Heterogene integratie van
III-V-halfgeleiderlichtbronnen van meerdere actieve media op siliciumfotonica

Heterogeneous Integration of
Multibandgap III-V Semiconductor Light Sources on Silicon Photonics

Andreas De Groote

Promotoren: prof. dr. ir. R. Baets, prof. dr. ir. G. Roelkens
Proefschrift ingediend tot het behalen van de graad van
Doctor in de ingenieurswetenschappen: fotonica

UNIVERSITEIT
GENT

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. B. Dhoedt
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2017 - 2018

ISBN 978-94-6355-044-4
NUR 959
Wettelijk depot: D/2017/10.500/79

UNIVERSITEIT
GENT

Promotoren:
Prof. Dr. Ir. Günther Roelkens
Prof. Dr. Ir. Roel Baets

Examencommissie:

Prof. Dr. Ir. H. Van Landeghem (voorzitter)	Universiteit Gent
Prof. Dr. Ir. G. Roelkens (promotor)	Universiteit Gent
Prof. Dr. Ir. R. Baets (promotor)	Universiteit Gent
Prof. Dr. Ir. J. Bowers	University of California, Santa Barbara
Prof. Dr. Ir. B. Corbett	Tyndall National Institute, University College Cork
Prof. Dr. Ir. Z. Hens	Universiteit Gent
Prof. Dr. Ir. D. Van Thourhout	Universiteit Gent

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur

Vakgroep Informatietechnologie
Photonic Research Group
Technologiepark-Zwijnaarde 15 iGent, 9052 Gent, België

Tel.: +32-9-264 3316
Fax.: +32-9-264 3593

Dit werk kwam tot stand in het kader van een specialisatiebeurs van het FWO-Vlaanderen (Fonds Wetenschappelijk Onderzoek).

Dankwoord

4 jaar (en een beetje) doctoreren. Volgens sommigen is het vier jaar de echte wereld uitstellen. Nog een beetje spelen, voor je moet gaan werken. Maar ja, zeggen ze dan, wat is vier jaar op een mensenleven? De ingenieur in mij antwoordt dan dat dat ongeveer 15% van mijn leven is, wat toch niet te verwaarlozen is. En daarbij, ik deed onderzoek in de echte wereld, om de echte wereld te verbeteren.

Het meest gebruikte woord in dit doctoraat is wellicht "We". Onderzoeken doe je namelijk niet alleen. De Photonics Research Group is een dynamische groep waarin getalenteerde mensen van overal samenkomen. Alleen al hiervoor verdient Roel Baets een welgemeende dankjewel. Naast een goede people-manager is Roel Baets ook een fantastische wetenschapper. Als doctoraatsstudent kan het soms pijnlijk zijn wanneer je in een vergadering een welgemikte vraag krijgt waardoor je begint te twijfelen aan de fundamenten van je onderzoek. Dit zorgt er onvermijdelijk voor dat de vergaderingen uitlopen en je soms met meer vragen vertrekt dan je gekomen was, maar het komt de kwaliteit van je onderzoek zeker ten goede. Naast rake vragen, komt Roel Baets ook met heldere antwoorden op mijn vragen. Het is geen twijfel dat het Roel was die me in 2010 overtuigde om een master in fotonica te volgen.

Gunther Roelkens is dan weer de promotor die oplossingen bedenkt voor je wist dat je vragen had (ik vermoed dat hij en Roel Baets daardoor zo'n sterk duo vormen). Gunther Roelkens is altijd bereikbaar en is bij wijze van spreken een wandelende encyclopedie. Dit wil helemaal niet zeggen dat het hem aan verbeelding ontbreekt. Gunther Roelkens, sorry voor al je gekke ideeën die ik niet in de praktijk bracht. Rest van de wereld, graag gedaan dat ik enkele van Gunther Roelkens' gekke ideeën in de praktijk kon brengen.

Naast mijn twee promotoren wil ik ook mijn stiefpromotor bedanken. John Bowers, I had the opportunity to work with you at University of California Santa Barbara for over a year. You are an enthusiastic and energetic motivator, pushing me and your entire *Optoelectronics Research Group* to higher throughput. When you see an opportunity, you go for it 100%. When Roel Baets, Gunther Roelkens and I suggested to collaborate, you did not only reply my email, you flew over to discuss it immediately!

Enkele promotoren en een doctoraatstudent. Dit is wat je vindt op het voorblad

van een doctoraatsthesis. Eigenlijk zouden we die boeken moeten printen op A0 papierformaat, want er zijn er zoveel meer die een prominente plaats verdienen.

Paolo Cardile, Ananth Subramanian en Danaë Delbeke, of Team Glucare, jullie hebben mee tranfer printing en de optische gepompte membranen ontwikkeld. Dit wil ook zeggen dat wij samen vaak met de handen in het haar zaten. Jullie optimisme was daarin zeker niet onbelangrijk. De meeste blokkeringen vergeet je gewoon achteraf, maar we hebben toch enkele memorabele stilstand-momenten meegemaakt. Dat die dan achteraf goedgemaakt worden door een grote sprong voorwaarts maakt natuurlijk veel goed. Wanneer Team Glucare vervelde naar het bedrijf Indigo, kwam continue glucosemonitoring weer dichterbij voor diabetici.

The transfer printing story was co-developed with X-Celeprint. They already had shown the technology for other platforms and were eager to develop it for integrated photonic circuit applications. Chris Bower was always free for input and his direct and clear correspondence made collaborating easy. Alin Fecioru was the cleanroom guy at the X-Celeprint side, a position he fills in expertly. Alin Fecioru is always in a good mood, a quality much needed when doing a multi-micron undercut with only nanometers of vertical tolerance.

Zoals je wel zal beseffen wanneer je dit boek uitgelezen hebt, heb ik mijn tijd hoofdzakelijk in de cleanroom doorgebracht. Het kloppend hart van onze UGent cleanroom kreeg een naam: Steven Verstuyft. Even schetsen hoe belangrijk Steven Verstuyft's aanwezigheid was voor mij. In het kantoor van de cleanroom hangt een grote kalender op, waar de Steven Verstuyft zijn zomervakanties en dergelijke aanduidt. Na de eerste zomer, waarin Gent geteisterd werd door falende RIE's en labiele mask aligners, besloot ik om mijn vakantie gewoon op hetzelfde moment te nemen. Wat mij betreft is Steven Verstuyft een magiër die zijn kalmte kan afstralen op de machines. Ik vermoed wel dat deze magiër van geluk kan spreken dat niet iedereen hem even veel vragen stelt als ondergetekende. Als je Steven Verstuyft een tovenaar noemt, veronderstel ik Liesbet Van Landschoot de goede heks is. Immer vrolijk verspreidt zij positieve vibes doorheen alle kille aparatuur. Maar Liesbet Van Landschoot is ook onze FIB/SEM-specialist. Zo zat zij er bij toen ik merkte dat de eerste transfer printing niet het verhoopte succes gaf. Ze was er echter ook toen diezelfde transfer printing wel succesvol was en Paolo Cardile en ik door het dak gingen. Liesbet Van Landschoot glunderde mee, maar bleef ook kalm om het beeld te vereeuwigen.

Mijn eerste dagen in de cleanroom moest ik echter doormaken zonder mijn twee tovenaars, aangezien zij weigerden me te volgen naar UCSB. Jon Peters was there however, to show me around. All the basic and more advanced skills (how do I hold my tweezers and how does that hold my sample?) I have learned from him. The fact that Jon Peters starts at 1 a.m. was unfortunate, but a small burden to bear. Mike - post-bond god - Davenport and Sid - lord of the intermixing - Jain also helped me get started. I will never forget the first and foremost of Sid Jain's rules of cleanroom processing: Always wear gloves, because otherwise you can forget about children.

Graag bedank ik ook mijn III-V collega's, zowel die van Gent als van over de grote plas. Alex Spott and Eric - do not forget to add the J when you google

my name - Stanton, I saw you suffer through the hard classes of the first year of a US PhD and you never denied me the pleasure of making fun of this. Sid Jain and his psychological twin Geza Kurczveil made clear that cleanroom processing is satisfying in an awful way, while Martijn Heck was the unexpected Dutchman in UCSB. Even I have to admit that I was glad to have had this Dutchman at my side to point out my interpretation errors. Yannick De Koninck nam me reeds als thesisstudent onder zijn vleugels. Yannick De Koninck is niet alleen een begenadigd processer, hij weet als geen ander hoe een FDTD simulatie op te zetten. Martijn Tassaert is dan weer een van de meest georganiseerde mensen die ik ken. Hij zette me op weg met high contrast membranes, waaruit uiteindelijk het idee van de optisch gepompte LED volgde. Amin Abbasi en Leila Abdollahi represent the country of Iran in our group. They always share local specialties and if you see them, you'd swear a big smile is one of them. Jing Zhang has tried processes in our cleanroom nobody ever has. I will always remember his solution consisting of three bondings in order be able to do ... bonding. Sören Dhoore is een rivaal, aangezien hij een fervent Club Bruggesupporter is. Gelukkig won deze voetbalclub slechts eenmaal de competitie in mijn doctoraat. Sören Dhoore heeft soms plagerijen moeten ondergaan. Kasper Van Gasse and Zhechao Wang were the mode-locked duo. Together they not only demonstrated world-class devices, but I mainly used them to bounce off processing ideas. Their input in this was greatly appreciated.

Het professionele leven is meer dan III-V echter (toegegeven, een minder bekende quote). Drie kantoren werden onveilig gemaakt in de tijdspanne beschreven in dit werk. The UCSB guys were already mentioned above as Alex Spott and Eric J Stanton were III-V guys as well as office mates. Ons kantoor in het legendarische technicum werd dan weer bestierd door ondergetekende, III-V guy Amin Abassi, Eva - hoe kunnen jullie nu niet door hebben dat ik zwanger ben als ik de ene zak nootjes na de andere eet - Ryckeboer, Abdul Rahim (or was it Rahim Abdul?), the silent Pijush Kumar Pal, advocaat Raphael Van Laer, Thijs Spuessens (de tweede Nederlander die ik genoodzaakt was te tolereren en zelfs te appreciëren), the brittle Haolan Zhao, Thomas Van Vaerenbergh die zelfs tegen M.K. niet nee durfde zeggen, dubbele meter Jérémy Wenger en jonkheer Artur Hermans. U merkt dat dit een groot kantoor was, met allerlei smaken aan mensen. Na de verhuis naar iGent werden we noodgedwongen gesplitst. Voortaan kon ik enkel the brittle and due to his wife now also stylish Haolan Zhao, jonkheer maandagmens Artur Hermans en Nan Ye (sorry to bother you) tot mijn inner office-circle rekenen. Het werd een kalmer maar niet minder aangenaam boeltje. Jammer van die subtiele hint om onze muurdecoratie weg te halen. Daarnaast wil ik ook nog even de lift uit het technicum vermelden. Er was veel gevloek toen ik met de trap van -4 naar +1 moest stappen, maar iGent leerde me hoe moeilijk het is voor een lift om degelijk te werken.

Na mijn III-V collega's en mijn kantoormaten wil ik ook nog enkele andere werkmensen bedanken. Daan Martens heeft mijn gezelschap reeds lang moeten verduren. Na twee jeugdige jaren werden we gescheiden voor vier jaar (we'll always have Paris). Ingenieur worden bracht ons weer samen en ingenieur zijn heeft

dat voorlopig nog niet veranderd. Geen theorie was te ver gezocht om te bespreken en Daan Martens leerde me dat je letterlijk de muren kan oplopen van de stress. Suzanne Bisschop is dan weer rustig en down to earth. Suzanne Bisschop is een fantastisch persoon, maar er is nog werk aan. Suzanne Bisschop heeft Pulp Fiction nog niet gezien, een euvel waar Daan Martens en ik een mouw moeten aan passen. Antonio Ribeiro werd stiekem collega. Niemand weet exact wanneer hij de overstap maakte van werknemer naar doctoraatsstudent (Wim Bogaerts wellicht nog het minst van al), maar hij doet lost computer- en serverproblemen op met verve. Michael Vanslembrouck is dan weer de duivel-doe-al van de groep. Van stukadoor tot automatic meausurement setups, Michael Vanslembrouck kan het allemaal. Ik ben ook zijn suggestie van Edelmire niet vergeten. Met Pieter Wuytens moest ik dan weer wedijveren voor kleurrijke figuur, een pleit dat ik toch in mijn voordeel beslechtte. Pieter Wuytens heeft wonderen verricht op een platform dat er eigenlijk nog geen was en zette zich tegelijkertijd in voor the greater good of the group. Ashwyn Srinivasan immortalized our meeting with a black bear by staying calm, take his camera, turn the camera on with the typical noise, turn off the flash and shoot the photo. The fact that before he insisted on walking in the middle as a precaution only makes it more impressive/incomprehensible.

Het beste werktuig om iemand nederigheid bij te brengen is administratie. Wie me kent weet dat dit bij mij een zeer efficiënte methode is. Ter illustratie: Een week voor de deadline merkte ik dat ik al twee jaar lang officieel geen doctoraatsstudent meer was. Aangezien het moeilijk is om af te studeren als je niet studeert begon ik als zot rond te bellen om dit nog snel in orde te maken. Dit is slechts een van vele anekdotes. Gelukkig waren er Ilse Van Royen, Kristien De Meulder, Ceanna Bowman en vele anderen die me hier door de wirwar aan formulieren hielpen.

Mens sana in corpore sano. Je kan enkel goed werk leveren als je ook buiten het werk goed in je vel zit. Familie is belangrijk voor mij en daarom houd ik er ook aan om hen even te vermelden. Eerst en vooral Dirk De Groote en Ann Van Buggenhout, papa en mama. Zij kneedden me tot wat ik nu ben, zij hebben me mijn capaciteiten laten ontplooien. Dat dit soms leidde tot een uitleg van een halfuur om 6u 's ochtends over waarom er zich een meteorengordel tussen Mars en Jupiter bevindt, namen ze er (graag) bij. Toen ik in het begin van mijn middelbare school buizen op godsdienst kreeg, kregen mama en papa het benauwd, maar ik bleef rustig. Toen daar Frans bij kwam, sloeg er lichte paniek toe bij mama en papa, maar ik bleef rustig. Toen er opeens een rood cijfer stond bij wiskunde, ging er een groot alarm af bij mama en papa, en toen werd ik toch ook wat onrustig. Een actieplan werd opgesteld, waarin voetbal op de eerste plaats stond, en ik besloot dan toch maar af en toe te studeren. Alles werd weer normaal en Dirk De Groote en Ann Van Buggenhout konden zich zorgen maken over mijn broers en zus. Of over de hond, Waldo, wiens niet aflatend enthousiasme ik ten zeerste bewonder.

Sebastiaan De Groote is een grote broer zoals een grote broer meestal is. Hij voert graag het hoogste woord en is verrast wanneer kleine broer Andreas De Groote dan gelijk blijkt te krijgen. Samen met Jana Baetens zorgde hij voor mijn eerste neefjes: Urbain De Groote en Ernest De Groote. De ene is een klein zeer

en de andere een boerke. Zie ik daar opnieuw het grote broer - kleine broer concept? Beide zijn ze echter heel lief en ik denk wel dat graafwerken De Groote een lang leven beschoren is. Ruben De Groote is een perfectionist, een eigenschap die de patiënten van deze uroloog zeker wel ten goede zal komen. Zijn natuur maakt het dubbel pijnlijk wanneer zijn jongere broer hem verslaat met armworstelen, zelfs wanneer die met zijn zwakkere rechterhand drukt. Die dag leerde Ruben De Groote een levensles: Onthoud altijd dat er zich soms nog een broer in het buitenland bevindt. De uroloog breidde hij het geslacht De Groote verder uit (pun intended). Samen met Sarah Poisson kreeg hij een zoon Jacob De Groote. Jacob De Groote zoekt naar nieuwe toonaarden wanneer hij papa wil zeggen. Hij laat geen gelegenheid liggen om op ontdekking te gaan, zelfs als dat wil zeggen dat hij samen met de neefjes de Ark van Noah ten gronde richt in de kerk. Hadewich De Groote is mijn jongere zusje. Jammer genoeg zijn er nu eenmaal minder foto's van het vierde kind, maar ik kan je verzekeren dat onze jeugd gelukkig was. Hadewich De Groote amuseerde zich toen ze mee ging kijken naar "de voetbal van de jongens", of toen ze gedwongen werd om tegen ondergetekende te schaken. Mede hierdoor groeide Hadewich De Groote op tot een veelzijdig en extreem weerbaar persoon. Toen ze de absurd hoge verwachtingen van onbekommerde universitaire studies niet kon waarmaken, zette ze toch door. Zelfs toen diezelfde faculteit haar thesis op zijn zachtst gezegd bemoeilijkte, gaf Hadewich De Groote niet op en schreef met goede moed naarstig voort. Ondertussen bewees ze aan de hele wereld (samen met Jana Baetens) haar creatieve kwaliteiten toen ze stylist-on-budget was in de kortfilm "Twist".

Op 26 november 2014 barste de weerbaarheid van Hadewich De Groote bijna. U herinnert het zich misschien nog: "Student valt door dak wintercircus Gent". De student in kwestie was Constant Van Grembergen, wederhelft van mijn zus. Constant Van Grembergen was niet dronken, deed geen helse toeren. Hij liep door een open deur, op een solide trap, om naar de zonsopgang te kijken. Constant Van Grembergen was goed en enthousiast over de kleine dingen in het leven. Hij was een creatieveling, een werker, ging er wel komen... indien hij de zonsopgang had gezien. Constant Van Grembergen stierf in nacht van 25 op 26 november en met hem stierf een beetje van mezelf en groot stuk van mijn zus. In een klap raakten we alle vertrouwen in rechtvaardigheid kwijt. "Als je goed bent en hard werkt, komt alles wel goed", bleek niet te kloppen. Constant Van Grembergen deed dit allemaal en nog veel meer maar het kwam niet goed. 2.5 jaar later is de schok voorbij, maar de leegte blijft. Ik kan alleen maar mijn grootste bewondering uitdrukken voor Hadewich De Groote, dat ze de herinnering aan Constant Van Grembergen eer aandoet. Ze werd geen lege huls van verdriet, maar ging door met leven met veel oog voor schoonheid en vreugde.

Valère Van Buggenhout heeft vele kilometers met me afgelegd. Hij pikte me elke middag op van de lagere school voor het eten. Na mijn gezonde boerenkoolsteetlust weer aangescherpt te hebben bracht hij me terug, om me enkele uren later terug op te pikken. Zoals u merkte, een heel ritueel, dat hij steeds stipt uitvoerde (als u de legendarische stiptheid van de familie Van Buggenhout zou kennen, zou

u de betekenis hiervan kunnen inschatten). Hij leerde me op negenjarige leeftijd hoe ik met een grote tractor en een nog grotere kar vol stro rond en niet op de botten moest rijden. Het gebrek aan servo en de diepe sleuven in het veld hielpen hier zeker niet. Lutgarde - Metie - Van Buggenhout zorgde dan weer voor een dak boven mijn hoofd toen ik het ouderlijk nest verliet. Deze flamboyante kapster is vurig als de Etna, maar tegelijkertijd zo zacht als een teddybeer. Ze houdt even veel van ons als van haar (huis)dieren en wanneer ze je zwangere vriendin met een boerenpaard vergelijkt, bedoelt ze dat zeer positief. Ze leerde me ook alles over schepjes tijdens het eten. Marc Piron heeft geen problemen om alle vertrouwen in een goede afloop te behouden, zelfs al ging mijn doctoraat wat overtijd. Annemarie De Groote en Daniël Goossens volgden me vanop een veilige afstand, je weet natuurlijk nooit welke schade lasers kunnen aanrichten. Tania Van Acker en Geert Uvin zijn familieleden die ik gratis bij mijn vriendin bij kreeg. Zij trachtten te begrijpen waar die nieuwe zonderling toch mee bezig was, een bewonderenswaardige onderneming waar zij soms nog in slaagden ook. Zij leerden me ook dat de afstand tussen scheikundigen en ingenieurs groter is dan je zou denken. Of dit aan de scheikundigen dan wel aan de ingenieurs ligt, is nog onderwerp van debat.

Sarah Uvin zal teleurgesteld zijn met plaats twee, maar ze zal het wel begrijpen aangezien haar dochter Irma De Groote op plaats één komt. Sarah Uvin is er steeds: in III-V meeting, aan de ontbijttafel, in de cleanroom, in de moestuin,... ”Is dit niet wat veel?”, vragen mensen dan. Ja, dit is veel, maar Sarah Uvin is dan ook een extreem veelzijdig persoon. Op het werk is Sarah Uvin een klankbord waartegen ik idiote ideeën kan aftoetsen. Ze stelt vlijmscherpe vragen, waarna ze eindigt met een bemoedigende zin. Slaan en zalven heet dat dan. Thuis is Sarah Uvin de persoon die met iedereen meevoelt. Mocht ik triestig zijn omwille van een tak die van de boom breekt, is ze triestig met mij. Sarah Uvin is een zorgdragend persoon: van haar boeken, van mij, van Irma. Soms vraag ik me hier de prioriteitsvolgorde af (vooral dan wie/wat op twee en drie staat). Sarah Uvin denkt en voelt mee met anderen, zelfs al ligt die leefwereld ver van de hare. Sarah Uvin kan bazig zijn, iets waar ik het nut natuurlijk niet kan/mag op papier zetten. Sarah Uvin kan veel, en nog meer. Omwille van al die Sarah Uvin’s is ze mijn vriendin en blijft ze dat.

In Irma De Groote zie je veel van Sarah Uvin terug. Mijn dochter kan gefocust zijn op een probleem en uiteindelijk zal ze het vakkundig oplossen. En baby’s hebben grote problemen, zoals ”hoe krijg ik deze wasknijper stuk?”. Irma De Groote is vrolijk, zeker wanneer ze buiten kan of opeens een paard ziet, waarna ze steevast Paaa zegt. Irma De Groote erfde haar interesse voor boeken. Het is dan ook mijn goede hoop en vrees dat ze dit boek zal verslinden. Ik wil Irma De Groote ook expliciet bedanken voor het aanbieden van een stukje komkommer tijdens de laatste loodjes van mijn doctoraat, hoewel ik liever de rijstwafel in haar andere hand had gehad.

4 jaar (en een beetje) doctoreren. Onderzoek dat bijna van vlees en bloed lijkt. Onderzoek om de wereld te verbeteren. Weet u nog wat u de voorbije vier jaar

gedaan heeft? Ik heb het geluk om dit neergeschreven te hebben. Zet u rustig neer in een comfortabele zetel en neem een glas wijn of water, want het gaat beginnen. Het boek waarop u allemaal gewacht heeft! (De film komt later)

Gent, Juni 2017
Andreas De Groote

Table of Contents

List of acronyms

B

BCB	Benzocyclobutene
BHF	Buffered hydrofluoride
BOX	Buried oxide

C

CCD	charge-coupled device
CGM	Continuous Glucose Monitoring
CMP	Chemical-Mechanical polishing

D

DFB	Distributed feedback
DRIE	Deep reactive ion etcher
DUT	Device under test
DVS-BCB	Divinyl Siloxane bis-Benzocyclobutene

E

EAM	Electro-absorption modulator

F

FCA	Free-carrier absorption
FDTD	Finite difference time domain
FIB	focused ion beam

I

IC	Integrate circuit
ICP	Inductively coupled plasma
IED	implantation enhanced disordering
IFVD	impurity-free vacancy disordering
IID	impurity induced disordering

L

LD	Laser diode
LED	Light emitting diode

M

MQW	Multi-quantum wells

N

NIR	Near infrared

O

OSA	Optical spectrum analyzer

P

PAID	photo-absorption induced disordering
PDMS	Polydimethylsiloxane
PECVD	plasma enhanced chemical vapour deposition
PIC	Photonic integrated circuit
PL	Photoluminescence

Q

QD Quantum dot
QW Quantum well
QWI Quantum well intermixing

R

RIE reactive ion etcher
RPM Rotations per minute

S

SAG Selective area growth
SCH Separate confinement heterostructure
SEM Scanning electron microscope
SLED Superluminescent light emitting diode
SMF Single mode fiber
SOI Silicon on insulator
SRH Shockley Read Hall
SSPD Superconducting single photon detector

T

TE Transverse electric
TLM Transmission line measurement
TM Transverse magnetic
TP Transfer printing
TRPL Time-resolved photoluminescence

W

WHO World Health Organization

Nederlandstalige samenvatting

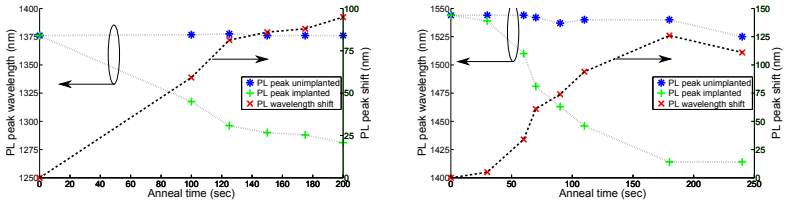
1 Breedbandige lichtbronnen op silicium

415 miljoen in 2016 en een toename tot 642 miljoen in 2040. Dat is hoeveel mensen op aarde zullen lijden aan diabetes. Of anders gezegd, 1 op de 10 mensen valt ten prooi aan deze ziekte. Onschuldig is diabetes allerminst: een te hoge suikerspiegel wordt beschouwd als de derde hoogste risicofactor voor een vroegtijdig overlijden. [1] Hoewel diabetes niet kan worden genezen, kan de hoeveelheid glucose in het bloed wel onder controle gehouden worden door het toedienen van synthetische insuline. Aangezien de suikerspiegel van nature sterk fluctueert, is het van groot belang om het glucosegehalte in het bloed van diabetespatiënten continu te meten.

Silicium-fotonica is een beloftevol technologiedomein dat de voorbije 10 jaar een snelle groei heeft gekend. Door handig gebuik te maken van de reusachtige investeringen reeds gedaan in de micro-electronicaindustrie, belooft siliciumfotonica chips die in grote hoeveelheden en goedkoop geproduceerd kunnen worden. Initieel werd vooral gekeken naar toepassingen voor de telecom- en datacomindustrie, maar later kwam de sensorwereld ook in het vizier. Het continu opmeten van bloeglucosewaarden is een perfect voorbeeld van een toepassing waar silicium-fotonica het verschil kan maken. Elke diabetespatiënt heeft zijn eigen glucosesensor nodig: deze moet dus goedkoop zijn en de potentiële afzetmarkt is gigantisch.

Silicium is een uitstekend materiaal voor het realizeren van passieve optische functies. Verder is het ook mogelijk om hoge kwaliteit optische detectoren te maken op dit platform. Helaas vormen lichtbronnen nog steeds een probleem. Aangezien silicium een indirecte bandkloof heeft, zijn andere materialen nodig om optische winst te genereren. Daarbij mogen we niet uit het oog verliezen dat de integratie van deze nieuwe materialen op het silicium platform schaalbaar moet zijn, willen we de massaproductie van silicium-fotonica chips mogelijk houden.

Heterogene integratie van III-V materialen op silicium past perfect in dit plaatje, wat dan ook het uitgangspunt van dit onderzoek was. In dit proefschrift gingen we op zoek naar breedbandige lichtbronnen geïntegreerd op silicium. Drie methodes voor het realizeren van verschillende actieve bandkloven op silicium door middel van heterogene integratie werden bestudeerd. Kwantumput vermenging (QWI - quantum well intermixing) is een diffusie methode die lokaal de samenstelling van de kwantumputten veranderd. Multiple die bonding is de integratie van verschil-



(a) O-band epitaxiale lagenstructuur. De anneal temperatuur was 700° C.

(b) C-band lagenstructuur. De anneal temperatuur was 725° C.

Figuur 1: PL piek golflengte verschuiving en netto verschuiving in functie van de anneal tijd.

lende epitaxiale lagenstructuren op eenzelfde silicium chip door middel van traditionele bonding en transfer printing is een nieuwe integratietechniek waarbij slechts een kleine coupon wordt verplaatst van het III-V bronmateriaal naar de silicium chip. Verder onderzochten we in dit werk 2 verschillende types van breedbandige bronnen. Elektrisch geïnjecteerde multi-bandkloof SLEDs genereerden een optisch spectrum van wel 300 nm breed. De optisch gepompte membraan-LEDs zijn een voorbeeld van eenvoud en efficiëntie gecombineerd in één bron.

2 De integratie van verschillende actieve bandkloven op silicium

De heterogene integratie van III-V materiaal op een siliciumgebaseerd fotonisch circuit (Si PIC) voegt het broodnodige winstmateriaal toe aan het silicium-fotonica ecosysteem. Maar voor veel toepassingen is één extra actieve bandkloof onvoldoende. In het kader van dit proefschrift onderzochten we drie uiteenlopende methodes om verschillende actieve bandkloven samen te brengen op een silicium chip.

Kwantumput vermenging

De vorm en samenstelling van een kwantumput bepalen de bijhorende bandkloof. Dus als we deze eigenschappen lokaal kunnen veranderen, dan ontstaat er een extra bandkloof. Dit is exact wat gebeurt als we atomen herordenen door middel van kwantumput vermenging (QWI). In dit onderzoek werd de herordening gestimuleerd door de aanwezigheid van onzuiverheden: fosforionen worden geïmplanteerd in een buffer laag. Tijdens een thermische anneal diffunderen deze ionen in de kwantumput laag, waar ze dienen als katalysator voor de diffusie van atomen aan de interface tussen de kwantumput en een barrière-laag. Dit zorgt ervoor dat de vorm van het energiediagram verandert, wat resulteert in een blauwverschuiving van de bandkloof. Experimentele resultaten van de verschoven fotoluminescentie (PL) piek zijn weergegeven in figuur 1. Op sommige plaatsen zijn

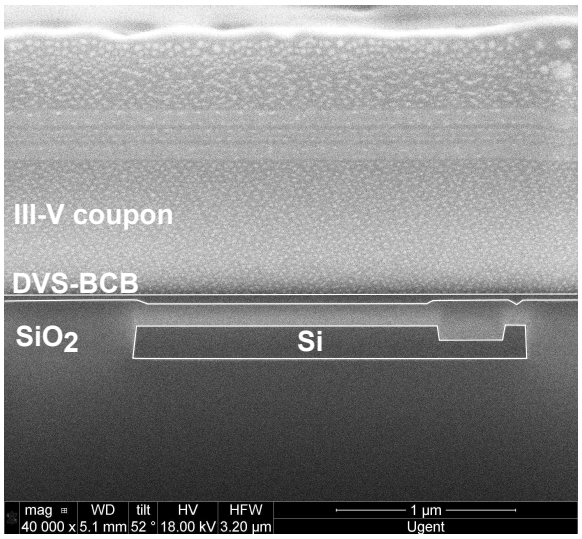
geen fosforionen aanwezig, aangezien de implantatie werd verhinderd door een masker. Deze regio's worden in de figuur aangeduid als unimplanted en we zien daar enkel een verschuiving wanneer de thermische anneal lang duurt. In dat geval zullen onzuiverheden reeds aanwezig in het gegroeide rooster verantwoordelijk zijn voor de herordening. We merken op dat de lagenstructuur in figuur 1a veel stabieler is dan die gebruikt in figuur 1b. Dit wordt bevestigd door een wiskundig model van de difussiecoëfficiënt gebaseerd op de vorm van de kwantumput en de PL verschuiving. Na de QWI wordt de bufferlaag verwijderd. In een volgende stap wordt traditionele, zij het gealigneerde, bonding gebruikt om het InP materiaal te integreren met een Si PIC. Met behulp van Fabry-Perot lasers als teststructuren, werd aangetoond dat de herordering geen effect heeft op de performantie van de lichtbron. De kwaliteit van de kwantumputten blijft dus hoog.

QWI is een krachtige techniek als we verschillende bandkloven willen maken spectraal dicht bij elkaar, aangezien de nettoverschuiving van de golflengte gelimiteerd is. Het fabricatieprocess van het III-V steunt op litografie, waardoor de spatiale resolutie grotendeels bepaald door de nauwkeurigheid van de bonding, wat ongeveer 1 μm is. Uitstekende controle van het fabricageproces is noodzakelijk. Elke imperfectie zal immers worden versterkt tijdens de thermische anneal, wat een goede bonding onmogelijk maakt.

Multiple die bonding

Het bonden van meerdere III-V dies lijkt de meest voor de hand liggende optie om meerdere actieve bandkloven samen te brengen op 1 chip. Wanneer we twee dies dicht bij elkaar plaatsen, ontstaan er echter een aantal nieuwe problemen. Aangezien de hoogte van de twee substraten niet dezelfde is, wordt een dun grafieten folie gebruikt om de toegepaste druk te verspreiden over de substraten. Verder moet de zone tussen de dies worden beschermd tegen de processtappen na de bonding, niet in het minst tegen de InP substraat verwijdering. Hiervoor wordt een conform oxide gedeponeerd en geopend op een zelf-gealigneerde manier. Met beide traditionele bonding technieken (moleculaire bonding en BCB bonding) werd op die manier een opening gecreëerd van minder dan 100 μm.

Het grote voordeel van multiple die bonding is de volledige vrijheid die we hebben in onze keuze van de lagenstructuren. Daar waar we bij QWI de vorm en samenstelling van de kwantumputten slechts lichtjes kunnen veranderen, laat multiple die bonding toe om volledig andere materialen te gebruiken en indien nodig zelfs de bulklagen te veranderen. Er is echter wel een praktisch probleem. Aangezien de minimale dimensie van een bondbaar stuk III-V enkele milimeters bedraagt, is de spatiale resolutie van deze techniek extreem laag. Dit wil zeggen dat wanneer we 2 dies integreren op een typische Si PIC chip, de ene helft volledig zal bedekt zijn met die A en de andere helft met die B. De weinig flexibele positie van de actieve bandkloof bemoeilijkt zo het design van het optisch circuit en een groot designoppervlak kan hierdoor verloren gaan. Wanneer we meer dan 2 III-V dies willen bonden op dezelfde Si chip, wordt de complexiteit snel onwerkbaar.



Figuur 2: Doorsnede van een coupon geïntegreerd op een planaire SOI chip gecoated met BCB door middel van transfer printing

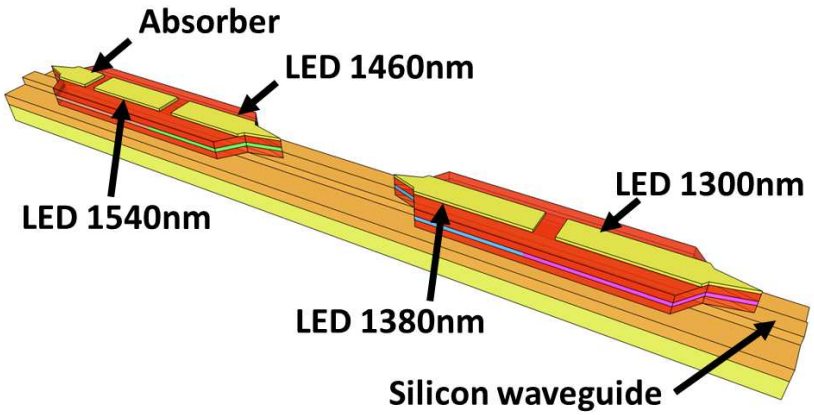
Transfer printing

Bij micro transfer printing, of kort gezegd transfer printing, worden de minimale dimensies die we kunnen verplaatsen gereduceerd tot micrometers. Met behulp van een rubberen stempel worden kleine stukjes III-V materiaal, genaamd coupons, opgepakt van het InP substraat en op een Si PIC geprint. Door gebruik te maken van stempels met verschillende oppik-posten, kunnen volledige rijen van coupons in één keer verplaatst worden, waardoor we in parallel kunnen werken. Deze techniek werd voor het eerst gebruikt voor een geïntegreerde fotonische toepassing in dit werk. In figuur 2 wordt een doorsnede getoond van de hechtingslaag na printen.

We zien dat de BCB bonding laag zeer dun is ($\pm$ 50 nm) zodat een sterke koppeling tussen het III-V materiaal en het Si gegarandeerd wordt. De kwaliteit van de bonding is hoog en kan de verdere processing van het geprinte materiaal te overleven.

Transfer printing steunt op de kinetische respons van de adhesie van de rubberen stempel aan het III-V halfgeleidermateriaal. Hoe hoger de snelheid van de stempel, hoe beter de adhesie. Om de sterke halfgeleiderinterfaces te overwinnen, wordt de volledige coupon ondergeëtst zodat de volledige structuur door niets meer wordt gesteund dan enkele tethers. Deze tethers zijn zo ontworpen dat ze makkelijk breken wanneer de coupon wordt opgepakt. In dit proefschrift werd het etsproces, nodig om de coupons vrij te maken, ontwikkeld daarbij rekening houdend met de nood voor een etsselectiviteit hoger dan 1000.

Net zoals bij de multiple die bonding kunnen we coupons met verschillende lagenstructuren op éénzelfde silicium chip plaatsen. We beschikken over een vol-



Figuur 3: Illustratie van een SLED op SOI bestaande uit 4 actieve bandkloven, gecreëerd door middel van QWI en multiple die bonding

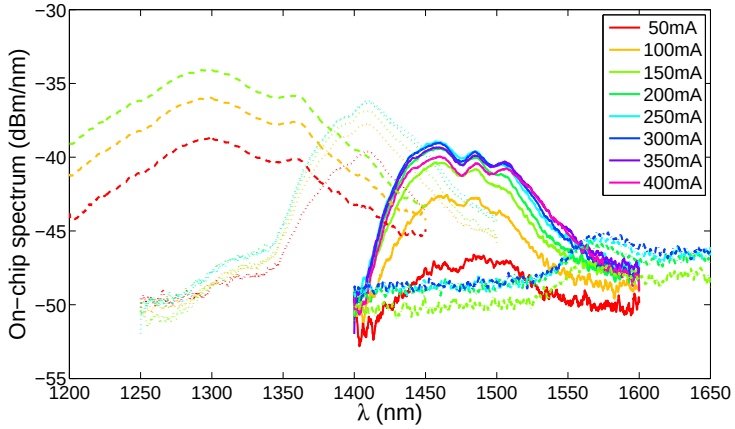
ledige vrijheid in onze keuze van de verschillende lagen in de III-V lagenstructuur. Omdat de coupons dimensies hebben in de orde van micrometers, hebben we niet langer problemen met de spatiale resolutie. We kunnen een coupon plaatsen op 1 μm nauwkeurig en verschillende coupons kunnen op een plaats naar keuze geprint worden. Deze nieuwe integratietechniek lijkt het beste van de 2 vorige technieken te combineren.

3 Elektrisch geïnjecteerde multi-bandkloof SLEDs

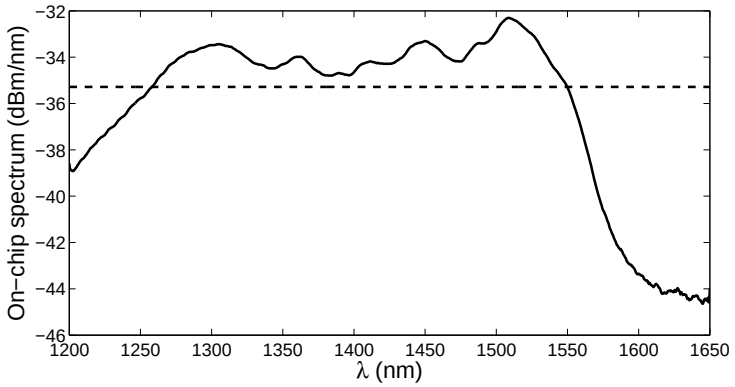
We demonstreerden elektrisch geïnjecteerde multi-bandkloof SLEDs op SOI door voor de allereerste keer QWI te combineren met multiple die bonding. De verschillende golflengtes werden gecombineerd zoals geïllustreerd in figuur 3.

De multi-bandkloof SLED bestaat uit twee sub-bronnen, afkomstig van twee verschillende gebonde InP dies. Elke die bevat 2 verschillende bandkloven als gevolg van QWI. De langste golflengtes propageren door de actieve regio's die kortere golflengtes generen. Aangezien de bandkloof groter en groter wordt, is elke sectie transparant voor alle eerdere secties. Dit wil uiteraard zeggen dat de bron slechts werkt in 1 richting. De spectra van de verschillende secties zijn als het ware gecombineerd in serie.

Na ontwerp en fabricage, werden de emissiespectra van de verschillende individuele secties opgemeten en gekarakteriseerd zoals getoond wordt in figuur 4a. De PL piek van de 4 secties is zodanig gekozen dat 2 opeenvolgende secties 80 nm van elkaar liggen. Door alle secties te pompen, verkrijgen we op die manier een optisch spectrum met een zeer brede 3 dB bandbreedte van 292 nm, zoals getoond in figuur 4b.

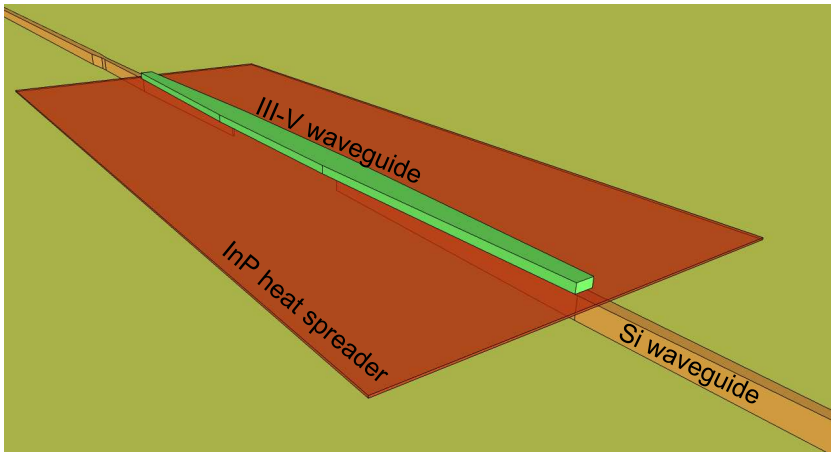


(a) Spectra verkregen door de verschillende bandkloven afzonderlijk te pompen. De gestreepte, gestippelde, volle en gestreept-gestippelde lijnen stellen de verschillende bandkloven voor.



(b) Spectrum verkregen door de elektrische stroom zodanig in de verschillende secties te injecteren dat de 3 dB bandbreedte van de SLED zo groot mogelijk is. De bandkloven van 1300 nm, 1380 nm, 1460 nm en 1540 nm werden gepompt aan respectievelijk 70 mA, 50 mA, 300 mA en 140 mA. De gestreepte lijn duidt de 3 dB bandbreedte aan.

Figuur 4: Spectrale karakterisatie van de elektrisch geïnjecteerde multi-bandkloof SLED

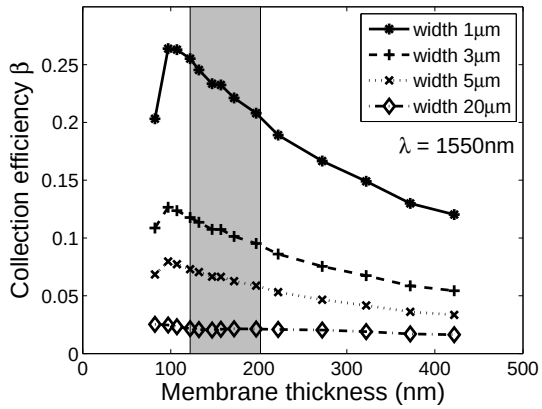


Figuur 5: Illustratie van een optisch gepompte membraan-LED

4 Optisch gepompte membraan-LEDs

De membraan-LEDs proberen vermogenefficiëntie te combineren met eenvoud en fabricage tolerantie. Het idee is ontstaan uit de observatie dat breedbandige bronnen kunnen worden opgedeeld in twee groepen. Golflengte-afstembare lasers en rijen van lasers kunnen zeer efficiënt zijn, maar hebben helaas vaak te lijden onder een grote gevoeligheid voor fabricage-imperfecties. Een excellente controle van het III-V proces is daarom onontbeerlijk, iets wat we hadden gehoopt te vermijden door te werken op een silicium substraat. SLEDs daarentegen zijn veel robuuster. Maar als gevolg van de kleine numerieke apertuur van een typische golfgeleidermode, kan de spontane emissie nooit efficiënt opgevangen worden. De koppelingsefficiëntie is typisch maar 0.1%.

Door de vereiste van een elektrische pomp te laten vallen, creëren we nieuwe ontwerpmogelijkheden die ons kunnen helpen om de spontane emissie efficiënt in de golfgeleider te collecteren. Aangezien we nu niet langer elektrische stroom kunnen gebruiken om de bron te pompen, moeten we overschakelen op een optische pomp. Dit nieuwe concept voor een breedbandige lichtbron is geïllustreerd in figuur 5. Veronderstel dat de optische pomp invalt op het membraan door middel van een silicium golfgeleider. De pomp koppelt van het silicium in het III-V met behulp van een dubbele taperstructuur. In de actieve lagen wordt het pomplicht sterk geabsorbeerd als gevolg van de hoge opsluiting van dit licht in de actieve regio. De gegenereerde electron-gat paren zullen vervolgens recombineren. Dit resulteert in isotrope spontane emissie. Als gevolg van het grote indexcontrast, is de numerieke apertuur van de golfgeleider heel groot. Daardoor wordt een groot deel van het gegenereerde licht gevangen in de golfgeleidermode. De gevangen fotonen kunnen nu koppelen naar de siliciumgolfgeleider door één van beide tapers. Indien een foton opnieuw wordt geabsorbeerd kan het hele verhaal herbeginnen:



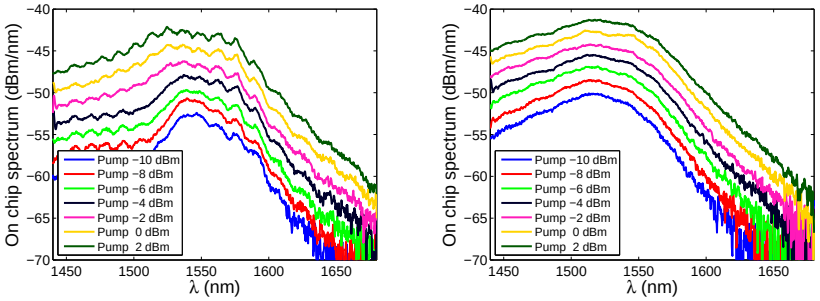
Figuur 6: Collectie-efficiëntie van de spontane emissie in functie van de membraandikte voor verschillende breedtes. De grijze zone duidt de diktes aan die experimenteel werden bestudeerd, gaande van 122 nm tot 202 nm.

er ontstaat een nieuw electron-gat paar dat op zijn beurt weer kan recombineren.

Een heleboel simulaties werden uitgevoerd om dit nieuwe idee te staven. Het kernidee van deze lichtbron is de afhankelijkheid van de collectie-efficiëntie van de dikte van het membraan, zoals wordt getoond in figuur 6. Een maximale koppeling van 26% werd bereikt voor een 1 μm-breed membraan met een dikte van 100 nm. De koppelingefficiëntie neemt af met toenemende breedte, maar blijft ruim boven de typische waarde van 0.1%. Een gelijkaardige grafiek kan worden gemaakt voor de opsluiting van het pomplicht in de actieve laag: een hoge opsluiting leidt tot efficiënte absorptie. De tapers werden eveneens bestudeerd. Als gevolg van de lage effectieve index van de golfgeleidermode van het III-V kunnen korte maar tolerante tapers gebruikt worden. Door het gebruik van een silicium nitride omhulsellaag, daalt de impact van opwarming van de bron.

Een eerste generatie LEDs toonde aan dat het principe achter de nieuwe lichtbron correct is en werkt. Zowel de bronnen met InGaAsP-kwantumputten als de LEDs op basis van chirped quantum dots (in het nederlands: kwetterende kwantumpunten) vertoonden een gelijkaardige prestatie. Helaas was de totale vermogensefficiëntie lager dan verhoopt met 0.1%. Dit gaf aanleiding tot een diepgaande studie van de LED waarbij zowel de optische als de elektrische eigenschappen van het membraan onder de loep werden genomen. Uiteindelijk konden we de efficiëntie een grootte-orde laten stijgen door de InGaAsP kwantumputten te vervangen door InAsP kwantumputten. Doordat de snelheid van de oppervlakterecombinatie afneemt, stijgt de totale efficië tot 1%. Figuur 7 toont de optische spectra van zowel de voorwaarts als de achterwaarts propagerende mode.

We denken dat de lichtbron nog verder kan worden geoptimalizeerd door de radiative levensduur te verkorten door middel van doping. Dit werd reeds getest op het InGaAsP materiaal, maar nog niet op de InAsP kwantumputten.



(a) Spectra in transmissie-mode *(b) Spectra in reflectie-mode*

Figuur 7: Karakterisatie van de optische LED met een InAsP actieve laag. De golflengte van de pomp is 1310 nm. Het pompvermogen op de chip staat vermeld in de legende.

In dit proefschrift werd het idee voor de optisch gepompte LEDs voor het eerst geopperd en werd het werkingsprincipe aangetoond. De efficiëntie van de tweede generatie van lichtbronnen was reeds een orde hoger dan die van de eerste. Aangezien dit type LED nog heel nieuw is zijn we ervan overtuigd dat in de toekomst nog meer progressie mogelijk is.

Referenties

[1] International Diabetes Federation. *IDF Diabetes Atlas - 7th Edition*, March 2017.

English summary

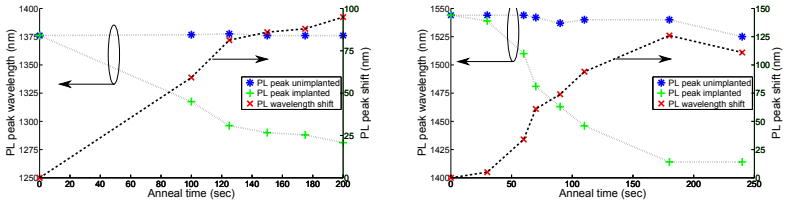
1 Broadband sources on silicon

415 million, rising to 642 million in 2040. This is how many people will suffer from diabetes worldwide. Because it is impossible to grasp numbers as great as these, we can also speak in relative terms and say that approximately 1 out of 10 suffers from the disease. Diabetes is not innocent either as high blood glucose is thought to be the third highest risk factor for premature mortality. [1] Although it is not curable, the blood sugar level can be kept within safe boundaries by applying synthetic insulin. As this level fluctuates heavily by nature, continuous monitoring is necessary.

Silicon photonics has made big waves in the past decade. Piggybacking on the massive investments already made by the microelectronics industry, silicon photonics has the promise of low-cost chips when high volumes can be produced. While initially mainly telecom and datacom were targeted as applications, recently also sensing sprang into view. Continuous glucose monitoring is a great example where silicon photonics can make a difference, because consumer (and hence low-cost) products are needed for a giant market.

While silicon has proven its worth on the passive optical functionality and also high quality detectors are in its portfolio, light sources remain a problem. Since silicon has an indirect band gap, other materials must provide gain. Thereby, we must never forget that only scalable solutions for light source integration will retain the mass production effect of silicon photonics.

Heterogeneous integration of III-V materials on the Si platform falls perfectly into this story, which was the starting point of our research. In this dissertation we have looked for integrated, broadband light sources on silicon. Three methods of creating multiple active band gaps on silicon by means of heterogeneous integration were studied. Quantum well intermixing is a disordering method which changes the composition of the quantum wells locally, multiple die bonding refers to the integration of more than one epitaxial layer stack using traditional bonding methods and transfer printing is a new integration technique where only small coupons are transferred from a source substrate to the silicon target. Furthermore, we have studied two types of devices. Electrically pumped multi-bandgap SLEDs achieved a broad optical spectrum of 300 nm, while the optically pumped membrane-based LEDs tried to marry efficiency and simplicity.



(a) O-band layer stack. The annealing temperature was 700° C.

(b) C-band layer stack. The annealing temperature was 725° C.

Figure 1: PL peak wavelength shift and net shift as a function of anneal time.

2 Creating multiple active band gaps on silicon

Bonding III-V sheet material to a silicon photonic integrated circuit (Si PIC) provides the much-needed gain material to the silicon photonics ecosystem. Often, one extra active band gap is not enough. We have looked into three different manners of creating multiple active band gaps on silicon

Quantum well intermixing

The shape and composition of the quantum wells dictate the band gap, so if these could be changed in certain locations an extra band gap is created. This is exactly what is done by atom disordering in quantum well intermixing (QWI). In our implementation, implantation enhanced disordering, phosphorous ions are implanted into a buffer layer. In a thermal anneal these are then diffused into the quantum well region, where they act as catalysts for the diffusion of atoms at the quantum well / barrier interface. The shape of the energy diagram changes, causing the band gap to blue-shift. Experimental data of the shifting photoluminesce (PL) peak can be seen in figure 1. In certain locations, no phosphorous ions are present as the implantation was masked. These regions are indicated in the figure as unimplanted and it can be seen that they only shift when the thermal anneal is very long. At this point impurities in the growth lattice aid in the disordering. It can be seen that the layer stack of figure 1a is much more stable than the one of figure 1b. This is confirmed by mathematical modeling of the diffusion coefficient based on the quantum well shape and the PL shift. After QWI, the buffer layer is removed and a traditional, be it aligned, bonding can be carried out to integrate the InP material with a Si PIC. The disordering has no effect on the device performance as shown by Fabry-Perot laser test structures, hence the QWs remain of high quality.

QWI is a very powerful technique if one wants to create multiple bandgaps which don't differ too much since the net shift is limited. Because the III-V process relies on lithography, the spatial resolution is mostly defined by the bonding accuracy, which is around 1 μm. Excellent fabrication process control is necessary since any imperfection will be highlighted by the thermal anneal and prevent

bonding.

Multiple die bonding

Bonding more than one III-V die seems the most straight-forward option to create multiple active band gaps. When placing two dies close together, a few new problems show up. Because the two substrates vary in height, the applied pressure must be spread out using graphite sheets. The region between the dies must be protected from the post-bond processing, not in the least the InP substrate removal. For this a conformal oxide was deposited and opened in a self-aligned manner. With both traditional bonding techniques (molecular and BCB bonding) we achieved a gap between the dies of below 100 μm.

The benefit of multiple die bonding is that we have complete freedom of choice with regards to the layer stacks. Where in QWI we could only change the shape and composition of the quantum wells slightly, here we can use different materials altogether and also change bulk layers. There is a practical problem though. Because the minimum size of a bondable III-V die is around a few millimeters, the spatial resolution of the technique is extremely low. When bonding two dies onto a Si PIC of a typical footprint, it means one half is completely covered with die A and the other half by die B. The optical circuit design may look awkward due to the inflexibility of the active band gap location and a lot of design space is offered to this. When more than two III-V dies are considered, the complexity quickly becomes intolerable.

Transfer printing

In micro transfer printing, or in short transfer printing, the transferable dimension is reduced to micrometers. Using a rubber stamp, small so-called coupons of III-V material can be picked up from the InP substrate and printed onto a Si PIC. By using multiple posts on a single stamp, a complete array of coupons can be transferred in one go, allowing for the parallelism. This technique was applied for photonic integrated circuit applications for the first time in this work. Figure 2 shows a cross section of the bonding interface after printing. It can be seen that the bonding BCB layer is very thin ($\pm$ 50 nm) which ensures a strong coupling between the III-V and the Si. The bond quality is high and is able to withstand further processing of the transferred material.

Transfer printing relies on the kinetic response of the adhesion of the rubber stamp to the III-V semiconductor material. The higher the velocity, the better the adhesion. The entire coupon is undercut, creating a free-hanging structure held only by a few tethers. These tethers can be engineered to break easily when a coupon is being picked. The release process was developed thereby adhering to the requirement of etching selectivities over 1000.

Similar to multiple die bonding, we can transfer coupons of different layer stacks to one silicon target. There is virtually complete freedom on the layers in the III-V layer stack. Because the coupons are only micrometer-sized, they can be freely placed as well, with a placement accuracy is within 1 μm and the different coupons can be placed at locations at will. This novel integration method seems to

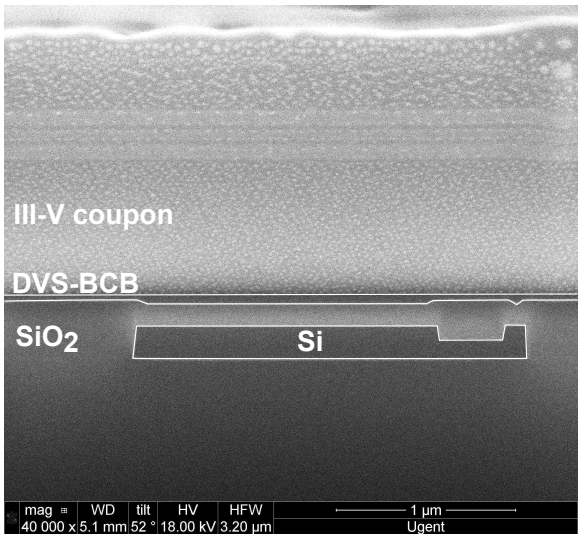


Figure 2: Cross section of a coupon transfer printed on BCB-coated planar SOI

combine the best of the previous two techniques.

3 Electrically pumped multi-bandgap SLEDs

Combining QWI and multiple die bonding for the first time, electrically pumped multi-bandgap SLEDs were demonstrated on SOI. The different wavelengths were combined as indicated in figure 3. The multi-bandgap SLED consists of two sub-devices, from the two bonded InP dies, each composed of two different bandgaps due to QWI. The longest wavelengths travel through the active regions of the shorter ones. Because the band gap widens progressively, each section is transparent to the previous ones. This also means that this is a one-directional device. One could say the spectra of the different sections are combined in series.

After design and fabrication, the emission of the different individual sections can be characterized as shown in figure 4a. The PL peaks of the four sections were chosen such that they are spaced 80 nm apart from each other. In this manner, when all sections are pumped, we can generate an optical spectrum with a very broad 3 dB bandwidth of 292 nm, shown in figure 4b.

4 Optically pumped membrane-based LEDs

The membrane-based LEDs try to link power efficiency to simplicity and tolerance. The idea arose from the observation that for broadband sources we generally have two categories. Tunable lasers and laser arrays can be very efficient,

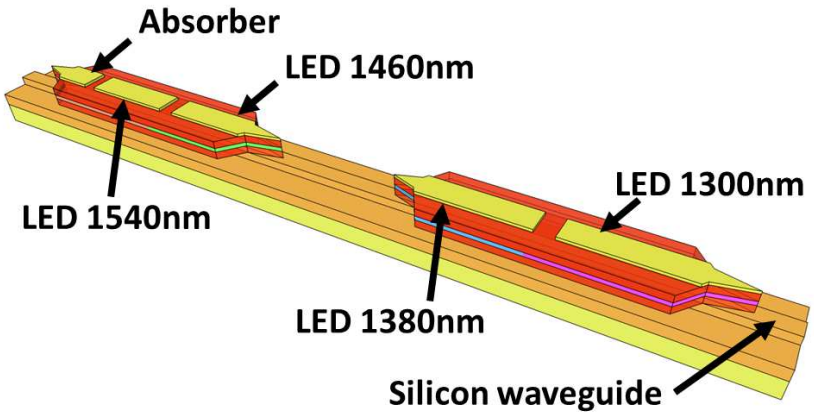
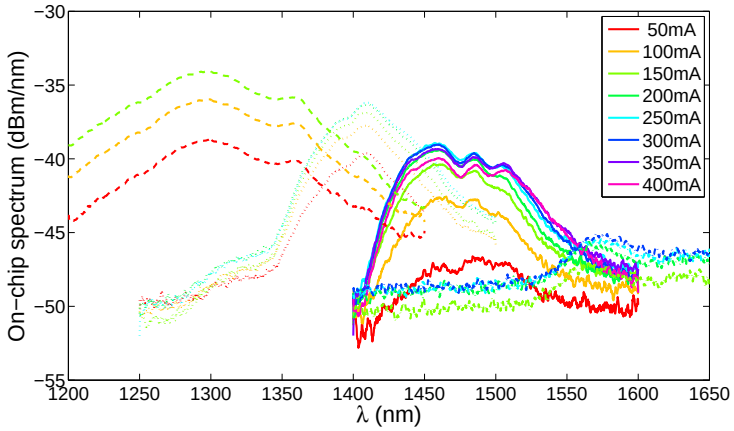


Figure 3: Illustration of a four-bandgap SLED on SOI, based on quantum well intermixing and multiple die bonding

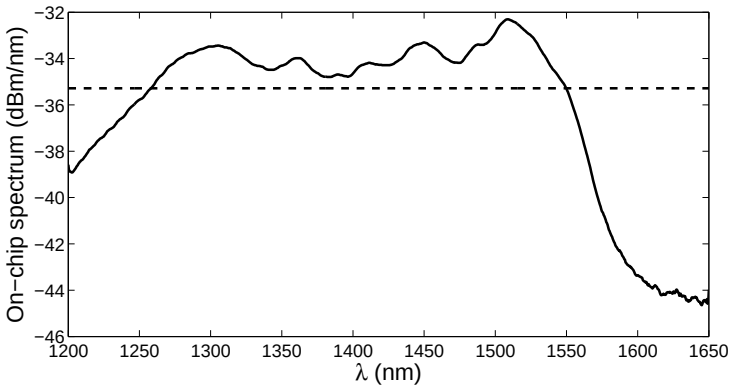
but are very susceptible to fabrication imperfections. Excellent control on III-V processing must be retained, which we had hoped to avoid by working on silicon substrates. SLEDs on the other hand are more robust. However, due to the small numerical aperture of the typical waveguide mode, spontaneous emission can never be efficiently collected. Typically the coupling efficiency is only 0.1%.

By letting go of the requirement of electrical pumping, we open up design space to boost the collection efficiency of spontaneous emission. No longer being able to electrically drive the device, optical pumping must be applied. The device concept is illustrated in figure 5. Suppose an optical pump is incident on the membrane through the silicon waveguide. The pump will couple up through the double taper structure and be strongly absorbed in the active region due to the strong confinement in the quantum wells. The generated electron-hole pairs will recombine, giving rise to isotropic spontaneous emission. Because of the high refractive index contrast, the numerical aperture of the waveguide mode is very large. A large fraction of the generate light is capture in the waveguide mode. The captured photons may now couple through either of the tapers to the silicon waveguide (co-propagating and counter-propagating). In case the photon is reabsorbed, the story recommences as an electron-hole pair was just created, which in turn can recombine.

A series of simulations was done to confirm our assumptions of this novel idea. Naturally, the core of the device concept is the dependence of collection efficiency on membrane thickness, shown in figure 6. A maximum coupling of 26% is achieved when the 1μm wide membrane is 100 nm thick. With increasing width the coupling decreases, but it remains well above the typical value of 0.1%. Similar graphs can be presented for the confinement of the pump, where a high confinement leads to efficient absorption. Also the tapers were looked at. Because of the low effective index of the III-V waveguide mode, a short but tolerant taper



(a) Spectra of the different band gap sections pumped individually. The dashed, dotted, full and dash-dotted lines symbolize the different sections.



(b) Spectrum when pumping the different sections of the SLED such that the 3 dB bandwidth is as broad as possible. The sections at 1300 nm, 1380 nm, 1460 nm and 1540 nm were pumped at 70 mA, 50 mA, 300 mA and 140 mA respectively. The dashed line indicates the 3 dB bandwidth.

Figure 4: Spectral characterization of the electrically pumped multi-bandgap SLED

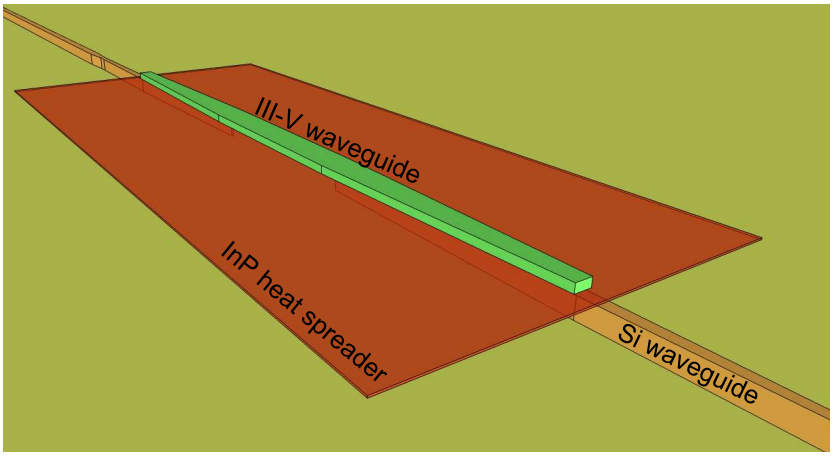


Figure 5: Illustration of the optically pumped membrane-based LED

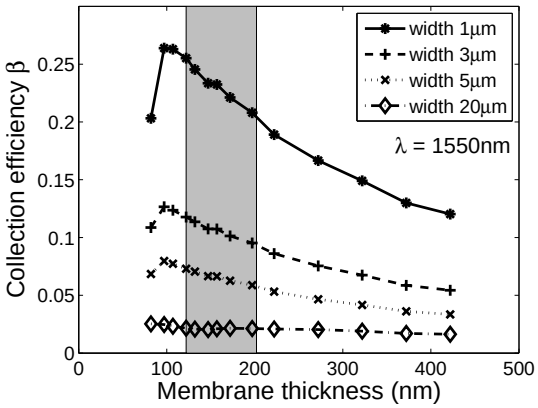
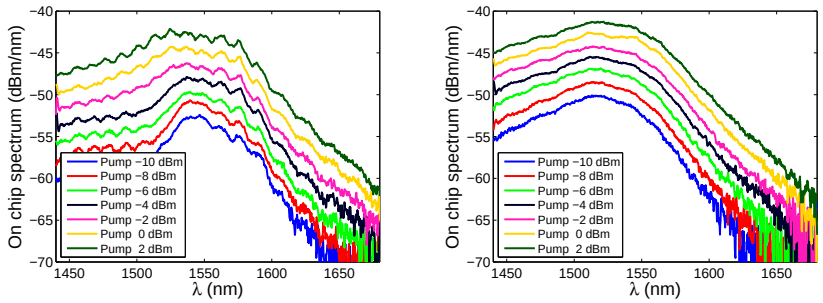


Figure 6: Spontaneous emission collection efficiency as a function membrane thickness for various device widths. The shaded area indicates the experimentally studied thicknesses, ranging from 122 nm to 202 nm.



(a) Spectra in transmissive mode *(b) Spectra in reflective mode*

Figure 7: Characterization of the InAsP-based LED. The pump wavelength was 1310 nm and the on-chip pump power is indicated in the legend.

can be realized. Working with silicon nitride capping layers reduces the impact of heat.

First generation devices showed the proof of principle of this novel device concept. Both InGaAsP-based quantum wells and InAs-based chirped quantum dots showed similar performance. The total power efficiency was below expectations though, at 0.1%. This sparked a thorough investigation whereby we looked at both the optical and electrical properties of our membranes. Ultimately, the efficiency was increased by one order by moving away from InGaAsP to InAsP-based material. The reduced surface recombination velocity boosts the efficiency to 1%. Figure 7 shows the optical spectra both in co-propagating and in counter-propagating mode. We think the device can further be optimized by using doping to reduce the radiative lifetime. This was tested on InGaAsP-based material, but not yet on InAsP quantum wells.

In conclusion, this novel idea of optically pumped LEDs has been proposed and the working principle proven. Our second generation of devices gained already one order in efficiency. Given how new these devices are, we are convinced more gains are to be made in further generations.

References

[1] International Diabetes Federation. *IDF Diabetes Atlas - 7th Edition*, March 2017.

1 Introduction

Diabetes is a chronic disease affecting 415 million people worldwide. Moreover, the prevalence of the affliction is projected to grow from 8.8% to 10.4% of the population in the next 25 years. In combination with the population growth, this means there will 642 million patients by 2040. The World Health Organization (WHO) estimates that globally, high blood glucose is the third highest risk factor for premature mortality, after high blood pressure and tobacco use. [1]

Patients with diabetes have a too high blood sugar level, so-called hyperglycemia. During a hyperglycemic event, the body cannot extract enough glucose from the blood so fat is broken down to use for energy. This process creates ketones as a waste product, which are toxic in high levels. Most of the ketones are excreted through the blood, but still a build-up may occur, having a coma or even death as a result. The blood sugar level is governed by a hormone called insulin. The root cause of diabetes can be a malfunction of the pancreas to produce enough insulin (type 1 diabetes) or when the body is insulin-resistant (type 2). To avoid the above-mentioned hyperglycemia, synthetic insulin must be applied to bring the high blood sugar level down. Care is needed though since now, due to too much insulin, hypoglycemia may occur, i.e. the blood sugar level is too low, which could again lead to a coma.

Clearly a tight control on the blood sugar level is needed. This reduces the health complications and the overall cost of a diabetes treatment. Continuous glucose monitoring (CGM) is crucial in this respect. While a drop of blood on a testing strip will tell you the blood sugar level at a certain time with high accuracy, the information gained is limited as no insight on the dynamics can be acquired.

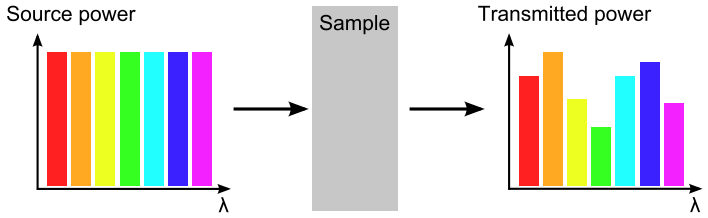


Figure 1.1: Principle of NIR absorption spectroscopy

CGM is the regular measurement of glucose concentration, the data of which can be used to predict events of hypo- or hyperglycemia. CGM however requires small, portable sensors. They should go unnoticed as not to hinder the patient and cannot in any way affect the human body.

Diabetes is only one example where small, affordable sensors can make a world of difference. The importance of measuring cannot be overstated in the 21st century. The proverb *"The numbers tell the tale"* has never been more appropriate.

1.1 Near-infrared absorption spectroscopy

There are many optical sensing techniques, which itself is only one of the many sensing methods. We would like to highlight near-infrared (NIR) absorption spectroscopy as it is highly suited for our application of diabetes.

NIR absorption spectroscopy measures the absorption fingerprint of a molecule of interest (e.g. glucose). The broadband light of a source is sent through a sample and the transmitted power spectrum is analyzed, shown in figure 1.1. The strength of the measured absorption gives us the needed information on the concentration of glucose present in the blood, while the shape ensures the required specificity. Many different molecules are present in the blood but only glucose can give this particular absorption fingerprint.

As can be seen from figure 1.1, there are a few basic components needed for this system. The light source has to encompass multiple wavelengths. A perfectly flat spectrum is not really needed as one can compare the transmitted light to the source spectrum, but it makes life easier. Once the light is generated, it is sent through the sample. A long interaction length will increase the effect the sample has on the transmitted light. At the receiver side, the power spectrum is measured. After separating the different wavelengths, a detector array captures the light.

Unless the incident light on the sample is of very high intensity, NIR spectroscopy is non-destructive. This and the fact that the operating principle is simple makes that it is a very attractive scheme for a sensor, in particular a glucose sensor.

1.2 Silicon photonics

The cost of a device is relative to the number of users. When a hospital buys an MRI scanner, they may pay $500 000. This is a big amount of money to spend, but they will also use it on countless patients. Therefore the cost per patient is still acceptable. In consumer products, we cannot make use of this scaling. One device will only serve one customer (and generally only for a limited time). A low cost will always be one of the targets for consumer products.

In order to bring the cost per product down, it is wise to look for other scaling fields. The reason a laptop is cheap is not because multiple people use it, but because there is only a little extra cost when going from a production of 10 laptops to 10 000. Since laptop manufacturers are able to reach the high volumes, the cost per laptop drops.

In this thesis we want to make the scaling effect possible through integration on a silicon platform. In the next paragraph we will outline what the reasoning behind silicon photonics is and what is needed for this scaling to happen. We will then look at which of these conditions are satisfied.

1.2.1 Concept

Silicon photonics maximally tries to make use of the scaling effect depicted above. Silicon has been the substrate material of choice for many decades in the electronics industry and the investments made in the fabrication tools are gigantic. Piggybacking on this history as much as possible, silicon photonics has grown to one of the the most dynamic and promising branches of photonics today. [2] Besides the scaling effects, the rationale for using silicon as the basic material for photonic components is based on other advantages. Silicon is transparent from a wavelength of 1.1 μm to 8 μm, encompassing the traditional telecommunication wavelengths around 1.3 μm and 1.55 μm, as well as a lot of sensing windows. Silicon wafers have the lowest cost per unit area and the highest crystal quality of any semiconductor material. Even 300 mm diameter silicon-on-insulator (SOI) wafers are cheaper than 2 inch diameter InP wafers. [3] The refractive index contrast between silicon and its oxide is very high, giving rise to high optical confinement and sharp waveguide bends. This makes the components and the complete circuit very compact, allowing to embed a lot of functionality per unit area.

In silicon photonics, it is very expensive to design and fabricate one chip. The main portion of the cost is in the design and process development though, meaning that when high volumes are reached the price per chip drops significantly. Moreover, silicon has been proven reliable when scaling up fabrication as is clear from the example of transistors in electronics. When the volumes are big enough, as is the case with diabetes, the silicon chips can become very inexpensive.

The story of great scaling in silicon photonics only holds when the yield of the

process is high. Optical chips are very complicated and consist of many processing steps. If every step only has a yield of 90%, the total yield of 30 steps is less than 5% and per wafer only 1 in 20 dies will work. The supreme processing quality of silicon luckily pushes this yield to higher levels. However, in the source there can be problems though.

1.2.2 The bottleneck of silicon: light sources

Of all the assets in the silicon portfolio, light generation is not one. The band gap of silicon is indirect, meaning that a photon can only be generated with the assistance of a phonon. As a consequence, an efficient light source in silicon is physically impossible. Optical circuits only make sense if there is light to be manipulated. While there have been lasers demonstrated in silicon, there is a consensus that III-V material have to be introduced to efficiently create light. We list a few options at hand below and go into detail on the III-V-on-Si heterogeneous integration.

Fiber coupling an external packaged light source to the silicon chip: The easiest way to achieve light source integration with a silicon chip is by fiber pigtailing an external III-V source to the chip. The advantages of such an approach are that a high-end III-V laser can be temperature-controlled separately from the silicon chip and the yield of the two components can be optimized separately. But it has a large footprint, degraded performance due to coupling loss and reflections at each fiber coupling point. As it is not an integrated solution, it does not scale well and the multiple packaging steps introduce extra cost.

Die-level integration of an external packaged source: Here, a prefabricated source is attached to the silicon-on-insulator (SOI) chip using flip-chip, micro-packaging [4] or a micro-optical bench [5]; and either mechanical stops or machine vision are used to passively align the source to the silicon waveguide. Also this option is not integrated and does not scale as such. There are tight mechanical alignment requirements, which make the packaging cost high, or alternatively the yield low.

Epitaxial growth of III-V layers: Epitaxial growth of III-V layers on silicon has been widely investigated. This approach is truly integrated, but so far the material quality has not been of high quality. Due to the lattice mismatch between silicon and InP/GaAs, defects tend to get created during the growth. In addition, the thermal budget of the current silicon photonics process flow does not allow the high temperatures needed for good growth conditions. While this is a highly promising approach, it is still in its infancy.

Wafer bonding or III-V-on-silicon heterogeneous integration: Wafer bonding involves bonding an unpatterned III-V die or wafer to an SOI chip, either by covalent bonds or by using an adhesive, and processing it further to fabricate the

integrated laser. Since the III-V material is provided rather than devices, the fabrication still happens on the silicon wafers, thereby retaining the scaling argument. Numerous light sources have been demonstrated in this manner [6]. Both molecular and adhesive bonding techniques were employed in this thesis, we go more into detail in the following section. Since this is also a fully integrated approach, the packaging cost stays low and it is highly suited for low-cost, high-volume applications. A disadvantage of this technique is the waste of III-V material, since the sources only cover a fraction of the provided active material.

Micro transfer printing: Micro-transfer-print (TP) is a technology where materials or devices can be selectively removed from their source wafer and transferred in parallel to a new platform as a printing process. This technology was first introduced early 2000's [7], but we were the first to apply it to on a silicon photonic IC. Thereby we had to overcome problems specific to silicon photonics, e.g. the thin bottom III-V layer could only be etched by nanometers since a thick adhesive would prevent optical coupling. Our interest in TP was sparked by the disadvantage of bonding, mentioned above. The waste of material originates in the fact that the smallest III-V die that can be bonded is still of millimeter scale. TP allows us to bring this down to tens of microns, or in other words only what is needed for the light sources. TP brings the added advantage that combining two different active materials becomes straight-forward. Since transfer printing can be seen as a bonding process on smaller scale, it is still highly suited for low-cost, high-volume applications.

1.2.2.1 Molecular or direct bonding

Homogeneous direct bonding is one of the most applied bonding techniques, as it is used to created silicon-on-insulator (SOI). Since 2000, high quality direct wafer bonded SOI substrates up to 300 mm in diameter have been commercially available, showing the manufacturability and yield of the fabrication process. A key challenge when bonding dissimilar materials is the thermal expansion mismatch. SOI is produced at temperatures above 800°C, a temperature at which III-V materials evaporate. Because lower temperatures are necessary, O_2 plasma-assisted bonding has been developed. Figure 1.2 illustrates the typical process flow.

After an intense cleaning procedure, the native oxides on the SOI and III-V stack are removed. Both surfaces are then activated in the O_2 plasma surface treatment. In this step the hydrophobic surfaces are converted to hydrophilic ones by growing a thin layer of oxide, resulting in a high density of -OH groups (i.e. hydroxyl groups). The two samples are then brought in contact at room temperature. Van der Waals forces ensure the initial bond. At this stage, the III-V die is bonded to the SOI, but can easily be detached again. It is only during the following anneal

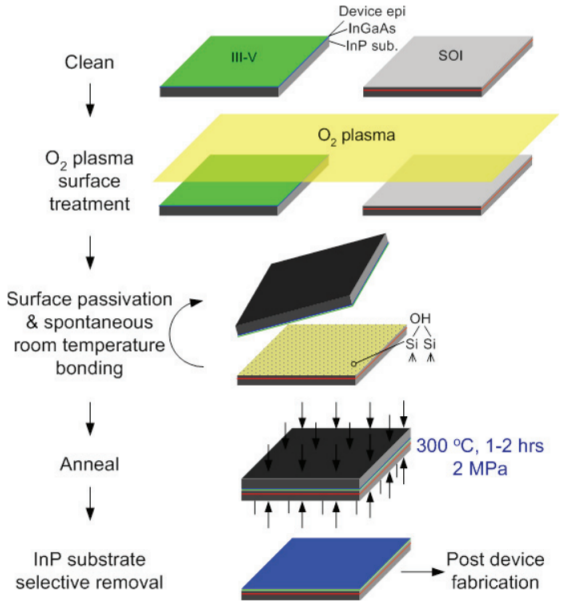


Figure 1.2: Operation principle of molecular bonding. [6]

that the covalent bonds are formed through the following chemical reactions.

$$\mathrm{Si{-}OH + M{-}OH \longrightarrow Si{-}O{-}M + H_2O(g)} \tag{1.1}$$

$$\mathrm{Si + 2\,H_2O \longrightarrow SiO_2 + 2\,H_2(g)} \tag{1.2}$$

where M stands for a metal with a low electronegativity, such as the group III and group V elements of the bonded III-V compound semiconductor. The anneal temperature in this process is now only 300°C, which all materials can easily handle. The gas byproducts of H_2O and H_2 of equations 1.1 and 1.2 can be a cause of concern as they can accumulate thereby creating a void at the bond interface. The outgassing problem is alleviated by etching so-called vertical outgassing channels in the SOI. A grid of square holes is etched through the silicon device layers into the buried oxide. The holes capture the gas byproducts, which can travel through the porous thermal buried oxide. Both the effect of vertical channels and void formation in the absence of them is clear in figure 1.3, where we see the bonded III-V die after InP substrate removal. Underneath one can see the SOI with waveguides and vertical channels patterned in it. A full stepper view is shown; at the bottom the next stepped Si pattern is already visible. The bulk of the image shows a voidless bond (the area which peeled off can be attributed to a different origin, discussed in section 2.3.2), i.e. the gas byproducts are channeled into the buried

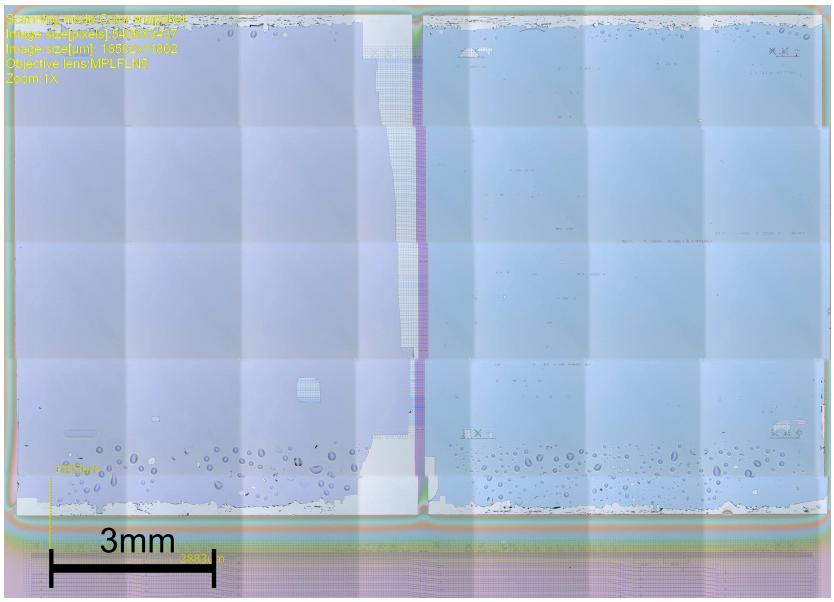


Figure 1.3: Void formation in molecular bonding and the effect of vertical channels. A top view is shown of two InP dies (left and right) bonded to SOI (visible in the middle). At the top and bottom of the figure, where the vertical channels are absent, bubbles are indicating void formation.

oxide preventing them from doing harm. At the top and bottom of the III-V die, numerous bubbles can be discerned. Because the stepper steps over this area, no vertical channels have been formed and therefore the gasses accumulate. The pitch and size of the vertical channels obviously have some influence. We used 8 μm by 8 μm holes in a 50 μm by 50 μm grid.

1.2.2.2 DVS-BCB bonding

Besides the direct bonding approach, adhesive bonding methods can also be used. The choice of adhesive is application dependent, for optical devices we selected DVS-BCB. An extensive review of adhesive wafer bonding can be found in [8].

Thermoplastic, elastomeric and thermosetting materials are typically used for adhesive bonding. Since thermoplastics melt and flow at the same melting temperatures each time they are thermally processed, the thermal budget is limited. The distinguishing characteristics of elastomers (often referred to as rubbers) are their ability to sustain large deformations with low stresses. Also here the thermal budget is limited, making it unsuitable for III-V processing (can go up to 400°C). Thermosetting materials such as DVS-BCB, epoxies, spin-on-glasses and polyimides do have the required thermal budget. During the thermal anneal they crosslink, creating a 3D network. Therefore they do not remelt during future thermal cycles. Polyimides create gaseous byproducts, resulting in large unbonded areas. Epoxies are often not optically transparent and spin-on-glasses tend to shrink excessively upon curing resulting in film stress. [6, 8]

DVS-BCB (divinylsiloxane-bis-benzocyclobutene), also referred to as BCB, does not suffer from these drawbacks. This thermosetting molecule is depicted in figure 1.4. A silicon backbone is terminated by two benzocyclobutene rings. These can be B-staged, i.e. upon curing the rings open to form a very reactive intermediate. A so-called Diels-Alder reaction finds place as the opened ring finds an available vinylsiloxane group. A three-dimensional network structure is formed. The most important take-away of figure 1.4 is that no byproducts are created and therefore one does not have to worry about accumulating gasses.

In order to couple from III-V to Si (and vice-versa) the bonding layer needs to be thin ($<$100 nm). This can be achieved by diluting the BCB with its solvent mesitylene. 50 nm thick bonding layers can be repeatedly achieved. Further diluting is possible, but a minimum thickness should be applied, determined by the remaining topography of the SOI waveguide circuit.

1.2.2.3 Van der Waals bonding

When bringing two flat materials close together, Van der Waals forces will ensure they bond (unless stronger forces counteract). The origin of Van der Waals forces

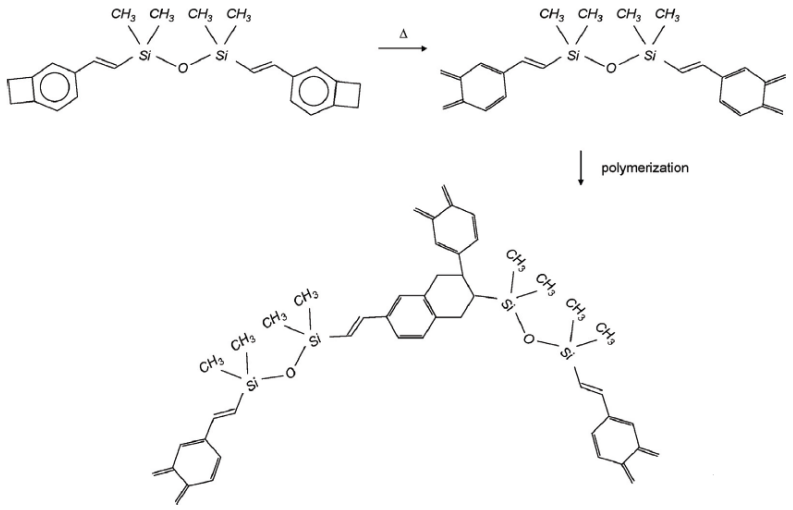


Figure 1.4: Polymerization reaction of DVS-BCB. [6]

is to be found in weak electrostatic forces between dipoles. We discern three types: dipole-dipole force, dipole-induced dipole force and dispersion force. [9]

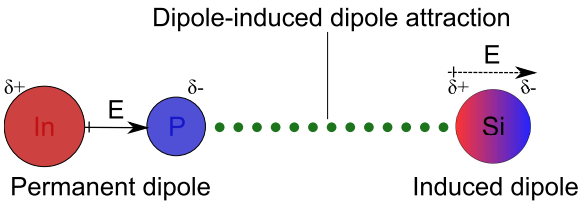
Although a molecule is electrically neutral as a whole, a charge distribution may occur, giving rise to a permanent dipole. The dipoles of two of these polar molecules will be attracted to each other through the electrostatic force. Therefore two polar molecules can bond through dipole-dipole force.

If no charge distribution occurs (e.g. in symmetrical molecules), no permanent dipole can occur. When a permanent dipole comes close however, it will induce a dipole moment in the non-polar molecule. This induced dipole in turn is attracted to the permanent one, giving rise to a dipole-induced dipole force.

In two non-polar molecules, no permanent dipole is present. The total dipole moment, averaged over the entire space, is zero, but the average of the square of the dipole moment is nonzero because the quantum mechanical wavefunctions are correlated. When two non-polar molecules do not interact, the wave function of the total is system is the product of ones of the separate molecules: $\psi_{\text{total}}(r_1, r_2) = \psi_1(r_1)\,\psi_2(r_2)$. In close contact, this is no longer true and the molecules behave as one system, thereby aligning themselves to reduce the system energy. [9]

Figure 1.5 shows two examples where Van der Waals forces become important. The first, figure 1.5a, shows a gecko holding on to glass. A gecko's feet contain nanofibers which ensure that he comes in close contact with the glass molecules. Since the nanofibers protrude to the glass irrespective of the glass roughness, he

(a) A gecko holding on to glass by means of Van der Waals forces. [10]



(b) Van der Waals forces in the III-V on Si platform.

Figure 1.5: Two examples of Van der Waals interaction.

can use the entire area to grip. This allows a gecko to run on walls, ceilings etc. of any material. Geckos' feet have been the subject of intensive research, in order to create a real-life Spiderman. Geckskin e.g., a project of University of Massachusetts Amherst, was named one of the top five science breakthroughs of 2012 by CNN Money [11].

Figure 1.5b shows the Van der Waals interaction in the III-V on Si platform. InP is polar due to the difference in electronegativity with the phosphorous atoms being the more negative ones. This means InP harbors a permanent dipole. Si is non-polar (because of symmetry), but the proximimity of the InP dipole induces a charge distribution as indicated in the illustration. As a consequence, an attractive dipole-induced dipoled force exists between the two molecules. The total bonding force can be large, provide the atoms can get very close together over the entire area. This means that surface roughnesses of less than 1 nm are needed.

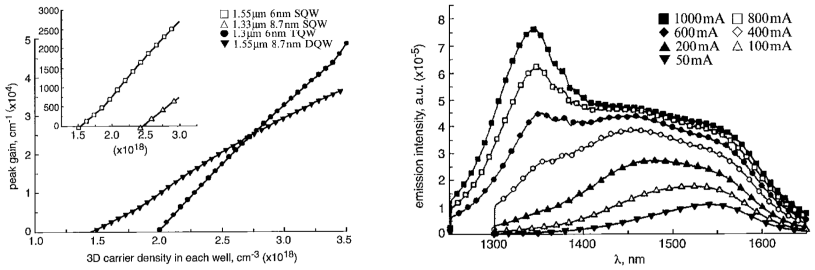
Also in molecular and BCB bonding, the initial bond (before the cure) is based on Van der Waals forces. This is a weak bond and a small kick to the substrate will overcome the bond energy. For transfer printing, the story is different.

1.3 Alternative broadband light sources

As optical sources with a high optical bandwidth are highly asked for, lots of research has been done on the subject. We have created sources based on quantum well intermixing, multiple die bonding and transfer printing, but also other ap-

proaches have been successful. We give a short overview below.

Dual quantum wells: In dual quantum wells, two different shapes of quantum wells are grown on top of each other. These different QW designs have different peak gains at different wavelengths. Typically one of the QWs will have a higher gain than the other at all carrier densities. By tuning the number of quantum wells, the gain curves of the two designs will cross at a certain point, at which the spectrum will be flat and the 3dB bandwidth is highest. This technique is illustrated in figure 1.6. Large bandwidths have been achieved with this technique, up to 280nm.



(a) Calculated peak gain in each MQW design. The inset shows the values for a single quantum well.

(b) Emission spectra of fabricated SLD at different injection currents

Figure 1.6: Dual quantum well design [12]

The typical output power is quite low however. The most prominent disadvantage is that this wide spectrum is only achieved at a sweet spot of carrier densities in the QWs. On top of this, defining this sweet spot in terms of current is very challenging. Because these structures have to be injected in series, the carrier densities are typically very non-uniform. When designing dual QWs one has to take into account factors such as QW transition energy, number and sequence of the different QWs, the thickness of the SCH layer, the selection of the dominant carrier, the ability of the QW to trap the 2D carriers, the uniformity of the 2D carriers within the QWs, etc. The control of some of these factors is not straight-forward and a small mistake might have detrimental consequences. Reordering the quantum wells could lead to the case where you never reach the sweet spot. [12–14]

Chirped quantum wells: Chirped quantum wells try to circumvent the current injection problem of dual quantum wells by filling up all the states in between. Even so, because of the difference in gain of the QWs and especially because of the complex current injection schemes only a limited range of currents can be designed for. The results are similar to the dual quantum well platform. [15]

Multi-state quantum wells: By widening the quantum wells one can allow for a second state in the conduction band. Again similar to the previous systems, one

can make sure to inject the system such that the gain of the ground state is equal to the higher order state. This requires reasonably high current injection and the broadening that can be achieved is typically below 100 nm. Precautions have to be taken in order to get around heating issues. This could be an even bigger challenge on the silicon photonics platform because of the buried oxide layer isolating the substrate from the waveguide. [16]

Single-state quantum wells: Contrary to the wide multi-state quantum wells, it is also an option to narrow the quantum wells. Thin quantum wells are used when one wants to make a broad spectrum at various currents. High current injection will broaden the spectrum considerably and because the upper state is pushed out, the gain will not abruptly increase at the blue side of the spectrum. The current injection of these thin wells is usually even higher than that of the wider wells mentioned above. Heating and saturation is a very big concern in these designs. [17]

Size inhomogeneity of quantum dots: Quantum dot research was originally driven in order to replace quantum wells as an extremely narrowband gain medium for lasers. Due to the non uniformity in the growth, they did not (yet) get to that stage. For an LED, these non uniform sizes of dots is a blessing though. Typically, also different sizes of quantum dots are grown on top of each other. The injection efficiency is therefore again rather complicated. A bandwidth of 100-200nm has been achieved. [18]

Supercontinuum due to non-linear effects: A completely different approach from the schemes above is to exploit non-linear optics. Third order nonlinearity in the susceptibility causes the spectrum of a high power laser to broaden and new bands far away from the original wavelength can be generated.[1] Usually one of the nonlinear effects such as four-wave mixing, self and cross phase modulation or Raman scattering dominates, but under the right conditions they can add up and create a so-called supercontinuum, as illustrated in figure 1.7. For the creation of a supercontinuum advanced dispersion engineering is usually needed and very high pump lasers over tens of watts peak power are required. At these high powers, extra loss factors such as two photon absorption come into play. With this technique, very high bandwidths [2] up to 1000nm can be achieved but an external pump laser is needed, making it not a fully integrated solution. [19] [20]

[1] Usually the second order nonlinearity is zero in crystalline materials because of inversion symmetry.

[2] 3dB band width is rarely quoted in nonlinear optics, because the pump is still many orders stronger than other wavelengths.

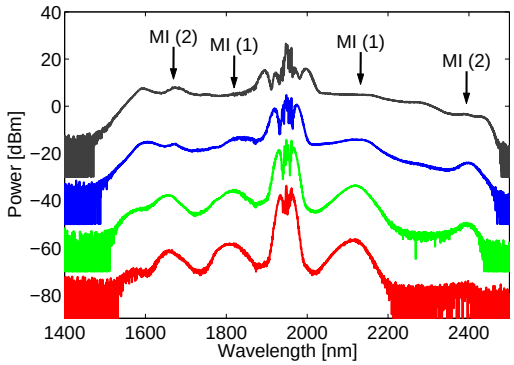


Figure 1.7: Build-up of a supercontinuum with increasing pump powers 11.2W, 12.7W, 16.6W, 28.1W. The spectra are vertically offset by multiples of 20dB for clarity. [19]

1.4 Outline

In this introductory chapter we started from the application side of the story. Sensors are of critical importance in today's world and often affordable, reliable, mass-produced devices are asked for. One such example is diabetes, a disease affecting 1 out 8 adults worldwide by 2040. An continuous sensor based on near-infrared spectroscopy will improve not only the life quality but also reduce the mortality rate of the patients. To realize this, our platform of choice is a silicon photonics one, where the mass fabrication techniques, developed for microelectronics, brings scaling arguments to the table. Making one sensor is expensive, but when making millions, the price per chip drops. Although the passive components and detectors show great specs, integrated light sources prove to be problematic. To circumvent this problem, we have introduced III-V material to the ecosystem in various manners to make broadband LEDs.

More specifically, in chapter 2 we study how to expand the bandwidth of the light source past the usual one. We investigated three techniques: quantum well intermixing, multiple die bonding and transfer printing. The former two were performed in the cleanroom of University of California Santa Barbara, in close collaboration with the Optoelectronics Research Group of John Bowers. The latter was developed in the Ghent University facilities.

In QWI, the band gap is locally modified by atom disordering. After an ion implantation and thermal anneal, atoms of the well and barrier diffuse to change the quantum well shape. We used experiments and simulations to identify the maximal achievable shift, which was around 120 nm in the blue direction. For this a Schrödinger's equation solver was created as described in appendix A. Test

structures were also fabricated to assess the material quality before and after intermixing.

The most obvious manner of creating multiple band gaps on one silicon chip is by bonding more than one III-V die. The problems this brings and how to solve them were investigated. Two different dies were bonded close together on an SOI chip. More than two would be feasible, but not desirable as the footprint of the layout would become too large. By combining multiple die bonding with quantum well intermixing, 4 different active band gaps were nonetheless achieved on one silicon chip. We have shown multiple die bonding to work equally well for molecular and BCB bonding.

A third technique was transfer printing. While the other two were already shown before this work (although not combined), transfer printing is completely new to integrated optics. It allows us to transfer a small sheet of III-V material (referred to as coupon) of only microns by microns and therefore efficient use of the expensive III-V material is possible. Next to this, it is straight-forward to transfer two different epitaxial layer stacks to an SOI chip by using two source wafers. It is also possible to transfer complete devices rather than material. Transfer printing relies on the kinetic response of an elastomer PDMS. When moving the PDMS stamp rapidly the adhesion to the coupon is greater than when moving slowly. This allows for a quick pick-up and a slow print. To create the right conditions, a release process was established. One of the key steps is the release etch, where a 40 μm wide coupon is etched horizontally while only nanometers vertically can be tolerated. The quality of the bond was investigated using FIB and optimized.

In chapters 3 and 4 devices were fabricated with the techniques mentioned above. The SLEDs of chapter 3 comprise of four different band gaps with a spectral spacing of 80 nm through the combination of quantum well intermixing and aligned multiple die bonding. A resulting 300 nm of 3 dB bandwidth was achieved with a serial design, where the red wavelengths travel through all the other sections. We go through the entire research cycle of these devices: After the initial idea, simulations were performed to identify the best designs. For the fabrication, the short loop experiments of chapter 2 were used to establish the correct process flow. It is therefore also obvious that these devices were realized at UCSB. Finally the devices and test structures were characterized and evaluated.

Where the SLEDs of chapter 3 were challenging mostly in fabrication and more traditional in design, the idea was to completely reverse this in chapter 4. By increasing the refractive index contrast of the active region with its surrounding, we pushed the spontaneous emission coupling efficiency from 0.1% to above 10%. This was possible by recognizing that an electrically pumped LED is not necessarily more efficient than a simple pump laser coupled to an optically pumped LED, which transfers the laser light to broadband emission. By removing the

layers needed for electrical injection, we opened the design freedom to push for high efficiency. We thoroughly looked at the feasibility of this: simulations were done regarding the spontaneous emission coupling factor, pump confinement factor, thermal design and taper design. All was geared towards simple, efficient devices. A first generation of devices showed the proof of concept, but also some problems. In an optimization run we looked at taper losses and pump absorption, non-radiative recombination through several ways, tunneling of carriers and sidewall recombination (with several mitigation strategies). Ultimately, InAsP based active material was found to improve the performance considerably.

In the conclusive chapter 5, the results of this works are summarized and an outlook to the future is provided.

1.5 List of publications

The results obtained within this work have been published in various papers and presented at various conferences. The research also gave rise to one patent.

1.5.1 Patents

- G. Roelkens, R. Baets, A. De Groote, A. Subramanian, P. Cardile, *On-chip broadband light source*

1.5.2 Publications in international journals

- A. De Groote, J.D. Peters, M.L. Davenport, M. Heck, R. Baets, G. Roelkens, J. Bowers, "Heterogeneously integrated III-V on silicon multi-bandgap superluminescent light emitting diode with 290nm optical bandwidth", *Optics Letters*, 39 (16), p.4784-4787 (2014)

- A. Subramanian, E.M.P. Ryckeboer, A. Dhakal, F. Peyskens, A. Malik, B. Kuyken, H. Zhao, S. Pathak, A. Ruocco, A. De Groote, P.C. Wuytens, D. Martens, F. Leo, W. Xie, U.D. Dave, M. Muneeb, P. Van Dorpe, J. Van Campenhout, W. Bogaerts, P. Bienstman, N. Le Thomas, D. Van Thourhout, Z. Hens, G. Roelkens, R. Baets, "Silicon and silicon nitride photonics circuits for spectroscopic sensing on-a-chip", *Photonics Research*, 5 (3), p.B47 (2015)

- G. Roelkens, A. Abbasi, P. Cardile, U.D. Dave, A. De Groote, Y. De Koninck, S. Dhoore, X. Fu, A. Gassenq, N. Hattasan, Q. Huang, S. Kumari, S. Keyvaninia, B. Kuyken, L. Li, P. Mechet, M. Muneeb, D. Sanshez, H. Shao, T. Spuesens, A. Subramanian, S. Uvin, M. Tassaert, K. Van Gasse, J. Verbist,

R. Wang, Z. Wang, J. Zhang, J. Van Campenhout, Y. Xin, J. Bauwelinck, G. Morthier, R. Baets, D. Van Thourhout, "III-V-on-silicon photonic devices for optical communication and sensing", *Photonics* (invited), 2 (3), p.969-1004 (2015)

- A. De Groote, P. Cardile, A. Subramanian, A. Fecioru, C. Bower, D. Delbeke, R. Baets, G. Roelkens, "Transfer-printing-based integration of single-mode waveguide-coupled III-V-on-silicon broadband light emitters", *Optics Express*, 24 (13), p.13754-13762 (2016)

- J. Zhang, A. De Groote, A. Abbasi, R. Loi, J. OĆallaghan, B. Corbett, A. Jose Trindade, C.A. Bowers, G. Roelkens, "A silicon photonics fiber-to-the-home transceiver array based on transfer-printing-based integration of III-V photodetectors, *Optics Express*, Accepted for publication

1.5.3 Publications in international conferences

- A. De Groote, J.D. Peters, M.L. Davenport, M. Heck, R. Baets, G. Roelkens, J. Bowers, "Broadband III-V on silicon hybrid superluminescent LEDs by quantum well intermixing and multiple die bonding", *IEEE Photonics Conference 2014*, United States, p.260-261 (2014)

- A. De Groote, P. Cardile, A. Subramanian, M. Tassaert, D. Delbeke, R. Baets, G. Roelkens, "A waveguide coupled LED on SOI by heterogeneous integration of InP-based membranes", *12th International Conference on GFP*, Canada, p.WG2 (2015)

- A. De Groote, P. Cardile, A. Subramanian, D. Delbeke, R. Baets, G. Roelkens, "A novel approach to an efficient LED on SOI", *IEEE Photonics Society Benelux*, Belgium, p.103-106 (2016)

- P. Cardile, A. De Groote, A. Subramanian, D. Delbeke, R. Baets, G. Roelkens, "Optically pumped broadband LED emission coupled to SOI waveguide", EMRS 2016, France, p.M.VI.2 (2016)

- A. De Groote, P. Cardile, A. Subramanian, A.M. Fecioru, C. Bower, D. Delbeke, R. Baets, G. Roelkens, "Integration of a III-V light emitter on a silicon photonic IC through transfer printing", *13th International Conference on GFP*, China, p.166-167 (2016)

- J. Zhang, A. Abbasi, A. De Groote, R. Loi, J. O'Callaghan, B. Corbett, A. Jose Trindade, C.A. Bower, G. Roelkens, "Silicon photonic transceiver array based on the transfer printing of III-V O-band photodetectors", *European Conference on Integrated Optics (ECIO)*, Netherlands, 2016

- J. Juvert, T. Cassese, S. Uvin, A. De Groote, B. Snyder, P. De Heyn, P. Verheyen, A.J. Trindade, C. Bower, M. Romagnoli, G. Roelkens, D. Van Thourhout, "Integration of III-V light sources on a silicon photonics circuit by transfer printing", *Group IV Photonics (GFP)*, Germany, 2017 (Accepted for oral presentation)

References

[1] International Diabetes Federation. *IDF Diabetes Atlas - 7th Edition*, March 2017.

[2] Richard Soref. *The past, present, and future of silicon photonics*. IEEE Journal of selected topics in quantum electronics, 12(6):1678–1687, 2006.

[3] Inc. University Wafer. *University Wafer, your leading supplier of semiconductor wafers*, April 2017.

[4] Peter De Dobbelaere, Ali Ayazi, Yuemeng Chi, Anders Dahl, Scott Denton, Steffen Gloeckner, Kam-Yan Hon, Steve Hovey, Yi Liang, Michael Mack, et al. *Packaging of silicon photonics systems*. In Optical Fiber Communications Conference and Exhibition (OFC), 2014, pages 1–3. IEEE, 2014.

[5] Bradley William Snyder. *Hybrid integration and packaging of grating-coupled silicon photonics*. PhD thesis, University College Cork, 2013.

[6] Günther Roelkens, Liu Liu, Di Liang, Richard Jones, Alexander Fang, Brian Koch, and John Bowers. *III-V/silicon photonics for on-chip and inter-chip optical interconnects*. Laser & Photonics Reviews, 4(6):751–779, 2010.

[7] E Menard, KJ Lee, D-Y Khang, RG Nuzzo, and JA Rogers. *A printable form of silicon for high performance thin film transistors on plastic substrates*. Applied Physics Letters, 84(26):5398–5400, 2004.

[8] Frank Niklaus, Göran Stemme, J-Q Lu, and RJ Gutmann. *Adhesive wafer bonding*. Journal of applied physics, 99(3):2, 2006.

[9] John W. Jewett and Raymond A. Serway. *Physics for scientists and engineers with modern physics*, chapter Molecules and Solids, page 1260. Thomson Learning Inc., 7 edition, 2008.

[10] Cleveland.com. *University of Akron's research into geckos' natural stickiness may pay off in companies and products.*, April 2017.

[11] University of Massachusetts Amherst. *Geckskin*, April 2017.

[12] Bing-Ruey Wu, Ching-Fuh Lin, Lih-Wen Laih, and Tien-Tsorng Shih. *Extremely broadband InGaAsP/InP superluminescent diodes*. Electronics Letters, 36(25):2093–2095, 2000.

[13] Jingcong Wang, Michael J Hamp, and Daniel T Cassidy. *Design considerations for asymmetric multiple quantum well broad spectral width superluminescent diodes*. Quantum Electronics, IEEE Journal of, 44(12):1256–1262, 2008.

[14] Ching-Fuh Lin, Yi-Shin Su, Chao-Hsin Wu, and Gagik S Shmavonyan. *Influence of separate confinement heterostructure on emission bandwidth of InGaAsP superluminescent diodes/semiconductor optical amplifiers with nonidentical multiple quantum wells*. Photonics Technology Letters, IEEE, 16(6):1441–1443, 2004.

[15] Takeshi Yamatoya, Shota Mori, Fumio Koyama, and Kenichi Iga. *High power GaInAsP/InP strained quantum well superluminescent diode with tapered active region*. Japanese journal of applied physics, 38:5121, 1999.

[16] JH Song, SH Cho, IK Han, Y Hu, PJS Heim, FG Johnson, DR Stone, and M Dagenais. *High-power broad-band superluminescent diode with low spectral modulation at 1.5-/spl mu/m wavelength*. Photonics Technology Letters, IEEE, 12(7):783–785, 2000.

[17] Ken Morito, Shinsuke Tanaka, Shuich Tomabechi, and Akito Kuramata. *A broad-band MQW semiconductor optical amplifier with high saturation output power and low noise figure*. Photonics Technology Letters, IEEE, 17(5):974–976, 2005.

[18] A Kovsh, A Gubenko, I Krestnikov, D Livshits, S Mikhrin, J Weimert, L West, G Wojcik, D Yin, C Bornholdt, et al. *Quantum dot comb-laser as efficient light source for silicon photonics*. In Photonics Europe, pages 69960V–69960V. International Society for Optics and Photonics, 2008.

[19] Sarah Uvin, Utsav D Dave, Bart Kuyken, Shankar Selvaraja, Francois Leo, and Gunther Roelkens. *Mid-infrared to telecom-band stable supercontinuum generation in hydrogenated amorphous silicon waveguides*. In Photonics Conference (IPC), 2013 IEEE, pages 380–381. IEEE, 2013.

[20] Bart Kuyken, Xiaoping Liu, M Richard Jr, Roel Baets, Günther Roelkens, and William MJ Green. *Mid-infrared to telecom-band supercontinuum generation in highly nonlinear silicon-on-insulator wire waveguides*. Optics Express, 19(21):20172–20181, 2011.

2

Technologies for multi-bandgap light source integration on silicon

Silicon photonics is great when working with passive components. The high refractive index contrast allows for very small structures, while the excellent processing tools provide for high yield and reproducibility. The large silicon-on-insulator wafers offer high volume and the prospect for low cost per chip. Silicon is however not suited for light generation and amplification. Efficiently creating light is impossible due to the indirect band gap. This requires the integration of III-V semiconductors on the silicon photonic platform. In several applications there is a need to integrate III-V materials with different bandgaps on a single chip, either for the implementation of multiple functions or to generate light over a broad wavelength range. In this chapter, we will move through the different options we have to introduce III-V structures with more than one band gap on a silicon chip. More in particular, we discuss quantum well intermixing, multiple die bonding and transfer printing in detail. We start with an overview of our options.

Foremost, we want to thank the entire research group headed by John Bowers. Integrating multiple dies by means of molecular bonding, as well as the intermixing process, was developed at the UCSB cleanroom with the knowledgeable help of Siddarth Jain and Jon Peters among others.

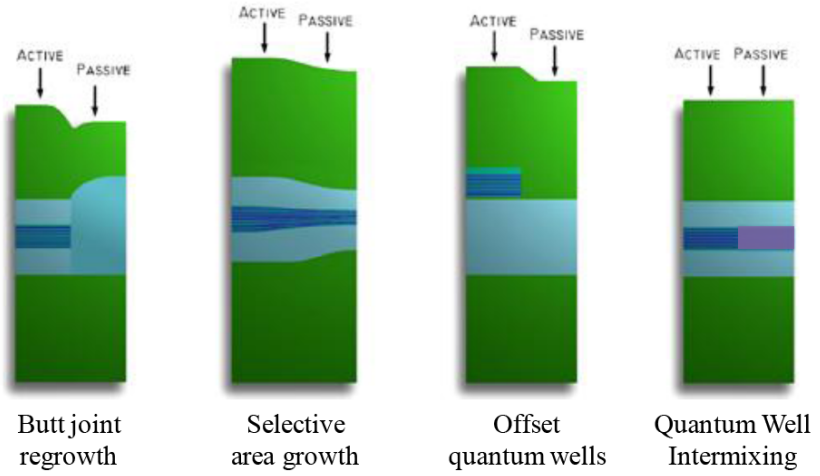


Figure 2.1: Techniques for creating multiple band gaps on the conventional InP platform. [1]

2.1 Overview of different techniques to create multiple band gaps on silicon

In a photonic circuit usually multiple band gaps are necessary. A complete circuit at least consists of a light source, a waveguide and a detector, meaning two different band gaps are necessary. When adding a modulator or a second light source, three or more band gaps come into play. Silicon has one band gap. It is indirect and corresponds to a bandgap wavelength of 1.1 μm. Heterogeneous integration has brought a second, active band gap into the picture. A III-V material sheet is bonded onto the silicon passive circuit and light sources, modulators or detectors can be created now. So far, most demonstrations of heterogeneous photonic integrated circuits have been limited by these two band gaps. As the bonding technology has matured, we are now looking to introduce the much-needed third-and-more band gap.

Multiple band gaps are common on the conventional InP platform. Techniques such as butt-joint regrowth and selective area growth (SAG) are widely used, while offset quantum wells and quantum well intermixing are also attractive options. These techniques are shown in figure 2.1. Butt-joint regrowth, SAG and offset quantum wells require some form of regrowth. During the butt-joint regrowth process, some layers are etched away at certain locations and other layers are grown in these places. For SAG, the grown material depends on the etched windows in which the new material is grown. Very high quality material can be created in this manner and regrowth can be considered the workhorse of the InP platform.

To compare these approaches and to understand the feasibility of such structures on a silicon platform, it is important to recognize the constraints of the heterogeneous integration of III-V and Si. Conventional p-i-n InP laser epitaxial design is grown p-side up, and are vertically quite symmetrical. This is no longer the case when integrating them with Si. Because the stack is flipped upside down after bonding and because we prefer the p-side to face up on the final device (due to lower hole mobility and highly doped InGaAs contact layers), the growth has to be done n-i-p. Since Zn diffusion is a constant worry, this is not at all preferred by growers. On top of that, the stack is now highly asymmetrical, since the n-side is typically very thin. This allows for the quantum wells to be very close to the silicon waveguide, which allows for efficient optical coupling between the III-V waveguide layer and the silicon layer. The regrowth techniques result in a non-uniform top surface, preventing a good bond interface. Regrowth after bonding can be done [2], but only under very strict conditions. The regrown material is not of the quality needed for active material and, due to the mismatch in thermal expansion coefficient between III-V and Si, there is a maximum III-V thickness of ± 300 nm. Above this III-V thickness, the material relaxes resulting in a lot of defects and therefore one can only use lateral P-N junctions, severely limiting the designs. Direct III-V growth on silicon followed by regrowth would solve these issues. Direct growth is still in its infancy however.

In quantum well intermixing, the band gap can be changed after the bonding process. Atom disordering causes the shape of the quantum well energy profile to change and therefore a blueshift in the bandgap wavelength is observed. The spatial resolution of this technique can be very good, as the location of the shifted band gap is lithographically defined. As no regrowth is needed, the surface of the III-V die remains flat and we can expect excellent bond interfaces. Note that the bonding needs to be aligned, as the created band gaps need to match up with a certain design on the silicon chip. Because the QWs are in essence damaged (be it in a controlled manner), there is a risk of performance degradation. However, if done properly, this is not necessarily true.

Heterogeneous integration of course also allows for bonding of multiple dies. Several epitaxial layer stacks are separately grown and thus complete design freedom can be utilized. The layer stacks are then bonded in one bonding step, and devices can be made on each of them. The main limit of this method can be found in the size of the bonded dies. III-V dimensions of a few millimeter by a few millimeter are common for bonding, meaning that a circuit of three III-V band gaps will be at least one centimeter long. The footprint grows rapidly due to this limitation. On top of that, part of the silicon circuit will be exposed during the post-bond fabrication process and precautions need to be taken not to damage these structures.

A last technique we want to bring up, is transfer printing. In this technique

small coupons of III-V material (typical dimensions are tens of micrometers), are transferred to the silicon PIC. This technique has all the advantages of multiple die bonding, but solves the spatial resolution issue. The small coupons can be printed where needed. They can be scattered throughout the wafer without impacting other III-V devices scattered throughout the wafer. In order to do transfer printing, the coupons need be released from the III-V growth substrate. The undercut of these structures is not trivial, but can be managed.

All techniques discussed here are summed up in table 2.1, listing their characteristics. In this chapter we will discuss the latter three techniques in detail. The process developed for each of them is outlined along with the problems encountered and how we solved them.

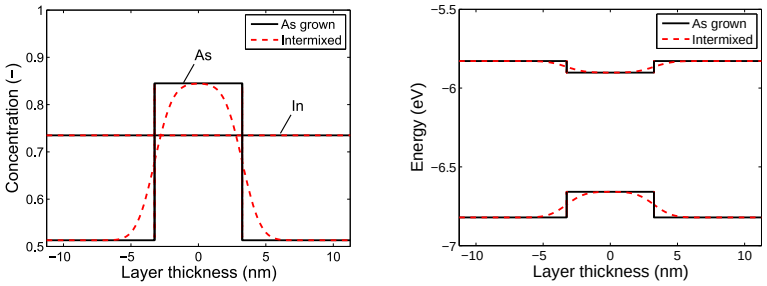
2.2 Quantum well intermixing

Using quantum well intermixing, it is possible to shift the band gap of a set of quantum wells to shorter wavelengths. This can be done in very specific locations, making high density integration an option. After the intermixing, the III-V die is bonded to a SOI waveguide circuit in an aligned manner, bringing the different band gaps to silicon. Note that intermixing after bonding is impossible due to the limited thermal budget of post-bond samples. In the following section, we will go over the basic principle and techniques to do intermixing. Furthermore, we outline our implementation and show some problems and solutions.

2.2.1 Basic principle and techniques

Quantum well intermixing is an atom disordering technique. Using thermal energy, the size and composition of the quantum wells and barriers are changed. Figure 2.2 shows an example of a quantum well design based on InGaAsP material (lattice matched to InP). As can be seen, the indium concentration is kept constant, while the arsenic concentration is varied to form a quantum well. The supplied thermal energy will allow atoms to come free of the lattice and effectively swap places with a corresponding one, in this case As will swap with P. The dashed curve shows the resulting profiles, where As atoms have diffused into the barriers. This alters the quantum well structure and thereby change the quantum levels.

Above we have identified thermal energy as the driving force of quantum well intermixing. Different catalysts are being used though, resulting in different techniques, although in essence they are all the same. Most common ones are impurity induced disordering (IID) [3], impurity-free vacancy disordering (IFVD) [4], photo-absorption induced disordering (PAID) [5] and implantation enhanced disordering (IED) [6].



(a) Concentration of different atoms *(b) Energy levels of the quantum well*

Figure 2.2: Illustration of quantum well intermixing of a quantum well

Technique	Butt-joint	SAG	Offset QWs	QWI	Multiple die bonding	Transfer printing
2+ band gaps?	Yes	Yes	No	Yes	Yes	Yes
Constraint for 2+ band gaps?	Multiple regrowths	None	Not possible	None	Footprint grows as bonded dies are mm-sized	None
# of regrowths	# of band gaps + 1	1	1	None	None	None
Planar bonding surface?	No	No	No	Yes	Yes	Yes
Considerations	Reflections, requirement of several regrowths	Epitaxial architecture fixed for all components	One QW design with bulk section	Potential performance degradation due to change in QW shape	Limited resolution	Structures have to be released
Advantages	Design flexibility, individual optimization of QWs	2+ QW band gaps in 1 regrowth	Simple fabrication technique	Excellent spatial resolution, post-bond process similar to single-die bonding process	Complete freedom in QW design	Complete freedom in QW design, good spatial resolution

Table 2.1: Comparison of different techniques creating multiple band gaps on a silicon chip.

IID depends on impurities/dopants to change the Fermi level, and hence the equilibrium dopant concentration is changed. This in turn promotes diffusion, leading to intermixing. The dopants are often electrically active though and can cause changes in the material properties. Higher optical loss, free-carrier absorption and high transparency currents in lasers are often observed.

In IVFD intermixing is done without introducing impurities in the lattice, as the name suggests. However, the defects are created by the use of capping layers. These interface defects diffuse into the quantum wells during the anneal. The thickness of the capping layer relates to the amount of defects created, controlling the extent of the intermixing. On InP common capping layers are SiO_2 or SiN_x, either sputtered or deposited by PECVD. Typically very high temperatures are needed for this technique ($>$800°C). It is believed that IVFD working at lower temperatures is actually IID, where Zn ions are diffused as a catalyst.

A bright laser is used in the PAID technique, also called laser induced disordering. By tuning the laser to the quantum wells absorption, heat can be generated in a very small spot. The lattice will expand and defects are created. As with all intermixing techniques, these defects then act as a catalyst in the following thermal anneal. This technique is dominated by the material thermal properties, more in particular lateral heat flow limiting the spatial resolution.

Finally, in IED the defects are caused by an ion implantation in a buffer layer away from the quantum wells. The thermal anneal then drives the defects into the quantum wells. As lithography can be used to mask part of the buffer layer, the resolution is determined by the lateral diffusion. In this work, we employed IED.

2.2.2 Our implementation

As stated, we opted to use IED. Two layer stacks were intermixed: one where the emission peak was changed from 1380nm to 1300nm and one where 1540nm was changed to 1460nm. The two epitaxial layer stacks are laid out in tables 2.2 and 2.3, which will first be intermixed and then bonded onto SOI. Both stacks were grown n-side up, in order to have a p-side up device after bonding. There is a sacrificial InP buffer layer and InGaAsP etch stop in which the implantation of the buffer layer will occur. The InGaAsP quantum wells do not contain aluminum atoms, as this would prevent intermixing as shown in [1].

Below we discuss the complete pre-bond process, first focusing on the intermixing step itself and then sketching the bigger picture.

#	Layer	Thickness	Doping	PL
1	Sacrificial InP buffer	450 nm	i	N/A
2	$In_{0.8423}GaAs_{0.3434}P$ etch stop	40 nm	i	1.1 μm
3	n-InP bottom device layer	17.5 nm	(n) 3×10^{18}cm^{-3}	N/A
4	$In_{0.85}GaAs_{0.327}P$/InP superlattice	(2x) 7.5/7.5 nm	(n) 3×10^{18}cm^{-3}	1.1 μm
5	n-InP contact	110 nm	(n) 3×10^{18}cm^{-3}	N/A
6	$In_{0.741}GaAs_{0.805}P$ well / $In_{0.9611}GaAs_{0.0025}P$ barrier	7x4.0 / 8x8.0 nm	i	1.38 μm
7	$In_{0.5236}AlAs$ e^- blocking layer	10 nm	i	N/A
8	p-$In_{0.5305}Al_{0.4055}Ga_{0.064}As$	250 nm	(p) 10^{17}cm^{-3}	N/A
9	p-InP	1.5 μm	(p) 10^{18}cm^{-3}	N/A
10	p-$In_{0.532}GaAs$	100 nm	(p) 10^{19}cm^{-3}	1.65 μm
11	p-InP substrate	350 μm	(p) 2×10^{18}cm^{-3}	N/A

Table 2.2: Epitaxial layer stack with PL peak at 1380nm

#	Layer	Thickness	Doping	PL
1	Sacrificial InP buffer	450 nm	i	N/A
2	$In_{0.711}GaAs_{0.611}P$ etch stop	40 nm	i	1.3 μm
3	n-InP bottom device layer	17.5 nm	(n) 3×10^{18}cm^{-3}	N/A
4	$In_{0.85}GaAs_{0.327}P$/InP superlattice	(2x) 7.5/7.5 nm	(n) 3×10^{18}cm^{-3}	1.1 μm
5	n-InP contact	110 nm	(n) 3×10^{18}cm^{-3}	N/A
6	$In_{0.735}GaAs_{0.845}P$ well / $In_{0.735}GaAs_{0.513}P$ barrier	7x6.5 / 8x8.0 nm	i	1.54 μm
7	p-$In_{0.768}GaAs_{0.504}P$	250 nm	(p) 10^{17}cm^{-3}	N/A
8	p-InP	1.5 μm	(p) 10^{18}cm^{-3}	N/A
9	p-$In_{0.532}GaAs$	100 nm	(p) 10^{19}cm^{-3}	1.65 μm
10	p-InP	200 nm	(p) 2×10^{18}cm^{-3}	N/A
11	p-$In_{0.532}GaAs$	50 nm	(p) 10^{19}cm^{-3}	1.65 μm
12	p-InP substrate	350 μm	(p) 2×10^{18}cm^{-3}	N/A

Table 2.3: Epitaxial layer stack with PL peak at 1540nm

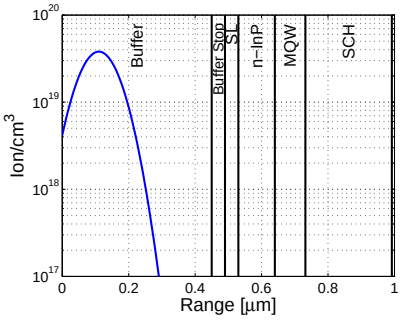


Figure 2.3: Simulation of the ion distribution after the phosphorous implant with energy 100 keV and dose $5 \times 10^{14} cm^{-2}$.

2.2.2.1 Intermixing

Simulation of implantation

Both stacks in tables 2.2 and 2.3 contain a sacrificial InP buffer layer and InGaAsP etch stop. The ions of choice were phosphorous, as these would introduce group V interstitial defects causing the group V elements to dominate the intermixing. As can be seen, we made sure the difference in As concentration between the barrier and quantum well was made as large as possible in both stacks. Figure 2.3 shows the ion distribution after the phosphorous implant (energy 100 keV and dose $5 \times 10^{14}\text{cm}^{-2}$), as simulated using *SRIM*. Note that the buffer layers are at the top, meaning that the implant energy can be very low. This minimizes the damage done to the lattice during the implantation. It is clear that the phosphorous ions are still far from the active region before the anneal and can therefore not do any damage. As the buffer layer is etched away before bonding, the defects in this layer are irrelevant. This is investigated in more detail in chapter 2 of [1].

Simulation of wavelength shift

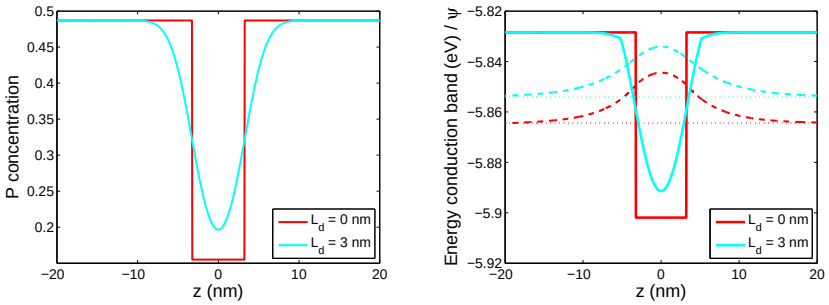
Diffusion is governed by Fick's law:

$$j = -D\nabla C \tag{2.1}$$

where j is the diffusion flux, D is the diffusion coefficient and C is the concentration. To estimate the final distribution of the atoms, we make use of the continuity equation:

$$\frac{\partial C}{\partial t} + \nabla \cdot j = G - R \tag{2.2}$$

with G and R the generation and recombination. These can be set zero for our case.If we now assume a delta peak to be the distribution at t=0, the solution to



(a) Concentration of phosphorous atoms *(b) Energy diagram*

Figure 2.4: Simulation of the effect of intermixing on the quantum well

this problem is a gaussian distribution in space at any given time $t > 0$:

$$C(z,0) = S\delta(z) \tag{2.3}$$

$$C(z,t) = \frac{S}{\sqrt{4\pi Dt}} \exp\left(-\frac{z^2}{4Dt}\right) = \frac{S}{\sqrt{2\pi}L_d} \exp\left(-\frac{z^2}{2L_d^2}\right) \tag{2.4}$$

where S is the atom dose, and L_d is the diffusion length. Note that the diffusion length is dependent on the defects present to assist the disordering.

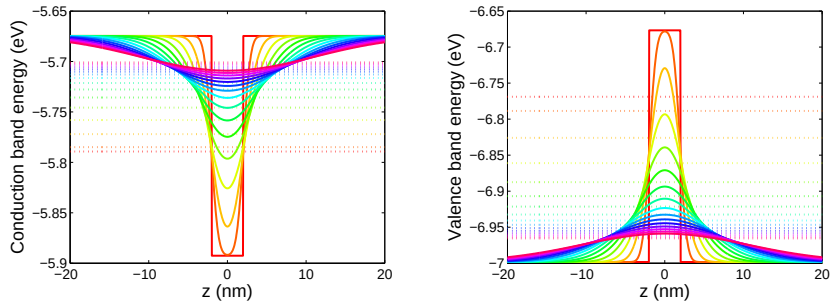
Because of this gaussian shape, we can devise a mathematical model for the atom diffusion using the error function. One could describe the concentration of phosphorous atoms (these are the P atoms as grown in the lattice, not to be confused with the implanted P ions which introduce the defects) after disordering as:

$$y = y_b + \frac{y_w - y_b}{2}\left[\text{erf}\left(\frac{\frac{w}{2} - z}{L_d}\right) + \text{erf}\left(\frac{\frac{w}{2} + z}{L_d}\right)\right] \tag{2.5}$$

$$\text{with} \quad \text{erf}(x) = \frac{1}{\sqrt{\pi}} \int_x^x e^{-t^2} \mathrm{d}t \tag{2.6}$$

where y_w and y_b is the concentration of phosphor in the quantum well and barrier before diffusion respectively, while w is the width of the quantum well. After calculating the new concentrations, the corresponding energy diagram can be drawn and the quantum states are to be calculated. Schrödinger's equations are solved using the transfer matrix method described in appendix A. In this method, the energy diagram is sliced up in space and the potential well is described by a piecewise function where all the sections are of constant potential. This allows us to solve Schrödinger's equation for an arbitrary potential, in particular for the quantum well after intermixing.

Figure 2.4 shows the intermixing of the layer stack of table 2.3. Equation 2.5 dictates the graph of figure 2.4a. The perfectly square quantum well gets rounded



(a) Conduction band *(b) Valence band*

Figure 2.5: Simulation for many different diffusion lengths, varying from 0 nm to 15 nm (steps of 1 nm).

and is filled towards the bottom. The concentration of phosphorous atoms was then translated to an energy diagram. Since changing the group V concentration also changes the strain of the wells/barriers, this was also taken into account. The conduction band is shown in figure 2.4b, which also shows the eigen energy in the dotted line. It can be seen that due to the disordering, the quantum state moves to a higher energy. As a result, a blue shift is observed. The wave function before and after intermixing is also shown, in dashed lines. The wave function becomes broader than before.

We can now simulate the quantum well intermixing for both the conduction and valence band (i.e. the heavy holes, which are made lighter by the compressive strain) for many different diffusion lengths, see figure 2.5. The wavelength shift as a function of diffusion length is plotted in figure 2.6 for the two layer stacks. The curves clearly show an S-shape. There is a slow start as long as the diffusion length is less than 1 nm. One can see in figure 2.5 that when the diffusion length is 1 nm, the bottom energy of the quantum well will start to move to higher energies, i.e. the minimum energy of the potential is no longer the same as the as-grown case. The tail at longer diffusion lengths indicates the quantum well is nearly smoothed out. There is only a very shallow and wide well left, close to the energy levels of the barriers. Most of the wavelength shift occurs between 1 nm and 6 nm. Next to this, the layer stack of table 2.2 with the as-grown PL peak at 1380 nm can shift much further than the one of table 2.3. This is obviously because the wells are much deeper. Given the difference between quantum well and barrier energy levels, there is much more room to bring the eigen state up in energy. There is more however as the shift is also much more sudden. E.g. for a diffusion length of 4 nm, the shift is 180 nm for the shorter wavelengths and only 82 nm for the longer ones. The reason for this can be found in the width of the quantum wells. A narrow quantum well, such as the 4 nm wide ones in table 2.2 will be subject to

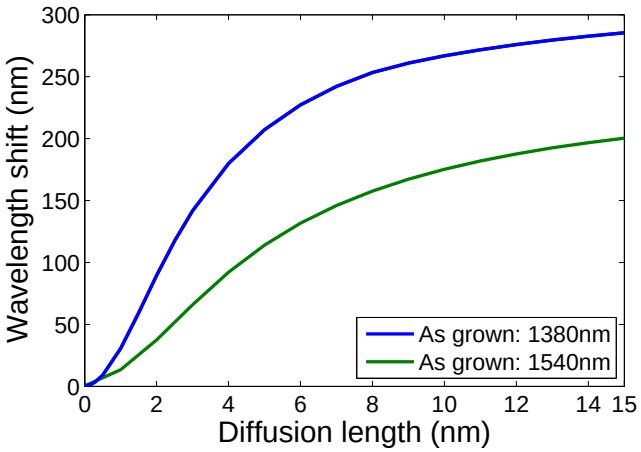


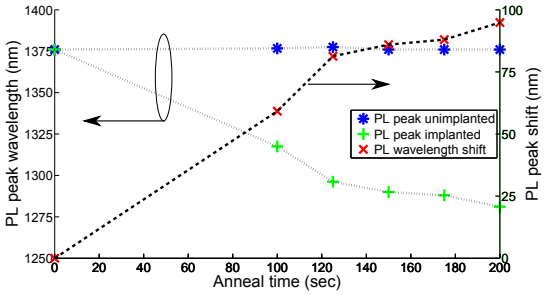
Figure 2.6: Wavelength shift as a function of diffusion length for the two layer stacks.

a greater effect when the shape slightly changes.

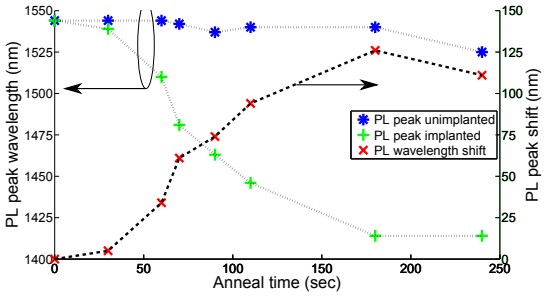
Test structures: blueshift as a function of anneal time

The subsequent rapid thermal anneal will drive the interstitial defects into the quantum wells and cause the disordering. Depending on the applied thermal energy, the diffusion length will be shorter or longer and the quantum states will shift correspondingly. Figure 2.7 shows the PL peak wavelength shift as a function of annealing time. The phosphorous ions were implanted with an energy of 100 keV and a dose of $5 \times 10^{14}\text{cm}^{-2}$. The buffer layer and its etch stop have been removed prior to the PL measurements in order to have a better signal. The stack emitting at 1380 nm is very stable. It can be seen in figure 2.7a that the unimplanted regions do not shift even at long annealing times. This was also observed by [1], who worked on the same material. The lack of parasitic shift is believed to be due a low defect density in this layerstack. When a lot of thermal energy has been added, the native defects will act as catalysts for the disordering. Furthermore, the implanted region shifts gradually towards shorter wavelengths, making it quite easy to hit the desired shift. Figure 2.7b shows a different picture however. It was observed in [7] that a 700 °C anneal would saturate at a shift of 80 nm. Since this is the value we target in this work, we increased the temperature to 725 °C. At longer anneal times, we here see a clear parasitic shift of the unimplanted regions, while the implanted ones have a sharp S-shaped wavelength shift curve. Hitting a net shift of 80 nm is more challenging, but still feasible if controlled well.

In the previous section we established a relation between wavelength shift and

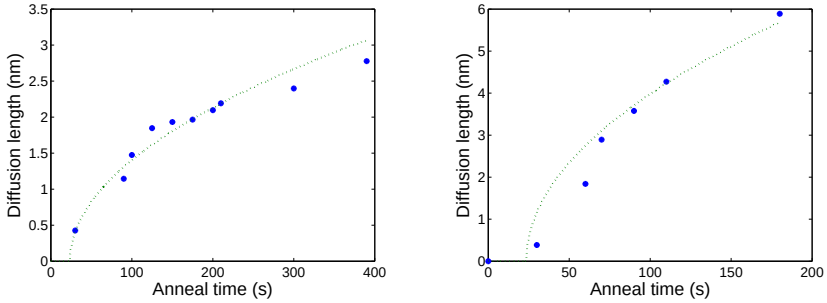


(a) O-band layer stack of table 2.2. Annealing temperature was 700° C.



(b) C-band layer stack of table 2.3. Annealing temperature was 725° C.

Figure 2.7: PL peak wavelength shift and net shift as a function of anneal time.



(a) Layer stack of table 2.2. Annealing temperature was 700°C.

(b) Layer stack of table 2.3. Annealing temperature was 725°C.

Figure 2.8: Diffusion length as a function of anneal time.

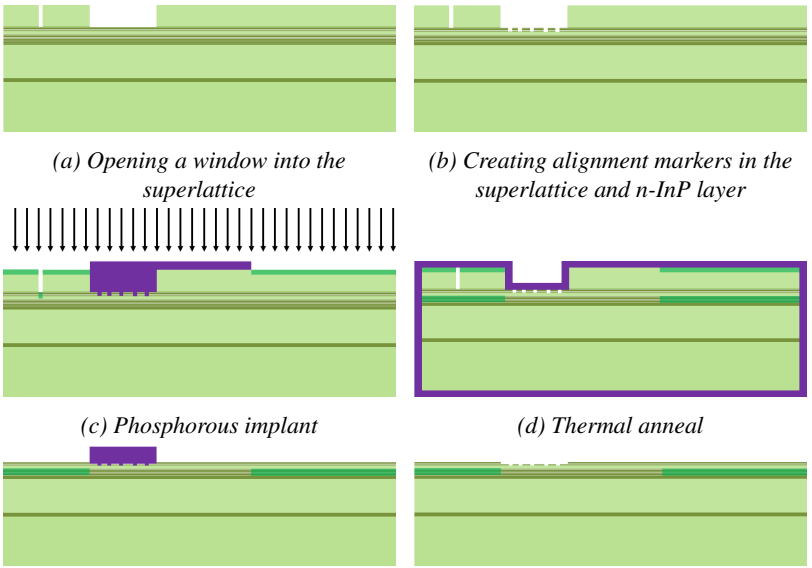
diffusion length. We can therefore rework figure 2.7 to find the diffusion length as a function of time, shown in figure 2.8. First of all, it is clear once more that the O-band stack can do the desired shift with a lot less diffusion, due to the narrower quantum wells. To achieve the shift of $\pm$ 100 nm, a diffusion length of only 3 nm is needed, contrasting with the 6 nm needed for the C-band stack. As shown in equation 2.4, the diffusion length, which determines the spatial extent of the gaussian concentration profile, has a square root dependence on time:

$$L_d = \sqrt{2Dt} \tag{2.7}$$

The dotted line in figure 2.8 show a square root fit. There is a dead time of approximately 20 seconds for both curves, which is probably due to the rapid thermal anneal furnace itself. It needs 20 seconds to go from room temperature to the desired 700 / 725 °C. From this fit we can also retrieve the diffusion coefficients D. For the O-band stack, we extracted $1.2 \times 10^{-16}\text{cm}^2/\text{s}$, for the C-band stack $2.0 \times 10^{-15}\text{cm}^2/\text{s}$. A difference this huge cannot be attributed to the temperature difference alone (700 °C and 725 °C respectively), but must also be a consequence of the growth. The O-band stack is extremely stable, as was also noticed by Sidharth Jain, who worked on the same material [1]. The stability of the O-band stack explain the absence of parasitic diffusion of the non-implanted regions, even at long annealing times.

2.2.2.2 Pre-bond fabrication process

Creating different band gaps in one epitaxial layer stack does not serve our purpose if it is not aligned to something and suitable for bonding. Figure 2.9 shows the processing steps done before integration on a silicon photonic IC. After this process, we end with a planar surface with only alignment marks visible (but they



(a) Opening a window into the superlattice

(b) Creating alignment markers in the superlattice and n-InP layer

(c) Phosphorous implant

(d) Thermal anneal

(e) Planarization, protecting alignment marks

(f) Intermixed epitaxial stack, ready for bonding

Figure 2.9: Processing steps of quantum well intermixing

do not protrude from the surface).

(a) **Window definition:** The process starts by etching through the buffer layer, into the superlattice. This is to ensure the to-be-defined alignment marks lie lower than the bond interface. This means that the alignment marks themselves will not actually bond to the silicon and they will be lost during the substrate removal. However, the other option is to risk them sticking out of the bond interface, causing the bonding of a big region to fail. Clearly the former is desirable over the latter. We defined rulers in the silicon to assess the bond accuracy, which is discussed in section 2.3.3 .

(b) **Alignment markers:** Next, we etch the alignment markers into this window, which will be used to align during the flip chip bonding. Note that we make sure that all the following steps are aligned to this step, not to the window definition.

(c) **Phosphorous implant:** We use a SiN_x hard mask to protect the regions we do not want to be intermixed. As there will be no implant in these regions, there are no defects to catalyze the disordering. As discussed above, the phosphorous implant is done at 100 keV with a dose of $5 \times 10^{14}\text{cm}^{-2}$, and will therefore only affect the buffer layer for now.

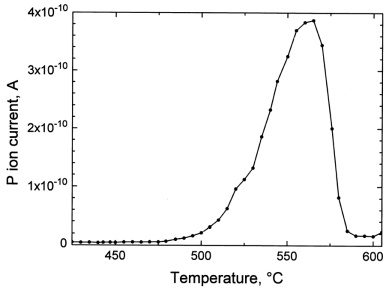


Figure 2.10: A typical P evaporation curve. The decline at higher temperature is attributed to P depletion. [8]

(d) **Thermal anneal:** The actual disordering occurs during the thermal anneal, when the defects are driven into the quantum wells. Because the samples are heated to 700/725 °C a suitable capping layer is necessary. After all, e.g. phosphorous atoms start to evaporate around 480 °C, as indicated in figure 2.10. By using a capping layer, an overpressure of phosphor can be created, preventing the evaporation of phosphor. This capping layer needs to be low-stress over the entire temperature range though, as it will otherwise crack, and thus a strain compensated SiN_x-SiO_2-SiN_x stack was deposited on all sides of the sample (including the back of the substrate). In addition the sample was place in a closed graphite crucible which was saturated with phosphorous ions before. In this manner we created a phosphorous atmosphere.

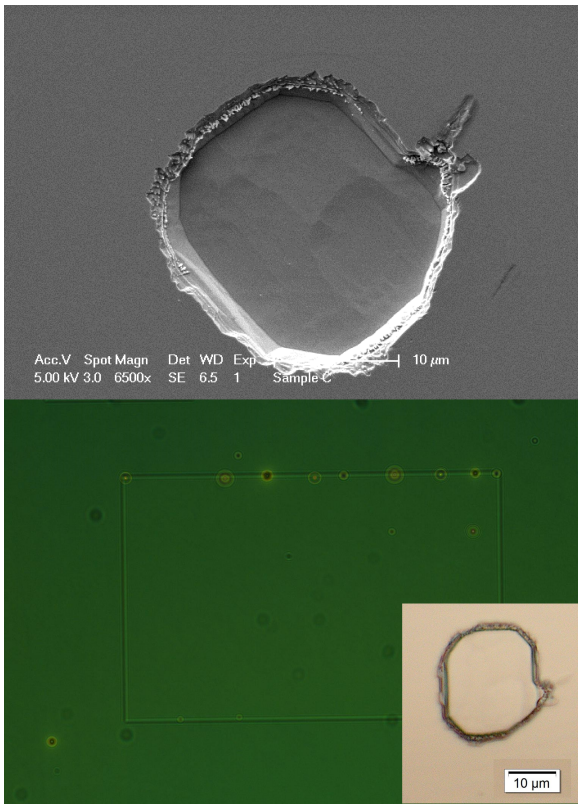
(e) **Planarization:** After the intermixing, it is time to prepare the bond surface, i.e. the die must be completely planar after this step. The buffer layer and its etch stop are etched away, but one must protect the alignment marks to prevent them from being washed away. This lithography (and all others) is done on a *GCA Autostep 200* stepper system, meaning that the area between two exposures is not illuminated as the stepper steps over it. Because we bond the entire die and not only the exposed part, we have to make sure that the sacrificial layers are also etched in this area. Therefore, it is important to use a negative resist.

The process above is both forgiving and terribly unforgiving at the same time. Since we still do bonding and a complete process to make devices afterwards, the alignment tolerances can be up to 5 μm. It also means of course the surface must remain bond-proof. Any particle deposited on the sample can be detrimental as it will cause a large bond fail if it is protruding the bond interface. In this respect, any hard mask definition is critical. As the nitride is deposited by PECVD at 250 °C, any particle is burned on and is very unlikely to ever come off. Before the thermal

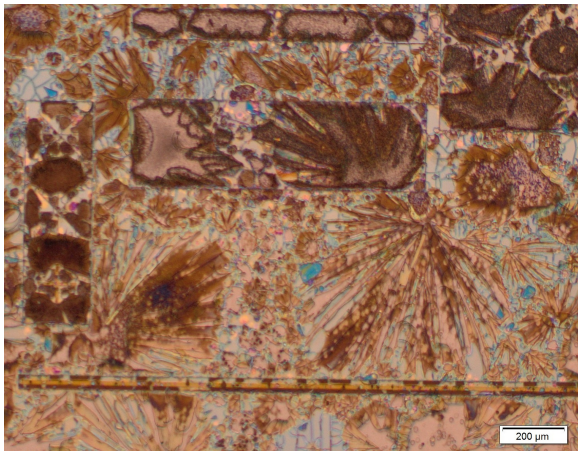
anneal, one has to be particularly careful at any step as any defect will not only be burned on but may also disturb the dielectric film.

The thermal anneal is a big milestone in the process. Any possible problem before will light up. The encapsulation layer is critical to create a phosphorous overpressure, to prevent outgassing. Even though an encapsulation layer and a graphite crucible were used, signs of phosphorous evaporation were observed, two examples of which are shown in figure 2.11. A so-called "etch pit" can be seen on the top of figure 2.11a. Here, phosphor was able to evaporate causing indium atoms to conglomerate. These are then washed away leaving these gaping holes, which are typically tens of microns in diameter. The lower picture of figure 2.11a show where these etch pits mostly occur. This picture was taken directly after the thermal anneal and thus one would not expect to see a pattern on it. However, the ion implantation causes the buffer layer to thin down by a few nanometer, which is then highlighted by the capping layer. It can be expected that this will thus be a weak spot in the capping and therefore evaporation is most likely to happen there. We also noticed that the pits tend to amass on one side of the structures, which remained a mystery. We suspected it might be due to the angle at which the implant is done in combination with the weakness of the capping layer (i.e. on one side more phosphorous atoms pile up under the weak spots), but we could never be conclusive about this. In terms of bonding, the etch pits are not a big problem. Since they have no height to it, they only affect the bonding at the exact location where they are. This means that unless a device happens to coincide with the pit, the pit will not be noticed in the final device.

Figure 2.11b shows a much more grim image of reality. The phosphorous outgassing has completely destroyed the dielectric capping layer and although most of the damage was seen to be in the to-be-removed buffer layer, other layers were also affected. Obviously bonding this sample was out of the question. It can also be noted that there is a clear difference between the implanted region, which has starlike defects, and the unimplanted regions, which have bubble-like defects. To find whether the cracking of the capping layer is due to stress in the dielectric film or the fact that the overpressure in the sample is simply too high, we did two sets of tests. In one, we used small pieces ($<$1cm^2) of our epitaxial material, which we capped and annealed as before. In the other, we used quarters of a 2 inch InP wafer to cap and anneal. The first test should indicate whether the capping layer can handle the pressure, while the second whether the stress is ok. To our surprise both tests came out fine. The only conclusion to be drawn is that is was a problem of a combination of these effects. A large InP sample can simulate the effect of lattice expansion and stress well enough, but in terms of phosphorous evaporation it is completely different since it behaves a huge phosphorous barrel, while epi has more restrictions in the vertical dimension (e.g. InGaAs and InGaAsP layers).



(a) Bondable sample with etch pits due to P evaporation



(b) Destroyed sample. The dielectric film cracked under the stress, after which P atoms were free to outgas.

Figure 2.11: Effects of phosphorous evaporation from the surface.

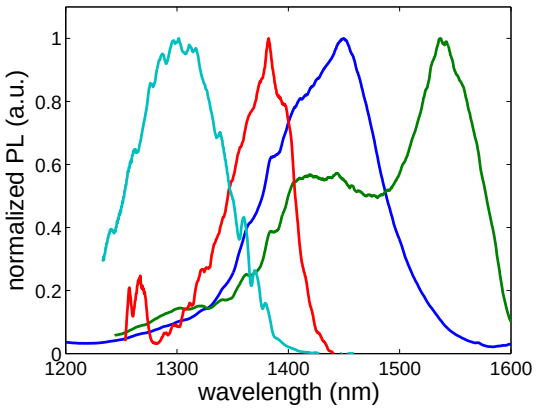
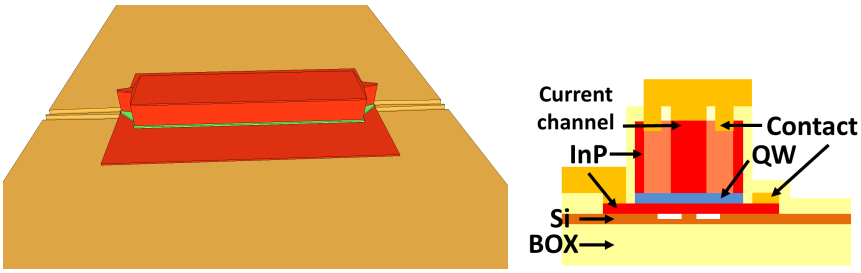


Figure 2.12: The photoluminescence spectra from the four generated band gaps transposed on one graph

2.2.3 80 nm spaced PL peaks and laser structures

The goal of our intermixing project was to create a superluminescent LED with an optical bandwidth of 300 nm, discussed in chapter 3. For this we need equally spaced band gaps, spaced approximately 80 nm. From the tests of figure 2.7, we deduced we should anneal one layer stack at 700 °C for 125 seconds to achieve a shift from 1377 nm to 1296 nm, while the other stack should be annealed at 725 °C for 90 seconds to shift from 1537 nm to 1463 nm. The four photoluminescence (PL) peaks are shown in figure 2.12. One can see that in this manner we have four band gaps each spaced 80 nm from each other.

Next to the location of the PL peak, for our device it is also necessary the material has remained of high quality. To assess this, we fabricated laser test structures. They are identical in every way, except that we have shifted the band gap by 80 nm with our quantum well intermixing technique. Figure 2.13 illustrates the design of the laser structures. An intermixed InP die was bonded on a patterned silicon waveguide circuit. In this transferred die, an optical amplifier was defined. The amplifier consists of a wide mesa, where current confinement is ensured by a proton implantation. The hybrid Si–III-V mode of the mesa couples to the fundamental mode of the silicon waveguide with a short taper. Finally the silicon waveguide is cleaved to create reflective facets. The structure acts as a Fabry-Perot laser, where the amplifier may or may not have been implanted with phosphorous ions before the thermal anneal. This means that also the non-intermixed lasers have been subject to the thermal anneal, but because of the lack of ion catalysts there was no change in band gap. Devices with and without quantum well intermixing



(a) 3D illustration containing an III-V amplifier coupling to silicon waveguides reflecting off polished facets.

(b) Cross section of the amplifier mesa.

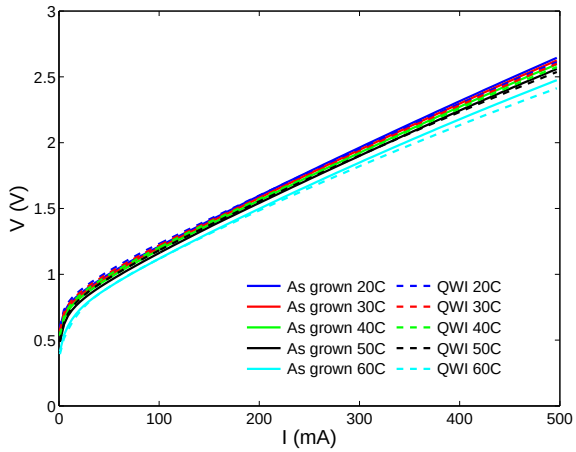
Figure 2.13: Illustration of the laser design to test the material quality after quantum well intermixing.

were fabricated on the same chip, to rule out fabrication variations a much as possible. The details of the fabrication process and the taper and mesa design can be found in section 3.2, as they are the same as the SLEDs.

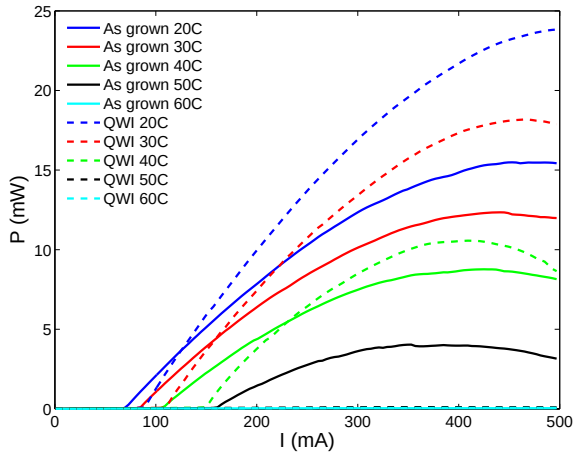
Figure 2.14 shows the characterization of one set of lasers in the C-band. The lasers in the as-grown material were operating around 1555 nm, the intermixed ones around 1485 nm. The voltage and output power was measured as a function of drive current for different temperatures, both for intermixed and non-intermixed lasers. Several observations can be made, the first being that the lasers are very similar. This confirms that there is no damage induced by the QWI process. Some minor differences can be spotted as well. For the blueshifted lasers, the threshold current has increased, as has the slope efficiency. The maximum power is higher and the temperature has a more pronounced effect. Considering that the intermixed quantum wells are shallower, the electrons can escape more easily due to thermal vibrations. A logical consequence of this is a larger temperature dependence. Shallower quantum wells also typically have a worse current injection efficiency, leading to the increased threshold current. The slope efficiency depends on the injection efficiency and the mirror and internal losses through the following equation.

$$\eta_d = \frac{\eta_i \alpha_m}{\alpha_m + \alpha_i} \tag{2.8}$$

with η_d the slope efficiency or differential quantum efficiency, η_i is the current injection efficiency, α_m and α_i the mirror and internal losses. A higher slope could come from higher mirror losses, provided the internal losses are not negligible. Physically, this means the outcoupling of the mirror is increased, while the internal losses of the cavity do not change as much. The reason for these higher mirror losses could be found in small differences in the polished facets of the silicon waveguides. Again, the main observation is that the lasers are very similar,



(a) Voltage as a function of drive current.



(b) Output power as a function of drive current.

Figure 2.14: Characterization of laser structures to asses material quality after quantum well intermixing (C-band).

indicating the quality of the QWs is preserved in the intermixing process.

2.2.4 Conclusion

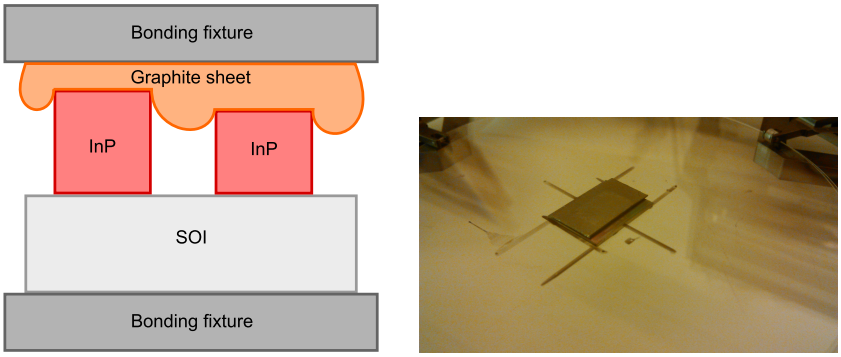
Using quantum well intermixing we can blue shift the emission of a MQW structure. Using implantation enhanced disordering (IED), we have successfully shifted the band gaps of two layer stacks, one in the O-band and one in the C-band. The phosphorous implantation is done in an InP buffer layer, which is later removed. During the thermal anneal these ions diffuse into the quantum wells catalyzing the disordering. We simulated the needed diffusion in order to achieve a sufficient shift of band gap. Next, an empirical relation was established between the anneal time and wavelength shift. Using both the model and empirical relation, we could characterize the material as well as the anneal furnace itself. The difference in diffusion coefficient explained why one layer stack was much more stable than the other, while the machine showed a dead time of around 20 seconds. A complete fabrication process was established rendering an intermixed epitaxial layer stack, which can be bonded in an aligned manner to a silicon PIC. Using the two epitaxial layer stacks, we could created 4 III-V band gaps on silicon, each spaced 80 nm apart from one another. To asses the quality of the intermixing, Fabry-Perot lasers were fabricated. The performance of these devices was very similar, with the intermixed lasers even outperforming the as-grown ones on some criteria. Therefore, it can be concluded intermixing does not have a damaging effect on the active region.

2.3 Multiple die bonding

"If bonding one III-V die offers you one band gap, than bonding two dies can offer you two band gaps, it's that simple." This is what a professor will tell you about multiple die bonding. And indeed, it seems the most straightforward manner of integrating multiple active regions on one silicon chip. The difference in thickness of the two III-V stacks might be problematic though, as we will incur different forces on both during the bonding process. There is also the practical problem of positioning the two dies and dealing with the exposed gap in between them. These are discussed below.

2.3.1 Difference in die thickness

When InP wafers are sliced from the ingot, the thickness can vary by $\pm 10\mu$m. When bonding two such dies and applying pressure with a rigid block, only the thicker die will feel it. In the most extreme case, the thinner die will delaminate while the thicker one will be cracked. The use of a graphite sheet can help to spread the pressure evenly though. It will compress on top of the thicker die and thus also contact the thinner one, indicated in figure 2.15



(a) Illustration of multiple die bonding with graphite sheet sheet to spread the pressure.

(b) Sample ready to load for bonding. Two InP dies are positioned on a SOI sample, with graphite sheet covering them. Everything is sandwiched between two glass wafers.

Figure 2.15: Use of graphite sheet to spread the applied pressure during bonding

In our case, the difference in die thickness was $\pm 5\mu$m, which can be successfully mitigated by the graphite sheet. The bond quality was as high as for single die bonding, both in the case of molecular die bonding and adhesive-assisted (BCB) bonding. These two bonding processes were discussed in more details in sections 1.2.2.1 and 1.2.2.2 respectively

2.3.2 Positioning of the dies, exposed gap and closest distance between two devices

Bonding two dies in the correct position requires some alignment, which admittedly can be very rough. The bonding is all about getting a sheet of epi in the right place, so as long as the correct active material is covering the correct SOI structures, it has done the trick. Typically alignment by eye will be satisfactory. On top of this, there is always a part of the SOI structures exposed after bonding. This should not necessarily be a problem, but most heterogeneous integration fabrication processes so far rely on the fact that all of the relevant SOI is protected by the die. This means that any dry etch might damage the exposed structures. To solve this problem without needing to change the entire process flow, we focused on two aspects: keep the gap as small as possible and provide an extra capping layer on the structures.

Again, we have tried this both by molecular bonding and BCB bonding. Figure 2.16 shows a picture of the two III-V dies bonded to silicon on the glass carrier of the bonding tool. To aid in the alignment, some lines were drawn on the glass

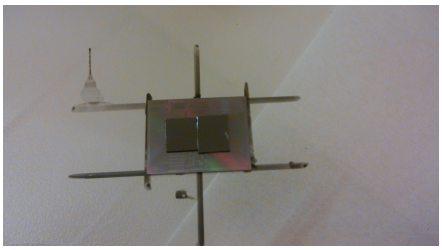
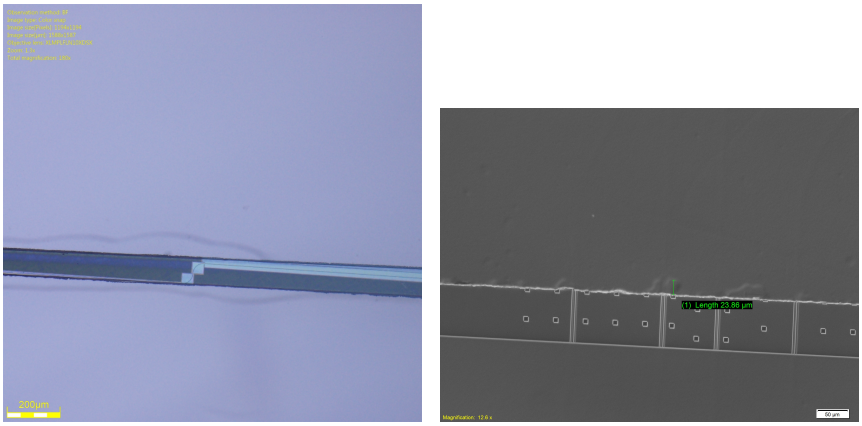


Figure 2.16: Multiple die BCB bonding using auxiliary lines

carrier indicating the positions of the SOI and gap region between the InP dies. Aligning these dies manually is very different when doing molecular bonding or BCB bonding. In case of partially precured BCB, the InP die tends to slide around after which repositioning is needed. In molecular bonding this is not the case, as during the first contact already a weak van der Waals bond is established. When dealing with one die this does not have a big impact: there is enough error margin such that it can always be positioned correctly and the worst that can happen is that one has to reposition the die to its original position (for BCB bonding). The story is different for multiple dies though. In the case of BCB bonding you can push the dies against each other thereby minimizing the gap between them. As this option is absent for molecular bonding minimizing the gap is more difficult. Because we combined multiple molecular die bonding with quantum well intermixing, we used a flip chip tool to accurately position the dies. Gaps as narrow as 60 μm were achieved for both molecular and adhesive bonding. Adding a safety margin, we designed the dies to be 100μm apart, shown in figure 2.17 for both bonding techniques.

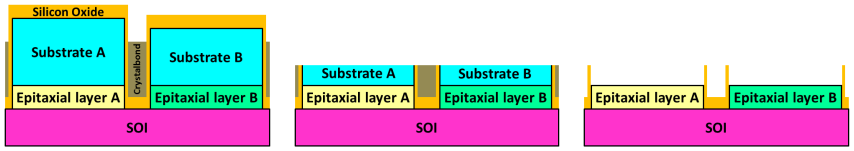
Even though the gap has been minimized to 100 μm, the fact remains that this region is exposed throughout the post-bond fabrication process. Any hard mask will pose problem as a slight overetch of the mask can affect the silicon waveguide layer. Obviously, decreasing the length over which this would be done was of paramount importance but nonetheless, if the waveguide is affected, even 100 μm waveguide length might introduce a large loss. Looking in literature a similar problem was encountered by G. Kurczveil [9]. There this was resolved by deposition of several protective dielectric layers, manual photolithography using a microscope and selective etching. As only a few waveguides were crossing the edge of the bonded InP die, this was still feasible. In our case, the high amount of waveguides crossing the gap region would make this process very time consuming. In addition, the InP bottom device layer was also attacked at the sides, causing a wicking effect as discussed below.

Therefore, we developed a more elegant solution, illustrated in figure 2.18. 1



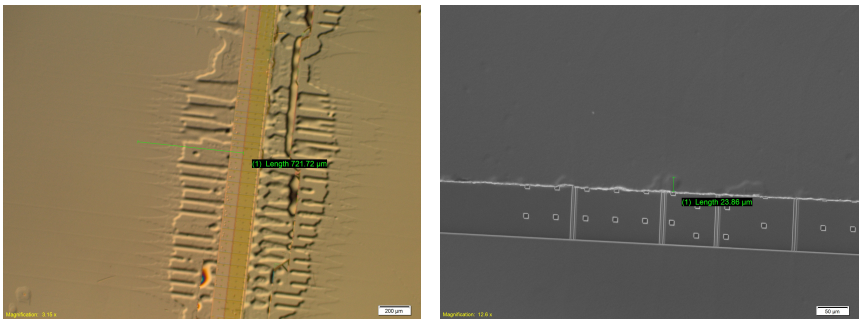
(a) Multiple BCB bonding *(b) Multiple molecular bonding*

Figure 2.17: Minimizing the gap region for both bonding techniques to 100 µm. The top and bottom of the picture show the two bonded InP dies. In the middle the exposed waveguide structures are visible.



(a) Combination of protective oxide and crystalbond *(b) Mechanical polishing* *(c) Wet substrate removal*

Figure 2.18: Schematic of the process to protect the gap region



(a) Wicking because of traveling wet etchants in the waveguide trenches. Cones of delamination, longer than 700μm, can be seen near the gap between the dies.

(b) Importance of the cleave quality to prevent wicking. The cones of delamination have clear, indicating the absence of traveling etchants.

Figure 2.19: Edges of the bonded InP die attacked during substrate removal. A SOI region can be seen, surrounded by two bonded InP dies.

μm of conformal silicon oxide was deposited using PECVD after the thermal anneal of the bonding (in which the covalent bond of molecular bonding is formed), but before the substrate removal. Crystalbond wax was then generously applied, as usual, protecting the SOI from dirt during the lapping. Particular care was taken to fill the gap region completely with Crystalbond wax. The mechanical polishing opens up the InP substrate, removing spill-over Crystalbond wax and oxide alike. In stead of the usual 100 μm, we leave only 50 μm of substrate. This was as far as we dared to go, fearing cracks because of the applied forces during grinding. After this, the remainder of the InP substrate is removed using hydrochloric acid. As can be seen in figure 2.18c, this self-aligned process leaves pillars of silicon oxide. These can potentially be very high. Strangely enough, we never observed them. They might be blown off when drying the sample after the wet etch. The mechanical polishing might also cause them to come of. After this process, 1 μm of SiO_2 is protecting the gap region without needing to apply any mask.

For the molecular bonding, the silicon waveguides are defined by etching air trenches. Therefore, wet etchants can travel in these trenches. More in particular, during the wet substrate removal, hydrochloric acid will penetrate the trenches. Unfortunately, the bottom device layer is also InP meaning that it is etched rapidly. Characteristic cones are observed, as indicated in figure 2.19a. Under these $\pm$ 700 μm long cones the bottom device layer has been etched and thus the epi above will come off during the process. Obviously no devices can be made here, but what is more, the SOI may again be exposed and thus the silicon circuitry may be attacked. Usually this is not a problem as the InP die is bigger than the silicon circuitry, but this does not hold here. In figure 2.19a the normal bonding was performed, i.e. no

additional protective oxide was deposited. Figure 2.19b shows what happens when the gap protection is applied. The waveguide trenches at the edge of the die have been filled with oxide and the wicking has been reduced to $\pm$ 25 μm. Upon careful inspection it can be seen that this time, the underetched region does not resemble a cone and they do not necessarily coincide with waveguide trenches. The bottom die indeed does not have any wicking at all. The cleave of the lower die was much better than the upper one, and clearly this is defining the quality now. Note that we cannot be sure the defects on the upper die are the consequence of the undercut by the wet etch. It is equally possible the bonding was never good at that point.

2.3.3 Multiple die bonding and quantum well intermixing

Multiple die bonding and quantum well intermixing are very complementary techniques. The first provides complete freedom on the design of the epitaxial layer stack, but severely limits you in terms of spatial resolution. Because the minimum bondable dies still has dimensions of a few millimeters, the components of different band gaps cannot be placed close together. The footprint of the circuit quickly becomes very large this way. It also means that a circuit might have to cross over from one III-V layer stack to the other multiple times. QWI on the other hand does offer this spatial resolution, as the different band gaps are defined by lithography rather than placement of bonded dies. The different gain media now have a very similar epitaxial layer stack, with only the shape of the quantum wells changed by diffusion. Using quantum well intermixing also requires an aligned bonding, as both the III-V and silicon chip are now patterned.

The multiple, aligned bonding of quantum well intermixed dies was done using a flip chip tool, the *finetech FINEPLACER lambda*. The specifications guarantee a placement accuracy of $\pm$ 0.5 μm, but the main limitation was the alignment itself, achieving a precision of 2 μm. [1]

During placement, we wanted to keep track of the gap region at all times. If a mistake was made, the to-be-placed die might hit the already bonded die. This would kick off the latter and ruin the bond interface of the former. An error at

[1] The best alignment can be achieved when the markers are perfectly centered to the lens of the flip chip bonder. This lens can only move in one dimension however, chosen along the gap region. One can put the markers bordering the gap region, but this is not advisable. The markers themselves are not bonded as we designed them to lie lower to the bond interface (in section 2.2.2.2). Since the borders of the die are exposed to a lot of processing during substrate removal, we do not want to put this weak point too close. The alignment markers were positioned 150 μm away from the edge of the die. This allows us to center the markers to the lens (in order to be able to go to high magnification), to see the gap region under low magnification and not to weaken the edge of the die with the non-bonded alignment markers. Note that it is far from straight-forward to center the alignment markers properly to the lens, as we can only move the optics in one dimension. We have to place the to-be-bonded die facing down on the table, while making sure that the alignment markers will be centered to the lens. Taking all the previous into account, we managed to achieve a bonding precision of 2 μm repeatedly.

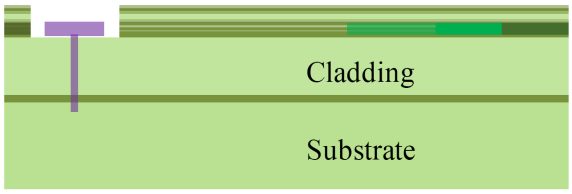


Figure 2.20: Deep alignment markers to check the bonding precision. Reproduced from [1].

this stage will thus have irreversible consequences. By keeping an eye on the gap region and being ready to hit the emergency stop, the risk of a fatal error is greatly reduced.

The distance between two III-V active regions is now not only defined by bonding precision and gap between the two bonded dies only. Because the III-V has several lithography steps on it, the structures cannot be at the edge of the chip. The edge bead of the photoresist is at least 1 mm, which would mean that the shortest distance between two hybrid devices would be over 2 mm. To overcome this, the sample was cleaved using a precision cleaver *Loomis LSD-100*, which has a precision of less than 10 μm. Note that the edge of the die is defined through crystal planes rather than lithographic features. In order to let the edge of the sample and the edge of the defined die coincide, it is of paramount importance that the first lithography defines the patterns parallel to the crystal planes. Otherwise a walk off will occur between the cleave and the patterns, in effect increasing the gap between two bonded dies. By precision cleaving the dies, solving the wicking during the substrate removal and aligning the dies carefully, the gap between them is entire defined by a safety margin of how close you want to bond them together.

The alignment markers themselves are not bonded, as we designed them to lie lower than the InP bottom device layer surface. After substrate removal, the III-V alignment marker used for the bonding cannot be discerned anymore. In [1], this was solved by defining an extra set of markers, which are now much deeper as illustrated in figure 2.20. They are still unbonded, but because of the sheer size of them and the fact that they are deposited on the bulk of the layer stack, the deep markers remain visible after bonding. Therefore, they can be used to estimate the bonding precision, but at a high cost. Three extra lithographic steps are needed, complicating the process considerably.

In order to limit the pre-bond processing, we used silicon rulers as shown in figure 2.21. Before bonding, we measure the precision of the cleave described above. After this, we perform the actual bonding and remove the substrate. At this moment, we can see the silicon rulers and the bonded epitaxial layer stack in one

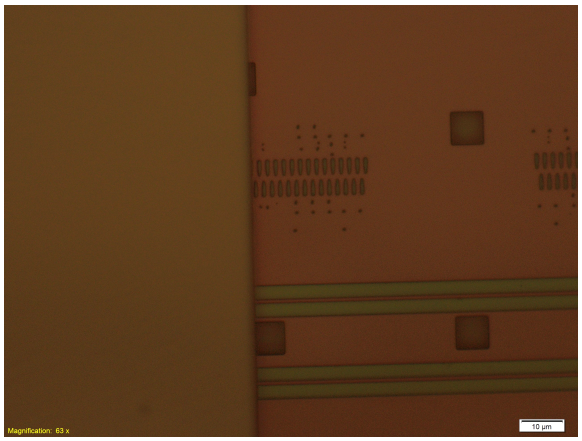


Figure 2.21: Silicon rulers to check the bonding precision.

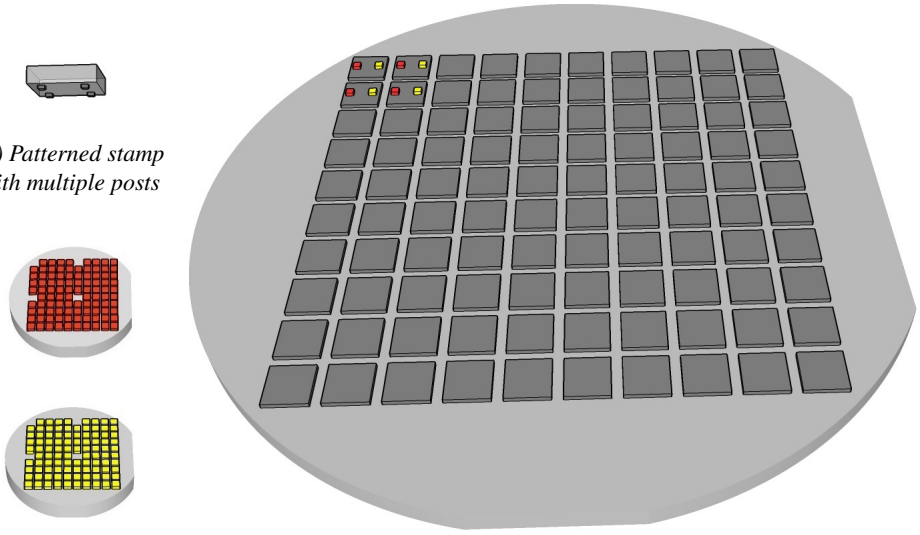
focus plane and estimate the bonding precision. In this manner, we could establish that the placement was always within $\pm 2\mu$m.

2.3.4 Conclusion

In this section we discussed multiple die bonding. Two III-V dies are bonded at the same time to a patterned SOI chip. The difference in substrate thickness was alleviated successfully by a graphite sheet. When bonding two dies, there is always some alignment needed. Auxiliary lines drawn on the glass carrier were of great help in the BCB bonding. The gap between the two dies was brought below 100 μm for both molecular and BCB bonding. To protect this gap from the post-bond processing, a self-aligned SiO_2 protection was deposited. This also solved the problem of traveling etchants, such that the hydrochloric etchant does not travel in the waveguide trenches. The wicking problem was reduced from $\pm$ 700 μm to essentially zero. Finally, we elaborated on the combination of multiple die bonding and quantum well intermixing, highlighting the bonding precision of 2 μm.

2.4 Transfer printing

The above multiple die bonding allows for the integration of different epitaxial layer structures onto SOI. However, a lot of design freedom is sacrificed because of the minimal size the die can be before it is impossible to handle to bond (typically around millimeters by millimeters). Additionally, the III-V is not used very efficiently, as a large sheet of material is transferred and only a few percent of that area ends up being used for devices.



(a) Patterned stamp with multiple posts

(b) Two source substrates with patterned coupons

(c) SOI target substrate with four printed coupons of each source in top left.

Figure 2.22: Transfer printing of four coupons of two different InP source wafers to SOI. The stamp with multiple posts ensures the high throughput, while the small coupons allow for spatial design freedom. The area magnification is readily apparent as the entire SOI wafer can be populated using the small InP wafers.

Transfer printing is a technique where only small coupons of material are transferred from one substrate to another. In this manner, we can mitigate the inefficient use of III-V by only providing it where necessary. Moreover, because the transfer sheet of material is now only micron-sized, different materials can be integrated without loss of design freedom. By using arrayed stamps, we can transfer arrays of coupons at the same time. Figure 2.22 shows the transfer printing of four coupons to silicon of two different InP wafers. The four coupons were transferred at the same time, allowing for a high throughput process. The area magnification is made very clear, as the entire SOI wafer can be populated with the small InP wafers. Depending on the size of the coupons, this can even get more drastic, where tens of large silicon wafers only consume one small InP wafer.

Below, we discuss how transfer printing works and show the road of development to make it suitable for our applications.

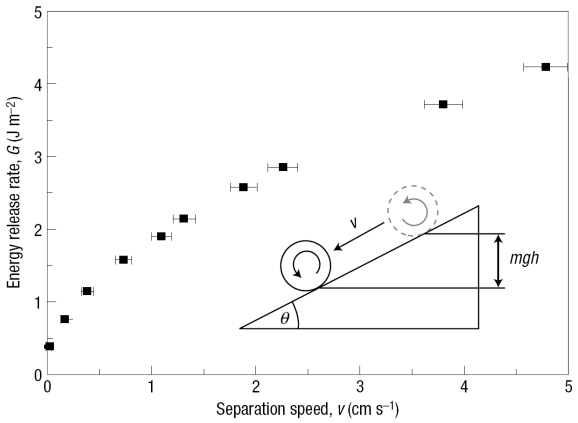


Figure 2.23: Kinetic response of the adhesion strength of PDMS. Reproduced from [10].

2.4.1 Working principle of transfer printing

Transfer printing relies on a kinetic response of the adhesion strength of the coupon (here InP-based) to the elastomer stamp (here polydimethylsiloxane, PDMS). The physics of this response has its origins in the viscoelastic effect in PDMS. When stress is applied to PDMS in a very fast manner, it is much more stiff and will act as an elastic material. Inversely, it is viscous when the stress changes very slowly. It means that a ball of PDMS which is thrown on a table will bounce back at you, while if that same ball is left to lie on the table overnight, it will spread itself and conform to all the irregularities of the table surface (so-called creep). The same effect is now used with a PDMS stamp and a semiconductor surface. When the stamp is pulled away quickly, the adhesion of the PDMS to the semiconductor will be high and when retracted slowly the adhesion is low. This was quantified by Meitl et.al. [10], the key result reproduced in figure 2.23. The separation energy G_{PDMS} depends on the speed of delamination v, i.e. $G_{PDMS} = G_0\,[1 + \phi(v)]$, where ϕ is a monotonically increasing function of v and G_0 is the energy release rate at standstill. In general ϕ can be approximated as $\left(\frac{v}{v_0}\right)^n$ with v_0 the speed at which the adhesion energy is double the standstill value and n a fitting parameter. This renders G_{PDMS}:

$$G_{\text{PDMS}} = G_0 \left[1 + \left(\frac{v}{v_0}\right)^n\right] \tag{2.9}$$

The separation energy of the to-be-picked coupon to the substrate $G_{\text{Substrate}}$ does not depend on v, as semiconductors are elastic materials. If the speed of movement is great enough, the adhesion of the coupon to the stamp might overcome the one to the native substrate, i.e. G_{PDMS} is bigger than $G_{\text{Substrate}}$. In that case the coupon would be picked up. The reverse can then occur when the speed becomes low and

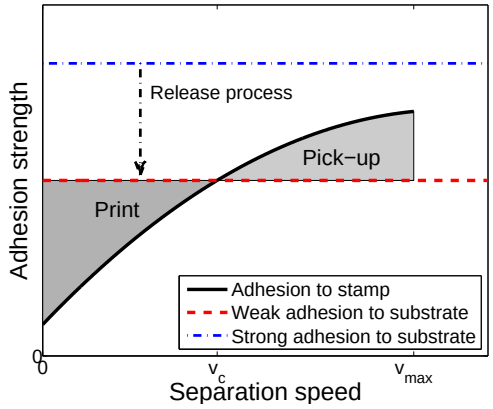


Figure 2.24: Schematic diagram of critical energy release rates

the coupon is printed on a new substrate. This is illustrated in figure 2.24 The printing and pick-up regions are indicated and separated by the critical speed v_c at which the adhesion to the stamp and the substrate is equal. This critical velocity easily computes to

$$v_c = v_0 \left(\frac{G_{\text{substrate}} - G_0}{G_0} \right)^{\frac{1}{n}} \tag{2.10}$$

So far we have assumed that this critical speed is lower than the maximum reachable speed $v_c < v_{\text{max}}$, but this is not necessarily true. Even more, when working with epitaxially grown interfaces, as we would do in semiconductors, the chemical bonds are far more strong than the weak bonds to the PDMS stamp. In order to make the coupons pickable, we have to weaken the adhesion to the substrate. This is done by undercutting the coupon and having it suspended in air. The released coupons are held on the substrate by tethers, the strength of which can be engineered.

Finally, also temperature also plays an important role in the viscoelastic kinetics controlling transfer printing. From [11], it is confirmed we can account for temperature effects by adding a temperature shift factor:

$$G_{\text{PDMS}} = G_{\text{PDMS}} \left(v a_T \right) \tag{2.11}$$

which can be approximated as

$$\log_{10} a_T = -17.6 \frac{T - T_g}{52 + T - T_g} \tag{2.12}$$

where T_g is the glass transition temperature of PDMS. It follows that the critical velocity becomes larger at higher temperature. It is therefore beneficial to pick up coupons at low temperatures and print at high.

#	Layer	Thickness	PL
0	InP substrate	360 μm	
1	InP buffer layer	150 nm	
2	$In_{0.532}GaAs$ release layer	1 μm	
3	InP	50nm	
4	InGaAsP (Q=1.35μm)	10 nm	
5	$In_{0.532}GaAs$	8 nm	1580 nm
6	InGaAsP (Q=1.35μm)	10 nm	
7	$In_{0.532}GaAs$	8 nm	1580 nm
8	InGaAsP (Q=1.35μm)	10 nm	
9	$In_{0.532}GaAs$	8 nm	1580 nm
10	InGaAsP (Q=1.35μm)	10 nm	
11	$In_{0.532}GaAs$	8 nm	1580 nm
12	InGaAsP (Q=1.35μm)	10 nm	
13	InP	20 nm	
14	$In_{0.532}GaAs$ sacrificial layer	100 nm	
15	InP sacrificial layer	1 μm	

Table 2.4: Epitaxial layer stack suitable for transfer printing

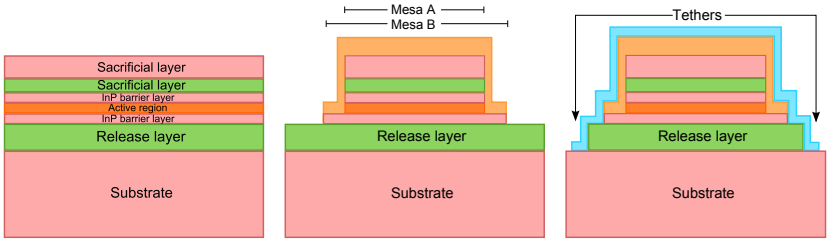
Armed with this knowledge we can now depict the transfer printing process. First we need to process the InP source material to release the coupons. Pick-up is then done by rapidly retracting the PDMS stamp after coming into contact with the coupon. If necessary we can cool down the system to decrease the critical velocity. Printing finally occurs by slowly retracting the stamp from the (if needed heated) target substrate. The adhesion to the new substrate is now greater than to the stamp and transfer printing is successful.

2.4.2 Process Development

2.4.2.1 The process established

As discussed in the previous section, it is necessary to release the coupon from its donor substrate. This can be realized, e.g. for the layer stack in table 2.4, being a layer stack that will be used later in this work to realize transfer printed light emitting diodes on silicon. To release the coupons from the substrate, layer #2 will be etched away in the layerstack of table 2.4. The to-be-transferred coupon will consist of layers 3 to 15 and hence will have to be protected during the release etch. Figure 2.25 shows a schematic of the process flow. There are 4 mask steps.

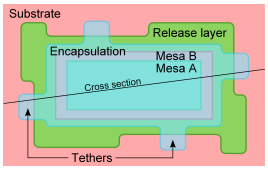
- **Mesa A definition:** The first step is to etch a mesa through most of the coupon layers. Ideally, we want to stop inside the bottom device layer #3. As in our cause this layer is only 50 nm thick, we stopped on top by a selective



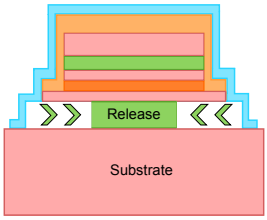
(a) InP starting layer stack

(b) Coupon patterning

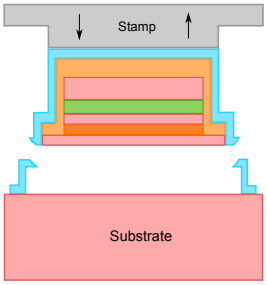
(c) Encapsulation and tether definition



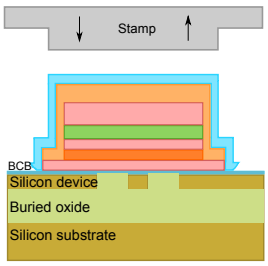
(d) Top view after encapsulation and tether definition. The cross section of the other figures is indicated.



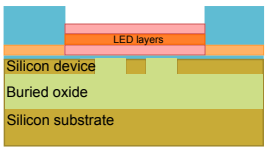
(e) Release etch



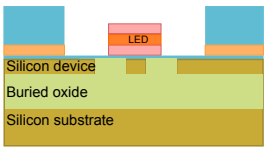
(f) Pick-up of the III-V coupon from the InP source substrate



(g) Printing of the III-V coupon to the SOI target substrate



(h) Removal of encapsulation and of sacrificial layers



(i) Post-bond processing

Figure 2.25: Process flow of transfer printed, optically pumped LED.

wet etch. This etch is done with a hard mask followed by a combination of dry and wet etching.

- **Mesa B definition:** Next, a second and wider step is etched. Here we want to stop level with the bond interface. Since our final encapsulation will sit on the layer we expose in this etch, it is important not to overetch this layer. If we do so, the photoresist encapsulation along the coupon will be lower than the bond interface and may affect the bonding quality. This etch is done with a highly selective wet etch and using a silicon oxide hard mask. This hard mask is not removed before the transfer printing and during the release etch it will act as an extra protection for the layers that were exposed during the Mesa A definition. This is shown in figure 2.25b. Strictly speaking, this step is redundant and could be replaced by conformal deposition of a dielectric followed by a maskless directional etch. In this manner the sidewalls would also be protected by the dielectric, but extra risk is introduced in the process. If there would be a small defect in the encapsulation during the release etch, Mesa B offers an extra safety distance between the edge of the coupon and the active layers. Because we can afford to have this extra mesa definition, it was always done so.

- **Sacrificial layer etching:** Tethers will have to be formed to keep the free hanging structure on the InP substrate. These tethers need secure places to anchor, i.e. they need to anchor to the substrate. Therefore we etch through the sacrificial layers and into the substrate where the tethers will later be defined. Note that we keep the release layer in between the tethers and along the coupons. This ensures that the release etchant will have enough places to penetrate the release layer.

- **Encapsulation and release etch:** The entire coupon is encapsulated with photoresist and tethers are formed to hold up the coupon. Subsequently, a suitable release etchant removes the release layer #2. The release layer is etched horizontally over the width of the coupon, while the bottom InP device layer acts as a vertical etch stop. Note that the selectivity required here is approximately 1000, as microns are etched horizontally while we can only tolerate nanometers vertically (to ensure a good bonding interface).

- **Pick-up and print of coupon:** Once the free hanging structure has been established, the actual transfer can happen. Pick-up is done by retracting the PDMS stamp quickly from the coupon, thereby exerting a large force on the coupon. After initial contact with the coupon, some overdrive is applied as instantaneous speed is impossible. The stamp then accelerates up by $1g$, as indicated in figure 2.26. A typical overdrive distance of $\pm$ 30 μm is applied and hence the speed at which the coupon is picked is $\pm$ 2.4 cm/s. This

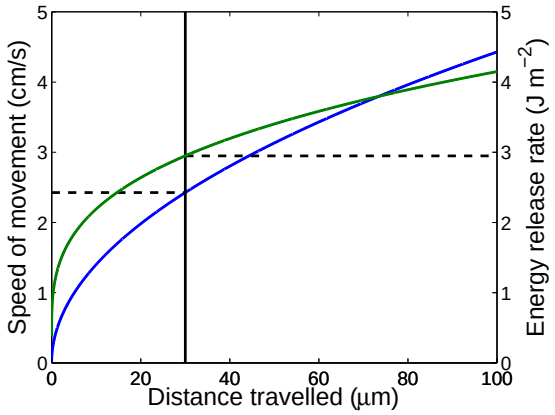


Figure 2.26: Acceleration and corresponding adhesion energy of the stamp during pick-up. An overdrive distance of 30 μm is indicated by the black vertical line.

corresponds to a release energy of almost 3 J m^{-2}. If more is necessary, one must either increase the acceleration or, more conveniently, the overdrive distance. As the post making contact is 50 μm high, this is an upper limit of the workable overdrive distance. If more is used, the bulk of the stamp is touching down and accidental pick-up of neighbouring coupons may occur.

The printing is done at a stamp velocity 0.02 mm/s, meaning that the energy to be overcome is only $\pm$ 0.38 J m^{-2}. This can be done without adhesion layer by most target samples we are interested in: pure Si, glass, aircladded SOI waveguides, SOI waveguides with oxide cladding in the trenches. Because a thin adhesion layer did improve the bonding quality, we ended up using a 50 nm thick BCB layer.

- **Planarization:** During the device processing, exposed SOI might be damaged due to the dry etches. To protect the Si PIC, we deposit 200 nm of PECVD SiO_2, followed by spinning 1.6 μm of BCB. The BCB is then etched back in the RIE, thereby exposing the top of the coupons in a self aligned manner.

- **Removal of sacrificial layers:** The layer stack contains a sacrificial layer pair of 1 μm InP and 100 nm InGaAs. These ensure the mechanical stability of the free-hanging and picked coupon. Since the device layers only total to a thickness of 150 nm, the chance of bending and buckling is too high, especially with the strained dielectric hard mask of Mesa B on top. After removing the dielectric, both sacrificial layers are etched with a combination of dry and wet etching. Because we have to do lithography on the remaining

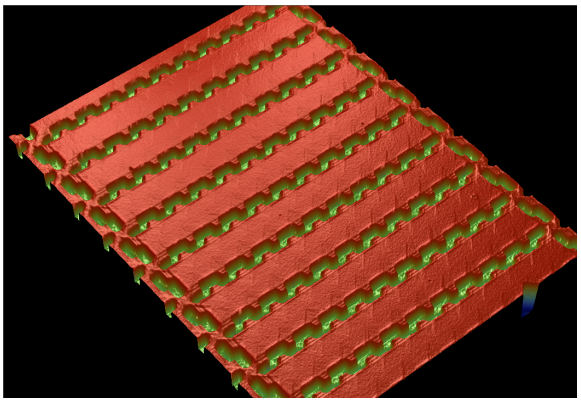
Etchant	Observations
H_2SO_4:H_2O_2:H_2O (1:1:18)	Partial undercutting
H_3PO_4:H_2O_2:H_2O (1:1:20)	Partial undercutting
Chromium etchant	Partial undercutting
HF:H_2O_2:H_2O (1:1:10)	Full undercutting
	Poor selectivity
	Encapsulation delaminates
Citric acid : H_2O_2 (1:10)	Full undercutting
	Slow etching speed
	Anchors delaminate
Tartaric acid : H_2O_2 (1:1)	Full undercutting
	Selectivity to InP >500
$FeCl_3$:H_2O (3 mol/l) @ RT	Full undercutting
	Selectivity to InP >500
$FeCl_3$:H_2O (3 mol/l) @ 5°C	Full undercutting
	Selectivity to InP >2000

Table 2.5: Etch tests for releasing coupons

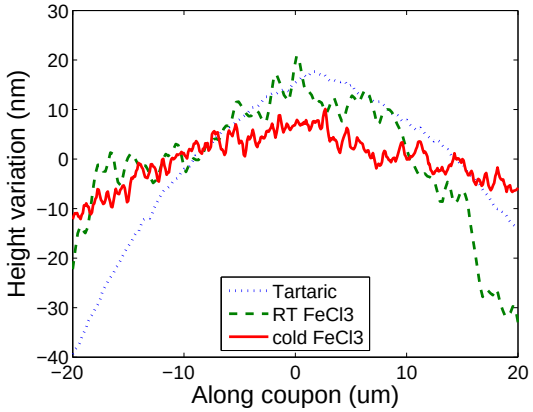
device layers, we have to make sure the BCB-well is not too deep. In the previous step, we do a sufficient overetch to end up with a step of a few hundreds of nanometers. The result is shown in figure 2.25h, with the final device processed in figure 2.25i.

2.4.2.2 Development of the release etch

The release etch is one of the critical steps in the process. We have to etch the release layer (InGaAs in our case) along the width of the coupon (typically 20 μm or more). Meanwhile, we can only etch a few nanometers vertically into the III-V device layer. At the same time, photoresist must also be able to withstand the etch as the coupon is protected and held up by photoresist tethers. We tried multiple etchants, recounted in table 2.5. The sulfuric acid-based and phosphoric acid-based etchants did not render a complete undercut, as was the case for the chromium etchant. This is because of anisotropic etching, where a slow etching crystal plane is being exposed (in these experiments the coupons were oriented with the short side along the (110) crystal axis). The hydrofluoric acid-based and citric acid-based mixtures were able to undercut the coupon completely. However, in both cases the encapsulation and tethers were delaminated. The HF based etchant was also not very selective, while the citric acid was etching very slowly. The tartaric acid-based and iron(III)chloride solution one were able to undercut the structure with suitable selectivity. To accurately estimate the selectivity, we examined the donor substrate (i.e. the InP substrate onto which the coupons were anchored, after the transfer printing of the coupons) using a Veeco Wyko micro-



(a) 3D image of donor substrate.

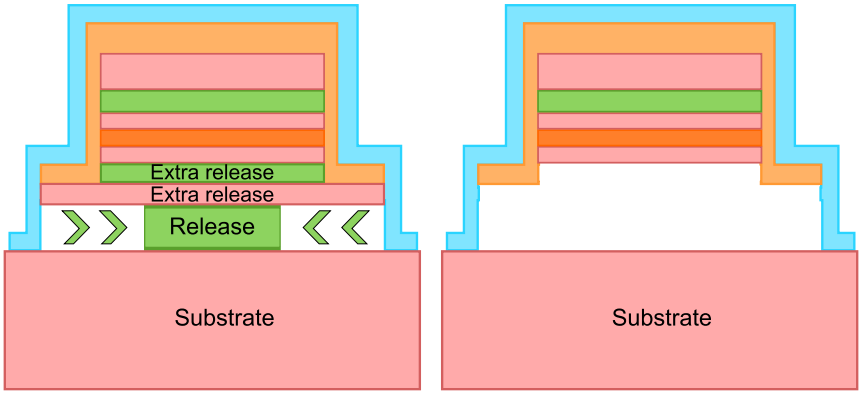


(b) Height variation along the width of the coupon for different etchants

Figure 2.27: Measuring etch selectivity for the release etch

scope system. Figure 2.27a shows the 3D image taken. One can clearly see the etched places where the tethers have anchored the coupons (which are already gone on this image). The red surface is a mirror image of the bottom layer of the coupon. It is analyzed in figure 2.27b, showing that a coupon of 40 μm wide was undercutted with a height variation of tens of nanometers. The tartaric acid based and iron(III)chloride mixtures at room temperature both gave a height variation of 60 nm. When cooling down the latter to 5°C, this was improved to 20 nm. Based on this, we deduced that with the cold $FeCl_3$:H_2O mixture we could release a coupon in 1.5 hours using a bottom InP device layer of only 50 nm.

Because the requirements on selectivity are so high, we also looked into a



(a) Free-hanging structures are created by removing the release layer

(b) The remaining sacrificial layers are etched vertically from the bottom. The Mesa B hard mask is now lower than the bottom device layer.

Figure 2.28: Illustration of the release process in multiple etch steps.

second approach of releasing, illustrated in figure 2.28. In stead of relying on one etchant to do the complete release, while maintaining a good uniformity of the InP bottom device layer (i.e. having a selectivity of over 1000), a multi-step release was devised. To this end, we added an extra InP (100 nm) / InGaAs (50 nm) layer pair in between the release layer #2 and the InP bottom device layer #3 of table 2.4. When etching Mesa B, we also etch through the InP bottom device layer and the InGaAs extra layer, as indicated in figure 2.28a. The release is still done on the 1 μm thick InGaAs layer, but now the selectivity is more relaxed as the 100 nm InP on top is no longer the bottom device layer. A dishing up to 100 nm can be tolerated, as long as it can be remedied afterwards by etching the InP / InGaAs sacrificial layers. Following the release, the extra layers are etched from the bottom, as in figure 2.28b. This can either be done on the source substrate or after pick-up. The biggest advantage of this method is that the requirements on selectivity are greatly reduced. In the release etch, we can tolerate 100 nm while etching tens of microns, so the selectivity can be as low as 100. In the next etches we etch vertically so the selectivity only has to account for the overetch. In worst case this would be 100 nm (from the dish made earlier), so the selectivity required is again only 100 (which can be spread over two etchants).

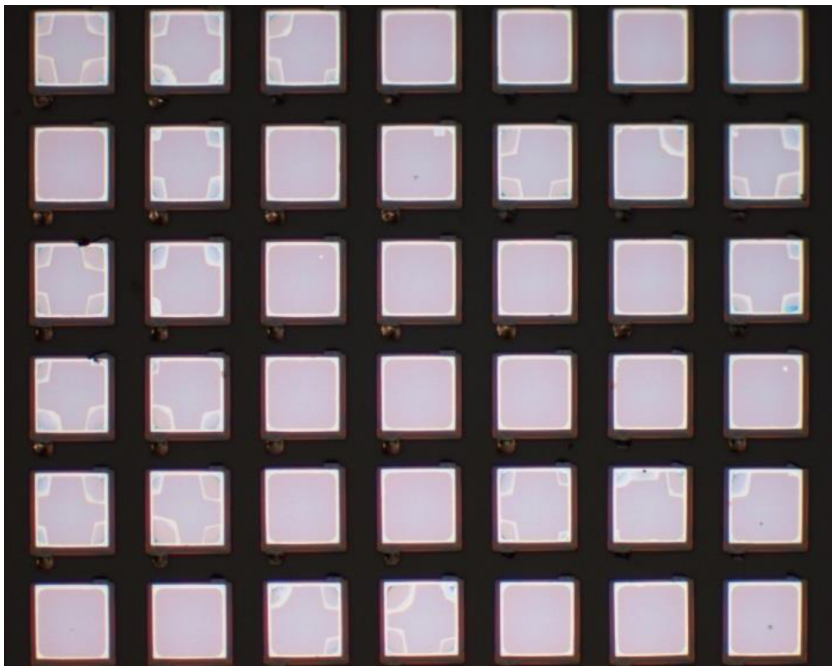
We first tried this after pick-up, because it is a more conservative use of coupons. This was very unsuccessful as the PDMS stamp degraded in the HCl-based etchant (and probably in the $FeCl_3$ as well). The elastomer redeposited on the sample and bonding interface. The contamination was so severe, this was a showstopper for this approach.

When etching all the coupons at once on the InP substrate, we successfully created the coupon we wanted. The pick-up and printing process is identical to the single step released coupons, but the yield was lower. The reason for this can be found in the encapsulation. In the single step release, the encapsulation covers the complete Mesa B and rests on the release layer as shown in figure 2.25d. This means the encapsulation at the sides is level with the bonding interface. This is no longer true for the multi step release because of the extra InP/InGaAs layer pair. The side of the encapsulation lies 150 nm below the bond interface, as can clearly be seen from figure 2.28b. When printing, this encapsulation must be broken causing the initial bonding to be much weaker. Furthermore, the broken encapsulation might be trapped under the coupon, causing bond failure.

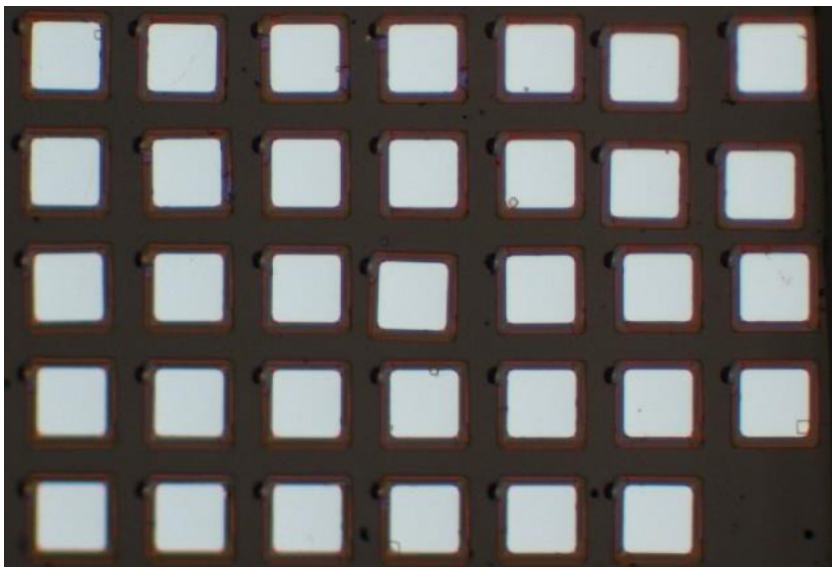
Because of the excellent results of the cold aqueous iron(III)chloride, we discarded the multi step approach and favored the single step approach.

Finally, we also want to highlight to importance of the encapsulation during the release etch. As indicated in figure 2.25b, the SiO_2 hard mask of Mesa B is only removed after the printing of the coupons. It acts as an extra protection of the ternary and quaternary layers in the device stack, which would also be etched during the release etch. The importance of this extra protection is apparent from figure 2.29. Approximately half the coupons of figure 2.29a, where the SiO_2 hard mask from Mesa B was removed, show that the corners have been attacked during the release etch. Although one could argue that some undercut can be tolerated, the situation shown here can not. The area lost is $\pm 40\%$ for these small coupons of 50 μm by 50 μm. Figure 2.29b on the other hand shows the situation when the Mesa B hard mask was not removed. No undercut can be discerned, i.e. the device layers are properly encapsulated.

During these tests, we also noticed some defects in the coupon, shown in figure 2.30a. The origin of these defects in the coupon can be explained by stepping back to the picture after the Mesa A/B etch, shown in figure 2.30b. After the QW etch of Mesa A (using H_2SO_4:H_2O_2:H_2O 1:1:18), defects appeared randomly throughout the sample. After removing the 60 nm thick InP bottom device layer these remained, indicating that the release layer was attacked during the quantum well removal. Growth defects form a channel through which the piranha etchant tunnels through the InP layer and attacks the InGaAs release layer. During the release etch, the channel now works in the opposited direction, allowing the $FeCl_3$:H_2O release etchant to access the quantum wells from below. No measures against this phenomenon can be taken accept ensuring a defectless epistack. Luckily, this was only encountered in one of our 6 different grown epitaxial layer stacks.

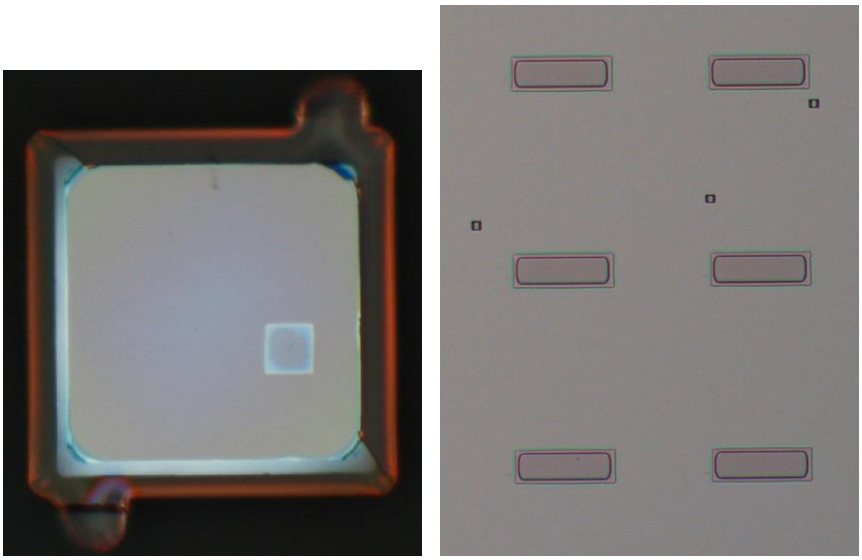


(a) The quantum wells are attacked due to insufficient protection by the encapsulation.



(b) No defects can be seen because of the extra SiO_2 encapsulation.

Figure 2.29: The importance of a good encapsulation during the release etch.



(a) Defects in the coupon after release etch. *(b) Growth defects in the InP bottom device layer showing after Mesa B etch.*

Figure 2.30: Growth defects leading to defects in the coupon during release etch.

2.4.2.3 Tether design

The tethers provide the adhesion of the coupon to the source substrate during the release etch and during pick-up as well. This means that they have to be strong and weak at the same time. Luckily, the tethers also offer us one of the easiest engineering options during the whole transfer printing process.

Figure 2.25d shows the simplest tether designs, i.e. just a simple rectangle. Even with these simple tethers, we can vary a lot of the design parameters: the period and duty cycle of the tethers. A duty cycle of 100 % describes a situation where all of the coupon is flanked by tether, meaning no etchant would be able to penetrate the release layer under the coupon. At the same time a duty cycle of 0 % would mean the coupon is not held on the substrate at all. Figure 2.31 shows the profile of the InP bottom device layer after release etch. The tether pitch was 40 μm and the width 25 μm, so the duty cycle was 62.5 %. A clear V-shape can be discerned. The etchant only has limited space to access the release layer, as indicated by the position of the tethers. This bottleneck is in place both for the etchant trying to get to the unetched layer as for the reaction products of the etch. In normal situations, the iron(III)chloride etch is not transfer limited, but rather reaction limited. However, with a thorough bottleneck due to a too dense placement of tethers, we did observe the etching to slow down.

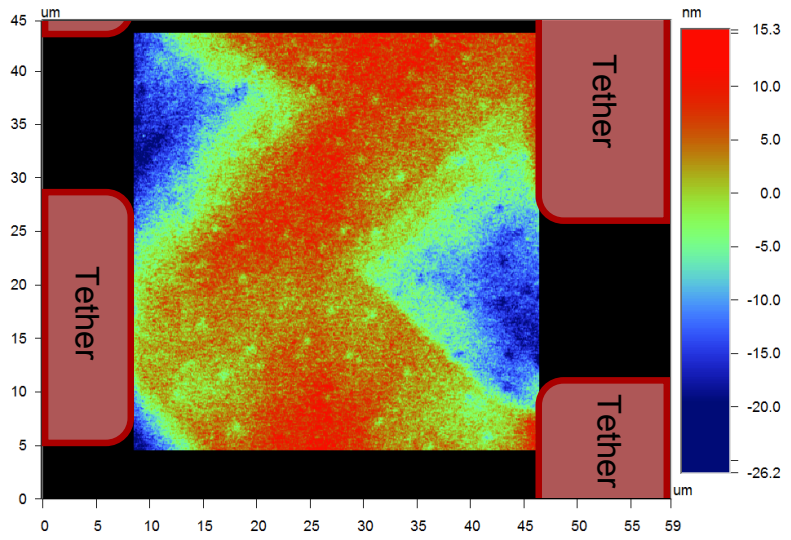


Figure 2.31: Bottom device layer profile after release etch for tethers with duty cycle of 62.5 %.

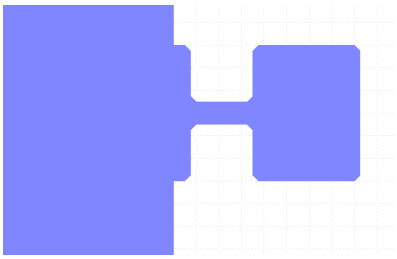


Figure 2.32: Tether with engineered breaking point

Other duty cycles were tried, ranging from 50 % to 17 %. All of these proved to withstand the release etch. Furthermore, no clear yield difference was noticed. The tethers all break at stamp contact. Although the rectangular tethers mostly break close to the coupon (which is where we want them to break), more clever designs were also implemented in which the location of the breaking point was engineered, shown in figure 2.32. The breaking point was engineered to be close to the coupon by making a skinny section of only 2 μm wide. While this worked, the rectangular ones worked equally well.

Finally, we would also like to highlight the importance of the photoresist thickness to engineer the tether strength. As can be gathered from table 2.4, there will be a topography of approximately 2.5 μm on the sample when the photoresist for tether formation is spun. We found spincoating the sample with *TI 35E* at 4000

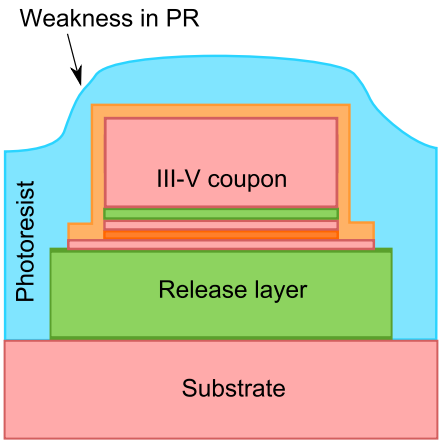


Figure 2.33: The photoresist is not planar after spincoating. Because of this, there is a natural weakness in the tether

RPM gave us the best results. On a planar sample this would give us a resist thickness of $\pm$ 2.8 μm, but because of the topography the resist is no longer planar here. This means at the corner of the Mesa B the photoresist is weak, where the tether is likely to break, as indicated in figure 2.33. Spinning this resist at 3000 RPM renders a 3.6 μm thick layer, which is able to planarize the sample. This mean that the described weakness will be strengthened considerably and figure 2.34 shows the resulting printed coupon. During pick-up, the tethers do not break but rather delaminate from the source substrate, i.e. the adhesion between the photoresist and the substrate fails. This means that the coupon and complete tether are both picked by the PDMS stamp and during printing the tether prevents a good contact. Thinning this resist down by 600 nm with an RIE oxygen plasma weakened the tether enough to break it at the desired location.

2.4.2.4 Development of the bond interface

As the printed coupons still have be processed into devices, it is of critical importance that the bond is of high quality. We examined this using FIB cross sections. We compared 4 different target substrates: unpatterned SOI, airclad patterned SOI circuits, patterned SOI circuits with oxide in the trenches and BCB coated planarized SOI circuits. Figure 2.35 shows a cross section of coupons printed on these four targets.

The bonding on the unpatterned SOI is solid. The two flat surfaces are ideal for the van der Waals forces and hence no delamination was observed with FIB. Also for the airclad patterned SOI the bond interface looked excellent. The coupon bonds to the silicon waveguide and the silicon bulk area, suspending over the

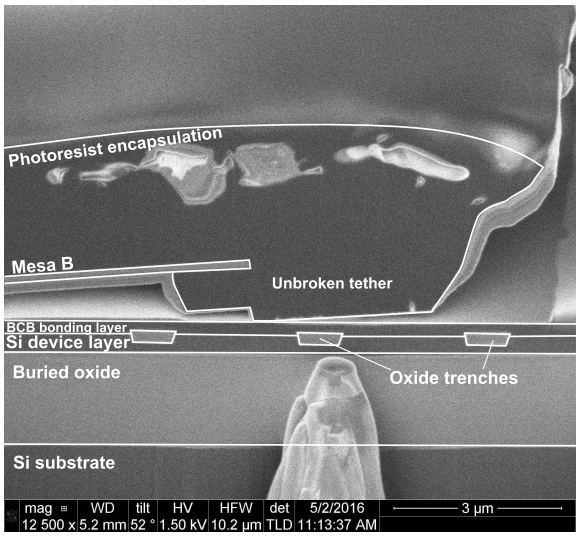
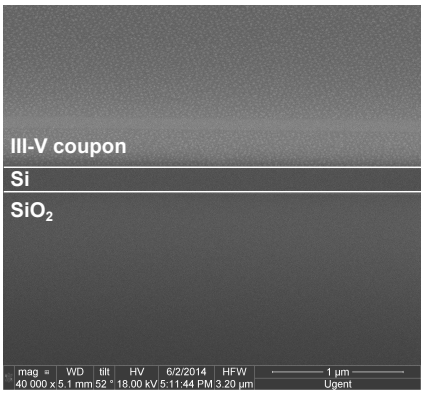


Figure 2.34: Failed print due to unbroken tether. The photoresist tether was too thick and therefor did not break, but delaminated from substrate.

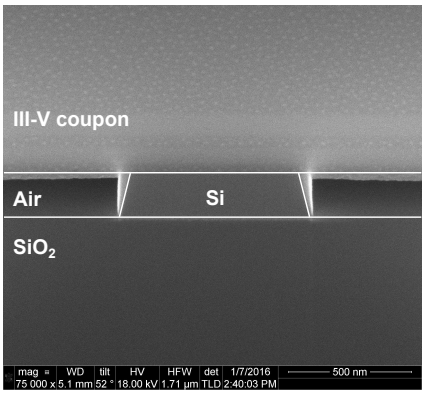
trench. This means that the same wicking problem could occur during the post-processing as discussed in section 2.3.2 for molecular bonded III-V dies. Here, the problem is more severe though, as the distance from the edge of the coupon to the relevant structures is negligible. Figure 2.35c shows the cross section of a coupon printed on patterned SOI, where the trenches are filled with oxide. This time, there are clear delamination areas, which could cause complete bond failure later in the process. The origin of this delamination can be found at the waveguide edge. During the processing it is not possible to get the SOI chips completely planar. The trenches are filled with a thick oxide which is then polished back using CMP. The stopping point is the silicon nitride hard mask of the waveguides which is then removed. This creates a step of approximately 15 nm, greatly diminishing of the strength of the van der Waals forces. Finally adding a thin DVS-BCB layer planarizes this unintentional step. It can be seen in figure 2.35d that there is no delamination. Note that figure 2.35d also shows extra oxide between the BCB and the Si device layer. This was deposited with PECVD for a project-specific reason and does not impact our analysis here.

2.4.3 Conclusion

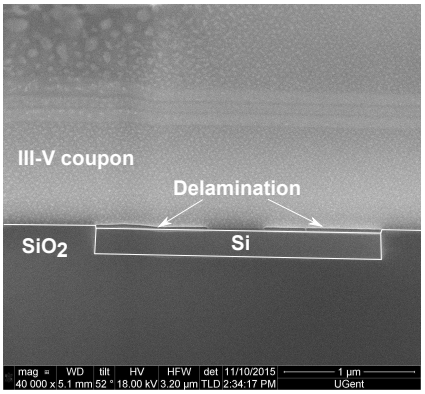
In this section we have discussed the development of transfer printing of III-V onto silicon for optical components. The III-V is used very efficiently, as it is only provided where needed. Because of the small size of the coupon, we do



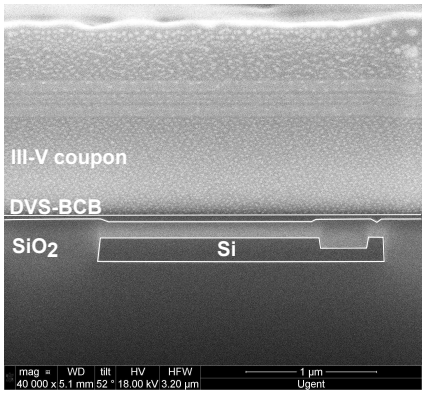
(a) Unpatterned SOI



(b) airclad patterned SOI



(c) Patterned SOI with oxide in the trenches



(d) BCB coated patterned SOI (with extra oxide)

Figure 2.35: FIB cross sections of coupons printed on various target substrates

not sacrifice Si design freedom for the freedom of epitaxial layer stack, while the multiple posts on the stamp ensure a high throughput process.

TP relies on the kinetic response of the adhesion strength of the coupon to the elastomeric stamp. PDMS is a viscoelastic material and therefore the adhesion energy to the coupon increases with increasing speed of stamp movement. If the speed is greater than the critical speed, the coupon is picked up from the source wafer. Next the coupon is printed on the silicon PIC by retracting the stamp with a speed below the critical speed. Because normally the adhesion between the source substrate and the coupon is too high, released free-hanging structures have to be fabricated.

A fabrication process was developed where a sacrificial release layer is etched away while polymer tethers ensure the coupon does not float away by anchoring the coupons to the substrate. For the identification of the release etchant, several chemical solutions were tested. In the end, cold aqueous iron(III)chloride was identified as the best, having a selectivity of over 2000. This means that microns of the InGaAs release layer can be etched horizontally, while only nanometers of the InP bottom device layer was removed. Sideways etching of the layers inside the coupon was avoided by using a SiO_2 - photoresist encapsulation, deposited by PECVD, covering the sidewalls of the coupon. The growth has to be of a high enough quality to avoid defects in the coupon. The tethers were optimized so they would easily break, while at the same time sturdy enough for the release etch. Especially the duty cycle of the tether framework was important. If it was more than 50 % (more tether than free space), a zig-zag etched pattern could be discerned in the bottom of the the device layers. Finally also the bond interface was optimized. Since during the device fabrication we will use aggressive etchants, the bond has to be strong enough. Moreover, as this is an optical interface, we can only use very thin bond interfaces. Relying on Van Der Waals forces was an option for the unpatterned SOI as well as the airclad patterned SOI. When the trenches are filled with oxide however, clear delamination can be observed. Spinning a 30 nm thin BCB layer solved this issue.

Finally, we can release InP-based coupons reliably and print them onto planarized SOI. Once this is done, we can fabricate our devices in the small sheet of material transferred.

2.5 Overview and comparison

Photonic circuits need multiple band gaps. InP photonic platforms mostly use regrowth techniques to achieve these. Regrowth is however not an option for heterogeneous integration of III-V to Si as the surface of the III-V does not remain planar. We have developed three different, attractive approaches to integrated multiple bandgap III-V material on a silicon PIC in this chapter.

First of all, quantum well intermixing allows one to blueshift the quantum wells at certain locations of the InP die. Relying on lithography allows implanting phosphorous atoms with great precision. These phosphorous atoms then act as catalysts for disordering when they are driven into the quantum wells. Using an empirical relation and a mathematical model, we could show that the diffusion coefficient for atom disordering of the two layer stacks used in this work were significantly different. We attributed this to the growth quality, where parasitic interstitial defects can aid the diffusion. After establishing the fabrication process, we made lasers to assess the quality of the intermixed quantum wells. Fabry-Perot lasers showed very similar results, where the blueshifted ones would even outperform the as-grown ones. Hence, the QWs remain of high quality.

Second, we looked at multiple die bonding. On first sight this is the most straight-forward option, since the die-to-wafer bonding process has matured now. We overcame the height difference of the two InP substrates using a graphite sheet, and used conformal oxide to fill the gap between the two bonded dies. The dimension of this gap was brought below 100 μm for both molecular and BCB bonding. The conformal oxide opened in a self-aligned manner, protects the silicon waveguides during the post-bond process and also fills the waveguide trenches. Since we are using airclad trenches for the molecular bonding, HCl-based etchants can travel underneath the die and attack the bottom device layer, causing so-called wicking. By using the conformal oxide and having a good cleave this wicking was reduced from $\pm 700\mu$m to zero. Finally, we also combined multiple die bonding with quantum well intermixing, in order to create four III-V band gaps on a Si PIC.

Last, transfer printing was developed for photonic applications. In transfer printing a rubber stamp is used to pick and print a III-V coupon to a Si PIC. In order to be able to pick up the coupons, the coupon has be made free-hanging. To do this, first the correct release etchant needed to be identified. As it needs to undercut multiple microns but can only etch nanometers vertically, selectivity of over 1000 was needed. We settled on cold $FeCl_3$:H_2O, which could undercut our coupons completely with a selectivity of over 2000. The encapsulation needs to be adequate, as otherwise the device layers would also be etched during the iron(III)chloride etch. Therefore, a combination of SiO_2 and photoresist was used. The tethers were photoresist and break easily during the picking. They cannot be too wide, as then the release etchant does not have enough access points. Although for unpatterned SOI and airclad patterned SOI we could rely on Van Der Waals forces to have a proper bond, we saw delamination for planarized SOI,our platform of preference. A thin layer of BCB at the bonding interface resolved this issue, and a strong bond was established that could withstand the further processing.

Comparing these three techniques, we would first of all like to state that QWI can be combined with the other two. QWI and multiple die bonding in particular

are very complementary. The former allows you to create multiple band gaps within a spatial resolution below 1 μm, while the latter provides you a lot of design freedom for the different layer stacks. The combination of QWI and TP is also an option, but the benefits are lowers. TP already has a decent spatial resolution, and therefore the argument of resolution drops from the equation. One advantage of QWI is that the different band gaps can be created within one III-V mesa. This means that you do not have to couple down to the silicon waveguide and back up again. The main disadvantage is the tricky fabrication. Especially the high temperature anneal can be a cause for major concern and can completely ruin the sample.

When comparing multiple die bonding and TP, the latter would come on top for most projects. The minimum size of the bonded dies of a few millimeters limits the design freedom considerably. E.g. suppose one wants to make a laser - EAM modulator - detector circuit. The laser and detector can be in the same material, while the EAM modulator is a different one. If one bonds two dies, U-shaped circuits need to be made, where two devices are spaced at least millimeters apart. Transfer printing does not have this constraint. Multiple die bonding can be interesting though if the density of the III-V devices is still very high. It might be preferable to work around the design constraints than to do the extra release processing.

References

[1] Siddharth Jain. *Hybrid Silicon Photonic Integration using Quantum Well Intermixing*. PhD thesis, University of California Santa Barbara, 2013.

[2] Shinji Matsuo, Takuro Fujii, Koichi Hasebe, Koji Takeda, Tomonari Sato, and Takaaki Kakitsuka. *Directly modulated buried heterostructure DFB laser on SiO 2/Si substrate fabricated by regrowth of InP using bonded active layer*. Optics express, 22(10):12139–12147, 2014.

[3] Dennis Glenn Deppe and N Holonyak Jr. *Atom diffusion and impurity-induced layer disordering in quantum well III-V semiconductor heterostructures*. Journal of applied physics, 64(12):R93–R113, 1988.

[4] Ju Han Lee, Sang Kee Si, HB Moon, Eui Joon Yoon, and Sung June Kim. *Bandgap tuning of In 0.53 Ga 0.47 As/InP multiquantum well structure by impurity free vacancy diffusion using In 0.53 Ga 0.47 As cap layer and SiO 2 dielectric capping*. Electronics Letters, 33(13):1179–1181, 1997.

[5] Andrew McKee, CJ McLean, Giuseppe Lullo, A Catrina Bryce, Richard M De La Rue, John H Marsh, and Christopher C Button. *Monolithic integration in InGaAs-InGaAsP multiple-quantum-well structures using laser intermixing*. IEEE journal of quantum electronics, 33(1):45–55, 1997.

[6] Sylvain Charbonneau, Emil S Koteles, PJ Poole, JJ He, GC Aers, J Haysom, M Buchanan, Y Feng, A Delage, F Yang, et al. *Photonic integrated circuits fabricated using ion implantation*. IEEE Journal of selected topics in quantum electronics, 4(4):772–793, 1998.

[7] Matthew N. Sysak. *Monolithically integrated wavelength converters using a dual quantum well integration platform*. PhD thesis, University of California Santa Barbara, 2005.

[8] Ferenc Riesz, L Dobos, C Vignali, and C Pelosi. *Thermal decomposition of InP surfaces: volatile component loss, morphological changes, and pattern formation*. Materials Science and Engineering: B, 80(1):54–59, 2001.

[9] Géza Kurczveil. *Hybrid Silicon AWG Lasers and Buffers*. PhD thesis, University of California Santa Barbara, 2012.

[10] Matthew A Meitl, Zheng-Tao Zhu, Vipan Kumar, Keon Jae Lee, Xue Feng, Yonggang Y Huang, Ilesanmi Adesida, Ralph G Nuzzo, and John A Rogers. *Transfer printing by kinetic control of adhesion to an elastomeric stamp*. Nature materials, 5(1):33–38, 2006.

[11] Matthew Alexander Meitl. *Transfer printing and micro-scale hybrid materials systems*. PhD thesis, University of Illinois at Urbana-Champaign, 2007.

3

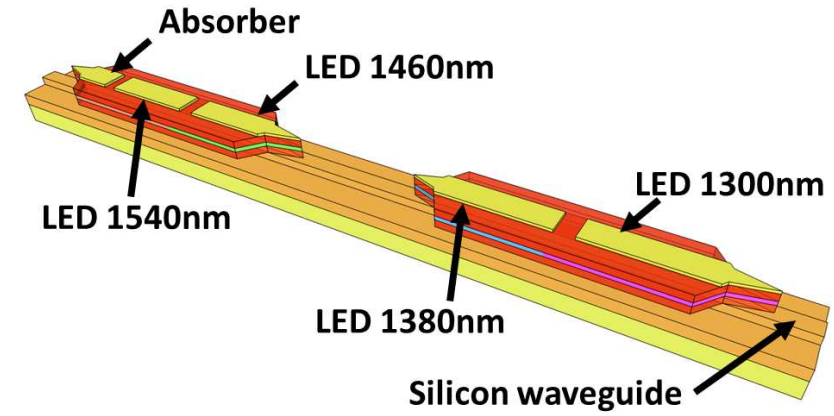
Electrically pumped multi-bandgap SLEDs

3.1 An SLED consisting of four active regions

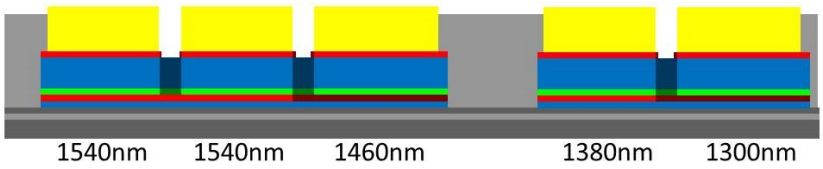
The goal of this project is to create a superluminescent LED which spans over the O, E, S and C band of the optical spectrum. In other words, we aim for a light source on silicon with a bandwidth of 300 nm, ranging from 1260 nm to 1560 nm. Since a typical III-V epitaxial layer stack covers 70-80 nm, the most prominent question is how to span four times that number. We tackled this problem with a two-step solution.

First of all, two different layer stacks were bonded. This multiple die bonding allows us to create two light sources with an epitaxy of choice, which should then be combined to create a broadband LED. We retain complete freedom over the quantum well design, as long as one can bond the stack. Obviously, in order to process them at the same time, they have to be similar. In order to do bonding, a minimum III-V die of several millimeters by millimeters is required. This means the spatial resolution of multiple die bonding is limited and, keeping a reasonable silicon footprint, the number of bonded dies is limited. Therefore, we applied quantum well intermixing to each layer stack creating two bandgaps within each die. In this manner the footprint of the device is acceptable, while we manage to create four different band gaps on chip. Both techniques are profoundly discussed in chapter 2.

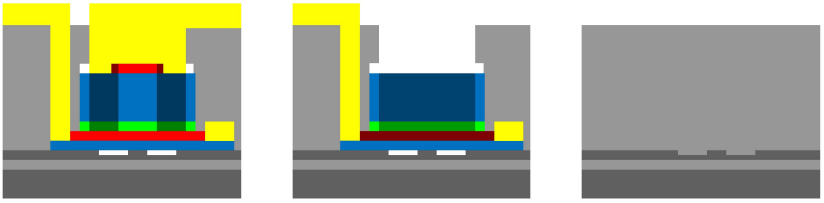
To combine the different spectra, we investigated a series design of four active



(a) 3D illustration



(b) Sideview



(c) Cross section of the mesa

(d) Cross section of the isolation section

(e) Cross section of the gap region

Figure 3.1: Illustration of the broadband SLED. The dark gray depicts silicon, the light gray silicon oxide. InP is blue, while the active region and SCH are red and green. The shade indicates the effect of the proton implantation and the metals are yellow.

regions, depicted in figure 3.1. Each broadband SLED consists of two subdevices, one in each epitaxial layer structure. Each subdevice in turn has two different sections, one intermixed and one with the as-grown band gap. We have organized the band gaps such that each section is transparent to the previous, i.e. the red side of the spectrum travels through the blue side. There is an extra absorber section for the longest wavelengths, to prevent lasing. Figure 3.1b clearly show the location of the different band gaps. This is a monodirectional design where light is propagating from the left to the right. Generated light traveling to the left

(e.g. from the 1300 nm region to the 1380 nm region) will be absorbed. The absorbed photon at 1300 nm gives rise to an excited carrier which may give rise to useful light at 1380 nm propagating to the right and is therefore not necessarily lost. Figure 3.1c and 3.1d show a cross section of the mesa in a pumped region and in one of the isolation sections in between two metals. We used a wide mesa to avoid sidewall losses and recombination. To confine the carriers, we used a proton implantation. Figure 3.1e shows the gap region, which is covered with the protective oxide as described in section 2.3.2.

This work was carried in close collaboration with the Optoelectronics Research Group of prof. Bowers of the University of California Santa Barbara (UCSB). All fabrication was carried out in the UCSB cleanroom, both the pre- and post-bond process. In this respect, we want to thank Mike Davenport and Jon Peters for their fabrication insights, and Martijn Heck and prof. John Bowers for their input in the design phase.

3.2 Modeling and design of the multi-bandgap SLED

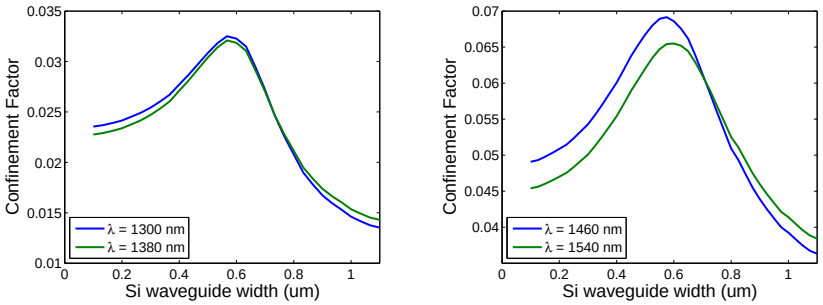
3.2.1 Mesa design and confinement

The bulk of the device consists of two mesas. To achieve efficient radiative carrier recombination in the active region, the mesa design must combine a low non-radiative recombination with good current confinement.

The confinement factor depends on many different parameters. Obviously the layer stack itself plays a huge role, but also the mesa width, silicon waveguide width, thickness and etch depth, etc. The parameter space was not explored completely by simulation, because we rather wanted to rely on experiments previously carried out in the group [1].

The layer stacks were already depicted in section 2.2 in tables 2.2 and 2.3. As the results on the quantum well intermixing were extremely satisfying for us, we preferred not to change the active layers to achieve marginal gain in confinement factor.

The silicon waveguide width is a parameter easy to change in the heterogeneous integration platform. The confinement factor in the pumped QW active region (a 4 μm wide current channel) was simulated for the different wavelengths and layer stacks by using a mode solver *Fimmwave* from *Photon Design*. The silicon device layer was assumed 500 nm thick and etched 250 nm. The mesa width was chosen 24 μm, where the confinement is calculated in the center 4 μm as that is the region being pumped (confined by the proton implantation). Since the lateral optical confinement is ensured by the silicon waveguide, there is an optimal silicon waveguide width, around 600 nm. Three different mode profiles are shown



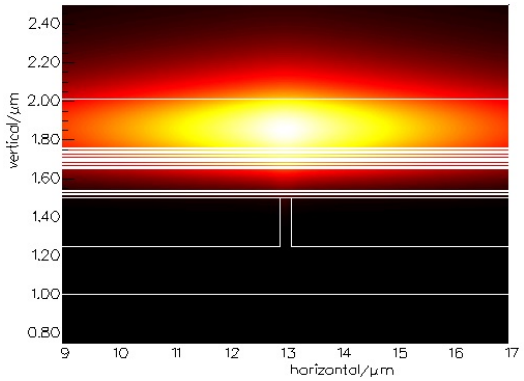
(a) Confinement factor in layer stack designed for 1380 nm.

(b) Confinement factor in layer stack designed for 1540 nm.

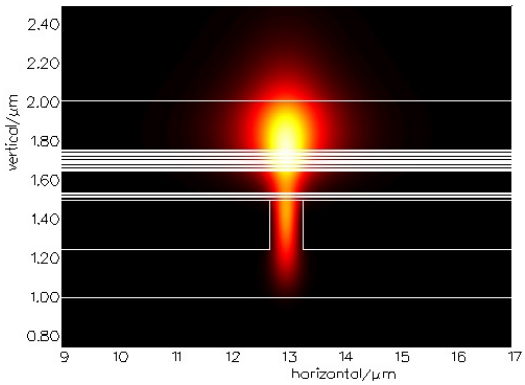
Figure 3.2: Simulation of confinement factor in the pumped QW active region versus silicon waveguide width

in figure 3.3. In figure 3.3a the silicon waveguide is very narrow and the mode is pushed up. Because of the lack of lateral confinement, the fundamental mode is much wider than the 4 μm wide current channel. In figure 3.3c on the other hand, the 1 μm wide high index silicon waveguide pulls the mode down, leaving only an evanescent tail in the active region. Figure 3.3b shows the sweet point of a 600 nm width. It can still be seen the mode mostly resides in the upper SCH layer. Because of the quantum well intermixing, we prefer not to introduce a lower SCH layer however.

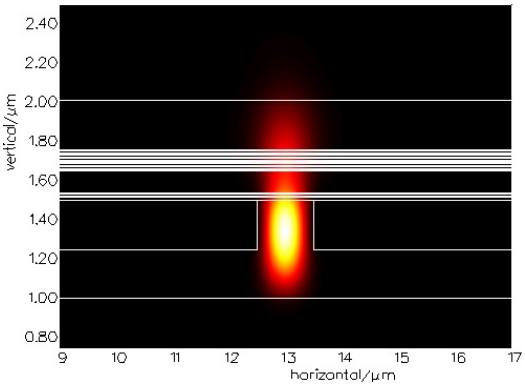
To avoid excessive sidewall recombination, a wide mesa of 24 μm is used. The mesa is implanted with protons to create a 4 μm current channel. Although many modes are present in the mesa, the fundamental one is pumped most efficiently due to a supreme overlap. The proton implant profile was simulated using *SRIM*, taking care that the total ion density remains higher than 2.5×10^{18} cm^{-3} in the regions that need to be highly resistive [2]. In the QWs, the implant will lead to non-radiative recombination centers and thus the ion concentration should be much lower. Because of the difference between the stacks, i.e. the sacrificial InP - InGaAs (200 nm - 50 nm) layer pair still present on one of the two bonded dies, there is a tradeoff to be made here. To minimize channeling of the ions, the implant was performed at a 7° angle with respect to the surface normal. The table of implantation energies and doses can be found in table 3.1. All these implantation steps are done consecutively, the resulting implant profiles are shown for both layer stacks in figure 3.4. Note that the total ion density (depicted by the fat blue line) is above the critical concentration in the P-doped layers and falls rapidly in the quantum wells. To form the current channel, the p-metal is used to mask the implant. Also the sides of the mesa have to be protected though, as an implant of



(a) Si waveguide width of 200 nm.



(b) Si waveguide width of 600 nm.

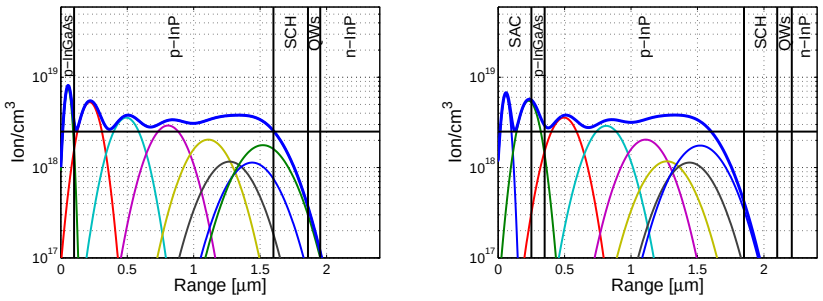


(c) Si waveguide width of 1 µm.

Figure 3.3: Mode profiles of the fundamental mode in the C-band stack at 1540 nm. The influence of the silicon waveguide width is highlighted.

Implantation energy	Implantation dose
5 keV	5×10^{13} cm^{-2}
25 keV	1×10^{14} cm^{-2}
65 keV	1×10^{14} cm^{-2}
110 keV	1×10^{14} cm^{-2}
150 keV	8×10^{13} cm^{-2}
170 keV	5×10^{13} cm^{-2}
190 keV	5×10^{13} cm^{-2}
200 keV	8×10^{13} cm^{-2}

Table 3.1: Implantation energy and dose of the proton implantation



(a) Proton implantation profile in the layerstack from table 2.2.

(b) Proton implantation profile in the layerstack from table 2.3.

Figure 3.4: Proton implantation profiles in which different layers are indicated, as is the critical concentration above which the electrical resistance increases. The different curves indicate different implantation steps with the energies and doses outlined in table 3.1. The fat blue line depicts the overall ion density.

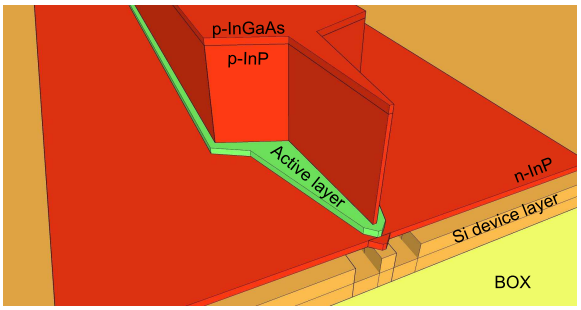


Figure 3.5: Illustration of the taper going from the III-V to the silicon waveguide.

the n-InP layer will cause the sheet resistance of the n-contact to be huge. This causes two parallel current paths which run parallel to the desired one. Since the highly doped p-InGaAs is conducting current efficiently even with the high proton density, electric current can be fed from the p-metal to the parasitic channels. Because of this, we etched the InGaAs layers on the sides of the mesa, disconnecting the parasitic channels.

As more than one band gap will reside in our mesa, due to the quantum well intermixing, we also need to be able to contact the different sections separately. To isolate these, we also implanted the 20 μm long section in between the two LEDs. In addition, we etched the p-InGaAs, providing us with an isolation of 30 kΩ. The optical losses in this section can be neglected, as we have intermixed it. Therefore, it is transparent to the light traveling through in our monodirectional SLEDs.

3.2.2 III-V to Si tapers

After light is generated in the quantum wells and has been coupled to the hybrid Si–III-V mode, it has to couple down to the silicon waveguide mode. On top of that, the light generated in the first subdevice must do this transition twice again to travel through the second subdevice (see figure 3.1a). The taper is illustrated in figure 3.5 and actually consists of a series of different tapers. The p-InP first quickly tapers down from 24 μm wide to 4 μm, while at the same time the active region goes from 26 μm wide to 16 μm. This first taper is only 5 μm long and hardly has any optical effect. In a second stage taper, the p-InP goes from 4 μm to 500 nm, with the active region narrowing from 16 μm to 3 μm. In this 20μm long second stage taper the silicon waveguide widens also from 600 nm to 2 μm. In the third stage, the p-InP has stopped and the quantum wells shrink to 500 nm over a length of 1.2 μm. The last step is the tapering of the n-InP, which so far has stayed wide. It transitions from 3 μm wide to 500 nm, all the while the silicon waveguide stayed 2 μm broad.

Figure 3.6 shows a simulation of the transmission through the various III-V

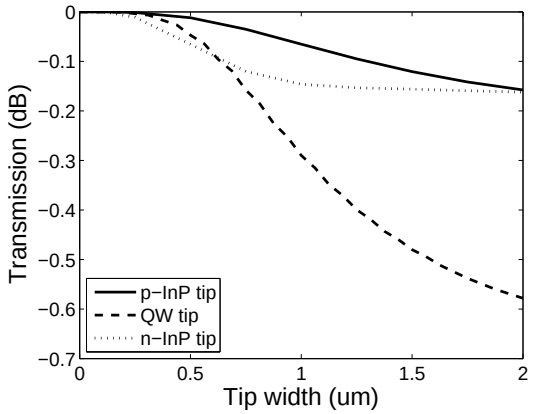


Figure 3.6: Simulation of the transmission through the taper tips at 1540 nm.

taper tips (the wavelength was 1540 nm). It is clear that the most critical tip is the one of the quantum well and SCH layer (indicated by QW tip in the legend). Even so, the loss in transmission is low for all tip widths. On top of that, as described in the process flow in section 3.3.2, the QWs will be wet etched and therefore the taper tip will be narrower than defined on the mask. Figure 3.7 shows a simulation of the transmission through the tapers themselves as function of their length, assuming all taper tips to be 500 nm wide. The adiabatic regime of the p-InP taper is obtained from a length of about 60 μm. Once it is longer than 20 μm however, the variations are less than 0.5 dB. The dashed line represents the QW taper, the transmission of which varies from -0.35 dB to -0.2 dB as the length increases from 0 to 5 μm. If this taper would be 5 μm long however, the tip would no longer be pumped. The quantum wells are intermixed in the taper, meaning that the light originating from the original band gaps will not experience absorption loss. The blue-shifted light will be absorbed by an unpumped taper though. Therefore, we limited this length to 1.2 μm. It can be seen that the n-InP taper has almost no influence whatsoever.

Remembering that there will always be a lot of non-radiative recombination at the sidewalls, it is very important to keep the taper as short as possible. We know these devices may suffer severely from sidewall recombination, as that is why we chose to work with a 26 μm wide mesa. We designed the p-InP taper to be 20 μm long and the QW and n-InP ones 1.2 μm long, the behaviour of which is shown in figure 3.8 for various wavelengths. The loss is below 1 dB for the entire spectrum. The blue side of the wavelength range only loses 0.13 dB of power and the red side 0.85 dB. The fall-off towards the red side cannot be attributed to one contribution alone, illustrated by figure 3.9 which is showing the power in the fundamental

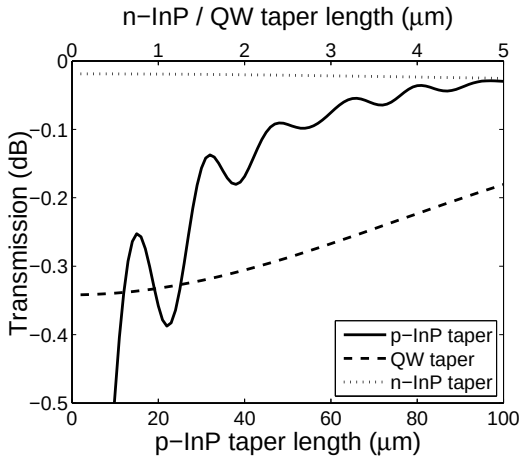


Figure 3.7: Simulation of the transmission at 1540 nm through the various tapers as a function of taper length.

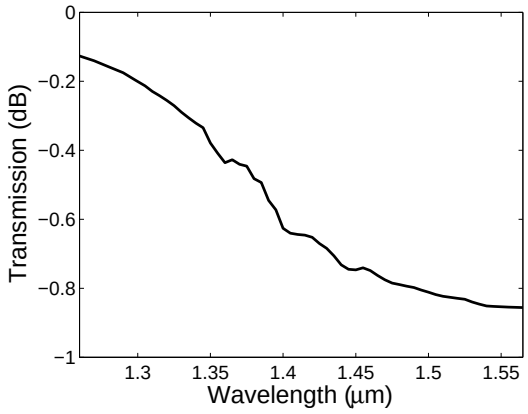


Figure 3.8: Simulation of the transmission through the complete taper as a function of wavelength.

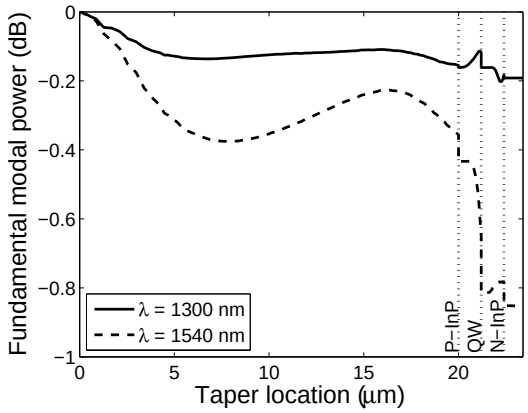


Figure 3.9: Power in the fundamental mode throughout the complete taper for 1300 nm and 1540 nm.

mode throughout the taper at 1300 nm and 1540 nm. For both wavelengths the P-InP taper is not completely adiabatic, as we already established in figure 3.7. Also the interfaces are not completely ideal, especially for the longer wavelength. The extra loss of the red side of the spectrum cannot be attributed to one interface or taper in particular, but is rather spread out over many contributions. We could alleviate this by making the tapers longer and the tip narrower, but as the taper is acceptable as it is and can be fabricated reliably this taper was elected best. Also, from other projects, measurements have shown it to work better than the longer tapers [3].

3.3 Post-bond fabrication on silicon

3.3.1 Mask design: Test structures

The first steps in fabricating such structures is to design a proper set of lithographic masks. Figure 3.10 shows an image of the full mask layout including the silicon layers (pink and red) and the III-V layers (other colors). Except for the SLEDs, the mask also contains a number of test structures:

1. **Alignment markers for silicon and post bond processing:** We use both global and local markers. The global markers are a skinny double cross etched in the silicon, whereas the local ones are INSITU markers, which employ the vernier effect. We have these INSITU markers both in the silicon and in the III-V. These markers are widely used on the *GCA Autostep 200* stepper system used for this project. Note that after bonding, all markers of

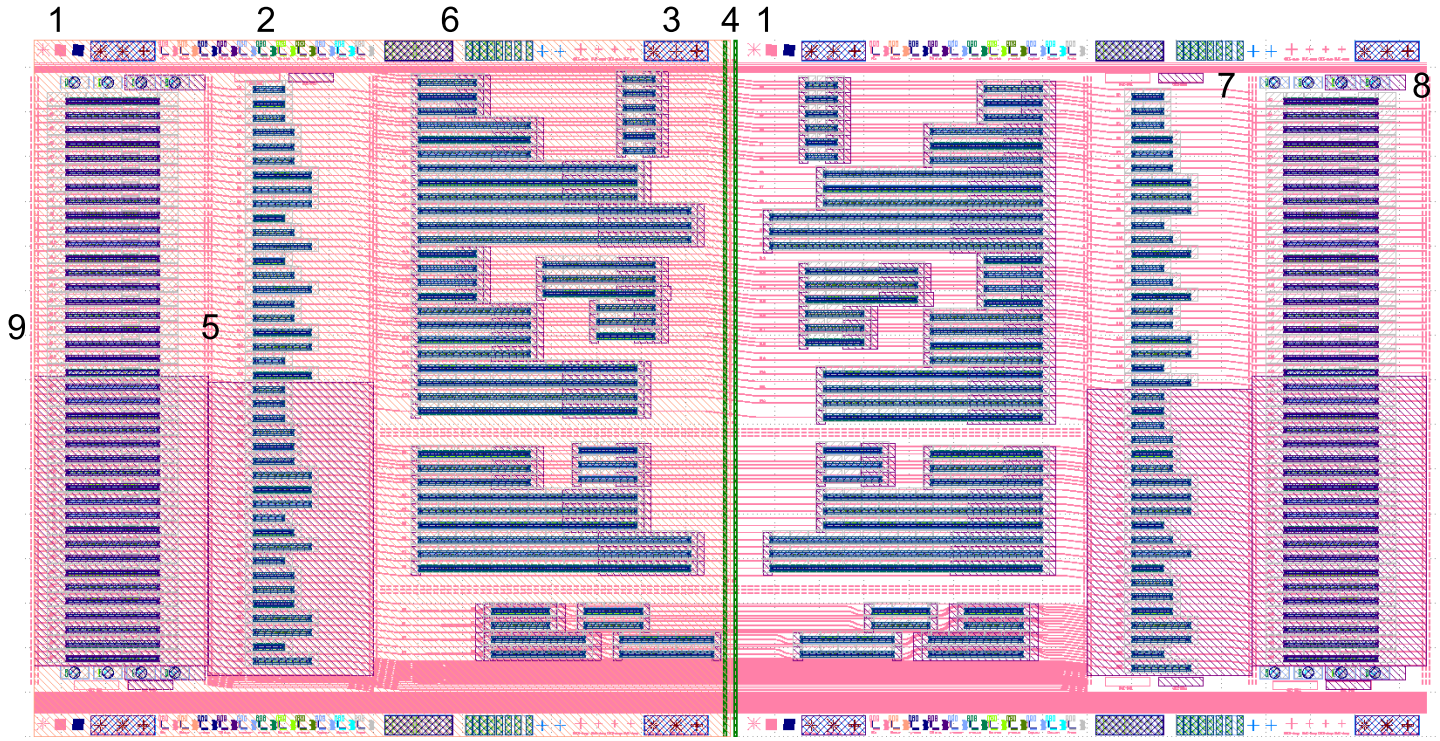


Figure 3.10: Full mask layout for the electrically pumped, broadband SLEDs, including the silicon layers (pink en red) and III-V layers (other colours).

the silicon are covered by the epi, meaning that they cannot be seen by the system. To overcome this, we define a matrix of 3 by 3 dies on the SOI and bond III-V on the middle one. To align, we step through the 8 uncovered dies and estimate the exact position of the middle die.

2. **Vernier structures for post bond processing:** The alignment of every lithographic exposure is checked using vernier markers. In this manner, we can detect up to 100 nm misalignment. Typically we achieved under 200 nm misalignment.

3. **Alignment markers for bonding:** The bonding itself is also aligned, because both the III-V and SOI are patterned. As explained in section 2.3.3, they need to be close to, but at a safety margin of the gap region. We used similar markers as the global ones mentioned above, but also a thicker simple cross. This redundancy was put in, because the aggressive III-V etches before bonding (removing the buffer layer and its etch stop) may damage the skinny markers. These markers are also the alignment markers for the III-V processing. Because the bonding precision is 2 μm there was no need for INSITU markers.

4. **Cleave lines:** To limit the distance between two mesas from the two bonded dies, we need to cleave the III-V just prior to bonding. Otherwise the edge bead of the photoresist prevents pushes the proximity of two mesas to around a millimeter. Because the location of the exact cleave is very important - to limit the width of the gap between the bonded dies and to estimate the bonding precision - clear cleavelines are laid out. Note that these act as guidelines as a cleave allways occurs along crystal planes rather than lithographic features. As the cleaveline is lower than the bonding surface, itself will not be bonded. To prevent long unbonded regions, a dashed line is preferred over a full one.

5. **Dice lines:** After the processing, the sample is diced and polished. To avoid mistakes beyond repair, we have put clear dicelines, indicating both the width of the dicing blade and the location. To prevent traveling etchants, we have made also these lines dashed.

6. **TLM structures:** Metal contacts may be the most common reason optoelectronic devices fail and therefore having sufficient test structures is important. The TLM (transmission line method) test structures allow us to measure the contact resistivity and sheet resistance of both n- and p-contact in detail. The distance between two contacts varies from 4 to 50 μm.

7. **PL test structures:** We work with four different band gaps, all of them may have shifted (un)intentionally during the processing. We designed various

locations to do photoluminesce (PL) measurements if needed. In this manner we can always confirm that we did not do some mistake in the processing (e.g. we could have swapped the two dies during bonding).

8. **Vertically emitting diodes:** These test structures allow us to eliminate to coupling to the fundamental waveguide mode and following III-V to Si taper. Measuring these structures confirms electric current is injected and light is generated.

9. **Laser test structures:** To confirm the quality of the complete process, we have made Fabry-Perot lasers as test structures. These characteristics can be compared to other, identical lasers processed in the research group. Also possible deterioration of the material due to the QWI was characterized in this manner.

Because the III-V die is flipped during the bonding, we need to mirror the masks of the quantum well intermixing in one dimension. It is important to subsequently also flip the die over this dimension during bonding.

3.3.2 Process flow

In this section we discuss the processing of the silicon chip prior to bonding, as well as of the entire sample after bonding. As the III-V dies are being intermixed, they have already seen their share of processing at the time of bonding. For this, we refer to section 2.2.2. We will be making devices on both bonded dies at the same time. As can be seen from tables 2.2 and 2.3, there are a few subtle differences in layers other than the quantum wells. This calls for extra processing steps or at least more complicated ones where this would otherwise not be necessary. Figure 3.11 illustrates the post-bond process.

1. **Silicon waveguides:** We start the process with unpatterned SOI with a silicon device layer of 500 nm. The buried oxide (BOX) is 2 μm thick. A photoresist mask is applied, followed by a dry etch in the deep reactive ion etcher (DRIE). This provides us with clean, vertical sidewalls. A dummy is always used to estimate the exact etch speed, as we have to stop blind after 250 nm. Luckily, this 250 nm is not too critical as no grating structures or so are being used. We repeatedly etched within $\pm$ 15 nm of the target.

2. **Vertical channels:** Vertical channels are squares etched through the silicon device layer into the BOX throughout the sample. During the bonding they harness the water vapor that is being formed. Without these channels, the water cannot escape and causes delamination. The etch is again done in the DRIE.

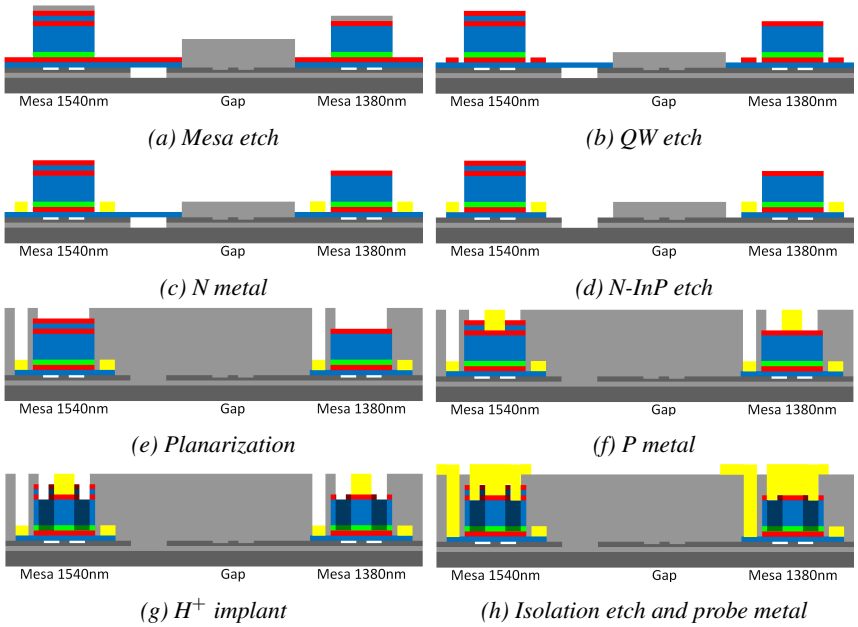


Figure 3.11: Overview of the different steps of the post bond process.

3. **Bonding:** Prior to bonding the silicon is thoroughly cleaned with a heated H_2SO_4:H_2O_2 3:1 piranha solution. After this, the Si sample is dipped in BHF to remove the native oxide. This dip also enlarges the vertical channels. The InP samples have been cleaved along the dedicated cleave lines and the protective photoresist can be removed. The alignment marks have been protected with a SiN hard mask, which can now be removed with a BHF dip. Both the InP samples and the SOI can now be activated with an oxygen plasma. During this step, hydroxyl radicals are formed on the surface of all samples. This is discussed more into detail in section 1.2.2.1

   Using the flip chip bonder, we bring the III-V dies in close contact with the SOI one after another. Special care is taken in order not to disturb the already positioned die when placing the second one. It is important to flip the dies over the correct dimension, in correspondence with the mirroring of the masks. Next, a graphite sheet is used to alleviate any thickness variation and a graphite fixture applies the bonding pressure. The whole is annealed for 1 hour at 300 °C. During this anneal the initial, weak bond based on Van Der Waals forces is converted into a strong covalent bond. When this covalent bond is established, water vapor is formed, which is channeled to the buried oxide through the vertical channels. After unloading the sample from the graphite fixture, it is thoroughly rinsed in acetone, isopropanol and

water to remove any graphite sheet residue.

As the waveguides between the two bonded dies will be damaged during the rest of the processing if unprotected and as the edge of the dies will be attacked because of traveling etchants, a protective 1 μm of SiO_2 is deposited with PECVD. Next, crystal wax is generously applied on the SOI, paying close attention to the gap region. This crystal wax protects the chip from particles during the subsequent mechanical grinding. The III-V substrates are thinned until 60 μm is remaining. In order to monitor this, we made sure one part of the SOI is not covered with crystal wax. This should be an unimportant part however, as it will be covered with particles after the polishing. After removing the crystal wax, the substrate removal is finished with a wet etch in hydrochloric acid. Because the wet etch is anisotropic, so-called mountains are formed parallel to the minor flat of the InP wafer, i.e along (011) crystal plane. To remove these, we first spin protective photoresist and then remove them mechanically with a razor blade. During the QWI process, we made sure the orientation of the sample was such that the mountains now appear on the top and bottom of the bonded die and not along the gap region. The size of the InP die was chosen such that the mountains do not lie above vertical channels. Because of this, the bond is very poor at these location and not a lot of force is needed to remove the mountains. After stripping the photoresist, the sacrificial etch stop layer is removed in diluted piranha and the bonding precision is measured.

4. **Mesa definition:** A 100 nm thick SiO_2 hard mask is used to define the mesa. Etching is done in a RIE using methane, hydrogen and argon. A laser trace is used to monitor the layers etched. The goal is to etch through the InGaAs contact layer and p-InP and stop in the SCH layer. For the O-band layer stack this SCH is an InGaAlAs layer, for the C-band stack, this is an InGaAsP layer. The latter also has an extra InP-InGaAs sacrificial layer pair on top of the p-contact layer. This is only so because the O-band stack was ordered before the benefit of the extra layers was established. Upon reordering of the epi, one would include the sacrificial layers on both stacks. These extra layers mean that the SCH of the O-band stack will be reached much earlier than the other one, but the etch rate in InAlGaAs is far slower than in InGaAsP. Therefore, we monitor the C-band die and stop when the etch is sufficient there. The fast etch rate of InGaAsP means that we do not have a lot of margin to stop in. On the one hand we have to make sure that we are through the p-InP layer across the entire die, on the other hand we cannot do a large overetch. Figure 3.12 shows an etch test on a dummy sample. The different layers are indicated. Because of the small period change between the SCH and p-InP layers and the etching inhomogeneity, stopping high in

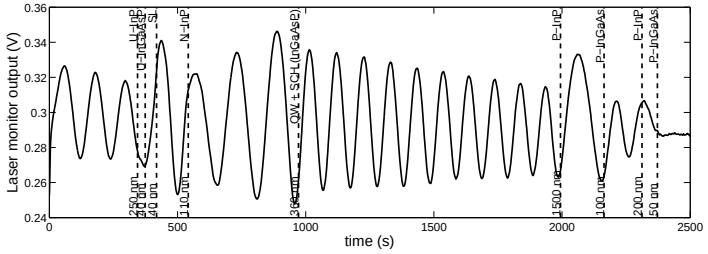


Figure 3.12: RIE etch test on a dummy sample. As can be seen, the period change between the SCH and p-InP is only fully clear when one sees at least a complete period. Note that in the normal process, one starts etching from the other side.

the SCH layer is not straight-forward.

5. **QW etch:** The QW etch is probably the most important step to define the III-V to Si taper. The two layer stacks differ significantly and the previous dry etch has done its share to enlargen the differences. Therefore, we process them separately, protecting one die with photoresist when etching the other. For the O-band stack, the SCH InGaAlAs layer is first etched with a H_3PO_4:H_2O_2:H_2O mixture while the active region is done with a diluted piranha mixture H_2SO_4:H_2O_2:H_2O. The C-band stack can be etched completely with the latter etchant.

6. **N-metal:** In stead of using a negative resist or image reversal for the metal liftoff, we used a bilayer photoresist process. The n-InP is still covered with the QW and SCH layers were the metal has to come. This ensures we are depositing the n-metal on a fresh surface for an improved contact. After lithography, the phosphoric and sulfuric etchants mentioned above are used to remove these sacrificial layers (now the two dies can be processed together) and the n-metal is evaporated. The metal stack used is Pd:Ge:Pd:Au 10:110:25:1000 nm.

7. **Island etch:** The remaining n-InP contact layer and supperlattice are removed in a short RIE dry etch identical to the mesa etch. A third, weak taper stage is added. In order to keep the losses low, the overetch is minimal.

8. **Planarization and via definition:** The chip is passivated by depositing a thick SiO_2 layer and vias are etched to expose the p-contact layer and n-metal. This etch is done in short steps, to minimize the over etch. In this manner, we limit the damage on the p-contact layer and prevent P and N to be shorted.

9. **Sacrificial etch:** As mentioned in the mesa etch, the C-band die has an extra InP-InGaAs sacrificial layer pair. This is there to protect the contact layer from any damage. Because the O-band stack was ordered before the benefit of the extra layers was clear, it does not contain them. Thus, we protect the O-band die and wet etch the sacrificial layers.

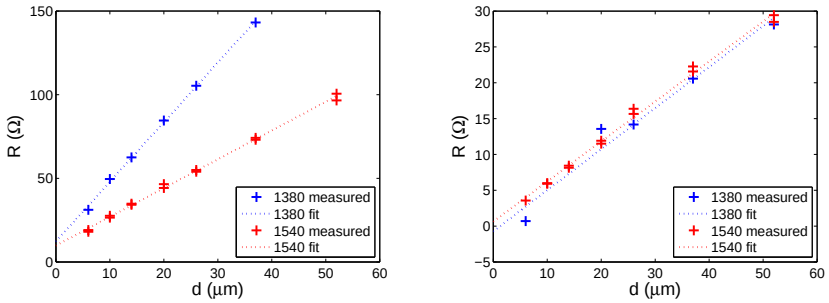
10. **P-metal:** A similar bilayer lithography as with the n-metal is used. A short diluted HCl etch is done to remove any native oxide and the metal can be deposited (Pd:Ti:Pd:Au 10:20:20:1500 nm). Both N and P contacts are then annealed at 350 °C for 30 seconds in forming gas. The p-metal only covers a 4μm wide strip on top of the mesa ridge. Only at the taper, the metal is wider, so as to contact it as good as possible. The oxide planarization and subsequent via etch make sure the device is not shorted. The p-metal acts as a hard mask in the next step and thus, by making the metal at the taper wider, we prevent electrically isolating the n-InP layer of the taper. Three different contacts are made on the mesa: one for the intermixed section, one for the as-grown section and one for an as-grown section acting as a back absorber to avoid lasing.

11. **Proton implant:** As we are using a wide mesa, we need to confine the current by proton implantation. Photoresist is applied to protect everything except the mesa top. A misalignment error needs to be taken into account and thus also the sides of the mesa top is protected. The proton implant is then performed as discussed in section 3.2.1. The p-metal acts as a metal hard mask, creating a self-aligned current channel under the metal.

12. **Isolation etch:** The different sections (intermixed, as grown and back absorber) still need to be electrically isolated further. The highly doped p-InGaAs contact layer is not affected sufficiently through the proton implant and must therefore be removed between the contacts. Also the parasitic current channels at the side of the mesa are disconnected by this etch. A large margin of 3 μm between the isolation opening and the p-metal is kept. If the undercut of the phosphoric acid based etchant would reach the metal, it would quickly be lifted off. In this manner an isolation of 30 kΩ is achieved.

13. **Probe metal:** Big probe pads are deposited with Ti:Au, for ease of use. They also act as a heat spreader.

14. **Dicing and polishing:** Finally, the sample is diced along the dice lines. Facets of optical quality are achieved by polishing carefully with a polishing pad with diamond particale size of 6 μm, 1 μm, 300 nm and 20 nm. Crystal wax was again used to keep the sample clean.



(a) P contacts *(b) N contacts*

Figure 3.13: Characterization of the contacts using TLM structures.

3.4 Characterization of the electrically pumped SLEDs.

3.4.1 Electrical contacts characterized

The transmission line method (TLM) allows us to characterize the electrical contacts and sheet resistance of the contact layers. The specific contact resistance and sheet resistance can be extracted by plotting the measured resistance for various distances in between contacts, as is done in figure 3.13 for the N and P contacts of both III-V dies. More in particular the following equation applies:

$$R_{\text{tot}} = 2R_c + R_{\text{sh}}\frac{L}{w} \tag{3.1}$$

where R_{tot} is the measured resistance (the probe resistance is calibrated out or a four point probe technique is used), R_{sh} is the sheet resistance, L is the distance between the two measured contacts and w is the width of the contact pad. R_c is the contact resistance, which is different for vertical and lateral conduction. In case of lateral conduction we write:

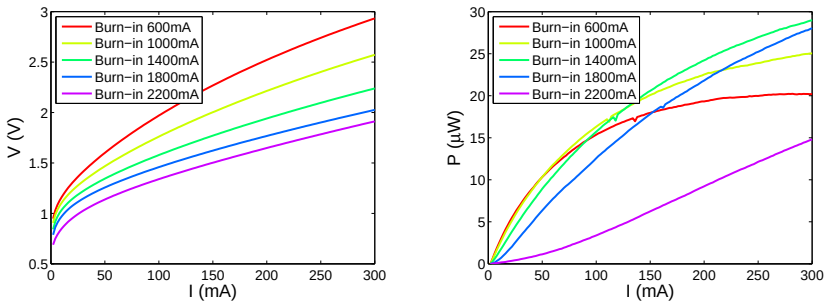
$$R_c = \frac{\sqrt{R_{\text{sh}} r_c}}{w} \tag{3.2}$$

where r_c is the specific contact resistance.

From the slope and y-intercepts of the straight line fits of figure 3.13, we extract the specific contact resistance and sheet resistance as noted in table 3.2. As can be seen in figure 3.10, there are both at the top and at the bottom TLM structures. In this manner we account for fabrication variation. The specific contact resistance is around $10^{-5}\Omega\text{cm}^2$ for the P-contact and around $10^{-7}\Omega\text{cm}^2$ for the N-contact, which are typical values of this fabrication process. Moreover, the two dies behave almost identical. The sheet resistance for the N-contacts is around 120 Ω / square for both dies, but the sheet resistance of the P-contact varies from 770 to 370 Ω / square. However, of the four characteristics, this one is the least important as the

	R_{sh} [Ω/square]	r_c [Ω cm^2]
O-band - P contact	771	2.3×10^{-5}
C-band - P contact	372	3.0×10^{-5}
O-band - N contact	124	5.0×10^{-7}
C-band - N contact	121	4.1×10^{-7}

Table 3.2: Sheet resistance and specific contact resistance of both III-V dies, as extracted from TLM structures



(a) The IV behaviour improves with higher burn-in current.

(b) When monitoring the output power, an optimal burn-in current can be identified.

Figure 3.14: Because of the tradeoff between optical losses and electrical resistance, there is an optimum in the local annealing current.

current is conducted vertically through the P-layers. Because of the use of a wide mesa, the sheet resistance of the n-InP should be dominating.

However, when testing the actual devices, the p-contact of the O-band stack showed to be dominating the IV curve. The resistance was much larger and not consistent throughout the chip. We attributed this to punch-through damage of the mesa dry etch through the hard mask. The C-band stack had a sacrificial layer pair, shielding the contact from harm. The devices near the edge of the die seemed to behave better than the others and as the TLM structures both are completely at the side, they could not provide extra insight, although the high sheet resistance might be attributed to this.

To improve the contact, we performed a local anneal. The current is slowly increased, allowing the contact to be annealed and the resistance to drop. This sintering process causes the metal to diffuse into the semiconductor material thereby improving the contact. If the diffusion is too severe however, the overlap of the optical mode with the lossy metal increases. Additionally, the contact cannot be assumed to be uniform anymore. Figure 3.14 shows this trade-off during the burn-in process. The same device was consecutively annealed at higher burn-in currents

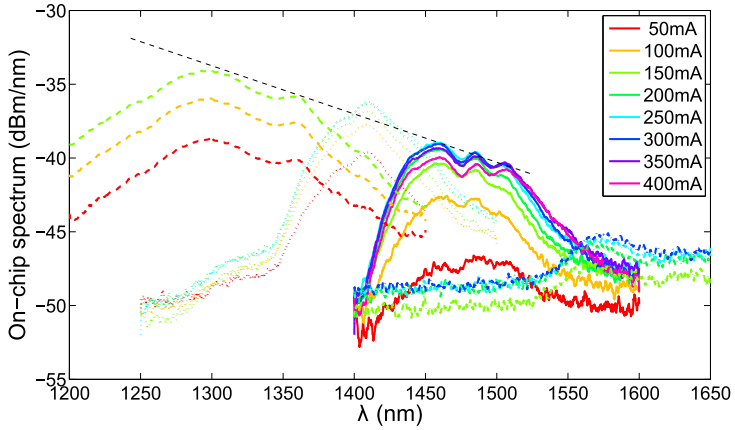


Figure 3.15: Spectra of the different sections pumped individually. The dashed, dotted, full and dash-dotted lines symbolize the different sections.

after which the voltage and output power were monitored. Note that the burn-in current was not applied all at once, but a slow ramp was needed to prevent blowing up the device. As can be seen in figure 3.14a, the resistance monotonically improves with stronger burn-in. Figure 3.14b on the other hand takes both electrical and optical behaviour into account and it can be seen that after an optimal current the device deteriorates. It was found that the optimal current was 1.750 A for a 1 mm long device. Even so, the performance is not as good as the TLMs might have indicated.

3.4.2 Evaluating the different active region separately

Because we have created different P-contacts for the sections emitting at different wavelengths, we can measure them first separately. After characterizing this, we manipulate the currents to level the spectrum.

Figure 3.15 shows the spectra of the different sections as measured through the same facet. Note that the longest wavelengths travel through all the other sections, as will be the case for the broadband SLED. The four different band gaps can clearly be discerned. We indeed created the correct band gaps at the correct locations. The power at the longest wavelengths is very low though. If we want to include this band gap, and if the spectrum of the separate sections is representative for the cumulation of the different sections, we have to bring the output of all the other sections to this level.

Furthermore, an angled line along the other three band gaps is added to figure 3.15, indicating higher losses for longer wavelengths. This can be explained by

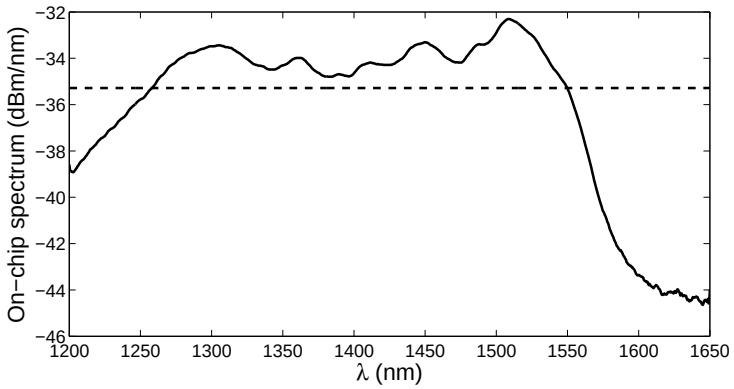


Figure 3.16: Spectrum when pumping the different sections of the SLED such that the 3 dB bandwidth is as broad as possible. The sections at 1300 nm, 1380 nm, 1460 nm and 1540 nm were pumped at 70 mA, 50 mA, 300 mA and 140 mA respectively. The dashed line indicates the 3 dB bandwidth.

linear losses, as the longer wavelengths have to travel further. For example, as is evident from figure 3.1a, the light from the LED at 1540 nm has to travel through the 1460 nm one, thereby experiencing linear loss from the hybrid mode as well as free-carrier losses. Then it tapers to the silicon waveguide with the corresponding coupling losses. In the passive silicon waveguide sidewall roughness induces propagation loss (2 dB/cm over a distance less than 2 mm). While in silicon, the light passes through the gap between the bonded dies, which underwent extra protection from processing as described in section 2.3.2. An imperfect protection can lead to extra losses in the silicon waveguide. The light now tapers to the second mesa, propagates through it and tapers back down.

3.4.3 Generating a broad emission spectrum

In figure 3.16 we pumped the different sections such that the 3 dB bandwidth is as broad as possible. The result is almost 300 nm of 3 dB bandwidth. More specifically, the bandwidth was 291 nm while the total power was -8.0 dBm. The flatness of the top is around 1.5 dB.

Looking more closely at the figure 3.16, one notices there is no light visible at 1540 nm but there is a peak at 1500 nm (which we did not design for). In figure 3.17 we varied the current of the section at 1540 nm, and the peak at 1500 nm is greatly affected by that. To verify the band gap, we can check the emission of the back absorber, also shown in figure 3.17. This absorber emits around 1540 nm, confirming our PL measurements. Upon close inspection a bump at 1540 nm can be noticed in the combined spectra. From this, we gather that the light

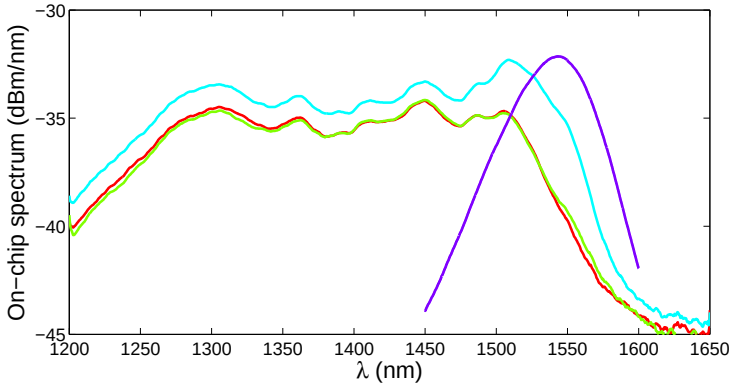


Figure 3.17: Spectrum when pumping the entire SLED as in figure 3.16. The injected current in the section at 1540 nm was 60, 100 or 140 mA. The emission from the back absorber is also shown and was normalized for illustration purposes.

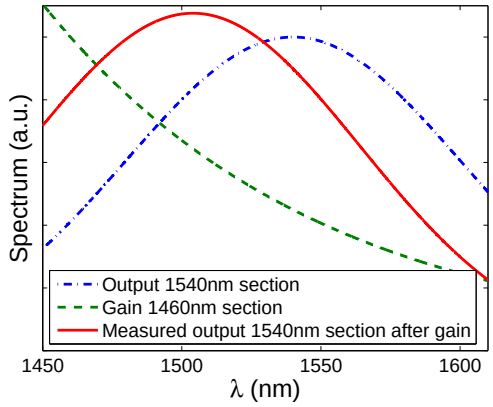


Figure 3.18: Illustration of how an asymmetrical gain profile can shift the light output from the 1540 nm to 1500 nm.

is being generated at the expected wavelengths, but the short side is amplified by the intermixed section. The long wavelength tail of the gain spectrum is very wavelength dependent and figure 3.18 illustrates how this type of gain will shift the measured emission peak. If we go back to figure 3.15, this also explains why the longest wavelengths are so weak. The gain experienced when pumping the intermixed section turns to loss when left unpumped and the emission is severely quenched.

3.5 Conclusion

In this chapter we proposed and realized a broadband III-V/Si SLED based on the combination of multiple die bonding and quantum well intermixing. The SLED consists of two mesas, each consisting of two band gaps. In this manner, we could create light in 4 band gaps, which are spaced 80 nm apart. The different active regions were combined in a series design where the longer wavelengths travel through the other sections.

We discussed the mesa and taper design. The mesa was kept wide to avoid non-radiative recombination, counting on a proton implant to confine the current. The confinement is optimized by choosing the correct silicon waveguide width. The tapers are deliberately made as short as possible, again to counter the sidewall recombination. In three stages, the hybrid mode is converted to a pure silicon mode.

The fabrication process and corresponding mask designs were laid out. We tried to highlight points of interest, while not going too far into detail. Several test structures were added to the mask, among which the TLMs showing decent contact resistances. Nonetheless, a local anneal had to be performed to improve the P contact on one of the dies. This contact was most likely damaged during the mesa etch.

Finally, we characterized the SLEDs, first separately and then as a whole. A 3 dB optical bandwidth of nearly 300 nm was achieved. This was limited on the red side by the long wavelength tail of the gain spectrum, shifting the emission from 1540 nm to 1500 nm.

In this chapter we opted to combine the light in a serial manner. The photons generated at the longest wavelengths travel through the other active sections, which makes it simple in terms of design. They are not directly absorbed, but still suffer from free-carrier absorption, linear losses etc. In a next generation, combining the light in a parallel manner should be looked into. This implementation could alleviate the above-mentioned losses. Of course, some coupling losses will have to be taken into account when combining the different parallel arms.

References

[1] Geza Kurczveil, Paolo Pintus, Martijn JR Heck, Jonathan D Peters, and John E Bowers. *Characterization of insertion loss and back reflection in passive hybrid silicon tapers*. IEEE Photonics Journal, 5(2):6600410–6600410, 2013.

[2] Geza Kurczveil. *Hybrid silicon AWG lasers and buffers*. PhD thesis, University of California Santa Barbara, 2012.

[3] M.L. Davenport. private communication, November 2012.

4

Optically pumped membrane LEDs on silicon

4.1 The need for a novel device concept

In this chapter we introduce a novel device structure: optically pumped LEDs based on single-mode high contrast III-V membranes.

The need for a device like this originated from sensing applications. As discussed silicon photonics can be an excellent platform for integrated optical spectroscopic sensors because of the low-loss high-contrast waveguides. By making a very long spiraled waveguide, the to-be-sensed material is in contact with the probing light over a great distance. After this, high-quality spectrometers are used to resolve the absorption spectrum.

This technique requires an integrated broadband light source, where SLEDs, tunable lasers and laser arrays are the traditional options. Above all, the source has to be efficient and should be manufacturable with high yield. In these characteristics, the traditional sources have both their merits and drawbacks:

The SLEDs, e.g. the SLED described in chapter 3, are relatively simple devices. They are very robust against any reflections or other parasitic feedback effects, especially if there is an absorbing layer as we have used. The layer stacks not only have an active region, but also extra layers are needed to accommodate the electrical pumping. Because of this, the fundamental mode will always be an extended mode and therefore, isotropic spontaneous emission cannot couple efficiently to the waveguide. Typically the coupling efficiency is only 0.1%. The

efficiency of the SLED can only be high in a very high gain regime, in which case it is also much more sensitive to feedback.

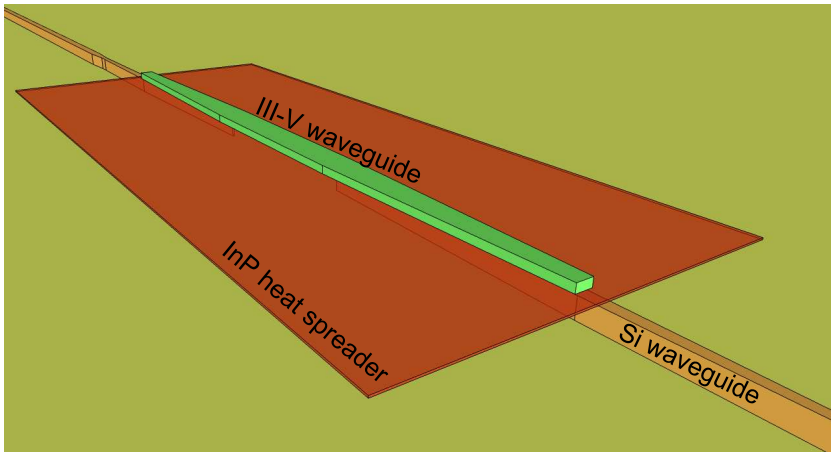
Lasers on the other hand can be very efficient. Spontaneous emission is negligible and virtually all of the radiative recombination is a consequence of stimulated emission. Contrary to spontaneous emission, all the photons resulting from stimulated emission couple to the waveguide mode. Fabrication variations can have detrimental effects on lasers though. Lasers are also very susceptible to reflections from the circuit, due to the lack of integrated isolators. The yield of tunable lasers currently is therefore typically low, making it not a very attractive option for integration in a low-cost sensing circuit. DFB lasers are less prone to these parasitics and an array could make for a broadband source. Unfortunately making multiple of them reduces the overall yield again considerably as all of them have to work ($\text{yield}_{\text{total}} = \text{yield}_{\text{single}}^{\#\text{lasers}}$).

The optically pumped LEDs tries to retain the simplicity of the SLEDs, while boosting the efficiency greatly. We rely on spontaneous emission and not on stimulated emission, which should lead to a robust device operation. By releasing the requirement of electrical pumping, we open design space to enhance the efficiency. The P- and N-InP layers are removed and a high refractive index contrast is created, which confines the optical mode tightly. As a result, the collection efficiency of spontaneous emission is improved by orders of magnitude.

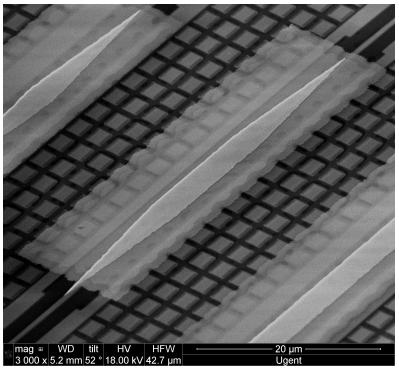
In a complete sensing circuit, a laser (which can now be simple as it only needs to provide pump power) could pump one or more optically pumped LEDs to probe the spectrum of interest.

In this chapter, we discuss the optically pumped LED in depth. The working principle and operating modes are first introduced. Next we simulate the different sides of the LED story. First and foremost the coupling of the spontaneous emission to the fundamental waveguide mode is studied, as this is the heart of the single-mode LED idea. Next the pump confinement is optimized for efficient pump absorption, the thermal effects are discussed and optical coupling structures are devised. We present a first generation of InP-based QW devices in section 4.5. By chirping the size of InAs quantum dots we also fabricate InP-based broadband LEDs in this section. In order to improve the efficiency of the first generation of devices, several approaches were studied. The optical and electrical design is reviewed. The taper losses and pump absorption are under review, while the non-radiative recombination is tackled in multiple ways. We also try to improve the radiative recombination rate. By making use of InAsP-based quantum wells, we manage to boost the power conversion efficiency to 1%.

Some of the devices fabricated in this section were realized by heterogeneous integration by means of bonding, while for most of them transfer printing was



(a) Illustration of the device structure.



(b) SEM image of one of the fabricated devices.

Figure 4.1: The optically pumped, single-mode LED.

employed.

4.2 A novel device concept explained

In this chapter we depict a novel device: a single-mode LED. By removing the layers needed for electrical injection, design freedom is created to enhance the coupling of the spontaneous emission to the fundamental waveguide mode. This means however that the device needs to be optically pumped.

The device geometry is illustrated in figure 4.1. The silicon chip consists of a 220 nm thick silicon device layer on a 2μm buried oxide layer. All the desired passive structures for e.g. sensing are etched into this chip in IMEC's 200 mm

pilot line. A thin III-V membrane is now bonded or transfer printed to this Si PIC. A high refractive index contrast waveguide is etched, which can boost the collection efficiency of the spontaneous emission. A double taper structure (both the silicon and III-V waveguide taper) ensures a short and efficient coupling between the silicon and III-V.

An optical pump is incident through the silicon waveguide (e.g. coming from the bottom right in figure 4.1a). The pump couples to the III-V waveguide, where it is strongly absorbed. The generated carriers can now radiatively combine by means of spontaneous emission. The spontaneous emission is efficiently captured by the co or counter propagating fundamental waveguide mode. If the generated photon is not absorbed, it couples back through the double tapers to the silicon waveguide. If it is absorbed, the generated electron-hole pair can again radiatively recombine, a process which is called photon recycling. Depending on whether the co or counter propagating mode is used, we will talk about transmissive or reflective operation mode. Note that by keeping the pump powers low, we want to keep the non-radiative recombination (Auger recombination) to a minimum and the efficiency high.

4.3 Simulating the optically pumped LED

In a typical (S)LED, metal contacts are needed which would induce loss if they come close to the optical mode. As a consequence we cannot achieve high contrast in the vertical direction. Since we no longer aim to pump the LEDs electrically, our design freedom considerably opened up, which is explored below.

4.3.1 Coupling efficiency of the spontaneous emission

The coupling efficiency of spontaneous emission to the fundamental waveguide mode is usually very low (order of 0.1%). The numerical aperture of such a waveguide mode is typically very small, meaning that little spontaneous emission can be coupled in this mode. However, by reducing the thickness of the membrane, we can make a high contrast waveguide in the III-V. This was simulated using *Lumerical FDTD Solutions*. A cross section of the simulated structure is shown in figure 4.2, where we see a set of 4 InGaAsP quantum wells and barriers with InP SCH layers above and below. A dipole source was placed at various locations in each of the quantum wells, and after propagating a few microns the coupling to the fundamental mode was calculated using an overlap integral method. To ensure the incoherent nature of the emission, this was done for all dipoles separately, the average of which indicates the spontaneous emission collection efficiency. The efficiency as a function of membrane thickness is shown in figure 4.3, for various widths of the membrane. To vary the thickness, the quantum barrier height was

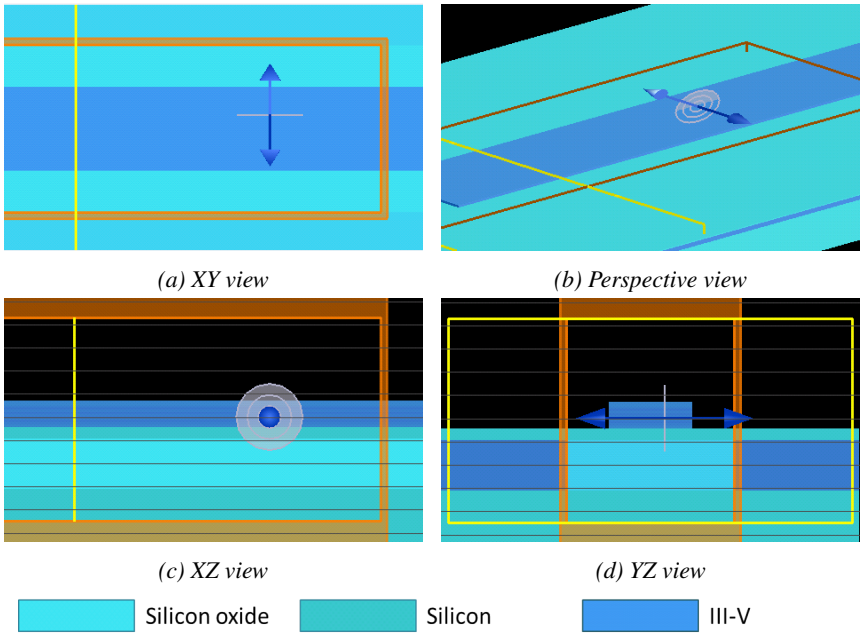


Figure 4.2: Illustration of the FDTD simulation. The arrow indicates the dipole source, along TE direction. At the yellow window, the light is measured and overlapped with the fundamental mode profile.

changed.

The simulation shows a clear optimum at 100 nm. At this optimum 12% couples to the fundamental mode of the 3 μm wide device (another 12% is collected in the fundamental mode propagating in the other direction). For the 1 μm wide device, this goes up even further to 26%. For reasons of pump confinement (below), the thinnest membrane we worked with was 122 nm thick, rendering a slightly lower number. In other tests, the thickness was further varied up to 202 nm, where 9.6% of the isotropic emission couples to the fundamental waveguide mode of the 3 μm wide device. The shaded area on the graph indicates the range of membrane thicknesses studied in this work.

Note that the 20μm wide device is essentially a slab mode. In this case, the only refractive index contrast comes from the vertical dimension. Even for wide device as these, 2% is still coupled, greatly exceeding traditional structures.

4.3.2 Confinement of the optical pump

A similar reasoning holds for the quantum well confinement factor. By introducing the high contrast waveguide, we can push the confinement factor to higher values.

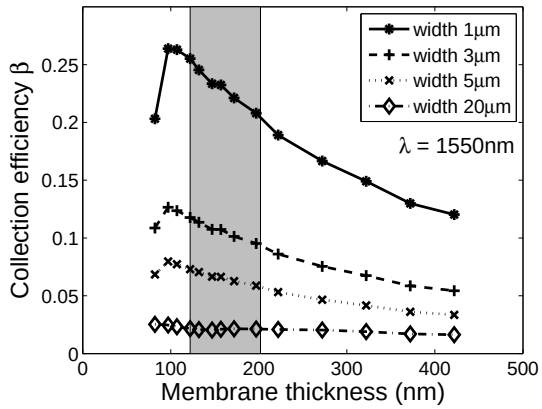


Figure 4.3: Spontaneous emission collection efficiency as a function membrane thickness for various device widths. The shaded area indicates the experimentally studied thicknesses, ranging from 122 nm to 202 nm. The data was simulated using a FDTD method.

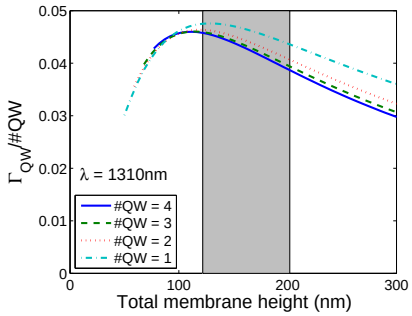


Figure 4.4: The confinement factor per quantum well as a function of membrane thickness. The shaded area indicates the expiremental studied thicknesses, ranging from 122 nm to 202 nm.

Figure 4.4 shows the confinement of the fundamental mode in the quantum wells as a function of the membrane thickness, as simulated by *Fimmwave* of *Photon Design*. An example of the mode profile of the fundamental mode of the membrane is shown in figure 4.5.

As in the simulation of efficiency, also this graphs shows a clear optimum: around 120nm the confinement per quantum well is 4.7%. Again the shaded area indicates the membranes experimentally studied here. The total pump confinement goes from 17.4% to 15.5% when 4 quantum wells are used. Comparing this to the 3-6% achieved with the electrically pumped SLED in figure 3.2 really highlights

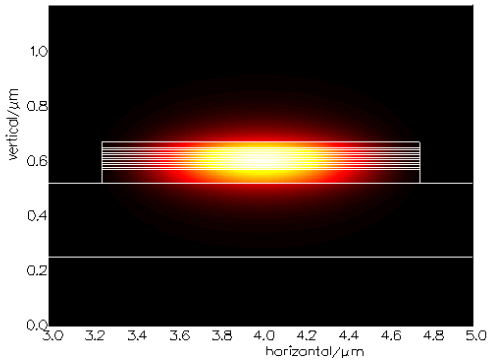


Figure 4.5: Example of a mode profile of a 122 nm thick, 1.5 µm wide III-V membrane.

the impact the high refractive index contrast has.

4.3.3 Thermal design

The device we depict now is very thin and compact, and hence heat generation can become a problem. We first take a look at the effect heat will have on the operation and then how we can reduce its effect by proper designs.

4.3.3.1 Influence of temperature on device operation

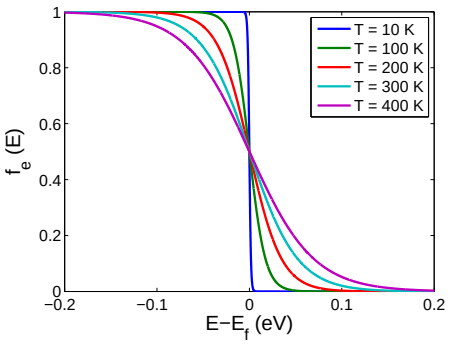
The carrier distribution for holes and electrons is governed by the Fermi-Dirac distribution, which is temperature dependent:

$$f_h(E) = \frac{1}{e^{\frac{E_{f,h}-E}{k_B T}} + 1} \tag{4.1}$$

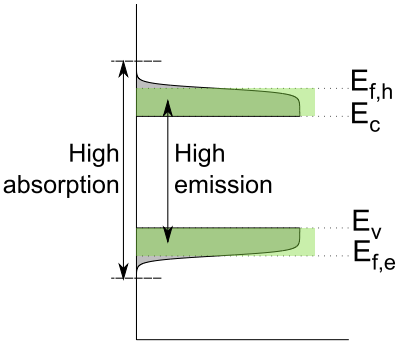
$$f_e(E) = \frac{1}{e^{\frac{E-E_{f,e}}{k_B T}} + 1} \tag{4.2}$$

where $E_{f,h}$ and $E_{f,e}$ is the Fermi energy of the holes and electrons respectively and k_B is Boltzmann's constant. E is the energy of the emitted/absorbed photons and T is the operation temperature.

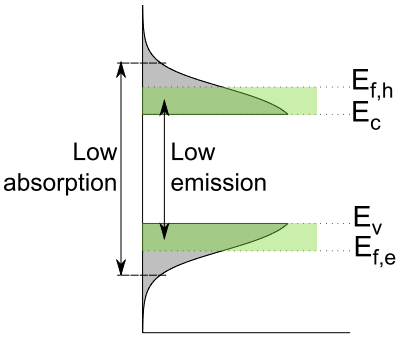
Figure 4.6a shows the behaviour of the Fermi-Dirac distribution for different temperatures. At low temperature the function is almost step-like, whereas in hotter conditions, the curve is smoothed out. After multiplying with the density of states, which can be treated as temperature independent, the carrier distribution can be visualized, as in figures 4.6b and 4.6c. For low temperatures the absorption is high as there are lots of electrons in the valence band and only few holes in the conduction band for the pump photon energy. There is also high emission as all



(a) Fermi-Dirac function for different temperatures



(b) Absorption and emission at low temperature illustrated.



(c) Absorption and emission at high temperature illustrated.

Figure 4.6: Influence of temperature on optically pumped devices.

the excited electrons conglomerate at the low energies and vice versa for the holes. A high temperature smears out the carrier profile. Carriers appear at higher energy levels, reducing the absorption and making the emission spectrum more smeared out.

In these optically pumped devices, the degrading effect of temperature hits you twice. There are less carriers because of the reduced absorption and those excitons that are created do not occur at the desired photon energies.

4.3.3.2 Reducing temperature effects

It is clear that we should minimize the heating of the device during operation. The first, and often forgotten, improvement is to limit the amount of dissipated power. Any design effort to improve the efficiency of the device, will reduce the power dissipated and therefore also the operating temperature. This lower temperature will in turn again improve the efficiency, as shown in figure 4.6, doubling the pay-off. Next to this, we have to spread the generated heat as much as possible, expressed by the thermal impedance. In this paragraph, we will focus on the thermal design.

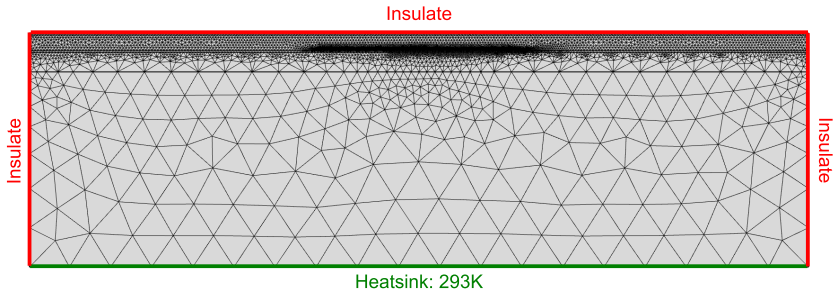
The thermal properties of the membranes were studied using a two-dimensional finite element solver from *COMSOL Multiphysics*. We assume conduction to be the only heat transfer mechanism, with radiation and convection being negligible. Therefore the only equation to be solved is Fourier's law of thermal conduction:

$$\vec{q} = -k\nabla T \tag{4.3}$$

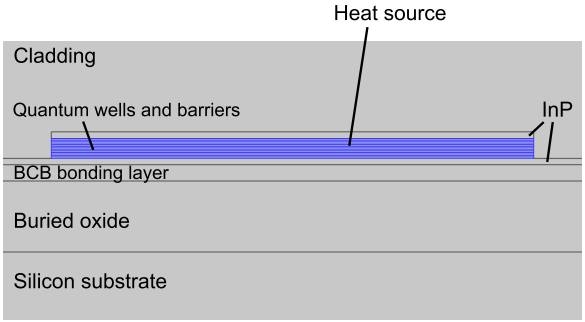
with $\vec{q}$ the local heat flux density, k the material's thermal conductivity and ∇T the temperature gradient. In addition to this equation, we define a boundary conditions at the heatsink (constant ambient temperature $T = 20°\text{C}$) and the simulation edges (no heat transfer through the boundary $q_{\text{normal}} = 0$). Figure 4.7 shows the simulation setup. The thermal conductivity of air is extremely low (0.024 W/m· K) and therefore we could also take the upper boundary to be completely insulating. Obviously one has to make the simulation window wide enough for the insulating sidewalls to be physical, 80 μm was sufficient. The substrate was taken 20 μm thick, as thicker had no influence thanks to the excellent conductivity of silicon. The distributed heat source spreads over the quantum wells and barriers, as all are absorbing the pump and therefore generating heat. There is a constant flux q_{heat} at the heat source of 10^{12} W/m^3. After the simulation, the maximum temperature can be measured and we can define a length specific thermal impedance R_{th}:

$$R_{\text{th}} = \frac{T_{\text{max}} - T_0}{q_{\text{heat}} w_{\text{heat}} h_{\text{heat}}} \qquad \left[\frac{K \cdot \text{cm}}{W}\right] \tag{4.4}$$

with w_{heat} and h_{heat} the width and height of the heat source.



(a) Overview of mesh with boundary conditions



(b) Distributed heat source in thermal simulations.

Figure 4.7: Overview of mesh with boundary conditions

Material	Thermal conductivity k (W / m · K)
Silicon	130
SiO_2	1.24
Si_3N_4	30
DVS-BCB	0.3
Air	0.024
InP	68
InGaAsP (wells)	5
InGaAsP (barriers)	10

Table 4.1: Thermal conductivity of various materials. [1, 2]

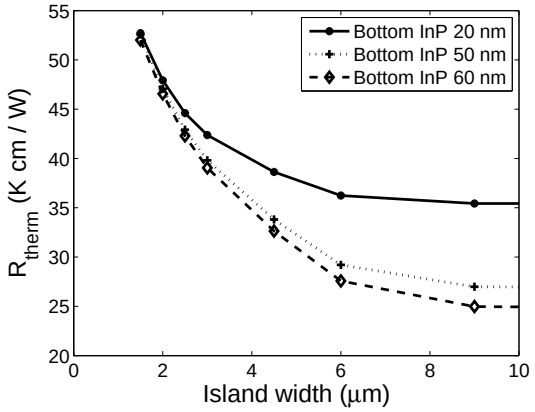


Figure 4.8: Evolution of the specific thermal resistance for various island widths of a BCB clad, 122 nm thick, shallow-etched membrane.

Let us first consider a 122 nm thick membrane, consisting of two 20 nm thick InP layers with a set of InGaAsP quantum wells and barriers in between them. A 1.5 μm wide ridge was etched through the top InP and active region. The membrane is bonded with a 50 nm thick BCB bonding layer to a SOI PIC, which at the location of our cross section consists of a 6 μm wide trench filled with oxide, etched in a 220 nm thick silicon device layer. The buried oxide is 2 μm thick. 1 μm of BCB has been spun over the device. Table 4.1 shows the thermal conductivity of the various materials used here.

Figure 4.8 shows the specific thermal resistance of the membrane for various island widths. Three different membranes were investigated, with varying bottom InP thicknesses. This island width refers to the width of the bottom InP layer. This means the first simulated point (where the island is 1.5 μm wide) depicts a fully etched membrane. The specific thermal resistance drops rapidly as the island becomes wider. The fully etched membrane has an R_{therm} of 52 $\frac{K \cdot \text{cm}}{W}$, while this

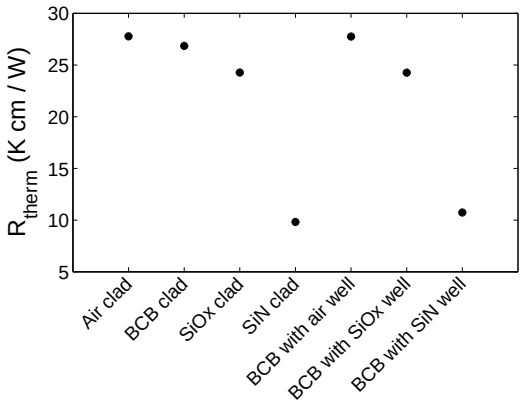


Figure 4.9: Specific thermal resistance of the 152 nm thick membrane for various claddings.

has dropped by 33% to 35.4 $\frac{K \cdot \mathrm{cm}}{W}$ for a 9 μm wide island. Note that the InP bottom layer is only 20 nm thick and has almost no optical influence. The influence on the thermal behaviour is huge though, more so for the thicker InP layers.

The effect of the cladding layers is studied in figure 4.9. Because we also consider shapes typical of our transfer printing process, only the 152 nm thick membrane is shown (this has a 50 nm lower InP cladding and 20 nm upper one). An island width of 20 μm was assumed. The first four options are the traditional uniform claddings. The three latter refer to the transfer printed LEDs which are processed as discussed in section 2.4.2.1. After transfer printing, the chip is planarized with BCB, after which the coupons are opened up self aligned. Etching back the sacrificial InP / InGaAs layers and defining the LED leaves us with a 20 μm wide well in the BCB where the coupon used to be. We can leave this open, or clad it with a dielectric. Figure 4.9 shows the power of using silicon nitride. The thermal conductivity of Si_3N_4 is one and two orders of magnitude larger than SiO_2 and BCB. Furthermore the figure also shows us the uniform claddings and well shapes behave virtually the same. Coating the LED with Si_3N_4 after definition greatly reduces the thermal resistance of the device. The optical field will become slightly larger, but the effect is all in all limited as the refractive index contrast with the InP system is still very large. The larger optical field will reduce the coupling of spontaneous emission to the waveguide mode slightly.

Finally we also want to take a look at how small the oxide trench in the SOI can be. Since silicon is the best heat sink our material system has to offer, we should try make use of it as much as possible. However, bringing silicon close to the

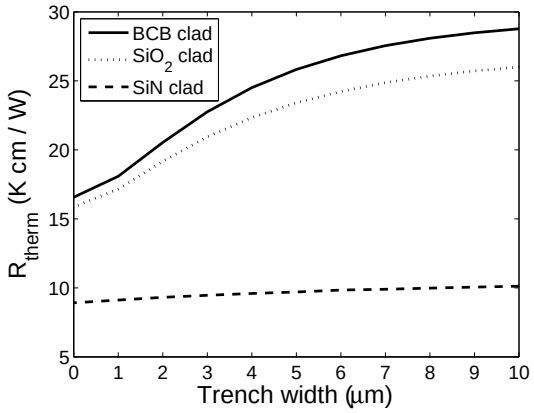
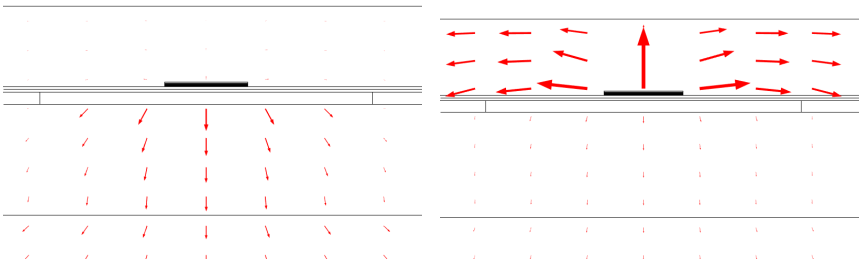


Figure 4.10: Specific thermal resistance as a function of trench width in the SOI for different claddings.

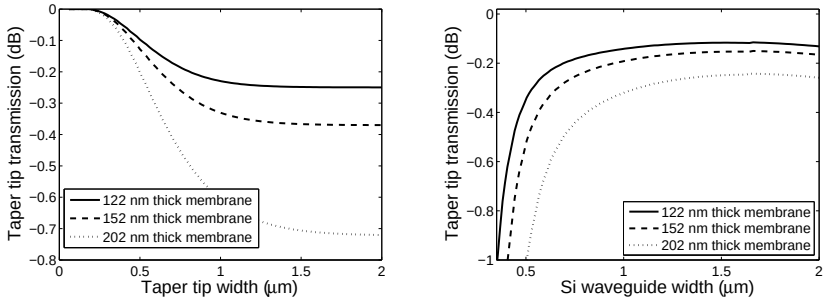


(a) The BCB cladding isolates the membrane from the top. This causes the dominant heat transfer path to go vertically down, through the trench oxide.

(b) The SiN cladding provides the option for the heat to spread horizontally and then go through the the silicon device layer.

Figure 4.11: Effect of the top cladding on the heat transfer path when the oxide trench is 6 μm wide.

LED carries a risk as the high refractive index of silicon disturbs the optical mode. Figure 4.10 shows the COMSOL simulation for three different claddings. The membrane had a 50 nm thick bottom InP layer and the island was 20 μm wide. BCB and SiO_2 behave similarly with a sharp increase between 1 μm and 5 μm after which the trend slows down. This is not very practical though as making a trench narrower than 6 μm increases the risk the optical mode leaks. The overcladding isolates the thermal source from the top though, making the path straight down through the trench the dominant one, shown in figure 4.11a. In the case of Si_3N_4



(a) Influence of the III-V taper tip. *(b) Influence of silicon waveguide width.*

Figure 4.12: Transmission through the taper tip of the different membranes.

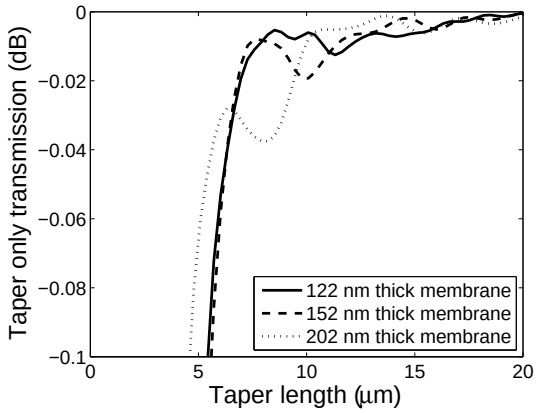


Figure 4.13: Transmission through the taper as a function of the taper length. The III-V tip was 700 nm wide for all membranes.

cladding, the heat is pulled into the Si_3N_4 where it can spread horizontally. After spreading the heat, it can propagate through the silicon device layer towards the substrate. Hence, the trench width is less important. Through all the devices, we used 6 μm wide trenches.

4.3.4 Coupling the III-V membrane to the silicon waveguide

Because we are working with thin membranes, the taper can be made short and tolerant at the same time. Due to the high contrast and limited dimensions, the effective index of the fundamental mode is only 2.3 (for a 1.5μm wide device), meaning that even with 220nm thick silicon waveguides, we can make an adiabatic taper for 1550 nm. We have used TE polarization throughout all our simulations.

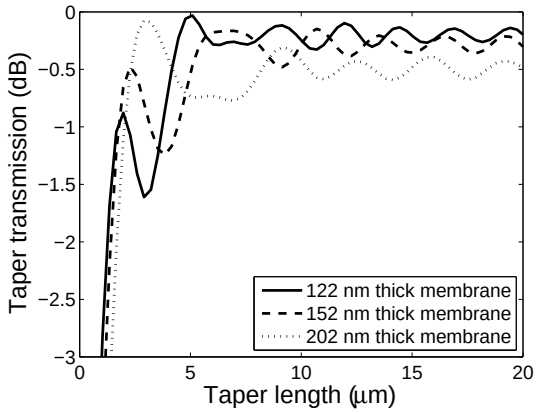


Figure 4.14: Transmission through the complete taper as a function of the taper length. The tip was 700 nm wide for all membranes.

Figure 4.12 shows the transmission through the III-V taper tip. The InP-based membrane can have wide taper tips, as can be seen from figure 4.12a (the silicon waveguide was 700 nm wide and the bonding layer is 50 nm thick). Even a 2 μm wide tip will only incur a loss of 0.25 dB for the 122 nm thick membrane. The effect of the silicon waveguide width is shown in figure 4.12b, with a III-V tip of 800 nm. As soon as the waveguide is wider than $\pm$ 600 nm, the width has only a negligible effect. At this point we essentially have a slab mode, so the mode does not change a lot anymore with changing waveguide width.

The effect of the taper length was studied in figure 4.13. To only look at the tapering section itself, we excited with the hybrid mode at the III-V taper tip and register the power in the hybrid mode at the silicon taper tip. Because of the strong coupling between the silicon waveguide and the III-V membrane, this taper can be extremely short. Already a 15 μm long taper is adiabatic when tapering the III-V from 800 nm to 1.5 μm and the Si from 700 nm to 150 nm (again the BCB layer was 50 nm thick. Note that the scale on the y-axis is extremely small in the figure. In our designs we typically used 18 μm long tapers.

Finally, bringing everything together, figure 4.14 simulates the entire taper, starting from the pure III-V mode and ending in the pure silicon mode. The III-V was tapering from 800 nm to 1.5 μm while the fully etched silicon waveguide width was going from 700 nm to 150 nm. The BCB bonding layer was assumed to be 50 nm thick. Tapers longer than 10 μm only show a maximum loss of 0.5 dB, which is mainly due to the III-V taper tip interface. Figure 4.15 shows a side view of the taper coupling light from the III-V to the silicon waveguide. The slight oscillation indicates the taper is not fully adiabatic, but as was discussed the losses are negligible.

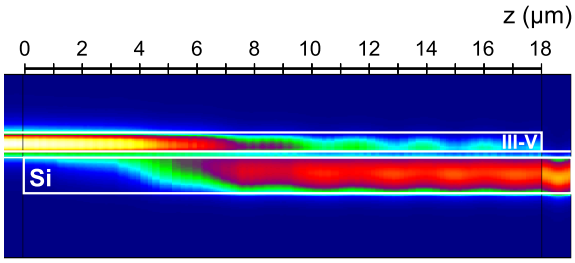


Figure 4.15: Sideview of the taper structure. 1550 nm light (TE polarized) is coupled from the III-V to the Si waveguide.

4.4 Fabrication process of optically pumped LEDs

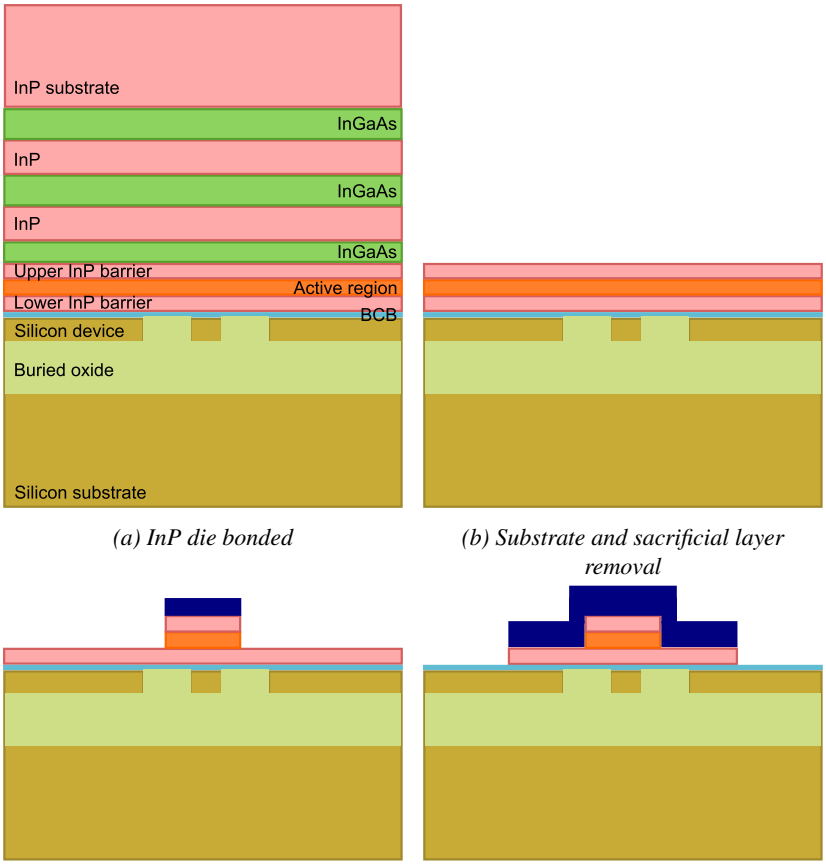
The optically pumped LED was designed to have a short and simple fabrication process. We have made LEDs both with the more traditional BCB bonding technology as well as with a transfer printing approach. Both fabrication processes are discussed below.

4.4.1 Based on BCB bonding

In die-to-die bonding technology it is of critical importance to work with clean samples as a parasitic particle will cause a bond failure which will cover a much larger area than the initial particle. The SOI was fabricated in imec's 200 nm pilot line, where a 220 nm silicon device layer was patterned with a partial etch of 70 nm and a full etch. After a oxygen plasma clean and a dehydration bake at 150°C, diluted BCB is spun on the sample. The dilution is chosen such 50 nm of BCB remain on top of the waveguide. A partial cure of 30 min at 180°C ensures flatness across the entire die. At the same time the III-V die is prepared by removing a sacrificial InP-InGaAs layer pair and deposition of a thin layer of silicon oxide. Both dies are now brought into contact in vacuum and 15 mbar of pressure is applied on a die of 4 mm by 8 mm. The BCB is hard cured at this stage. We now have an SOI sample with a InP-based epitaxial layer stack bonded to it, as shown in figure 4.16a.

The InP substrate is removed in a heated HCl wet etch and the sacrificial InGaAs and InP layer are removed in H_2SO_4:H_2O_2:H_2O and HCl respectively, depicted in figure 4.16b. Note the existence of many different sacrificial layers in figure 4.16a, which were there to develop the transfer printing process. They do not serve any purpose in the typical bonding process as one InGaAs etch stop layer suffices. At this stage, the III-V epitaxy has been provided to the silicon platform and the fabrication of the actual device can start.

Figure 4.16c shows how the III-V rib waveguide is etch using a silicon nitride



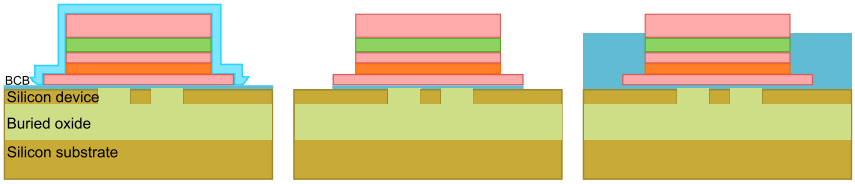
(a) InP die bonded

(b) Substrate and sacrificial layer removal

(c) Rib etch of the III-V waveguide

(d) Island etch of the III-V waveguide

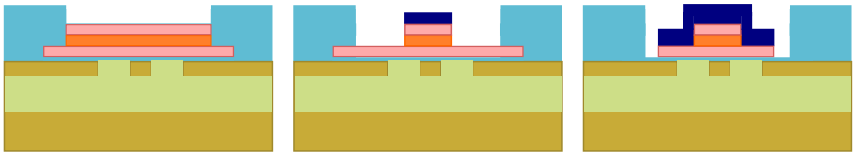
Figure 4.16: Fabrication process of the optically pumped LEDs based on BCB bonding.



(a) Transfer printed coupon

(b) Removal of the photoresist encapsulation and cure of the BCB bonding layer

(c) BCB protection of the rest of the SOI circuit



(d) Sacrificial layer removal

(e) Rib etch of the III-V waveguide

(f) Island etch of the III-V waveguide

Figure 4.17: Fabrication process of the optically pumped LEDs based on transfer printing.

hard mask. We prefer wet etch for this step (HCl:H_2O and H_2SO_4:H_2O_2:H_2O) but also dry etch was used. The sidewall contains less recombination centers with the wet etch, but the dry etch does not suffer from undercut.

Another silicon nitride hard mask is applied to do the wet island etch in figure 4.16d. Note that in the previous step the hard mask was used to define sharp features, whereas in this step the reason is mostly found in the thermal behaviour. A silicon nitride coating will greatly decrease the thermal resistivity of the devices, as discussed in section 4.3.3. Next to the thermal argument, this nitride also seals the active region from the environment, which in turn is beneficial to prevent sidewall recombination.

4.4.2 Based on transfer printing

In order to do transfer printing, the III-V coupons have to be released from the substrate, as described in section 2.4.2.1. The process flow of figure 4.17 starts when the coupons have been transfer printed.

The printed coupon consists of the LED layer as well as a sacrificial InP-InGaAs layer pair for mechanical stability during the transfer. The encapsulation consists of photoresist as well as a silicon oxide cap. The first step is to remove this photoresist, depicted in figure 4.17b in an extended oxygen plasma. This has to be done even before the BCB is hard cured as the photoresist would burn. Once the resist is gone it is safe to do the BCB bake at 250°C.

Next a thicker BCB planarization is done to protect the SOI circuit from further

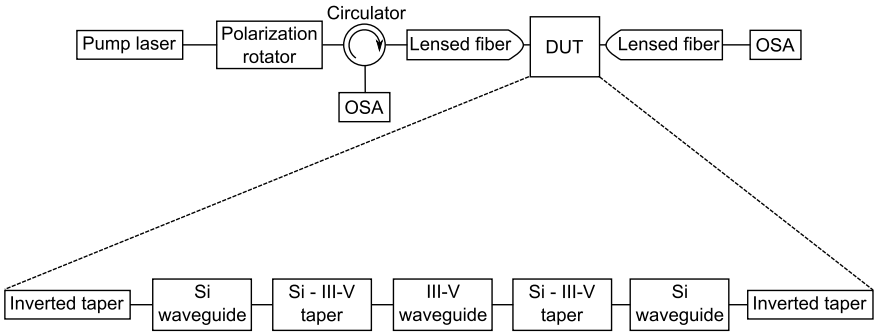


Figure 4.18: Characterization setup for optically pumped LEDs

processing. As shown in figure 4.17c, the BCB is etched back thereby exposing the coupon. In a self-aligned manner, only the coupon is not covered by BCB.

In figure 4.17d, the sacrificial InP-InGaAs layer pair is removed as well as the oxide cap (not depicted). Most of the 1μm thick InP is etched dry, only the last 70 nm is done wet in HCl:H_2O. During this wet etch a large overetch has to be taken into account since slow-etching crystal planes will be exposed. The 100 nm InGaAs layer is etched in H_2SO_4:H_2O_2:H_2O.

At this stage the III-V material is provided to the silicon platform and we can pattern the actual device as in the previous section. The III-V waveguide is etched in two steps, each with a nitride hard mask.

4.5 First generation of devices

4.5.1 QW-based devices

Taking all the design aspects discussed above into account we fabricated optically pumped LEDs with a III-V membrane only 122 nm thick. The characterization of these devices is reported in this section. We used a measurement setup as illustrated in figure 4.18. A 1310 nm pump laser is coupled through a polarization controller to a circulator. The output of the circulator is connected to a lensed fiber coupling to the device under test (DUT). On the other side of the chip, light is coupled out with a lensed fiber and directed to an optical spectrum analyzer (OSA) to measure the generated light co-propagating with the pump beam. The counter-propagating light is measured at the third port of the circulator. The DUT is a SOI chip, where the coupling to and off chip is realized by inverted silicon waveguides with an SU8 cladding, matching the fundamental mode of the lensed fiber. This broadband coupling structure is discussed in more detail in [3]. The Si – III-V taper and III-V waveguide have been discussed above.

Figure 4.19 shows the resulting spectra measured from a 150 μm long, 1.5 μm

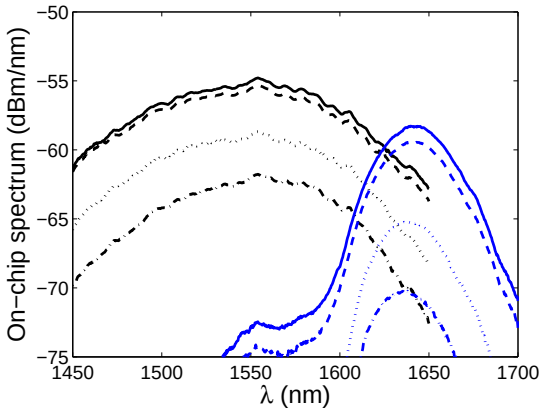


Figure 4.19: On-chip spectra of a 150 μm long, 122 nm thick membrane-based LED. The black graphs depict the counter-propagating generated light while the blue graphs refers to the co-propagating spectrum. The dash-dotted, dotted, dashed and full lines indicate pumping at -12 dBm, -8.6 dBm, -3.7 dBm and -2.7 dBm on-chip respectively.

wide LED. Note that two 18 μm long tapers have been included in this 150 μm length. The resulting counter-propagating spectra are shown in black, the different line styles depicting different pumping conditions as outlined in the figure caption. All spectra have a 3 dB bandwidth of approximately 120 nm, centered around 1550 nm. The co-propagating light is shown in blue. A clear redshift can be observed, with a 45 nm wide spectrum centered around 1640 nm. This can be attributed to incomplete pumping as the pump absorption is so high. As the LED is 150 μm long and the absorption coefficient is around 1000 dB/cm (discussed in section 4.6.1), the pump has been depleted by 15 dB by the end of the LED. This end is not well pumped anymore, especially for the blue side of the spectrum. A spectrum similar to the black graphs is generate, but the short wavelength side is reabsorbed while propagating through the LED. This effect is even more clear in figure 4.20, where spectra from LEDs of different lengths are displayed. The end of the long LEDs (200 μm and 250 μm) is not pumped anymore, thus very little spontaneous emission occurs and what was generated is the beginning of the LED is absorbed. Hence, almost no light can be measured, especially at shorter wavelengths. As we are looking at shorter LEDs, the long wavelengths first appear around 1640 nm (at a length of 150 μm). When the LED can be pumped completely (37 μm and 50 μm, the spectrum is centered at 1550 nm and resembles the spectrum in reflective mode.

An overview of the performance of the optically pumped LEDs is given in figures 4.21 and 4.22 for the co- and counter-propagating signals. Each figure shows the evolution of the 3 dB bandwidth, generated power and efficiency as a

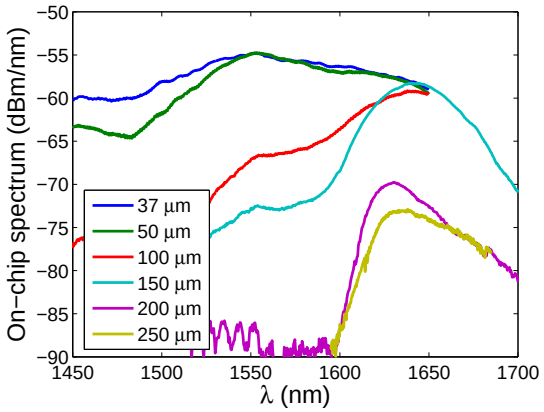


Figure 4.20: On-chip spectra of LEDs of different length in transmissive mode, pumped with -2.7 dBm at 1310 nm.

function of pump power for various lengths of LEDs. A coupling loss between the lensed fiber and the on-chip spotsize converter of 5dB was assumed. The rest of the setup was also calibrated out and, in order to be sure the pump is excluded, the power was integrated from the spectrum. In transmissive mode, the bandwidth decreases for increasing LED length due to incomplete pumping, as discussed in figure 4.20. A similar, but more modest behavior can be observed with rising pump power, as the end of the LED will be pumped better. The generated power rises with increasing pump power, to a maximum of -35 dBm for the shortest LED. Longer devices generate less power in transmission as this part of it will be absorbed by the unpumped end of the LED. This is why we observe a monotonic increase in efficiency as a function of pump power. The end of the LED is pumped better and better and absorbs less of the generated signal. Only for the shortest one (37 μm) this is different, where the optimal efficiency is achieved when pumping at -8.6 dBm, achieving a value of -30 dB or 0.1 %. The shortest LED is fully pumped and when excited more strongly, the extra carriers find other ways of recombining (e.g. Auger recombination), lowering the efficiency. In addition also the free carrier absorption increases.

In reflective mode, the bandwidth does not change dramatically as a function of pump power (note the difference of scale between figures 4.21 and 4.22). The generated power and efficiency also only show subtle differences. The efficiency is optimal at lower injection powers (-12 dBm and -8.6 dBm) at -29 dB. The hit in efficiency is more pronounced for the shortest LED than for the others. This might be due to the fact that there is no extra, previously unpumped, material contributing to the emission, or it could be related to the much narrower shape of the LEDs. All LEDs are 1.5 μm wide, but of course the tapers are not. In case of a 37 μm long

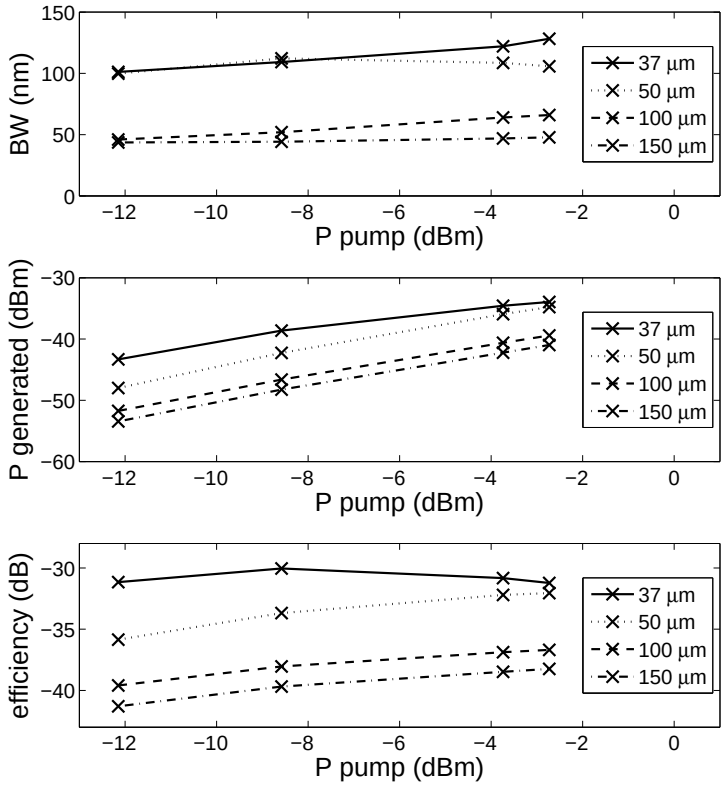


Figure 4.21: Overview of the performance of the 122 nm thick optically pumped LED in transmissive mode

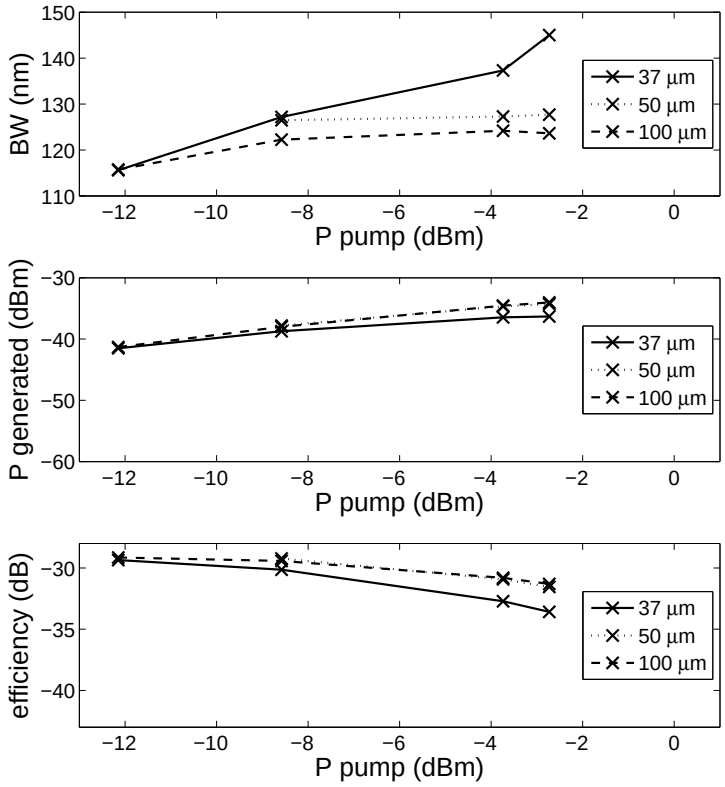
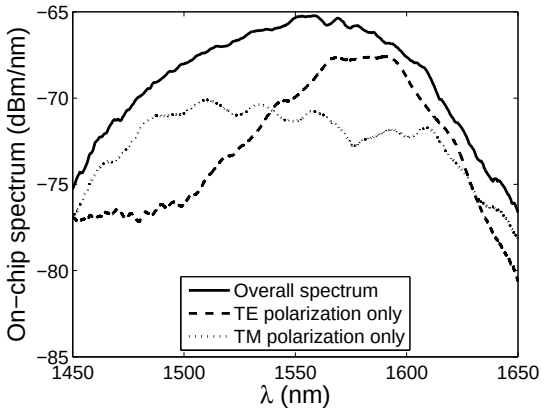
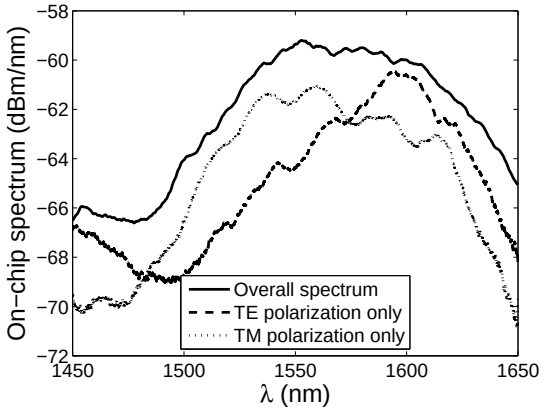


Figure 4.22: Overview of the performance of the 122 nm thick optically pumped LED in reflective mode



(a) The counter-propagating spectrum of a 50 μm long LED.



(b) Co-propagating spectrum of a 37 μm long LED.

Figure 4.23: Polarization dependence of the spectra, pumped with -8.6 dBm.

device, only 1 μm is actually 1.5 μm wide, with the rest tapering down. Hence, non-radiative recombination will come into play.

For both measurement modes, the efficiencies were much lower than anticipated. This is investigated in further detail in section 4.6.

Finally, the polarization dependence is investigated in figure 4.23. To do this, we added a polarizer before the OSAs in the setup of figure 4.18. For the LED in reflective mode, we show the 50 μm long LED in figure 4.23a. The TE polarized light is clearly dominating the longer wavelengths, while the TM polarization is

#	Layer	Thickness
0	InP substrate	360 μm
1	InGaAs etch stop	150 nm
2	InP	20 nm
3	InAs dot / InAlGaAs barrier	3 - 3.5 - 4 - 4.5 - 5 - 5.7 monolayers / 7 x 15.0 nm
4	InP	20 nm
5	InGaAs sacrificial layer	150 nm
6	InP sacrificial layer	150 nm

Table 4.2: Epitaxial layer stack containing chirped quantum dots for broadband emission.

more present on the short side of the spectrum (the effect of pump polarization was limited). Since all the layers are unstrained in our layer stack, the difference in these polarizations is related to the hole mass in the different valence bands. The TM light is dominated by the light holes, while the TE light is also dependent on the heavy holes. Because of the larger mass of the heavy holes, the quantum states are closer to the band edge. Therefore the light holes render higher energy photons and thus, the TM light is found at the short wavelength side. A similar behavior can be seen in figure 4.23b, depicting the LED in transmissive mode.

4.5.2 QD-based devices

Chirped quantum dots can offer very broadband optical spectra. Due to the size inhomogeneity, self-assembled quantum dots already offer a broadband medium, and this is further enhanced by growing several layers of different sizes. In collaboration with prof. Reithmaier in Kassel University, broadband optically pumped LEDs were realized by chirping quantum dots. The grown layer stack is outlined in table 4.2. The active region consists of the chirped InAs quantum dots in InAlGaAs barriers. By varying the layer thickness, the size of the self-assembled dots is chirped. Two 20 nm thick InP barriers are used to prevent carriers from recombining vertically. The final total membrane thickness is 160 nm. The other layers in figure 4.24 serve as sacrificial layers in the fabrication process.

In order to optimize the layer stack, test wafers were grown without the sacrificial layers. By removing the InGaAs layers, we can easily do photoluminescence measurements and work towards a broadband gain medium. Figure 4.25 shows the result of this exercise. The layer stack emits from 1300 nm to 1600 nm at 10 K and is consistent for different excitation power densities.

Heterogeneous integration requires high quality III-V materials. While the BCB bonding layer can alleviated imperfections on the die surface, the substrate removal will highlight any defects. Figure 4.26 shows a high density of oval defects observed on the III-V epi-structure. During the substrate removal, the heated

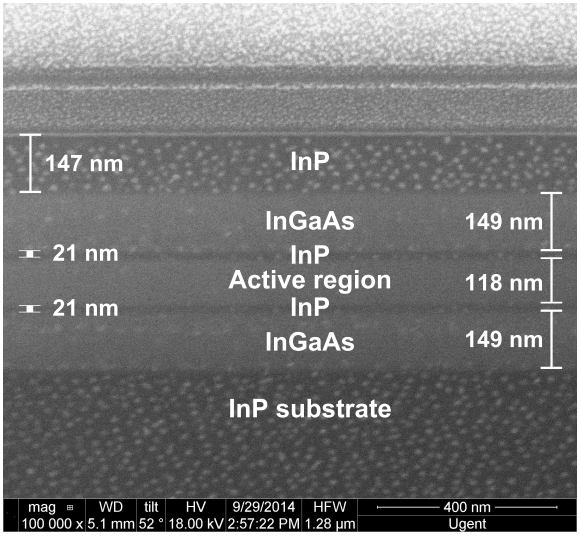


Figure 4.24: A cross section of the layer stack containing chirped quantum dots.

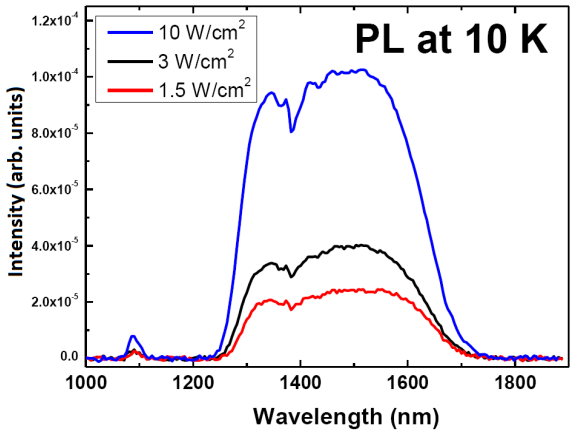


Figure 4.25: PL measurement of the chirped quantum dots at 10 K with different excitation powers densities.

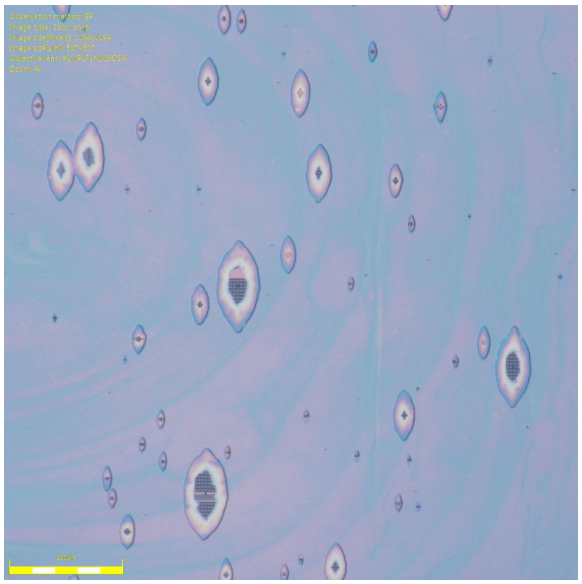
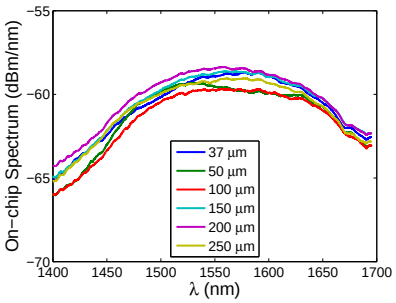


Figure 4.26: Oval defects in the chirped QDs, highlighted by the InP substrate removal.

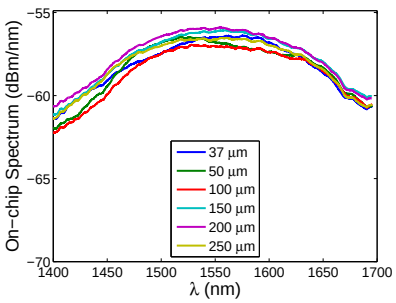
HCl penetrates through the defects and etches the device layers. We want to emphasize that this is not a consequence of the bonding itself, but the pre-existing, microscopic defects are enlarged to macroscopic sizes due to the wet etchants. Unfortunately the grown stack was suffering a lot from these defects, as illustrated in figure 4.26.

Nonetheless, we could make devices with them, the spectra of which are shown in figure 4.27. Figures 4.27a and 4.27b show the spectra in reflective mode. The length of the LED does not have a large influence, indicating that the devices are dominated by the emission from the start of the III-V waveguide. The variation visible is probably due to processing and measurement imperfections. The 3 dB bandwidth is around 220 nm for all, ranging from 1440 nm to 1660 nm. The co-propagating signals are studied in figures 4.27c and 4.27d. The shortest LED is clearly the best, with the blue side losing power rapidly with increasing length. The end of the long LEDs is not pumped and the generated signals are absorbed. The co-propagating result of the 37 μm long LED resembles the counter-propagating one, especially when pumping hard. This is expected, as the LED is pumped completely in this case.

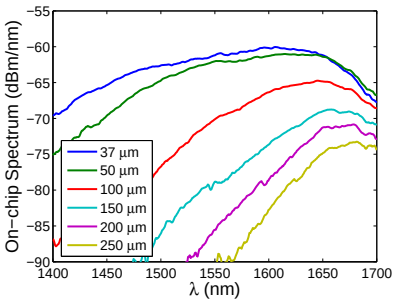
It is worth noting that although the high density of oval defects will degrade the performance, the efficiency could still be around -30 dB. This is comparable to the QW-based LEDs described above, which again confirms these were performing below the potential.



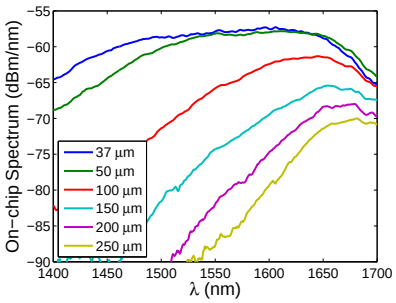
(a) Reflective mode, pumped by 0 dBm on-chip.



(b) Reflective mode, pumped by 4.4 dBm on-chip.



(c) Transmissive mode, pumped by 0 dBm on-chip.



(d) Transmissive mode, pumped by 4.4 dBm on-chip.

Figure 4.27: Spectra of optically pumped LEDs based on chirped quantum dots.

4.6 Identifying problems and optimization

Although simulation was predicting an efficient device, the measurement stayed below expectation. Tracking down the problem(s) limiting the output power was challenging because of the many possibilities at hand.

First of all, we need to confirm whether the optical coupling structures work as expected. Does the 1310 nm pump light couple from the silicon to the III-V waveguide? If it does, is it sufficiently absorbed and does it create electron hole pairs?

Once the pump photons have been converted to electron-hole pairs, these need to recombine radiatively. It is of critical importance that the non-radiative recombination lifetime is as long as possible. A reduction in non-radiative lifetime can be due to many effects. First of all, bulk defects have a devastating effect as they will provide recombination centers where most of the carriers reside. The most typical source of bulk defects are growth imperfections, but in these ultrathin membranes the processing may also introduce them. When the InP substrate is removed, the lattice may relax with the related defect formation. Plasmas can also induce damage to semiconductor material and since in these membranes the quantum wells lie 20 nm from the exposed interface, they could be susceptible.

Vertically, the InP barriers confine the carriers to the active region. Due to the processing and exposure to the air, surface states have been formed on the top and bottom of the device. If the carriers are able to tunnel through the InP barriers, they will recombine at these surface states and are not able to contribute to light generation.

The etched sidewalls may also contain many surface states. The carriers are mostly created in the center of the waveguide, where the pump optical mode is strongest. With no horizontal electron and hole barrier, diffusion can freely spread out these carriers horizontally as wide as twice the diffusion length. One manner of counteracting sidewall recombination is using a III-V waveguide wider than this diffusion length. Most of the carriers will not reach the sidewall before they recombine radiatively. Reducing the diffusion length is also an attractive strategy, but is usually only possible by changing the material.

The problem can also be tackled at the root, by removing or never creating the surface traps. Wet etching will generally cause less states to exist than dry etching does. Passivation can get rid of the remaining recombination centra.

Non-radiative recombination is only as important as the radiative transition lets it be. A short radiative lifetime is the best guarantee for efficient light emission. InP-based materials and other III-Vs are very good in this regard. Typical lifetimes are of the order of nanoseconds, giving the non-radiative recombination only a short time frame to be of importance.

Due to the lack of stimulated emission (there is no input signal), radiative re-

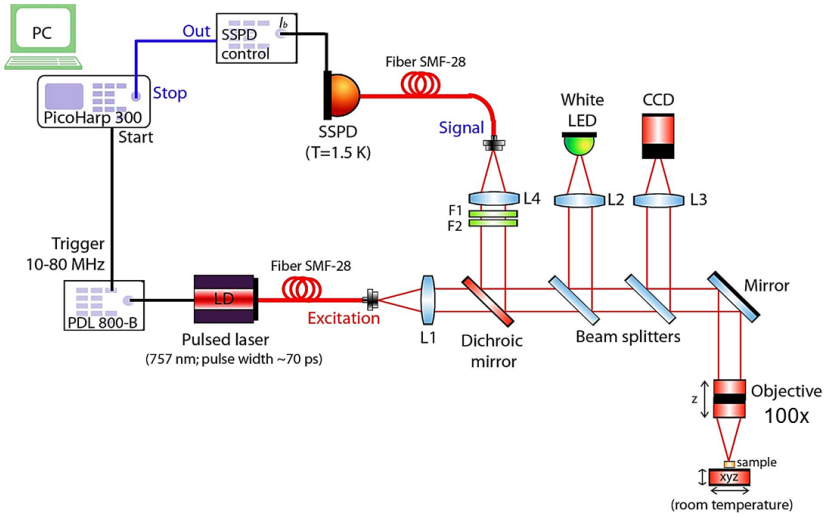


Figure 4.28: Time-resolved photoluminescence measurement setup

combination is mostly translated into spontaneous emission. This isotropic emission couples to the fundamental mode of the III-V waveguide. The high refractive index contrast ensures this is an efficient coupling and forms the core idea of this novel device concept. Finally the fundamental III-V mode is brought to the silicon waveguide with the same taper used by the pump.

Some pitfalls described above are to be investigated by characterizing the complete device, in particular the optical paths of the pump and signal. While the different causes for non-radiative recombination can also be determined in this manner, it is a very indirect way. A more direct method is time-resolved photoluminescence measurements. We did these measurements in collaboration with Bruno Romeira and Simone Birindelli of the Photonics and Semiconductor Nanophysics group from the Eindhoven University of Technology.

Figure 4.28 shows a schematic of the measurement setup used. A 757nm pulsed laser diode is exciting the sample. The pulse width is around 70 ps and the repetition rate can be changed from 10 MHz to 80 MHz. After passing through a dichroic mirror and some beam splitters a 100x objective is used to focus the laser spot on the sample. The diameter of the focused spot is around 1 μm. The photoluminescence is captured through the same objective and reflected on the dichroic mirror. After attenuating the pump further in two band pass filters (F1 and F2 on the figure), the signal is incident on a superconducting single photon detector (SSPD). The white LED and the CCD camera make sure we can align our excitation spot to our desired location. The PicoHarp 300 software in essence mea-

sures the delay between the excitation pulse trigger and the first incident photon on the SSPD. In this manner the delay between excitation and emission is measured from which the carrier lifetime can be deduced.

In the following sections we will go over the different potential efficiency killers of our LEDs. The taper losses and pump absorption are discussed, as well as the different sources of non-radiative recombination. These include tunneling vertically through the InP barriers, as well as sidewall recombination. Several solutions to the latter are investigated. We also take a close look on the radiative recombination and how to decrease the radiative lifetime. Finally, an overview and conclusion is formed.

4.6.1 Taper losses and pump absorption

We couple in a laser diode through a polarization rotator and measure the transmission through the chip. By comparing LEDs of different lengths and reference waveguides, we can estimate the taper losses and pump absorption/loss. Since absorption should be much larger than the linear loss, we can assume the pump loss will mostly be due to absorption.

Figure 4.29 plots the transmission loss as a function of LED length, with a straight line fitted to the measurements. The LEDs used here were the same ones as in section 4.5.1. The slope of the fitted line provides the absorption coefficient of the pump in the LED (around 1000 dB/cm), while the comparison to the silicon reference waveguide renders the taper loss. As designed, this taper loss is very low, in fact too low to measure in this manner. Finally, as we do not see a dependence of the absorption coefficient on the pump power, we do not observe a bleaching effect.

4.6.2 Non-radiative recombination due to bulk defects

Working with good quality material is of paramount importance for making efficient devices. Defects in the lattice cause mid-gap states, which act as non-radiative recombination centers. Regarding non-radiative recombination, we first need to check whether the material quality has been compromised. We describe our tests on bonding, plasma effects and growth defects.

4.6.2.1 Bonding and substrate removal

Heterogeneous integration of III-V on Si is a well established process, but we have been pushing it to the extreme. The remaining III-V material is only 122 nm thick, only one third of which is InP. Although we targeted unstrained material during the growth, there could be residual stress in the layers. When removing the InP

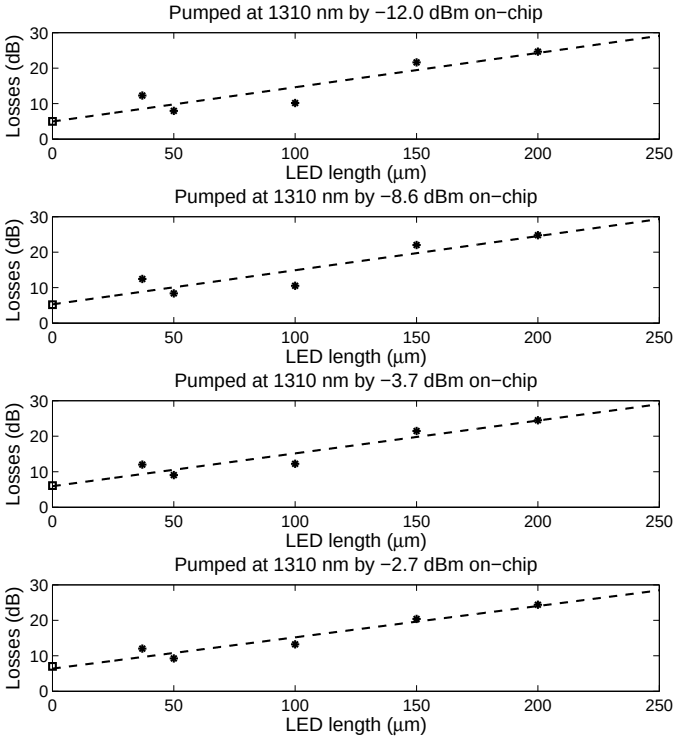


Figure 4.29: Loss of the 1310 nm pump laser through the optical chip for different pump powers. The absorption coefficient is found to be around 1000 dB/cm for all pump powers, while the taper loss is too low to estimate.

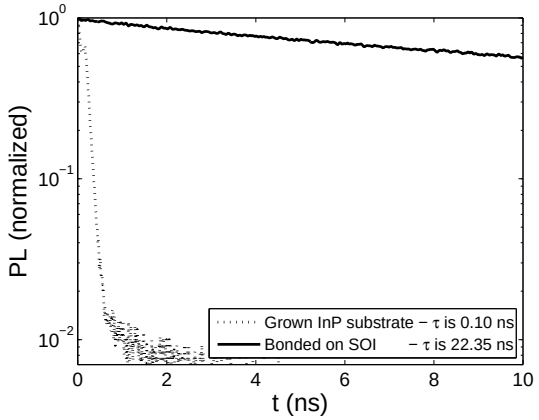


Figure 4.30: Time-resolved PL trace for a membrane grown on InP and a bonded one.

substrate and sacrificial layers, there may not be enough InP material left to force the strained material in the InP lattice and it may relax, thereby potentially causing defects or simply moving to a different lattice constant.

Figure 4.30 shows two time traces as measured using the time-resolved PL measurement setup. The fitted lifetimes have been added in the legend. One trace was measured when the membrane was still on the InP growth substrate. The top InGaAs-InP sacrificial layer pair has been etched away, as the InGaAs would interfere too much with the measurements. There is still a sacrificial InGaAs layer between the membrane and the substrate however. The other trace is taken from a membrane bonded on SOI. The substrate and sacrificial layers have been removed, but there has been no other processing. It is a 122 nm thick sheet of III-V material, BCB bonded to SOI.

The membrane on the InP substrate is showing an extremely short lifetime. This is surprising and the reason for this remains unknown.

More importantly, from figure 4.30 we see the bonded membrane sheet does have a long lifetime, indicating non-radiative recombination is absent. Therefore, we conclude the material quality is preserved during the bonding process.

4.6.2.2 Plasma damage

During a typical fabrication process, the samples can be subject to many different plasmas or accelerated ions. Dry etching is usually done by reactive ion etching (RIE) or inductively coupled plasma etching (ICP) and the most used deposition technique is plasma-enhanced chemical vapor deposition (PECVD). Photoresist residue is usually removed in extended oxygen plasmas. All these processing techniques involve accelerating charged ions onto the sample. The impact of these ions

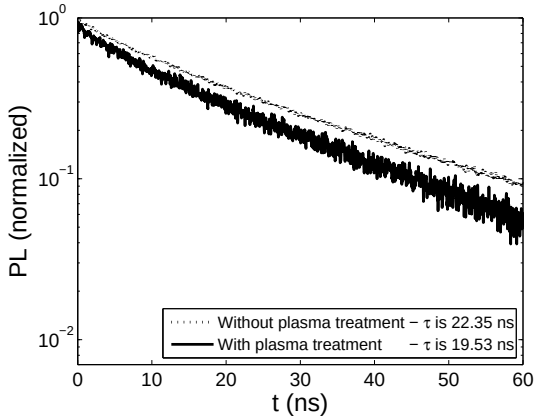


Figure 4.31: Time-resolved PL traces to investigate plasma-induced damage

can create both surface traps as well as bulk defect states [4]. Because of the limited thickness of the III-V membrane, this could also be a concern to us.

Figure 4.31 shows two time-resolved PL traces. The first is a bonded membrane sheet which has seen almost no plasma treatments (only a 20 second PECVD deposition of SiO_2 before bonding). The other trace is taken from an identical membrane, but this has been subject to 1 hour of RIE oxygen plasma. As can be seen from the figure, there is almost no change in the lifetime. Both are very long, around 20 ns.

From this, we can conclude that although the membrane is very thin, plasma-assisted damage does not occur.

4.6.2.3 Growth defects

The most obvious source of recombination centers are growth defects. These are presents because of imperfections during the growth and a problem here is not related specifically to the thin membranes described in this work.

Figure 4.32 shows three time-resolved PL traces from three different growths. The first contains quantum wells surrounded by two 20 nm thick InP layers, the second is identical but the InP layers are 60 nm thick. The third trace is taken from the quantum dot material used in section 4.5.2. The two QW membranes show long lifetimes of around 20 ns, but this is not the case for the QD membrane. Figure 4.26 oval defects were already discovered by doing the substrate removal. This high density of defects is again confirmed by the much shorter lifetime of 1 ns. The reduced lifetime of the quantum dots is in turn reflected in the low efficiency of the devices in section 4.5.2.

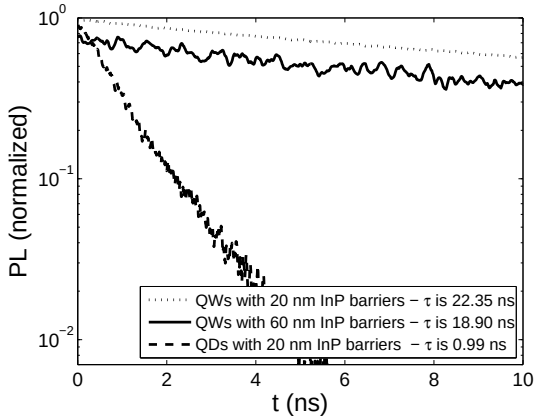


Figure 4.32: Time-resolved PL traces to investigate the effect of growth defects.

4.6.3 Sidewall recombination

Of all the possible origins of non-radiative recombination centers, sidewall recombination is probably the most cited. While carriers may have been created in the center, diffusion ensures that they will find their way to the etched sidewall, where they recombine in the numerous surface states available. There are several ways of mitigating this. The first is to work with devices so wide that only a very limited fraction of the carriers reaches the sidewalls. This approach was used in the SLEDs of chapter 3, where we employed a III-V mesa of 24 μm wide. A second approach is to reduce the number of surface traps by passivating them. In this regard, wet etching is also typically preferred over dry etching. The third option is to use materials with a lower mobility, making sure the carriers do not diffuse as far. We discuss all three approaches below in addition to vertical recombination by tunneling through the InP barrier.

4.6.3.1 Tunneling through the InP barrier

When the pump is absorbed, carriers are created in the entire active region as both the barriers and wells are absorbing the pump wavelength. Vertical recombination is prohibited by the 20 nm thick InP layer. If carriers would be able to penetrate through this barrier, surface recombination would reduce the non-radiative recombination enormously.

Next to the effect of growth defects, figure 4.32 also illustrates this does not occur however. The measured lifetime does not change when the InP layers are thicker.

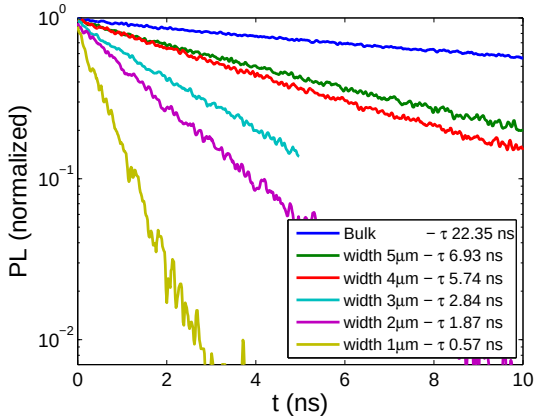


Figure 4.33: When reducing the width of the 122 nm thick membrane, the lifetime decreases significantly.

4.6.3.2 Effect of the membrane width

Due to lateral diffusion, the generated carriers propagate to the etched sidewall of the LED. Contrary to the vertical propagation, there is no grown electron or hole barrier to counteract this diffusion. Once the carriers reach the sidewall, they can recombine non-radiatively through the surface states. Figure 4.33 illustrates that this indeed an effect to be taken into account. While the 122 nm thick membrane sheet material (this is bonded epi to silicon) still exhibits a long lifetime of 22 ns, etching the waveguide clearly reduces the lifetime. For the 3 μm wide device, the lifetime is 2.8 ns, while for a width of 1 μm this has been further reduced to 0.5 ns.

When inspecting the device performance, no significant difference was found though. Figure 4.34 shows what happens when a printed coupon is etched. When the device is 20 μm wide, light is still coupled through the inverted silicon taper with 85% efficiency The III-V is so wide, it can be considered bulk (in fact the limited width comes from the transfer printing process), and hence the lifetime is long. We etch a 1.5 μm wide device with tapering structures in the coupon, but the performance is hardly affected. This was observed for all devices, as it was standard practice to measure the device after transfer printing. Note that the sample did not leave the cleanroom for these measurements. From the discussion on lifetimes above, we expect a performance improvement, while from the simulations shown in figure 4.3 we know the collection efficiency to decrease considerably. These two effect are counteracting one another, canceling the desired improvement.

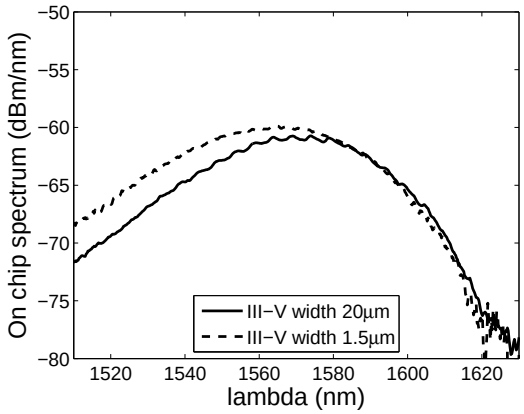


Figure 4.34: When reducing the device width, no performance change is observed. Note that the 20μm width can be considered bulk and the 1.5μm wide device also consists of a tapering structure.

4.6.3.3 S-Passivation of the sidewall

Passivation aims at reducing the surface recombination velocity. The non-radiative lifetime is kept high by making sure that the carriers present at the sidewalls do not recombine non-radiatively (whereas the previous paragraph aimed at reducing the amount of carriers at the sidewalls). If a lot of dangling bonds are present, a high recombination rate can be expected and vice versa. Therefore, a passivation steps tries to reduce the amount of dangling bonds. The best results in InGaAs(P) have been achieved with a treatment based on sulfur ions [5, 6]. In this passivation scheme the sample is first atomically cleaned by an iteration of oxygen plasma and a diluted wet etch. In this manner a few atomic layers are removed and a new, fresh surface is created. Next, a hot solution of diluted ammonium polysulfide $(NH_4)_2S$:H_2O is applied, the sulfur ions reacting to the dangling bonds. The formed amorphous S-film is then removed in a water rinse and the sample is capped in a PECVD SiO_2 layer.

As this method has been hugely successful in the InGaAs-based nanopillars of Higuera-Rodriguez et.al. [6], we collaborated with them to passivate our samples. Figure 4.35 shows the resulting lifetime before and after passivation of a 1.5 μm wide LED. In disagreement to previously reported, the lifetime has not increased but rather went down. The reason for this is unknown and previous attempts on similar devices have also been unsuccessful. Contrary to [3], a proven successful passivation recipe was used, but still the result was unsatisfactory.

Figure 4.36 shows a spectrum of an LED before and after passivation. In this case a grating coupler was used to couple out the light from the SOI chip. The

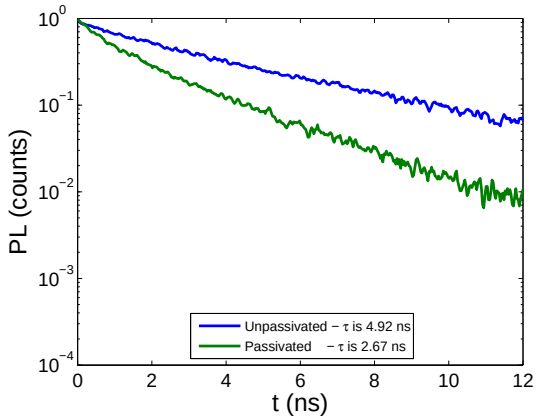


Figure 4.35: Effect of the S-passivation on the carrier lifetime.

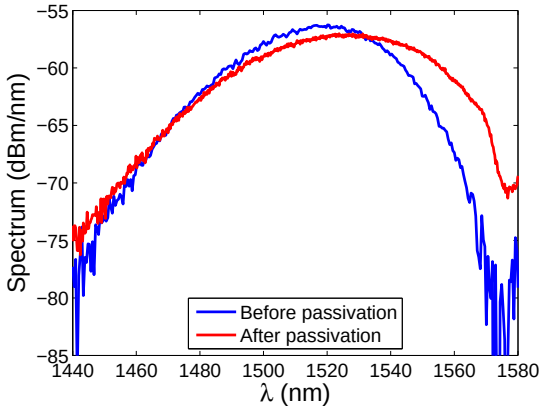


Figure 4.36: Effect of the S-passivation on the device performance.

#	Layer	Thickness	PL
0	InP substrate	360 μm	
1	InP buffer layer	150 nm	
2	$In_{0.532}GaAs$ release layer	1 μm	
3	InP	50nm	
4	InGaAsP (Q=1.2μm)	10 nm	
5	$InAs_{0.65}P_{0.35}$	6 nm	1520 nm
6	InGaAsP (Q=1.2μm)	10 nm	
7	$InAs_{0.65}P_{0.35}$	6 nm	1520 nm
8	InGaAsP (Q=1.2μm)	10 nm	
9	$InAs_{0.65}P_{0.35}$	6 nm	1520 nm
10	InGaAsP (Q=1.2μm)	10 nm	
11	$InAs_{0.65}P_{0.35}$	6 nm	1520 nm
12	InGaAsP (Q=1.2μm)	10 nm	
13	InP	20 nm	
14	$In_{0.532}GaAs$ sacrificial layer	100 nm	
15	InP sacrificial layer	1 μm	

Table 4.3: Epitaxial layer stack with InAsP-based quantum wells

advantage is that it is much easier and more accurate to measure the efficiency of the devices. It also means however, we cannot see the whole spectrum. The grating coupler was centered at 1550 nm so only the band around this wavelength is coupled out efficiently. It can be seen that the sulfur passivation has a negligible result on the device performance, if any.

4.6.3.4 InAsP-based quantum wells to reduce sidewall recombination rate

The third option to counteract lateral non-radiative recombination, is to use a different material altogether. The surface recombination velocity of InAsP quantum wells is lower than that of InGaAsP ones. Depending on the sidewall etching and passivation, this can be one order of magnitude. [7–10]

The layer stack of table 4.3 shows a layer stack similar than the ones already used, but the quantum wells are InAsP-based. Note that the quantum wells are now highly strained (2%). We have avoided this in the previous stacks because the thin membranes may relax after substrate removal (in case of die-to-die bonding) or when released (in case of transfer printing). If we want to use InAsP however, we do not have another option than to strain the layer (and use compensating strain in the barriers).

Figure 4.37 shows the measured lifetime for the bulk sheet material (transfer printed on a silicon substrate) and of two narrow LEDs. The width of the LEDs does not have an effect on the lifetime, which is 5-6 ns for all three, illustrating the

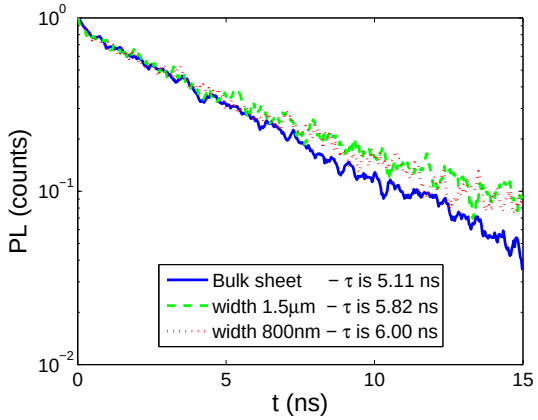


Figure 4.37: Time-resolved PL measurement of InAsP-based membrane. The lifetime is independent of LED width.

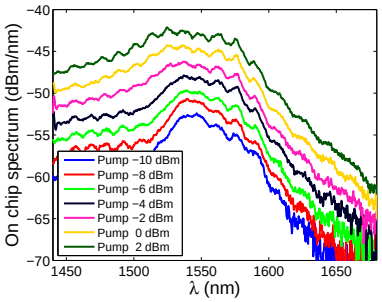
reduced surface recombination rate.

Supported by this promising result, 1.5 μm wide LEDs were fabricated, the measured spectra of which are shown in figure 4.38. There is a clear increase in output power compared to the other, previously fabricated optically pumped LEDs (e.g. figure 4.22). The efficiency of these devices therefore is also higher for both the co- and counter propagating mode. The former operating mode renders an efficiency of -24 dB for all pump powers, while the latter varies slightly from -20 dB to -23 dB. This decrease in efficiency could be due to auger recombination or free carrier absorption. The fact that the co-propagating efficiency does not change for different pump powers indicates that the LED is completely pumped, i.e. the end of the LED is not absorbing the already generated light. The redshift in the spectra is also minimal, confirming the complete pumping.

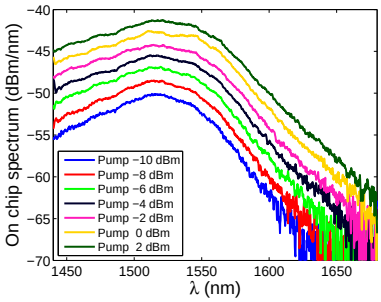
Introducing InAsP quantum wells has clearly improved our devices. The carrier lifetime of 6 ns is independent of the device width due to the low surface recombination velocity. The LED efficiency has increased by almost an order of magnitude.

4.6.4 Reducing the radiative recombination lifetime by means of doping

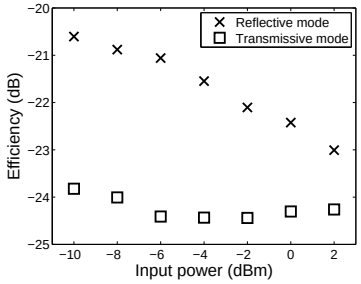
Radiative and non-radiative recombination processes are in constant competition. In the previous paragraphs we tried our best to make the non-radiative recombination lifetime long, i.e. in relation to the radiative lifetime. Typically, III-V compounds exhibit radiative lifetimes of the order of nanoseconds. However, as can be seen in the traces of figures 4.30, 4.32 and 4.33, the bulk sheet material



(a) Spectra in transmissive mode



(b) Spectra in reflective mode



(c) Efficiency as a function of pump power.

Figure 4.38: Characterization of a 50 μm long InAsP-based LED. The pump wavelength was 1310 nm and the on-chip pump power is indicated in the legend.

exhibits an unusually long lifetime of around 20 ns. We can assume the bulk sheet material lifetime is dominated by the radiative one (and otherwise the radiative lifetime is even longer). This means that the radiative lifetime in these devices is extremely long, making it very easy for the non-radiative lifetime to compete.

The reason for this can be found in the low carrier injection regime. Although usually depicted as a constant, the lifetimes depend on the carrier concentration:

$$R_{\text{sp}} = B_{\text{bimol}} NP = B_{\text{bimol}} N^2 = \frac{N}{\tau_{\text{rad}}} \tag{4.5}$$

$$R_{\text{non-rad}} = A_{\text{SRH}} N + C_{\text{Aug}} N^3 \tag{4.6}$$

$$\frac{1}{\tau_{\text{rad}}} = B_{\text{bimol}} N \tag{4.7}$$

$$\frac{1}{\tau_{\text{non-rad}}} = A_{\text{SRH}} + C_{\text{Aug}} N^2 \tag{4.8}$$

with R_{sp} and $R_{\text{non-rad}}$ the spontaneous emission rate and the non-radiative recombination rate, B_{bimol} the bi-molecular recombination coefficient, A_{SRH} the Shockley-Read-Hall recombination coefficient and C_{Aug} the Auger recombination coefficient. N and P are the electron and hole carrier density respectively, which are equal in intrinsic material due to charge neutrality. τ_{rad} and $\tau_{\text{non-rad}}$ are the radiative and non-radiative lifetimes. We can then plug the equation for the radiative and non-radiative lifetimes in a rate equation model. The following discretized equations are solved numerically to a steady solution.

$$P_{\text{pump,abs}}(z) = P_{\text{pump,in}} \frac{e^{-\alpha_{\text{pump}}(z+\Delta z)} - e^{-\alpha_{\text{pump}} z}}{\Delta z} \tag{4.9}$$

$$0 = \frac{P_{\text{pump,abs}}(z)}{h\nu_{\text{pump}} w_{\text{act}} t_{\text{act}}} - \frac{P_{\text{sig,stim}}(z)}{h\nu_{\text{sig}} w_{\text{act}} t_{\text{act}}} \frac{\Delta\lambda}{\Delta z} - \frac{N(z)}{\tau_{\text{nrad}}(z)} - \frac{N(z)}{\tau_{\text{rad}}(z)} \tag{4.10}$$

$$P_{\text{spont}}(z) = \beta \frac{N(z)}{\tau_{\text{rad}}(z)} h\nu_{\text{sig}} w_{\text{act}} t_{\text{act}} s_{\text{spont}} \tag{4.11}$$

$$\begin{aligned} P^+_{\text{sig}}(z+\Delta z) =& P^+_{\text{sig}}(z) + P_{\text{spont}}(z) + \Gamma_{\text{sig}} g_{\text{sig}}(z) P^+_{\text{sig}}(z) \\ & - \alpha_{\text{lin}} \Delta z P^+_{\text{sig}}(z) - \alpha_{\text{FCA}} \Delta z P^+_{\text{sig}}(z) \end{aligned} \tag{4.12}$$

$$\begin{aligned} P^-_{\text{sig}}(z+\Delta z) =& P^-_{\text{sig}}(z) + P_{\text{spont}}(z) + \Gamma_{\text{sig}} g_{\text{sig}}(z) P^-_{\text{sig}}(z) \\ & - \alpha_{\text{lin}} \Delta z P^-_{\text{sig}}(z) - \alpha_{\text{FCA}} \Delta z P^-_{\text{sig}}(z) \end{aligned} \tag{4.13}$$

$$\alpha_{\text{FCA}}(z) = \Gamma_{\text{sig}} \kappa_{\text{FCA}} \left(N(z) + N_{\text{Doping}} \right) \tag{4.14}$$

$$P_{\text{sig,stim}}(z) = \left(P^+_{\text{sig}}(z) + P^-_{\text{sig}}(z) \right) \Gamma_{\text{sig}} g_{\text{sig}}(z) \Delta z \tag{4.15}$$

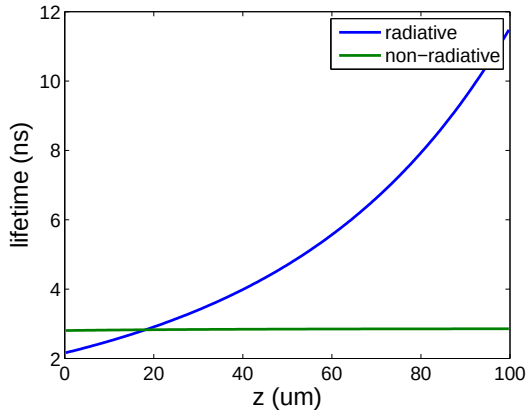
Equation 4.9 describes the absorbed pump power $P_{\text{pump,abs}}$ as a function of input pump power $P_{\text{pump,in}}$ and the absorption coefficient α_{pump}. Because the absorption coefficient is around 1000 dB/cm, all the pump loss can be attributed to direct

absorption. Equation 4.10 describes the excited carriers N in the LED. The first term is the pumping term, with $h\nu_{\text{pump}}$ the pump energy, w_{act} and t_{act} the width and thickness of the active region. In the second term $P_{\text{sig,stim}}$ is related to both the stimulated recombination of the signal wavelengths as well as the absorption. Therefore this term will become negative if population inversion is not achieved. In this case, the term will aid in the pumping of the membrane. The third and fourth term of equation 4.10 are the non-radiative and radiative recombinations. Note that the lifetimes τ_{nrad} and τ_{rad} also depend on the carrier concentration N through equations 4.8 and 4.7 (undoped) or equations 4.17 and 4.16 (doped). Equation 4.11 depicts the coupled spontaneous emission P_{spont} wherein β is the collection efficiency which was boosted by the high refractive index contrast. s_{spont} is a shapefunction for the spontaneous emission. The total integral of this shapefunction is 1 and it brings the carrier-dependent wavelength spread of the signal in the equations. Equations 4.12 and 4.13 describe the forward and backward propagating signals P^{+}_{sig} and P^{-}_{sig}. They each consist of a propagation term, a collected spontaneous emission term, a term related to stimulated emission and direct absorption. The gain factor g_{sig} was approximated by polynomials as in [11]. The two last terms are due to the linear loss α_{lin} and the free carrier absorption α_{FCA}, which in turn is described in equation 4.14. Equations 4.12 and 4.13 are also linked to each other through a front and back reflectivity (not shown in the equation). In equation 4.14, Γ_{sig} is the confinement of the signal in the active region and κ_{FCA} the free carrier loss coefficient. The last equation 4.15 calculates the power absorbed or in stimulated emission of the forward and backward propagating signals.

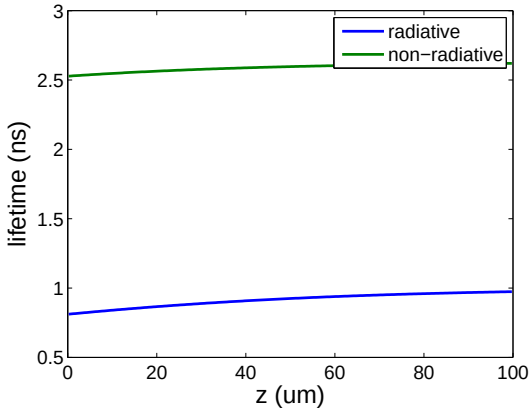
Equations 4.9 to 4.15 are solved numerically and must be iterated through until there is a stable solution. The discretization of both the LED spatial segments and spectral distribution must be fine enough to resolve the necessary features. A too fine mesh will result in a long computation time though. The values of the used independent parameters are listed in table 4.4

In electrically injected devices the carrier density will be pushed to 10^{18} carriers/cm^3, which renders a radiative lifetime τ_{rad} of around 1 ns. In the optically pumped LEDs we depict however, we do not create as many carriers. As a result, the radiative lifetime becomes long, letting the non-radiative recombination dominate the carrier population. This is illustrated in figure 4.39a, which shows a simulation of the LED along its length, in particular the carrier lifetimes. Because the pump light is depleted by the end of the LED, the radiative lifetime increases and surpasses the non-radiative one.

The radiative lifetime can be decreased by doping the active layers. Equations



(a) Undoped active region: the non-radiative lifetime is dominating.



(b) N-doped active region: the radiative lifetime is dominating.

Figure 4.39: Simulated effect of doping on the (non-)radiative carrier lifetime.

Symbol	Parameter	Value
$P_{\text{pump,in}}$	Input pump power	-10 dBm
L_{LED}	LED length	100 μm
α_{pump}	Pump absorption coefficient	1000 dB/cm
λ_{pump}	Pump wavelength	1310 nm
A_{SRH}	Shockley-Read-Hall recombination coefficient	3.5 $\times 10^{8} s^{-1}$
B_{bimol}	Bimolecular radiative recombination coefficient	1 $\times 10^{-15} m^{3}/s$
α_{lin}	Linear loss	3 dB/cm
w_{act}	width of the active region	1 μm
t_{act}	thickness of the active region	40 nm
β	Spontaneous emission collection efficiency	12%
R_{front}	Power reflectivity at front facet	0.01
R_{back}	Power reflectivity at back facet	0.01
κ_{FCA}	Free carrier loss coefficient	1 $\times 10^{-18}$ cm^2

Table 4.4: Independent parameters in the rate equation model of the optically pumped LED. [12]

4.7 and 4.8 alter into the following equations:

$$\frac{1}{\tau_{\text{rad}}} = B_{\text{bimol}} \left(N + N_D\right) \tag{4.16}$$

$$\frac{1}{\tau_{\text{non-rad}}} = A_{\text{SRH}} + C_{\text{Aug}} \left(N + N_D\right)^2 \tag{4.17}$$

where N_D is the doping concentration. Note that if N_D is 0, the equations reduce to the undoped case. By increasing the doping concentration, we can decrease the radiative lifetime as we wish. The non-radiative lifetime will decrease as well due to increased Auger recombination, but this will only be significant when doping strongly. Note that the optical losses will also increase, which can be modeled as free carrier absorption. Figure 4.39b shows the lifetimes throughout the LED when the active layer is doped. The input power was -10 dBm and the N-doping concentration was 10^{18} cm^{-3}. The radiative lifetime now stays below 1 ns for the entire device. There is still a small z-dependence because the generated carrier concentration close to $z = 0$ is not negligible compared to the doping concentration, i.e. N cannot be neglected compared to N_D in the equations.

Figure 4.40 shows the time-resolved PL measurements of InGaAsP QWs with three different doping levels. The membranes were 152 nm thick and we always measured in the middle of sheet material of 40 μm by 100 μm, which was transfer printed on silicon. One membrane was nominally undoped, the second had an N-dopant concentration of 2×10^{17}cm^{-3} and the third had an N-dopant concentration of 1×10^{18}cm^{-3}. The effect on the lifetime is clear, there is a decrease from 30 ns over 23 ns to 12 ns. Even though this is a huge improvement, 12 ns is still a high value. Higher doping concentrations would bring this down further, but the optical

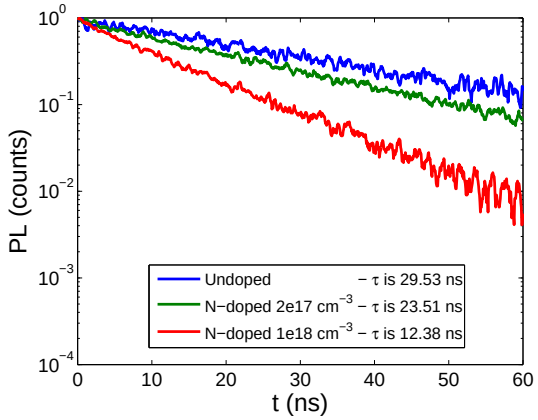
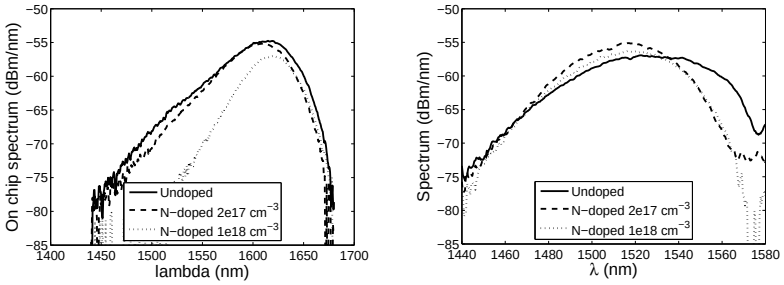


Figure 4.40: Time-resolved PL measurement illustrating the effect of doping on the radiative lifetime



(a) 40 μm wide LED *(b) 1.5 μm wide LED*

Figure 4.41: LEDs with different doping concentrations behave almost identical.

losses would also increase annihilating the positive effect.

Figure 4.41 shows the effect of doping on the LEDs. Identical coupons as the ones from figure 4.40 were printed on SOI, where a grating coupler could be used to extract the efficiency of the LEDs. Both very wide LEDs (these were actually the coupons as printed) and narrower devices with proper III-V tapers were characterized. The devices behave more or less identical however, independent of doping concentration. For the narrow LED, this could be expected since the non-radiative recombination rate will still exceed the radiative one by far. This argument does not hold for the wide one though, where we have measured the carrier lifetime to be long.

4.6.5 Summarizing the second generation of devices

After the lower-than-expected performance of the first generation of devices, a lot of tests were carried out. The power efficiency was eventually boosted to -20 dB. We summarize and compare the tests below.

Transmission measurements confirmed the quality of the optical paths. The pump absorption was estimated around 1000 dB/cm, with the taper losses too low to measure.

Since the optics behave as expected, we took a closer look at the carrier dynamics. The bulk material quality remains high. The process of bonding and removing the substrate did not have an effect on the carrier lifetime nor did extensive submission to oxygen plasma leave an impact. Growth defects did show detrimental effects, as was shown in the QD-based membrane. Other non-radiative recombination centers are located at the sidewalls. Vertically there is no issue, as the 20 nm InP barrier is shown sufficient to contain the carriers. Horizontally, the time-resolved PL setup learns us that the width of the device has a critical impact on the carrier lifetime. Devices of different width showed similar performance though, as the reduction in collection efficiency canceled the longer lifetime. After a failed passivation attempt, InAsP based quantum wells were able to bring down the effects of the sidewall recombination centers. The efficiency of the devices is boosted to -20 dB and -24 dB in reflective and transmissive mode respectively. The carrier lifetime was 6 ns, independent of device width.

As the radiative lifetime was measured to be exceptionally long, we doped the InGaAsP-based membranes. TR-PL measurements confirmed the effect of doping on the carrier lifetime, but the improvement was too limited to measure in devices. It is possible doping could increase the efficiency of the InAsP-based membranes further since the reduced recombination velocity makes the effects on radiative lifetime more outspoken. Due to lack of time, this route was not yet explored.

4.7 Conclusion

In this chapter we introduced a novel device concept: optically pumped LEDs based on single-mode high contrast III-V–on–Si membranes. This device tries to marry the high efficiency of lasers and the simplicity of SLEDs. The depicted LED is tolerant and simple making it possible to achieve the high yield needed for an integrated light source for low-cost spectroscopic sensing applications.

No longer requiring the electrical pumping opens up previously inaccessible design freedom. We can use this flexibility to boost the coupling efficiency of the isotropic spontaneous emission to the fundamental mode. By utilizing the high refractive index contrast of a thin III-V membrane in oxide/nitride environment, we

can achieve up to 12% coupling into the waveguide mode of a 3 μm wide device. For very narrow devices, 1 μm width, this number increases to 26%. Having removed the layers necessary for electrical pumping, an optical pumping scheme must now be applied. Because the confinement of the pump is very high and the tapers provide near-perfect coupling, the pump can be absorbed efficiently. The resulting device is both simple and tolerant. A broad range of simulations supports the idea outlined in this chapter.

The first generation of devices shows the proof of principle of these devices. Both quantum well-based and quantum dot-based membranes were integrated on SOI. These devices were characterized in detail. The large optical bandwidth results from the broadband nature of spontaneous emission, and was even further boosted in the quantum dot-based devices where the different layers of dots were chirped. The bandwidth was 220 nm and the efficiency around -30 dB.

The efficiency of the first generation of devices was below expectations, and hence a large range of experiments were carried out to find the culprit. The optics was found to work, with taper losses too low to measure and absorption of around 1000 dB/cm. If the non-radiative carrier lifetime is much lower than the radiative one, none of our generated electron-hole pairs will yield photons. Bulk defects could be ruled out, since neither plasma damage nor bonding and substrate removal affect the lifetime. The influence of growth defects is obvious, but only the quantum dot-based membranes suffer from this. Looking at sidewall recombination, the lack of a vertical recombination path is also confirmed by looking at different InP barrier layers. The effect of the waveguide width was made clear with the time-resolved photoluminescence measurements. Device measurements support this, where no different in performance was measured even though the collection efficiency of slab structures is much lower. Despite using a proven recipe, a S-passivation of the sidewalls was ineffective, as established by both TR-PL and device measurements. When changing from InGaAsP quantum wells to InAsP-based ones, the desired efficiency improvement was achieved. The lower recombination velocity of InAsP ensures a reduced sidewall recombination, thereby boosting the performance to -20 dB. Other than increasing the non-radiative recombination lifetime, we tried to decrease the radiative one. Doping the InGaAsP material did the trick, but to a limited effect if one wants avoid excessive optical doping losses.

After the extensive series of optimization experiments a discrepancy between the theoretical and measured efficiency remains. Where the theoretical optimum lies around 10%, the best measured power efficiency is 1%. The InAsP-based membranes clearly improved the performance considerably, highlighting the importance of non-radiative recombination. The last order of magnitude has not been accounted for however. The origin of this last difference is not clear, and needs to be looked into further. It is our belief, the non-radiative recombination is still limiting the device. With this understanding, combined with the effects of doping

on radiative lifetime, we think doped InAsP-based membranes to be an attractive route forward and the next step in this story.

References

[1] Stevan Stankovic. *Hybrid III-V/Si DFB lasers based on polymer bonding technology*. PhD thesis, Ghent university, 2012.

[2] Accuratus. *Silicon nitride Ceramic properties*, May 2017. http://www.accuratus.com/silinit.html.

[3] Martijn Tassaert. *Leveraging high index contrast for optical signal processing applications*. PhD thesis, Ghent University, 2014.

[4] J Weber. *Defect generation during plasma treatment of semiconductors*. Physica B: Condensed Matter, 170(1-4):201–217, 1991.

[5] Rama Raj, Fabrice Raineri, Alexandre Bazin, and Paul Monnier. *Subduing or Harnessing Surface Recombination for CW Nanolaser or Fast Optical Switch Heterogenously Integrated on SOI*. In Integrated Photonics Research, Silicon and Nanophotonics, pages IM2A–5. Optical Society of America, 2013.

[6] Aura Higuera Rodriguez, Bruno Romeira, Simone Birindelli, Lachlan Black, Barry Smalbrugge, Erwin Kessels, Meint Smit, and Andrea Fiore. *Ultra-low surface recombination for deeply etched III-V semiconductor nano-cavity lasers*. In Integrated Photonics Research, Silicon and Nanophotonics, pages ITu2A–2. Optical Society of America, 2016.

[7] ON Oliveira Jr, Maria Raposo, A Em Dhanabalan, and HS Nalwa. *Handbook of Surfaces and Interfaces of Materials*. Handbook of Surfaces and Interfaces of Materials, 4, 2001.

[8] Jingqing Huang, Se-Heon Kim, Jonathan Gardner, Philippe Regreny, Christian Seassal, Pablo Aitor Postigo, and Axel Scherer. *Room temperature, continuous-wave coupled-cavity InAsP/InP photonic crystal laser with enhanced far-field emission directionality*. Applied Physics Letters, 99(9):091110, 2011.

[9] Larry A Coldren, Scott W Corzine, and Milan L Mashanovitch. *Diode lasers and photonic integrated circuits*, volume 218. Wiley. com, 2012.

[10] Georgiana Dagnall, Jeng-Jung Shen, Tong-Ho Kim, Robert A Metzger, April S Brown, and Stuart R Stock. *Solid source MBE growth of InAsP/InP quantum wells*. Journal of electronic materials, 28(8):933–938, 1999.

[11] J Leuthold, M Mayer, J Eckner, G Guekos, H Melchior, and Ch Zellweger. *Material gain of bulk 1.55 μm InGaAsP/InP semiconductor optical amplifiers approximated by a polynomial model*. Journal of Applied Physics, 87(1):618–620, 2000.

[12] Michael J Connelly. *Wideband semiconductor optical amplifier steady-state numerical model*. IEEE Journal of Quantum Electronics, 37(3):439–447, 2001.

5 Conclusion and outlook

In today's globalized world, diseases also occur at a global scale. This means that many hundreds of millions of people (will) suffer from diabetes, tuberculosis and suchlike and that a solution must be of similar magnitude. The silicon platform can couple mass production and low cost to each other through its unique scaling laws. It is because of this reason silicon photonics targets these 'great' diseases.

The scaling story only holds when all the steps follow the same scaling laws. Because of the indirect band gap of silicon, efficient, integrated light sources are problematic and hence, the entire scaling story is threatened. To mitigate this, we have introduced InP-based material to the silicon photonics ecosystem. In this chapter we want to formulate an overview of our research on broadband light sources on silicon as well as provide the reader with the future prospects we anticipate.

5.1 Conclusion

Our research (and therefore also this dissertation) was focused both on integration technology as well as device designs. For clarity, the conclusions are also kept separate.

5.1.1 Technology

Silicon photonics has need of external materials which are able to emit broadband light in an efficient way. Moreover, multiple band gaps on-chip allow for

wavelength multiplexing as is needed for many sensing applications. We looked at three different techniques of introducing more than one active band gap to silicon: quantum well intermixing, multiple die bonding and transfer printing.

Quantum well intermixing allows one to blueshift the quantum wells at certain locations of the InP die. In our implementation of QWI, implantation enhanced disordering, a sacrificial buffer layer is implanted with phosphorous ions. During the thermal anneal, these then act as catalysts to change the composition of the well. After creating the different band gaps on the InP chip, it was bonded to a Si PIC to create light sources. This molecular bonding was aligned, so as to provide the right band gaps to the right location. Using an empirical relation and a mathematical model, we could show that the diffusion coefficient for atom disordering of the two epitaxial layer stacks used in this work were significantly different. We attributed this to the growth quality, where parasitic interstitial defects can aid the diffusion. After establishing the fabrication process, we made lasers to assess the quality of the intermixed quantum wells. Fabry-Perot lasers showed very similar results, where the blueshifted ones would even outperform the as-grown ones. Therefore we can conclude the QWs remain of high quality during the disordering. This technique was then also applied to create broadband electrically pumped SLEDs.

Multiple die bonding seems the most straight-forward option to create multiple band gaps op chip, since the die-to-wafer bonding process has matured for both molecular and BCB bonding. Obstacles to overcome were the height difference of the two InP substrates and the protection of the exposed area between the two dies. The first was mitigated using a graphite sheet, for the latter we applied a conformal oxide to fill the gap between the substrates. The dimension of this gap was brought below 100 μm for both molecular and BCB bonding. The conformal oxide, which is opened in a self-aligned manner, protects the silicon waveguides during the post-bond process and also fills the waveguide trenches. Since we are using airclad trenches for the molecular bonding, HCl-based etchants can travel underneath the die and attack the bottom device layer, causing so-called wicking. By using the conformal oxide and having a good cleave this wicking was reduced from $\pm 700\mu$m to zero. Finally, we also combined aligned multiple die bonding with quantum well intermixing, in order to create four III-V band gaps on a Si PIC and a wideband SLED source.

Transfer printing is a new integration technique altogether. It was developed for photonic integrated circuit applications for the first time in this work and allows for more efficient use of III-V. In transfer printing a rubber stamp is used to pick and print a III-V coupon to a Si PIC. These coupons are only pickable if they are free-hanging. To do this, an entire fabrication process was developed, starting from the identification of the correct release etchant. Cold $FeCl_3$:H_2O provides the needed selectivity of over 2000, while all the vital layers were encapsulated

by a combination of SiO_2 and photoresist. The photoresist tethers hold the entire structure and break easily during the picking. Although for unpatterned SOI and airclad patterned SOI we could rely on Van Der Waals forces to have a proper bond, we saw delamination for planarized SOI, our platform of preference. A thin layer of BCB at the bonding interface resolved this issue, and a strong bond was established that could withstand the further processing. Transfer printing was used to fabricate the optically pumped LEDs, although one could say there was no particular reason for this. The fact that we preferred transfer printing over traditional bonding methods (once the TP process was established) shows how reliable and easy the integration by means of transfer printing already is.

Three techniques of creating multiple active band gaps were described and a comparison is in order.

It is worth noting that QWI can be combined with the other two. Especially QWI and multiple die bonding are very complimentary. The power of QWI lies in the fact that multiple band gaps can be created within a spatial resolution of 1 μm, giving the designer a lot of flexibility in the location of the band gap. The downside is that only the shape of the quantum wells can be changed, and even those within a limited range. Multiple die bonding provides full freedom in the layer stack, but since the minimum dimension of a bondable die is a few millimeters, the positioning of the band gaps has a lot of restraints. This can make the design of the Si PIC very awkward with a large footprint as a result. Clearly the one technique shows strength where the other does not.

Also QWI and transfer printing could be combined, the reason why is difficult to see. Transfer printing already offers a spatial resolution equal to that of quantum well intermixing and therefore the resolution argument no longer holds. By using different donor epitaxy, completely different layer stacks can be transfer printed. One reason to consider QWI in this context anyway, is that several band gaps can be created within one mesa, i.e. there is no need to couple to the silicon waveguide and back up through the tapers. Only in very specific projects will this outweigh the tricky fabrication where especially the high temperature anneal is a potential showstopper.

Of the two integration techniques, i.e. multiple die bonding and TP, the latter would come out on top for most projects. The minimum bondable die size of a few millimeters is a design limitation hard to overlook. With most Si PIC footprints this comes down to two bonded dies per photonic circuit. The left half of the PIC is covered with layer stack A, the right half with layer stack B. For a laser (stack A) - modulator (stack B)- detector (stack A) circuit, U-shaped designs are inevitable. For more complex circuits, the routing becomes cumbersome very quickly. In TP, the coupons can be as small as needed and they can be placed at will. Very little thought of the specific location of the band gaps is needed, allowing the designer

to prioritize other requirements. The maturity of multiple die bonding should be considered higher than that of transfer printing though. Multiple die bonding lies very close to the traditional bonding methods, which have had the chance to grow reliable over more than a decade. The different thicknesses of the III-V material is a minor change in that respect. Transfer printing is newer and is really a different technique, but it can recuperate a lot of the bonding knowledge already achieved. Therefore we can expect it to mature very quickly. The choice between die bonding (single or multiple) and transfer printing, will mostly come down to density of III-V material. If a high density of III-V material is needed, the traditional bonding process may be chosen over the release process of TP. If not, a single release of coupons can populate a lot of silicon PICs, making it more convenient. We did not see a difference in performance between the different methods.

5.1.2 Devices

We studied two different LED approaches in this dissertation: more traditional electrically pumped SLEDs and a radical novel approach of optically pumped membrane-based LEDs.

The electrically pumped SLEDs emitted a very broad light spectrum due to the integration with QWI and multiple die bonding. In essence we created four sub-devices, each with its own band gap, which were combined in series. By pumping each appropriately, a 3 dB bandwidth of almost 300 nm was achieved. Both the design and fabrication were discussed in detail.

Using very thin membranes, the collection efficiency of spontaneous emission in the fundamental waveguide mode was boosted by two orders of magnitude to 12 %. This observation was the basis to propose a novel device structure: the opically pumped LED. The idea hinges on the fact that there is need of an efficient light source on silicon, which is simple and tolerant enough to be produced with high yield. The target application is low-cost sensing, where the mass production of silicon can make a difference, lowering the cost of a single chip because of the high volumes. After proposing the initial idea, a series of simulations supported our initial assumptions. A first generation of devices with InP-based quantum wells and InAs-based QDs was realized, showing a proof of principle of our new device architecture. The efficiency of this first generation stayed below expectations though, at 0.1 %. A thorough investigation was launched to find the origin of this lower efficiency, culminating in an improved efficiency of 1 % for the InAsP-based QWs. These suffer less from sidewall recombination due to the reduced recombination velocity. While this is still one order below the theoretical maximum, the improvement is clear. We believe the remaining gains are to be found by improving the radiative recombination rate.

Both LED device structures are interesting to study and have their own place in the silicon photonics ecosystem. The electrically pumped SLED can only be efficient if it is pumped sufficiently to population inversion. In that case, stimulated emission becomes the dominant phenomenon boosting the efficiency. The spontaneous emission collection efficiency is inherently low, making it a poor process to rely on. For the optically pumped LED the opposite is true. The high collection efficiency makes spontaneous emission an interesting recombination process to exploit. Keeping the carrier density low avoids non-radiative recombination such as Auger recombination, and keeps the stimulated emission low. The thin membranes will scare some of the industrial players. Because the membrane is only 100-200 nm thick, damage might easily occur. In our tests, we did not find any bulk defects nor vertical recombination even though we specifically looked for it. Still more tests may be required to convince a company to fully invest in it. The electrically pumped SLED has been around for so long, that this is not an issue for them.

5.2 Outlook

Since 2012 I have focused on the research of broadband sources on silicon by means of III-V integration. The achievements are listed above, I would also like to present a (read 'my') window to the future. Again the division between technology and devices must be made.

5.2.1 Technology

Quantum well intermixing, multiple die bonding and transfer printing are three techniques of creating multiple band gaps in III-V, integrated on Si. While all three are interesting and worth further investigating, our preference goes to transfer printing. It is the new kid on the block, but we expect it to quickly become mainstream. Of the three, it has the biggest growth margin while it already proved to be reliable in our optically pumped LEDs. TP has multiple advantages over the other techniques, such as efficient use of III-V and flexibility in both the layer stack and the location it is provided. To illustrate our confidence in TP: We used it extensively in the second generation of optically pumped LEDs. This was not because of any of the above-mentionned benefits, but solely because it was the more convenient and time-saving approach of integrating the III-V membranes with SOI. While the InP release process takes approximately a week of fabrication, we could repeat the transfer printing process very well and one released sample can last for more than one PhD. We only had to do new releases when a new material stack was introduced.

In this dissertation we used transfer printing to provide InP-based sheet material to the Si PIC. There are many more possibilities with transfer printing to be looked into, both in terms of material as in terms of coupon design.

First of all, GaAs material has been shown to release very easily. SOI on the other hand could also be released by etching the buried oxide layer. As this would be an HF-based etchant, encapsulating the entrails of the coupon is not straightforward. PECVD-deposited amorphous silicon could do the trick, but the removal of that layer could then again be problematic. As for the target substrate, passive SOI was used so far. While this is interesting, printing to a so-called full platform silicon photonic chip (where germanium detectors, heaters and so forth have been integrated) opens up much more options. This means printing in an etched well, but initial tests do not show problems here.

Regarding the coupon design, ours was aimed at providing III-V sheet material. One could also fabricate devices on the source wafer, release them and print them onto a Si PIC. This requires highly accurate printing, where the specification of 1 μm misalignment (3σ value) can be enough. Alignment tolerant structures have been designed for this purpose. Transfer printing now resembles less the traditional bonding, but leans towards flip-chip integration. The main advantage here is the parallelism a multi-post stamp brings. A matrix of devices can be transferred in one go, retaining high throughput. In terms of devices we think of pure III-V-based light sources, but also III-V–on–Si ones. The coupling scheme may be adiabatic, but also butt coupling can be interesting.

The bottom line is that transfer printing is very flexible and only a very small subset of the photonic capabilities have been looked into. At the same time, it rapidly became very reliable and well-controlled. This combination is a guarantee for (potential) success and therefore ideal for further research.

5.2.2 Devices

The two proposed device structures, the multi-band gap electrically pumped SLED and the optically pumped membrane-based LED, are both interesting research topics. Moreover, we believe they both have their applications.

The electrically pumped SLED was a milestone in multi-bandgap integration. Never before have so many different active band gaps been integrated on one Si chip. The series design of LEDs is elegant since no multiplexing structures are needed to collect the light in one silicon waveguide. A parallel approach should still be investigated though. The power of silicon photonics does not lie in the III-V but in the silicon passive structures. One should try to exploit this strength as much as possible and that is where parallel sources come into play. Keeping all the sources parallel allows to optimize them separately much more and one could avoid losses like free carrier absorption.

The optically pumped LED was only proposed for the first time in this dissertation and therefore is brand-new. While the optical circuitry is proven to be tolerant to imperfections, the behaviour of the generated electron-hole pairs is not. The sidewall recombination was overcome by using InAsP-based QWs, and therefore further research should focus on this material. The efficiency now lies around 1% and there is still room for improvement. Also the effect of doping on the radiative lifetime is promising. A clear reduction of radiative lifetime of InGaAsP-based membranes was observed, but this did not translate into improved device efficiency. In the InAsP-based membranes the radiative recombination is more dominant and thus also the effect of doping could be more prominent. Furthermore, given the extremity of the membrane thickness, an industrial player will want more research on bulk defects and other damages. To be clear, this is no critique on our own research, but multiple, independent researches will be necessary to convince the industry.

A

Solving Schrödinger's equation with the transfer matrix method

This appendix outlines how to solve the Schrödinger's equation for a random potential profile. It makes use of the transfer matrix method and is based on the work of Jonsson and Eng. [1]

This solver was implemented because the intermixed active region gives rise to potential wells dominated by a diffusion length. Therefore, section 2.2.2.1 makes extensive use of this paragraph, where the relation between diffusion length and PL peak wavelength shift was established.

A.1 Theory

The properties of particles in a semiconductor material can be found by solving the Schrödinger equation:

$$j\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\cdot\nabla^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t) \tag{A.1}$$

where the variables $\mathbf{r}$ and t denote space and time coordinates. The particle mass is m, $\hbar$ is the reduced Planck constant and j is the imaginary unit. The potential in which the particle is moving is described by $V(\mathbf{r},t)$. As we are mostly interested in steady-state quantum well profiles, we can safely assume this function to be time independent. If we also adopt the effective-mass approximation and only

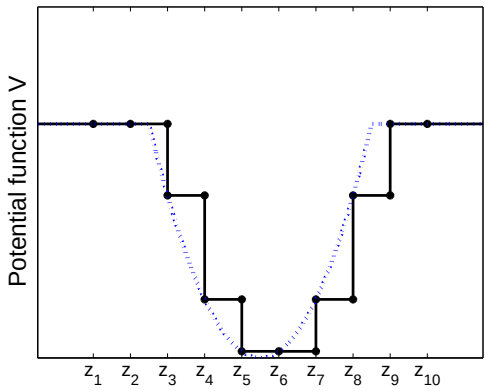


Figure A.1: Approximating the potential profile by a piecewise constant profile consisting of N layers.

steady-state problems in one dimension, the above equation reduces to the time-independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2}\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{m(z)}\frac{\mathrm{d}}{\mathrm{d}z}+V(z)\right]\psi(z)=E\psi(z) \tag{A.2}$$

where $\Psi(z,t)=\psi(z)e^{-j\frac{Et}{\hbar}}$ describes the link between the two above equations. $m(z)$ is now the particle's effective mass in the semiconductor lattice and E is the energy.

Equation A.2 can only be solved analytically for a very restricted amount of problems, e.g. a free particle, the hydrogen atom, infinite and finite square quantum wells and the harmonic oscillator. The potential profiles after quantum well intermixing however are not as nicely behaved and therefore have to be solved numerically. In the transfer matrix method, the potential profiles $V(z)$ is approximated by a piecewise function where within each segment it is constant, shown in figure A.1. [1] There are in total N layers in this piecewise potential. Layer $\#0$ takes care of the region $-\infty < z <= z_1$, while layer N describes $z_{N-1} <= z < +\infty$. Within each layer i, the potential function $V_i(z)$ and the effective mass $m_i(z)$ are constant. Hence, within each layer i, we are describing a free particle, which as the following solution:

$$\psi_i(z)=F_ie^{jk_iz}+R_ie^{-jk_iz} \qquad (z_i <= z < z_{i+1}) \tag{A.3}$$

$$k_i=\frac{\sqrt{2m_i(E-V_i)}}{\hbar} \qquad (E>V_i) \tag{A.4}$$

$$k_i=j\frac{\sqrt{2m_i(V_i-E)}}{\hbar} \qquad (E<V_i) \tag{A.5}$$

If E is larger than V_i, k_i is real, which means $\psi_i(z)$ consists of a forward and backward propagating wave. If E is smaller than V_i on the other hand, k_i is purely imaginary, meaning that $\psi_i(z)$ does not contain a propagating wave but only exponenially decaying waves. Note that so far, we have not said anything about the shape of the wave function except that is must be possible to write it as a linear combination of complex exponential function as in equation A.3.

The boundary conditions at the border between two layers determine the complex wave function coefficients F_i and R_i. Both the wave function itself and its derivative (scaled by the effective mass) must be continuous:

$$\psi_i(z_{i+1}) = \psi_{i+1}(z_{i+1}) \tag{A.6}$$

$$\frac{1}{m_i}\frac{\partial\psi_i(z_{i+1})}{\partial z} = \frac{1}{m_{i+1}}\frac{\partial\psi_{i+1}(z_{i+1})}{\partial z} \tag{A.7}$$

The first expresses the continuity of the wave function across a boundary, the second can be found by integrating both sides of the time independent Schrödinger's equation A.2 over an infinitesimal range across the boundary:

$$E\int_{-\frac{\delta}{2}}^{\frac{\delta}{2}}\psi(z)\mathrm{d}z = \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}}\left[-\frac{\hbar^2}{2}\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{m(z)}\frac{\mathrm{d}}{\mathrm{d}z} + V(z)\right]\psi(z)\mathrm{d}z$$

$$\left[E - \frac{1}{V_>}\right]\psi(z=0)\delta = -\frac{\hbar^2}{2}\left[\frac{1}{m(z)}\frac{\mathrm{d}}{\mathrm{d}z}\psi(z)\right]_{z=-\frac{\delta}{2}}^{z=\frac{\delta}{2}}$$

In the limit of $\delta \to 0$ one finds the boundary condition A.7.

Equation A.3 must now be plugged into the boundary equations A.6 and A.7:

$$F_i e^{jk_i z_{i+1}} + R_i e^{-jk_i z_{i+1}} = F_{i+1}e^{jk_{i+1}z_{i+1}} + R_{i+1}e^{-jk_{i+1}z_{i+1}} \tag{A.8}$$

$$\begin{aligned} F_i\frac{jk_i}{m_i}e^{jk_i z_{i+1}} - R_i\frac{jk_i}{m_i}e^{-jk_i z_{i+1}} &= F_{i+1}\frac{jk_{i+1}}{m_{i+1}}e^{jk_{i+1}z_{i+1}} \\ &\quad - R_{i+1}\frac{jk_{i+1}}{m_{i+1}}e^{-jk_{i+1}z_{i+1}} \end{aligned} \tag{A.9}$$

When we divide equation A.8 by $e^{jk_i z_{i+1}}$ and equation A.9 by $\frac{jk_i}{m_i}e^{jk_i z_{i+1}}$, we get:

$$F_i + R_i e^{-j2k_i z_{i+1}} = F_{i+1}e^{j(k_{i+1}-k_i)z_{i+1}} + R_{i+1}e^{-j(k_{i+1}+k_i)z_{i+1}} \tag{A.10}$$

$$\begin{aligned} F_i - R_i e^{-j2k_i z_{i+1}} &= F_{i+1}\frac{k_{i+1}}{k_i}\frac{m_i}{m_{i+1}}e^{j(k_{i+1}-k_i)z_{i+1}} \\ &\quad - R_{i+1}\frac{k_{i+1}}{k_i}\frac{m_i}{m_{i+1}}e^{-j(k_{i+1}+k_i)z_{i+1}} \end{aligned} \tag{A.11}$$

To bring out F_i and R_i, we can now subtract and add the two equations above

respectively. We then bring the factors to the right side:

$$F_i = F_{i+1}\left(\frac{1}{2}+\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{j(k_{i+1}-k_i)z_{i+1}} + R_{i+1}\left(\frac{1}{2}-\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{-j(k_{i+1}+k_i)z_{i+1}} \tag{A.12}$$

$$R_i = F_{i+1}\left(\frac{1}{2}-\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{j(k_{i+1}+k_i)z_{i+1}} + R_{i+1}\left(\frac{1}{2}+\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{-j(k_{i+1}-k_i)z_{i+1}} \tag{A.13}$$

As the name suggests, in the transfer matrix method it is customary to write these equation in matrix form:

$$\begin{pmatrix} F_i \\ R_i \end{pmatrix} = M_i \begin{pmatrix} F_{i+1} \\ R_{i+1} \end{pmatrix} \tag{A.14}$$

with

$$M_i(1,1) = \left(\frac{1}{2}+\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{j(k_{i+1}-k_i)z_{i+1}} \tag{A.15}$$

$$M_i(1,2) = \left(\frac{1}{2}-\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{-j(k_{i+1}+k_i)z_{i+1}} \tag{A.16}$$

$$M_i(2,1) = \left(\frac{1}{2}-\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{j(k_{i+1}+k_i)z_{i+1}} \tag{A.17}$$

$$M_i(2,2) = \left(\frac{1}{2}+\frac{k_{i+1}m_i}{2k_i m_{i+1}}\right)e^{-j(k_{i+1}-k_i)z_{i+1}} \tag{A.18}$$

Finally, we obtain the total transfer matrix by cascading each matrix in sequence

$$\begin{aligned} \begin{pmatrix} F_1 \\ R_1 \end{pmatrix} &= M_1 M_2 ... M_{N-1} \begin{pmatrix} F_N \\ R_N \end{pmatrix} = T \begin{pmatrix} F_N \\ R_N \end{pmatrix} \\ &= \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} F_N \\ R_N \end{pmatrix} \end{aligned} \tag{A.19}$$

A.2 Scattering and tunneling problems

The transfer matrix method reduces the irregular potential function to a black box problem governed by the transfer matrix M. We can now solve a typical scattering or tunneling problem for transmission or reflection, as depicted in figure A.2.

As there is no wave incident from the right, we can set R_N = 0 in equation A.19. This allows us to write F_N and R_1 as a function of T_1:

$$\begin{pmatrix} F_1 \\ R_1 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} T_N \\ 0 \end{pmatrix} \tag{A.20}$$

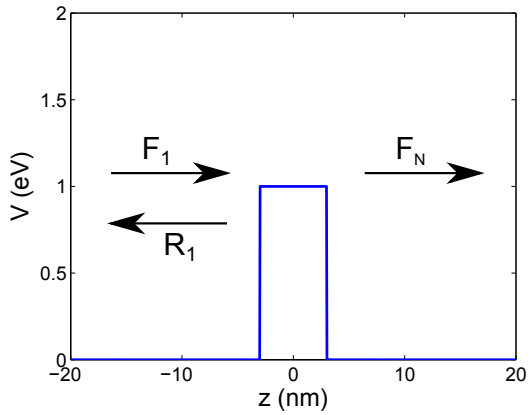


Figure A.2: A typical quantum mechanical tunneling problem

$$F_1 = T_{11} F_N \tag{A.21}$$

$$R_1 = T_{21} F_N = \frac{T_{21}}{T_{11}} F_1 \tag{A.22}$$

The transmission and reflection can now readily be calculated:

$$R_{\text{Reflection}} \equiv \frac{|R_1|^2}{|F_1|^2} = \left| \frac{T_{21}}{T_{11}} \right|^2 \qquad (E > V_1) \tag{A.23}$$

$$T_{\text{Transmission}} \equiv \frac{m_1 k_N}{m_N k_1} \frac{|F_N|^2}{|F_1|^2} = \frac{m_1 k_N}{m_N k_1} \frac{1}{|T_{11}|^2} \qquad (E > V_1 \text{ and } E > V_N) \tag{A.24}$$

The condition on the energy levels in equations A.23 and A.24 ensure that there are the usual propagating waves.

The transfer matrix method was applied for the tunneling problem here described. The resulting reflection and transmission are shown in figure A.3. The barrier height was 1 eV and width was 6 nm. The particle mass was 0.067 times the free electron mass. One clearly sees the tunneling just below 1 eV and the resonances afterwards, causing the transmission not to be a flat one above 1 eV.

A.3 Bound state problems

The transfer matrix method can also be used to find the bound states of e.g. a quantum well. A problem is only bound if the energy is lower than the potential

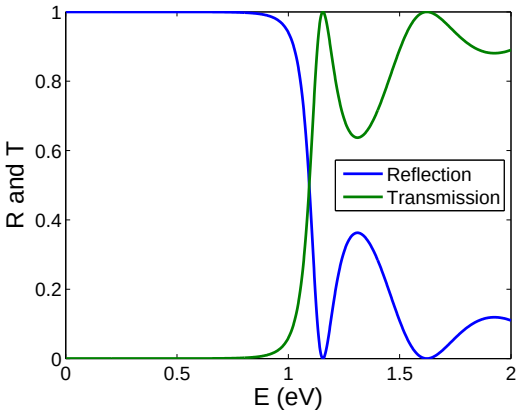


Figure A.3: Reflection and transmission in tunneling problem

on the left and right of the structure. Only exponentially decaying tails are allowed. In the transfer matrix method this means:

$$F_1 = 0 \qquad (E < V_1) \tag{A.25}$$

$$R_N = 0 \qquad (E < V_N) \tag{A.26}$$

If this is not the case, the wave function will go to infinity and is clearly not bound. This means:

$$0 = F_1 = T_{11} F_N \tag{A.27}$$

The above can only be true if F_N is zero, which is the trivial solution of a wave function being completely zero, or if T_{11} is zero. As T_{11} depends on the energy E, this is only true for certain energies and we have found discrete bound states.

$$T_{11}(E) = 0 \quad (E < V_1 \text{ and } E < V_N) \tag{A.28}$$

The states themselves can be found by using equation A.3.

We can illustrate this by finding the states in a square quantum well. Figure A.4 shows a quantum well of 1 eV deep and 6 nm wide. The electron mass is assumed 0.067 times the free electron mass. The quantum states are at 0.0987 eV, 0.3852 eV and 0.8124 eV above the bottom of the well. The typical shapes of the wave functions can also be observed.

The real power of this method of course lies in the fact that the eigenstate of any arbitrary potential function can be found. Figure A.5 shows a quantum house in a quantum well. Because the potential function resembles two neighbouring quantum wells, two eigenstates can be found. The chimney causes a substantial disruption of the two wells, causing the eigen energies to move apart further. It

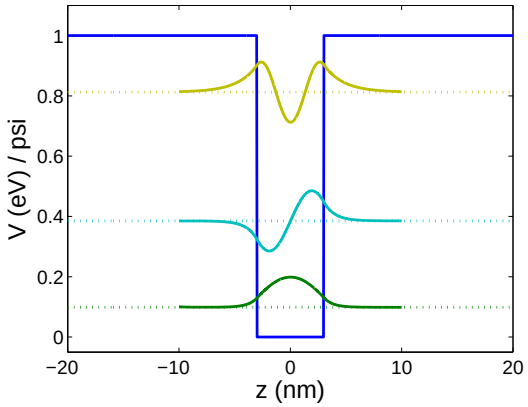


Figure A.4: Eigen energies and states in a square finite quantum well

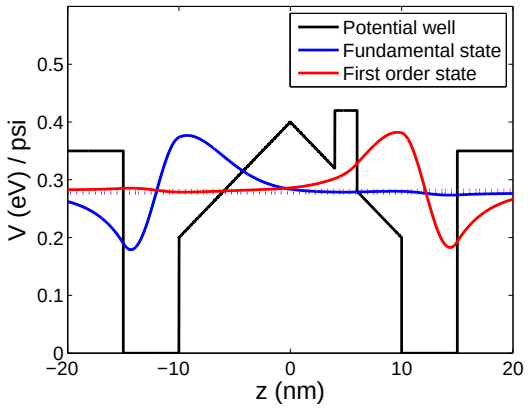


Figure A.5: Eigen energies and states of a quantum house in a quantum well

can be seen that in the fundamental state the wave function becomes zero at the chimney, whereas this is definitely not the case for the first order wave function.

References

[1] Björn Jonsson and Sverre T Eng. *Solving the Schrodinger equation in arbitrary quantum-well potential profiles using the transfer matrix method*. IEEE journal of quantum electronics, 26(11):2025–2035, 1990.

List of Figures

List of Tables