

Optical interconnects

As the computational power of integrated circuits continues to increase, more and more data needs to be pushed around a chip. Until now, electrical connections have been able to provide for the necessary bandwidth to perform this task. There is a problem however: although transistors become faster when they are further miniaturized, this is not the case for electrical connections. The result is that the electrical interconnects are becoming a bottleneck. An integrated optical solution might offer the bandwidth required for next-generation computing [1], as these are not limited by the same constrains as electrical interconnects. To allow close integration with electronic systems, the optical technology to achieve this should be CMOS compatible. An excellent candidate in this perspective is the silicon-on-insulator waveguide platform, as the materials and processing involved are essentially the same as for electronics. To achieve an optical interconnect however, also active devices are required, which is not trivial as silicon is an indirect band gap semiconductor. Currently the most successful approach to solve this problem is the heterogeneous integration of a III-V epitaxial layer stack on top of the silicon-on-insulator waveguide circuits [2, 3, 4].

In on-chip interconnects, the requirements on integrated lasers and modulators are very strenuous however, as they need to operate at elevated temperatures, with a low power consumption and small device footprint. To meet these requirements, it is proposed in this work to leverage the high index contrast between a bonded III-V waveguide and the surrounding oxide and polymer cladding layers.

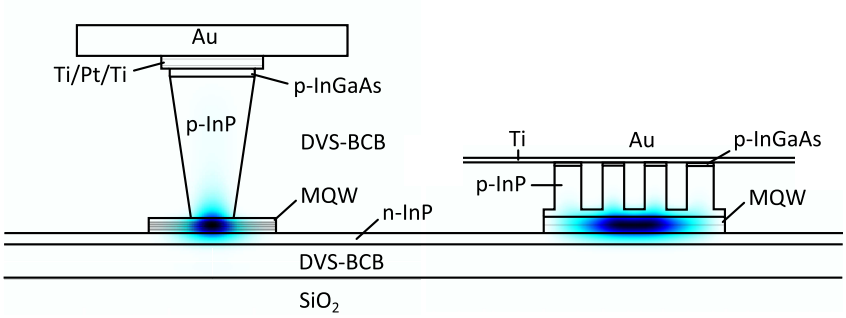


Figure 4: Proposed electrically injected waveguides: the V-waveguide (left) and the gallery waveguide (right).

Two waveguide structures are highlighted. The gallery waveguide structure, which is a ridge waveguide with trenches etched into the top p-InP cladding layer, is proposed as a means to optimize the optical confinement inside the active region of the waveguide. Because of the etched trenches, the effective index of the top cladding layer is reduced and therefore the confinement in the active region is enhanced (see figure 4). It is shown that using this technique, the net modal gain can be doubled compared to the classical ridge waveguide structure. Such a waveguide would therefore be very suitable to use in an on-chip microlaser, as the high net gain allows for more compact devices, with lower threshold currents.

A second waveguide structure that is proposed in this work is the V-waveguide. In this electrically injected waveguide, the top p-InP cladding layer is etched in such a way that a trapezoid shape is achieved (see figure 4). This is very useful, as it allows us to create a narrow p-InP pillar of $500\,\mathrm{nm}$ wide near the optical mode, while the top contacting layer remains $1\,\mu\mathrm{m}$ broad, which means that it can still be defined using simple contact lithography. Another advantage of this waveguide structure is that it intrinsically has a small specific capacitance. For this reason, an electro-absorption modulator was designed based on the V-waveguide structure. It was shown that after optimization of the active region, a high performance modulator can be made with a very low power consumption. The proposed modulator requires a driving peak-to-peak voltage of only $0.45\,\mathrm{V}$ to achieve an extinction ratio of more than $10\,\mathrm{dB}$ without the need for a DC bias. As a result, a very low power consumption of $48\,\mathrm{fJ/bit}$ at a bit rate of $40\,\mathrm{Gbit/s}$ is predicted. Furthermore a high bandwidth of $26.5\,\mathrm{GHz}$ can be achieved

using a lumped electrode design. It should be pointed out that these are still results of a simulation in which fabrication and material quality is assumed ideal, and therefore the absolute values given above should be taken with a grain of salt. Nevertheless, the potential of a modulator based on the V-waveguide structure is clear.

Optical networks

Also in optical networks the required bandwidth is growing rapidly, due to the advent of new services like video-on-demand and cloud computing. As a result, the so-called 'last mile' between the providers and users which historically uses the copper wiring from cable or telephone networks, is being replaced by optical fiber. Currently, mostly passive optical networks are deployed. This network architecture has the advantage that it is very cheap to implement, as several users share a single fiber to the provider connected to a simple power splitter. However, there are also several disadvantages. First of all, they are not very power efficient, as the power is just split between the different connected users. Furthermore, as each user also receives the downstream data meant for other users, a lot of power is wasted in the high-speed processing of data which is going to be discarded anyway.

To address this problem, an all-optical switch architecture is proposed. In such an architecture, each optical network unit only receives useful data, and therefore a lot of power can be saved, e.g. by using a sleep mode when no data is incoming. Next to this, in a switched architecture the insertion loss could be made lower than the intrinsic splitter loss.

To be able to realize such a switch, first an all-optically controlled gate on SOI is developed. The all-optical control is important, as the switches will have to be placed in locations where typically no electrical power is present. Because of the fact that no electrical injection is required, there is a complete freedom in the design of the waveguide structure. Therefore, the waveguide can be optimized to maximize the confinement in the QWs. This should lead to a strong light-matter interaction, and therefore a high absorption and gain in a compact device. It is found that this is the case when the waveguide is a membrane of only $100\,\mathrm{nm}$ thick (see figure 5). In a first experiment, it is shown that such a membrane waveguide has indeed a large potential for gain, as a pulsed gain of more than $8\,\mathrm{dB}$ is demonstrated in a device of only

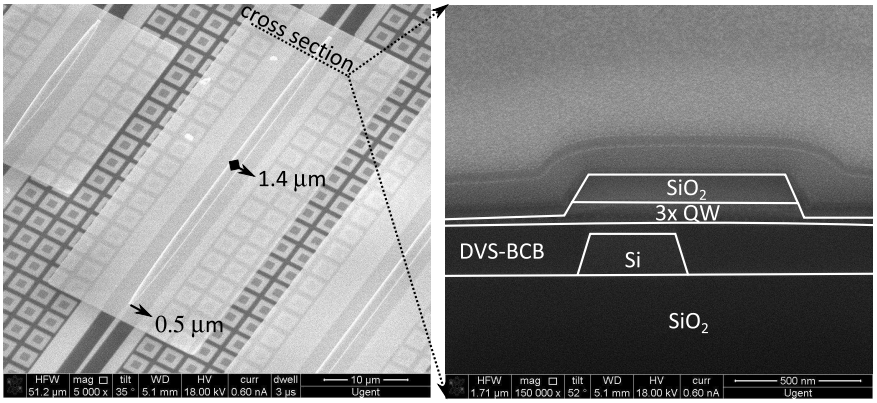


Figure 5: SEM micrograph of a membrane waveguide and cross section through the taper section.

$100\,\mu\mathrm{m}$ long. In CW however, the demonstrated gain remained limited to approximately $2\,\mathrm{dB}$, due to the self-heating of the device, even after optimization of the waveguide structure. It is found however that the main limitation on the device gain is formed by the quality of the etched sidewalls, as this has a direct influence on the surface recombination rate. As a result, the non-radiative carrier lifetime is of the order of $500\,\mathrm{ps}$. Because of this short carrier lifetime, a lot of heat is generated by non-radiative recombination. By improving the passivation of these sidewalls, the lifetime can be greatly increased. This is shown in simulation, and the first fabrication attempts show a $20\,\%$ improvement in the carrier lifetime. Based on results found in literature [5], the carrier lifetime can be increased to more than $2\,\mathrm{ns}$ with an optimized passivation recipe. Such a long carrier lifetime would allow for a much lower power consumption, as population inversion would be achieved at a lower input power. As a result, the self-heating problem can be avoided, and a high CW gain should be possible.

Using the designed membrane gate, a 1x4 all-optical switch is developed (see figure 6). To assess the performance of the gate as a switching element in a larger switching fabric, two experiments are performed. First of all, it is shown that a single gate allows for a very high extinction ratio of more than $30\,\mathrm{dB}$, and that furthermore there is no power penalty related to the transmission through the gate. Secondly, it is shown that the transparent band for wavelengths longer than the bandgap wavelength can be used to broadcast data, such as television services, or even to broadcast a clock signal over the access

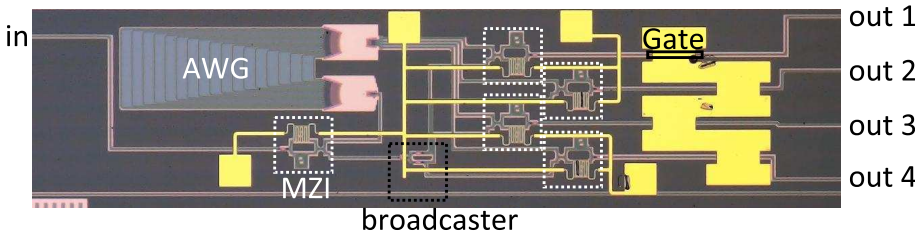


Figure 6: Optical microscopy image of fabricated 1x4 switch.

network. In this way, the optical network units become much less complex, as no clock recovery circuitry is required anymore. Finally, also a full 1x4 switch is demonstrated. The broadcast-and-select switch architecture is used, as this is the most robust architecture for variations in the processing of the passive silicon components. Using this architecture, transmission to all the output ports is achieved without a measurable power penalty. Furthermore, a high port isolation of more than $25\,\mathrm{dB}$ is demonstrated over a broad signal wavelength range. However, insertion loss was too high: the excess loss of the switch compared to a simple passive splitter is determined to be $20\,\mathrm{dB}$. This loss is related to the use of both imperfect passive components and the poor performance of the used gates. However, it is shown in simulation that this large excess loss can be turned into a net gain, by using state-of-the-art passive components in combination with a properly passivated gate.

Next to the application in access networks, the membrane gate could also be used as a component in optical packet switches in metro networks. Currently there is a large research effort directed towards the realization of optical switches that could replace the currently used electrical switches in metro networks, in an effort to reduce the power consumption and delays. To achieve this, complex optical signal processing circuits are necessary. One of the components that could play a role here, is an on-chip regenerator. It seems that the membrane waveguide is ideal to perform this task: because of the strong light-matter interaction, the device has a very non-linear transmission. When a signal is incoming with an average power within the non-linear region of transmission, the membrane will absorb the zero-level much stronger than the one-level of the signal. As a result, the ER of the signal is improved. Regeneration in this way is demonstrated at $2.5\,\mathrm{Gbit/s}$, with an ER improvement from $2\,\mathrm{dB}$ at

the input up to $6.2\,\mathrm{dB}$ using a $100\,\mu\mathrm{m}$ long device. This leads to a receiver sensitivity improvement of $3.6\,\mathrm{dB}$ over the entire C-band, up to $4.5\,\mathrm{dB}$ at wavelengths $1530-1540\,\mathrm{nm}$. Although the device does not need to be pumped externally, the regeneration still comes at a cost: for a higher extinction ratio improvement, also the device insertion loss becomes higher.

In conclusion, it was shown in this work that there is still a lot of headroom for improvement of III-V on silicon devices, by relatively simple changes in the device waveguide structure.

List of Acronyms

AC	Alternating current
ASE	Amplified spontaneous emission
AWG	Arrayed waveguide grating
BER	Bit error rate
B2B	Back-to-back
CMOS	Complementary Metal-Oxide-Semiconductor
CO	Central office
CW	Continuous wave
DC	Direct current
DFB	Distributed feedback
DI	Deionized
DVS-BCB	Divinyltetramethyldisiloxane-benzocyclobutene
EAM	Electro-absorption modulator
EDFA	Erbium doped fiber amplifier
ER	Extinction ratio
FCA	Free carrier absorption
FCD	Free carrier dispersion
FSR	Free spectral range
FTTH	Fiber to the home
ICP	Inductively coupled plasma
MMI	Multimode interference coupler
MQW	Multiple quantum well
MZI	Mach Zehnder interferometer
NRZ	Non-return to zero

ONU	Optical network unit
OOK	On-off keying
OSNR	Optical signal-to-noise ratio
PECVD	Plasma enhanced chemical vapor deposition
PL	Photoluminescence
PON	Passive optical network
PRBS	Pseudorandom binary sequence
QCSE	Quantum confined stark effect
QW	Quantum well
RIE	Reactive ion etch
RF	Radio frequency
SCH	Separate confinement heterostructure
SEM	Scanning electron microscope
SOI	Silicon-on-insulator
SRH	Shockley-Read-Hall
TCAD	Technology computer aided design
TDM	Time division multiplexing
TDMA	Time division multiple access
TE	Transverse electric
TM	Transverse magnetic
UV	Ultra violet
WDM	Wavelength division multiplexing

Chapter 1

Introduction

Power efficiency and device footprint are important for many applications. In the following, three important applications for heterogeneously integrated III-V membranes are highlighted where relevant results were obtained in this phd thesis: on-chip optical interconnects, optical signal processing and all-optical switches.

1.1 On-chip optical interconnects

The current tendency in high performance computing is directed towards parallelism on all levels. As the number of cores, executional units and transistors increases, the amount of data that needs to be sent across a chip increases as well. Until now, electronic interconnects were sped up in order to meet this increasing demand, however the speed of electronics is fundamentally limited. In [6], an expression for the fundamental limit of the bandwidth in an electrical interconnect is given:

$$B \propto B_0 \frac{A}{L^2}$$

Here, A is the cross sectional area of the electrical wiring, L is the length of the wire, and bit rate B_0 lies in the range of $10^{15} - 10^{18}\mathrm{bit/s}$, depending on the interconnect type that is being considered. From this formula it can be seen that the interconnect bandwidth is proportional to A/L^2. While transistors can be sped up by miniaturization, this is not the case for electrical interconnects as both the length and the cross sectional area will decrease by a similar amount. For this reason, an alternate solution needs to be found. Similar as what has happened in the past in long haul communication networks, an optical solution might

also offer the way forward to meet the bandwidth demand in future high performance electronic chips [1]. Optical interconnects have many advantages over their electrical counterparts. First of all, the bandwidth limitation given above does not hold for optical links. Next to this, they do not suffer from frequency dependent loss and cross talk. They are also immune to electromagnetic noise, and impedance matching and reflection problems are easier to solve. The reason why the transition to optical interconnects has not occurred yet, is due to the fact that such interconnects should have a very low operating power of the order of $1\,\mathrm{pJ/bit}$ [7] and preferably be even lower for the shortest on-chip interconnects [8] in order to be practical for the integration on a chip.

To allow close integration with electronic systems, the optical technology to achieve this should be CMOS compatible. An excellent candidate in this perspective is the silicon-on-insulator waveguide platform, as the materials and processing involved are essentially the same as for electronics. As a proof of its potential, recently chips with both electronics and optics integrated on the same die were demonstrated [9].

Silicon-on-insulator is an excellent platform for the realization of passive optical functions. Low-loss waveguides have been demonstrated and the high index contrast between the silicon core and oxide cladding allows the creation of very compact circuits [10]. However, the lack of an easily integrated light-emitting material is an important challenge. To that end several possible solutions are pursued. A first way to solve this problem, is to try to engineer the energy bands of silicon or another material that can be grown on silicon, to create a direct band gap material on silicon. Strained germanium is one such candidate that could lead to a future monolithically integrated laser on silicon [11].

Another strategy that is being pursued is the direct growth of III-Vs on silicon, as these material systems are typically used for light emission and detection. This is a very challenging task however, as the lattice constant of silicon ($5.43\,\text{Å}$) differs a lot from the lattice constants of light-emitting III-V alloys ($5.7 - 6.3\,\text{Å}$). This has the consequence that a lot of defects will be present in III-Vs grown directly on silicon. These act as non-radiative recombination centers, which greatly reduce the material's capability to emit light. To resolve this, research has been focusing on creating barriers that can avoid that the defects in the seed layer can propagate through the grown material. One of the most promising techniques is to grow the III-V in narrow, deep trenches, as the defects typically propagate under an angle of 45 degrees with the

material surface. Therefore, most of the propagating defects in the material can be terminated on the sidewall of the trench and defect-free III-V can be obtained at the surface. Although lasing devices have been obtained using this technique [12, 13], further work is still required to obtain a practical laser with high performance on silicon.

A third way to overcome this problem is by using colloidal quantum dots. These quantum dots are created chemically in a reaction vessel and can therefore be manufactured easily in large numbers. Furthermore, their properties such as the emission wavelength can be tuned by changing the reaction conditions and reaction time. The idea is then to deposit these quantum dots selectively on a waveguide circuit and use them as an active material. Optically pumped lasers have already been demonstrated [14]. Furthermore, an integration scheme on SOI using the CMOS-compatible silicon nitride as a matrix material has been shown in [15]. However also in this case, further work is required to obtain commercially viable devices.

Finally, the currently most successful way of creating active devices on silicon is by heterogeneous integration of III-Vs on SOI. In this approach, an epitaxially grown III-V layer stack is bonded on top of a SOI waveguide circuit through either die-to-wafer or wafer-to-wafer bonding. This can be achieved by either direct bonding [16], or adhesive bonding using DVS-BCB as an intermediate adhesive between the SOI circuit and the III-V layers [17]. After bonding, the III-V layers are then further processed to create an active device on top of the SOI circuit. Lasers [18, 2, 19], modulators [20, 3, 21] and photodetectors [22, 4] have been demonstrated using this technique, however in these demonstrations the full potential of the material system was not yet used. In this work, we propose to use the high index contrast between the bonded III-V layer and the surrounding DVS-BCB/oxide cladding layers, to increase the efficiency and reduce the footprint of these active devices by optimizing the geometry of the III-V waveguide.

1.1.1 Microlaser

A first important component which is necessary for the realization of on-chip optical interconnects is a microlaser. In a demanding application as on-chip optical interconnects, this laser must be of extraordinary quality. First of all, the power consumption should be minimal, as there is only a limited thermal budget: the dissipated power density of the highest performing electronic chips is already approaching that in a

nuclear reactor [23]. Related to this, the lasers should also be able to operate at elevated temperatures. Next to this, the density of optical links that can be achieved is an important metric, and therefore also the device footprint should be minimal. For these reasons, a compact laser should be created with a very low threshold current. To be able to achieve this, the waveguide providing the gain in the laser should be designed in such a way that the net modal gain is maximized. This problem is addressed in section 2.2.

1.1.2 Modulator

In the class of very high speed optical interconnects where direct modulation is not an option anymore, on-chip modulators are required. Also for this device, the same metrics apply as for the microlaser: both real estate and power consumption should be minimized. To achieve this, the design of a heterogeneously integrated electro-absorption modulator is investigated. It is found that in this case the specific capacitance of the modulator waveguide should be minimized, as this allows the active region to become thinner without losing the potential to achieve high bit rates. Because a thinner active region correponds to a lower driving voltage, the power consumption could be greatly reduced in this way. This is further discussed in section 2.3.

1.2 Access networks

The advent of optical fiber networks has revolutionized communications across the world. Thanks to the huge bandwidth of optical networks in combination with a low operating cost, the Internet as we know it today was made possible. Until recently only the core and metro networks of the Internet service providers were using optical fiber, while the so called 'last mile' between the customer and the provider was using historically present copper infrastructure from obsolete cable and telephone services. As more and more bandwidth hungry services such as video-on-demand and cloud computing are being deployed however, the demand for bandwidth is still continuously increasing. Therefore, the last mile is now also being replaced by fiber, to form fiber-to-the-home networks. To keep the cost of the network as low possible, typically passive optical networks are being deployed. In this concept, a number of users share a single fiber to the central office

by using a passive splitter in combination with time-division multiplexing to combine all the signals from the different users on the fiber.

However, it is expected that the relentless increase in demand for bandwidth is not going to stop any time soon [24]. For even higher bandwidth, the solution offered by passive optical networks does not work well anymore, as transferring a higher bandwidth requires more power, and therefore reduces the possible splitting ratio and network reach of a PON. Furthermore, because the optical network units at the user side receive also the downstream data meant for other users, a lot of power is wasted in the high-speed processing of this data which is going to be discarded anyway.

In this context, an alternative network solution which uses an active all-optical switch instead of a passive splitter is proposed in this work. Such a switched network architecture has several advantages. First of all, the insertion loss of the switch can be made lower than the intrinsic splitter loss in a PON. The gained power budget can then be used to achieve higher bandwidths, connect more users or increase the distance between the users and the central office. Secondly, the optical network units only receive useful data, and therefore the power consumption can be made much lower.

1.2.1 All-optical gate

To realize such a switch, a heterogeneously integrated optically controlled gate on SOI is proposed. Because of the high index contrast in bonded III-V films between waveguide core and cladding, a very strong light-matter interaction can be obtained in an ultrathin III-V membrane waveguide. The design, fabrication and characterization of these optical gates is discussed in chapter 3.

Next to the application in an all-optical switch, the membrane gate can also be used in other signal processing applications. In chapter 4, we show that the strong light-matter interaction in the device makes the device ideal for the realization of a passive optical signal regenerator. Although signal regenerators are typically used in long-haul telecommunications, where they are deployed at a regular distance to keep the signal integrity, they also could become an important building block on a much smaller scale. There is currently a large research effort directed towards the realization of optical switches that could replace the currently used electrical switches in metro networks. To achieve

this, complex optical signal processing is required, and in this context on-chip regenerators could play a role.

1.2.2 All-optical switch

Finally, in chapter 5 the different possible switching architectures that use the developed all-optical gate for the realization of an all-optical switch are explored. The broadcast-and-select switching fabric is selected, and a 1x4 switch using this architecture is demonstrated.

Chapter 2

Electrically driven devices

In this chapter we will investigate the operation of heterogeneously integrated electrically injected p-i-n diodes, and how their shape can be optimized to achieve a much stronger light-matter interaction. Two different applications will be highlighted. First, a waveguide will be designed to optimize the gain in a bonded amplifier. Secondly, an integrated electro-absorption modulator is designed and simulated, where we show that the decreased footprint and therefore decreased device capacitance allows a high bandwidth and low power consumption to be achieved.

2.1 Introduction

Using bonding technology for heterogeneous integration of III-Vs on SOI, several modulators and lasers have already been achieved [20, 3, 21, 25, 26, 2, 19]. However, in these approaches the top p-InP current injecting layer was typically more than $1\,\mu\mathrm{m}$ thick to avoid excessive absorption losses for the optical mode in the top contact layers. Furthermore, to optimize the confinement of light in the active region, SCH layers of a high index material are typically added around it. Although this leads to working devices, this complicates the device design, as this also increases the effective index of the mode in the III-V, making adiabatic coupling between the III-V waveguide and (thin) underlying SOI waveguides more difficult.

In this chapter, we will show that by using the high index contrast between the III-V core of the waveguide and the low-index DVS-BCB/oxide cladding, the light can be confined more efficiently in the

active region, while a relatively low effective index can be obtained. A first method to achieve this, is by etching narrow trenches in the top p-InP layer and filling these with a low-index polymer. In this way, the effective index of the upper p-InP cladding can be reduced, pushing the mode down into the active region. Furthermore, the position of the remaining p-InP pillars can be chosen in such a way that current is only injected where the optical mode is present, increasing the efficiency of the current injection. Such a waveguide structure is particularly useful for the realization of a microlaser. In the rest of this work, this waveguide structure is called a gallery waveguide. To assess the performance of an amplifier based on the gallery waveguide, we will compare the performance with a classical p-i-n diode ridge waveguide. For that end, an amplifier with an active region that contains a single QW will be designed using both the gallery and the classical waveguide structure, and both will be simulated using the FDM modesolver in Fimmwave [27] from Photon Design and the TCAD software package ATLAS from Silvaco International [28]. This software supports fully coupled thermal, electrical and optical simulation of semiconductor devices, and allows us to compare the performance of both waveguide structures.

A second method to improve confinement in the active region uses the sidewall slope under anisotropic wet etching of InP with a diluted HCl solution. If one takes care to orient the bonded III-V film in the correct orientation, a negative etching slope can be achieved, which can be used to create a narrow p-InP base of the order of $500\,\mathrm{nm}$ wide, using a mask of $1\,\mu\mathrm{m}$ wide. The resulting waveguide confines the mode very well, and has the additional advantage that it has a low electrical capacitance, making it potentially useful for the realization of a compact, high-speed electro-absorption modulator. To assess its performance, the optical response and bandwidth of a modulator using such a waveguide structure are simulated using ATLAS.

2.2 High-gain amplifier for low-threshold microlasers

A first application that is highlighted is the design of a gain section meant for the realization of a low-threshold microlaser. For the realization of on-chip optical interconnects, a low-threshold laser with a small device footprint and high tolerance for variations in device tempera-

ture is an important prerequisite. Also for the realization of biosensors, such as an in-vivo optical glucose sensor as proposed in [29], an integrated microlaser array integrated on SOI is necessary. As this sensor is implanted in the subject, a low power consumption is required in order to avoid the need for a large battery.

In order to realize a microlaser with an as low as possible device footprint and threshold current, the modal gain that can be achieved for a certain number of QWs needs to be maximized, because in that case the gain required to start lasing will be reached at a lower threshold current. Furthermore, a higher modal gain also allows more compact devices. A drawback is that the saturation output power of the laser will be lower than in a laser with low confinement, however for the applications mentioned above an output power of less than $1\,\mathrm{mW}$ is more than enough [29], so this is not an issue.

2.2.1 Optical optimization

Methodology

To be able to compare the performance of the proposed gallery waveguide structure with the classical waveguide structures, the design of both waveguides needs to be optimized. To be able to make a fair comparison, the material composition of the QW and SCH layers is the same. For the QW, an $8\,\mathrm{nm}$ $In_{0.47}Ga_{0.53}As_{0.88}P_{0.12}$(Q1.55) is used, sandwiched between two $In_{0.47}Ga_{0.53}As_{0.46}P_{0.54}$(Q1.2) barriers, that also form the SCH layers. Furthermore, the width and thickness of several layers is set to a fixed value as the influence on the performance of both waveguides is similar. The value of these are given in table 2.1.

To be able to fully optimize the design, the gain and losses in each section of the waveguide need to be known as well. Therefore, also the doping profile is kept fixed, and a QW material gain of $1000\,/\mathrm{cm}$ is assumed. From this doping profile, also the loss due to FCA can be calculated, as this loss is linearly proportional to the doping concentration, with proportionality constants of $1 \times 10^{-18}\,\mathrm{cm}^2$ in n-InP, $24 \times 10^{-18}\,\mathrm{cm}^2$ in p-InP, and $40 \times 10^{-18}\,\mathrm{cm}^2$ in InGaAsP (see section A.6.3). When pumping the device to reach a material gain of $1000\,/\mathrm{cm}$ however, a number of free carriers will also be present in the SCH layers, and this will give rise to an additional loss due to FCA. To estimate this loss, we simulated the electro-optical performance of an unoptimized classical waveguide structure. Using this simulation, the loss due to FCA in the SCH layers was determined to be $12\,/\mathrm{cm}$ at a material gain of $1000\,/\mathrm{cm}$,

Table 2.1: Dimensions for the waveguide structure, given bottom up.

Region	thickness (nm)	width ($\mu\mathrm{m}$)
SiO_2	2000	-
BCB	450	-
n-InP	150	25
SCH	-	3
QW	8	3
p-InP	-	3
p-InGaAs	50	3
Ti	40	3
Au	400	3

Table 2.2: Material properties for the waveguide structure, given bottom up.

Region	index	doping ($/\mathrm{cm}^3$)	gain ($/\mathrm{cm}$)
SiO_2	1.444	-	-
BCB	1.54	-	-
n-InP	3.18	1×10^{18}	-1
SCH	3.38	2×10^{15}	-12
QW	3.5	2×10^{15}	1000
p-InP	3.18	graded	graded
		0.2×10^{18}	-5
		to 2×10^{18}	to -50
p-InGaAs	3.4	-	-7000
Ti	3.7	-	-364830
Au	0.56	-	-797827

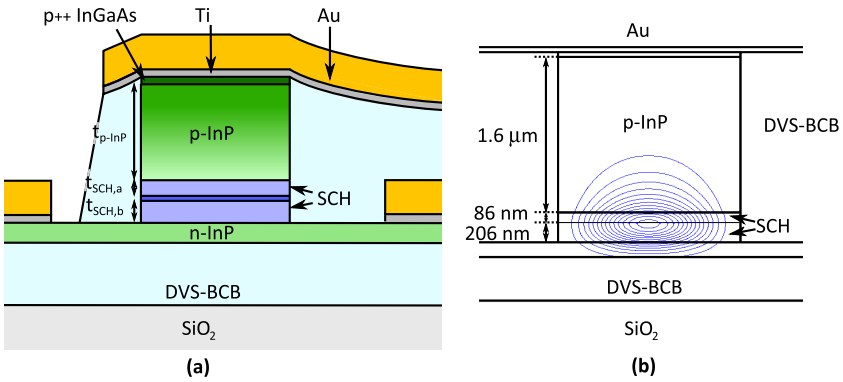


Figure 2.1: (a) Basic layout of the classical waveguide structure used on the heterogeneously integrated III-V/SOI platform. (b) Mode profile in the optimized waveguide.

and this value is therefore used for the optimization of the waveguide structures.

The values for all fixed material parameters are given in table 2.2. The waveguide structures are then optimized optically for the remaining parameters using a script that performs mode profile calculations in Fimmwave. In this simulation the mode profile is calculated for each combination of possible parameters, and the confinement in the QW, losses due to metalization and FCA, and the net modal gain are extracted.

Classical waveguide structure

A schematic of the classical bonded p-i-n waveguide can be seen in Fig. 2.1. It is a rib waveguide, which is etched through the active layers to expose the bottom n-InP layer. The n-type contact is then made on the sides, while the p-contact is made on the thicker center part of the waveguide. The higher index active region is positioned as close as possible to the interface with the DVS-BCB layer, as this brings the fundamental mode to the bottom part of the waveguide. This is done to separate the mode as far as possible from the strongly absorbing top contacting layers. A second advantage of this configuration, is that the interaction with underlying SOI structures is made stronger. This simplifies the coupling of the fundamental TE-mode to an underlying silicon waveguide and in case of a DFB laser allows for an adequate interaction with a distributed bragg reflector fabricated in the silicon

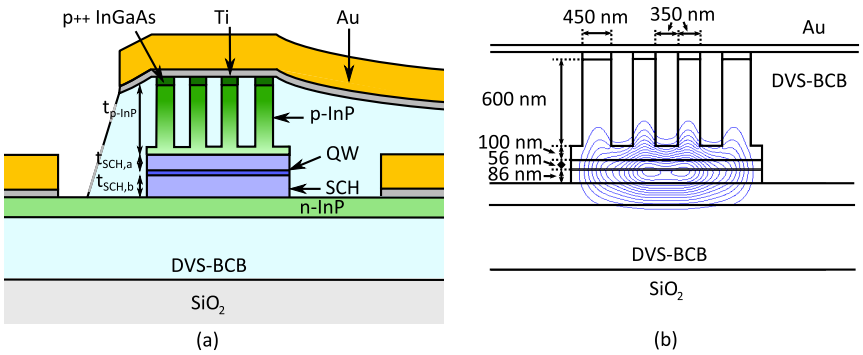


Figure 2.2: (a) Basic layout of the gallery waveguide structure used on the heterogeneously integrated III-V/SOI platform. (b) Mode profile in the optimized waveguide.

waveguide layer [19]. To avoid excessive losses in the top contact layers, the p-InP layer is chosen thick enough and the active layers are surrounded by high index SCH layers.

The parameters used for the optimization procedure are the total SCH layer thickness, the position of the QW in this SCH layer, and the top p-InP thickness. It is not straightforward to optimize this last parameter, as a thicker p-InP layer always seems better from an optical perspective. However, with increasing p-InP thickness the resistance of the device increases as well, and therefore a true optimization would require a fully coupled electro-optical optimization procedure. As this is quite tedious, we choose to add a criterion for the determination of the p-InP thickness: the optimal value is the thinnest p-InP that yields a loss due to the p-contact of $< 2\,/\mathrm{cm}$. The resulting optimized waveguide structure and mode profile can be seen in figure 2.1(b). In this waveguide structure a confinement of $1.8\,\%$ is achieved in the QW.

Gallery waveguide structure

The basic layout of a gallery waveguide structure is shown in figure 2.2. The waveguide has a similar structure as the classical design. The difference is that narrow trenches are etched in the top contacting layers, which are afterwards planarized by spincoating DVS-BCB, and etching it back to the p-type contact layer. If the widths of the trenches and pillars are sufficiently smaller than the wavelength of light, the top cladding behaves as a uniform material with an index of less than 2.5. Therefore, a substantially higher index contrast is achieved between the

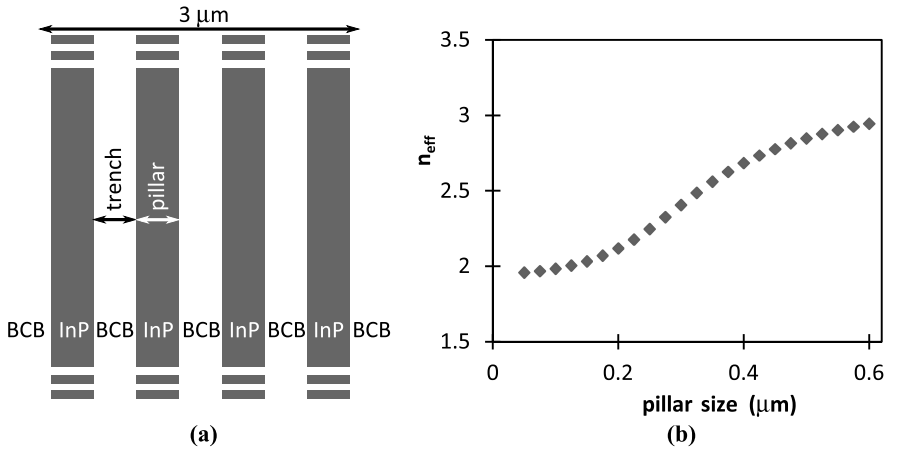


Figure 2.3: (a) Slab structure used to assess the effective index of the gallery layer. The number of pillars is always the maximal number that fits within $3\,\mu\mathrm{m}$. (b) Effective index as a function of pillar width. The associated trenches have the same width as the pillars.

top cladding and the waveguide core. Note that a $100\,\mathrm{nm}$ thick layer of p-InP is kept above the waveguide, as this facilitates current spreading.

Two very important parameters are the width of the trenches and pillars in the gallery layer. To assess the effect of these, we first calculate the effective index of a $3\,\mu\mathrm{m}$ thick slab waveguide of alternating InP and BCB slabs of the same thickness. In figure 2.3(a) the waveguide that is being simulated is schematically drawn, and in figure 2.3(b) the effective index of the slab TM-mode, which is the mode corresponding to the TE-mode in the gallery waveguide, is shown. In the best case, the trenches and pillars are infinitesimally small, as in this case the slab mode has a maximal overlap with the lower index DVS-BCB layer, leading to the smallest effective index for the gallery layer. For an increasing pillar and trench width however, the mode stays more and more confined in the high index pillars, leading to an increasing effective index. However, although the narrowest trenches and pillars lead to the lowest effective index, there are fabrication limits that need to be taken into account. It would actually be beneficial to find the largest dimensions for which the gallery concept still works well, so that it can be fabricated more easily. More specifically, it would be useful if the structure can be fabricated using an i-line projection lithography stepper, as these are much cheaper than deep-UV or e-beam lithography systems. In [30], the resolution limits of an i-line stepper are explored. It is found

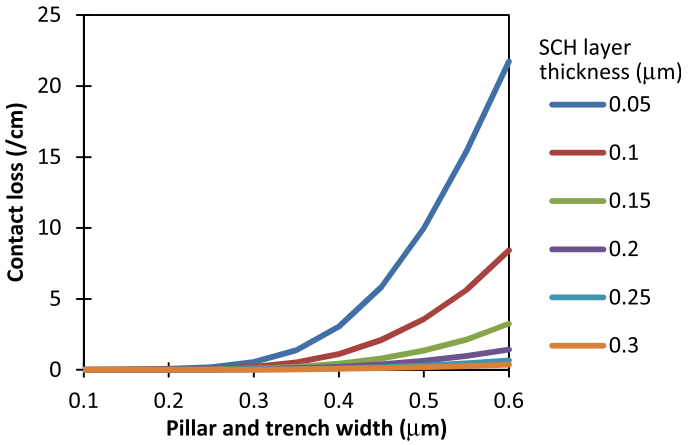


Figure 2.4: Losses due to absorption in the p-contact as a function of pillar and trench width, and SCH layer thickness.

that a dense line pattern with lines and gaps of $250\,\mathrm{nm}$ broad are still obtainable. To explore the effect of changing pillar and trench widths, a simulation was executed on a waveguide with dimensions as in table 2.1, and top p-InP thickness of $700\,\mathrm{nm}$ with $600\,\mathrm{nm}$ deep trenches. In figure 2.4 the loss in the waveguide due to metalization and doping is plotted, and in figure 2.5 the confinement in the QW is plotted, both as a function of pillar and trench width and for different SCH layer thicknesses. These results indeed confirm that a smaller trench and pillar width leads to the best results: as the top gallery layer confines the mode very well in this case, a thick SCH layer is not required in order to confine the mode to the active region. Therefore, the SCH layer thickness can be chosen in such a way that the confinement in the QW is maximized. However, even for a pillar and trench width of less than $350\,\mathrm{nm}$ the SCH layer still does not have to be very thick to achieve a low p-contact loss. For a width of $350\,\mathrm{nm}$, we have 4 pillars in total. As these dimensions are well within the limits of what is possible with i-line lithography, we select this width for the further optimization of the structure.

The parameters used for the further optimization procedure are the total SCH layer thickness, the position of the QW in this SCH layer, and the top p-InP thickness. We use the same criteria as for the classical waveguide structure. In a final optimization step, the outer two pillars of the gallery layer are broadened as much as possible to decrease the contact resistance without increasing the optical loss. The resulting

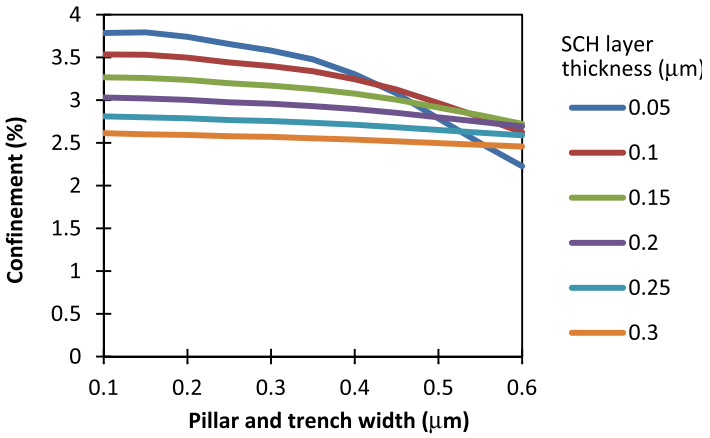


Figure 2.5: Confinement in the QW as a function of pillar and trench width, and SCH layer thickness.

optimized waveguide structure and mode profile can be seen in figure 2.2(b). In this optimized structure, the confinement in the QW is 3.1%, which is 70% higher than in the classical design. Furthermore, the top p-InP layer thickness is reduced to only $700\,\mathrm{nm}$ thick, as the mode does not extent as far from the active region as in the classical waveguide structure.

2.2.2 Electro-optical simulation

Methodology

The optimized waveguides are then simulated using ATLAS (TM) [28], a commercial package capable of a full thermal, electrical and optical simulation of semiconductor devices. This package uses analytical models for each physical effect, which can be activated separately and for which the details can be found in the manual [31]. Furthermore, the package contains an extensive library of material parameters. An issue with ATLAS however, is that it only contains a package to calculate lasing devices, and not optical amplifiers. However, this problem is easily circumvented by choosing a low reflectivity at both laser facets ($R = 1\,\%$) as this prevents lasing to occur and therefore only optical amplification is simulated. In the simulation, first the Schrödinger equation is solved for the QWs, and the band diagram of the device structure is calculated (see figure 2.6). After this, an electrical bias is ramped and the effects on the carrier injection, thermal behaviour and optical

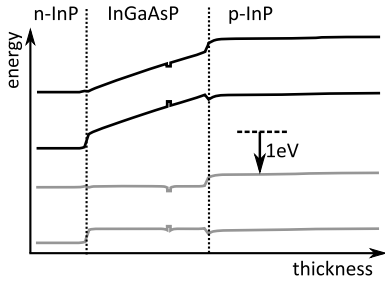


Figure 2.6: Band diagram at $0\,\mathrm{V}$ (top) and $1\,\mathrm{V}$ forward bias (bottom).

properties are simulated. To realistically assess the performance of the waveguide structures, the following effects are taken into account.

Electrically, SRH recombination (appendix A.4.2), Auger recombination (appendix A.4.2), surface recombination (appendix A.4.2) and radiative recombination terms (appendix A.4.1) are considered. To model carrier mobility saturation effects, the carrier mobility was modeled to depend on the local electric field, with a saturation velocity. Fermi-Dirac statistics are everywhere used. To simplify the simulation, the p-type and n-type semiconductor-metal interfaces are not explicitly simulated, but are modeled as Ohmic contacts instead.

Optically, losses due to material absorption, sidewall scattering and FCA (appendix A.6.3) are taken into account. The sidewall scattering loss was estimated using the volume current method, which models the sidewall roughness as an equivalent current distribution [32]. The fields generated by this current distribution can then be used to calculate the scattering loss. In this calculation an approximation is made to allow an analytical treatment of the problem: the medium surrounding the current distribution is considered homogeneous, which allows the use of the magnetic vector potential to derive an analytical expression for the scattering loss [33]. Although this estimation is not exact, the results obtained for the TE-mode in a $500\,\mathrm{nm} \times 200\,\mathrm{nm}$ and a $300\,\mathrm{nm} \times 300\,\mathrm{nm}$ high index contrast silicon wire match with an error margin of respectively only $4\,\%$ [34] and $20\,\%$ [35] with the experimentally measured loss. In the simulation an rms sidewall roughness of $10\,\mathrm{nm}$ and rms surface roughness of $3\,\mathrm{nm}$ with correlation length of $50\,\mathrm{nm}$ was assumed, both prudent estimations for InP waveguides etched by inductively coupled plasma etching [36, 37]. Both spontaneous emission and stimulated emission are considered, where the gain is calculated using the QW model developed in [38] which only consid-

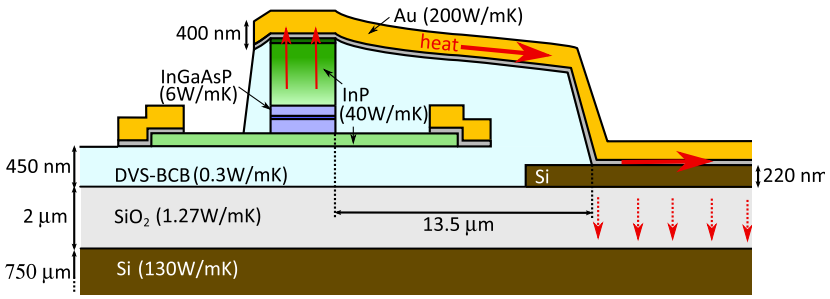


Figure 2.7: Thermal design used in the simulation.

ers a single valence band. Note that this calculated gain is everywhere the small signal gain for which no saturation effects have occurred. This optical calculation is performed for a broad wavelength range, but in the following only the achieved peak gain will be considered.

The device temperature is calculated using a simple model in which the thermal conductivities of all materials are assumed constant. Both joule heating and heating due to recombination of carriers are considered. For both waveguide designs it is assumed that the back of the SOI sample has a constant temperature of $20\,^{\circ}\mathrm{C}$ and that furthermore the bonded III-V waveguides use the top contacting layer as a thermal shunt to avoid excessive heating because of the low thermal conductivity of the DVS-BCB bonding layer, as was proposed in [39]. In figure 2.7 the device layout for the thermal simulation is shown, and in table 2.3 the thermal conductivities used for the different materials in the simulation are given. The resulting temperature is fed back to all electrical and optical models which are temperature dependent. This simulation is performed for a device of $200\,\mu\mathrm{m}$ long, using both the optimized classical and gallery waveguide structures.

Electrical performance

In figure 2.8(a) the current is plotted as a function of applied voltage. The gallery design has a differential resistance of $6.6\,\Omega$ for a $200\,\mu\mathrm{m}$ long device, 50% higher than the classical waveguide structure. This increase is caused by the fact that the top contact in the gallery design is only $1.5\,\mu\mathrm{m}$ wide, compared to $3\,\mu\mathrm{m}$ in the classical design. On the other hand, the gallery waveguide structure has a thinner p-InP layer, which compensates partly for the increased resistance.

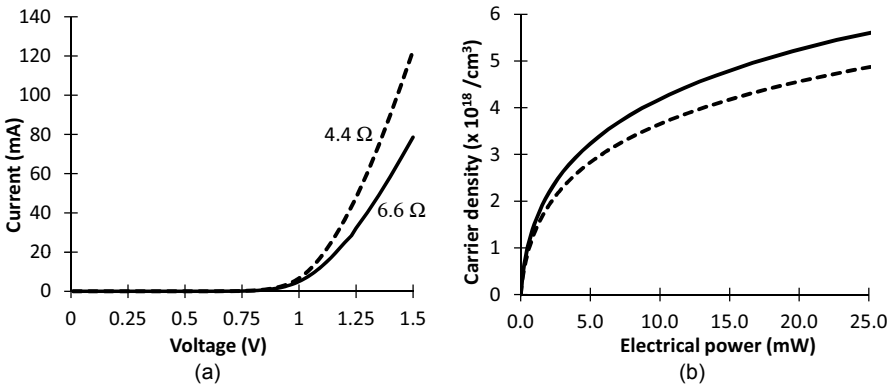


Figure 2.8: I(V) curve for the classical (dashed line) and gallery (solid line) design. The gallery has a higher series resistance, because of the reduced top contact area. (b) Achieved carrier concentration in the QW as a function of electrical power for the classical (dashed line) and gallery (solid line) design.

Table 2.3: Device parameters.

Thermal conductivity	Value	Unit
κ_{Si}	130	$\mathrm{Wm^{-1}K^{-1}}$
κ_{SiO_2}	1.27	$\mathrm{Wm^{-1}K^{-1}}$
κ_{BCB}	0.3	$\mathrm{Wm^{-1}K^{-1}}$
κ_{InP}	40	$\mathrm{Wm^{-1}K^{-1}}$
$\kappa_{InGaAsP}$	6	$\mathrm{Wm^{-1}K^{-1}}$
κ_{Au}	200	$\mathrm{Wm^{-1}K^{-1}}$
Electrical properties	Value	Unit
$\rho_{c,n-type}$	10^{-6}	$\Omega/\mathrm{cm^2}$
$\rho_{c,p-type}$	10^{-5}	$\Omega/\mathrm{cm^2}$
τ_{p0}	20	ns
τ_{n0}	1	ns
B_{InP}	1.5×10^{-10}	$\mathrm{cm^3s^{-1}}$
$B_{InGaAsP}$	1.5×10^{-10}	$\mathrm{cm^3s^{-1}}$
C_{InP}	9×10^{-31}	$\mathrm{cm^6/s}$
$C_{0,InGaAsP}$	3.5×10^{-30}	$\mathrm{cm^6/s}$
$K_{InGaAsP}$	3.65	-
$v_{s,InP}$	8×10^3	cm/s
$v_{s,InGaAsP}$	1×10^5	cm/s

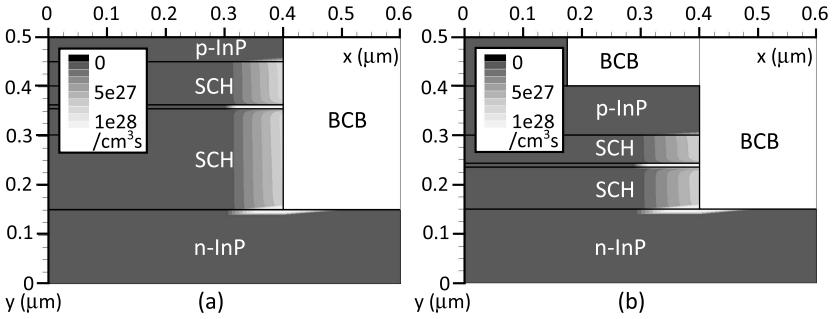


Figure 2.9: SRH and surface recombination rate ($\mathrm{cm}^{-3}\mathrm{s}^{-1}$) in the (a) classical design and (b) the gallery design.

In figure 2.8(b) the achieved carrier density in the QW is shown as a function of applied electrical power. It seems that although the gallery structure has a higher series resistance, a higher carrier density is achieved in the QW than in the classical design. In figure 2.9 the SRH and surface recombination are plotted. As can be seen the rate of recombination is locally nearly the same in both waveguide structures, however because the InGaAsP SCH layers are thicker in the classical waveguide, the total integrated recombination rate is much larger in the classical waveguide. The reason for this is that the InGaAsP SCH layers have a higher surface recombination rate ($1 \times 10^5 \mathrm{cm/s}$) than InP ($8 \times 10^3 \mathrm{cm/s}$). The consequence is that the carrier lifetime will be reduced in the classical design, leading to the lower carrier concentration in the classical waveguide even though the device resistance is lower.

Optical performance

To assess the optical performance of the device, the net small signal peak gain as a function of the carrier density is plotted in figure 2.10(a) for a device of $200\ \mu\mathrm{m}$ long. The net gain in the gallery design has more than doubled compared to the classical design. This improvement can be attributed to the higher optical confinement in the QW, combined with a small reduction in FCA. Note that the apparent threshold in the curve is related to the fact that when the device has not reached transparency yet: the peak 'gain' corresponds there to the loss for light with a photon energy below the QW band gap, caused by scattering and FCA. In figure 2.10(b) the FCA profile in the gallery waveguide is shown. From this it can be seen that the main contributions to FCA are due to the high concentration of holes in the top p-InP layer and in the

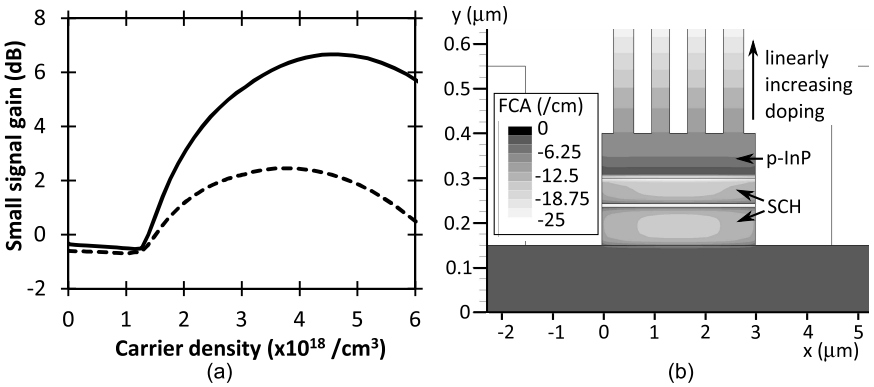


Figure 2.10: (a) Achieved small signal net gain as a function of the carrier concentration in the QW for the classical (dashed line) and gallery (solid line) design with a device length of $200\,\mu\mathrm{m}$ long. The improved design of the gallery waveguide structure leads to a net gain that is two times higher than in the classical structure. (b) FCA in the gallery waveguide structure for an applied voltage of 1 V.

SCH layers. In the gallery waveguide, the mode tail is not extending as far in the top p-InP layer as in the classical waveguide. This leads to a lower FCA contribution from this layer, as the confinement in this p-InP layer in gallery structure is only 26%, compared to 40% in the classical structure. Furthermore, the confinement in the SCH layers is the same in both the classical and gallery waveguide, leading to a net reduction of FCA.

For higher carrier densities, the gain is saturating as a result of thermal roll-over. This is because the number of free states in the QW is decreasing. This leads to a higher carrier density in the SCH layers, which in turn increases the FCA and through non-radiative recombination leads to a higher heat dissipation. This results in an increased device temperature, leading to both a lower QW gain due to a smoothing of the Fermi-Dirac carrier distribution, and an increase in the efficiency of non-radiative recombination processes, especially Auger recombination (see section A.4.2). Eventually the increased losses and degraded QW gain lead to a decline in the gain curve.

When comparing the gallery with the classical design however, it seems that this thermal roll-over occurs sooner in the classical design than in the gallery design. Both waveguide structures have a similar thermal resistance of $4.5\,\mathrm{Kcm^2kW^{-1}}$, so the thermal design is not the cause for this difference. The cause seems to lie with the higher

non-radiative recombination as described in section 2.2.2: because of a higher intrinsic surface recombination in the classical design, the temperature rises more quickly, degrading the performance faster than in the gallery design.

A drawback of the gallery design is the higher scattering loss: when considering only the scattering due to the sidewalls a 32-fold increase of the loss is observed, from a simulated $0.19\,\mathrm{dB/cm}$ to $6\,\mathrm{dB/cm}$. Additionally, the vicinity of the bottom of the etched trenches to the optical mode adds an excess loss of $1.3\,\mathrm{dB/cm}$, leading to a total scattering loss of $7.3\,\mathrm{dB/cm}$. However, in a $200\,\mu\mathrm{m}$ long device, this only leads to an excess loss of $0.15\,\mathrm{dB}$. Compared to the gain improvement, this has only a marginal impact on the device performance, and this higher loss is therefore not visible in the net gain curves in figure 2.10.

2.2.3 Conclusion

An electrically pumped amplifier with high gain was designed and simulated. By using the high index contrast between the bonded III-V film and the surrounding DVS-BCB and oxide cladding layers, a waveguide was designed that maximizes the confinement in the active region and at the same time keeps the mode away from the lossy contact layers, in this way doubling the net gain that can be achieved compared to the classical waveguide structure. This type of amplifier could be used for the realization of a low threshold microlaser.

2.3 Integrated QCSE modulator

Although several types of modulators have already been demonstrated on the silicon-on-insulator platform, the performance of these devices is no match to the intrinsically higher performance of III-V based electro-absorption modulators (EAMs). Electro-absorption modulation in III-Vs is most commonly based on the quantum confined stark effect (QCSE) in a MQW active region. This physical effect allows the realization of high speed, high extinction ratio modulators. Attempts have already been made to use the same effect in Germanium based QWs, as these can be directly grown on silicon. Good results have been obtained, however the fabrication technology is not mature yet in comparison with III-V EAM technology. Therefore, it is still advantageous to bond III-Vs on the SOI platform to create III-V EAMs integrated on SOI.

However, instead of just copying an existing modulator design used in III-V technology, the design can be improved by leveraging the high index contrast between the bonded III-V and surrounding DVS-BCB and SiO_2 cladding. As was shown in section 2.2, a very high confinement can be achieved in bonded III-V waveguides. In the case of a modulator, this high confinement can be used to decrease the size of the modulator, which leads to a direct reduction of the device capacitance. Therefore, modulators based on bonded III-V waveguides, should be able to obtain higher speeds and/or lower operating voltages than existing pure III-V solutions. In this work, the aim is to design a modulator that is driven with an as low as possible operating voltage, but that can still achieve a bandwidth of at least $10\,\mathrm{GHz}$ when driven by simple lumped electrodes.

2.3.1 The state of the art

Silicon modulators

A first class of modulators achieved on the SOI platform, are the pure silicon modulators. As silicon is transparent in the telecom wavelength range, all pure silicon modulators operate through the application of phase shifts in a resonator or interferometric device. This phase shift can be achieved by the application of the plasma dispersion effect, which describes the dependence of the material refractive index on the free carrier density. This modulation of the free carrier density can be achieved in several ways, however carrier depletion is the most successful method so far, due to the high speed that can be obtained. In carrier depletion modulators, the modulation of the applied voltage changes the size of the depletion region of a p-n junction inside the waveguide. In this way, the free carrier density inside the waveguide is changed, resulting in a local change of the refractive index.

In [40], a silicon ring modulator is demonstrated based on the carrier depletion effect. In this work, shifting of a resonance of a ring resonator is used to achieve modulation. Using this modulator, a bit rate of $25\,\mathrm{Gbit/s}$ was achieved using a peak-to-peak voltage of $1\,\mathrm{V}$ resulting in an ER of $> 5\,\mathrm{dB}$. A big problem of resonator based modulators however, is that they need to be thermally tuned to work properly. In this work the modulation energy required was only $7\,\mathrm{fJ/bit}$, however the thermal tuning leads to an additional power consumption of $2.6\,\mathrm{pJ/bit}$ for a full FSR tuning range. As the heater was in this work already integrated inside the ring resonator, the only option to further reduce this

power consumption is either by increasing the speed, or by increasing thermal insulation. In [41] a very high thermal tuning efficiency was obtained after underetching of the device, with a tuning power of only $2.4\,\mathrm{mW}$ for tuning over the complete FSR, which is equivalent to a tuning energy of $96\,\mathrm{fJ/bit}$ at a bit rate of $25\,\mathrm{Gbit/s}$. A second challenge in resonator based devices is that higher bandwidths are difficult to obtain, as the modulation speed then also becomes limited by the photon lifetime in the resonators. In [42], a bit rate of $50\,\mathrm{Gbit/s}$ was obtained by driving the ring modulator using pre-emphasis to compensate for the reduced modulation efficiency.

Alternatively, a Mach-Zehnder interferometer can be used, where in one or both of the arms of the interferometer the phase is changed to achieve modulation. In [43], a modulator with $V_\pi L$ product of $1.4\,\mathrm{Vcm}$ was achieved. With a $1\,\mathrm{mm}$ phase shifter, the device operates at a modulation speed of $12\,\mathrm{GHz}$ and bias of $6\,\mathrm{V}$. The main issue with this kind of device, is the large footprint required to create a device with a low driving voltage and high ER. Related to this high footprint, also the speed when driven with lumped electrodes is limited, as the device capacitance is usually larger than for a small ring resonator. The advantage however is that thermal tuning is usually not necessary, as these devices are sufficiently broadband. Furthermore, the applied bias voltage can also be used to tune the modulator to the right operating wavelength.

Germanium based modulators

Germanium is one of the materials that can be integrated monolithically on silicon. It has the interesting feature that it shows absorption in the telecom wavelength range, and can therefore be used to create electro-absorption modulators. The main challenge here however is that although growth is possible, there is still a large mismatch in lattice constant between germanium and silicon, which complicates the fabrication of high-quality devices. This is usually resolved by using a buffer layer as a virtual substrate to relax the strain and capture threading dislocations, however this poses a limitation on the waveguide structures that can be achieved and still does not remove the strain problem completely.

In [44], electro-absorption modulation is achieved by using the Franz-Keldysh effect in a germanium waveguide integrated on a $3\,\mu\mathrm{m}$ silicon waveguide platform. To apply a field across the waveguide,

a horizontal p-i-n diode structure is created in the germanium waveguide. Using this device, an ER of 4 to $7.5\,\mathrm{dB}$ is achieved over a wavelength range of $1610 - 1640\,\mathrm{nm}$, with an electrical bandwidth of $30\,\mathrm{GHz}$. The power consumption of this device however is still relatively high: a reverse bias of $-4\,\mathrm{V}$ with a peak-to-peak voltage of $4\,\mathrm{V}$ is required to drive the modulator at this high speed. In the paper, the required modulation energy consumed by the device is calculated to be $100\,\mathrm{fJ/bit}$ at $25\,\mathrm{Gbit/s}$, however this does not take into account the energy consumption caused by the large required bias voltage in combination with the generated photocurrent (see section 2.3.4). Assuming a $7\,\mathrm{dB}$ ER at $25\,\mathrm{Gbit/s}$ and a $5\,\mathrm{dBm}$ laser input power at a wavelength of $1620\,\mathrm{nm}$, calculation shows that the high bias voltage leads to an additional $260\,\mathrm{fJ/bit}$. Only the optical insertion losses are neglected in this case. Therefore, the actual total energy consumption of the device is probably closer to $360\,\mathrm{fJ/bit}$ at $25\,\mathrm{Gbit/s}$.

Next to devices using a bulk germanium layer, also devices with germanium QWs have been proposed. In this case QCSE, which is the stronger counterpart of the Franz-Keldysh effect in QWs, is used to achieve modulation. In [45], an ER of $9\,\mathrm{dB}$ is achieved using a $90\,\mu\mathrm{m}$ long waveguide with Ge/SiGe MQW active region, with an electrical bandwidth of $23\,\mathrm{GHz}$. These results were obtained using a reverse bias voltage of $-3.5\,\mathrm{V}$, with a peak-to-peak driving voltage of only $1\,\mathrm{V}$, resulting in a dynamic power consumption of $107\,\mathrm{fJ/bit}$. Although the driving peak-to-peak voltage is much lower than in [44], the same dynamic power consumption is achieved due to the fact that the capacitance of this device is much higher. Also in this case an additional contribution due to the bias should be taken into account, which will be of the same order of [44], leading to a total energy consumption of approximately $360\,\mathrm{fJ/bit}$ at $25\,\mathrm{Gbit/s}$.

Heterogeneously integrated III-V modulators

Although electro-absorption modulators based on SiGe MQWs work, they do not perform as well as modulators based on InP. This is related to the indirect bandgap of germanium. First of all, the absorption tail is longer, which causes unwanted absorption, and therefore increases the modulator insertion loss. Next to this, the absorption is lower than in InP QWs, so more QWs are required to achieve the same effect.

Therefore, a third approach to realize modulators on SOI is by heterogeneously integrating an InP epitaxial layer stack on top of

a SOI waveguide circuit, and post-processing this layer stack in a III-V fab. Although this is not the easiest solution as it brings a lot of post-processing steps, the intrinsic high performance of InP MQWs allows the fabrication of state-of-the-art modulators on silicon. In [3], a bandwidth of over $67\,\mathrm{GHz}$ was demonstrated by using traveling wave electrodes to drive the modulator. Furthermore, at $50\,\mathrm{Gbit/s}$ an ER of $9.6\,\mathrm{dB}$ was obtained using a peak-to-peak drive voltage of $2.2\,\mathrm{V}$ and a reverse bias of $-4\,\mathrm{V}$. Next to designs optimized for a maximal bandwidth, also efforts are made to reduce in the power consumption. In [21], a modulator with $30\,\mathrm{GHz}$ bandwidth is demonstrated which requires a peak-to-peak drive voltage of only $1\,\mathrm{V}$, and does not require any cooling. In this way, a dynamic power consumption of only $20\,\mathrm{fJ/bit}$ is obtained. However, because the bias voltage is of the order of $-3\,\mathrm{V}$ in this demonstration, the total power consumption is still $112.5\,\mathrm{fJ/bit}$ with an input laser power of $5\,\mathrm{dBm}$ and bit rate of $40\,\mathrm{Gbit/s}$.

Design guidelines

In order to be able to compete with electrical interconnects on short distances, the total energy budget should be limited to $1\,\mathrm{pJ/bit}$ [7] and preferably be even lower for the shortest on-chip interconnects [8]. This includes all power: light generation, modulation, detection, driving circuitry and any cooling required to allow the devices to run. Therefore, all components should operate with the lowest power consumption possible. For the design of the modulator, this means that

- The optical insertion loss should be minimized, as this directly leads to a reduction of the required laser output power. Therefore, this also leads to a reduction of the generated photocurrent, which also has an impact on the device power consumption (see section 2.3.4).

- The peak-to-peak voltage swing required to achieve modulation should be minimized in order to reduce the power consumption. Furthermore, it would be an advantage to be able to drive the modulators using a simple CMOS circuit. If the modulator is working with a peak-to-peak driving voltage of less than $1\,\mathrm{V}$, a simple inverter can be used to drive the modulator [46]. To take into account the offset between a simulation result and an actual device with all its imperfections due to the fabrication process,

the aim is to design a modulator that can be driven with a driving voltage of less than $0.5\,\mathrm{V}$.

- The bias voltage should be low, as a higher bias also leads to an additional power consumption because of the generated photocurrent.
- The bandwidth should be large enough, as a higher bit rate also reduces the energy required per bit.
- The ER should be high enough to avoid a large power penalty at the receiver. For an ER of more than $9.5\,\mathrm{dB}$, the power penalty at the receiver is reduced to less than $1\,\mathrm{dB}$.

2.3.2 The QCSE and its use for optical modulation

The QCSE describes the shifting of the absorption edge of a set of QWs towards longer wavelengths under influence of an electric field applied perpendicularly to the QWs. This shift is also called the stark shift. The reason this happens can be understood by considering the band diagram of a set of QWs under a high electric field as depicted in figure 2.11(a). When applying a large electric field the conduction band and valence band are warped in such a way that the resulting hole and electron energy levels in the QWs are brought closer together. This causes the red shift of the absorption edge of the QWs. The resulting absorption curve is not a purely red shifted version of the original absorption curve however. From Fermi's golden rule (appendix A.6.2) it follows that the absorption is proportional with the overlap of the hole and electron wave functions. As these wave functions are shifted away from each other because of the presence of the electric field, the absorption in the QWs will be reduced while it is shifted to longer wavelengths. The changing absorption is schematically shown in figure 2.11(b).

The energy shift due to the QCSE can be calculated analytically by considering the QWs as infinite potential wells. Then, the following relation for the change of the QW band gap is found [47]

$$\Delta E_g \propto W^2 E^2 \tag{2.1}$$

Two important conclusions follow from this: first of all, the band gap shift is proportional to E^2. Therefore, in order to avoid excessively high voltages, the structure of the modulator should be adapted such that a small voltage results in a large field over the QWs. This can be

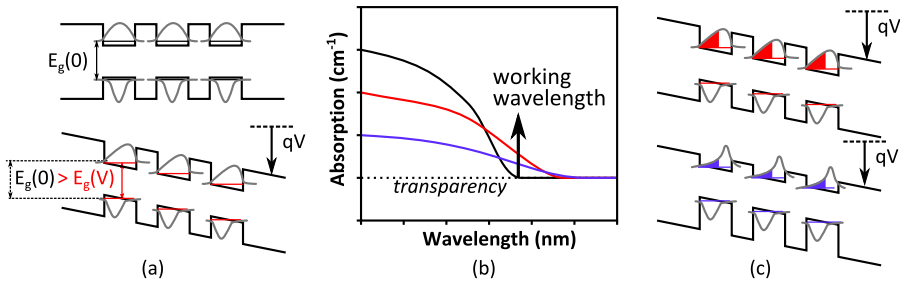


Figure 2.11: Principle of QCSE. (a) The reduction of the band gap and the shifting of the wave functions (grey) when applying a high electric field. (b) The absorption profile for the original case (black), and for an applied electric field for a large (red) and small (blue) conduction band discontinuity. (c) The band diagram under electric field for large and small conduction band discontinuity. The shaded area suggests the amount of overlap between the hole and electron wave functions.

done by reducing the thickness of the undoped active region t_{act}, as the field is proportional to V/t_{act}. However, this comes at the cost of a higher capacitance, as the capacitance scales with $1/t_{act}$ as well. Secondly, the shift is proportional to W^2. This means that broader QWs will show larger shifts. However, also in this case there is a trade-off. For broader QWs, the electron and hole wave functions will also shift more, and the overlap of these wave functions will therefore decrease more drastically, leading to a lower absorption. The result is that even though the absorption shifts faster to longer wavelengths for QWs that are too broad, the drop in the absorption value more than compensates this advantage. Furthermore, broader QWs have a worse performance, as these behave more like bulk material compared to narrower QWs. Hence, the QW width is an important parameter that needs to be optimized.

In reality the QWs are not infinitely deep potential wells, but have a certain barrier height. An important characteristic of a well/barrier material combination, is the valence and conduction band offset between these two materials, as these determine the barrier height experienced by the QW. It is important to note here that due to the generally higher effective mass of holes compared to conduction band electrons, the holes are better confined in a QW than the electrons. Furthermore, the widely used InGaAsP system for telecom wavelengths shows a higher valence band offset than conduction band offset ($\Delta E_c = 0.39\Delta E_g$). The result is that the electron wave function will shift much more rapidly

than the hole wave function with increasing electric field, leading to a faster reduction of the absorption. However, this is not the case in all well/barrier material combinations: in strained InGaAlAs MQW structures a conduction band offset of $0.72\Delta E_g$ is obtained [48]. Therefore larger electric fields can be applied before the overlap of hole and electron wave functions drastically decreases. In [49] a comparison was made, resulting in a 30% improvement of the extinction ratio when using InGaAlAs QWs. An artists impression of the band diagram under applied electric field is shown in figure 2.11(c) for a large and a small conduction band offset, with their corresponding absorption curves shown in figure 2.11(b).

The QCSE can be harnessed in several ways to achieve optical modulation. One way is the use of the fact that absorption is reduced when a strong electric field is applied because of the shifting of the electron and hole wave functions [50]. However, most commonly the shifting of the band edge is used [49, 3, 21]. In this approach, the signal has a photon energy that is slightly lower than the QW band gap. Therefore, at zero bias the absorption is very low. When a voltage is applied however, a strong electric field in the QWs reduces the effective band gap below the photon energy of the signal. Therefore, the signal is now strongly absorbed by band-to-band absorption and modulation is achieved.

2.3.3 Bandwidth considerations in QCSE modulators

The bandwidth of QCSE modulators is usually limited by the speed at which the voltage can be applied across the device's MQW structure. When the device is driven as a lumped component, the device bandwidth will be determined by the bandwidth of the RC network formed by the device capacitance C and resistance R, and the load of the source (usually $50\,\Omega$). This then leads to the limitation

$$\Delta f_{RC} = 1/(2\pi(R + 50\,\Omega)C) \tag{2.2}$$

Usually, this forms a hard limitation for the bandwidth of QCSE modulators. However, even if the resistance and capacitance can be sufficiently reduced, a lumped configuration will always be limited in speed, because of the transit time of the optical mode through the waveguide. If the frequency of modulation is so high that the voltage is changed before the optical wave with previously applied voltage has left the device, modulation will not be possible. For a $400\,\mu\mathrm{m}$ long device, the transit time τ_{tr} is of the order of $5\,\mathrm{ps}$, limiting the bandwidth to $\Delta f_{tr} = 1/(2\pi\tau_{tr}) = 30\,\mathrm{GHz}$.

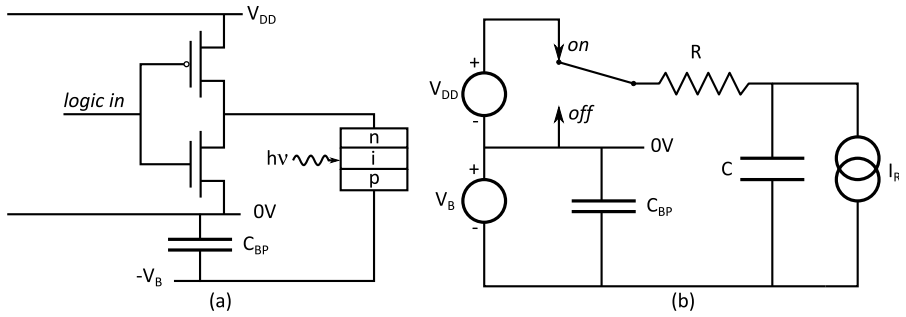


Figure 2.12: Based on [8] (a) Driving circuit for QCSE modulator. (b) Equivalent circuit.

This limitation can be lifted however by using a traveling wave configuration. In this electrode configuration, the electrical wave travels along with the modulator to drive the device. This has several advantages. First of all, the resistance and capacitance of the device can be distributed, and therefore the RC-limitation on the bandwidth is lifted. Furthermore, if the phase velocity of the electrical signal is matched with the group velocity of the optical mode, the transit time limitation is also completely lifted. Although it would seem that an infinite bandwidth would be possible in this case, other limitations cap the maximal obtainable bandwidth. The most important limitation is formed by the electrical loss in the traveling wave electrodes, as this loss increases with increasing bandwidth. This is due to the skin effect, which describes that electrical current is increasingly confined to the surface of the electrodes. At frequencies above $10\,\mathrm{GHz}$, the skin depth where current still flows is already less than $1\,\mu\mathrm{m}$, and decreases further with higher frequencies.

2.3.4 Power consumption

There are three important contributions to the power consumption of a modulator. The discussion below is a summary of the results presented in [8]. A first contribution is formed by the optical power that is being absorbed during modulator operation. Even in an ideal modulator, at least half of the incoming light should be absorbed. However, in a realistic modulator there are usually additional insertion losses that increase the amount of lost optical power.

The second contribution comes from the electronic circuit (see figure 2.12). When the modulator changes from an on-state to an off-state, the

charge stored in the modulator capacitance changes. Therefore, an electric current that carries the energy stored in the modulator capacitance needs to run through the circuit. As the circuit has a certain resistance, the energy carried by this current is dissipated as heat. Assuming a NRZ-OOK modulation scheme, in half of the bit periods the capacitor will have to charge or discharge. The second loss contribution is therefore given by

$$\Delta E_{bit,c} = \frac{1}{4} C V_{DD}^2 \tag{2.3}$$

See [8] for a more thorough derivation. Here $E_{bit,c}$ is the energy lost per bit due to the capacitance, C is the modulator capacitance and V_{DD} is the voltage swing applied by the source on the modulator. Contrary to what one would expect, there is no contribution due to the DC bias, as any energy flowing into the circuit can be recuperated by adding a bypass capacitor C_{BP} to the source as shown in figure 2.12.

A third contribution is formed by the photocurrent that is generated by the absorption of the signal beam. For each photon absorbed in the device, an electron-hole pair is generated, which is quickly transported out of the active region because of the high electric field. Therefore, this electron-hole pair does not have the time to recombine before it is extracted. In this case, the bias voltage does have an important impact on the dissipated energy:

$$\Delta E_{bit,i}\left(V_{tot}\right) = E_{abs}\left(V_{tot}\right)\left[1 + \frac{V_{tot}}{\hbar\omega/q}\right] \tag{2.4}$$

where $E_{bit,i}$ is the energy lost per bit due to the photocurrent, E_{abs} is the energy absorbed per bit and V_{tot} is the total reverse bias with the convention that a positive value equals a reverse bias. From this equation, one can conclude that a large DC bias will lead to a large waste of energy. However, such DC bias is usually necessary to achieve a good performance in the modulation, as the shift due to a changing electric field scales with E^2 (see equation 2.1). This energy waste can be largely avoided however by using a thinner active region, as the field in the active region scales with V_{tot}/t_{act}. Furthermore, the built-in electric field in a thin ($< 200\,\mathrm{nm}$) active region will already reach a value of $> 4 \times 10^4\,\mathrm{V/m}$, further reducing the need for a DC bias. In [8], it is even suggested to use an active region with thickness of the order of $80\,\mathrm{nm}$, as the built-in electric field is then sufficiently large to provide for an adequate stark shift. Modulation can then be obtained by applying a small positive bias that counteracts the stark shift due the built-in

electric field. If this positive bias is small enough to avoid generating a meaningful current, the generated photocurrent in such a device would lead to an amount of generated energy instead of dissipated energy.

In summary, both the DC bias and the magnitude of the voltage swing applied over the modulator have to be minimized to keep the optical power consumption as low as possible. Both can be achieved by using a sufficiently thin intrinsically doped active region, as large electric fields can then be created using small voltages.

2.3.5 V-waveguide structure

Basic layout

The basic layout of the V-waveguide structure is shown in figure 2.13. It is a regular p-i-n waveguide structure, however the top p-InP layer is etched with a negative angle of $-10°$ to $-8°$, resulting in the V-shaped profile. Such an angle can be obtained by wet etching the p-InP layer with a HCl:H_2O solution, when a stripe pattern is oriented parallel to the (01-1) crystal plane. The (01-1) crystal plane is easily recognizable, as it is one of the directions along which a standard InP wafer with (100) orientation can be cleaved. The reason a range of angles is given rather than one exact value, is because this is based on actual measurements of the etch angle rather than a theoretical derivation from the etch process. As there is an uncertainty on both the alignment of the crystal with the etch pattern, and the accuracy of the measurement tool (dual beam FIB/SEM), this measured range of angles is used to have a realistic idea of what can be fabricated.

This V-shaped profile has several advantages. First of all, it allows us to realize a narrow p-InP base near the optical mode. Such a narrow base has a similar effect as the etched trenches in the top p-InP in the gallery waveguide structure: the effective index of the top cladding is reduced, increasing the confinement in the active region and keeping the mode away from the top contacting layers, as is shown in figure 2.14. However due to the increasing width towards the top contact, in this case this advantage does not come at the cost of a reduced top contact area. Furthermore, the capacitance of the V-waveguide can be much lower than in a classical waveguide structure, which will be discussed in the next section. Next to this, the fabrication of such a structure is quite straightforward. Because the top stripe width is of the order of $1\,\mu\mathrm{m}$, contact lithography can be used, while a very narrow p-InP base and narrow taper tips can be achieved using wet etching.

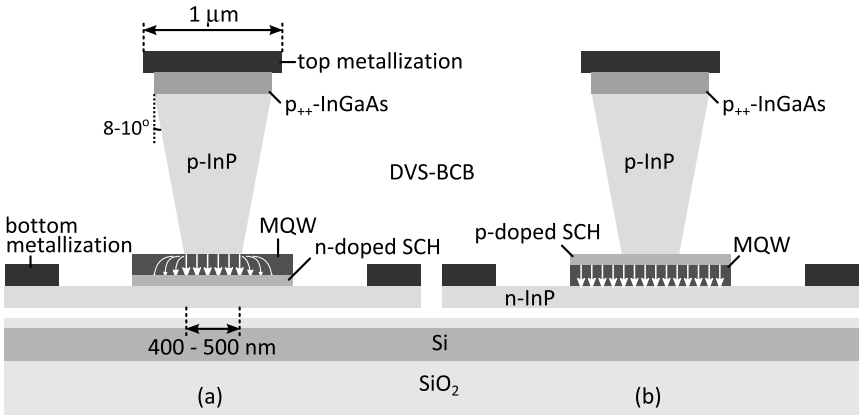


Figure 2.13: Basic layout of the V-waveguide structure optimized for modulator operation. (a) Without p-doped SCH layer, leading to a very low capacitance, but non-uniform electric field. Instead of a p-doped SCH layer, an n-doped SCH layer can be used in this case to achieve a certain active region thickness. (b) With p-doped SCH layer.

This type of waveguide structure has already been used in the group to achieve laser devices [51]. In this case, the V-shaped profile was mainly useful to achieve very narrow taper tips using simple contact lithography, leading to an adiabatic taper transition.

Capacitance and power consumption

There is an important design choice that has to be made. For the lowest capacitance per length unit possible, the design as shown in figure 2.13(a) is optimal. This can be understood by considering the p-i-n structure as a parallel plate capacitor, where the p-InP and n-InP regions form the two plates, and the active region with SCH layers forms the dielectric. As the top p-InP layer can be made very narrow, the resulting capacitance will also scale down. A drawback of this however is that only the part under the p-InP base will experience a uniform, high electric field. Therefore, only this part of the QW volume can be used for modulation. Alternatively, the design as shown in figure 2.13(b) is possible. In this design a thin part of the top SCH layer is p-doped to allow for current spreading and a uniform electric field across the complete active region, at the cost of a higher capacitance per length unit. In this case the full active region can be used for modulation. It is not obvious which of these two designs will lead to the lowest device ca-

pacitance, as a modulator based on the V-waveguide without p-doped SCH layer will have to be longer to achieve the same modulation depth than a modulator based on the waveguide with p-doped SCH layer.

Another parameter that has a direct influence on the device capacitance is the (undoped) active layer thickness (t_{act}). As discussed in section 2.3.2, this parameter represents a trade-off between the capacitance and the driving voltage of the device. Therefore, the optimal value for t_{act} is determined by the specifications of the envisioned application. In our case, we aim to design a more efficient modulator that can still operate at sufficiently high speeds (bandwidth at least $10\,\mathrm{GHz}$). To help determining the active layer thickness without having to completely simulate the device performance, published experimental results can be used to make this design choice. In [20], a lumped QCSE modulator integrated on SOI is achieved with a bandwidth of $10\,\mathrm{GHz}$, DC bias of $4\,\mathrm{V}$ and voltage swing of the order of $0.8\,\mathrm{V}$. In this design, the intrinsic active region was $287\,\mathrm{nm}$ thick, with a width of $3\,\mu\mathrm{m}$. Based on this, we set the intrinsic active region thickness to $180\,\mathrm{nm}$ as this will significantly reduce the DC bias as well as the required voltage swing to $0.5\,\mathrm{V}$. To achieve a sufficiently high confinement in such a thin active region, the base of the V-shaped p-InP pillar should be of the order of $500\,\mathrm{nm}$. This has the additional advantage that the capacitance can be reduced by approximately a factor of 6 compared to the waveguide in [20], if the top SCH layers remain undoped. In case the top SCH layer is p-doped however, the width of the active region determines the capacitance. If one uses the top metal mask as a mask for dry etching, a $1\,\mu\mathrm{m}$ broad active region can be relatively easily obtained, leading to a reduction of the capacitance with a factor of 3 compared to [20]. Both these designs should therefore allow to achieve a bandwidth of at least $15\,\mathrm{GHz}$, with a very low required modulation and DC bias voltage.

Optical simulation of the waveguide structure

Several parameters need to be determined. As an active region in this simulation, a MQW structure is used which consists of 7 compressively strained ($+0.41\,\%$) InGaAlAs QWs (Q1.55) of $11\,\mathrm{nm}$ thick, with tensile strained ($-0.4\,\%$) InGaAlAs (Q1.19) barriers of $7\,\mathrm{nm}$ thick in between. This stack is sandwiched by slightly tensile strained InGaAlAs SCH layers to achieve complete strain compensation in the stack. Three parameters in particular are important: the thickness of the p-InP pillar, the thickness of the p-InGaAs contacting layer, and the thickness of the

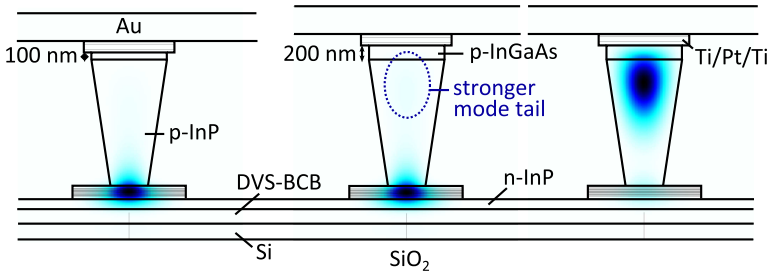


Figure 2.14: (left to right) Mode profile of fundamental TE mode in the unmatched case ($t_{InGaAs} = 100\,\mathrm{nm}$, $t_{SCH} = 32\,\mathrm{nm}$), fundamental TE mode in the matched case ($t_{InGaAs} = 200\,\mathrm{nm}$, $t_{SCH} = 32\,\mathrm{nm}$), and corresponding lossy mode with similar effective index.

SCH layers. In the optical simulations described below, an operation wavelength of $1.55\,\mu\mathrm{m}$ is assumed.

As already briefly discussed in the previous section, the base of the V-shaped p-InP pillar should have a width of approximately $500\,\mathrm{nm}$ to achieve a high confinement in an active region of only $150\,\mathrm{nm}$ thick. Therefore, assuming a wet etching angle of $8\,^\circ$, the p-InP pillar height should be at least $1.8\,\mu\mathrm{m}$ to get a starting top width of $1\,\mu\mathrm{m}$ (a $1\,\mu\mathrm{m}$ isolated line pattern is readily achieved using contact lithography).

The two other parameters, the p-InGaAs thickness and SCH layer thickness should be chosen in such a way that the device performance is not too strongly dependent on the exact width at the bottom of the p-InP pillar. The higher the tolerance on this critical width, the easier it becomes to fabricate the actual devices. Although it would seem that the p-InGaAs thickness is not important because of the already thick p-InP cladding above the active region, one should take care to avoid effective index matching of the fundamental TE-mode and the lossy mode centered in the p-InGaAs layer. In figure 2.14, a waveguide structure is shown where these modes have a similar effective index, leading to an odd- and even-like mode shape. It is clear that this causes a higher confinement of the fundamental TE-mode in the lossy contacting layers, and therefore a high modal loss. Also the SCH layer thickness is an important parameter: it also influences this effective index matching as described above, but next to this also the active region confinement, extend of the evanescent mode tail in p-InP, and effective index of the fundamental mode are strongly influenced by this parameter.

In figure 2.15(a), the modal loss of the fundamental TE-mode due to the contacting layers as a function of the top metal width is plot-

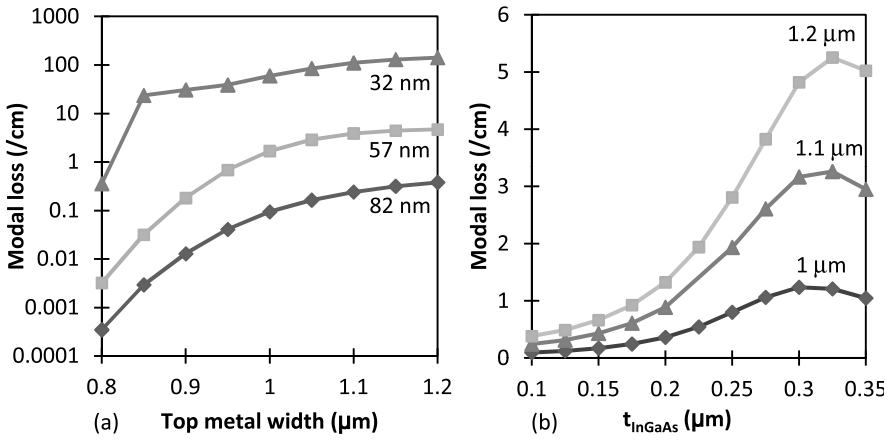


Figure 2.15: (a) Modal loss due to the contacting layers of the fundamental TE-mode as function of top metal width, for different values of the bottom SCH layer thickness. (b) Modal loss due to the contacting layers of the fundamental TE-mode as function of t_{InGaAs}, for different values of the top metal width. The loss has a peak for a InGaAs thickness of $300\,\mathrm{nm}$, due to the effective index matching between the lossy mode near the metalization layers and the fundamental TE-mode.

ted for different bottom SCH layer thicknesses. For this simulation, $t_{InGaAs} = 100\,\mathrm{nm}$ and $t_{SCH,top} = 32\,\mathrm{nm}$ were used. Only the loss due to the metalization and the highly doped InGaAs layer were considered in this simulation. The reason only the bottom layer thickness is changed, is because this allows us to dope a part of the SCH-layer to keep the thickness of the intrinsic active layers the same, without losing the very low capacitance due to the narrow p-InP base. It is clear that a thicker SCH layer drastically decreases the loss, and therefore increases the robustness of the waveguide design. However, the SCH layer should not be made too thick, or the confinement bonus in the MQW region due to the low effective index of the top p-InP pillar structure will be lost. Next to this, also the effective index increases with a thicker bottom SCH layer, leading to a more difficult coupling to the underlying silicon waveguide circuit. As for $t_{SCH,bot} = 82\,\mathrm{nm}$ the loss due to the top contacting layers is negligible, and the design is robust for variations in the lithography as well, this value is selected. In figure 2.15(b), the influence of t_{InGaAs} is shown on the modal loss of the fundamental TE-mode. Due to the use of a thicker bottom SCH layer, the influence of t_{InGaAs} on the loss is limited. Nevertheless, a thickness of less than $100 - 200\,\mathrm{nm}$ is advised to minimize the losses.

In the selected waveguide structure, a confinement of $21.5\,\%$ is achieved in the MQW layers with a fundamental TE-mode effective index of 3.07, while the loss due to the contacting layers is negligible.

2.3.6 Coupling section design

To couple light between the V-waveguide and the underlying silicon waveguide circuit, an inverted taper coupler is proposed. Because of the relatively high effective index of the fundamental TE-mode in the V-waveguide ($n_{eff}(1.55\,\mu\mathrm{m}) \approx 3.0$), a shorter taper with taper tips that do not need to be as narrow can be achieved if a thicker underlying silicon waveguide of $400\,\mathrm{nm}$ is used instead of the standard $220\,\mathrm{nm}$ thick silicon waveguides. This is the case because it is easier to match the effective index of the fundamental TE-mode in a $400\,\mathrm{nm}$ thick silicon wire to the effective index in the corresponding mode in the V-waveguide.

To achieve an as short as possible taper structure, a multi-step taper structure is used, as shown in figure 2.16. In the section between II and III, the n-InP is tapered in. Then, between III and IV the top p-InP is tapered slightly open, from the minimal attainable width ($\approx 800\,\mathrm{nm}$) to the required width for the design (1 to $1.2\,\mu\mathrm{m}$). At the same time, also the active layer width is tapered, by using a separate wet etching mask that is underetched, to achieve a narrow tip at the V-waveguide tip. Furthermore, the silicon wire underneath is tapering down to push the mode up. Between IV and V, the silicon tapers down to a $150\,\mathrm{nm}$ broad tip, while the active layer tapers open to its designed width in the waveguide ($> 1\,\mu\mathrm{m}$).

Note that in the simulations presented below, a DVS-BCB bonding layer thickness of $60\,\mathrm{nm}$ is everywhere assumed. This is the bonding layer thickness that is reproducibly achieved using machine bonding. Because this layer thickness is always the same, the robustness of the taper design towards this parameter is not taken into consideration.

silicon to silicon/III-V transition

The quality of the transition from a broad silicon wire to a broad silicon wire with III-V taper tip on top can have a large effect on the performance of the total taper structure (the transition from II to III in figure 2.16). In figure 2.17(a), the effect of the active region width in the III-V taper tip on the transmission is shown. This simulation is performed for two cases: for a III-V waveguide with the optimized design as above ($t_{act} = 200\,\mathrm{nm}$) and for a III-V waveguide with thicker active region

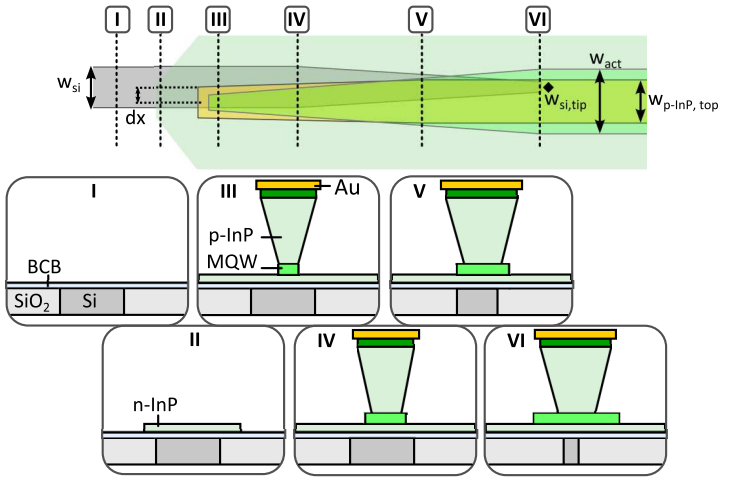


Figure 2.16: Basic layout of the inverted taper structure used to couple between the V-waveguide and the silicon wire. The parameter dx is the misalignment in the taper, which can occur during processing. In the ideal case, this should be $0\,\mathrm{nm}$.

($t_{act} = 300\,\mathrm{nm}$) and therefore an intrinsically higher effective index. A silicon wire width of $1.25\,\mu\mathrm{m}$ and DVS-BCB bonding layer thickness of $60\,\mathrm{nm}$ is assumed. From this simulation it is clear that a narrow tip ($< 500\,\mathrm{nm}$) is important to achieve a lossless taper structure, however this is even more important when the active region is thicker. In that case the transmission quickly drops below $-1\,\mathrm{dB}$ for a width of the order of $1\,\mu\mathrm{m}$, while in the optimized case the design is much more robust to a broader taper tip with a loss limited to $-0.55\,\mathrm{dB}$. To make this tip as narrow as possible, a wet etching scheme can be used, where a broader mask pattern is underetched to achieve a tip of less than $500\,\mathrm{nm}$. For this reason, the active region width w_{act} can be different than the top p-InP width in the III-V waveguide (see figure 2.16).

Next to this, also the width of the silicon wire has an important influence on the performance: the light is better confined in a broader silicon wire, and therefore also the loss due the sudden transition in the waveguide is lower. In figure 2.17(b), the transmission is shown as a function of silicon wire width using a III-V taper tip of $500\,\mathrm{nm}$. Although a broader wire always yields a higher transmission, this also has a drawback as it makes the tapering from this broad silicon to a narrow silicon tip longer, which results in a longer device length. As it

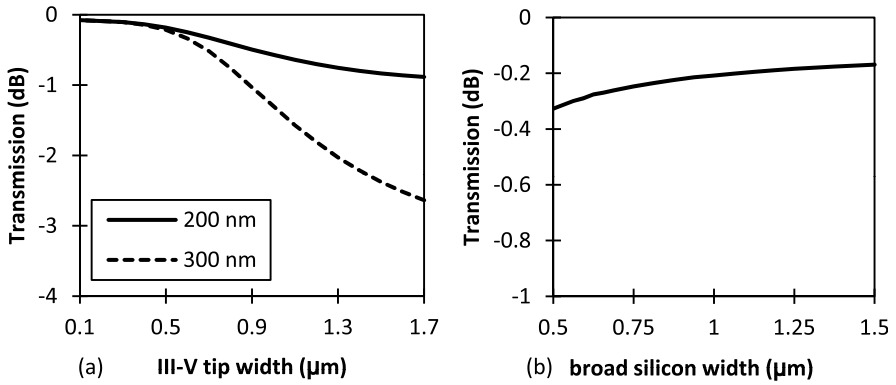


Figure 2.17: (a) Effect of the active region tip width on the device transmission of the fundamental TE-mode for a silicon width of $1.25\,\mu\mathrm{m}$. (b) Effect of the width of the silicon wire underneath the III-V tip on the transmission of the fundamental TE-mode for a III-V tip of $500\,\mathrm{nm}$ wide.

is our goal to limit the device capacitance and hence the device length, a moderate silicon width of $1.25\,\mu\mathrm{m}$ is selected.

silicon/III-V to III-V transition

The second abrupt transition in the taper structure is much easier to design: because of the availability of high resolution deep-UV lithography for the definition of the silicon layer, taper tips of the order of $100\,\mathrm{nm}$ are achievable. Therefore, losses can be made extremely low. In figure 2.18, the transmission through this transition is plotted as a function of silicon tip width. For a silicon tip of less than $200\,\mathrm{nm}$ broad, the loss is negligible. Therefore, in the design a width of $150\,\mathrm{nm}$ is selected.

silicon/III-V taper

In the inverted taper part, the silicon needs to taper from a width of $1.25\,\mu\mathrm{m}$ to $0.15\,\mu\mathrm{m}$, while the III-V should taper from a top metal width of $0.8\,\mu\mathrm{m}$ to $1.2\,\mu\mathrm{m}$ and an active region width of $0.4\,\mu\mathrm{m}$ to $1.7\,\mu\mathrm{m}$. In the simulation, two steps are performed. First the length of the total structure is varied, where a perfectly linear taper is assumed (see figure 2.19(a)). As an extra parameter, the alignment accuracy of the top III-V waveguide to the underlying silicon is considered, to be able to make a more robust design choice. From this simulation, it follows that a taper length of $35\,\mu\mathrm{m}$ is required to achieve an adiabatic transition in the per-

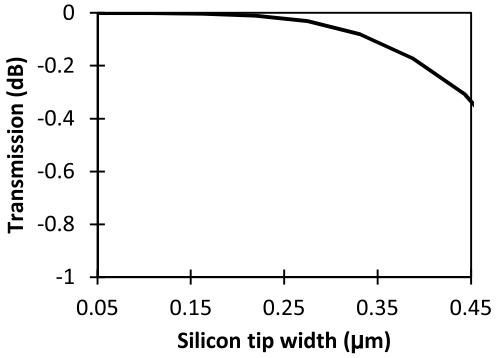


Figure 2.18: Effect of the silicon tip width on the device transmission of the fundamental TE-mode.

fectly aligned case, while this already becomes more than $50\,\mu\mathrm{m}$ when there is a $250\,\mathrm{nm}$ alignment error. For an alignment error of $500\,\mathrm{nm}$, the tapering does not work well anymore.

In a second step, the relative scale of the three different taper sections (see figure 2.16, from III to VI) is optimized, and the transmission through the taper as a function of total taper length is simulated again. The result of this simulation is shown in figure 2.19(b). In this case, a taper length of $25\,\mu\mathrm{m}$ is enough to attain adiabatic coupling when the III-V waveguide is perfectly aligned to the underlying silicon circuit, and a length of $35\,\mu\mathrm{m}$ ensures adiabatic coupling even when the III-V waveguide is misaligned by $250\,\mathrm{nm}$. However also in this optimized case, an alignment error of $500\,\mathrm{nm}$ is too much for the tapering scheme to work well. The optimized taper is not exactly linear anymore. Instead, it is a piecewise linear structure, which is slightly slower changing in taper sections III to IV and V to VI, where the mode is transforming from a mode mainly confined either in silicon or in III-V, to a 'hybrid' mode, which is confined in both waveguide cross sections.

2.3.7 Electro-optical simulation

Methodology

A similar method is used as described in section 2.2.2, however in this case scattering losses and thermal effects are not taken into account. Instead, an operation temperature of $300\,\mathrm{K}$ is assumed. When applying a voltage, ATLAS automatically calculates the associated stark shift in the band diagram. Then, using the same QW model as for the cal-

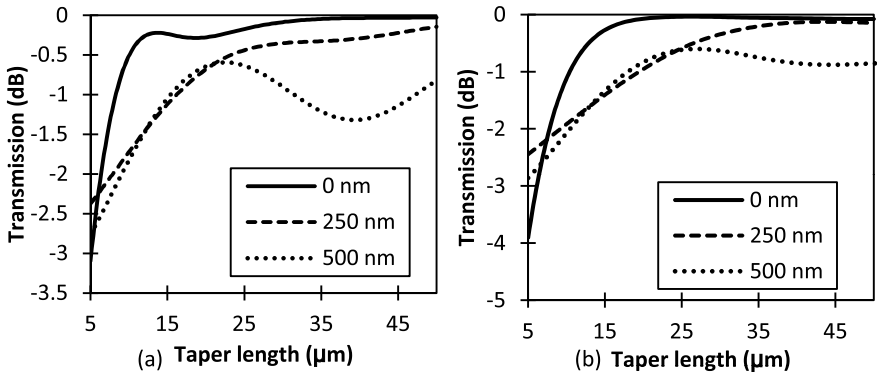


Figure 2.19: Transmission for the fundamental TE-mode through respectively (a) a linear taper design and (b) a piecewise linear taper design, for an alignment error of $0\,\mathrm{nm}$, $250\,\mathrm{nm}$ and $500\,\mathrm{nm}$, as a function of taper length.

culation of gain, the QW absorption is calculated. By performing this calculation for different applied reverse biases, the DC transmission as a function of both wavelength and applied bias can be calculated. Next to this, ATLAS also allows us to determine the small signal response of the device at a certain bias point. In this way, the capacitance of the device can be accurately calculated, and together with the knowledge of the device resistance the projected electrical bandwidth can be determined.

Because the InGaAlAs material system is not well supported in the Silvaco material library, a MQW stack based on the InGaAsP material system was used in the simulation. One of the consequences is that the simulated performance will be worse than anticipated as the conduction band offset of InGaAsP materials is lower than in InGaAlAs (see section 2.3.2). Next to this, a staircase approximation of the trapezoid shaped p-InP layer is made. Although ATLAS allows for the definition of trapezoid shapes, this leads to a huge increase in number of grid points due to the simple grid algorithm used in ATLAS. To save simulation time, the result of an electrical simulation using an exact shape was compared to the result of a simulation using a staircase approximation, as the latter still allows for a simple grid to be used. It was found that there was barely any difference in the simulation results, and therefore in the following the staircase approximation is everywhere used.

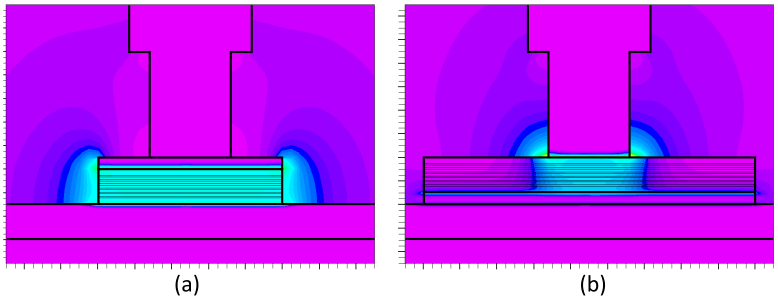


Figure 2.20: (a) Electric field distribution in the p-waveguide. (b) Electric field distribution in the n-waveguide.

Static performance

In figures 2.21 and 2.22, the absorption spectrum of the MQW structure for the fundamental TE-mode is shown at different reverse bias voltages. In figure 2.21, this is shown for the waveguide structure with an extra n-doped SCH layer in the waveguide structure which will be referred to as the n-waveguide, while in figure 2.22 this is shown with an extra p-doped SCH layer which will be referred to as the p-waveguide. The electric field distribution in the active region of both waveguide structures is shown in figure 2.20. Note that the p-waveguide requires a well-controlled active region width in order to keep the capacitance low. An active region width of $1\,\mu\mathrm{m}$ is assumed. In the n-waveguide, the active region width does not determine the device performance, and therefore a broader width of $1.7\,\mu\mathrm{m}$ is assumed.

The QCSE is clearly visible in both cases: while the absorption band edge shifts to longer wavelengths, also the QW absorption drops due to the reduced overlap of hole and electron wave functions. Because of the lack of a p-doped SCH layer in the n-waveguide, a slightly weaker absorption modulation is obtained, as the strongest part of the applied electric field is only present in the central part of the MQW structure, situated under the p-InP pillar. This is also clear from the fact that the absorption at shorter wavelengths remains stronger in the n-waveguide, because the QW absorption on the sides of the p-InP pillar is more weakly shifted. As the fundamental TE-mode is mostly confined in the central section however, the difference in performance to the p-waveguide remains limited to $37\,\%$, and therefore to obtain a similar ER, the device will also have to be $37\,\%$ longer.

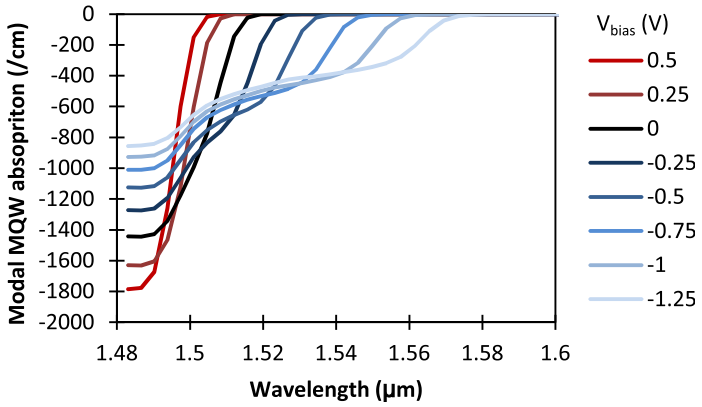


Figure 2.21: Modal absorption spectrum of the MQW structure with n-doped SCH layer as a function of wavelength, for different applied bias voltages across the device.

Another interesting feature is visible from these simulations: apparently the built-in electric field in the p-i-n junction is sufficiently strong to create a small stark shift without any externally applied bias. Therefore, the device can be operated at zero bias, or even a small forward bias. As discussed in section 2.3.4, this leads to a large reduction of the power consumption. A reverse bias can still be necessary however in order to tune the modulator to the right operating wavelength with a tuning rate of $40\,\mathrm{nm/V}$.

In figure 2.23(a) and (b), the resulting modulation of the fundamental TE-mode in respectively a $40\,\mu\mathrm{m}$ long and a $55\,\mu\mathrm{m}$ long device is shown for different signal wavelengths as a function of the bias voltage for both waveguide designs. In both cases a large ER is obtained for a very small peak-to-peak voltage: an ER in the range of 8 to $18\,\mathrm{dB}$ can be obtained in the wavelength range of $1510 - 1555\,\mathrm{nm}$ depending on the bias voltage for a voltage swing of the order of $0.45\,\mathrm{V}$.

Dynamic performance

To calculate the device response at higher frequencies, a small signal approximation is made. The equivalent circuit of the AC source with $50\,\Omega$ input impedance and connected modulator, which is driven with lumped electrodes is shown in figure 2.24.

First the specific device resistance (R_s) is calculated in a static simulation, where a current is driven through the device and the resistivity

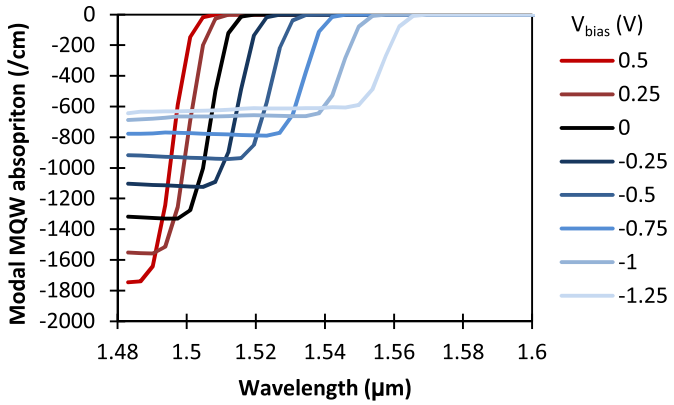


Figure 2.22: Modal absorption spectrum of the MQW structure with p-doped SCH layer as a function of wavelength, for different applied bias voltages across the device.

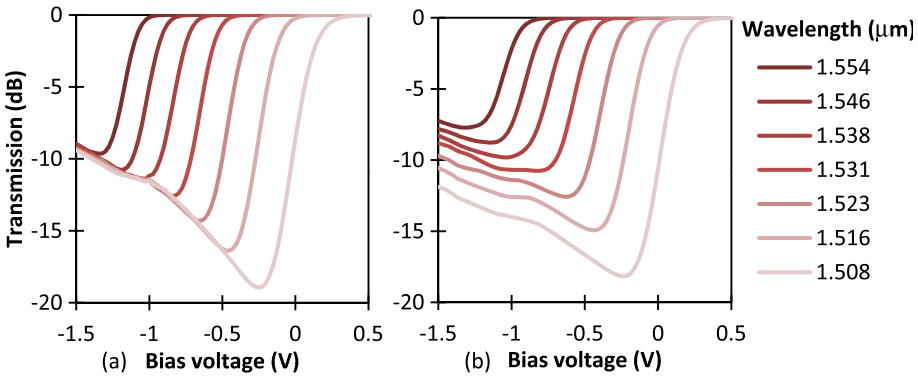


Figure 2.23: Simulated transmission through a modulator of respectively (a) $40\ \mu\text{m}$ long using the p-waveguide, and (b) $55\ \mu\text{m}$ long using the n-waveguide.

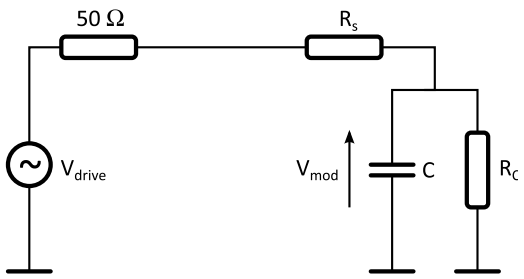


Figure 2.24: Small signal schematic of modulator driven by lumped electrodes.

of each device section can be calculated. The specific resistance of both waveguide designs is the same, and a value of $3\,\Omega\mathrm{mm}$ is obtained. It is also clear that the device resistance is mainly the result from the p-doped InP pillar, which has a specific resistance of $2.3\,\Omega\mathrm{mm}$. Increasing the doping in this region could improve the device resistance greatly, however at the cost of a higher insertion loss due to FCA.

Then, the device conductance and capacitance per unit length are calculated using a small signal simulation. As in principle no reverse bias is required to drive the device, a bias voltage of $0.0\,\mathrm{V}$ is selected. At low frequencies, the device conductivity ($1/R_c$) is nearly zero, while a specific capacitance (C) of $0.79\,\mathrm{pF/mm}$ in the p-waveguide and $0.59\,\mathrm{pF/mm}$ in the n-waveguide are obtained. This includes the parasitic capacitance to the silicon substrate. At higher frequencies however ($> 10\,\mathrm{GHz}$), the majority charge carriers which determine the device operation do not follow the changes in the field quasi-statically anymore, and as a result both the capacitance and conductance change: a higher conductance but lower capacitance are obtained. Taking into account that a $37\,\%$ longer device is needed when the n-waveguide is used to achieve the same ER as with the p-waveguide and adding the required taper length which amounts to a total of $70\,\mu\mathrm{m}$, the device bandwidth of both waveguide designs can be compared. The transfer functions which take into account the frequency dependance of the capacitance and conductance are shown in figure 2.25, for a $110\,\mu\mathrm{m}$ long p-waveguide and a $125\,\mu\mathrm{m}$ long n-waveguide. A device bandwidth of respectively $22.5\,\mathrm{GHz}$ and $26.5\,\mathrm{GHz}$ are obtained. By applying a bias of $-1\,\mathrm{V}$, the bandwidth can be slightly increased to $25\,\mathrm{GHz}$ and $29.6\,\mathrm{GHz}$ for p-waveguide and n-waveguide respectively. This is due to the fact that the depletion region width becomes larger when a reverse bias is applied, decreasing the specific capacitance of the waveguide.

Power consumption

Because a reverse bias is not necessary for the device operation, the power consumption can be very low for both the n-waveguide and the p-waveguide. Assuming an input power of $5\,\mathrm{dBm}$ as in [21], an ER of $10\,\mathrm{dB}$, a bit rate of $40\,\mathrm{Gbit/s}$ and a bias of $0\,\mathrm{V}$, the total calculated power consumption is for both waveguide designs approximately $48\,\mathrm{fJ/bit}$ (see section 2.3.4). However, this includes the insertion loss due to the modulator absorption, while in the cited results in section 2.3.1, the

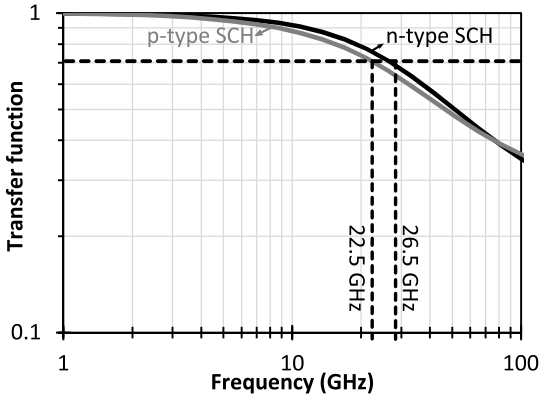


Figure 2.25: Device transfer function for a device length of $110\,\mu\mathrm{m}$ and $125\,\mu\mathrm{m}$, for p-waveguide (grey) and n-waveguide (black) respectively.

power consumption related to the optical loss was not included. As no DC bias is applied, the contribution due to the photocurrent is very small, as V_{tot} is in this case only $0.225\,\mathrm{V}$. Therefore we get a power consumption excluding the optical loss of only $13\,\mathrm{fJ/bit}$.

In the case the modulator response needs to be shifted to a longer wavelength, and e.g. a reverse bias of $-1\,\mathrm{V}$ is applied, the power consumption will increase. In this case, a total power consumption of $87\,\mathrm{fJ/bit}$ is achieved including the optical loss, and $52\,\mathrm{fJ/bit}$ excluding this loss.

Clearly, the thin active region and low capacitance of the device pays off. Compared to the state-of-the-art, the power consumption is predicted to be less than half of what is achieved in [21].

Fabrication considerations

From the above, we can conclude that both the p-waveguide and n-waveguide are viable waveguide structures for the realization of an integrated QCSE modulator on SOI. Both waveguide structures have similar electrical and optical performance. Therefore, the ease of fabrication can be the determining factor in the comparison between these two choices.

Both waveguide structures need a well-controlled lithography and etching step to define the p-InP pillar, as the position and width of this pillar determines the mode-shape and alignment to the silicon waveguide circuit. Even when the etched active region is misaligned

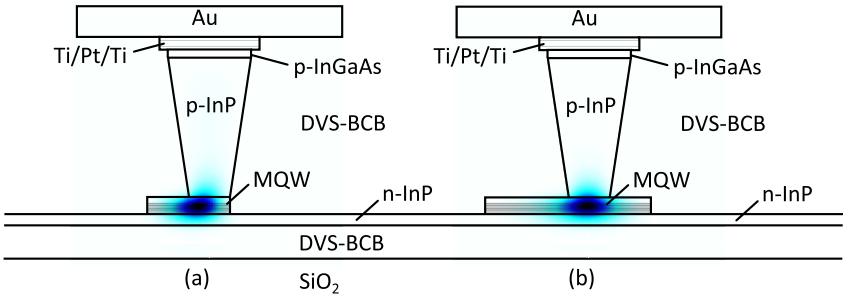


Figure 2.26: (a) p-waveguide of $1\,\mu\mathrm{m}$ broad with $250\,\mathrm{nm}$ misalignment of active region to p-InP pillar position. (b) n-waveguide of $2\,\mu\mathrm{m}$ broad with $250\,\mathrm{nm}$ misalignment of active region to p-InP pillar position.

compared to the top p-InP pillar, the mode shape will be nearly unaffected, as long as the etched sidewalls of the active region are far enough from the mode center (see figure 2.26). However, the p-waveguide has the disadvantage that the capacitance is determined by the active region width, while the capacitance in the n-waveguide is determined by the width of the pillar base only. This means that in the p-waveguide the active region should be narrow enough and that therefore also the alignment becomes important. As a result of a minor alignment error of $250\,\mathrm{nm}$ as in figure 2.26, the mode in the $1\,\mu\mathrm{m}$ wide p-waveguide is already deformed, while the mode in the $2\,\mu\mathrm{m}$ broad n-waveguide remains unaffected. For larger misalignments, this effect is even stronger. As a result, the fabrication process is more stringent for the p-waveguide than for the n-waveguide.

2.3.8 Conclusion

A high performance QCSE modulator integrated on the SOI waveguide platform was designed. By using wet chemical etching of InP, smooth sidewalls with a negative slope can be easily obtained. This allows us to define a narrow p-InP pillar of the order of $500\,\mathrm{nm}$ wide on top of the active region using simple contact lithography with a mask stripe width of $1\,\mu\mathrm{m}$. Because of the low effective index of such a narrow p-InP pillar, the fundamental TE-mode is very well confined in the active region, and furthermore the doping profile in the SCH layers of the active region can be used to reduce the specific capacitance to only $0.59\,\mathrm{pF/mm}$.

In the optimized design, we were able to show through simulation that modulation with an ER of over $10\,\mathrm{dB}$ can be obtained in a device of $55\,\mu\mathrm{m}$ long over a wide wavelength range, using a peak-to-peak driving voltage of only $0.45\,\mathrm{V}$. Next to this, the device has a high bandwidth of $26.5\,\mathrm{GHz}$ even when it is driven by simple lumped electrodes. Furthermore, because of the small thickness of the active region, the built-in electric field is high enough to avoid the need for a bias voltage. As a result, a very low power consumption of $48\,\mathrm{fJ/bit}$ is predicted. A DC bias is only needed to tune the wavelength of operation of the modulator, which is tunable at a rate of $40\,\mathrm{nm/V}$. The application of a reverse bias can also help speed up the device even more, allowing a bandwidth of up to $30\,\mathrm{GHz}$.

2.4 Summary

In this chapter, it was investigated how the high index contrast in the bonded III-V/SOI platform can be used for the optimization of the performance of integrated active devices. It was shown that by careful design of the top p-InP contact region, the confinement in the active region of the fundamental TE-mode can be maximized. This can be used to create very high gain amplifiers useful for the realization of compact microlasers. Next to this, also the device capacitance can be minimized, which can be used to create highly performing, low-voltage QCSE modulators integrated on SOI.

Chapter 3

Optically Pumped Membrane Gates

3.1 Introduction

In the previous chapter, the potential of waveguide structures that provide a high optical confinement as well as electrical pumping were discussed. However for some applications, electrical pumping is not required and might even be undesirable. Because of this, the current injecting layers, such as the highly doped n-InP and p-InP layers as well as the metal layers do not need to be included into the waveguide design anymore. The only thing that needs to remain, is the active region of the device. This allows us a lot more design freedom. We will show in this chapter that by shrinking the active waveguide to a $100\,\mathrm{nm}$ thick stripe, the optical confinement in the QWs can be maximized and a very high light-matter interaction can be achieved. Therefore, devices can be made more compact, leading to a lower energy consumption and higher possible integration density, making this kind of devices an ideal candidate for all-optical signal processing applications. Using these devices, we will demonstrate all-optical regeneration (see chapter 4), and an all-optical packet switch will be created (see chapter 5).

3.2 Membrane waveguide design

3.2.1 Optical confinement

As the electrical contact layers can be removed from the design, there are no metalization or doping losses anymore that have to be consid-

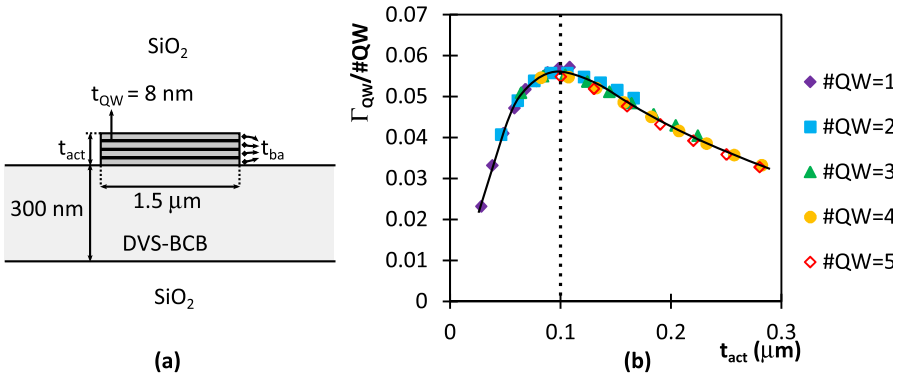


Figure 3.1: (a) Simulated waveguide structure: the barrier thickness t_{ba} as well as the number of QWs in the structure are parameters that are changed in the optimization procedure. The QW material is InGaAs, the barrier material is InP. (b) Confinement per QW.

ered when designing the waveguide structure. This allows for a much more compact waveguide design, with the only trade off that scattering losses will be higher. To explore the possibilities, a simulation was done on the waveguide structure depicted in figure 3.1(a). For each number of QWs, the thickness of the QW barrier layers t_{ba} was changed, which also changed the total waveguide thickness t_{act}. In this simulation the QW was an $8\,\mathrm{nm}$ thick InGaAs layer, with InP barrier layers. In figure 3.1(b) the achieved confinement per QW is plotted as a function of the total waveguide thickness for the fundamental TE-mode at a wavelength of $1550\,\mathrm{nm}$. From this it is clear that there is an optimal waveguide thickness of $100\,\mathrm{nm}$ to achieve a maximal confinement for any number of QWs in the device structure. Therefore, throughout the chapter we will only consider ultrathin membranes of $\approx 100\,\mathrm{nm}$ thick.

3.2.2 Quantum Wells

As an active region for the membrane waveguide, a MQW structure is used. The advantage of this, is that by tuning the dimensions and changing the material of the QWs, the QW response can be tuned to fit the application. In the following chapters, two different applications will be discussed. For the optical regenerator, the optimal response is a QW with PL peak at $1.58\,\mu\mathrm{m}$, while for the all-optical switch a QW with PL peak at $1.55\,\mu\mathrm{m}$ should be used.

A second design choice that needs to be made, is the choice of the material combination used for the QW and barrier layers. In this chapter, we will use different combinations. As a starting point an InGaAlAs/ InGaAlAs QW/barrier will be used. This type of QWs has been found to be performing very well for electrically pumped amplifiers, as it has an increased conduction band offset compared to the InGaAsP/ InGaAsP material system [49]. As discussed in section 2.3.2 this is a very important advantage for the realization of QCSE modulators, however also for the realization of amplifiers it can be an advantage. Because of the increased conduction band offset, it will be less likely for free electrons in the QW to be thermally excited to escape the QW. Therefore, amplifiers in the InGaAlAs/ InGaAlAs typically perform better at higher temperatures. However, the QWs in the membrane waveguide structure are not electrically pumped. Therefore, the barrier material can be made in an arbitrarily large band gap material. A second candidate that will be considered is therefore the InGaAsP/ InP material system, as the barrier that needs to be crossed by thermal excitation is in this case also very large. Furthermore, this material system is better for the thermal design as will be discussed in section 3.2.4. In electrically pumped MQW structures this material system is not used, as the QWs are typically embedded in a lower band gap SCH (e.g. p-InP/ InGaAsP(Q1.2)/ n-InP), as this heterostructure forms an additional barrier for injected electrons and holes, such that they are more likely to be captured into the QWs in the MQW structure.

A third design choice is the number of QWs that will be used in the MQW structure. From the simulation result shown in figure 3.1(b), it also follows that the total optical confinement in the QWs scales linearly with the number of QWs. In order to achieve the most compact device possible, it is therefore best to take the highest amount of QWs that fit in a $100\,\mathrm{nm}$ thick layer. Although this delivers the highest possible integration density and lowest total loss due to scattering, a drawback is that power will be dissipated on a smaller surface area, leading to a higher operating temperature and therefore worse device performance.

3.2.3 Optical pumping

There are two ways to optically pump a MQW structure. A first way is to pump the structure with a beam with a photon energy that is slightly larger than the QW band gap (see figure 3.2(a)). Because of the band-

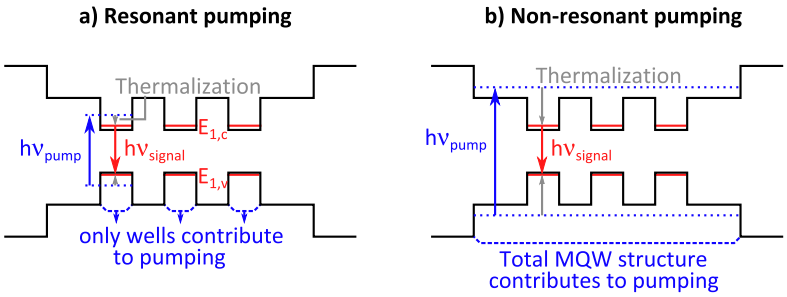


Figure 3.2: (a) Resonant optical pumping: only the QWs contribute. Note that the blue dotted line does not (necessarily) correspond to a higher QW level, but to a higher energy within the first QW level. (b) Non-resonant optical pumping: the total MQW structure contributes, however a lot of energy is lost as heat due to thermalization.

to-band absorption of this beam, free carriers are created. As the pump beam increases in power, more and more free carriers are created and stimulated emission becomes more important. In the limit, the absorption of the QWs for the pump beam drops to zero when an equal number of free electrons is present in the conduction band as there are valence electrons in the valence band. When a free electron/hole is created by band-to-band absorption, this free carrier will quickly relax to the bottom of the conduction/valence band because of thermalization, e.g. by non-radiative interactions with phonons. Therefore, when the MQW structure is strongly pumped such that the pump beam absorption becomes nearly zero, there will be population inversion for photon energies smaller than the photon energy of the pump beam. This is called resonant pumping. An advantage of this technique is that it is potentially more efficient, as ideally only the energy difference between the absorbed pump photon and emitted signal photon will be dissipated as thermal energy inside the device during device operation. This is the reason why this technique has been used for the realization of high power optically pumped lasers [52]. A drawback of resonant pumping however, is that it is more difficult to reach transparency as the pump beam absorption is bleached itself when free carriers are created. By increasing the difference in pump photon and QW band gap energy, this can be improved at the cost of a lower pumping efficiency.

The second way is by pumping the complete MQW structure by using a pump beam with a photon energy that is larger than the barrier band gap (see figure 3.2(b)). This is called non-resonant pumping.

This technique has the advantage that the pump beam will be strongly absorbed by the complete MQW structure, even when the QWs are already transparent for the signal beam. Carriers generated in the barrier layers are captured in the QWs by thermalization. A disadvantage is therefore that a larger portion of the pump photon energy will be lost as heat, and as a consequence a device might be more susceptible to thermal roll-over. To limit the amount of energy lost, the barrier height should not be too high. This pumping technique is typically used when the optical pumping beam is incident from the top of the device, as a very high pump beam absorption is then required to generate a sufficient amount of carriers without going to a very high pump power [53].

Although optical pumping by itself is in theory an efficient process, the pump photons at some point need to be generated by electrical pumping. Therefore, the true efficiency of optical pumping is a lot lower, and is limited by the wall-plug efficiency of the pump laser. However, in some cases optical pumping is still the best way to pump the device. For the realization of high power lasers [52], it has the advantage that the amount of dissipated power can be limited, improving the performance of the laser in the desired wavelength region, while a well established laser at a single wavelength with high wall-plug efficiency [54] can be used as pump laser. In some cases the device that needs to be pumped is in a remote location that does not have electrical power present (see chapter 5). Furthermore, when the pump laser is replaced by a signal, the carrier dynamics in the device can be used to perform complex signal processing functions (see chapter 4).

3.2.4 Thermal design

As the membrane is a very thin and compact device, the generation of heat during device operation can become a problem.

Influence of temperature on device operation

The carrier distribution in a semiconductor is governed by Fermi-Dirac statistics. Under pumping, we get the following distributions for elec-

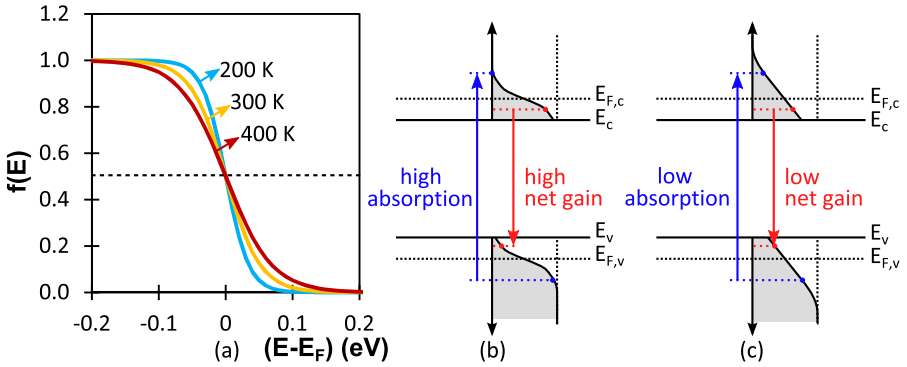


Figure 3.3: (a) The influence of temperature on Fermi-Dirac statistics. (b) Gain and absorption for at low temperatures with $E_{F,c} > E_c$ and $E_{F,v} < E_v$. (c) Gain and absorption for high temperatures with $E_{F,c} > E_c$ and $E_{F,v} < E_v$.

trons and holes respectively

$$f_c(E) = \frac{1}{\exp\left[(E - E_{F,c})/(k_B T/q)\right] + 1} \tag{3.1}$$

$$f_v(E) = \frac{1}{\exp\left[(E - E_{F,v})/(k_B T/q)\right] + 1} \tag{3.2}$$

As explained in the appendix (see section A.6.2), the optical gain for a transition from an energy level E_j to E_i can be calculated using Fermi's golden rule [55]

$$G(E_{ij}) = \frac{q^2 h}{2m_0^2 \epsilon_0 n c} \frac{|M(E_{ij})|}{E_{ij}} \rho_r(E_{ij}) \left(f_c(E_j) - f_v(E_i)\right) \tag{3.3}$$

When the factor $(f_c(E_j) - f_v(E_i))$ is positive, population inversion is achieved which results in a net gain for a photon energy equal to the transition energy E_{ij}. When this factor is negative however, this corresponds to net absorption. This factor is dependent on the temperature, as the Fermi-Dirac distribution has a strong temperature dependence. This can clearly be seen in figure 3.3(a) where the Fermi-Dirac distribution is plotted for a temperature of $200\,\mathrm{K}$, $300\,\mathrm{K}$, and $400\,\mathrm{K}$. For very low temperatures, the Fermi-Dirac distribution becomes nearly a step function that is ≈ 1 for $E < E_F$ and ≈ 0 for $E > E_F$. This case corresponds to the situation where all free carriers relax to the lowest available energy states. This situation is schematically shown in figure 3.3(b). Because of the nearly stepwise nature of the Fermi-Dirac distribution, a relatively small amount of free carriers is required to achieve

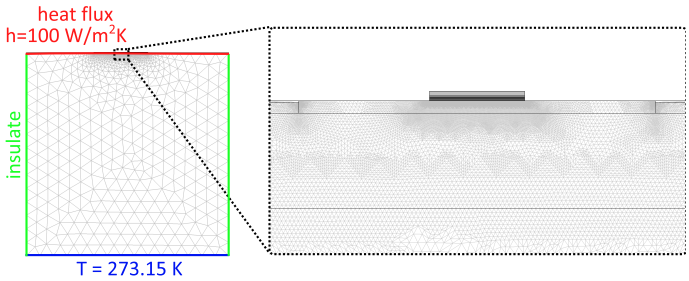


Figure 3.4: Mesh used for the thermal simulation of the membrane devices. Also the boundary conditions used in the simulation are shown.

a strong population inversion. Furthermore, it is possible that a net gain is achieved for a photon energy E_{signal} while there is still a strong absorption for a pump photon energy E_{pump}. At a high temperature however, the Fermi-Dirac distribution becomes much more smooth. This situation is schematically shown in figure 3.3(c). Therefore, a lot more carriers are required to achieve a strong population inversion, as they are spread out over a wider energy range. Furthermore, when a net gain is achieved for a photon energy E_{signal}, the absorption at the pump photon energy E_{pump} will also be strongly bleached because the carriers are more spread out due to the smoothness of the Fermi-Dirac distribution. When using optical pumping, a higher temperature has thus two times a negative impact on the device performance.

Design optimization

It is clear that we should minimize heating of the device during operation. A first way to avoid heating is obviously to limit the amount of dissipated power per unit area. Any design effort that improves the efficiency of the device, will therefore also limit the heating of the device. Next to this, one can also use a smaller amount QWs in the device, as this increases the device size to achieve the same performance. Therefore the power will be dissipated over a larger area, leading to a lower device temperature. However the drawback of this is a lower possible integration density. In section 3.7, experimental results for both epitaxial structures with three and four QWs will be presented.

Another important design aspect is the thermal resistance of the device, as the device temperature scales linearly with the thermal resistance. To assess the thermal resistance, a 2D thermal simulation was performed in COMSOL. In this simulation, the membrane device is

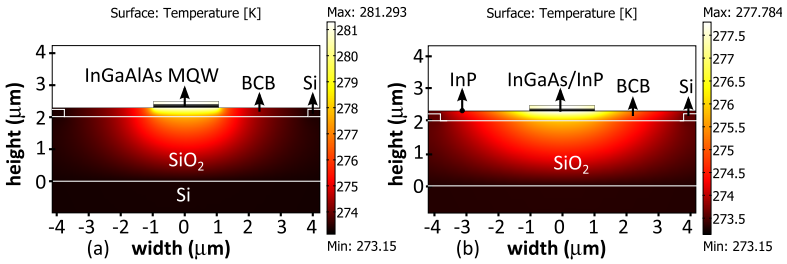


Figure 3.5: (a) Thermal profile for the stripe waveguide. (b) Thermal profile for the ridge waveguide.

considered to be infinitely long, and therefore only the heat flow perpendicular to the cross section of the device is evaluated. To simulate the actual temperature rise in a $100\,\mu\mathrm{m}$ long device, a heat source is defined in the QWs, such that the total integrated dissipated power over $100\,\mu\mathrm{m}$ with three QWs amounts to $1\,\mathrm{mW}$. In this way, we can calculate the thermal resistance in units of $\mathrm{K/mW}$. We assume furthermore that the backside of the silicon die is kept at a constant temperature of $273.15\,\mathrm{K}$. In order to simulate the effect of heat spreading across the die, a simulation window of $750\,\mu\mathrm{m} \times 753\,\mu\mathrm{m}$ is considered. The top part of the die is assumed to be in contact with air, which has a thermal heat transfer coefficient of $< 100\,\mathrm{Wm^{-2}K^{-1}}$, however the cooling effect is so small that considering thermal insulation instead did not alter the simulation results. The mesh is defined using advancing front triangular meshing, such that the grid size is smallest near the III-V membrane and coarse far away from the membrane. The mesh and boundary conditions are shown in figure 3.4.

The basic design considered in this simulation is a $2\,\mu\mathrm{m}$ broad strip waveguide that consists of three $8\,\mathrm{nm}$ thick InGaAlAs QWs, separated by $10\,\mathrm{nm}$ InGaAlAs barrier layers, with a $14\,\mathrm{nm}$ InGaAlAs and $10\,\mathrm{nm}$ InP cladding layer on top and bottom of the MQW stack. The thermal conductivities of the different materials in the design are listed in table 2.3. In figure 3.5(a) the temperature profile in this waveguide structure is plotted. The thermal resistance is found to be $8.14\,\mathrm{K/mW}$. This high value is mainly the result of the insulating BCB and SiO_2 layers.

Due to the strong insulation provided by the BCB and SiO_2 layer, there is very little heat spreading. However, this can be circumvented: if a ridge waveguide shape is used instead of a stripe, the shallowly etched III-V will act as a heat spreader, as III-Vs have a much larger thermal conductivity than both BCB and SiO_2. To maximize this effect,

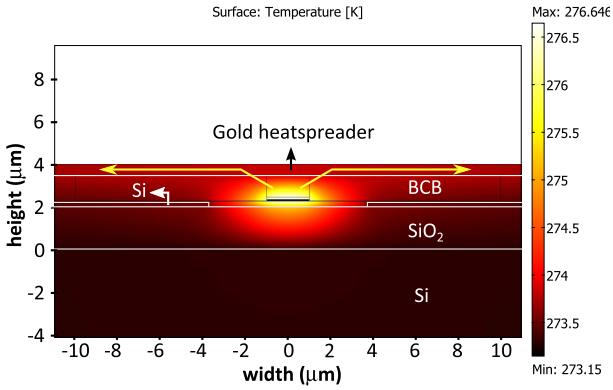


Figure 3.6: Thermal profile for the ridge waveguide with top gold heat spreader. The yellow arrows show the additional pathway along which generated heat can escape.

it would furthermore be beneficial to use a pure InP heat spreader, as InP has a thermal conductivity of $40\,\mathrm{Wm^{-1}K^{-1}}$, while the thermal conductivity of materials typically used for the active region such as InGaAlAs and InGaAsP is $< 10\,\mathrm{Wm^{-1}K^{-1}}$. Therefore, in this improved design we switch to a MQW system with $8\,\mathrm{nm}$ InGaAsP QWs separated by $10\,\mathrm{nm}$ InP barriers. This MQW stack is then sandwiched by two $29\,\mathrm{nm}$ InP cladding layers. The ridge is shallow etched such that only the bottom InP layer remains, with a ridge width of $20\,\mu\mathrm{m}$. The resulting thermal profile is plotted in figure 3.5(b). From the shape of the profile, it is clear that the heat is spread out much better, and this results in a nearly $50\,\%$ reduced thermal resistance of $4.43\,\mathrm{K/mW}$.

In the current design, the heat can only escape by going down or sideways from the heat source in the QWs. The resistance can therefore be further improved by providing a pathway on top of the device. This can be done by depositing a thick BCB layer on top of the membrane device, and depositing a thick gold layer on top of that. Like the ridge waveguide, this gold layer functions as a heat spreader. The BCB layer needs to be thick enough, to make sure that there is no unwanted optical absorption due to the presence of the gold layer. In figure 3.6, the thermal profile is plotted for a device with a $1\,\mu\mathrm{m}$ thick top BCB cladding, and a $500\,\mathrm{nm}$ thick gold heat spreader with a width of $50\,\mu\mathrm{m}$. The thermal resistance is further reduced to $3.5\,\mathrm{K/mW}$.

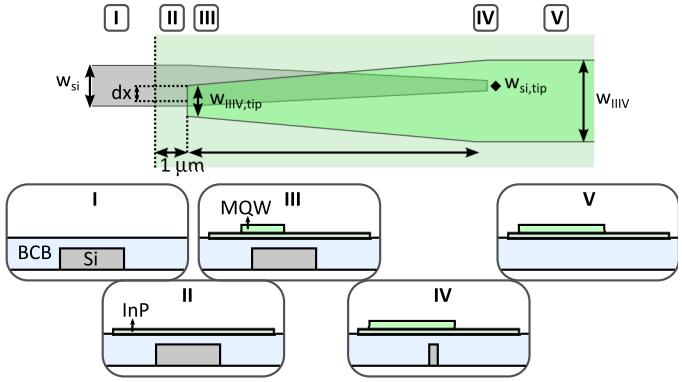


Figure 3.7: Schematic of the inverted taper coupler.

3.3 Silicon design

3.3.1 Coupling section design

As a coupling section, we will use an inverted taper coupler. In the following, we will show that the III-V membrane waveguide has the ideal shape to easily couple the fundamental mode adiabatically from the silicon wire to the III-V membrane waveguide and visa versa using an inverted taper coupler.

The general shape of the inverted taper coupler is schematically shown in figure 3.7. It can be divided into three parts. First there is a transition from a pure silicon waveguide to a silicon waveguide with III-V waveguide on top. The second part is the tapering from a broad silicon waveguide to a narrow silicon waveguide, and optionally simultaneously from a narrow III-V waveguide to a broad III-V waveguide. The third part is the transition from a silicon/III-V waveguide to a pure III-V waveguide.

Silicon to silicon/III-V transition

A first advantage of the III-V membrane waveguide is that the effective index of the fundamental mode for TE polarization is quite low ($n_{eff} \approx 2$ at $1550\,\mathrm{nm}$). As the effective index in a $220\,\mathrm{nm}$ thick silicon wire can easily be made larger than this, it is not so important to achieve a narrow taper tip in the III-V membrane waveguide in order to avoid reflections, as the fundamental mode shape in a sufficiently broad silicon wire will only slightly change shape with or without III-V

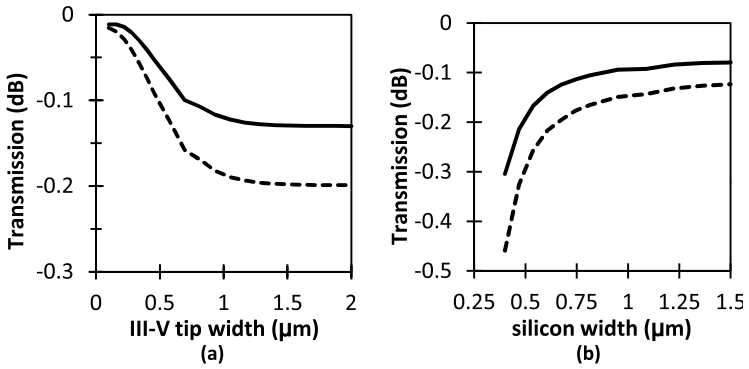


Figure 3.8: (a) Transmission as a function of III-V tip width for a silicon wire width of $700\,\mathrm{nm}$. (b) Transmission as a function of silicon wire width for a III-V tip width of $1\,\mu\mathrm{m}$. The dashed curve corresponds to the situation with a strip waveguide, while the solid curve corresponds to the situation with a ridge waveguide.

membrane on top. In figure 3.8(a) the transmission for the fundamental mode at the interface between a silicon wire and a III-V membrane is plotted as a function of III-V membrane tip width, and in figure 3.8(b) as a function of silicon wire width. Two design cases are being simulated: for a III-V strip waveguide, there is an abrupt transition from the silicon wire to silicon/III-V. However, for the III-V ridge waveguide, there is an opportunity to first transition from the silicon wire to the silicon wire with shallowly etched III-V on top, before going to the full silicon/III-V. This is schematically shown in figure 3.7. This can be considered as a form of vertical tapering, and will therefore also reduce reflections and improve transmission. From these simulations, it is clear that even for a III-V tip of $1\,\mu\mathrm{m}$, which is easily obtainable using contact lithography, the loss at the transition can be limited to $< 0.15\,\mathrm{dB}$ for a moderately wide starting silicon width of $> 700\,\mathrm{nm}$.

Silicon/III-V to III-V transition

Because of the high index of silicon, the silicon taper will need to end in a narrow tip to avoid reflections and losses. In figure 3.9, the transmission at the interface of the silicon/III-V to III-V waveguide is plotted as a function of silicon taper width. As the silicon circuit is defined using a deep-UV lithography stepper, dimensions of the order of $100\,\mathrm{nm}$ can be readily achieved, and therefore losses can be avoided.

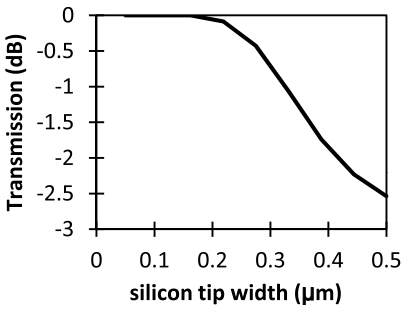


Figure 3.9: Transmission as a function of silicon tip width for a III-V width of $1\,\mu\mathrm{m}$.

Adiabatic mode conversion in the taper

Another advantage of the membrane waveguide for the coupling section design is the proximity of the fundamental mode to the DVS-BCB bonding interface. Because of this, there will be a considerable overlap between the fundamental modes of a silicon wire and a III-V membrane waveguide when one is put on top of the other. Such a large overlap is beneficial for all kinds of evanescent coupling designs, as this speeds up the coupling between the modes. In case of an inverted taper, the large overlap leads to shorter taper lengths in order to adiabatically couple the mode.

There are two especially important parameters in the design of the taper. First of all, the DVS-BCB bonding layer thickness will determine the distance between the modes that need to be converted, and will therefore also determine the minimal taper length required for adiabatic coupling. As for the first generation of devices a manual bonding recipe was still used and therefore the DVS-BCB layer thickness was not well controlled, the taper design should be tolerant towards a thicker DVS-BCB thickness. For devices from second generation and later, the DVS-BCB layer thickness was always approximately $60\,\mathrm{nm}$ thick. A second important parameter is the alignment accuracy dx of the contact lithography process used to define the III-V membrane on top of the silicon waveguiding layer. Contact lithography has a typical alignment accuracy of about $500\,\mathrm{nm}$, and therefore for a robust design, the taper should still work well for an alignment error of about $500\,\mathrm{nm}$. When the III-V wire is not perfectly aligned, the overlap between the two fundamental modes will be lower, and therefore the taper will need

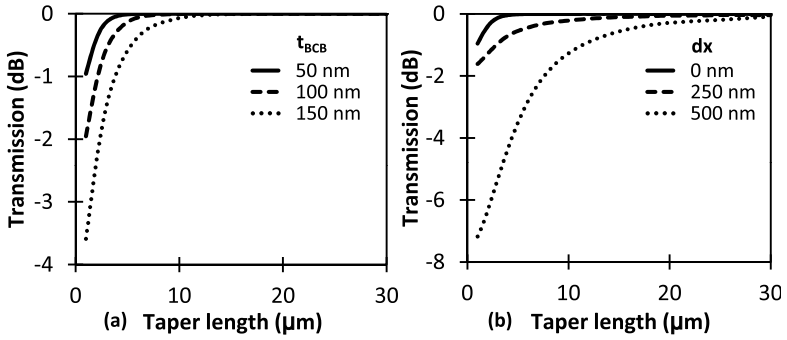


Figure 3.10: (a) Influence of the DVS-BCB thickness on the required taper length for adiabatic mode conversion ($dx = 0$). (b) Influence of the alignment accuracy on the taper length for adiabatic mode conversion ($t_{BCB} = 50\,\mathrm{nm}$).

to be longer. Furthermore, coupling to higher order modes will be more likely due to the breaking of the symmetry in the taper devices.

In figure 3.10(a), the transmission of the taper is plotted as a function of taper length, and is shown for different DVS-BCB bonding layer thicknesses. It is clear that an increasing thickness leads to a longer taper length. In figure 3.10(b), the transmission is plotted as a function of taper length, but is shown in this case for different alignment errors. For a taper length of $12\,\mu\mathrm{m}$, the mode conversion works well up to an alignment error of $250\,\mathrm{nm}$ and a BCB thickness $150\,\mathrm{nm}$. However, for a length of $18\,\mu\mathrm{m}$, alignment errors up to $500\,\mathrm{nm}$ can be tolerated. Both these cases will be used in the fabricated design.

3.3.2 Fiber-to-chip coupling

In this work, two different fiber-to-chip coupling strategies are being used. Grating couplers are the most convenient way of coupling light. However for some applications, a lower insertion loss and higher bandwidth is required. In those cases, a horizontal coupling scheme can be used.

Grating couplers

Grating couplers are the most convenient way of coupling light [56]. These structures can be probed from the top of the die, and except for a shallow silicon etching step which can be done with high precision in a CMOS fab, no further post-processing is required for their operation.

Therefore, they are especially useful for process development and the measurement of test structures.

A grating coupler is actually a very simple structure. In its most basic form, it consists of a taper to a broad waveguide, that has a fundamental mode width that matches the mode width in an optical fiber, and a shallow grating etched in this broad waveguide, which scatters the light out of the waveguide. Because the grating is a periodic structure, the scattered light does not radiate away in random directions, but instead interferes constructively in a specific direction. When positioning the fiber above the grating in the right direction, the light can be picked up in the fiber. Using such grating couplers, a coupling efficiency of $30\,\%$ over a $3\mathrm{dB}$ bandwidth of $\approx 65\,\mathrm{nm}$[56] can be achieved in the standard $220\,\mathrm{nm}$ passives silicon waveguide platform, and even an efficiency of $70\,\%$ over a $3\mathrm{dB}$ bandwidth of $80\,\mathrm{nm}$ [57] in the thicker $380\,\mathrm{nm}$ advanced passives silicon waveguide platform [58].

Horizontal couplers

A second way of coupling light between fiber and chip, is by using horizontal couplers. In our group, we have developed a fabrication process for horizontal couplers based on an intermediate SU-8 polymer waveguide, which is a commonly used method in literature [59]. In this coupling scheme, a lensed fiber is brought in close proximity to a cleaved facet of the SU-8 polymer waveguide. The polymer waveguide has a size of $\approx 3\,\mu\mathrm{m} \times 2.4\,\mu\mathrm{m}$, which has a fundamental mode that matches well with the mode size in the focus of the lensed fiber. Then, the light in the SU-8 waveguide is coupled to a silicon wire using an inverted taper coupler. Because of the low effective index of the fundamental mode in the SU-8 waveguide $n_{eff} \approx 1.5$, it is very important that the silicon tip in the inverted taper coupler is as narrow as possible. In figure 3.11, the transmission and reflection of the transition between a SU-8 waveguide and SU-8/silicon is plotted as a function of silicon tip width. The tip width should therefore be less than $150\,\mathrm{nm}$ to avoid losses. As the process allows tips of the order of $100\,\mathrm{nm}$, the SU-8 taper should perform very well.

3.4 Device simulator

In order to get a deeper understanding of the physical mechanisms involved in the static and dynamic behaviour of optically pumped mem-

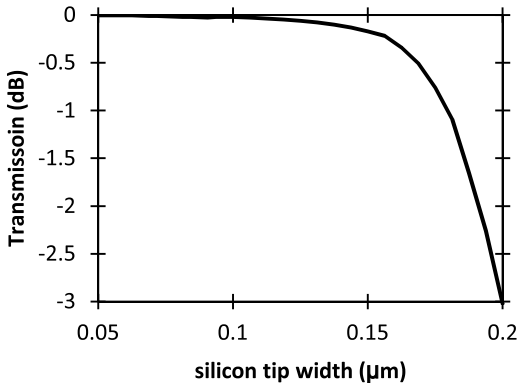


Figure 3.11: Influence of the silicon tip width on the SU-8 taper performance.

branes, a rate equation model was constructed. In the following, the considered physical mechanisms and their rate equation representation will be presented.

3.4.1 Rate equations

For the simulation of the free carrier concentration, we consider the following rate equation:

$$\frac{\partial N}{\partial t} = \int \frac{\widetilde{P}_{abs}(\lambda, \vec{r}, t)}{h\nu} d\lambda - \frac{N(\vec{r}, t)}{\tau_c} \tag{3.4}$$

$$- BN(\vec{r}, t)^2 - CN(\vec{r}, t)^3 + D\nabla^2 N(\vec{r}, t) \tag{3.5}$$

In this equation, the term on the left is the rate of change in time t of the free carrier concentration $N(\vec{r}, t)$ in the QWs. The first term on the right hand side corresponds to the rate of change of the free carrier concentration due to generation by the absorption of optical power: the absorbed power in the QWs per unit volume $\widetilde{P}_{abs}(\lambda, \vec{r}, t)$ is converted to a number of photons by division by the photon energy $h\nu$, and integrated over all wavelengths to find the total generation rate at position $\vec{r}$ and time t. The second term is the rate of non-radiative recombination due to traps and defects with a non-radiative lifetime τ_c (see section A.4.2). In the third term, the rate of radiative recombination is given, where B is the radiative recombination coefficient (see section A.4.1). The fourth term corresponds to the rate of Auger recombination where C is the Auger recombination coefficient (see section A.4.2).

The last term governs carrier diffusion, where D is the effective diffusion constant experienced by electrons and holes (see section A.3.1). The boundary conditions are governed by surface recombination (see section A.4.2).

This equation can be readily simplified however. Because of the narrow devices used, we can assume that due to carrier diffusion in the lateral direction a nearly uniform carrier density is achieved in each QW. Furthermore, the close proximity of the QWs leads to a similar modal overlap of the pump beam with each QW, which allows us to assume that the carrier density in all QWs will be the same. To include the influence of surface recombition at the device edges, the non-radiative carrier lifetime τ_c is replaced by the effective non-radiative carrier lifetime τ_{eff}, which can be determined by calculating the diffusion and recombination of carriers in the membrane cross section a single time, using an initial carrier density with the shape of the mode profile inside the membrane. By assuming that the variation in the carrier density is spatially slow enough along the device (z), we can neglect the influence of carrier diffusion in this direction as well. Next to this, the pump and signal beams can be considered as perfect monochromatic sources, which allows us to approximate the integral in equation 3.5 as a sum of two terms. This all leads to the following:

$$\frac{\partial N}{\partial t} = \frac{P_{abs,\lambda_p}(z,t)}{h\nu_p A_{QW}} + \frac{P_{abs,\lambda_s}(z,t)}{h\nu_s A_{QW}} - \frac{N(z,t)}{\tau_{eff}} - BN(z,t)^2 - CN(z,t)^3 \quad (3.6)$$

Optically, we consider material absorption and gain in the QWs, FCA and scattering. We consider two different equations for the pump and signal beams which have a wavelength of respectively λ_p and λ_s. Next to this, we assume that the concentration of free carriers in the barriers is negligible compared to the concentration in the QWs. Furthermore, as the carrier density is assumed laterally uniform in the QWs, also the absorption and gain are only dependent on one spatial direction z, leading to the following equation

$$\frac{\partial P_\lambda(z,t)}{\partial z} = \left(-\Gamma_{QW}\alpha_{mat}(\lambda,z,t) - \Gamma_{QW}\alpha_{FCA}(z,t) - \alpha_{scat}\right) P_\lambda(z,t) \quad (3.7)$$

where Γ_{QW} is the optical confinement of the optical mode in the QWs. From the solution of this equation, $\widetilde{P}_{abs,\lambda}$ can be determined as follows

$$\widetilde{P}_{abs,\lambda}(z,t) = -\Gamma_{QW}\alpha_{mat}P(z,t) \tag{3.8}$$

The material absorption in the QWs used in the simulations is either based on a simulated QW response using the same software and methodology as in section 2.2.2, or is the following linear model fitted to experimental data

$$\alpha_{mat}(\lambda,z,t) = \alpha_0(\lambda)\left(1 - \frac{N(z,t)}{N_t}\right) \tag{3.9}$$

with α_0 the material absorption obtained when the device is not pumped, and N_t the transparency carrier density. The free carrier absorption is calculated using

$$\alpha_{FCA}(z,t) = (k_{FCA,n} + k_{FCA,p})N(z,t) \tag{3.10}$$

where we assume an equal amount of free holes and electrons, with $k_{FCA,n}$ and $k_{FCA,p}$ the FCA coefficients in the QW material for electrons and holes respectively.

3.4.2 Numerical integration

The differential equations are not strongly coupled, and therefore can be easily solved by differentiating in both space (z) and time (t). Associated with the grid points z where the power is calculated, also the sections of length Δz in between these points are numbered k. In these sections, the carrier concentration and absorption is calculated.

For each time step first the material absorption and free carrier absorption for signal and pump beams is updated, using the previously calculated carrier concentrations $N(z,t)$. Then the analytical solution of equation 3.7 is used to calculate the optical power at each location z:

$$P_{\lambda_p}(z+\Delta z,t) = P_{\lambda_p}(z)e^{\left(\Gamma_{QW}\alpha_{mat}(\lambda_p,z,t)+\Gamma_{QW}\alpha_{FCA}(z,t)+\alpha_{scat}\right)\Delta z} \tag{3.11}$$

$$P_{\lambda_s}(z+\Delta z,t) = P_{\lambda_s}(z)e^{\left(\Gamma_{QW}\alpha_{mat}(\lambda_s,z,t)+\Gamma_{QW}\alpha_{FCA}(z,t)+\alpha_{scat}\right)\Delta z} \tag{3.12}$$

From these, also the absorbed power in each grid section k can be calculated:

$$P_{abs,\lambda_p}(k,t) = -\frac{\Gamma_{QW}\alpha_{mat}(\lambda_p,k,t)}{\Gamma_{QW}\alpha_{mat}(\lambda_s,k,t)+\Gamma_{QW}\alpha_{FCA}(k,t)+\alpha_{scat}} \cdot \left(P_{\lambda_p}(z_k+\Delta z)-P_{\lambda_p}(z_k)\right) \tag{3.13}$$

$$P_{abs,\lambda_s}(k,t) = -\frac{\Gamma_{QW}\alpha_{mat}(\lambda_s,k,t)}{\Gamma_{QW}\alpha_{mat}(\lambda_s,k,t)+\Gamma_{QW}\alpha_{FCA}(k,t)+\alpha_{scat}} \cdot \left(P_{\lambda_s}(z_k+\Delta z)-P_{\lambda_s}(z_k)\right) \tag{3.14}$$

Finally, the carrier equation is solved using the fifth order Adams-Bashforth method:

$$f_N(k,t) = \frac{P_{abs,\lambda_p}(k,t)}{h\nu_p A_{QW}\Delta z} + \frac{P_{abs,\lambda_s}(k,t)}{h\nu_s A_{QW}\Delta z} - \frac{N(k,t)}{\tau_{eff}} - BN^2(k,t)) - CN^3(k,t) \tag{3.15}$$

$$\begin{aligned} N(k,t+\Delta t) = N(k,t) &+ \Delta t\frac{1901}{720}f_N(k,t) - \Delta t\frac{1387}{360}f_N(k,t-\Delta t) \\ &+ \Delta t\frac{109}{30}f_N(k,t-2\Delta t) - \Delta t\frac{637}{360}f_N(k,t-3\Delta t) \\ &+ \Delta t\frac{251}{720}f_N(k,t-4\Delta t) \end{aligned} \tag{3.16}$$

The initial value at $t = 0$ of $N(k,0)$, $P_{\lambda_p}(z,0)$, and $P_{\lambda_s}(z,0)$ is a matrix of zeros. When the pump and signal are turned on at a time t_{on}, the boundary condition at $z = 0$ for $P_{\lambda_p}(0, t \geq t_{on})$ and $P_{\lambda_s}(0, t \geq t_{on})$ is changed to the pump and signal power P_{pump} and P_{signal} respectively.

Two important parameters that need to be set are the time discretization Δt and space discretization Δz. Because the time trace is calculated explicitly, these should satisfy the Courant-Friedrichs-Lewy stability condition [60], to make sure the time trace is accurate. This condition states that Δt should be shorter than the time required for the optical beam to travel between two adjacent space discretization points. To satisfy this condition, a Δz of $5\,\mu\mathrm{m}$ is used, in combination with a Δt of $25\,\mathrm{fs}$. For the simulation of static curves, the same algorithm is used, until the steady state solutions are reached. In this case, the changes as a function of time are not important: only the convergence of the algorithm matters. Therefore a smaller spatial discretization is used in order to improve the accuracy of the static solution, in combination with a larger time step to keep simulation time at bay. In this case a Δz of $2\,\mu\mathrm{m}$ and a Δt of $100\,\mathrm{fs}$ are used.

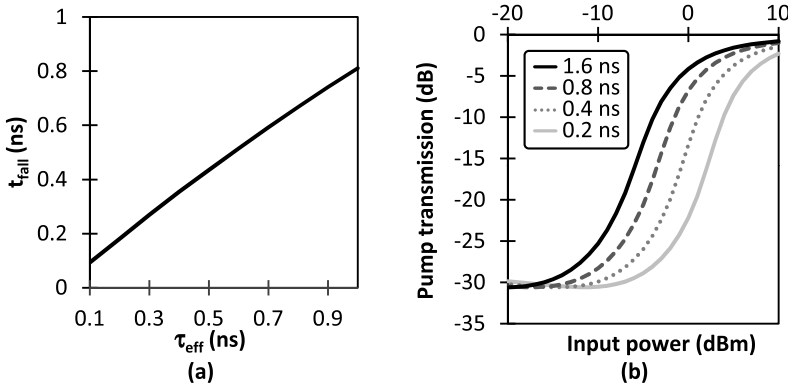


Figure 3.12: (a) Influence of τ_{eff} on the device fall time t_{fall}. (b) Influence of τ_{eff} on the pumping efficiency of the device.

3.5 Device speed and efficiency

Two parameters in particular determine the device speed and have in this way also a strong influence on the pumping efficiency. In the following, first the speed limiting factor in the device dynamics will be determined using the simulation engine as described in section 3.4.

3.5.1 Engineering the device speed

In figure 3.12(a) the device fall time t_{fall} is plotted as a function of τ_{eff} and in figure 3.12(b) the static device transmission is plotted for a changing τ_{eff}. In these simulations, the device consists of a MQW structure that contains three $8\,\mathrm{nm}$ thick InGaAs QWs separated by InP barriers, with a device width of $1\,\mu\mathrm{m}$ and a length of $100\,\mu\mathrm{m}$. The QW response was calculated using Silvaco as in section 2.2.2. It is clear that the device speed is limited by τ_{eff}: by reducing τ_{eff}, t_{fall} can be proportionally reduced. However, there is a drawback to a faster device. In order to pump the device, the amount of free carriers that needs to be present in the QWs remains the same. So, the amount of energy that needs to be delivered during a time τ_{eff} is also constant. Therefore, when making the device k times faster by reducing the lifetime to τ_{eff}/k, the power consumption will also go up with a factor of k. The conclusion is that we should try to engineer the carrier lifetime in the QWs, such that the speed of switching is just fast enough for the application.

The non-radiative recombination described by the effective non-radiative lifetime is the result of two different physical effects: Shockley-Read-Hall recombination (see section A.4.2) and surface recombination (see section A.4.2). Shockley-Read-Hall is present in the complete QW structure, and therefore determines the upper limit for the effective non-radiative lifetime. It is a measure for the semiconductor quality of the QW structure, as the recombination velocity scales with the number of defects present in the QW structure. Surface recombination on the other hand is only present at etched surfaces of the QW, and is described using the surface recombination velocity v_s. This velocity is specific for the etched QW interface, and can be changed to higher and lower values using the appropriate fabrication techniques.

Calculation

To investigate the influence of the membrane width and surface recombination velocity on τ_{eff}, the carrier diffusion along the device width is calculated explicitly in the device cross section using the following equations and boundary conditions:

$$\begin{aligned} \frac{\partial N(x,t)}{\partial t} &= G(x,t) + D\frac{\partial^2 N(x,t)}{\partial x^2} - \frac{N(x,t)}{\tau_c} \\ G(x,t) &= H(-t)G_0(x) = H(-t)\frac{N_{gen}}{\sqrt{\pi}w_{mode}}\exp\left[-\frac{x^2}{w_{mode}^2}\right] \\ \frac{\partial N(0,t)}{\partial x} &= 0 \\ -v_s N(w/2,t) &= D\frac{\partial N(w/2,t)}{\partial x} \end{aligned} \tag{3.17}$$

Here, $G(x,t)$ represents the carrier generation by absorption of a pump beam. It is modeled as a Gaussian $G_0(x)$ fitted to the intensity profile of the fundamental mode in the membrane waveguide with w_{mode} the $1/e$-width of the profile, multiplied with the heaviside function $H(-t)$, which means that the pump that generates the carriers is turned off for $t \geq 0$. As there is a steady state situation for $t < 0$, $N(x,0)$ is equal to the static solution of equation 3.17, which is easily obtained by putting all derivatives to t to zero. The equations are solved for $w/2 \geq x \geq 0$, as the problem is symmetric around $x = 0$.

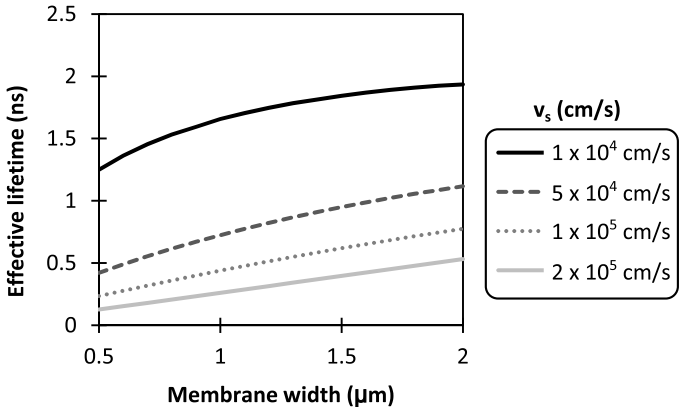


Figure 3.13: Influence of the surface recombination velocity v_s and the device width w on the effective carrier lifetime.

To calculate the effect of the decreasing carrier density $N(x, t > 0)$ on the mode in the membrane waveguide, the overlap of the mode with the carrier density needs to be considered:

$$N_{eff}(t) = \frac{\int N(x,t)G_0(x)dx}{\int G_0(x)dx} \tag{3.18}$$

This 'effective' carrier density can then be fitted to an exponential decay, to obtain the effective non-radiative carrier lifetime.

Adjusting the effective carrier lifetime

Using the above, we can investigate the influence of the device width and surface recombination velocity on the effective carrier lifetime. In figure 3.13, the effective carrier lifetime τ_{eff} is plotted as a function of the width, for different values of the surface recombination velocity v_s. The non-radiative carrier lifetime τ_c inside the QW was set at $2.5\,\mathrm{ns}$, and forms the upper limit for the effective carrier lifetime. The width was swept from $500\,\mathrm{nm}$ to $2\,\mu\mathrm{m}$, while v_s was set at values ranging from $1 \times 10^4\,\mathrm{cm/s}$, which corresponds to a well passivated interface [61], to $2 \times 10^5\,\mathrm{cm/s}$ which corresponds to an interface with a lot of dangling bonds. Both parameters have a clear influence on the value of τ_{eff}, however the effect of v_s is dominant. For large values of v_s, the non-radiative recombination is dominated by surface recombination, and very short effective carrier lifetimes can be obtained. On the other hand, if the QW interface is passivated, surface recombination becomes

less important and the effective carrier lifetime is therefore closer to the value for the unetched MQW stack. In the limit of $v_s \rightarrow 0\,\mathrm{cm/s}$, which is the case for a perfectly passivated surface, τ_{eff} is equal to τ_c for any device width. Next to this, also the width plays an important role. When the device is narrower, the distance free carriers have to travel to reach the etched surfaces of the QWs is shorter. As the non-radiative recombination is enhanced at these interfaces, the effective carrier lifetime is shorter for narrower devices.

A second way to reduce the effective carrier lifetime, is by increasing the importance of SRH recombination. This can be achieved by damaging the QWs by irradiating the device with highly energetic ions [62, 63]. These can introduce new traps and defects in the semiconductor lattice that can act as efficient recombination centers, and thereby reduce the non-radiative carrier lifetime τ_c. As τ_c forms the upper limit for τ_{eff}, reducing τ_c will also proportionally reduce τ_{eff}.

Measuring the effective lifetime

In order to investigate the different techniques to adjust τ_{eff}, we need a reliable technique to measure this parameter. Assuming the absorption α in the device under investigation with length l is uniform, the measured output power for an injected probe beam can be expressed as

$$P(t) = P_0 \exp\left[-\alpha(t) l\right] \tag{3.19}$$

Using the expression for the material absorption in equation 3.9, and assuming losses due to free carrier absorption and scattering are negligible, we obtain

$$P(t) = P_0 \exp\left[-\Gamma_{qw}\alpha_0 \left(1 - N(t)/N_t\right) l\right] \tag{3.20}$$

This approximation is only valid for wavelengths sufficiently below the band gap wavelength, as then the quantum well absorption is the dominant loss mechanism. The most convenient way to bring the effective carrier lifetime in this equation, is by using an impulse response for the carrier density. This can be done by pumping the device using a femtosecond modelocked laser with sufficient energy in each pulse to bleach the device under investigation. Then, the carrier density is nearly instantaneously changed to a value (N_0) of the order of the transparency carrier density, while at the same time the pump light

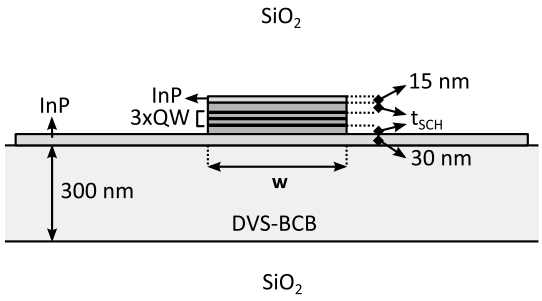


Figure 3.14: Schematic of waveguide that will be simulated.

is nearly immediately turned off. Assuming such a pumping source, the transmitted power becomes

$$P(t) = P_0 \exp\left[-\Gamma_{qw}\alpha_0 \left(1 - N_0/N_t \exp(-t/\tau_{eff})\right) l\right] \tag{3.21}$$

$$= P_0 \exp\left[-\Gamma_{qw}\alpha_0 l\right] \cdot \exp\left[\Gamma_{qw}\alpha_0 N_0/N_t l \cdot \exp(-t/\tau_{eff})\right] \tag{3.22}$$

$$= C_1 \cdot \exp\left[C_2 \cdot \exp(-t/\tau_{eff})\right] \tag{3.23}$$

where we assume that single carrier non-radiative recombination is the dominant recombination process. To determine the effective carrier lifetime, the following needs to be calculated from the measured $P(t)$-trace:

$$\ln\left[\ln\left(P(t)\right) - \ln(C_1)\right] = \ln\left(C_2\right) - t/\tau_{eff} \tag{3.24}$$

where $\ln(C_1)$ can be determined by averaging the calculated values for $\ln\left(P(t)\right)$ before the excitation of the device by the pulse. By linearly fitting the resulting data from the calculation in the left hand side of equation (3.24), a value for τ_{eff} can be obtained. If the obtained data is not linear, also higher order carrier interactions (radiative recombination and Auger recombination) need to be considered.

3.5.2 Optimizing the device efficiency

Until now, an arbitrarily chosen device width of 1.5 or $2\,\mu\mathrm{m}$ was assumed (see section 3.2.1). However, also this parameter can be optimized. In figure 3.14 the simulated waveguide structure is shown, with free parameter w. For this simulation, t_{SCH} is set to zero. The triple QW structure uses $8\,\mathrm{nm}$ InGaAs QWs separated by $10\,\mathrm{nm}$ InP barrier layers. In figure 3.15(a) the optical confinement is plotted as a function of the

membrane width w. As expected, a narrower membrane waveguide has a worse confinement than a broader waveguide, and it is found that the QW confinement monotonically decreases with decreasing device width. Therefore, to obtain similar device characteristics, narrower membrane waveguides should be longer:

$$\Gamma_{QW,n} l_n = \Gamma_{QW,b} l_b \tag{3.25}$$

However, it seems that the confinement does not scale up linearly with the device width. Instead, it has a nearly constant value for a membrane with widths broader than $1\,\mu\mathrm{m}$, and decreases rapidly when going to device widths narrower than $700\,\mathrm{nm}$.

On the other hand, the total QW volume of a device scales linearly with the device width. Therefore, assuming perfectly passivated QW surfaces as discussed in section 3.5.1, also the amount of power required to pump the device scales with the width, leading to the following for devices of equal length

$$\frac{P_n}{w_n} = \frac{P_b}{w_b} \tag{3.26}$$

If the surfaces are not perfectly passivated however, the carrier lifetime in the narrower device will be shorter than in the broader device. As the carrier density in the devices should be the same to obtain the same device response, the power density Q delivered during a time τ_{eff} to the device should be constant:

$$\tau_{eff,n} Q_n = \tau_{eff,b} Q_b \tag{3.27}$$

Integrating this power density over the entire device leads to

$$\frac{\tau_{eff,n}}{w_n} P_n = \frac{\tau_{eff,b}}{w_b} P_b \tag{3.28}$$

Furthermore, to obtain the same output characteristics the device should be longer when the width is narrower as the confinement of the signal beam is lower, and therefore using equation 3.25, we obtain

$$\begin{aligned} \frac{\Gamma_{QW,n}\tau_{eff,n}}{w_n} P_n &= \frac{\Gamma_{QW,b}\tau_{eff,b}}{w_b} P_b = C \\ P_n &\propto \frac{w_n}{\Gamma_{QW,n}\tau_{eff,n}} \end{aligned} \tag{3.29}$$

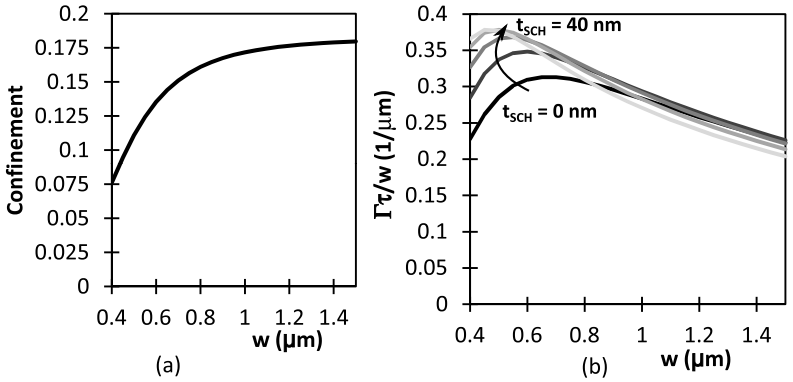


Figure 3.15: (a) Confinement as a function of membrane width w. (b) Simulated efficiency $\Gamma\tau/w$ as a function of the membrane width w and SCH layer thickness t_{SCH} around the triple QW structure for $v_s = 1 \times 10^4\,\mathrm{cm/s}$.

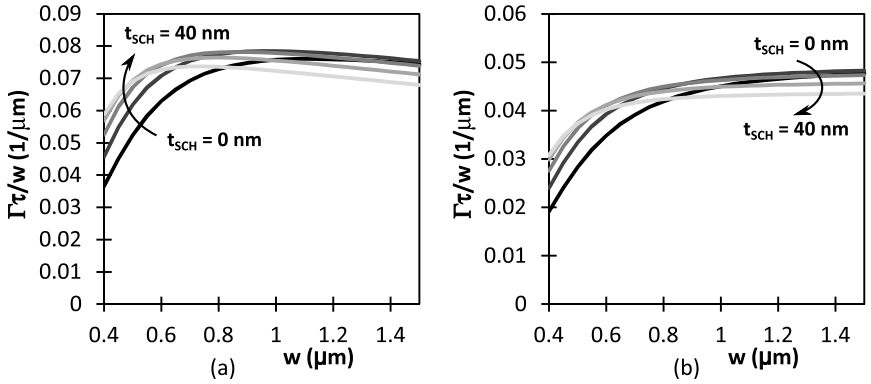


Figure 3.16: Simulated efficiency $\Gamma\tau/w$ as a function of the membrane width w and SCH layer thickness t_{SCH} around the triple QW structure (a) for $v_s = 1 \times 10^5\,\mathrm{cm/s}$, and (b) for $v_s = 2 \times 10^5\,\mathrm{cm/s}$.

To minimize power consumption, the device design should be optimized to maximize the ratio $\frac{\Gamma_{QW}\tau_{eff}}{w}$. As discussed above, the confinement Γ_{QW} will decrease with decreasing width w. However, this can be partially compensated by using a slightly thicker membrane with a higher index barrier material (InGaAsP(Q1.2)) to increase the effective index of the waveguide core. Therefore, the device scheme in figure 3.14 is slightly adapted to a MQW structure with $8\,\mathrm{nm}$ QWs separated by $10\,\mathrm{nm}$ thick InGaAsP(Q1.2) barriers. In the simulation, both w and t_{SCH} are free parameters to be optimized. In figure 3.15(b) the ratio $\frac{\Gamma_{QW}\tau_{eff}}{w}$ is plotted for different SCH layer thicknesses, assuming a good passivation with $v_s = 1 \times 10^4\,\mathrm{cm/s}$ and $\tau_c = 2.5\,\mathrm{ns}$. From this we can conclude that for a passivated device, narrower waveguides are nearly always better: reducing the width from $1.5\,\mu\mathrm{m}$ to $750\,\mathrm{nm}$ improves the efficiency with $\approx 50\,\%$ for any choice of the SCH layer thickness. In figure 3.16(a) and 3.16(b) the same simulation is performed for respectively $v_s = 1 \times 10^5\,\mathrm{cm/s}$ and $v_s = 2 \times 10^5\,\mathrm{cm/s}$. In the unpassivated case, the situation is different: for $v_s = 1 \times 10^5\,\mathrm{cm/s}$, the efficiency can be maximized by using an $800 - 1000\,\mathrm{nm}$ broad device, however the efficiency gains that can be obtained are marginal. For $v_s = 2 \times 10^5\,\mathrm{cm/s}$, the negative influence of the surface recombination on the carrier lifetime is so pronounced, that devices become more efficient as the membrane becomes broader, even though a larger QW volume needs to be pumped.

3.6 Fabrication

3.6.1 Preparation and bonding

Fabrication of the SOI waveguide circuits was done using a $193\,\mathrm{nm}$ deep UV lithography stepper and dry etching on a $200\,\mathrm{mm}$ wafer in a CMOS pilot-line [58]. After dicing of the SOI wafer, the separate dies are cleaned in a standard clean-1 (SC-1: $1NH_3$:$1H_2O_2$:$5H_2O$) solution for 15 minutes at $70 - 80\,^\circ\mathrm{C}$. In the mean time, the III-V dies are prepared for bonding by removal of the InP/InGaAs sacrificial layer pair, using pure HCl($37\,\%$) and a $1H_2SO_4$:$1H_2O_2$:$18H_2O$ solution respectively. It is important to let the $1H_2SO_4$:$1H_2O_2$:$18H_2O$ solution cool down to $< 23^\circ\mathrm{C}$ to achieve an etching rate of $\approx 300\,\mathrm{nm/min}$. To improve adhesion to the DVS-BCB adhesive layer, a $15\,\mathrm{nm}$ thick layer of SiO_2 is deposited on the III-V dies using PECVD. After the SC-1 treatment, the SOI dies are rinsed with DI water, blow dried, and left for

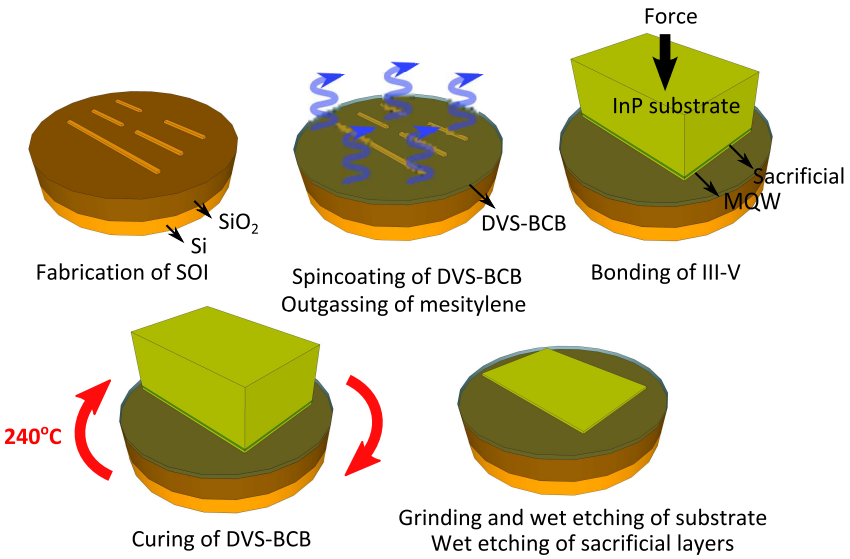


Figure 3.17: Fabrication recipe: preparation and bonding.

5 minutes on a $150\,^{\circ}\mathrm{C}$ hot plate to dehydrate the SOI surface. After this, the adhesion promoter AP3000 is spincoated on the SOI dies at $3000\,\mathrm{rpm}$, immediately followed by the spinning of a DVS-BCB solution diluted with mesitylene. Depending on the SOI topography, DVS-BCB dilutions ranging from 1:5 to 1:1 have been used in this work. This large variation is mainly related to the inclusion of tiling structures in the later silicon designs by Epixfab, which increased the volume that needs to be filled by BCB considerably [58]. After the spincoating, the samples are baked at $150\,^{\circ}\mathrm{C}$ to evaporate the mesitylene. Then, the dies are brought into close contact in a controlled environment using a bonding pressure of $1.25\,\mathrm{MPa}$ at a temperature of $150\,^{\circ}\mathrm{C}$. To cure the DVS-BCB, the temperature is slowly ramped ($1.6\,^{\circ}\mathrm{C}/\mathrm{min}$) to $240\,^{\circ}\mathrm{C}$ and kept at this temperature for 1 hour. After the bonding of the dies, the InP substrate is removed using a combination of mechanical grinding and chemical etching using pure HCl($37\,\%$) until the InGaAs etch stop layer is reached. Subsequently the etch stop layer is removed using a $1H_2SO_4$:$1H_2O_2$:$18H_2O$ solution.

3.6.2 Patterning and etching of the membrane

Several patterning possibilities were explored. The general process is shown schematically in figure 3.18.

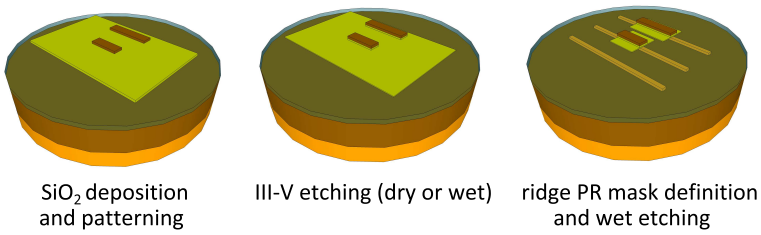


Figure 3.18: Fabrication recipe: patterning and etching of the membrane.

Dry etching

First a $100\,\text{nm}$ thick SiO_2 hard mask is deposited. This hard mask is patterned using contact lithography to create the core of the rib waveguide. After dry etching the hard mask using RIE(SF_6($15\,\text{sccm}$):O_2($50\,\text{sccm}$)), the bonded III-V film is etched using either the RIE or ICP. In both cases, a CH_4:H_2 gas mixture is used to etch the III-V, alternated with a O_2 plasma to remove deposited hydrocarbons during the etching. The main difference is that the ICP recipe runs in chamber with temperature of $100\,^\circ\text{C}$, while the RIE runs in a chamber with temperature of $20\,^\circ\text{C}$. In case a ridge waveguide is created, the III-V is etched until $20 - 40\,\text{nm}$ remains in the etched parts. Afterwards, another contact lithography step is performed to define the larger III-V islands and the remaining III-V is etched using a $NaClO_3$($0.3\,\text{g}$):HCl($9\,\text{ml}$):CH_3COOH($15\,\text{ml}$):H_2O($90\,\text{ml}$) solution, which is an isotropic etch mixture for InP/InGaAsP. In case a stripe is created, the III-V is etched completely by dry etching.

Wet etching

Here a resist mask is defined using contact lithography to create the core of the rib waveguide. Then, the core of the waveguide is defined by subsequent etching using $3HCl$:$2H_2O$ and $1H_2SO_4$:$1H_2O_2$:$18H_2O$, to remove the InP and InGaAsP layers. In case a rib waveguide is created, the final InP layer is etched with HCl using a larger resist mask to define the III-V island.

3.6.3 Passivation

As mentioned in section 3.5.1, an important parameter that determines the device efficiency is surface recombination velocity. This recombina-

tion velocity is associated with the quality of the etched QW surface: if a lot of dangling bonds are present, a high recombination rate can be expected, and vice versa. Therefore, usually an extra passivation step is added to the fabrication procedure, which has the goal of reducing the amount of dangling bonds. For the passivation of InGaAs/GaAs type semiconductors, the best results so far are obtained using a treatment based on free sulfur ions [64, 65, 66, 67, 5]. In this passivation scheme, the sample is in most cases first prepared by a short wet etching treatment to remove the semiconductor native oxides from the surface. Subsequently, the sample is put into a hot solution of ammoniumpolysulfide ($(NH_4)_2S_x$, $x > 1$), which is made by dissolving pure sulfur in commercially available ammonium(mono)sulfide ($(NH_4)_2S$) solutions. After the treatment, the amorphous sulfur film that remains on the sample is removed by a DI-water rinse, or by sublimation in a vacuum, and the treated surface is immediately covered by a cap layer to avoid reoxidation. This treatment has proven to be successful for surface preparation before metalization [64, 65], the reduction of the dark current in photodiodes [66, 67], and increasing the non-radiative carrier lifetime in thin film devices with etched through QWs [5].

Although there are a lot of publications available on the topic, most are not clear on the exact methodology. In some publications, e.g. in [66], the concentration of the ammoniumsulfide is not mentioned. In other publications, e.g. in [67], it is written that a ammoniumsulfide solution of a certain concentration is used, but the concentration of the additional elemental sulfur is not mentioned. Furthermore, the reaction time, temperature and solution concentration vary across the different papers covering the subject. Therefore, it is not a trivial task to reproduce the results reported in these papers.

In [5], the exact same application is pursued: increasing the non-radiative carrier lifetime of a membrane device. In this paper, this is even more important due to the use of a photon crystal defined in the membrane. This means that the membrane is pierced at regular intervals and that the distance for a free carrier to travel to a etched QW surface is therefore only a few hundreds of nanometers. They managed to increase the carrier lifetime to $2.2\,\mathrm{ns}$, using the following treatment. After dry etching, first the QW surfaces are cleaned, by dipping the sample for 10 seconds in a $1H_2SO_4$:$8H_2O_2$:$5000H_2O$ solution. The reason such a high dilution is used, is because the size of the holes etched into the membrane is critical. It is assumed that this etching step removes the native oxides and the outer damaged layers due to the dry

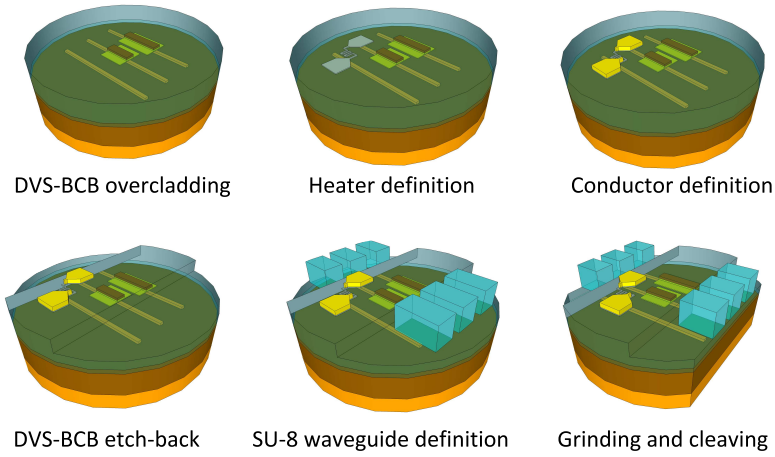


Figure 3.19: Fabrication recipe: heaters and couplers.

etching. After this, the sample is put in a unspecified $(NH_4)_2S_x$ solution at a temperature of $45\,^{\circ}\mathrm{C}$ for 10 minutes. After this treatment, a DVS-BCB solution is immediately spin coated and cured as a capping layer to stabilize the passivation.

Several attempts were done to improve the device characteristics. The fabrication details and measurements are given in section 3.7.3.

3.6.4 Further post-processing

In principle the devices are now ready, however also heaters and horizontal couplers can still be created. The associated processing steps are schematically shown in figure 3.19.

Heaters

In case heaters are to be integrated, first a $1.2\,\mu\mathrm{m}$ thick DVS-BCB layer is spinned and cured on the chip. Then, using an image reversal lithography pattern, the heater layer that consists of a $100\,\mathrm{nm}$ titanium layer covered by $3\,\mathrm{nm}$ gold to prevent oxidation is deposited and selectively lifted off. The same process is repeated for the definition of the conductor layer, which consists of a $40\,\mathrm{nm}$ titanium layer, covered by $300\,\mathrm{nm}$ gold. In case horizontal couplers need to be defined in a later step, the DVS-BCB layer is selectively etched back $600\,\mathrm{nm}$ near the horizontal coupler positions.

Horizontal couplers

In a final step, horizontal couplers can be defined. These are created by direct photo lithographic patterning of a $2.4\,\mu\mathrm{m}$ thick UV-curable SU-8 polymer layer. After the definition, the SU-8 layer is further cured at $200\,^{\circ}\mathrm{C}$ for 10 minutes in a nitrogen environment. Then, the silicon die is thinned down by grinding to $250\,\mu\mathrm{m}$, and cleaved through the SU-8 waveguide. This cleaving step is critical, and is the main reason why there is a difference in the performance of the horizontal couplers from chip to chip.

3.7 Experimental results

A lot of different devices were fabricated in the course of this work. However, these can be classified in three generations. In this section, the results from pump-probe experiments are summarized.

3.7.1 First generation

The first generation of devices used an InGaAlAs MQW stack, that consisted of four $7\,\mathrm{nm}$ thick $Al_{0.05}Ga_{0.29}In_{0.66}As$ QWs seperated by $10\,\mathrm{nm}$ $Al_{0.13}Ga_{0.42}In_{0.45}As$ barrier layers, sandwiched between two $10\,\mathrm{nm}$ InP cladding layers. This MQW structure was etched completely through the III-V membrane using dry etching in an ICP etcher to form a strip waveguide. The exposed QW surfaces were not passivated. The device used in the measurements was $100\,\mu\mathrm{m}$ long, and had a III-V width of $2\,\mu\mathrm{m}$. Grating couplers with at peak efficiency an insertion loss of $5\,\mathrm{dB}$ were used for fiber-to-chip coupling.

Continuous wave experiment

In a first experiment, a continuous wave pump-probe measurement was performed, in order to determine the net gain as a function of pump power. In this experiment, a CW pump beam at $1545\,\mathrm{nm}$ was combined with a CW probe beam using a 1:99 coupler, and sent through the chip. At the output side, a narrow band pass filter was used to remove the pump beam and spontaneous emission from the signal, and the remaining light was measured using a power detector. By comparing the probe transmission through the device with the transmission through a reference silicon waveguide, the net gain in

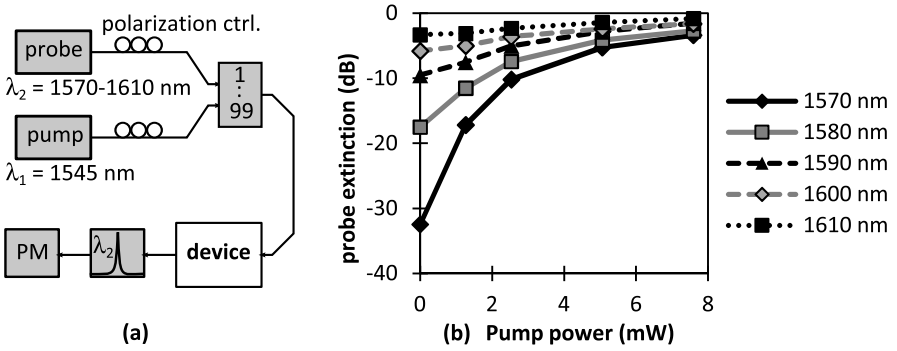


Figure 3.20: (a) Pump-probe setup. (b) Probe transmission as a function of pump power for different probe wavelengths.

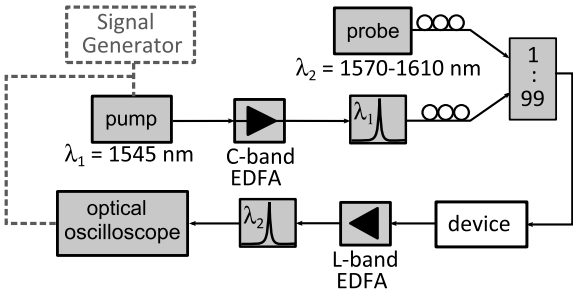


Figure 3.21: Setup of the pulsed experiment.

the III-V membrane waveguide can be extracted. A schematic of this experiment is shown in figure 3.20(a). Both the pump power and the probe wavelength are varied. The results of this experiment are shown in figure 3.20(b). Note that the pump power given in the graph is the on-chip power. When the pump power is increased, the absorption of the probe decreases as expected and eventually starts saturating. However, the point where we have net gain is not reached. This lack of net gain can be attributed to thermal effects, which degrade the device performance as discussed in section 3.2.4.

Pulsed experiment

To test our hypothesis that thermal effects cause the lack of gain, we have also performed a pulsed experiment. In this experiment, the CW pump laser is replaced by a directly modulated distributed feedback

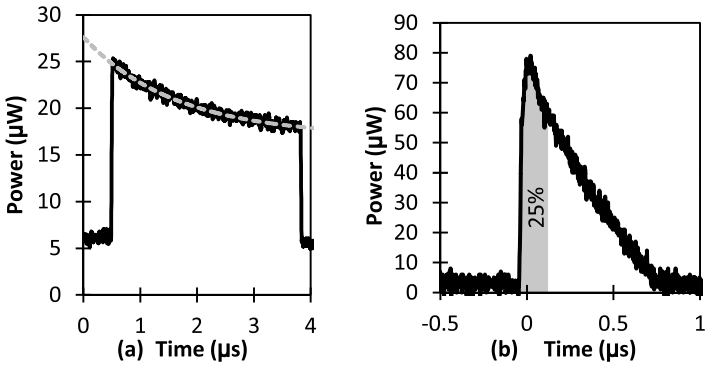


Figure 3.22: (a) Effect on the probe transmission of a $50\,\%$ duty cycle pump pulse train, with fitted exponential ($\tau \approx 1.5\,\mu\mathrm{s}$). (b) Pump pulse shape used in the experiment. The gray area shows the time during which the corresponding probe modulation is averaged for calculation of the net gain.

laser which also operates at $1545\,\mathrm{nm}$. To boost the pump power, the pump beam is sent through a C-band EDFA, and afterwards filtered by a band pass filter to avoid that the ASE emitted by the EDFA would pump the device. The pump pulse train is sent together with a CW probe beam in the device. Because of the pulsed character of the pump signal, the CW probe signal is amplified only during the duration of the pump pulse. After transmission through the device, the probe is amplified using a L-band EDFA and filtered in a band pass filter, which also removes the remaining pump. The resulting signal is then measured using a DC-coupled optical oscilloscope, which was referenced to the zero level before measurement. The setup for the measurement is schematically shown in figure 3.21.

To check the pulse duration that is required to avoid thermal effects, first a well-behaved pulse train with duty cycle of $50\,\%$, consisting of rectangularly shaped $3\,\mu\mathrm{s}$ long pulses was sent through the device, and the effect on the probe beam was measured using the oscilloscope. The trace is shown in figure 3.22(a). From this we can conclude that the thermal time constant of the device is of the order of a few $\mu\mathrm{s}$, and therefore our pump pulse should be $< 1\,\mu\mathrm{s}$ long to avoid thermal effects. It is unfortunately not possible to derive a value for the thermal time constant with high precision, as the recorded time during the thermal roll-off was too short. The pulse that was used in the experiment is shown in figure 3.22(b). The directly modulated pump laser produces nearly tri-

angularly shaped pump pulses of $\approx 0.75\,\mu s$ long, in a pulse train with a period of $10\,\mu s$. The triangular shape is caused by driving the signal generator at a duty cycle close to the limit of the equipment, however this was required to reach a short enough pulse duration. Because of the triangular shape, the net gain is extracted from the output pulse by averaging the transmitted power over the peak $100\,ns$ only, and comparing this power to the transmitted power in a reference waveguide. In figure 3.23(a) an example of a measured signal pulse is shown together with the recorded reference level. To have an idea of the pump power required in a CW scenario where thermal effects can be avoided, the pump power during the pulse was estimated by measuring the average pump power, and extracting the power during the peak $100\,ns$ of the pump pulse, considering the fact that $25\,\%$ of the pulse energy is present during the peak $100\,ns$.

The results for all wavelengths are shown in figure 3.23(b). Note that the pump power given in the graph is the on-chip power. From these results, it is clear that thermal effects must be the cause for the lack of gain in CW regime. A gain of $8\,dB$ is achieved in pulsed regime for a peak pump power of $8.5\,dBm$, while in CW regime the device does not reach transparency for a comparable input power. The peak gain lies at a wavelength of $1580\,nm$ in this case. This is to be expected, as the band edge lies around $1600\,nm$, and the device is pumped and bleached at $1545\,nm$. It seems that the gain is already saturated for the lowest used pump power at $1590\,nm$, while it is still increasing a little for the shorter wavelengths. This increase is small however, because of absorption saturation of the pump beam. To increase the gain even more, one could move to a shorter pump wavelength as the absorption for shorter wavelengths is not completely saturated yet. However, in the current setup the limited bandwidth of the grating couplers limits the separation between pump and probe wavelength.

3.7.2 Second generation

As the performance in the first generation of devices was limited by thermal effects, we tried to resolve this in the second generation by using a better thermal design. As discussed in section 3.2.4, changing the waveguide structure from a strip waveguide to a ridge waveguide can greatly reduce the thermal resistivity. Furthermore, the InGaAlAs MQW structure with four QWs was replaced by a InGaAs/InP structure with only three QWs. This means that the dissipated power per

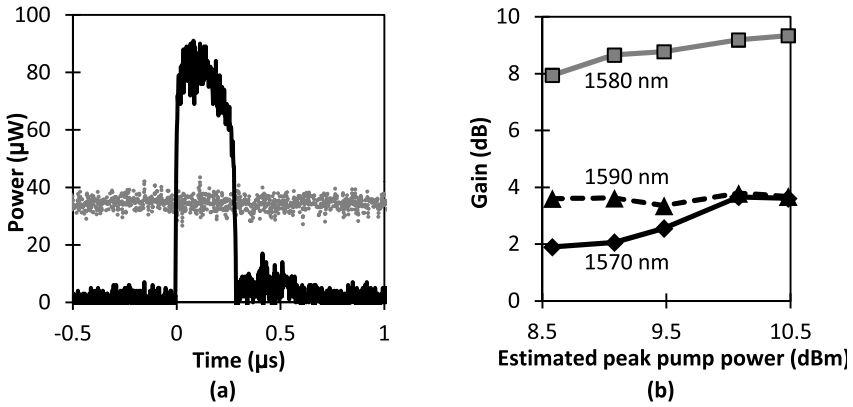


Figure 3.23: (a) Modulation in the probe signal (black) and reference trace (grey). (b) Net gain as a function of (estimated) pump pulse power.

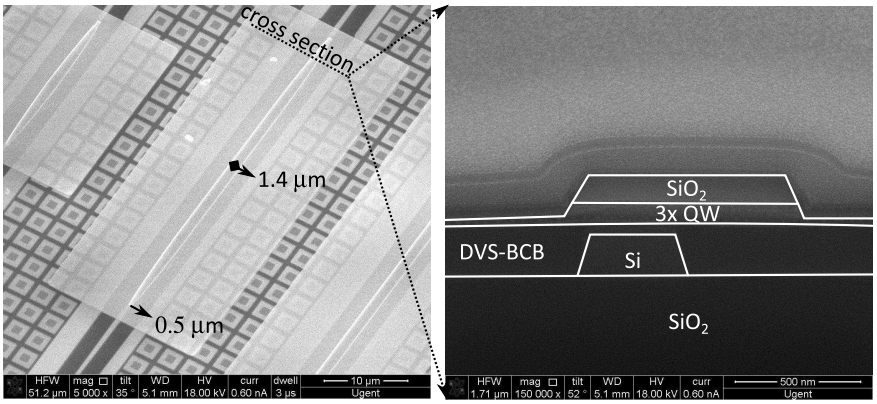


Figure 3.24: SEM micrograph of a device with cross section through the taper.

unit area decreases by $25\,\%$. Furthermore, InP performs an order of magnitude better in conducting heat than InGaAlAs, which increases the capability of the ridge waveguide to spread heat. To increase this capability even further, the InP cladding is made $25\,\mathrm{nm}$ thick. A SEM picture and cross section of this device are shown in figure 3.24.

The tested device was $150\,\mu\mathrm{m}$ long and had a core width of $1.4\,\mu\mathrm{m}$ and tapers down to a taper tip of $\approx\ 0.5\,\mu\mathrm{m}$. The core of the ridge waveguide was defined using dry etching in an RIE etcher, and leaving $\approx 20\,\mathrm{nm}$ in the shallow etched regions. The exposed QW surfaces were not passivated. Grating couplers with at peak efficiency an inser-

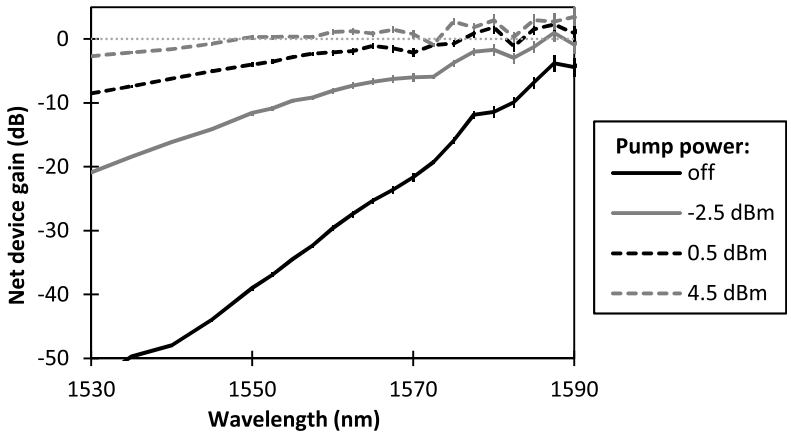


Figure 3.25: Probe transmission as a function of the probe wavelength, under different pumping conditions.

tion loss of $6.5\,\mathrm{dB}$ were used for fiber-to-chip coupling. This reduced efficiency was the result of a minor design error.

To assess the performance of this improved design, a CW pump-probe experiment was performed. The setup was similar as in figure 3.20(a), however related to the different MQW structure that was used, the pump wavelength was changed to $1505\,\mathrm{nm}$, and the filter with power meter was replaced by an optical spectrum analyzer, as this allowed us to measure automatically once the fibers had been aligned. The results are shown in figure 3.25. The error bars on this graph were obtained by calculating the standard deviation on the measurement of five different reference waveguides. These display that the error on the measurement for wavelengths $> 1570\,\mathrm{nm}$ starts increasing to the order of $1.5\,\mathrm{dB}$ around $1590\,\mathrm{nm}$, due to the fact that these measurement points lie on the falling edge of the grating coupler. This explains the wavy shape of the transmission curve near these wavelengths. From these measurements we can conclude that the improved thermal design has lead to a small net gain of a few dB in CW regime. Furthermore, only $4.5\,\mathrm{dB}$ on-chip power is required to achieve this gain. This is a clear improvement on the previous results.

3.7.3 Third generation

To improve the device characteristics, a better passivation of the etched QWs is required, as such a treatment can increase the carrier lifetime

Table 3.1: Measurements of the effective carrier lifetime, measured on $50\,\mu\mathrm{m}$ long devices.

Device width	$1.0\,\mu\mathrm{m}$	$1.5\,\mu\mathrm{m}$	$2.0\,\mu\mathrm{m}$
No passivation (reference)	443 ps	533 ps	551.6 ps
Passivation, DI rinse	138.5 ps	204.5 ps	*no data*
Passivation, vacuum bake	684.7 ps	613.1 ps	730.1 ps
Passivation, no vacuum bake (same sample as reference)	491.6 ps	578.8 ps	742.0 ps

Table 3.2: Error margin on the measurements of the effective carrier lifetime, measured on $50\,\mu\mathrm{m}$ long devices. The error margin given is the standard deviation on the carrier lifetime, resulting from the uncertainty of the fit on the raw data (see section 3.5.1 for the methodology).

Device width	$1.0\,\mu\mathrm{m}$	$1.5\,\mu\mathrm{m}$	$2.0\,\mu\mathrm{m}$
No passivation (reference)	$\pm$1.4 ps	$\pm$1.7 ps	$\pm$2.4 ps
Passivation, DI rinse	$\pm$0.6 ps	$\pm$0.6 ps	*no data*
Passivation, vacuum bake	$\pm$1.9 ps	$\pm$1.7 ps	$\pm$1.4 ps
Passivation, no vacuum bake (same sample as reference)	$\pm$1.4 ps	$\pm$1.6 ps	$\pm$1.9 ps

by nearly an order of magnitude. In all experiments, the devices were defined by RIE etching using a $CH_4 : H_2$ gas mixture. As this was attempted near the end of this work, only three samples were processed.

In a first experiment, we attempted to reproduce the results obtained in [5] as described in section 3.6.3. To do this, the recipe described in the paper was followed, where we used a $(NH_4)_2S_x$ solution that was prepared by adding $0.05\,\mathrm{g}$ of pure sulfur to $10\,\mathrm{mL}$ of a $20\,\%$ $(NH_4)_2S$ solution on a hot plate at $45\,^\circ\mathrm{C}$. This corresponds to a solution with $5\,\%$ of free sulfur. When the extra sulfur is dissolving, the solution turns from light yellow to a deep red, nearly brown color, which is expected for this solution. First a die with fabricated devices was dipped for 10 seconds in a $1H_2SO_4$:$8H_2O_2$:$5000H_2O$ solution to remove a few nanometers from the QW surface, as this surface is expected to be damaged by the dry etching. This is immediately followed by a treatment in the ammoniumsulfide solution for 10 minutes. Afterwards,

the die is rinsed with DI water to remove the deposited amorphous sulfur layer from the sample surface. The reason we do this, is because this film looks quite rough, and could lead to increased scatter losses. Subsequently the sample is dried with a nitrogen gun, and finally BCB (Cyclotene 3022-35) is spinned at 3000 rpm and cured in a nitrogen environment. Unfortunately, the carrier lifetime obtained in this way was even lower than before the passivation attempt (see results in table 3.1 and error margin in table 3.2).

After this negative result, the authors of [5] were contacted, and it was suggested that the DI rinsing step removes the passivation effect, although several publications use the DI rinse without any comment on its effect on the device performance. The above described recipe was repeated, however this time no additional sulfur was added as the authors did not mention this as well. Two samples were treated: in the first sample, the remaining amorphous sulfur film after treatment was removed by baking the sample in a vacuum environment, causing the sulfur to sublimate. Afterwards, BCB was immediately spinned and cured. In the second sample, first a reference measurement was performed. Then, the sample was treated as well, however this time BCB was spinned immediately on top of the treated surface, even though the amorphous layer was present in this sample as well. The results of this experiment are presented in table 3.1. It is clear that both samples show a similar improvement of about $10 - 20\,\%$ compared to the prior situation. We are on the right track, however further development is necessary to reach to the $2\,\mathrm{ns}$ effective carrier lifetime that should be possible.

3.8 Conclusion

In this chapter we have introduced the membrane waveguide as a component suitable for optical signal processing applications. A simulator for the device was written, and the theoretical insights required to optimize speed and efficiency of the device were derived. In the following two chapters, the membrane will be used in a high speed signal processing application as a regenerator, and in a lower speed gating application, as a component in an all-optical packet switch.

Chapter 4

Regeneration

The membrane gates that were introduced in the previous chapter have several applications. One of them is their use for all-optical signal processing. In principle, all signal processing that can be performed with an electrically pumped optical amplifier, should also be possible using the membrane device. However, because it is not trivial to achieve a high gain in these membrane devices, a signal processing application is chosen that does not require signal gain, but that instead mainly relies on the high confinement that is achieved in the active region of the membrane. In the following, first the optical regeneration application is introduced. Then, the advantages and design guidelines of the membrane for this specific application are highlighted. Finally, the experimental demonstrations are presented.

4.1 The need for optical regeneration

A regenerator is in essence a device that improves the quality of an input signal. Such a device is very useful in networks, where losses, dispersion, noise sources and jitter continuously deteriorate the SNR of a propagating signal. A typical application for an optical regenerator is therefore in long-haul fiber-optic links, where one places several regenerators along the path of propagation to make sure that the signal is still of acceptable quality at the receiver side. It is only by using this technique that e.g. high speed transatlantic optical links are made possible.

However, also on a smaller scale, this component could become very important, especially since a large research effort is currently

directed towards optical processing, where instead of electrons, photons are used in a computational process. In current electrical packet switches, the input optical signals are transfered to the electrical domain. Then, using well-established electronics, the address information is processed and the packets are transfered to the corresponding output port. At this port, the electric version of the packet is then again transfered to the optical domain. This is of course not the most efficient strategy. Although electrical switches work very well, the electro-optical transitions are an energy cost that could in principle be avoided. Furthermore, electronics are limited in speed, and electrical switches cause therefore considerable delays in the transmission of optical signals. Therefore, if an all-optical alternative to these switches could be created, gains in power consumption and switch latency could possibly be obtained.

To achieve this, several optical components will need to be designed and cascaded to altogether replace the electrical processing. However, with each cascaded component, one can expect that the signal that is propagating through this chain of components deteriorates in quality due to losses and noise addition. For this reason, optical regenerators might also be important on a small scale for the realization of complex optical processing networks.

4.2 Regeneration in literature

Regenerators are classified into three categories depending on their capabilities.

4.2.1 1R regeneration

1R regeneration corresponds to a reamplification of the input signal. Although the signal quality itself is not improved, it suffices in a lot of cases to just compensate for the propagation losses. 1R regenerators are therefore just optical amplifiers. A typical example is the use of EDFA stages in long haul communications.

4.2.2 2R regeneration

2R regeneration corresponds to a reamplification in combination with a reshaping of the input signal. In this case, also the OSNR of the signal is improved. This can be achieved in many ways.

In [68], it is proposed to insert saturable absorbers in long haul fiber links to reduce the required input power and to decrease the required channel spacing to still achieve error free transmission. This is achieved by using the highly non-linear transmission of a saturable absorber to bring down the zero-level of the input signal. In this way the accumulated ASE in the signal zero level from passing through the EDFA stages in the long haul link can be removed, increasing the signal ER and OSNR. The proposed device consists of a MQW layer stack inside a cavity, and operates in reflection using a circulator. The device is operated at $20\,\mathrm{Gbit/s}$ using RZ signal modulation, at an input power of $14\,\mathrm{dBm}$, and features an insertion loss (including circulator loss) of $8\,\mathrm{dB}$ and saturation power of $7\,\mathrm{dBm}$. The main advantage of this type of regenerator, is that it is a passive device: no electrical current or optical control is required for its operation.

In [69], it is shown that DFB lasers that are injected by an optical signal exhibit hysteresis in the transmission characteristics of the device. By injecting a CW holding beam with input power in the range of this hysteresis, the laser can be brought into a bistable state. When a signal beam is injected in this device, the laser will switch between a transmission state when the signal is sufficiently high and a blocked state when the signal is sufficiently low, leading to an improvement of the ER of the input signal. In addition to this, also the noise is strongly suppressed. Furthermore, because of the presence of the hysteresis the device actually switches between two decision levels, which means that the device can actually lower the noise floor of an input signal. Instead of having an insertion loss, this regeneration scheme additionally also reamplifies the input signal. In the experiments, the device was operated at $10\,\mathrm{Gbit/s}$ using NRZ signal modulation, at a holding beam power of $5\,\mathrm{dBm}$, signal beam power of $5 - 7\,\mathrm{dBm}$ and a laser bias current of $150\,\mathrm{mA}$. As a result, the ER of the input signal could be improved by $2\,\mathrm{dB}$, while the OSNR was improved by $4\,\mathrm{dB}$.

In [70], the same principle is applied using microdisk lasers. Because of the smaller scale however, the operating power is much lower. In this case, a signal with input power of $3.5\,\mathrm{dBm}$ is injected into the grating couplers that connect to the device, leading to an input power in the laser of less than $0\,\mathrm{dBm}$. Furthermore a bias current of only $4\,\mathrm{mA}$ was applied. An ER improvement from $13\,\mathrm{dB}$ to $15.9\,\mathrm{dB}$ is demonstrated at a bit rate of $10\,\mathrm{Gbit/s}$.

In [71], an all-active Mach-Zehnder interferometer with optical amplifiers in each arm is used as a regenerator. In this scheme, the de-

graded input signal is converted to a different wavelength determined by the CW beam in the system. In this way, a high output OSNR of approximately $40\,\mathrm{dB}$ can be achieved for an input OSNR in the range of $15 - 27\,\mathrm{dB}$. The power consumption of this regeneration scheme is high however, as it requires the driving of 6 SOAs and 1 laser diode.

Next to these demonstrations, regeneration has also been shown using self-phase modulation [72], injection locking of a laser diode [73], and four wave mixing [74].

4.2.3 3R regeneration

In 3R regeneration, in addition to the above also the signal is retimed. Therefore, these regenerators consist of a clock recovery circuit, and a way to completely retransmit the input signal. E.g. in [75], this achieved by cascading an electro-absorption modulator that converts the input wavelength to a fixed operating wavelength for the circuit using the cross-absorption modulation effect, a high speed photodiode with an electrical clock recovery circuit, and finally another electro-absorption modulator stage that reconverts the wavelength to the input wavelength and is gated at the same time using the recovered clock to achieve 3R regeneration. In principle each stage where the optical signal is converted to the electrical domain and back again, can also be considered as a 3R regenerator, as the signal is completely reconstructed.

4.3 Design of a regenerator based on membrane technology

4.3.1 Operation principle

Because of the high confinement of light in the active region of the membrane waveguide, the membrane gate is ideal to use as a saturable absorber. This high confinement actually means that there is a strong light-matter interaction. This translates here for light with photon energy above the device's band edge into a strongly non-linear transmission through the membrane waveguide. This can be used to create a regenerator, as illustrated in figure 4.1. By choosing the appropriate input power for the signal, one can make sure that the input signal '1' level (In_1) is nearly completely transmitted by the device ($Out_1 \approx 0\,\mathrm{dB}$), while the device is still strongly absorbing at the input signal '0' level

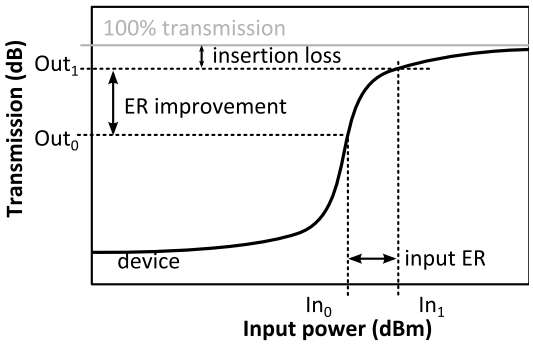


Figure 4.1: Regeneration concept: due to the non-linear transmission of the membrane device, the extinction ratio (ER) of the input signal can be improved.

($In_0 \rightarrow Out_0$), leading to an improvement of the extinction ratio by $Out_1 - Out_0$. Of course, the signal '1' level will never be completely bleached, and the remaining absorption corresponds to the insertion loss of the device. This is not a new concept [68], however existing implementations are larger scale, intended for use in fiber-optics, rather than integrated devices on chip.

4.3.2 Design guidelines

The confinement

As already stated above, the optical confinement in the QWs is an important parameter. To quantify how important the parameter is exactly, we will take a look at the equations governing the transmission through the membrane waveguide. The device absorption can be roughly modeled as

$$\alpha(N) = \Gamma_{QW}\alpha_{mat}(N) = \Gamma_{QW}\alpha_0\left(1 - \frac{N}{N_{tr}}\right) \tag{4.1}$$

The differential modal absorption $\partial\alpha/\partial N$ is therefore proportional to the confinement in the QWs. As a consequence, the device absorption will rapidly change with changing carrier concentrations. Moreover, these carriers are generated by the absorption of the signal beam itself:

$$N \propto P_{abs} \propto \exp\left(\alpha(N)\right) \tag{4.2}$$

These two combined effects lead to a strongly non-linear transmission.

Device speed

For the regeneration application, it is important that the membrane can follow the changes in the input signal. To have a useful device, the regenerator should be able to at least process signals with a bit rate of $2.5\,\mathrm{Gbit/s}$ and ideally speeds up to $10\,\mathrm{Gbit/s}$ should be obtained. As the device speed is completely limited by the effective carrier lifetime, increasing the speed corresponds to decreasing the carrier lifetime. However, as discussed in section 3.5.1, there is a trade-off between the device speed and power consumption, and therefore the design effort should be directed to reducing this lifetime in a controlled fashion, such that a device does not become too fast for the envisioned application as this comes with a higher power consumption.

To decrease the effective carrier lifetime, the non-radiative recombination rate should be increased. One way to do this, is by damaging the QW surfaces as much as possible, to create a maximal amount of dangling bonds that will act as efficient surface recombination centers. Next to this, these surfaces can be brought closer to the optical mode by moving to a narrower waveguide. This is also clear from figure 3.13.

Next to this, also the non-radiative carrier lifetime in the bulk of the QW can be reduced by irradiating the sample with highly energetic ions.

4.4 Simulation of regeneration performance

4.4.1 Methodology

To assess the potential regeneration performance of the membrane devices, the device simulator as described in section 3.4 was used to simulate the performance that can be obtained. To have a more realistic view of the possible device performance, the absorption characteristics used by the simulator were based on measured absorption data of a membrane waveguide that uses a triple QW stack that contains $8\,\mathrm{nm}$ InGaAs QWs with a band gap wavelength of $\approx 1585\,\mathrm{nm}$, separated by $10\,\mathrm{nm}$ InP barrier layers. This measurement data was fitted using the absorption model in equation (3.9). The absorption at low carrier density $\alpha_0(\lambda)$ was determined using the cut-back method commonly used to determine waveguide losses. In this case the device transmission was measured at a low signal power ($-10\,\mathrm{dBm}$ before coupling to the chip) to avoid pumping for devices of different lengths,

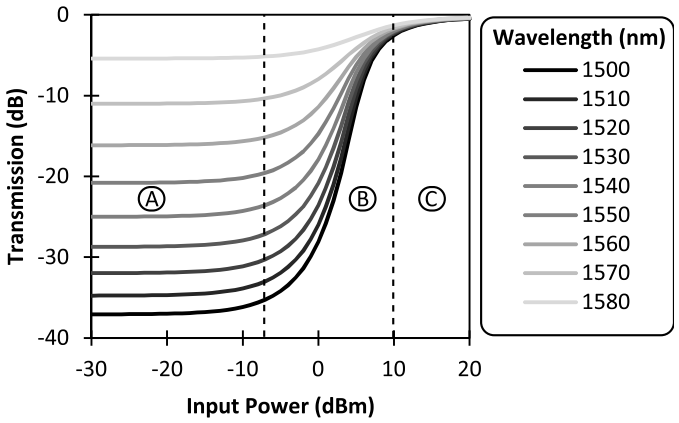


Figure 4.2: Transmission as a function of signal input power and signal wavelength.

and the transmission was correlated with the device length to obtain $\alpha_0(\lambda)$. For ease of use in the simulation engine, this data was then fitted to a quadratic curve which yields for the points $1530\,\mathrm{nm}$, $1550\,\mathrm{nm}$ and $1570\,\mathrm{nm}$ respectively $-5800\,/\mathrm{cm}$, $-4230\,/\mathrm{cm}$ and $-2175\,/\mathrm{cm}$. To determine the transparency carrier density N_t, the pump transmission as a function of input pump power was compared to a simulation of the transmission. First, for each value of the input power, the power absorbed by the QW was calculated. Then, using a measured carrier lifetime of $600\,\mathrm{ps}$, the average carrier density in the device could be estimated as a function of absorbed pump power. By linearly fitting this data to the QW absorption experienced by the probe, the transparency carrier density N_t could be found as a function of probe wavelength. It was found that the obtained values had a nearly linear relation with the wavelength, and therefore the data was fitted linearly to obtain $N_t(\lambda) = -(6.7 \pm 0.4) \times 10^{15}\lambda(\mathrm{nm}) + (1.20 \pm 0.06) \times 10^{19}$.

As we aim for a regenerator that works at $10\,\mathrm{Gbit/s}$, the carrier lifetime is assumed to be of the order of $100\,\mathrm{ps}$, unless stated otherwise.

4.4.2 Static non-linearity

In figure 4.2, the static transmission through a $1\,\mu\mathrm{m}$ broad membrane of $100\,\mu\mathrm{m}$ long is shown for different wavelengths. The membrane shows a large non-linear response. There are three regimes in the device operation. For low pump powers (A), only a small amount of carriers is generated, and therefore the device absorption is barely bleached. As

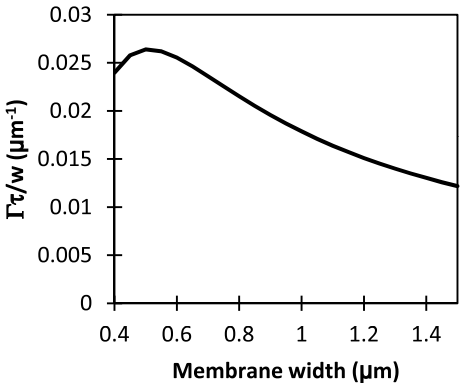


Figure 4.3: Efficiency of the regenerator as a function of the membrane width.

a consequence the output power varies linearly with the input power in this region, and hence the transmission is constant. When power increases however (B), the amount of generated carriers starts reducing the device absorption and a $\partial P_{out}/\partial P_{in}$ slope of more than one is achieved. In this region, the membrane could act as a regenerator, as a small change in the input power is converted to a larger change in the output power, which is useful to enhance the ER of an input signal. For wavelengths further away from the band gap wavelength, the achieved slope is very steep, up to a value of 4 for wavelengths shorter than $1530\,\mathrm{nm}$. To achieve the best performance, the band gap energy of the chosen QW stack should therefore be at least $20\,\mathrm{meV}$ lower in energy than the photon energy of the signal that needs to be regenerated. For high signal powers (C), the device is completely bleached and the output power is almost equal to the input power.

The membrane width is also an important parameter for the device operation, as it has an influence on the power consumption, the confinement factor, and the device speed. To study its influence on the device efficiency only, the efficiency parameter $\Gamma\tau/w$ as defined in section 3.5.2 is calculated for a fixed lifetime of $100\,\mathrm{ps}$. In figure 4.3 this parameter is plotted as a function of membrane width. Apparently there is an optimal width of approximately $500\,\mathrm{nm}$ to obtain the highest device efficiency, assuming that the effective carrier lifetime can be controlled to the desired value in another way (e.g. by passivation to increase the lifetime, by ion implantation to decrease the). In this work however, we are limited by the resolution of contact lithography, making features $< 1\,\mu\mathrm{m}$ difficult to obtain.

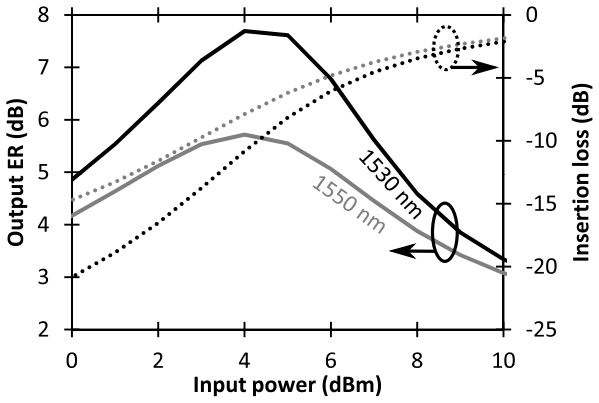


Figure 4.4: Output ER (solid line) and insertion loss (dotted line) of a $1530\,\mathrm{nm}$ (black) and $1550\,\mathrm{nm}$ (grey) signal with an input ER of $2\,\mathrm{dB}$.

4.4.3 Working point

As discussed in the previous section, injecting a signal with a low ER with a power level in the non-linear region of the membrane, will improve the ER of the signal. However, how much the ER improves exactly depends on the slope of the non-linear transmission function at the signal input power. In figure 4.2, it can be seen that this slope is continuously changing, and therefore the input signal power can be optimized to have a maximal improvement of the ER.

The choice of the signal input power also determines the insertion loss of the device. The insertion loss is here defined as the loss for the '1' signal level only, as any losses for the '0' signal level are desired as these improve the ER of the signal. When a higher signal input power is chosen, the device is bleached more, and therefore the insertion loss is lower. However, there is a clear trade-off between the regeneration performance and the insertion loss of the device. When a high input power is chosen to minimize the insertion loss, the slope of the non-linear transmission curve is low, and therefore the amount of regeneration is limited.

In figure 4.4, the output ER and insertion loss of a $1\,\mu\mathrm{m}$ broad membrane of $100\,\mu\mathrm{m}$ long is shown. The input signal has an input ER of $2\,\mathrm{dB}$, and is imprinted on a single mode carrier at a wavelength of $1530\,\mathrm{nm}$ or $1550\,\mathrm{nm}$. This clearly shows the trade-off: while there is an optimal input power to achieve the highest possible ER improvement, the insertion loss is minimized by using an as high as possible input

power. The optimal input power is then determined by the amount of insertion loss that can be tolerated.

4.4.4 Dynamic operation

As the device has to operate dynamically, continuously changing the absorption depending on the power in the input signal, also the dynamic operation needs to be simulated. To that end, a PRBS of 2^{10} bits is generated and imprinted on a carrier using non-return-to-zero on-off keying with an ER of $2\,\mathrm{dB}$. We assume that the modulation is infinitely fast, as we only want to study the dynamics of the membrane device. The simulation is performed at a carrier wavelength of $1530\,\mathrm{nm}$, and an input power of $7\,\mathrm{dBm}$ is chosen, as this limits the insertion loss to approximately $5\,\mathrm{dB}$ but still leads to an output ER of $5.6\,\mathrm{dB}$. In figure 4.5 and 4.6 simulated eye diagrams are shown at a bit rate of $10\,\mathrm{Gbit/s}$ and $20\,\mathrm{Gbit/s}$ respectively. The stepwise transitions in the both the eyes are the result of the original modulation, while the slower transitions are caused by the dynamics of the membrane. From this it is clear that the speed is limited by the carrier lifetime. In the eye at $10\,\mathrm{Gbit/s}$ the static values for the output power are just reached at the end of each bit period, which means that the predicted maximal ER improvement is achieved in a narrow time window. In the $20\,\mathrm{Gbit/s}$ eye on the other hand, the static values are not at all achieved, and the result is that the eye is closing. There is still a small improvement in the ER, however the time window during which an improvement is achieved, is rather small compared to the $10\,\mathrm{Gbit/s}$ eye.

4.5 Experimental results

The devices described below were made using the fabrication process as described in section 3.6. They are both using SU-8 inverted taper couplers as a means for fiber-to-chip coupling.

4.5.1 First generation

In the first generation of regenerators, the design of the membrane was not specifically optimized for this application. In fact, a device meant for a different application (see chapter 5) was used for these first proof-of-concept experiments. The first generation consists of a $150\,\mu\mathrm{m}$ long

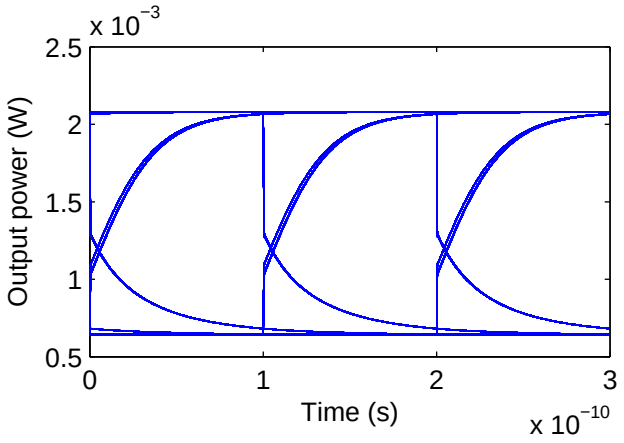


Figure 4.5: Simulated eye diagram at a bit rate of $10\,\mathrm{Gbit/s}$.

shallow etched membrane waveguide, coupled to the underlying silicon waveguide circuit using $18\,\mu\mathrm{m}$ long inverted taper couplers. The membrane waveguide is a $2\,\mu\mathrm{m}$ broad ridge waveguide, which consists of three $8\,\mathrm{nm}$ InGaAs QWs separated by $10\,\mathrm{nm}$ thick InP barriers, sandwiched between two $25\,\mathrm{nm}$ thick InP cladding layers. The ridge waveguide was defined using the standard dry etching process (CH_4:H_2/O_2) in the RIE etcher.

Static non-linearity

To assess the potential regeneration performance of this device, the static transmission curves have been measured at different wavelengths. The measurement results are shown in figure 4.7 and the corresponding simulation results are shown in figure 4.8. These were obtained by setting τ_c to a measured value of $600\,\mathrm{ps}$. The experimental results match well with the simulated results, proving the validity of the simulation engine, despite the use of simple modeling equations. At shorter wavelengths, there is a discrepancy between the simulation results and the measurement results. This is probably caused by a varying fiber-to-chip coupling loss, while this was assumed constant at $6.5\,\mathrm{dB}$ in all measurements here. Because the reference waveguides on the sample were damaged, this coupling value was determined by measuring the transmission at a wavelength of $1.6\,\mu\mathrm{m}$ instead, for which there is no material absorption.

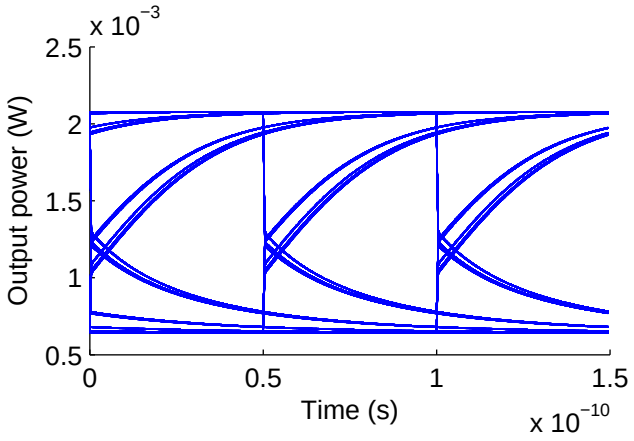


Figure 4.6: Simulated eye diagram at a bit rate of $20\,\mathrm{Gbit/s}$.

Regeneration performance

In order to establish the regeneration operation of the membrane, a BER measurement was performed using the setup depicted in figure 4.9. In these experiments a PRBS of $2^{31}-1$ bits was used, imprinted on the CW light carrier using NRZ-OOK. In a first experiment the achieved ER and associated insertion loss were measured as a function of input power, for an input signal with an ER of $2\,\mathrm{dB}$ at a bitrate of $1\,\mathrm{Gbit/s}$. The results of this experiment are shown in figure 4.10. As discussed in section 4.4.3, the optimal working point can be determined from these curves. In this particular experiment, we choose the working point at each wavelength to achieve a maximal regeneration: e.g. at $1530\,\mathrm{nm}$ an input power of $-1.5\,\mathrm{dBm}$ is chosen which leads to an ER improvement from $2\,\mathrm{dB}$ to $6.2\mathrm{dB}$, with an associated insertion loss of approximately $10\,\mathrm{dB}$.

The optimal input power for regeneration was determined for all tested wavelengths and a BER measurement at $1\,\mathrm{Gbit/s}$ was performed to check the above results. According to the theory, the improvement of the signal ER should have an immediate effect on the BER performance of the optical link, as the power penalty at the receiver will be reduced. The power penalty (δ) is a parameter defined as the difference between the required power incident on the receiver to achieve error-free operation (BER $< 10^{-9}$), between the case with an ideal, infinitely high ER, and the case with a limited ER. From the measurements pre-

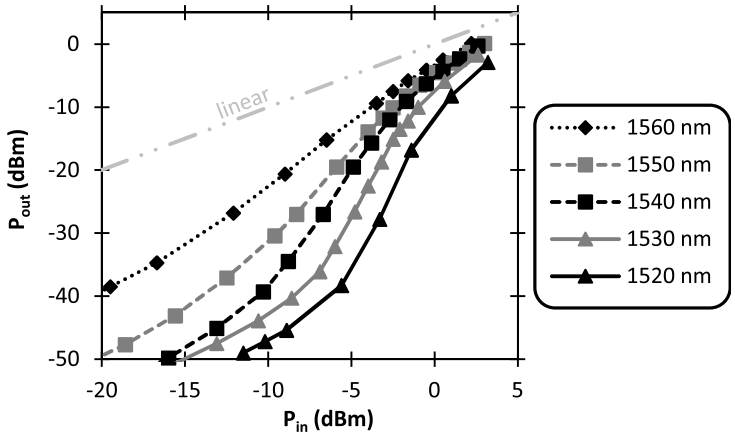


Figure 4.7: Static transmission through a $150\,\mu m$ long membrane.

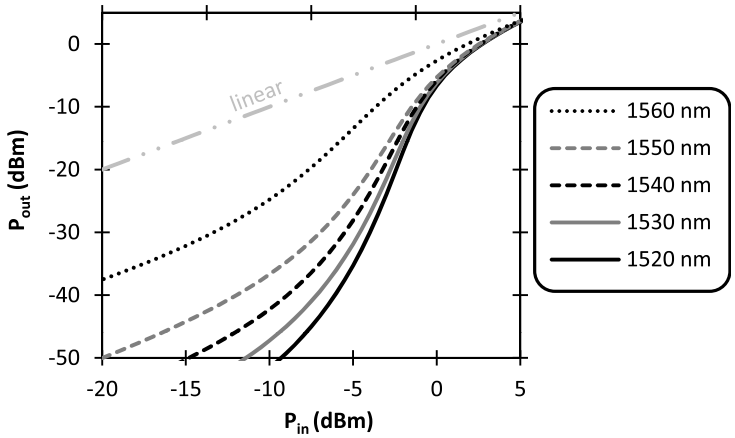


Figure 4.8: Simulation of static transmission through a $150\,\mu m$ long membrane.

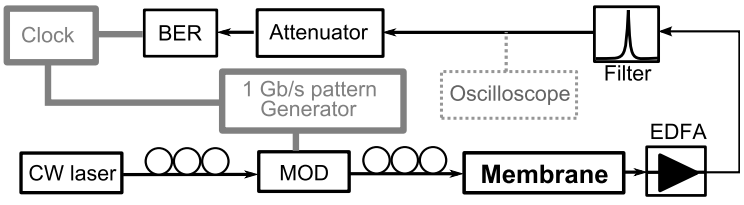


Figure 4.9: Experimental setup of BER measurement for the first generation devices.

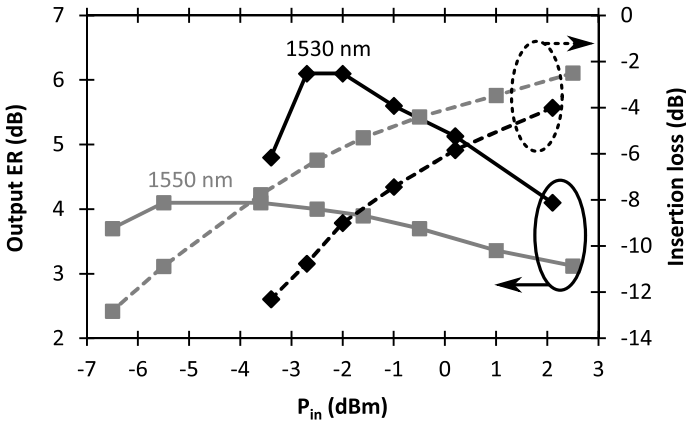


Figure 4.10: Insertion loss (dashed line) and output ER (solid line) as a function of input signal power for signals at $1530\,\mathrm{nm}$ and $1550\,\mathrm{nm}$.

sented in figure 4.10, the expected reduction of the power penalty can be calculated, as the power penalty and ER are related by [76]

$$\delta[\mathrm{dB}] = 10\log_{10}\left(\frac{1+r_{ER}}{1-r_{ER}}\right) \tag{4.3}$$

where r_{ER} is the linear extinction ratio. Then the expected improvement is equal to $\delta_{in} - \delta_{out}$, with δ_{in} the penalty of the input signal and δ_{out} the penalty of the output signal, after transmission through the regenerator. For an improvement from $2\,\mathrm{dB}$ to $6.2\mathrm{dB}$, we can expect a penalty reduction of $4.3\,\mathrm{dB}$. To verify this experimentally, a back-to-back (B2B) measurement was compared to a measurement on the regenerated signals for different wavelengths. The results are shown in figure 4.11. It is clear that the achieved ER improvement also leads to the expected reduction in power penalty. This can also be seen in the eye diagram in figure 4.12. Furthermore, the speed limitation by the carrier lifetime ($600\,\mathrm{ps}$) is clearly visible in the falling edge of the eye. As expected, the lower improvement of the ER at longer wavelengths also leads to a lower reduction of the power penalty.

Discussion

These first results clearly show the potential of the membrane as a regenerator. However, there are some flaws in the design that could be improved. First of all, the device only works well at $1530\,\mathrm{nm}$. It would be better if a single device would give similar results across complete

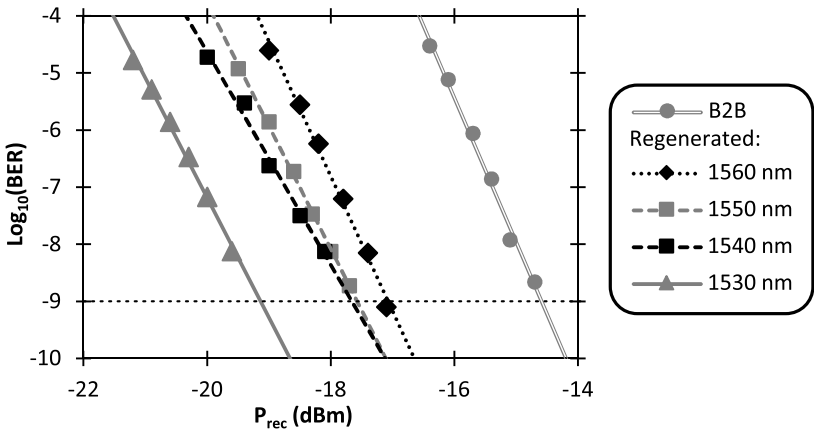


Figure 4.11: Comparison of the B2B BER measurement and the regenerated BER measurement in the 1530-1560 nm wavelength range.

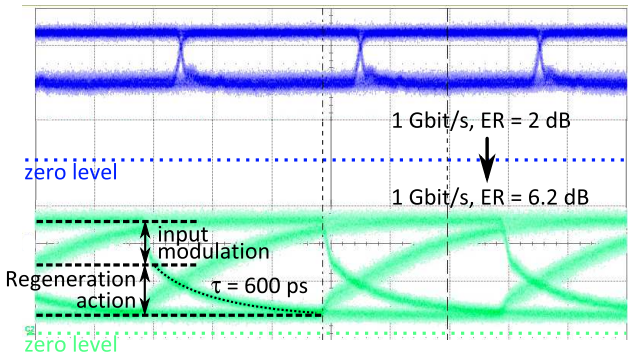


Figure 4.12: Measured eye diagram at 1 Gbit/s before and after regeneration at optimum regeneration conditions at a wavelength of 1530 nm.

C-band. Next to this, the speed of the current device is too low to be useful in a real life application.

4.5.2 Second generation

In the second generation of devices, some optimizations where made to the device design to make it more suitable as a regenerator. First of all, a different epitaxial layer stack was used: a triple QW structure with $10\,\mathrm{nm}$ InGaAs QWs separated by $8\,\mathrm{nm}$ InP barriers. The reason for this, is that this new MQW stack will allow good regeneration performance across the complete C-band, and in this way lead to a more useful device. The increased QW width also means that the etched QW surface area is slightly larger, which should enhance the surface recombination. Furthermore, also the optical confinement in the QWs is increased to a value of $22\,\%$, which more than compensates the lower material absorption of broader QWs. The combination of the higher absorption of the MQW stack in the C-band with the higher confinement leads to a reduction of the device length from $150\,\mu\mathrm{m}$ to $100\,\mu\mathrm{m}$.

Next to this, the membrane width is reduced from $2\,\mu\mathrm{m}$ in the previous design to only $1\,\mu\mathrm{m}$ in this second generation design. As discussed in section 3.5.1, this increases the influence of surface recombination, and in this way speeds up the device. Furthermore this improves the efficiency of the device as we showed in section 4.4.2, assuming however that the effective carrier lifetime in the device does not become too short for the bit rate the regenerator should operate at.

To increase the surface recombination rate even more, a high temperature ICP etching recipe was used to etch through the QWs. This recipe should lead to more sidewall damage, and should in this way increase the surface recombination velocity.

Static non-linearity and device speed

In figure 4.13(a) the CW device transmission is shown for different wavelengths across the C-band. The transmission curves have a similar shape and slope in the non-linear region. Therefore, we can expect similar regeneration performance across the C-band.

To assess the device speed, a femtosecond pulse train with high peak power and center wavelength at $1550\,\mathrm{nm}$ was sent together with a weak CW probe bream at a wavelength of $1590\,\mathrm{nm}$ through the device. When such a femtosecond pulse is transmitted through the device, a large number of carriers is created quasi-instantaneously. This causes

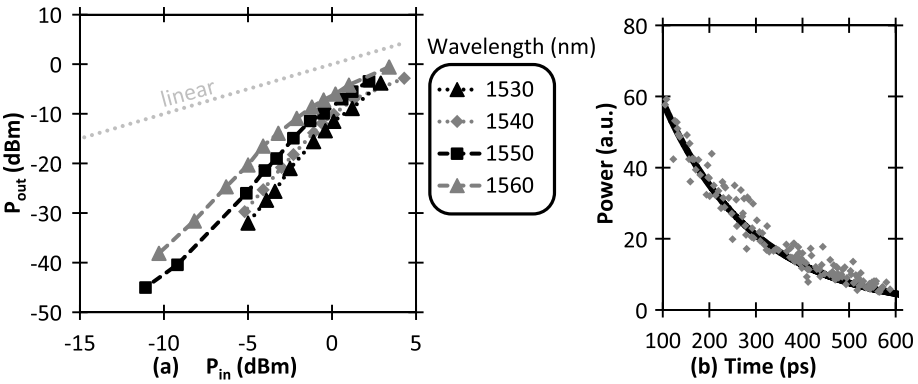


Figure 4.13: (a) Static transmission through the 2nd generation device. (b) Measurement of the roll-off of the probe transmission (grey), and fitted exponential (black) with $\tau_{eff} = 200\,\mathrm{ps}$.

an immediate bleaching of the absorption of the probe signal, which can be monitored on an optical oscilloscope. By fitting the exponential roll-off of the probe transmission due to the recombination of the created carriers, the carrier lifetime was determined to be $200\,\mathrm{ps}$ (see figure 4.13), an improvement of a factor of three compared to the first generation design. To check whether the cause of this lower lifetime lies with the enhancement of the surface recombination, the lifetime of a $1.5\,\mu\mathrm{m}$ broad device was also measured. We found an effective carrier lifetime of $310\,\mathrm{ps}$. When we look back at the simulation results presented in figure 3.13, we can see that these data points both lie on the curve with a surface recombination velocity of $2 \times 10^5\,\mathrm{cm/s}$. This means that the surface recombination velocity is indeed enhanced as intended. Furthermore, we can also conclude that a higher SRH recombination is not the cause for the shorter lifetime, as otherwise the lifetime would be less dependent on the device width. This means that the epitaxial layers used in this experiment are of similar growth quality as the epitaxy used for the first generation devices.

Regeneration performance

Similar as for the first generation of devices, a BER measurement is performed, using the setup as depicted in figure 4.9. First the output ER and insertion loss were determined as a function of signal input power, for a signal with $2\,\mathrm{dB}$ input ER and a bit rate of $2.5\,\mathrm{Gbit/s}$, at wavelengths of $1530 - 1560\,\mathrm{nm}$. In figure 4.14, the results are shown

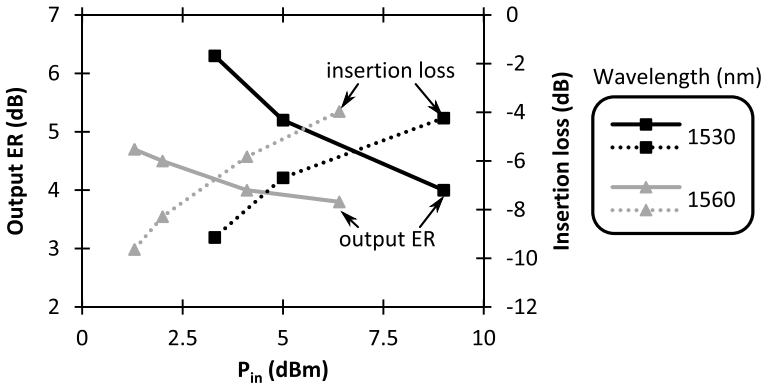


Figure 4.14: Insertion loss (dotted line) and output ER (solid line) as a function of input signal power for signals at $1530\,\mathrm{nm}$ and $1560\,\mathrm{nm}$.

for $1530\,\mathrm{nm}$ and $1560\,\mathrm{nm}$. These form the outer bounds for the results that can be expected in the C-band. Assuming we can afford a $10\,\mathrm{dB}$ insertion loss like before, we can obtain an enhancement of the output ER to $4.7 - 6.2\,\mathrm{dB}$ for wavelengths in the C-band. For these working points, also the BER was measured, and results are given in figure 4.15. It is clear that the enhanced output ER also leads to the associated receiver sensitivity improvement. The power penalty is this time reduced by $3.6\,\mathrm{dB}$ over the entire C-band, up to $4.5\,\mathrm{dB}$ at wavelengths $1530 - 1540\,\mathrm{nm}$.

4.6 Conclusion

In this chapter the heterogeneously integrated membrane gate has been successfully applied for a signal processing application. We have shown that the device can be used as a regenerator. Due to the large light-matter interaction in the device, the membrane works especially well for the regeneration of input signals with a low ER. As a demonstration, we were able to finally obtain a receiver sensitivity enhancement of $3.6\,\mathrm{dB}$ over the entire C-band, up to $4.5\,\mathrm{dB}$ at wavelengths $1530 - 1540\,\mathrm{nm}$, for a signal with an input ER of only $2\,\mathrm{dB}$ at a bit rate of $2.5\,\mathrm{Gbit/s}$. Next to this, also the speed limitation of the device was experimentally determined. By playing with the device design to enhance surface recombination, we were able to increase the speed by a factor of three compared to the initial unoptimized design. Furthermore, higher speeds can be obtained by increasing the bulk

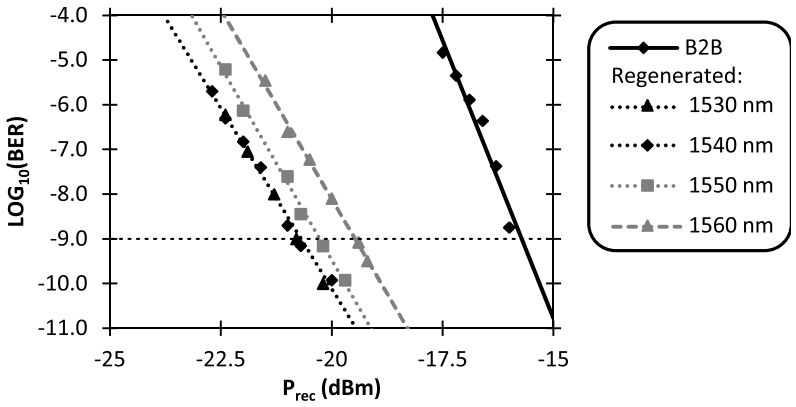


Figure 4.15: Comparison of the B2B BER measurement and the regenerated BER measurement in the 1530-1560 nm wavelength range.

non-radiative recombination by implanting the device with highly energetic ions [62, 63], however this was not further pursued in this work.

Chapter 5

All-optical switching

In the previous chapter it was shown that integrated membranes can be useful for high speed signal processing applications, due to the large light-matter interaction that is obtained in the device. In this chapter, we will focus on another application that can greatly benefit from this strong interaction, but that does not need to be fast: the realization of an all-optical packet switch for application in access networks.

5.1 The need for all-optical packet switches in access networks

The worldwide required bandwidth has been rising rapidly the passed years. The onset of video-on-demand, cloud computing and other online services has been the main driver for this bandwidth explosion, and the end does not seem near [24]. To fulfill this bandwidth requirement, the network backbone has been relying on optical fiber communication for years, however in most cases copper wire is still being used to bridge the last mile between operators and clients. The reason for this is obvious: copper wiring had been installed in most homes during the deployment of telephone and cable networks, and reusing the existing wiring was therefore the most cost effective way to connect homes to the network.

With the continuing increase of required bandwidth per user however, it will be inevitable to replace this so called last mile by fiber. There are several ways to implement such an access network. Currently, the network technology that is mainly used is the PON solution, because of its low cost and simplicity. In section 5.2, the advantages and

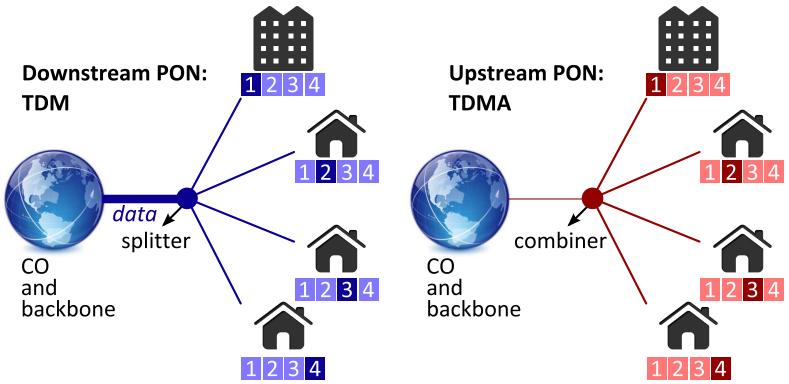


Figure 5.1: schematic of PON solution. A different wavelength is used for downstream and upstream data.

drawbacks of these PONs will be discussed, and competing technologies that aim to address some of these drawbacks will be introduced. One of these alternatives is the use of all-optical switches in the access network. In this work, we aim to use III-V membranes to implement such a switch. In section 5.3, the different possible switch architectures that can be used in combination with the III-V membrane technology will be briefly discussed. Finally, a full 1x4 switch is designed in section 5.4, and the characterization results for a first fabricated prototype are discussed in section 5.5.

5.2 The state of the art

In the following the most important currently available technologies to implement fiber to the home networks are discussed. Only point-to-multipoint technologies are considered, as point-to-point links, in which every client is connected with his own fiber, are currently very expensive due to the large amount of equipment required.

5.2.1 Passive technologies

Passive Optical Networks (PON)

Currently, most fiber to the home (FTTH) networks that are being rolled out make use of a PON architecture. Such a network consists of a number of ONUs that are connected to the CO which provides for the connection to the backbone network. In a PON, a large number of ONUs

(typically $32 - 64$ [77]) shares a 10 to $20\,\mathrm{km}$ long fiber link to the CO. At the end of this link, a passive splitter is placed, which splits the optical signal in equal parts (figure 5.1).

Although this approach to FTTH is easy to implement, these systems are not suitable as a long-term solution as they can only provide a limited bandwidth. This is due to the fact that the network splitting ratio, network reach and line rate are all determined by the power budget. Therefore, as the provided line rate increases, the available power budget decreases as well, limiting the number of users and distance from the CO. Next to this, clients connected to the same splitter have to compete with each other for bandwidth through TDM. This means that while more connected clients per node is cheaper for the service provider, the available bandwidth per user will go down at the same time. Furthermore, as all data packets are broadcast to all users, most of the high-speed and hence power consuming processing performed on the incoming data stream in an ONU is redundant, as only $1 - 3\,\%$ of the incoming data is actually meant for this particular user. Therefore, a lot of power is wasted. In [78], this problem is addressed by moving to a bit-interleaved TDM algorithm, but this makes it more difficult to dynamically assign bandwidth accross the different users. Additionally, the broadcasting of data brings the additional requirement that the data should be encrypted to prevent eavesdropping on confidential information, which increases the overhead at both the ONUs and CO. To accommodate upstream data communication, the TDMA protocol is used, where the CO specifies a time window for each client to transmit upstream data. The upstream data is imprinted on a different wavelength band (typically $1.3\,\mu\mathrm{m}$) than the downstream data (typically $1.49/1.55\,\mu\mathrm{m}$ for data and television services).

Wavelength Division Multiplexed Passive Optical Networks (WDM-PON)

To solve the main difficulties of PON, WDM-PON was proposed as an alternative. In this network architecture wavelength division multiplexing is combined with passive splitters to increase the number of possible clients connected to a single downstream fiber link [79], due to the low loss of wavelength division (de)multiplexers compared to splitters (see figure 5.2). Furthermore, because several wavelength channels are used at once within a single network, there is a higher available bandwidth per user than in a PON approach.

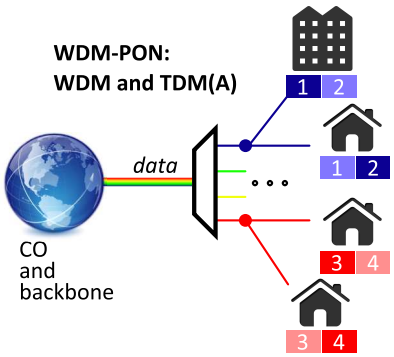


Figure 5.2: schematic of WDM-PON solution. Because of the WDM filtering the upstream data communication is more complicated, as the exact same wavelength needs to be used as for downstream communication.

Although this is a significant improvement over PON, different users linked to a certain downstream wavelength still have to compete for bandwidth through TDM for downstream and TDMA for upstream communication, and the problems mentioned above hence remain the same, but occur at higher speeds and a higher number of connected users than in a PON. Next to this, the upstream data communication is much more complicated than in a PON, as the upstream data has to pass through the wavelength multiplexer. The obvious solution is to use a (tunable) laser source at the ONU at the user side. However, sufficiently widely tunable lasers are currently too expensive to implement in all ONUs. The use of different, unique ONUs with a laser with a specific wavelength has the advantage that it is cheaper to produce, but on the other hand this brings a huge logistic cost as a provider would have to keep track of all the different ONUs and keep a large storage of all the different ONUs with their specific wavelength. Therefore, there has been a lot of research on possible 'colorless' alternatives, which use either a reflective amplifier [80], an injection locked Fabry-Perot laser [81] or a wavelength converter [82]. Although these solutions simplify the equipment required at the ONU side, a CW probe still needs to be sent by the CO to the ONU as a seed beam for the upstream signal.

5.2.2 Active switches

Instead of using a passive network, also active solutions using switches could be useful. Switches are already widely used in the network back-

bone, however in those cases the switch actually operates in the electrical domain, with a detector array and laser with modulator array to convert the optical signal to an electrical signal and vice versa [83]. Such a solution is too expensive to be cost effective for a few hundreds of connected users, and has a too high power consumption as well. Therefore, only electro-optical or all-optical solutions will be considered. The basic idea of a FTTH implementation using active switches is shown in figure 5.3.

While in traditional electrical switches the address for a packet of data is encoded inside the data packet itself, this address is now replaced by an optical signal that travels along with the packet. This so-called label is then used by the all-optical switch to route the data packet to the right output port(s). This kind of switching assumes therefore that the CO converts the electrical address to the appropriate label.

Using an active switch can have several potential advantages. Such an active solution can lead to a lower total power consumption, despite the fact that the switch itself consumes power. First of all, the switch can have a lower optical insertion loss than a passive solution. This increases the power budget available which can be used to connect more users, connect users which are further away from the CO, or increase the bandwidth of the fiber-optic link. Next to this also electrically the power consumption can be decreased, as the ONUs of the connected users only receive useful data. As explained above, a lot of power is wasted in traditional ONUs used for PON on the high-speed processing of incoming data which is going to be discarded anyway, as most of it is meant for other users. As this redundant data is already optically removed from the incoming data stream, there is an opportunity to achieve a lower power consumption of the ONUs, if one is able to find an efficient solution to tackle the burst-like nature of the incoming data stream. A second advantage is that the division of the available bandwidth is completely determined by the label: in principle the allocated packet length to a certain user can be changed continuously without much added complexity. A third advantage is the added security of a system with a switched architecture. While in PONs all connected users to a single splitter receive all packets sent on this FTTH link, in a switched architecture each user only receives his own data and therefore the risk of eavesdropping is lower.

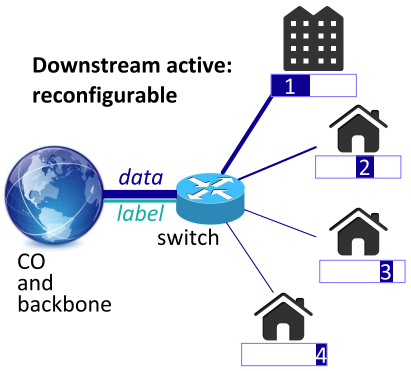


Figure 5.3: Schematic of typical active solution. Each connected ONU only receives the data meant for that ONU.

5.2.3 Electro-optical switches

Several electro-optical switches have been proposed in literature [84, 85, 86]. A first category of switches aims to increase the reconfigurability of existing networks. In [84], electrically controlled switches are used to choose between a simple broadcast or a more advanced WDM-based filtering of downstream packets, allowing a much more efficient use of the available bandwidth. The rationale of this research was to provide for a mechanism to upgrade existing PON to more advanced WDM-PON networks without having to shut down the PON operation. In this case, the optical label to switch between states is a high power pulse that only needs to be applied during the switching of the optical switching elements, optical latch switches, between bar and cross states. The power in this pulse is then converted into an electrical signal using a photovoltaic cell to drive the switch. Although this architecture provides for some reconfigurability, the access network will suffer from the same problems as WDM-PON.

A second category of switches, aims to implement a fully functional dynamic packet switch [85, 86]. In [85], the electro-optic effect in PLZT (Lead Lanthanum Zirconate Titanate) was used to create a packet switch capable of switching incoming packets to one of its output ports. Although these switches show a good extinction and are fast enough (switching speed $\approx 10\,\mathrm{ns}$) to do packet switching, they also have a number of disadvantages. In the first implementations the PLZT switches could only switch to one output at a time. As this was a major disadvantage for the broadcasting of television and the implementation of the basic client discovery process, a solution was found where the

switches were driven at an average voltage between cross and bar state to allow multicast operation [87]. Next to this, although the power consumption of the switches themselves is relatively low ($10\,\mathrm{mW}$), the required driving electronics pushes the total power consumption to more than $2\,\mathrm{W}$ [88]. In [86], a similar switching concept is used, but this time based on $LiNbO_3$ switches.

The main problem with electro-optic switching technologies that can be continuously reconfigured lies however with the requirement for electrical power, since traditionally PON splitter boxes are located in cabinets were no electrical power is present.

5.2.4 All-Optical switches

To overcome the lack of electrical power, a purely optical solution is pursued. In [89], silicon ring resonators are used as all-optical switch elements. In this approach, the label is a high power optical signal in the C-band ($P \approx 20\,\mathrm{dBm}$ before waveguide injection) which is tuned to one of the resonances of the ring resonator. When the label is present, a high intensity builds up in the ring, which creates a large number of free carriers due to two photon absorption of the label. Because of the plasma dispersion effect, the effective index of the resonator mode changes, thereby shifting the resonance and switching the data signal to the drop port. Although this kind of switch allows fast switching (switching time $< 1\,\mathrm{ns}$), a reasonable extinction ratio ($10\,\mathrm{dB}$) and insertion loss ($1 - 3\,\mathrm{dB}$), the requirement for accurate tuning of the ring's resonances and the high power of the optical label are important disadvantages.

In [90], photonic crystal cavities are used as resonators instead. The operating principle is similar as above, however the free carriers are created in this case by a combination of linear material absorption at the absorption edge and TPA in the InGaAsP photonic crystal. The result is that a phase shift is obtained by both the FCD (related to FCA) and the band filling dispersion (related to the linear absorption) effect, leading to a low energy consumption. In this case, a high switching speed was aspired, however by using proper passivation of the etched photonic crystal holes, the switching speed could in principle be reduced in favor of a lower operating power. The drawback of the need for careful tuning of the operating wavelength remains however.

Next to these resonator based solutions, also several broadband high-speed switching solutions have been proposed. In [91], all-optical

cross-phase modulation is used to create a phase shift in a phased array architecture. The main difficulties here were the high power required to drive the device, and the low ER obtained. In [92], four wave mixing is used to create very high speed all-optical switches by using a highly non-linear polymer in a silicon slot waveguide to optimize the non-linear efficiency. Also here, the high operating power renders the device impractical for use as a packet switch.

5.2.5 Design guidelines

From the above we can conclude that the following design goals should be met for the design of an all-optical switch:

- The insertion loss of the switch should be lower than for a purely passive solution (PON).
- The switching time should be of the order of a few nanoseconds, in order to keep the guard time as small as possible.
- The switching power should be low enough to be able to generate the label using a standard laser diode.
- The switch should be passive: only optical power can be used to drive the switch.
- The switched architecture should provide support for a sufficiently high number of connected clients. E.g. according to the specifications for XG-PON, a minimal amount of 64 ONUs is supposed, with possible upscaling to 128 to 256 users [77], so this should also be possible using the switched architecture.

5.3 Switching fabrics using membrane technology

In this section the different switching architectures for the realization of an all-optical switch that can make use of the characteristics of integrated III-V membranes are discussed in a nutshell.

5.3.1 Broadcast-and-select architecture

A first important architecture is the broadcast-and-select architecture [83]. The basic layout of such a switching fabric is shown in figure 5.4. In a switch based on such a architecture, the data packets are

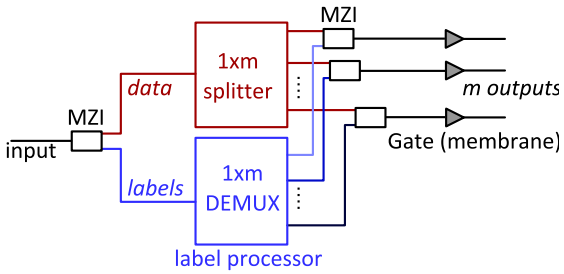


Figure 5.4: schematic of implementation of broadcast-and-select architecture using membrane gates. The MZIs are used to (de)multiplex the label and data wavelength bands.

broadcasted to all output nodes. In each of those nodes, an amplifier is present that acts as a gate that can either amplify the incoming packet or block it. These kinds of switches allow for unicast, multicast and broadcast operation, as one or multiple output amplifiers can be driven at once. Furthermore, they are inherently broadband, as the optical bandwidth will only be limited by the bandwidth of the splitters used in the broadcaster and the gain bandwidth of the output amplifiers. However, the switch is not very power efficient: similar as in a PON, the data signal is broadcasted. To allow for a higher bit rate, more connected ONUs and a larger network reach than in a PON, the output amplifiers provide for gain to compensate part of the loss due to the broadcasting of the data signal.

The membrane gates introduced in chapter 3 could be used to create an all-optical switch with this architecture. Because of the high confinement in the III-V membrane, a short device ($< 100\,\mu\mathrm{m}$) can already have a large enough absorption to achieve a high ER between the different output ports of the switch. E.g. in figure 3.25, an ER of $40\,\mathrm{dB}$ is obtained in a $150\,\mu\mathrm{m}$ long device, while actually an ER of $20\,\mathrm{dB}$ would already be more than sufficient. Furthermore, the membrane gates are inherently designed for optical pumping without the need for any electrical power supply, and therefore ideal for the use in an all-optical switch. From this design it also directly follows that the label that carries the address information, should be a shorter wavelength signal ($\lambda \approx 1500\,\mathrm{nm}$) with enough power to bleach the membrane gates, while the data should be positioned around $1550\,\mathrm{nm}$.

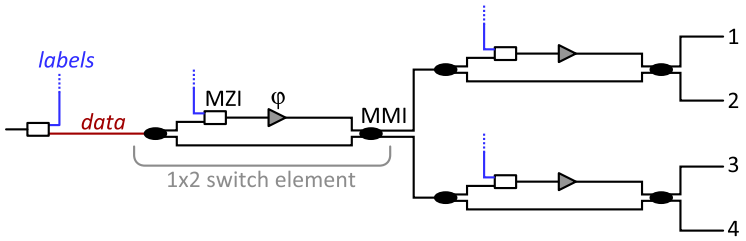


Figure 5.5: Schematic of implementation of concatenated MZI architecture using membrane gates as phase section. The MZIs in the variable phase arm are used to (de)multiplex the label and data wavelength bands.

5.3.2 Concatenated MZI architecture

In a concatenated MZI architecture, the switch is constructed using 1x2 switch building blocks. The basic layout of such a switching fabric using membrane gates is shown in figure 5.5. Here we assume a similar label processing scheme as shown in figure 5.4. Each of the 1x2 switches consists of a balanced 1x2 MZI, with a variable phase section in one of the two branches. By changing the phase, the output can then be tuned from one output port to the other.

This architecture can be realized using the membrane gate by leveraging the free carrier dispersion effect (FCD). To do this, some optimizations should be applied to the quantum well design. The quantum wells in the membrane should be designed such that they do not absorb at the data signal wavelength (e.g. in C-band), while they should absorb at the shorter label wavelength (e.g. in S-band), which is added using a 2x1 MZI before each phase section as depicted in figure 5.5. If this is the case, the free carriers created by the label absorption will cause a phase shift in the membrane due to the FCD effect, while the data signal will only experience a loss due to FCA. If the total accumulated phase shift amounts to π, the signal will be completely switched. It is important to point out that the phase change is caused by a slightly decreasing refractive index with increasing carrier density. Although this approach could work in theory, there are some difficulties. First of all, the heating of the device associated with optical pumping (see section 3.2.4) will counteract the phase change, as the refractive index will increase with increasing temperature. Next to this, the FCD effect is not very strong leading to a large footprint: for a π phase shift, a free carrier density of on average $9 \times 10^{17}\,/\mathrm{cm}^3$ is required in a device of $400\,\mu\mathrm{m}$ long, assuming a confinement in the QWs of $18\,\%$. Further-

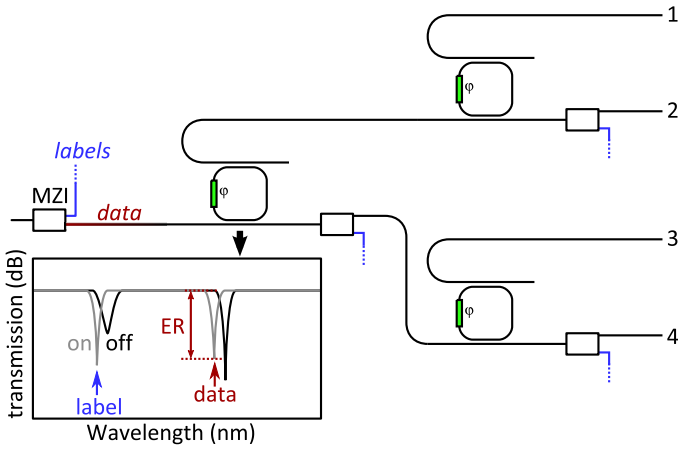


Figure 5.6: Schematic of implementation of resonator architecture using III-V membrane as phase section. The MZIs are used to (de)multiplex the label and data wavelength bands.

more, the combination of FCA and scattering in the device will lead to an insertion loss of approximately 1 dB for each 1x2 switch element, assuming a lossless passive waveguide design.

5.3.3 Resonant cavity architecture

Another architecture that has been getting a lot of attention [89, 90] is the resonant cavity architecture. In this architecture the data signal is switched by controlling the resonance of a photonic crystal, ring or racetrack resonator. This can be done by applying a small phase change in the resonator, which causes the characteristic transmission of the resonator to shift. This is schematically shown in the inset of figure 5.6. Similar as in the concatenated MZI architecture, this phase change can be provided by including a short III-V membrane in the resonator cavity, using the FCD effect. However, in this case the label and data wavelengths need to be aligned with the resonances for the switch to work. With each resonator, a 1x2 switch element can be formed, and these can be combined to form a 1xm switch as shown in figure 5.6.

The main advantage of resonator based switches is that only a small phase shift is required to achieve a large extinction between the two output ports, because of the narrowband nature of the switch. Therefore the power required to switch these devices will be quite low. However, the fact that the device is narrowband also has a drawback: any

change in the environment (e.g. temperature change) will cause a shift of the resonances. Furthermore, it is (currently) very challenging to reproducibly fabricate multiple resonators, especially considering that a III-V membrane needs to be processed on top. For these reasons, a heater is required near each resonator to thermally control the position of the resonances, which greatly increases the amount of power required to drive these switches.

5.4 Design of an all-optical switch

Due to its simplicity and the anticipated performance of the membrane as a gate, we choose to implement an all-optical switch using the broadcast-and-select architecture. Furthermore, a switch using this architecture could actually be designed to be backwards compatible with PON. We will discuss this in section 5.4.1. In section 5.4.2 the layout of the switch is described and the required passive silicon building blocks are designed.

Also the concatenated MZI and resonator architecture are viable switching fabrics and might even perform better when all building blocks are well developed. However, for both these architectures, it is much more important to achieve a good control and reproducibility of the passive structures to be able to create a working device. Due to its simple layout, the broadcast-and-select architecture is more forgiving.

5.4.1 Compatibility and clock distribution

Next to the technical performance of the switch, also other considerations could be important for the success of a particular switching architecture. One important aspect is the aspect of logistics. In case a network operator wants to upgrade an existing PON architecture, it would be interesting if the ONUs connected to the network could be temporarily connected to the new switch without loss of functionality. This is important because otherwise all the ONUs have to be replaced at the same time as the installation of the switch to avoid clients to go offline, or the clients have to be connected one by one to the new switch, which has to operate in that case in parallel for some time with the PON.

One of the advantages of the broadcast-and-select architecture is the similarity to a PON. Like in a PON, the data signal is broadcasted to all connected clients, however in this case the switch can control which

clients actually will receive the data. This allows for two ways to maintain backwards compatibility. One possibility is by pumping the output gates that are connected to the legacy ONUs at all times, using TDM and TDMA as in PON. However, a more interesting possibility is to use a shorter wavelength band for the data signal in the switched architecture. Then, the output gates can be designed in such a way that they absorb at the switched data signal, while they are transparent for the broadcasted data signal. Next to providing backwards compatibility, this transparent band can also be used for broadcasting services such as television.

A drawback of a switched architecture is that the data signal at the ONU consists of a series of short bursts, as the data signal is only transmitted if it is meant for this particular user. As a result, a more challenging burst mode receiver is required at the ONUs, as clock recovery is more difficult to achieve on such a signal. However, the broadcasting possibility at longer wavelengths can be used to avoid the need for complex clock recovery electronics. By broadcasting the clock to all ONUs at a longer wavelength, a simple WDM splitter and extra detector at the receiver side could replace the electronics that would otherwise be required. On the other hand, the burst mode shape of the incoming data also means that there is an opportunity to design a receiver in which a large part of the electronics can be turned off during the time no signal is incoming, in order to reduce the power consumption of the ONUs. For instance the remaining label power which is sent along with the data packets might be used as a trigger to power up the power-hungry high-speed electronics.

5.4.2 Switch layout

A scheme of the architecture is shown in figure 5.7. For the separation of label, and clock and data wavelength bands, a MZI is used. This MZI is constructed using a 1x2 coupler, a 2x2 coupler and appropriate phase section. As a label processor, an AWG is used. Depending on the label wavelength, the label will then only be sent to one of the different output gates. For the broadcaster, a splitter tree is used that is composed of 1x2 couplers.

From the measurements described in section 3.7.2, we can conclude that the wavelengths of the labels should ideally be placed in the S-band ($\approx 1500\,\mathrm{nm}$), while the data should be put in the C-band ($\approx 1550\,\mathrm{nm}$) and the clock in the L-band ($> 1600\,\mathrm{nm}$) as the device

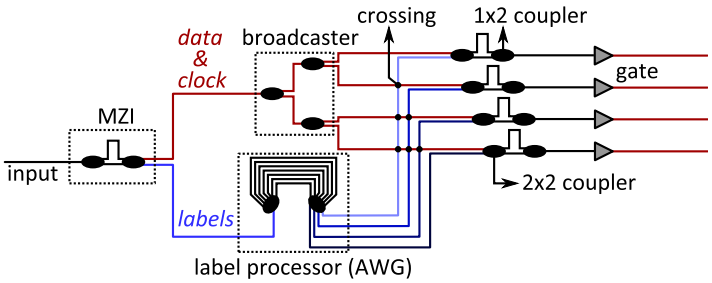


Figure 5.7: Detailed schematic of implementation of 1x4 broadcast-and-select switch using membrane gates.

absorption drops to zero there. The designed passive components should support these wavelength bands.

Another component that is required is the waveguide crossing: it is important to achieve a low loss and cross talk to maintain a good performance. The design of the passive components described below, is mainly the result of the efforts of (past) colleagues in the group that have built up a vast knowledge base.

1x2 coupler

The 1x2 coupler is used for both the broadcaster and the MZI, and its performance therefore has a large impact on the complete switch. The specifications for the design of the 1x2 coupler are a $50/50$ splitting ratio, equal phase at the output, and low excess insertion loss. To achieve the first two requirements, we choose to use a symmetric 1x2 MMI. Because of the device symmetry a nearly perfect phase and splitting ratio can be achieved. For the 1x2 coupler in the broadcaster, the standard 1x2 MMI of the picazzo toolbox [93] was used with length of $9.9\,\mu\mathrm{m}$ and width of $2.9\,\mu\mathrm{m}$, as this component is optimized for operation around $1550\,\mathrm{nm}$. However, for the MZI, a slightly shorter MMI section of $9.5\,\mu\mathrm{m}$ was used, as this should increase the performance for the shorter label wavelengths. This is important as there should be enough power remaining at the output to be able to bleach the membrane gate. Both MMIs showed an insertion loss of approximately $1\,\mathrm{dB}$ at the wavelength they were optimized for.

The reason for this loss seems to be related to the tolerance on the silicon thickness across different wafers, as this component has worked with lower insertion loss on older fabrication runs. Recently, the MMI from the library used in this work has been replaced by a more robust

design with a lower loss and less variation. This was done by moving to a larger design with a width of $5\,\mu\mathrm{m}$ and length of $25\,\mu\mathrm{m}$.

2x2 coupler

In an initial design, also for the 2x2 coupler the standard 2x2 MMI of the picazzo toolbox [93] was used. This component has a width of $2.9\,\mu\mathrm{m}$ and length of $38\,\mu\mathrm{m}$. However, measurements on the MZIs using this 2x2 MMI revealed that the output phase difference was not $\pi/2$ as it should be, leading to a high insertion loss ($\approx 3\,\mathrm{dB}$) and poor ER in the MZI (de)multiplexing.

Therefore, in the second run a directional coupler was used instead. This component has the advantage that the phase difference is always $\pi/2$ and that the insertion loss is negligible, however the bandwidth where a splitting ratio of $50/50$ is achieved, is much smaller than in a MMI. To determine the directional coupler design, measurement data from the 'gollum' design sweep was analyzed. To account for the fact that the passive components will be coated with DVS-BCB, an oxide clad sample was used. From this data, it was extracted that the optimal directional coupler with bend radius of $5\,\mu\mathrm{m}$ should have a gap width of $200\,\mathrm{nm}$ and straight coupler length of $10\,\mu\mathrm{m}$. To account for process variations and the fact that the sample is clad by DVS-BCB instead of oxide, designs with lengths of 8, 10 and $12\,\mu\mathrm{m}$ were put on the mask.

Next to this, also for the 2x2 coupler an improved 2x2 MMI was recently designed. In this case, the width is $5\,\mu\mathrm{m}$ and length is $100\,\mu\mathrm{m}$.

Mach Zehnder Interferometer (MZI)

As pointed out above, the label should be situated around $1500\,\mathrm{nm}$, while the data should be in the band around $1550\,\mathrm{nm}$. Therefore, the 1x2 MZI for (de)multiplexing the label and data is designed with a FSR of $35\,\mathrm{nm}$ as this separates the data and label bands by 1.5 times the FSR, but still allows for additional freedom in case the two bands should be put closer together. The arm length difference required to obtain this FSR can be calculated by

$$\Delta\lambda_{FSR} = \frac{c}{n_g l} \tag{5.1}$$

where $n_g \approx 4$ is the group index of the silicon wire. From this, an additional arm length of $16.4\,\mu\mathrm{m}$ was obtained.

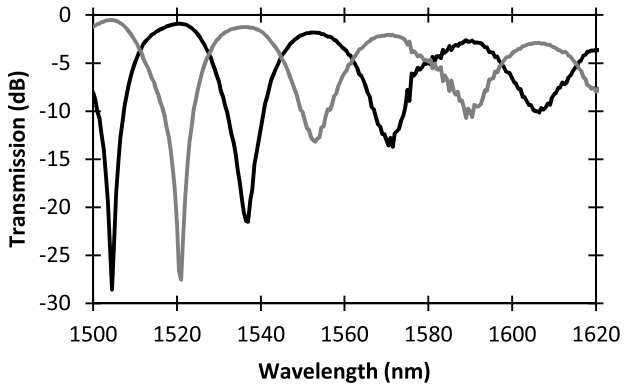


Figure 5.8: Transmission of the MZI design using a $10\,\mu\mathrm{m}$ long directional coupler as 2x2 coupler.

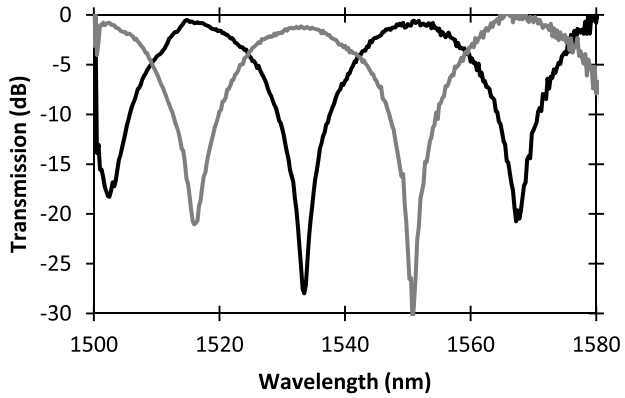


Figure 5.9: Transmission of the MZI design using improved MMI as 2x2 coupler.

Furthermore, to increase flexibility of the circuit, both arms of the MZI are equipped with a $355\,\mu\mathrm{m}$ long spiral to allow a simple top heater to tune the MZI response over a full 2π tuning range. As described above, the coupling sections of the MZI consist of a 1x2 MMI on one side and a directional coupler on the other side. The through and cross transmission response of the fabricated MZI are shown in figure 5.8. An ER of more than $10\,\mathrm{dB}$ with an insertion loss of $1-2\,\mathrm{dB}$ was obtained in the wavelength range $1500-1600\,\mathrm{nm}$, with the best performance around the label wavelength. The loss of the MZI is caused by an imperfect design in combination with the limited bandwidth of the 1x2

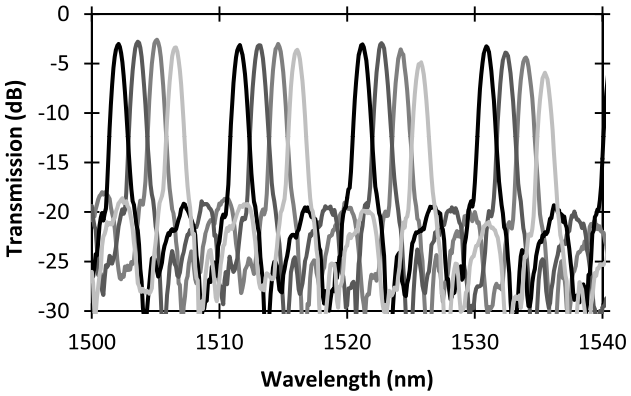


Figure 5.10: Transmission of the cyclic AWG design.

MMI. This design was used for the measurements described in section 5.5.3.

Using the recently improved 1x2 MMI and 2x2 MMI as coupling sections, the MZI response could be improved. Next to a high ER over a large bandwidth, also the insertion loss is kept below $1\,\mathrm{dB}$. The measurement of the transmission of the improved design is shown in figure 5.9. This design was not used yet in a full switch, but could be useful in future experiments.

Arrayed Waveguide Grating (AWG)

The AWG plays the role of label processor. The idea is that each output channel of the switch corresponds to a certain label wavelength within the label band, and that the AWG demultiplexes the incoming labels and sends them to the right output channels. The AWG was designed to operate at a center wavelength of $1505\,\mathrm{nm}$, with a channel spacing of $200\,\mathrm{GHz}$. As the AWG design does not allow for straightforward thermal tuning, the FSR was designed to be as low as $9.6\,\mathrm{nm}$ as this allows for some crude wavelength flexibility for the labels. The design was generated using IPKISS [93] as described in [94], using dispersion data for oxide clad silicon waveguides. The response of the AWG is shown in figure 5.10. The measured insertion loss was $3\mathrm{dB}$ with a channel cross talk of at least $-17\,\mathrm{dB}$, which provides for sufficient channel isolation considering the strongly non-linear absorption of the membrane gates.

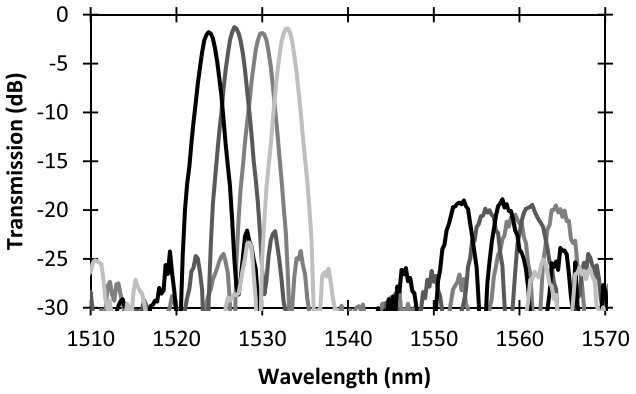


Figure 5.11: Transmission of the improved AWG design.

Next to this, also an improved AWG was designed and fabricated based on [95] in order to achieve a lower insertion loss. The drawback of this design in our case is the higher FSR, which makes it more important to have a good control of the fabrication processes used. The transmission of the improved AWG with $400\,\mathrm{GHz}$ channel spacing is shown in figure 5.11. The insertion loss could be reduced to $< 2\,\mathrm{dB}$, while the channel crosstalk remained at the same level of $-17\,\mathrm{dB}$. This design was not used yet in a full switch, but could be useful in future experiments.

Waveguide crossings

For the waveguide crossings, the design as described in [96] was used. This yields a cross talk level of $-40\,\mathrm{dB}$, and an insertion loss of only $-0.16\,\mathrm{dB}$ per crossing.

5.5 Measurements

As a full packet switch is a complex device, first the operation of the single gate is experimentally verified. In the following, the gating performance is evaluated in two proof-of-principle experiments. Then, in section 5.5.3 the full switch is tested.

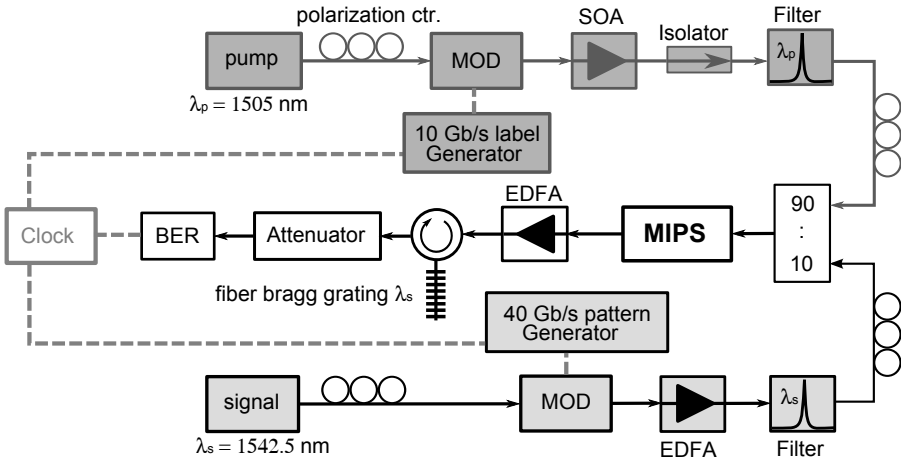


Figure 5.12: Setup for single gating experiment.

5.5.1 Single gate performance

To prove that the III-V membrane is suited to be used as a gate for implementation in the all-optical switch, we first have to show that the OSNR of a gated data signal is not severely degraded by the gate operation. To do this, we have used the experimental setup depicted in figure 5.12. The gate used in this experiment had a length of $150\,\mu\mathrm{m}$ and a width of $1.4\,\mu\mathrm{m}$, for which the basic transmission measurements already have been described in section 3.7.2. A data signal was sent at a wavelength of $1542.5\,\mathrm{nm}$ with an average on-chip power of $-12\,\mathrm{dBm}$ together with a pump beam at a wavelength of $1505\,\mathrm{nm}$ and on-chip power of $2\,\mathrm{dBm}$. For this pump power, the measured gate exhibits an extinction ratio of over $30\,\mathrm{dB}$ (see section 3.7.2) and has an insertion loss of $4.5\,\mathrm{dB}$ due to incomplete bleaching of the device absorption. The device was coupled to the outside world using two grating couplers, showing a total measured setup insertion loss of $14.5\,\mathrm{dB}$. Using this setup, three experiments were performed.

First the performance of the switch was characterized under constant pumping. The data signal was modulated at a bitrate of $20\,\mathrm{Gbit/s}$ using two different PRBS of respectively 2^7 and 2^{31} bits and transmitted through the device, to test if there is any pattern dependence of the switch operation. The BER was then determined at the output of the device and compared to a back-to-back BER measurement for these two different PRBS. The results are shown in figure 5.13(a). As there is

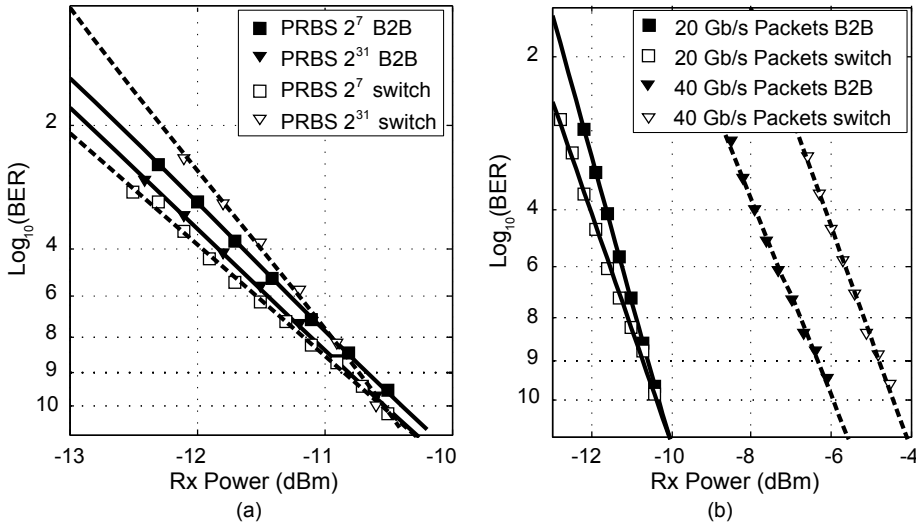


Figure 5.13: BER measurements through a single gate. (a) BER curves for a back-to-back and through the gate measurement for different pattern lengths. (b) BER curves for back-to-back and through the gate measurement at both $20\,\mathrm{Gbit/s}$ and $40\,\mathrm{Gbit/s}$.

no significant difference between the measured BER curves, it is clear that the membrane gate shows no measurable pattern dependence.

To be able to serve as an all-optical packet switch, the device needs to switch fast enough between the steady state situations with and without pump. Therefore, in a second experiment the switch on/switch off times of the device were measured by applying a realistic switching signal at a wavelength of $1505\,\mathrm{nm}$, which is on during a period of $195\,\mathrm{ns}$ and off during a switching window of $5\,\mathrm{ns}$. By monitoring the resulting modulation of an injected CW probe signal at $1542\,\mathrm{nm}$ with an optical oscilloscope, the switching times can be extracted. In figure 5.14(a) traces of the rising and falling edge of such a switching window are shown. From this the switch on and the switch off time are determined to be respectively $400\,\mathrm{ps}$ and $1.3\,\mathrm{ns}$. Using a switching window of $5\,\mathrm{ns}$, this is more than fast enough to ensure that the achieved ER in the CW experiment is also achieved in a dynamic switching experiment. To demonstrate this, in figure 5.14(b) a trace of a switching window between two switched packets is shown.

In the last experiment, the pattern generator for the data signal was programmed to create $195\,\mathrm{ns}$ long packets with a $5\,\mathrm{ns}$ switching window in between at data rates of both $20\,\mathrm{Gbit/s}$ and $40\,\mathrm{Gbit/s}$. During

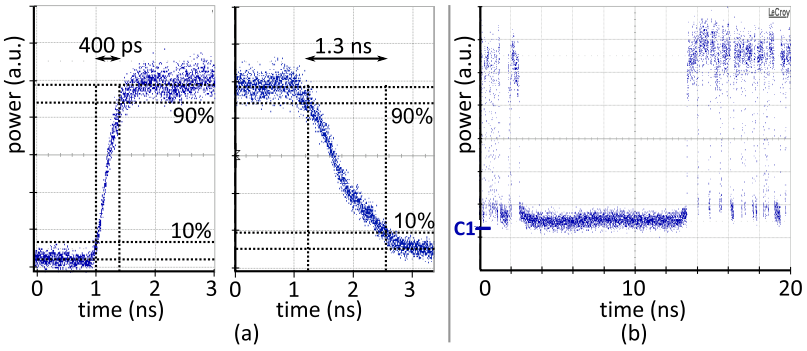


Figure 5.14: (a) Trace of gate switching on and off. (b) Trace of two packets with a switching window in between.

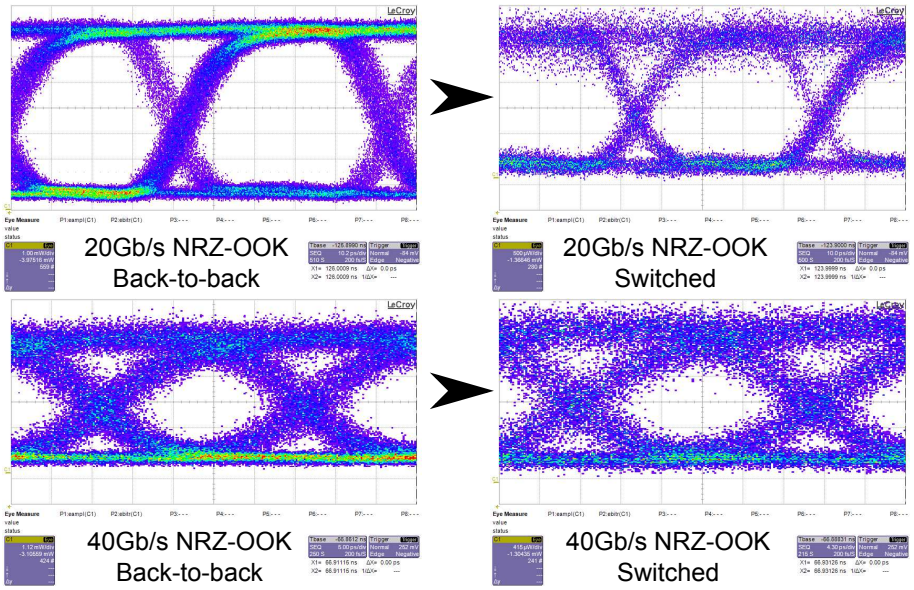


Figure 5.15: Eye diagrams for both the back-to-back and through the switch case at both $20\,\mathrm{Gbit/s}$ and $40\,\mathrm{Gbit/s}$.

this switching window, the pump was switched off and on again, to emulate actual switching of the gate. Both BER curves and eye patterns were recorded and compared to a back-to-back measurement. The resulting BER curves can be found in figure 5.13(b) and the eye patterns in figure 5.15. It is clear that for the $20\,\mathrm{Gbit/s}$ signal there is no switching related receiver sensitivity penalty. For the $40\,\mathrm{Gbit/s}$ signal however, a penalty of $1.5\,\mathrm{dB}$ is observed. This seems to be related to a combination of the low efficiency of the used grating couplers and the excess losses in the setup. Because of the high insertion losses in the setup, the data signal had to be amplified by an C-band preamplier EDFA to be able to measure the BER. Unavoidably, this adds noise to the signal due to the amplified spontaneous emission (ASE) emitted by the EDFA. Due to this noise, the OSNR of the data signal is degraded, leading to a power penalty. This addition of noise can be clearly seen in figure 5.15 in both the eye patterns for the $20\,\mathrm{Gbit/s}$ and $40\,\mathrm{Gbit/s}$ signals. The reason why this effect is only visible in the BER curve for the $40\,\mathrm{Gbit/s}$ signal is that the extinction ratio at the point of generation of this signal was much lower than for the $20\,\mathrm{Gbit/s}$ signal, making the signal more sensitive to a reduction of the OSNR.

5.5.2 Gating with clock transmission

Setup

In the second experiment, we want to demonstrate the possibility of broadcasting a clock signal to all connected clients by imprinting the clock on a longer wavelength carrier, for which the proposed gate is transparent. This allows for the use of a simple receiver with two detectors instead of a single receiver with complex, high-speed clock recovery electronics. To prove that this concept can work, a data link between central office and end user was emulated using a single gate instead of a full-switch. The full setup is shown in figure 5.16. Compared to the setup of the previous experiment, a few changes were made. First of all, a length of $25\,\mathrm{km}$ of single mode fiber was added before the chip, as in reality the packet switch would also be placed in a remote location, close to the clients (see section 5.2.1). Next to this, the wavelengths of pump and data beam are changed to fit the channels of a cyclic AWG pair, which will be used for (de)multiplexing the different wavelengths. The pump and data are moved to $1511.65\,\mathrm{nm}$ and $1549.32\,\mathrm{nm}$ respectively. Furthermore, the pump is now only on during one out of two sent packets. By setting the BER tester to expect only zeros during the

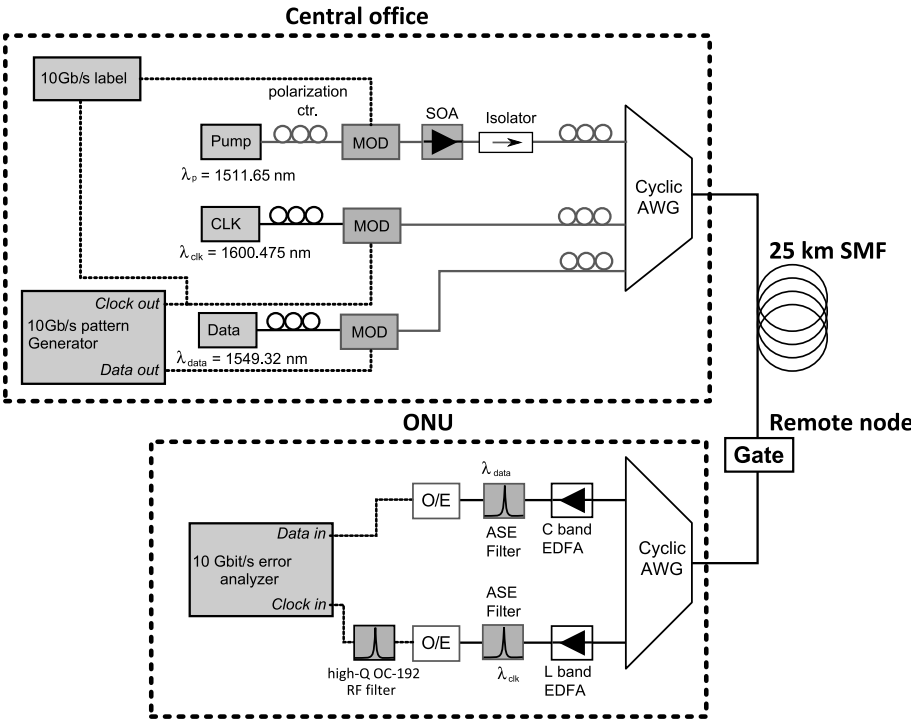


Figure 5.16: Setup for single gating experiment with clock distribution.

time when the pump is switched off, the influence of cross talk can be included in the experiment. This cross talk can be the result of a bad extinction of the pump beam or due to a lack of absorption of the gate when it is switched off. Finally, also the clock signal is transmitted through the device, at a wavelength of $1600.475\,\mathrm{nm}$.

Because a broad spectrum of wavelengths has to be coupled to the chip, a new chip had to be fabricated using the horizontal couplers as described in section 3.3.2 instead of grating couplers. The fiber-to-fiber coupling loss of this setup was measured to be $10\,\mathrm{dB}$ over a wide bandwidth. The gate used in this experiment had a length of $150\,\mu\mathrm{m}$ and width of $2\,\mu\mathrm{m}$. In figure 5.17, the measured CW transmission curves for the sample used in this experiment are shown. Unfortunately, the sample has worse characteristics than the sample used in the previous experiment. This seems to be related to a worse epitaxial layer quality, as the absorption ($26\,\mathrm{dB}$) is clearly not as a high as expected ($40\,\mathrm{dB}$) for a device of $150\,\mu\mathrm{m}$ long (see figure 3.25). The device has therefore a

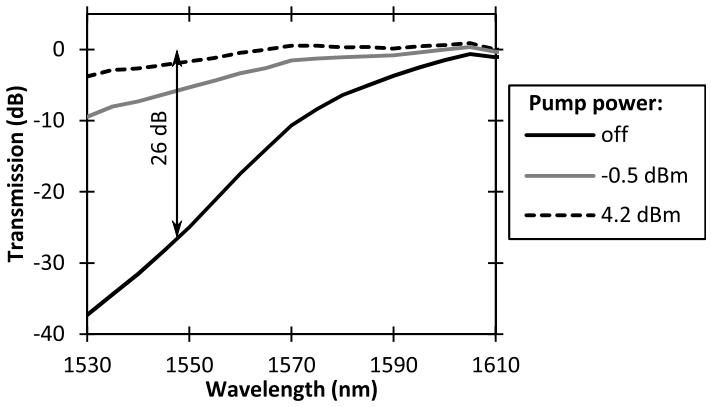


Figure 5.17: CW transmission through the gate used for the clock distribution experiment.

lower maximally obtainable ER, and at the same time a higher power is required to bleach the absorption.

At the input of the gate, the pump, data and clock signals have a power of $0\,\mathrm{dBm}$, $-9\,\mathrm{dBm}$ and $-20\,\mathrm{dBm}$ respectively. In figure 5.18, the input and output power spectrum is shown for the specified pump, data and clock signals. Using the transmission results given in figure 5.17, we can extract from this that an ER of only $12\,\mathrm{dB}$ is obtainable at the pump power level injected into the device. Note that the secondary peaks in the spectrum are made up of filtered ASE of the SOA for the pump beam by the cyclic AWG. At the receiver side, the transmitted signals are being split again using a cyclic AWG, so the presence of these peaks has no effect on the detection of data and clock signals. Due to a relatively high loss of the device, the output signal levels for data and clock both had to be amplified using two EDFA pre-amplifiers, which deteriorated the OSNR of these signals. After filtering of the ASE, the signals are both sent to an AC coupled detector. The detected clock signal is additionally filtered using a high-Q OC-192 RF filter and injected together with the detected data signal in the error analyzer to analyze the performance of the optical link. The use of AC coupled detectors is the reason why the data packet length was restricted to about $100\,\mathrm{ns}$, in order to avoid AC coupled eye closure during a switched-off packet in the error analyzer.

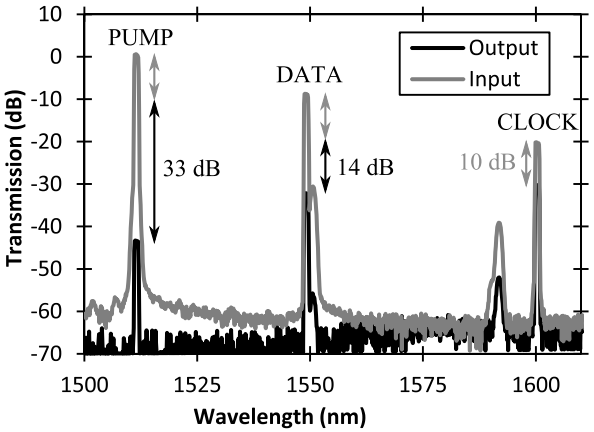


Figure 5.18: Input and output spectrum of pump, data and clock signal through single gate. The grey arrow marks the insertion loss, the black arrows mark the absorption loss in the gate.

Results

In figure 5.19 the data signal as detected by the AC coupled detector is shown after transmission through the membrane gate. Due to the fact that the pump signal is only on during one out of two packets, the membrane is only transmitting one out of two data packets. However, the achieved ER is much lower than what is potentially possible. This is due to a combination of the lack of power to bleach the device completely which limits the ER for the pump power used to $12\,\mathrm{dB}$, and the limited ER in the modulation of the pump signal, which reduces the ER even more. In a real application, seeing the very slow modulation speed needed ($5 - 10\,\mathrm{MHz}$), direct current modulation can be used to completely turn off the pump signal, thereby improving the undesirable cross talk during the off state. In figure 5.20 the bit error rate (BER) performance of the link is shown. The measured performance takes into consideration the extra $3\,\mathrm{dB}$ peak-to-average power ratio of the partially transmitted data stream. At a BER of 10^{-9} there is a power penalty of $1.5\,\mathrm{dB}$. This penalty is caused by the poor ER of the switch in this experiment as explained above, in combination with a reduced OSNR due to the required amplification by an EDFA after the switch. The fact that the slope of the BER curves differs, is related to the reduced OSNR as well. This can be understood from equation (5.2), which is reproduced from [76].

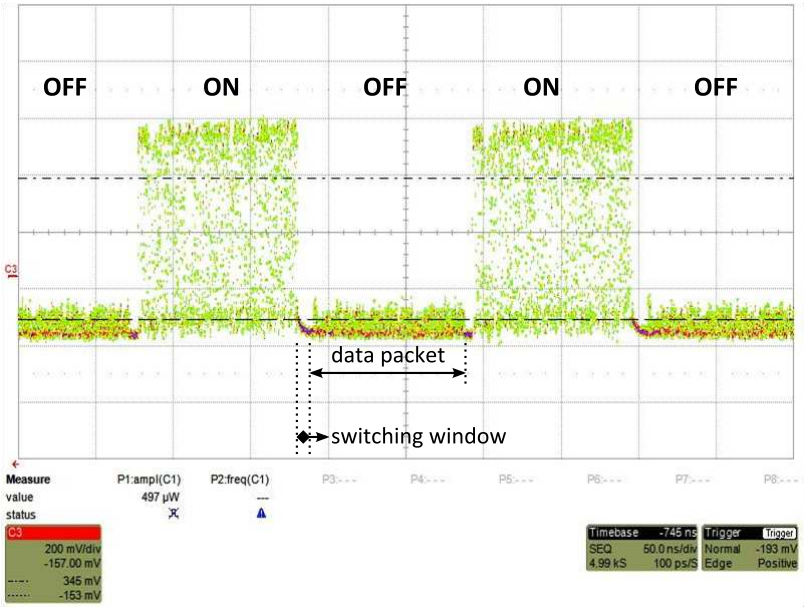


Figure 5.19: Trace of data packets after transmission through a switched on and switched off gate.

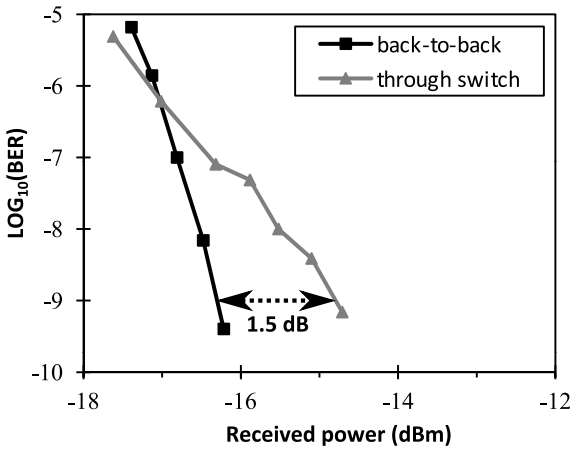


Figure 5.20: BER measurement results using the optical clock distribution.

$$\delta_I = -10 \log_{10} \left(1 - r_I^2 Q^2\right) \tag{5.2}$$

This equation relates the power penalty due to intensity noise δ_I to the Q parameter and the intensity noise parameter r_I and shows that the slope $\partial\delta_I/\partial Q$ increases with increasing intensity noise r_I. An uncommon feature in the BER curves is however that the back-to-back curve crosses the curve for transmission through the gate at high bit error rates. This seems to be related to the fact that the device is not completely bleached in its on state. Therefore, the device transmission for the data signal will be non-linear, causing the device to operate as a signal regenerator (see chapter 4). As a consequence, the ER of the data signal will be slightly improved and the corresponding power penalty at the receiver reduced, which leads to the crossing of the BER curves.

Although the device in this experiment suffered from a too high insertion loss in both setup and device causing the power penalty in the BER measurement, the usefulness of a transparent wavelength band for clock distribution could be shown, as clock recovery at the receiver could in this way be avoided.

5.5.3 1x4 switch experiment

Next to these experiments on stand alone gates, also a full 1x4 switch was fabricated. An optical microscopy image of the fabricated switch can be seen in figure 5.21. Because of the difficulties with the epitaxial layer quality of the devices made for the clock distribution experiment, we decided to use the epitaxial layers optimized for the regeneration application (see section 4.5.2). This is of course not an optimal solution, however much better gating results are obtained than in the clock distribution experiment. The drawback is that the devices have a lower carrier lifetime and also absorb at longer wavelengths, which leads to a lower efficiency and complicates the implementation of clock distribution for the full 1x4 switch. In the following, first the static performance of the switch is characterized. Then, also the dynamic switching is investigated.

Static performance

In a first experiment the static performance of the switch was characterized by injecting both a CW label corresponding to one of the output channels and a CW data signal for which the wavelength was swept

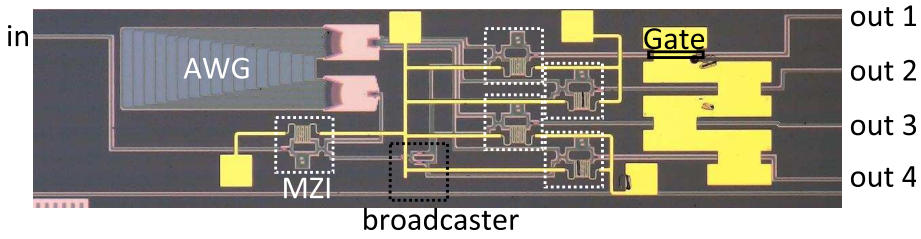


Figure 5.21: Optical microscopy image of fabricated 1x4 switch.

over the C-band. In this way, the intrinsic ER of the switch can be obtained as well as the effect of the cross talk in the AWG and MZIs. The label power was $10.6\,\mathrm{dBm}$, and the data signal power was $-0.6\,\mathrm{dBm}$ before coupling into the chip. The heaters on the MZIs are tuned to align the response with the chosen label and data bands. In figure 5.22, infrared images are shown, demonstrating the operation of the passive components which route the labels to the right output gate. Both the MMIs and AWG have a strong signature in the infrared image as these components unfortunately show a relatively high loss, as was characterized above in section 5.4.2.

In figure 5.23 we show the behavior at a single output port of the switch when alternating the label wavelength between the different output port labels. The used label wavelengths were at $1510.77\,\mathrm{nm}$, $1509.22\,\mathrm{nm}$, $1507.80\,\mathrm{nm}$ and $1506.35\,\mathrm{nm}$ for ports 1-4 respectively. In the inset, also transmission at the other output ports is shown. A high port isolation of more than $25\,\mathrm{dB}$ could be achieved in all the output ports across the C-band, up to more than $30\,\mathrm{dB}$ for $1550 - 1560\,\mathrm{nm}$, the intended wavelength range for the data signal. Furthermore, tuning the label wavelength to another output port nearly has the same effect on the transmission of the data signal as turning the label power off: the ER changes only by a few dB. Although the cross talk levels in the passive components (AWG and MZIs) are much worse than $30\,\mathrm{dB}$, the non-linear absorption in the membrane gate reduces its effect on the actual device cross talk.

The achieved insertion loss is of the order of $26\,\mathrm{dB}$, $20\,\mathrm{dB}$ higher than the intrinsic broadcasting loss. There are several causes for this loss. A first cause is the fiber-to-chip coupling loss which was determined to be $3.5\,\mathrm{dB}$ per transition in this chip. Furthermore the passive components used in the switch architecture were not ideal (see section 5.4.2), and excess losses due to MMIs and MZIs for the data signal ac-

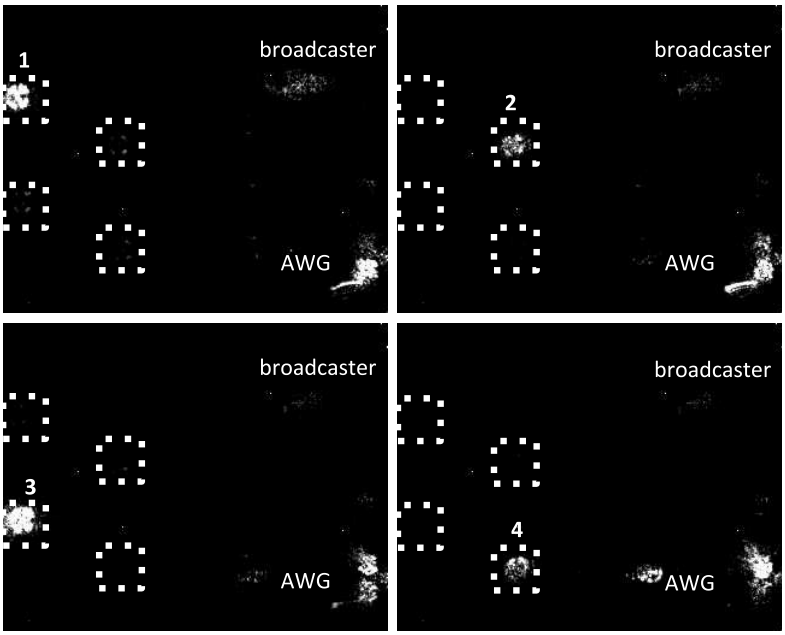


Figure 5.22: Static infrared images of label traversing the passives in the 1x4 switch. By changing the wavelength of the label, the label clearly ends up at the different MZI multiplexers at each output gate.

cumulated to approximately $4\,\mathrm{dB}$. Obviously also the broadcasting of the data signal yields a loss of $6\,\mathrm{dB}$. The remainder of the loss ($9\,\mathrm{dB}$) can be attributed to scattering losses and insufficient pumping of the membrane gates.

Dynamic performance

To demonstrate the operation of the switch, also a BER experiment was performed. The setup that was used is depicted in figure 5.24. To achieve a better ER in the label than in the clock distribution experiment, the label is modulated by directly modulating a C-band SOA, as the label can be completely turned off in this way. To better exploit the gain of the C-band SOA that was used to modulate the label, a different set of label wavelengths was chosen in comparison with the static measurement. The labels were put at $1535\,\mathrm{nm}$, $1536.6\,\mathrm{nm}$, $1538.2\,\mathrm{nm}$ and $1539.8\,\mathrm{nm}$. In the switch, the heaters were tuned to align the MZI filter curves to the chosen label and data bands. The data signal was created by modulating a laser at $1559.38\,\mathrm{nm}$ using with a $10\,\mathrm{Gbit/s}$ pattern

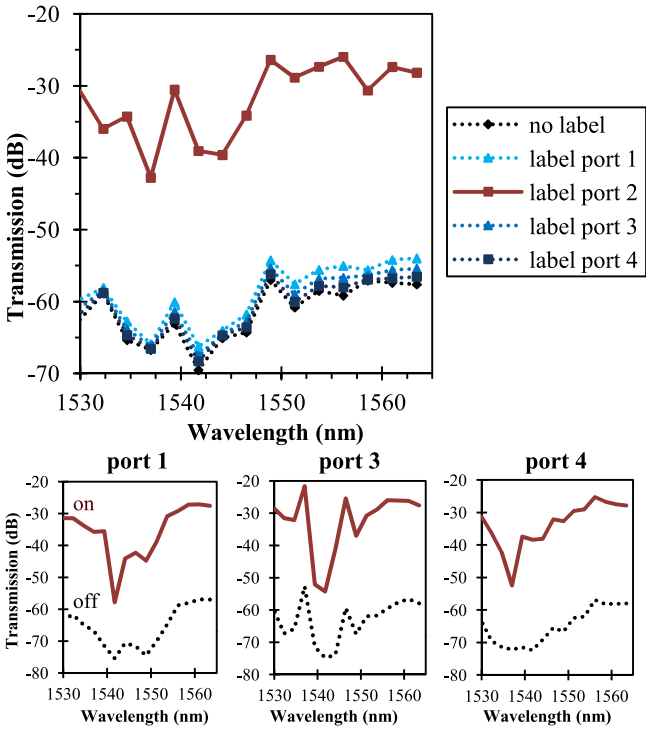


Figure 5.23: Static transmission through 1x4 switch. A high port isolation is demonstrated for all ports.

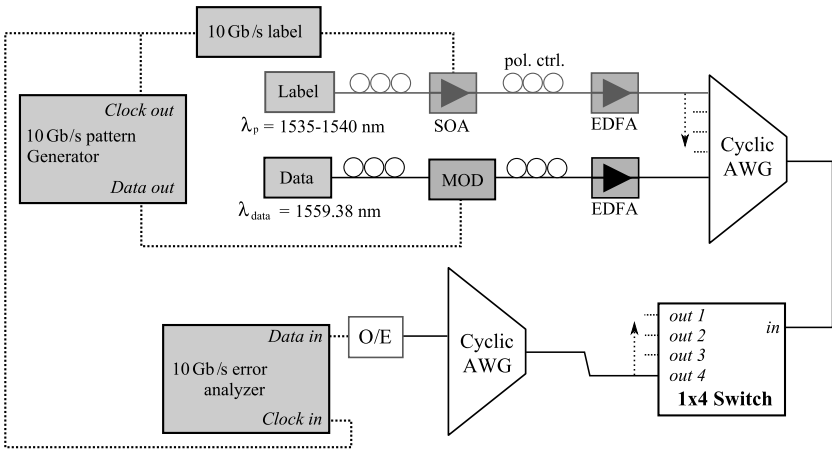


Figure 5.24: Setup for measurement of the BER of the 1x4 switch.

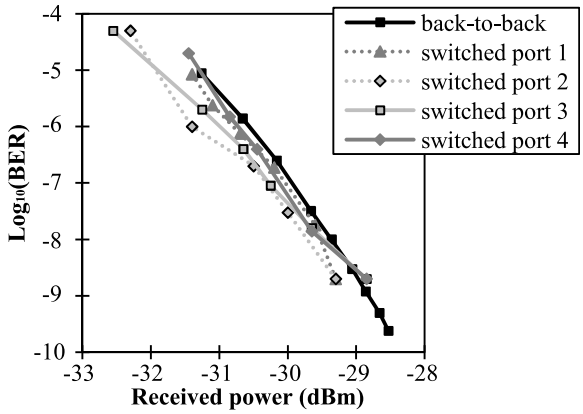


Figure 5.25: BER measurement for the 1x4 switch.

generator running a $2^{31} - 1$ PRBS in a NRZ-OOK modulation scheme. Due to the fact that this initial design still suffers from too high insertion losses, both the data signal and label had to be amplified using an EDFA to boost the input signal power to $4.5\,\mathrm{dBm}$ and $13\,\mathrm{dBm}$ respectively. After combining data and label in an AWG implemented on a planar lightwave circuit which is connected with single mode fiber, the signals are sent through the 1x4 switch. At the output, the light is split again using an AWG, and the data signal is detected by a detector and sent to the BER analyzer. The receiver used for these measurements was an AC coupled avalanche photodiode housed inside an XFP module with a typical receiver sensitivity of $-28\,\mathrm{dBm}$.

In figure 5.25 the measured BER is shown for a PRBS routed through each of the output ports, while keeping the label operating in CW. There is no measurable power penalty related to the transmission through the switch. Since the detector used in this experiment was AC coupled, it could not be used to measure the BER for packetized data. Therefore the AC coupled avalanche detector was replaced by a DC coupled Agilent 8192A $10\,\mathrm{Gbit/s}$ receiver. Also in this case, error free operation could be achieved, with a penalty of $0.8\,\mathrm{dB}$. In addition we show a time trace together with the eye diagrams of the packetized data for the back-to-back case and at the output of port 1 in figure 5.26. The label is modulated in such a way that only one out of four packets is being transmitted. The resulting time trace clearly shows the high ER that can be obtained with the switch. The eye diagrams showing the packetized data were captured with the DC coupled Agilent receiver. From

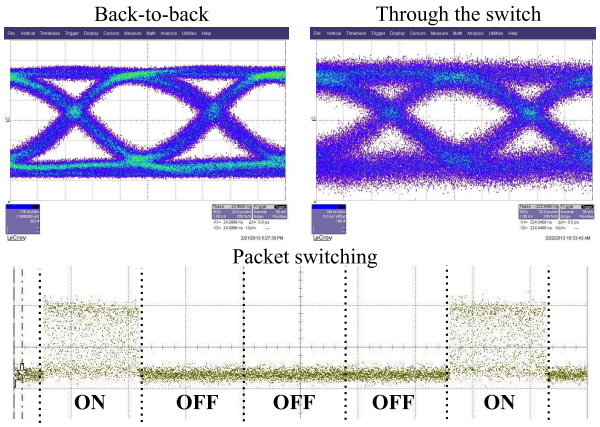


Figure 5.26: Top: eye diagram before and after switching. Bottom: trace of data where only one out of four packets is sent to the output port.

these, we can see that except for some additional noise both eyes are identical and open, confirming the result obtained with the BER measurement. For the eye after transmission through the switch, the optical input power at the receiver was much lower ($-18.5\,\mathrm{dBm}$ compared to $-10.2\,\mathrm{dBm}$ in the back-to-back case), which explains the increased electrical noise in the detected signal.

5.5.4 Summary

In the above experiments, we have shown the potential of the membrane gate as a switching element in a broadcast-and-select switching fabric. An ER of more than $30\,\mathrm{dB}$ was achieved in a membrane gate of only $150\,\mu\mathrm{m}$ long. The device is also more than fast enough to switch during a relatively short time window of $5\,\mathrm{ns}$. Furthermore, using a BER measurement, it was demonstrated that the membrane gate does not introduce a measurable power penalty, and shows no pattern dependence in its operation. Finally, the gate was implemented in a full 1x4 broadcast-and-select switch. Its operation was demonstrated using both static and dynamic experiments.

Next to its merits as a gate, the membrane waveguide also has the advantage that it is transparent for wavelengths with energy lower than the device's band gap energy. This is useful in network topologies where one wants to provide a broadcasting service (such as television) as well as switched packetized data. Furthermore, this can be

used to broadcast a clock signal across the network, which simplifies the receiver necessary in the ONU at the client side. Instead of needing a complex clock recovery circuit, a simple WDM splitter and extra detector suffice to get the clock signal at the receiver side.

5.6 Prospects for the future

Although the switch works as expected, the large excess insertion loss for the data signal ($20\,\mathrm{dB}$), and the related high pump power required ($> 13\,\mathrm{dBm}$) are currently limiting the applicability of the switch. In fact, the switch would actually only be considered as a viable alternative for a passive splitter, if its insertion loss would be less than the loss of a passive splitter. This is because the total cost of the network link is the most important variable to take into account. As the infrastructure for a switched access network is more expensive, such an architecture will only be a viable option, if the network reach and number of clients that can be connected to a single network link are increased as well. Therefore, the insertion loss should be as low as possible to allow for extra network reach and connected clients.

The switch presented in this paper does not reach these specifications, as this means that an insertion loss of less than $6\,\mathrm{dB}$ would be required. However, there is room for improvement in the design and fabrication of the used passives and membrane gates.

Most of the loss in the passive components can be avoided if the current state of the art passives in SOI are used. MMIs [97] have reached excess losses of only $0.06\,\mathrm{dB}$ at $1550\,\mathrm{nm}$ and remain limited to less than $0.2\,\mathrm{dB}$ in the band of $1500\,\mathrm{nm}$ to $1600\,\mathrm{nm}$. AWGs [95] have been demonstrated with an insertion loss of less than $1\,\mathrm{dB}$. Inverted taper couplers using a cleaved SU-8 waveguide for fiber-to-chip coupling have reached a coupling loss of only $0.66\,\mathrm{dB}$ [59]. These improvements reduce both insertion loss for data and clock signal, and the required label power to drive the membrane gate. Another solution to improve the losses of the passive circuit would be to move to the silicon nitride platform instead of using silicon-on-insulator. Although the devices have a larger footprint due to a lower index contrast in the silicon nitride platform, cross talk and insertion losses of passive components are generally better than in the silicon-on-insulator platform [98], especially for fiber-to-chip coupling. A drawback is that the lower index of silicon nitride compared to III-Vs makes the coupling to a bonded III-V waveguide more challenging.

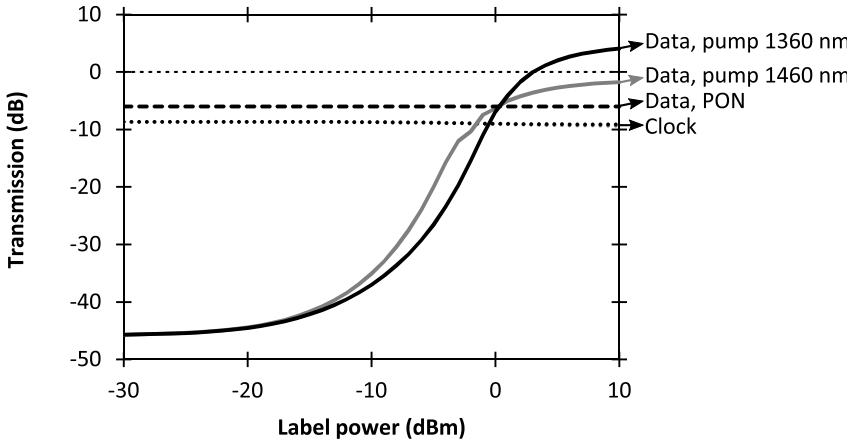


Figure 5.27: Simulation of the transmission through a 1x4 switch with state-of-the-art passives and well-passivated gates, both for pumping at $1460\,\mathrm{nm}$ and $1360\,\mathrm{nm}$.

Next to improving the passive components, also the fabrication of the membrane gate can be optimized. By further developing sulfur passivation techniques (see section 3.6.3), the effective carrier lifetime can be increased to the order of $2\,\mathrm{ns}$ [5]. Although the device becomes a lot slower, this is still fast enough to allow switching within a time window of $5\,\mathrm{ns}$. As the etched surfaces are well passivated, the device efficiency can be further optimized by moving to a narrower device width (see section 3.5.2): by reducing the width of the devices from $1\,\mu\mathrm{m}$ to $700\,\mathrm{nm}$ and including a thin additional SCH layer pair around the QWs, the amount of power required to pump the device can be reduced by approximately $30\,\%$. Both these improvements will reduce power consumption and therefore also heat generation in the device. Furthermore, assuming that the device will be fabricated using a more advanced lithography technique (e.g. i-line lithography), the alignment accuracy and feature size should be better controlled, allowing us to reduce the silicon to membrane taper length to $10\,\mu\mathrm{m}$.

To quantify the possible performance of such an improved device, a simulation was performed. For the passive components the characterization results from [97, 95, 59] were used. For the membrane gate, the device transmission was simulated using the simulation engine as described in section 3.4. To account for the improved passivation a carrier lifetime of $2.0\,\mathrm{ns}$ was assumed. Furthermore, the device will suffer much less from heat generation, and therefore a numerically simulated

QW response at $300\,\mathrm{K}$ is used in the simulation. For this simulation, the silvaco software was used, and a MQW structure with $8\,\mathrm{nm}$ InGaAs QWs was assumed.

In figure 5.27 the simulated switch response of a 1x4 switch with $150\,\mu\mathrm{m}$ long membrane gates can be seen under two pumping conditions. From this it can be seen that device performance is much better: the gate is already bleached at an off-chip label power of $-1\,\mathrm{dBm}$, while for a higher power the additional gain provided by the membrane compensates for insertion and splitter losses, and can even lead to a small off-chip gain. The gain that can be achieved is strongly dependent on the label wavelength however, as the highest free carrier density that can be generated at a particular label wavelength is the transparency carrier density, which is higher for shorter wavelengths. At a label wavelength of $1360\,\mathrm{nm}$, the result is $9\,\mathrm{dB}$ extra power budget compared to a PON. This additional power budget can be used to connect extra clients, increase the distance to the central office or increase the data rate transmitted over the network link.

In order to operate the switch in a realistic environment, more clients should be connected as discussed in the design requirements (see section 5.2.5). To do this the switch would have to be extended with extra output gates, and the AWG would have to be redesigned to demultiplex more label wavelengths. The effect on the data signal transmission is the same as in a PON: connecting more clients leads to a higher intrinsic splitting loss. However, to connect up to 128 or more ONUs, extending the switch is not straightforward anymore. In this case, a combination between WDM and the switched architecture could be used. For instance, four different data and label bands can be split using two consecutive MZIs for demultiplexing, after which the four separate bands can be sent to their corresponding switch of e.g. 32 output ports. In this way, 128 users could be connected.

Another important issue with operation in the field is that the incoming polarization of label, data and clock wavelengths is not known. One possible solution for this is by polarization diversity, where two identical circuits are used for both TE- and TM-polarization. One difficulty however, is that the absolute label power is important to achieve switching. Therefore, the generated label should consist of an equal amount of TE- and TM-polarized light, in order to receive also at the remote node an equal amount of power in both TE- and TM-polarization. This could be done using a quarter waveplate at the CO, to convert the linear polarization emitted by the lasers into circularly polarized light.

5.7 Summary

In this chapter we have concisely reviewed the different possible network architectures to implement a FTTH network. Currently the PON solution is most widely deployed, however using all-optical switches instead might offer several advantages. However, the realization of such a switch is not trivial, as the switches should be cheap to produce and have a low power consumption. Because membrane gates have to potential to become low power switching elements, a study was made which switching fabrics are possible using these devices. Of these architectures, the broadcast-and-select architecture was the most simple to implement. In a first stage a few proof-of-concept experiments were performed on the single gate. It was shown that the gate can provide for a very high ER, and is at the same time fast enough with a potentially low operating power. Next to this, a full 1x4 switch was designed, fabricated and experimentally characterized. This switch provided for more than $25\,\mathrm{dB}$ channel isolation, with no measurable power penalty in a BER measurement. Because of the high insertion losses however, practical application is not possible yet. However, there are several possibilities to further improve the performance: by using current state-of-the-art passives, and improving the passivation of the etched sidewalls of the membrane gate, the all-optical switch could outperform PONs in available power budget, and therefore become a viable alternative.

Chapter 6

Conclusions

6.1 Conclusions

In this work, several applications were highlighted where power efficiency is extremely important. For the realization of on-chip optical interconnects, it is found that the requirements for an integrated laser are very strenuous. Such a laser would need to have a low power consumption, a small footprint, and a high tolerance for temperature changes. For that end, a novel waveguide structure was developed in this work, that allows for the optimization of the optical confinement inside the active region, in order to increase the net modal gain that can be achieved. It was shown that the net modal gain could be doubled compared to the classical approach, leading to a reduction of the current density required to achieve a certain gain, and therefore potentially a more efficient laser. Using such a waveguide structure inside a micro laser cavity, would allow for a more compact device with a low threshold current.

A second important component for the realization of optical interconnects at the highest speeds is an integrated modulator. Also in this case, the device power consumption should be minimized. Furthermore, a high enough operation speed must be maintained. To achieve these features, a heterogeneously integrated III-V modulator was designed that works using the quantum confined stark effect, as it has already been demonstrated that these modulators have the best characteristics. A waveguide structure was developed that minimized the inherent capacitance of the modulator design. This allows obviously for a high bandwidth, as the bandwidth is limited by the RC-constant of the electrical circuit, but also for a lower power consumption. This

is possible because a QCSE modulator is driven by the electric field inside the active region. As the capacitance is quite low in the proposed waveguide design, we can afford to increase it a little as this leads to a lower voltage to achieve the same electric field in the active region. In this way, it was shown in simulation that such an optimized modulator could achieve a bandwidth of $26.5\,\mathrm{GHz}$ when driven by lumped electrodes. A peak-to-peak driving voltage of only $0.45\,\mathrm{V}$ is required to achieve an ER of more than $10\,\mathrm{dB}$. Furthermore, a reverse bias is not required, as the built-in electric field in the thin active region is already high enough. As a result, a very low power consumption of $48\,\mathrm{fJ/bit}$ at a bit rate of $40\,\mathrm{Gbit/s}$ is predicted.

Another application that was highlighted, is access networks in telecommunications. The growing bandwidth requirements generated by new services like video-on-demand and cloud computing are driving the move from a last mile over copper wiring towards fiber-to-the-home networks. Currently, mostly passive optical networks are deployed, however these suffer from several disadvantages. First of all, they are not very power efficient, as the power is just split between the different connected users. Furthermore, as the optical network unit there also receives the downstream data meant for other users, a lot of power is wasted in the high-speed processing of data which is going to be discarded anyway. To address this problem, an all-optical switch architecture was proposed. In such an architecture, each optical network unit only receives useful data, and therefore a lot of power can be saved, e.g. by using a sleep mode when no data is incoming. Next to this, in a switched architecture the insertion loss could be made lower than the intrinsic splitter loss.

To be able to realize such a switch, first a heterogeneously integrated gate on SOI was developed, which can be fully remotely controlled using a pump beam. This is important, as the switches will have to be placed in locations where typically no electrical power is present. It was shown that these devices have a large potential for gain, as a gain of more than $8\,\mathrm{dB}$ was demonstrated in a device of only $100\,\mu\mathrm{m}$ long, using a pulsed pump. In CW however, the demonstrated gain remained limited to approximately $2\,\mathrm{dB}$, due to the heating of the device. It was found however, that by improving the passivation of the etched QW sidewalls, the device performance could be greatly improved. This was proven in simulation, and the first fabrication attempts were made, with a $20\,\%$ improvement in the carrier lifetime.

Using this membrane device, a 1x4 all-optical switch was then developed. In a first experiment, it was shown that a single gate allows for a very high extinction ratio of more than $30\,\mathrm{dB}$, and that furthermore there is no power penalty related to the transmission through the gate. In a second experiment, it was shown that the transparent band for wavelengths longer than the bandgap wavelength can be used to broadcast data, such as television services, or even to broadcast a clock over the access network. In this way, the optical network units could become much less complex, as then no clock recovery circuitry is required anymore. Finally, also a full 1x4 switch was demonstrated. The selected switching fabric was the broadcast-and-select architecture, as this was the most robust architecture for variations in the processing of the passive silicon components. Using this architecture, transmission to all the output ports was achieved without a measurable power penalty. Furthermore, a high port isolation of more than $25\,\mathrm{dB}$ was achieved over a broad signal wavelength range. However, the excess insertion loss of the switch compared to a simple splitter was $20\,\mathrm{dB}$. This loss was related to the use of both imperfect passive components and the poor performance of the used gates.

The membrane gate can also be used in other signal processing applications. By optimizing the gate for a higher speed, and using an adapted MQW active region, a passive regenerator could be realized. The device works using the principle of saturable absorption, where the signal zero-level is being absorbed more than the signal one-level, and in this way the ER of the signal can be improved. Regeneration was demonstrated at $2.5\,\mathrm{Gbit/s}$, with an ER improvement from $2\,\mathrm{dB}$ at the input up to $6.2\,\mathrm{dB}$ using a $100\,\mu\mathrm{m}$ long device. This lead to a receiver sensitivity improvement of $3.6\,\mathrm{dB}$ over the entire C-band, up to $4.5\,\mathrm{dB}$ at wavelengths $1530 - 1540\,\mathrm{nm}$. In the design of this component, it was found that the most important trade-off was the achieved extinction ratio improvement compared to the insertion loss of the device.

6.2 Outlook

It would be very interesting to see an application of the novel waveguide structures for electrically pumped devices, especially the V-waveguide for the implementation of a QCSE modulator. This one has the advantage that the processing is quite standard: only contact lithography is required, and most of the etching can be done by wet chemical etching. Furthermore, a modulator based on this waveguide

structure is an improvement in every way on the more classical waveguides which are currently used [20, 3, 21].

The membrane waveguide has much potential as a possible low-power optical gate, but due to the immaturity of the fabrication process high CW amplification could not be shown. This could be solved on the short term however, as an optimized passivation procedure using the proposed ammoniumsulfide solution should lead to much better results [5]. Next to this, working at a shorter pump wavelength should also improve the gain.

The switch design in the demonstrated prototype was also suffering from the excess losses in the passive components. As we showed in section 5.6, using optimized SOI components which have already been demonstrated in literature, would lead to a significant improvement in the switch performance, to the point that it would actually have a lower insertion loss than a passive splitter. However there is also another option: the other proposed switch architectures which were dismissed in this work because of an even higher dependence on the performance of the passive components, might in the end lead to better results. Both the concatenated MZI and ring resonator architecture work as true switches, and can therefore in principle have a very low insertion loss, especially when the number of output ports becomes larger. Assuming that the passives can be well controlled, only the free carrier absorption in the membrane is intrinsically causing a small excess loss. Furthermore, when the passivation procedure for the membrane waveguide is developed, also these switching architectures will have the benefit of a lower operating power. Especially the ring resonator based architecture could have a very small operating power if thermal control can be avoided. This is because of the small phase change required to shift the resonances enough to achieve switching, as opposed to the concatenated MZI architecture where a π-phase shift is required. However, to achieve that a very good control of the fabrication of both the ring resonator and the III-V on top would be needed.

Appendix A

Physics

In this chapter we give an overview of the different physical mechanisms in a III-V membrane that are used within this thesis.

A.1 Electrical properties of a III-V membrane

A III-V membrane typically consists of a stack of layers that is especially designed for a certain function. In case one wants to create a amplifier, it is very important that carriers can be injected efficiently into the III-V membrane, and that they are trapped inside the active region. Therefore, typically an electrical confinement structure is defined within the active region of the device.

The explanation of different physical concepts below is greatly simplified. The only purpose is to provide for sufficient knowledge to understand the account in the thesis.

A.2 Electrical confinement of free carriers

Electrical confinement of free carriers is very important for most applications. Without structures to achieve confinement of free carriers, pumping a semiconductor to achieve gain would be nearly impossible, as the carriers would diffuse and drift away continuously. In the following, heterojunctions and QWs are discussed.

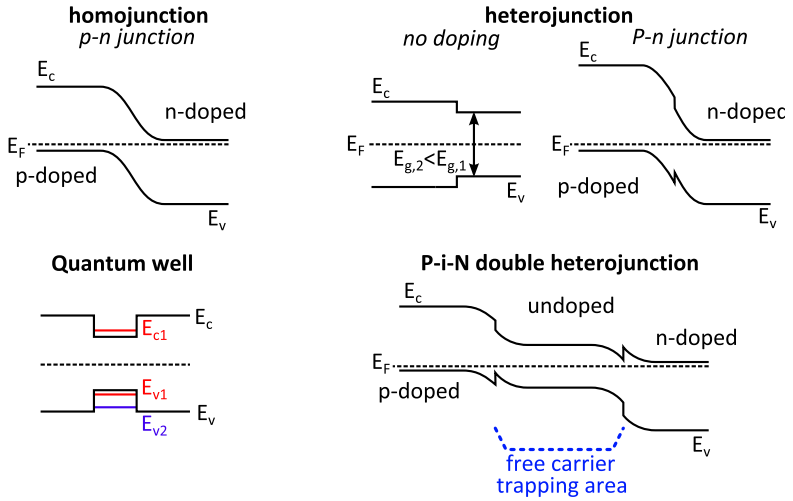


Figure A.1: Energy diagram of electrical confinement structures, showing conduction and valence band energies of the semiconductor materials.

A.2.1 Heterojunctions

Heterojunctions are junctions formed by bringing two materials together with a different band gap. Similar as in a junction formed by one material with different doping on either side of the junction, the fermi energy of both materials forming the junction are aligned in steady state. In the case of a heterojunction, the resulting band diagram can be engineered to create an energy barrier for both electrons and holes (see figure A.1, top). Furthermore, if two heterojunctions are brought together such that the center material has the smallest band gap, electrons and holes can be confined in both directions, allowing efficient pumping of this center semiconductor region (see figure A.1, bottom).

A.2.2 Quantum Wells

Quantum wells are in fact a special case of an undoped double heterojuction as described above. By reducing the size of the center semiconductor, the electrons and holes will be confined in a smaller space. When this size approaches the Broglie wavelength of an electron however, the classical 'electron as a particle' picture is not valid anymore and the wave character of the electrons have to be taken into account. By solving the Schrödinger equation, it can be derived that a quan-

tum well only supports certain allowed energy levels for the carriers trapped inside the well (see figure A.1). The lowest allowed energy level ($E_{c,1}$ and $E_{v,1}$ for conduction and valence band respectively) is determined by the width of the quantum well, which is very useful as it allows us to engineer the band gap of a quantum well merely by changing the geometry. Next to this, also other properties such as the density of states (= how many electrons and holes fit in per unit volume) are different from bulk material. One of the consequences is that quantum wells have a higher differential gain than a bulk active region. In other words a change in the carrier concentration will have a bigger effect on the response of a quantum well than of bulk material.

A.3 Carrier transport

A.3.1 Diffusion

When carriers are not confined, e.g. by a heterojunction or etched sidewall, diffusion will play a role. The equation governing carrier diffusion is similar to the diffusion of heat or macroscopic particles (e.g. ink in water). This is because the concept of diffusion is related to random movement, which is present in all these systems. Therefore, the current as a result of carrier diffusion can be described by Fick's law

$$J_n = -D_n \nabla N(\vec{r}, t) \tag{A.1}$$

$$J_p = -D_p \nabla P(\vec{r}, t) \tag{A.2}$$

where D_n and D_p are the diffusion constant for respectively electrons (N) and holes (P). This equation actually describes that charge carriers will move from areas with higher concentration to areas with lower concentration, at a speed proportional to the gradient in the concentration. The diffusion constants can be derived using the Einstein relation:

$$D_{p,n} = \frac{\mu_{p,n} k_B T}{q} \tag{A.3}$$

However, as both electrons and holes are diffusing at the same time and these have an opposite charge, the diffusion processes for electrons and holes are not decoupled. Because of electrostatic attraction, the diffusion of electrons and holes will actually occur at the nearly the same speed. The large electric field that quickly builds up as electron and hole concentrations are not balancing each other, cause drift currents in

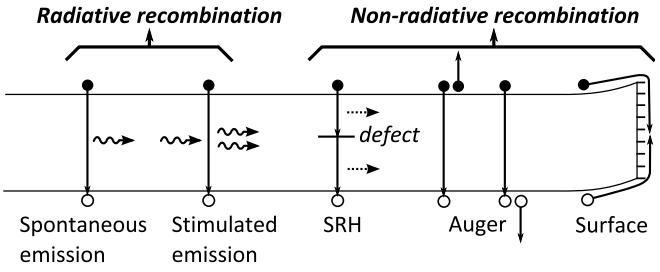


Figure A.2: Schematic overview of the different carrier generation and recombination processes.

the opposite direction, which slows down the diffusion of the electrons and speeds up the diffusion of the holes. This effect is called ambipolar diffusion [99], and the associated ambipolar diffusion coefficient that is valid for both electrons and holes can be calculated as follows

$$D = \frac{D_n \mu_p + D_p \mu_n}{\mu_p + \mu_n} \tag{A.4}$$

A.3.2 Drift

A second way for carriers to be transported is by drift current. Drift currents are the result of the presence of an electric field that attracts the charge carriers towards a particular direction. These currents can be expressed by

$$J_n = -q\mu_n EN \tag{A.5}$$
$$J_p = q\mu_p EP \tag{A.6}$$

A.4 Carrier recombination

In this section, an overview will be given of the different carrier recombination processes. In figure A.2 a schematic overview is given of the different physical processes involved. They can clearly be divided into radiative and non-radiative processes.

A.4.1 Radiative recombination

A first way by which free carriers can recombine is by radiative recombination, with the generation of a photon. This can be both a spontaneous and a stimulated process.

Spontaneous emission

In spontaneous emission two free carriers recombine to generate a photon with an energy that corresponds to the band gap energy of the semiconductor material. Because this is a one-step process, the spontaneous emission rate can be described as

$$R_{spont} = BN^2 \tag{A.7}$$

This process however is only significant in direct band gap semiconductors. The reason for this, is that in the case of an indirect band gap, a phonon has to be generated in the process as well to conserve momentum, which strongly reduces the probability of the process to occur. Therefore in indirect band gap materials such as silicon, non-radiative recombination processes are dominant.

Stimulated emission

In stimulated emission an electron and a hole recombine, stimulated by the presence of another photon with energy larger than the semiconductor band gap to generate an exact copy of this stimulating photon. This process is very important, as it is the underlying physical mechanism for optical gain. The effect will be discussed in more detail in section A.6.2 on optical gain.

A.4.2 Non-radiative recombination

Shockley-Read-Hall recombination

Shockley-Read-Hall recombination is a non-radiative process that is assisted by a defect in the semiconductor crystal. In this process, the free carrier is trapped by the defect, which sits within the band gap of the semiconductor, and recombines non-radiatively before it is released again in the conduction band. It is therefore a two-step process that can be modeled by the following equation [31]:

$$R_{SRH} = \frac{PN - n_i^2}{\tau_{p0}\left[N + n_i \exp\left(\frac{E_{trap}}{k_B T}\right)\right] + \tau_{n0}\left[P + n_i \exp\left(\frac{-E_{trap}}{k_B T}\right)\right]} \quad \text{(A.8)}$$

where τ_{n0} and τ_{p0} are the electron and hole lifetimes respectively, and E_{trap} is the dominant defect energy relative to the fermi level of the semiconductor. However, usually this equation is simplified to

$$R_{SRH} = \frac{N}{\tau_c} \quad \text{(A.9)}$$

where τ_c is the non-radiative lifetime of a free carrier in the semiconductor. Although two free carriers are involved in the process, the rate is proportional to N and not N^2, as the process occurs in two steps. Because the growth quality of III-V materials is usually very good, the defect density is low, and therefore the contribution of this effect is limited.

Auger recombination

Auger recombination is a non-radiative process that involves three free carriers. In this process, an electron-hole pair recombines and transfers its energy to another electron or hole, which is then excited within its energy band. Subsequently, the energy is released under the form of phonons. This process is usually described as

$$R_{Auger} = CN^3 \quad \text{(A.10)}$$

where C is the Auger recombination coefficient of the semiconductor, and N the free carrier concentration. As Auger recombination is a strongly temperature dependent process, a better description is given by the following equations [31]:

$$R_{Auger} = C_n(PN^2 - Nni^2) + C_p(NP^2 - Pni^2) \quad \text{(A.11)}$$

$$C_n = C_{0,n}\left(\frac{T}{300}\right)^{K_n} \quad \text{(A.12)}$$

$$C_p = C_{0,p}\left(\frac{T}{300}\right)^{K_p} \quad \text{(A.13)}$$

Because three free carriers are involved in the process, it is especially important in regions where high free carrier concentrations are reached.

Surface recombination

Similar to SRH recombination, traps formed by dangling bonds at semiconductor interfaces form non-radiative recombination centers. Although this type of recombination is only present at semiconductor interfaces, this effect can be the dominant non-radiative recombination process in some compounds. The rate of recombination is usually described in the same way as SRH recombination (see equation A.8), however with an adapted effective carrier lifetime [31]:

$$\frac{1}{\tau_{eff,n}} = \frac{1}{\tau_n} + \frac{A}{V} v_{s,n} \tag{A.14}$$

$$\frac{1}{\tau_{eff,p}} = \frac{1}{\tau_p} + \frac{A}{V} v_{s,p} \tag{A.15}$$

where A is the exposed surface area, V the semiconductor volume and $v_{s,n}$ and $v_{s,p}$ respectively the surface recombination velocities for electrons and holes. The surface recombination velocity is strongly material dependent and ranges from $< 10^4\,\mathrm{cm/s}$ for InP to $10^5\,\mathrm{cm/s}$ for InGaAs. Furthermore, it is dependent on the surface condition. It has been found that treatment of InGaAs with sulfur compounds can drastically reduce the surface recombination rate (see section 3.6.3).

A.5 Carrier generation

Free carriers can be generated in two ways: they can either be brought into the device by an electrical current (electrical pumping), or they can be created by the absorption of one or more photons (optical pumping).

A.5.1 Electrical pumping

When a current is sent through a semiconductor device, free carriers are injected. These can then be trapped inside the device by using electrical confinement structures such as heterojunctions and quantum wells (see section A.2) to make the pumping more efficient.

P-i-N heterojunction

To inject current, typically a P-i-N heterojunction is used, with quantum wells embedded into the smaller bandgap intrinsic region of the hetero-

junction. This semiconductor stack forms a diode, and has several advantages. First of all, a low electrical resistance can be obtained in the p and n-doped regions by properly doping them. Typically a graded doping profile is chosen for the p-doped region, as optical losses due to FCA are higher in a p-doped semiconductor. Therefore, a lower doping is used near the optical mode, while a higher doping is used near the electrical contacts to improve the conductivity. Next to this, the middle intrinsic region of the heterojunction efficiently traps injected carriers, such that they can nearly all be captured by the quantum wells, where they can provide for gain.

Ohmic contacts

Next to the structure of the diode, also the metallic contacts that connect the diode are important. In the ideal case, the resistivity of these contacts should be in such a way that they behave as a small ohmic resistor. When this is the case, the contacts are called 'ohmic'. To achieve this, the band diagram of the contact area needs to be engineered in such a way, that no potential barriers need to be crossed by carriers to be injected into the device. For an ideally terminated semiconductor interface, the barrier height can be calculated using the following equations:

$$\Phi_{B,n} = \Psi_M - \chi - \Delta\Phi \tag{A.16}$$

$$\Phi_{B,p} = E_G/q - \Psi_M + \chi - \Delta\Phi \tag{A.17}$$

where Ψ_M is the metal work function, χ is the semiconductor affinity, E_G is the semiconductor band gap and $\Delta\Phi$ is the barrier-lowering potential which is related to the applied electric field. From these equations, one can conclude that in order to achieve a low barrier height, a low metal work function is required for a n-type contact, as it should be as close as possible to the electron affinity. For a p-type contact on the other hand, a higher metal work function is required as next to the affinity, also the band gap of the semiconductor needs to be crossed. This situation is visually shown in figure A.3.

However the requirement for ideally terminated semiconductor surfaces is in a lot of cases not met. In that case, the surface states at the semiconductor interface completely determine the band bending in the semiconductor at the interface with the metal. This phenomenon is also referred to as Fermi level pinning. The consequence is that the

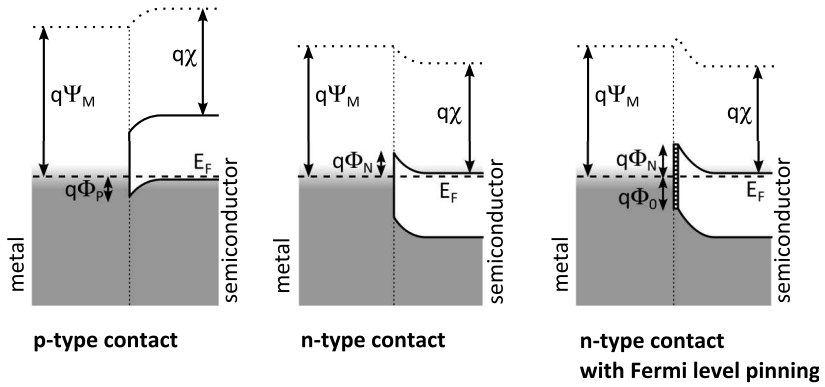


Figure A.3: Band diagram for metal-semiconductor contacts: p-type contact, n-type contact and n-type contact with Fermi level pinning.

metal work function does not influence the barrier height anymore. The equations for the barrier height then become

$$\Phi_{B,n} = E_G/q - \Phi_0 - \Delta\Phi \tag{A.18}$$

$$\Phi_{B,p} = \Phi_0 - \Delta\Phi \tag{A.19}$$

where $q\Phi_0$ is the energy above the valence band edge where the Fermi level is pinned. In this case, the work function of the metal is not important. The lowest resistance can then be obtained by using a small band gap semiconductor with high doping.

A.5.2 Optical pumping

Next to electrical pumping, also optical pumping is possible. The principle, its advantages and disadvantages are thoroughly discussed in section 3.2.3.

A.6 Optical properties of a III-V membrane

A.6.1 Mode profile and optical confinement

By properly etching the III-V membrane in a certain shape, an optical waveguide can be formed. III-V semiconductors are high index materials ($n > 3$), and therefore a high index contrast is achieved with the surrounding cladding, which consists of either SiO_2 ($n = 1.44$), BCB

($n = 1.54$) or air ($n = 1$). Furthermore, because of the similar refractive index with silicon, it is possible to phase match the modes in silicon and III-V waveguides. Therefore, different options to couple light between silicon and III-V are possible. Most commonly, the light is coupled using inverted taper couplers, or by evanescent coupling of a part of the light in an active III-V waveguide to the underlying silicon circuit.

An important parameter of an optical mode is the optical confinement in a certain waveguide region. The optical confinement of a particular mode is a measure for the light-matter interaction for that mode that is obtained in this waveguide region. Usually, the confinement is calculated in the active region of the device, e.g. the MQW region. It can be calculated by taking the ratio between the power flow in the active region to the total power flow in the waveguide.

A.6.2 Quantum well absorption and gain

Band-to-band absorption

When a photon with energy larger than the band gap travels through a semiconductor, it can be absorbed by a valence electron to create a free electron-hole pair. This process is called band-to-band absorption. This process can be direct, which corresponds to a photon interaction, and indirect, which corresponds to a combination of a photon and phonon interaction (see figure A.4). Direct band-to-band absorption is actually the equivalent process of stimulated emission. The fact that band-to-band absorption and stimulated emission are competing processes, is also the reason why population inversion is necessary in a semiconductor to have gain. This will be further discussed in the next section. Direct band-to-band absorption of an optical pump beam is the physical process which is used in this work to generate free carriers in optical pumping.

Optical gain

When a beam of photons is impinging on a piece of direct band gap semiconductor material, there are two competing physical processes which determine whether there will be a net gain or absorption in the material. On one hand, the incoming photons can be absorbed by band-to-band absorption, where a valence electron is moved to the conduction band, while on the other hand the incoming photon could also

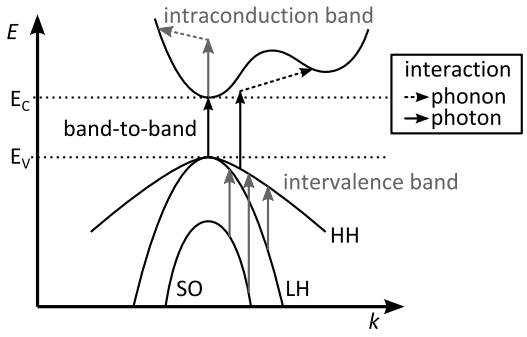


Figure A.4: Simplified band diagram of a direct semiconductor, showing the different absorption mechanisms.

stimulate a free electron in the conduction band to recombine and create an extra photon through stimulated emission. When the semiconductor is not pumped, there will always be a net absorption.

When a semiconductor is pumped hard enough however, the amount of valence electrons near the band edge available for absorption can become smaller than the amount of free electrons in the conduction band of the material. When this is the case, the probability that an incoming photon will be absorbed becomes smaller than the probability that an extra photon will be emitted by stimulated emission. This situation in the carrier distribution is called population inversion, and as a result a net optical gain is achieved in the material. The absorption and gain in a material can be rigorously calculated using Fermi's golden rule:

$$G(E_{ij}) = \frac{q^2 h}{2m_0^2 \epsilon_0 nc} \frac{|M(E_{ij})|}{E_{ij}} \rho_r(E_{ij}) \left(f_c(E_j) - f_v(E_i)\right) \tag{A.20}$$

A.6.3 Parasitic loss mechanisms

Free carrier absorption

Free carrier absorption is a loss mechanism where free carriers in the valence or conduction band are excited by an incoming photon to a higher energy within the valence or conduction band respectively. The different absorption processes are schematically shown in figure A.4. The most important contribution to this loss is due to free carrier absorption in the valence band, as this band actually consists of three sep-

arate energy bands: the heavy hole, light hole and split-off band. Therefore, there are more available energy levels to excite a hole to a higher energy within one of the valence bands. Furthermore a vertical transition within the band diagram is possible, which means that momentum is conserved without the need for interaction with a phonon. This leads to a higher free carrier absorption loss due to holes than due to electrons. The loss scales approximately linearly, and can be expressed as

$$\alpha_{FCA} = k_n N + k_p P \tag{A.21}$$

where k_n is of the order of $1 - 2 \times 10^{-18}\,\mathrm{cm}^2$ and k_p is of the order of $20 - 60 \times 10^{-18}\,\mathrm{cm}^2$ for bulk InP/InGaAsP at $1550\,\mathrm{nm}$[100].

Metal absorption

In contrast with semiconductors, metals do not have a band gap between the valence and conduction band, and therefore there is a strong absorption at any wavelength in the telecom wavelength range. This is the reason why metals should be kept far away from the optical mode in electrically pumped devices.

Scattering

Scattering of light is the result of imperfections in the waveguide. A first type of scattering is volume scattering, which is the result of inhomogeneities in the material through which the light travels. Next to this, light can also scatter because of roughness at a material interface. This can be a quite important loss mechanism when the optical mode is close to the interface, especially in high index contrast waveguides, as the surface scattering scales with both electric field and refractive index difference at the interface [32].

Appendix B

Publications

B.1 International Journals

1. E. Lambert, M. Fiers, S. Nizamov, M. Tassaert, S. G. Johnson, P. Bienstman, W. Bogaerts, Python bindings for the open source electromagnetic simulator MEEP, Computing in Science and Engineering, 13(3), p.53-65 (2011).

2. M. Tassaert, D. Van Thourhout, G. Roelkens, Strategies to increase the modal gain in heterogeneously integrated III-V amplifiers on Silicon-On-Insulator, Optical and Quantum Electronics, p.683-689 (2011).

3. M. Tassaert, G. Roelkens, H.J.S. Dorren, D. Van Thourhout, O. Raz, Bias-free, low power and optically driven membrane InP switch on SOI for remotely configurable photonic packet switches, Optics Express, 19(26), p.B817-B824 (2011).

4. M. Tassaert, S. Keyvaninia, D. Van Thourhout, W.M.J. Green, Y. Vlasov, G. Roelkens, An optically pumped nanophotonic InP/InGaAlAs optical amplifier integrated on a SOI waveguide circuit, Optical and Quantum Electronics (invited), 44(12), p.513-519 (2012).

5. M. Tassaert, H.J.S Dorren, G. Roelkens, O. Raz, Passive InP regenerator integrated on SOI for the support of broadband silicon modulators, Optics Express, 20(10), p.11383-11388 (2012).

6. M. Tassaert, G. Roelkens, H.J.S. Dorren, O. Raz, Optically Reconfigurable 1x4 Remote Node Switch for Access Networks, Journal of Lightwave Technology, (2013).

B.2 International Conference Proceedings

1. M. Tassaert, G. Roelkens, D. Van Thourhout, R. Baets, All-optical low-power switch based on III-V/SOI heterogeneous integration, Photonics in Switching 2009, Italy, p.FrI2-4 (2009).

2. G. Roelkens, O. Raz, W. green, S. Assefa, M. Tassaert, S. Keyvaninia, K.T Vandoorne, D. Van Thourhout, R. Baets, Y. Vlasov, Towards a low-power optically-pumped nanophotonic semiconductor amplifier heterogeneously integrated with SOI waveguides, Group IV Photonics, China, p.WB4 (2010).

3. G. Roelkens, D. Vermeulen, L. Liu, T. Spuesens, R. Kumar, P. Mechet, K. Huybrechts, S. Keyvaninia, S. Stankovic, M. Tassaert, P. De Heyn, K. Komorowska, S. Selvaraja, D. Van Thourhout, G. Morthier, R. Baets, R. Halir, III-V/silicon photonic integrated circuits for FTTH and optical interconnect, IB2COM (invited), Spain, (2010).

4. M. Tassaert, S. Keyvaninia, D. Van Thourhout, G. Roelkens, Strategies to increase the modal gain in electrically pumped quantum dot based amplifiers integrated on a Silicon-On-Insulator waveguide circuit, Photonics Society Benelux meeting, Netherlands, (2010).

5. G. Roelkens, Y. De Koninck, S. Keyvaninia, S. Stankovic, M. Tassaert, M. Lamponi, G. Duan, D. Van Thourhout, R. Baets, Hybrid Silicon Lasers, SPIE Photonics West 2011 (invited), United States, (2011).

6. M. Tassaert, S. Keyvaninia, D. Van Thourhout, W.M.J. Green, Y. Vlasov, G. Roelkens, A nanophotonic InP/InGaAlAs optical amplifier integrated on a Silicon-On-Insulator waveguide circuit , Information Photonics 2011, Canada, (2011).

7. O. Raz, M. Tassaert, G. Roelkens, D. Van Thourhout, H.J.S. Dorren, Bias-Free, Low Power and Optically Driven InP on SOI

Switch for Remotely Configurable Photonic Packet Switches, ECOC 2011, Switzerland, (2011).

8. O. Raz, M. Tassaert, G. Roelkens, H.J.S. Dorren, InP Membrane Based Broadband Regenerator for silicon-based optical interconnect applications, IEEE Optical Interconnects Conference 2012, United States, p.paper MB3 (2012).

9. G. Roelkens, S. Keyvaninia, S. Stankovic, M. Tassaert, N. Hattasan, A. Gassenq, P. De Heyn, Y. De Koninck, P. Mechet, R. Kumar, M. Muneeb, D. Vermeulen, G. Morthier, R. Baets, D. Van Thourhout, III-V-on-silicon membrane photonics for near-infrared and mid-infrared applications, 16th European Conference on Integrated Optics (ECIO 2012) (invited), Spain, (2012).

10. Y. De Koninck, M. Tassaert, G. Roelkens, R. Baets, A. Bazin, F. Raineri, A novel electrically pumped III-V on silicon micro-laser, 16th European Conference on Integrated Optics (ECIO 2012), Spain, p.paper 147 (2012).

11. O. Raz, M. Tassaert, G. Roelkens, H. Dorren, All-optical membrane InP switch on silicon for access applications, Asia Communications and Photonics Conference (ACP2012), China, p.paper AS1B5 (2012).

12. M. Tassaert, G. Roelkens, H.J.S. Dorren, D. Van Thourhout, O. Raz, A Heterogeneously Integrated III-V Membrane on Silicon-On-Insulator for All-Optical Signal Processing Applications, Proceedings of the 2012 Annual Symposium of the IEEE Photonics Society Belenux Chapter, Belgium, p.37-40 (2012).

13. O. Raz, M. Tassaert, G. Roelkens, H. Dorren, Optically reconfigurable 1x4 SOI remote node switch for access networks, OFCNFOEC 2013 (Optical Fiber Communication Conference and Exposition/National Fiber Optic Engineers Conference), United States, p.PDP5C.2 (2013).

14. M. Tassaert, H.J.S. Dorren, G. Roelkens, O. Raz, Broadband Passive InP Membrane Regenerator for Silicon-based Optical Interconnect Applications, OFC/NFOEC, United States, p.OW3D.4 (2013).

15. O. Raz, M. Tassaert, G. Roelkens, H.J.S. Dorren, Membrane InP Saturable Absorbers on Silicon as Building Blocks for Transparent Optical Networks, 15th International Conference on Transparent Optical Networks (ICTON 2013) (invited), Spain, p.paper Tu.D2.3 (2013).

16. G. Morthier, M. Tassaert, P. Mechet, O. Raz, H. Dorren, D. Van Thourhout, G. Roelkens, III-V on silicon components for packet switching, 18th OptoElectronics and Communications Conference held jointly with 2013 International Conference on Photonics in Switching (OECC/PS) (invited), Japan, (2013).

17. G. Roelkens, S. Keyvaninia, Y. De Koninck, P. Mechet, T. Spuesens, M. Tassaert, S. Stankovic, G. Morthier, R. Baets, D. Van Thourhout, G. Duan, J.M. Fedeli, M. Smit, Heterogeneous III-V/silicon photonic integrated circuits, (invited) publication in Asia Communications and Photonics Conference 2013.

18. Y. De Koninck, M. Tassaert, G. Roelkens, R. Baets, Design of an Electrically-Injected Hybrid Silicon Laser with Resonant Mirrors, IEEE Photonics Conference 2013 (IPC), United States, p.paper TuH2.2 (2013).

19. S. Keyvaninia, M. Muneeb, S. Stankovic, S. Ghosh, M. Tassaert, D. Van Thourhout, G. Roelkens, DVS-BCB Adhesive Bonding for Silicon Photonics, Conference on Wafer Bonding for Microsystems 3D- and Wafer Level Integration (WaferBond 13) (invited), Sweden, (2013).

B.3 National Conference Proceedings

1. M. Tassaert, D. Van Thourhout, G. Roelkens, Towards a quantum dot mode-locked laser integrated on the Silicon-On-Insulator platform, Doctoraatssymposium, Belgium, (2010).

List of figures

Bibliography

[1] D. Miller, "Rationale and challenges for optical interconnects to electronic chips," *Proceedings of the IEEE*, vol. 88, pp. 728–749, June 2000.

[2] S. R. Jain, M. N. Sysak, G. Kurczveil, and J. E. Bowers, "Integrated hybrid silicon DFB laser-EAM array using quantum well intermixing," *Optics Express*, vol. 19, pp. 13692–13699, 2011.

[3] Y. Tang, J. D. Peters, and J. E. Bowers, "Over 67 GHz bandwidth hybrid silicon electroabsorption modulator with asymmetric segmented electrode for 1.3 mu m transmission," *Optics Express*, vol. 20, pp. 11529–11535, May 2012.

[4] Z. Sheng, L. Liu, J. Brouckaert, S. He, and D. Van Thourhout, "InGaAs PIN photodetectors integrated on silicon-on-insulator waveguides," *Optics Express*, vol. 18, pp. 1756–1761, January 2010.

[5] R. Raj, F. Raineri, A. Bazin, and P. Monnier, "Subduing or harnessing surface recombination for cw nanolaser or fast optical switch heterogenously integrated on soi," in *Advanced Photonics 2013*, p. IM2A.5, Optical Society of America, 2013.

[6] D. Miller, "Physical reasons for optical interconnection," *International Journal of Optoelectronics*, vol. 11, no. 3, pp. 155–168, 1997.

[7] A. V. Krishnamoorthy, K. W. Goossen, W. Jan, X. Zheng, R. Ho, G. Li, R. Rozier, F. Liu, D. Patil, J. Lexau, H. Schwetman, D. Feng, M. Asghari, T. Pinguet, and J. E. Cunningham, "Progress in Low-Power Switched Optical Interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, pp. 357–376, March-April 2011.

[8] D. A. B. Miller, "Energy consumption in optical modulators for interconnects," *Optics Express*, vol. 20, pp. A293–A308, March 2012.

[9] J. Proesel, B. G. Lee, C. W. Baks, and C. Schow, "35-gb/s vcsel-based optical link using 32-nm soi cmos circuits," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, p. OM2H.2, Optical Society of America, 2013.

[10] W. Bogaerts, Liu Liu, S. Selvaraja, P. Dumon, J. Brouckaert, D. Taillaert, D. Vermeulen, G. Roelkens, D. Van Thourhout, and R. Baets, "Silicon nanophotonic waveguides and their applications," in *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 7134, p. 71341O (13 pp.), 2008.

[11] B. R. Dutt, D. S. Sukhdeo, D. Nam, B. M. Vulovic, Z. Yuan, and K. C. Saraswat, "Roadmap to an Efficient Germanium-on-Silicon Laser: Strain vs. n-Type Doping," *IEEE Photonics Journal*, vol. 4, pp. 2002–2009, October 2012.

[12] H. Kataria, C. Junesand, Z. Wang, W. Metaferia, Y. T. Sun, S. Lourdudoss, G. Patriarche, A. Bazin, F. Raineri, P. Mages, N. Julian, and J. E. Bowers, "Towards a monolithically integrated III-V laser on silicon: optimization of multi-quantum well growth on InP on Si," *Semicondutor Science and Technology*, vol. 28, September 2013.

[13] Z. Wang, B. Tian, M. Paladugu, M. Pantouvaki, N. Le Thomas, C. Merckling, W. Guo, J. Dekoster, J. Van Campenhout, P. Absil, and D. Van Thourhout, "Polytypic InP Nanolaser Monolithically Integrated on (001) Silicon," *Nano Letters*, vol. 13, pp. 5063–5069, November 2013.

[14] C. Dang, J. Lee, K. Roh, H. Kim, S. Ahn, H. Jeon, C. Breen, J. S. Steckel, S. Coe-Sullivan, and A. Nurmikko, "Highly efficient, spatially coherent distributed feedback lasers from dense colloidal quantum dot films," *Applied Physics Letters*, vol. 103, October 2013.

[15] B. De Geyter, K. Komorowska, E. Brainis, P. Emplit, P. Geiregat, A. Hassinen, Z. Hens, and D. Van Thourhout, "From fabrication to mode mapping in silicon nitride microdisks with embedded colloidal quantum dots," *Applied Physics Letters*, vol. 101, October 2012.

[16] E. Augendre, J.-M. Fedeli, D. Bordel, B. Ben Bakir, C. Kopp, L. Grenouillet, J.-M. Hartmann, J. Harduin, P. Philippe, N. Olivier, M. Fournier, M. Zussy, K. Lefebvre, J. Sturm, L. Di Cioccio, L. Fulbert, and L. Clavelier, "Direct bonding for silicon photonics," in *2010 Photonics Global Conference (PGC 2010)*, p. 5, IEEE, 2010.

[17] G. Roelkens, J. Brouckaert, D. Van Thourhout, R. Baets, R. Notzel, and M. Smit, "Adhesive Bonding of InP/InGaAsP Dies to Processed Silicon-on-Insulator Wafers using DVS-bis-Benzocyclobutene," *Journal of Electrochemical Society*, vol. 153, no. 12, pp. G1015–G1019, 2006.

[18] J. Van Campenhout, P. R. A. Binetti, P. R. Romeo, P. Regreny, C. Seassal, X. J. M. Leijtens, T. de Vries, Y. S. Oei, R. P. J. van Veldhoven, R. Notzel, L. Di Cioccio, J.-M. Fedeli, M. K. Smit, D. Van Thourhout, and R. Baets, "Low-Footprint Optical Interconnect on an SOI Chip Through Heterogeneous Integration of InP-Based Microdisk Lasers and Microdetectors," *IEEE Photonics Technology Letters*, vol. 21, pp. 522–524, April 2009.

[19] S. Stankovic, R. Jones, M. N. Sysak, J. M. Heck, G. Roelkens, and D. Van Thourhout, "Hybrid III-V/Si Distributed-Feedback Laser Based on Adhesive Bonding," *IEEE Photonics Technology Letters*, vol. 24, pp. 2155–2158, December 2012.

[20] Y.-h. Kuo, H.-W. Chen, and J. E. Bowers, "High speed hybrid silicon evanescent electroabsorption modulator," *Optics Express*, vol. 16, pp. 9936–9941, June 2008.

[21] Y. Tang, J. Peters, and J. Bowers, "Energy-efficient hybrid silicon electroabsorption modulator for 40-gb/s 1-v uncooled operation," *IEEE Photonics Technology Letters*, vol. 24, no. 19, pp. 1689–1692, 2012.

[22] J. Brouckaert, G. Roelkens, D. Van Thourhout, and R. Baets, "Compact InAlAs-InGaAs Metal-Semiconductor-Metal photodetectors integrated on silicon-on-insulator waveguides," *IEEE Photonics Technology Letters*, vol. 19, pp. 1484–1486, September-October 2007.

[23] E. Pop, "Energy dissipation and transport in nanoscale devices," *Nano Research*, vol. 3, no. 3, pp. 147–169, 2010.

[24] "Cisco visual networking index: Forecast and methodology, 2012-2017," 2013.

[25] Hyundai Park, A. Fang, and J. Bowers, "An electrically pumped hybrid silicon evanescent amplifier," in *OFCNFOEC*, pp. 387–9, IEEE, 2007.

[26] B. Ben Bakir, A. Descos, N. Olivier, D. Bordel, P. Grosse, E. Augendre, L. Fulbert, and J. M. Fedeli, "Electrically driven hybrid si/iii-v fabry-pérot lasers based on adiabatic mode transformers," *Optics Express*, vol. 19, no. 11, pp. 10317–10325, 2011.

[27] PhotonDesign, "Fimmwave, http://www.photond.com/," 2013.

[28] Silvaco, "Atlas, http://www.silvaco.com/products/device simulation/atlas.html," 2013.

[29] E. Ryckeboer, N. Hattasan, G. Morren, D. Delbeke, W. Bogaerts, and R. Baets, "Power budget considerations for in vivo continuous glucose monitoring using absorption spectroscopy," in *Proceedings of the 2010 annual symposium of the IEEE/LEOS Photonics Benelux Chapter*, pp. 253–256, TNO, 2010.

[30] ASML, "i-line lithography, http://www.asml.com/doclib/product andservices/94105.pdf?."

[31] Silvaco, "https://dynamic.silvaco.com/dynamicweb/jsp/downloads/entryaction.do?action=silen-menu&key=23846&format=4," 2013.

[32] C. Ciminelli, F. Dell'Olio, V. M. N. Passaro, and M. N. Armenise, "Fully three-dimensional accurate modeling of scattering loss in optical waveguides," *Optical and Quantum Electronics*, vol. 41, pp. 285–298, March 2009.

[33] C. Ciminelli, V. M. N. Passaro, F. Dell'Olio, and M. N. Armenise, "Three-dimensional modelling of scattering loss in InGaAsP/InP and silica-on-silicon bent waveguides," *Journal of the European Optical Society-Rapid Publications*, vol. 4, 2009.

[34] K. Lee, D. Lim, H. Luan, A. Agarwal, J. Foresi, and L. Kimerling, "Effect of size and roughness on light transmission in a Si/SiO2 waveguide: Experiments and model," *Applied Physics Letters*, vol. 77, pp. 1617–1619, September 2000.

[35] H. Yamada, T. Chu, S. Ishida, and Y. Arakawa, "Si photonic wire waveguide devices," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, pp. 1371–1379, November-December 2006.

[36] J. Bae, W. Zhao, J. Jang, I. Adesida, A. Lepore, M. Kwakernaak, and J. Abeles, "Characterization of sidewall, roughness of InP/InGaAsP etched using inductively coupled plasma for low loss optical waveguide applications," *Journal of Vacuum Science & Technology B*, vol. 21, pp. 2888–2891, November-December 2003.

[37] C. Lee, D. Nie, T. Mei, and M. Chin, "Study and optimization of room temperature inductively coupled plasma etching of InP using Cl-2/CH4/H-2 and CH4/H-2," *Journal of Crystal Growth*, vol. 288, pp. 213–216, February 2006.

[38] R. Yan, S. Corzine, L. Coldren, and I. Suemune, "Corrections to the Expression for Gain in GaAs," *IEEE Journal of Quantum Electronics*, vol. 26, pp. 213–216, February 1990.

[39] S. Stankovic, G. Roelkens, D. Van Thourhout, R. Baets, R. Jones, M. Sysak, and B. Koch, "Hybrid III-V/Silicon laser based on DVS-BCB bonding," in *Proceedings of the 2008 annual symposium of the IEEE/LEOS Photonics Benelux Chapter*, pp. 139–142, 2008.

[40] G. Li, X. Zheng, J. Yao, H. Thacker, I. Shubin, Y. Luo, K. Raj, J. E. Cunningham, and A. V. Krishnamoorthy, "25gb/s 1v-driving cmos ring modulator with integrated thermal tuning," *Optics Express*, vol. 19, pp. 20435–20443, October 2011.

[41] P. Dong, W. Qian, H. Liang, R. Shafiiha, D. Feng, G. Li, J. E. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "Thermally tunable silicon racetrack resonators with ultralow tuning power," *Optics Express*, vol. 18, pp. 20298–20304, September 2010.

[42] T. Baba, S. Akiyama, M. Imai, N. Hirayama, H. Takahashi, Y. Noguchi, T. Horikawa, and T. Usuki, "50-gb/s ring-resonator-based silicon modulator," *Optics Express*, vol. 21, pp. 11869–11876, May 2013.

[43] N.-N. Feng, S. Liao, D. Feng, P. Dong, D. Zheng, H. Liang, R. Shafiiha, G. Li, J. E. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "High speed carrier-depletion modulators with 1.4v-cm vπl integrated on 0.25μm silicon-on-insulator waveguides," *Optics Express*, vol. 18, pp. 7994–7999, April 2010.

[44] N.-N. Feng, D. Feng, S. Liao, X. Wang, P. Dong, H. Liang, C.-C. Kung, W. Qian, J. Fong, R. Shafiiha, Y. Luo, J. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "30ghz ge electro-absorption modulator integrated with 3μm silicon-on-insulator waveguide," *Optics Express*, vol. 19, pp. 7062–7067, April 2011.

[45] P. Chaisakul, D. Marris-Morini, M.-S. Rouifed, G. Isella, D. Chrastina, J. Frigerio, X. Le Roux, S. Edmond, J.-R. Coudevylle, and L. Vivien, "23 GHz Ge/SiGe multiple quantum well electro-absorption modulator," *Optics Express*, vol. 20, pp. 3219–3224, January 2012.

[46] X. Zheng, J. Lexau, Y. Luo, H. Thacker, T. Pinguet, A. Mekis, G. Li, J. Shi, P. Amberg, N. Pinckney, K. Raj, R. Ho, J. E. Cunningham, and A. V. Krishnamoorthy, "Ultra-low-energy all-cmos modulator integrated with driver," *Opt. Express*, vol. 18, no. 3, pp. 3059–3070, 2010.

[47] R. J. Ram, "Semiconductor Optoelectronics. Lecture 13: Quantum Confined Stark Effect." Available online: http://www.rle.mit.edu/sclaser/6.973%20lecture%20notes/Lecture%2013c.pdf.

[48] Chung-En Zah, R. Bhat, B. Pathak, F. Favire, Wei Lin, M. Wang, N. Andreadakis, D. Hwang, M. Koza, Tein-Pei Lee, Zheng Wang, D. Darby, D. Flanders, and J. Heieh, "High-performance uncooled 1.3-mum Al xGa yIn 1-x-yAs/InP strained-layer quantum-well lasers for subscriber loop applications," *IEEE Journal of Quantum Electronics*, vol. 30, pp. 511–23, February 1994.

[49] J. Shimizu, M. Aoki, T. Tsuchiya, M. Shirai, A. Taike, T. Ohtoshi, and S. Tsuji, "Advantages of optical modulators with InGaAlAs/InGaAlAs MQW structure," *Electronics Letters*, vol. 38, pp. 821–822, July 2002.

[50] M. Islam, R. Hillman, D. Miller, D. Chemla, A. Gossard, and J. English, "Electroabsorption in GaAs/AlGaAs coupled quantum-well wave-guides," *Applied Physics Letters*, vol. 50, pp. 1098–1100, April 1987.

[51] S. Keyvaninia, S. Verstuyft, S. Pathak, F. Lelarge, G. H. Duan, D. Bordel, J. M. Fedeli, T. De Vries, B. Smalbrugge, E. J. Geluk,

J. Bolk, M. Smit, G. Roelkens, and D. Van Thourhout, "III-V-on-silicon multi-frequency lasers," *Optics Express*, vol. 21, pp. 13675–13683, June 2013.

[52] S. Beyertt, M. Zorn, T. Kubler, H. Wenzel, M. Weyers, A. Giesen, G. Trankle, and U. Brauch, "Optical in-well pumping of a semiconductor disk laser with high optical efficiency," *IEEE Journal of Quantum Electronics*, vol. 41, pp. 1439–1449, December 2005.

[53] Y. de Koninck, F. Raineri, A. Bazin, R. Raj, G. Roelkens, and R. Baets, "Experimental demonstration of a hybrid III-V-on-silicon microlaser based on resonant grating cavity mirrors," *Optics Letters*, vol. 38, pp. 2496–2498, July 2013.

[54] T. Kimura, N. Tsukiji, J. Yoshida, N. Kimura, T. Aikiyo, T. Ijichi, and Y. Ikegami, "1480-nm Laser Diode Module with 250-mW Output for Optical Amplifiers (FOL 1404QQ Series)," *Furukawa Review*, vol. 19, 2000.

[55] Coldren, L A and Corzine, S W, *Diode Lasers and Photonic Integrated Circuits*. New York: Wiley, 1995.

[56] D. Taillaert, F. Van Laere, M. Ayre, W. Bogaerts, D. Van Thourhout, P. Bienstman, and R. Baets, "Grating couplers for coupling between optical fibers and nanophotonic waveguides," *Japanese Journal of Applied Physics Part 1-Regular Papers Brief Communications*, vol. 45, pp. 6071–6077, 2006.

[57] D. Vermeulen, S. Selvaraja, P. Verheyen, G. Lepage, W. Bogaerts, P. Absil, D. Van Thourhout, and G. Roelkens, "High-efficiency fiber-to-chip grating couplers realized using an advanced CMOS-compatible Silicon-On-Insulator platform," *Optics Express*, vol. 18, pp. 18278–18283, August 2010.

[58] Epixfab, "www.epixfab.eu," 2004.

[59] M. Pu, L. Liu, H. Ou, K. Yvind, and J. M. Hvam, "Ultra-low-loss inverted taper coupler for silicon-on-insulator ridge waveguide," *Optics Communications*, vol. 283, pp. 3678–3682, October 2010.

[60] R. Courant, K. Friedric, and H. Lewy, "On partial Difference Equations of Mathematical Physics," *IBM Journal of Research and Development*, vol. 11, no. 2, pp. 215–&, 1967.

[61] K. Tai, T. Hayes, S. L. McCall, and W. Tsang, "Optical measurement of surface recombination in ingaas quantum well mesa structures," *Applied Physics Letters*, vol. 53, no. 4, pp. 302–303, 1988.

[62] J. Mangeney, H. Choumane, G. Patriarche, G. Leroux, G. Aubin, J. Harmand, J. Oudar, and H. Bernas, "Comparison of light- and heavy-ion-irradiated quantum-wells for use as ultrafast saturable absorbers," *Applied Physics Letters*, vol. 79, pp. 2722–2724, October 2001.

[63] J. C. Delagnes, P. Mounaix, H. Nemec, L. Fekete, F. Kadlec, P. Kuzel, M. Martin, and J. Mangeney, "High photocarrier mobility in ultrafast ion-irradiated In0.53Ga0.47As for terahertz applications," *Journal of Physics D - Applied Physics*, vol. 42, October 2009.

[64] H. Oigawa, J.-F. Fan, Y. Nannichi, H. Sugahara, and M. Oshima, "Universal passivation effect of $(nh_4)_2s_x$ treatment on the surface of iii-v compound semiconductors," *Japanese Journal of Applied Physics*, vol. 30, no. Part 2, No. 3A, pp. L322–L325, 1991.

[65] D. B. Suyatin, C. Thelander, M. T. Bjoerk, I. Maximov, and L. Samuelson, "Sulfur passivation for ohmic contact formation to InAs nanowires," *Nanotechnology*, vol. 18, March 2007.

[66] M. R. Ravi, A. DasGupta, and N. DasGupta, "Effect of sulfur passivation and polyimide capping on ingaas-inp pin photodetectors," *IEEE Transactions on Electron Devices*, vol. 50, no. 2, pp. 532–534, 2003.

[67] D. Sheela and N. DasGupta, "Optimization of surface passivation for InGaAs/InP pin photodetectors using ammonium sulfide," *Semiconductor Science and Technology*, vol. 23, March 2008.

[68] F. Seguineau, D. Rouvillain, H. Choumane, G. Aubin, J. Oudar, P. Brindel, B. Lavigne, and O. Leclerc, "Regeneration capabilities of passive saturable absorber-based optical 2R in 20 Gbit/s RZ DWDM long-haul transmissions," *Electronics Letters*, vol. 39, pp. 857–859, May 2003.

[69] K. Huybrechts, T. Tanemura, K. Takeda, Y. Nakano, R. Baets, and G. Morthier, "All-Optical 2R Regeneration Using the Hysteresis

in a Distributed Feedback Laser Diode," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, pp. 1434–1440, September-October 2010.

[70] P. Mechet, T. Spuesens, N. Olivier, J.-M. Fedeli, P. Regreny, D. Van Thourhout, G. Roelkens, and G. Morthier, "All-optical, low-power 2r regeneration of 10gb/s nrz signals using a iii-v on soi microdisk laser," in *18th Opto-Electronics and Communications Conference held jointly with 2013 International Conference on Photonics in Switching (OECC/PS)*, pp. TuPO–4, 2013.

[71] D. Wolfson, A. Kloch, T. Fjelde, C. Janz, B. Dagens, and M. Renaud, "40-gb/s all-optical wavelength conversion, regeneration, and demultiplexing in an soa-based all-active mach-zehnder interferometer," *Photonics Technology Letters, IEEE*, vol. 12, pp. 332–334, March 2000.

[72] M. Rochette, L. Fu, V. Ta'eed, D. J. Moss, and B. J. Eggleton, "2R optical regeneration: An all-optical solution for BER improvement," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 4, pp. 736–744, 2006.

[73] A. Kuramoto and S. Yamashita, "All-optical regeneration using a side-mode injection-locked semiconductor laser," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, no. 5, pp. 1283–1287, 2003.

[74] H. Simos, A. Bogris, and D. Syvridis, "Investigation of a 2R all-optical regenerator based on four-wave mixing in a semiconductor optical amplifier," *Journal of Lightwave Technology*, vol. 22, no. 2, pp. 595–604, 2004.

[75] T. Otani, T. Miyazaki, and S. Yamamoto, "40-gb/s optical 3r regenerator using electroabsorption modulators for optical networks," *Journal of Lightwave Technology*, vol. 20, no. 2, pp. 195–200, 2002.

[76] G. P. Agrawal, *Fiber-optic communication systems*, p. 169. Wiley-interscience, 2002.

[77] Huawei, "Next-generation pon evolution," 2010.

[78] C. V. Praet, H. Chow, D. Suvakovic, D. V. Veen, A. Dupas, R. Boislaigue, R. Farah, M. F. Lau, J. Galaro, G. Qua, N. P. Anthapadmanabhan, G. Torfs, X. Yin, and P. Vetter, "Demonstration of low-power bit-interleaving tdm pon," *Optics Express*, vol. 20, pp. B7–B14, December 2012.

[79] J. Theodoras and K. Grobe, "100gbe - the future of ethernet," 2010.

[80] P. Healey, P. Townsend, C. Ford, L. Johnston, P. Townley, I. Lealman, L. Rivers, S. Perrin, and R. Moore, "Spectral slicing WDM-PON using wavelength-seeded reflective SOAs," *Electronics Letters*, vol. 37, pp. 1181–1182, September 2001.

[81] H. D. Kim, S.-G. Kang, and C.-H. Le, "A low-cost wdm source with an ase injected fabry-perot semiconductor laser," *IEEE Photonics Technology Letters*, vol. 12, no. 8, pp. 1067–1069, 2000.

[82] Z. Al-Qazwini, M. Thollabandi, and H. Kim, "Colorless optical transmitter for upstream wdm pon based on wavelength conversion," *Journal of Lightwave Technology*, vol. 31, no. 6, pp. 896–902, 2013.

[83] R. S. Tucker, "Scalability and Energy Consumption of Optical and Electronic Packet Switching," *Journal of Lightwave Technology*, vol. 29, pp. 2410–2421, August 2011.

[84] J. H. Lee, K.-M. Choi, and C.-H. Lee, "A remotely reconfigurable remote node for next-generation access networks," *IEEE Photonics Technology Letters*, vol. 20, pp. 915–917, May-June 2008.

[85] K. Tokuhashi, K. Ashizawa, D. Ishii, Y. Arakawa, N. Yamanaka, and K. Wakayama, "Secure and scalable optical access network using plzt high-speed optical switches," in *High Performance Switching and Routing, 2009. HPSR 2009. International Conference on*, pp. 1–6, 2009.

[86] F. Gomez-Agis, H. Hu, J. Luo, H. C. H. Mulvad, M. Galili, N. Calabretta, L. K. Oxenløwe, H. J. S. Dorren, and P. Jeppesen, "Optical switching and detection of 640gbits/s optical time-division multiplexed data packets transmitted over 50km of fiber," *Optics Letters*, vol. 36, pp. 3473–3475, September 2011.

[87] K. Ashizawa, K. Tokuhashi, D. Ishii, S. Okamoto, N. Yamanaka, and E. Oki, "Efficient singlecast / multicast method for active optical access network using plzt high-speed optical switches," in *High Performance Switching and Routing (HPSR), 2010 International Conference on*, pp. 14–19, 2010.

[88] K. Nashimoto, "Nano-Speed PLZT Optical Switches," tech. rep., Nozomi Photonics Co., Ltd., September 2006. ONT3 Workshop 2006, Japan.

[89] B. Lee, A. Biberman, J. Chan, and K. Bergman, "High-performance modulators and switches for silicon photonic networks-on-chip," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 1, pp. 6–22, 2010.

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

[91] J. Feng, R. Akimoto, S. ichiro Gozu, T. Mozume, T. Hasama, and H. Ishikawa, "Ultrafast all-optical switch with cross-phase modulation by area-selective ion implantation in ingaas/alassb coupled double quantum wells," *Optics Express*, vol. 20, pp. B279–B287, December 2012.

[92] C. Koos, P. Vorreau, T. Vallaitis, P. Dumon, W. Bogaerts, R. Baets, B. Esembeson, I. Biaggio, T. Michinobu, F. Diederich, W. Freude, and J. Leuthold, "All-optical high-speed signal processing with silicon-organic hybrid slot waveguides," *Nature Photonics*, vol. 3, pp. 216–219, April 2009.

[93] Ipkiss, "www.ipkiss.org," 2013.

[94] W. Bogaerts, P. Dumon, E. Lambert, M. Fiers, S. Pathak, and A. Ribeiro, "IPKISS: A parametric design and simulation framework for silicon photonics," in *2012 IEEE 9th International Conference on Group IV Photonics (GFP)*, (Piscataway, NJ, USA), pp. 30–2, IEEE Photonics Soc., IEEE, 2012.

[95] S. Pathak, D. Van Thourhout, and W. Bogaerts, "Design trade-offs for silicon-on-insulator-based AWGs for (de)multiplexer applications," *Optics Letters*, vol. 38, pp. 2961–2964, August 2013.

[96] W. Bogaerts, P. Dumon, D. Van Thourhout, and R. Baets, "Low-loss, low-cross-talk crossings for silicon-on-insulator nanophotonic waveguides," *Optics Letters*, vol. 32, pp. 2801–2803, October 2007.

[97] Z. Sheng, Z. Wang, C. Qiu, L. Li, A. Pang, A. Wu, X. Wang, S. Zou, and F. Gan, "A Compact and Low-Loss MMI Coupler Fabricated With CMOS Technology," *IEEE Photonics Journal*, vol. 4, pp. 2272–2277, December 2012.

[98] D. Dai, Z. Wang, J. F. Bauters, M. C. Tien, M. J. R. Heck, D. J. Blumenthal, and J. E. Bowers, "Low-loss Si3N4 arrayed-waveguide grating (de)multiplexer using nano-core optical waveguides," *Optics Express*, vol. 19, pp. 14130–14136, July 2011.

[99] J. F. Young and H. M. van Driel, "Ambipolar diffusion of high-density electrons and holes in ge, si, and gaas: Many-body effects," *Physics Review B*, vol. 26, pp. 2147–2158, August 1982.

[100] J. Piprek, *Semiconductor Optoelectronic Devices. Introduction to Physics and Simulation.* San Diego: Elsevier, 2003.